Non-covalent interactions in biomolecules studied by ^{17}O NMR and MD simulations

---

Thesis submitted to the
Board of the Faculty of Biological Sciences
in partial fulfilment of the requirements
for the Degree of Doctor of Philosophy
at the University of Oxford

Vincent J. Lemaître

St–Hugh's College, Oxford

Michaelmas Term 2004

ABSTRACT

Vincent J. Lemaître — Doctor of Philosophy

St–Hugh's College, Oxford — Michaelmas Term, 2004

Non-covalent interactions in biomolecules studied by ^{17}O NMR and MD simulations

PART I:

PART II:

LIST OF PUBLICATIONS

1. Contera, S, Lemaître V, de Planque MR, Watts A, Ryan J. 2005. Membrane–Peptide Interactions Studied by Atomic-Force Microscopy and Molecular Dynamic Simulation. (To be submitted).
2. Lemaître V, Yeagle P and Watts A 2005. Molecular Dynamic Simulation of Retinal in Rhodopsin: From the Dark-Adapted State Towards Lumirhodopsin. (To be submitted).
3. Lemaître V and Watts A 2005. Conformational Studies of Retinal in Rhodopsin. Ann. Rev. of Biophys. Submitted (To be submitted).
4. Lemaître V, Watts A, Fischer WB. 2004. Steered Molecular Dynamics Study of Ion Permeation across Vpu Ion-Channels. Submitted.
5. Lemaître V, Willbold D, Watts A, Fischer WB. 2004. Full length Vpu from HIV-1: Combining molecular dynamics simulations with NMR spectroscopy. Submitted.
6. Lemaître V, Kim CG, Fischer D, Lam YH, Watts A, Fischer WB. 2004. Defining drug interactions with the Vpu ion channel from HIV-1. In Viral membrane proteins: structure, function and drug design, ed. WB Fischer. In press.
7. Lemaître V, de Planque MR, Howes AP, Smith ME, Dupree R, Watts A. 2004. Solid state ^{17}O NMR as a probe for structural studies of proteins in biomembranes. *J. Am. Chem. Soc.* In press.
8. Lemaître V, Smith ME, Watts A. 2004. A Review of Oxygen-17 Solid state NMR of Organic Materials. Solid State NMR, 26: 215–235.
9. Lemaître V, Ali R, Kim CG, Watts A, Fischer WB. 2004. Interaction of amiloride and one of its derivatives with Vpu from HIV-1: a molecular dynamics simulation. *FEBS Lett* 563: 75–81.
10. Pike KJ, Lemaître V, Kukol A, Anupõld T, Samoson A, et al. 2004. Solid state ^{17}O NMR of Amino Acids. *Journal of Physical Chemistry B* 108: 9256–9263.
11. Lemaître V, Pike KJ, Watts A, Anupõld T, Samoson A, et al. 2003. New Insights into the Bonding Arrangements of *L*- and *D*-Glutamates from Solid State ^{17}O NMR. *Chem. Phys. Lett.* 371: 91–97.
12. Sramala I, Lemaître V, Faraldo-Gomez JD, Vincent S, Watts A, Fischer WB. 2003. Molecular dynamics simulations on the first two helices of Vpu from HIV-1. *Biophys J* 84: 3276–3284

.

ACKNOWLEDGEMENTS

During the last three year and a half, this project has been carrying me in various places, countries, and even continents, providing me the opportunity to great and inspiring people.

I would sincerely like to thank my employer Nestec S.A. for providing me the resources and the freedom to do scientific research. In particular, I am indebted to Dr Sébastien Vincent both for convincing me to do a Ph.D. and giving me a chance to do a D.Phil. He inspired the philosophy behind this work at every level. I am also grateful to the Nestlé Research Center (NRC) for hosting me, in particular to its successive directors Prof. Andrea Pfeifer and Prof. Peter van Bladeren for the positive and continuous support they provided to this project. I am also grateful to my successive heads of department, Dr Jean-Richard Neeser, Dr Marcel-Alexandre Juillerat and Dr Laurent-Bernard Fay for giving me a total freedom in my work.

I would sincerely like to thank Prof. Anthony Watts for taking me on board in his research group, his overall supervision and continuous support in the past year. This work is the result of the freedom he gave me. I was never scolded for failing, even when it involved a health and safety officer. He also gave me many opportunities to learn more than just science.

In my journey through scientific research, various people have taught me various skills. I would like to take here the opportunity to acknowledge my teachers. I am indebted to Prof. Alexandre Bonvin (University of Utrecht) for teaching me the art of molecular modeling. This was a valuable skill I have used daily in the last three years. I am also indebted to Dr Eric Hughes (NRC), Dr René Verel and Dr Scott Goodall for teaching me solid state NMR, and Dr Dahlia Fischer for teaching me solid-phase peptide synthesis.

All the work presented in this thesis results from collaborations with inspiring people. All the work related with oxygen-17 was carried in a close collaboration with Prof. Ray Dupree, Prof. Mark Smith, Dr Andy Howes and Dr Kevin Pike from the University of Warwick. I also enjoyed working with Dr Maurits de Planque, spending weeks in RAL setting-up ad acquiring data on the 800 for the

^{17}O-labelled peptide work. I am also thankful Dr Jacco van Beek for the time we spent at the University of Southampton, unfortunately being unsuccessful. The work about rhodopsin was initiated by the enthusiasm of Prof Phil Yeagle from the University of Connecticut. I am indebted to Phil and his wife Prof. Arlene Albert for their warm hospitality in Connecticut while setting-up their X-Server G4 Cluster. The WALP work is directly related to the pleasure I had to work with Dr Maurits de Planque and Dr Sonia Contera, trying to simulate AFM pulling experiments. I would like to thank Dr Wolfgang Fischer for letting work on his favorite protein and trusting me for teaching a few of his part II students. I would also like to thank all the member of the "Watts" group over the last few years for their general help and support, and particularly Dr Paul Spooner, for his deep knowledge of biophysics, his inquiring mind and the truly inspiring discussions we had together about retinal and rhodopsin, certainly improving the quality of my work. I would also like to thank Pete Fischer and Liz Mitchell for their daily help.

The university of Oxford offers his students to share their academic life between their College and Department. I would like to thank the people from St-Hugh's, my college, which harboured me during all my stay in Oxford and the frame for my social life during my D.Phil. I met great people and built great memories, especially with the St-Hugh's Boat Club and in the frame of my office of Jr Dean. I want to remember all the moment spent with the member of the St-Hugh's MCR during the last three years, especially with Dominic, Dragana, Mika, Gabriel, Will, Samira and her family.

I would like to finish this acknowledgement by thanking my family and friends for supporting me, visiting or hosting me from time to time back home (Arash, Maxime, Eric, Susana, Marisa, Sabin, Pascal, Francesca and Corinne), keeping my closest and dearest for the end: my parents, Father and Mother and my sister Marie, always there for me and me for them.

ABBREVIATIONS

1D, 2D, 3D, …	one-dimensional, two-dimensional, three-dimensional, etc.
3QMAS	triple quantum magic-angle spinning
AFM	atomic-force microscopy
Ala	alanine
AM	amiloride
Arg	arginine
B_0	Static external magnetic field
BSI	blue-shifted intermediate (retinal)
CP	cross polarization
CSA	chemical shift anisotropy
CD	circular dichroism (analytical technique) or cluster of differentiation (co-receptor)
CD4	cluster of differentiation 4 (co-receptor)
Cys	cystein
DARR	dipolar-assisted rotational resonance
DAS	double angle spinning
DMPC	1,2-dimyristoyl-*sn*-glycero-3-phosphocholine
DPPC	1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine
DSPC	1,2-disteraoyl-*sn*-glycero-3-phosphocholine
DFT	density functional theory
DNA	deoxyribonucleic acid
DOR	double orientation rotation
e	elementary charge
EFG	electric field gradient
ER	endoplasmic reticulum
F	force
FDA	federal drug administration
fs	femtosecond (10^{-15} s)
FTIR	Fourier-transform infrared spectroscopy

Glu	glutamic acid
Gly	glycine
GPCR	G protein glutamate receptor
HMA	hexamethyleneamiloride
ICH	international conference on harmonization
H	Hamiltonian operator
H_{II}	non-lamellar phase in lipids
HIV	human immunodeficiency virus
Hz	Hertz
K_d or k_d	harmonic spring constant
k	Boltzmann constant
kDa	kiloDalton (atomic mass unit)
kJ	kiloJoule
Leu	leucine
LINCS	
LJ	Lennard-Jones
nm	nanometer
MAS	magic-angle spinning
MD	molecular dynamics
MQ MAS	multiple quantum magic-angle spinning
MS	mass-spectrometry
MPI	message-passing interface
NMR	nuclear magnetic resonance
NPT	isothermal-isobaric ensemble
NQR	nuclear quadrupole resonance
ns	nanosecond (10^{-9} s)
PDB	protein databank
Phe	phenylanaline
PME	particle mesh Ewald

ppm	part per million
POPC	1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine
ps	picosecond (10^{-12} s)
Q	scalar quadrupole moment
RAM	random-access memory
RMSD	root mean-square deviarion
RMSF	root mean-square fluctuation
RR	rotational resonance
ROS	rod outer segment
Ser	serine
SHAKE	
SMD	steered molecular dynamic simulation
SPC	single point particle (theoretical model for water)
T	absolute temperature
TM	transembrane
Trp	tryptophan
Tyr	tyrosin
UV-VIS	ultraviolet-visible spectroscopy
VMD	visual molecular dynamics (software for visualization of molecular models)
Vpu	Viral protein U
WALP	synthetic peptide mimicking TM segment, made of tryptophan, alanine and leucine residues

TABLE OF CONTENTS

Introduction

STRUCTURAL APPROACH

The central premise of structural biology is that the knowledge of the three-dimensional structure of biomolecules such as proteins and nucleic acids is the key to the proper understanding for their respective mechanism of action and interaction (structure-function relationships).

A suitable approach in gaining understanding of the function of biological molecules and macromolecular assemblies follows a logical progression from the determination of the chemical composition of the system under study, followed by generation of an accurate three-dimensional structure. Once this picture is established, the characterization of mobility and dynamics of the system leads to an understanding of the mechanism of action, finally leading to a complete knowledge of its biological function in the best cases.

This way of thinking has driven biological and pharmaceutical research in the last 15 years and was accompanied by the rapid development of experimental techniques and theoretical methods aimed at establishing the structural and mobility characteristics of the various component molecules, and then to relate these to an understanding of biological function.

The tools of structural biology include both experimental techniques (NMR, electron microscopy and of course X-ray crystallography), and theoretical methods related with the building of three-dimensional models of biomolecules based on the knowledge built from gathered experimental data (secondary structure prediction methods, homology modelling, molecular dynamic simulations, docking simulation). Often experimental techniques and theoretical methods are closely related, in the sense that the interpretation of the experimental data relies on theoretical methods, due to the complexity of the data sets themselves (e.g. structure determination by solution NMR relies on simulated annealing, a MD technique, in order to interpret the experimentally-derived distance restraints).

The application of the traditional structural biology tools (X-ray crystallography and NMR) to membrane proteins is shown to be extremely challenging. Membrane proteins are abundant, estimated to constitute a third of the genomes, and their biological importance extreme, as they are involved in almost every process in the cell and account for the activity of the majority of all prescription drugs (3, 4). Despite being such important and abundant, there is a real shortage of structural information available. Less than 100 atomic-resolution structures of membrane proteins has been deposited in the Protein Data Bank, compared to several thousands of water-soluble proteins. The experimental difficulties encountered using classical techniques have led to the development of new biophysical and computational methods (5).

In the 1990's, the number of experimental techniques developed to collect structural information regarding membrane-associated proteins has greatly increased. Some of them are an evolution of techniques traditionally used for globular proteins, like X-ray crystallography (new crystallographic methods – single-particle methods) and Solution NMR (new pulse sequences developed for the use of micelle-solubilized proteins). In addition, new techniques have emerged, including solid state NMR for selected accurate distance measurements up to 5-10 Å, site-directed spin-labelling electron paramagnetic resonance for longer distance constraints, Atomic-Force Microscopy (AFM) and electron microscopy for membrane imaging, allowing in some cases to suggest aggregation models for membrane protein oligomerization.

On the other hand, the use of theoretical computational methods like secondary-structure prediction of membrane protein and molecular dynamics simulations (MD) has also made an important impact, taking advantage of the increasing computational capability available, improvement of the algorithms and the slow growth of experimentally-derived information.

Here, attention has been focused on two of these methods, with a special emphasis on trying to link them together, using experimentally derived information as the seed for building molecular models which are subsequently investigated by MD simulations.

A more complete introduction to each technique is given at the beginning of each of the dedicated sections of the thesis.

NON–COVALENT INTERACTIONS IN BIOLOGY

"The essence of biological processes—the basis of the uniformity of living systems—is in its most fundamental sense molecular interactions; in other words, the chemistry that takes place between molecules. Biochemistry is the chemistry that takes place within living systems. To truly understand biochemistry, we need to understand chemical bonding" (2).

The strongest bonds that are present in organic molecules are covalent bonds. A covalent bond is formed by the sharing of a pair of electrons between adjacent atoms, where a typical carbon-carbon (C-C) covalent bond has a bond length of 1.54 Å and the energy involved in such bond is approximately 356 $kJ \cdot mol^{-1}$, meaning considerable energy must be expended to break covalent bonds (6).

Beyond covalent interactions that define molecules, many weaker non-covalent interactions govern the properties of reversible interactions involving biomolecules (2). The formation and dissociation of specific non-covalent interactions between various macromolecules play a crucial role in the function of biological systems. Such weak, non-covalent forces play essential roles in vital biological processes like replication of DNA, the folding of proteins, the specific recognition of substrates by enzymes and the detection of molecular signals. Indeed, all biological structures and processes depend on the interplay of non-covalent interactions as well as covalent ones (2).

There are three fundamental non-covalent bond types: electrostatic interactions, hydrogen bonds and van der Waals interactions. These different types of interactions differ in geometry, strength, and specificity. Furthermore, they are greatly affected in different ways by the local environment of the atoms involved in such interactions (1).

Electrostatic interactions depend on the electric charges on the atoms and the interaction strength is given by Coulomb's law. The electrostatic interaction between two atoms bearing single opposite charges separated by 3 Å in water has an energy of 5.9 $kJ \cdot mol^{-1}$ (2).

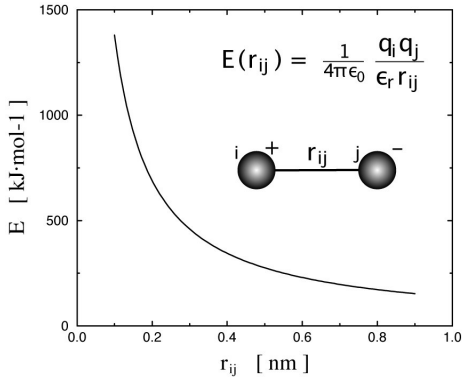


Figure 1: Illustration of the Coulomb's law. r_{ij} is the distance between the two electrical charges, $E(r_{ij})$ is the energy, q_i and q_j are the electrical charges on the two atoms, ε_r is the dielectric constant, which accounts for the effects of the intervening medium, and $1/4\pi\varepsilon_0$ is a proportionality constant.

Hydrogen bonds are weaker interactions, but nonetheless crucial for the structure of biological macromolecules such as DNA and proteins, since they are directly in charge of the secondary structure of these molecules and effect their interactions with other molecules (e.g. solvent like water). The hydrogen atom in a hydrogen bond is partly shared between two relatively electronegative atoms such as nitrogen or oxygen. Hydrogen bonds are fundamentally electrostatic interactions and involve three players: the hydrogen atom itself, the hydrogen-bond donor, which is the atom to which the hydrogen is covalently bonded, and the hydrogen-bond acceptor, which less tightly linked to the hydrogen atom. The relatively electronegative atom to which the hydrogen atom is covalently bonded pulls electron density away from the hydrogen atom so that it develops a partial positive charge. Thus, the hydrogen-atom can interact with a second atom having a partial negative charge through an electrostatic interaction. Hydrogen bonds are much weaker than covalent bonds and involve energies of 4–13 kJ·mol^{-1} compared to approximately 418 kJ·mol^{-1} for a carbon-hydrogen covalent bond (6).

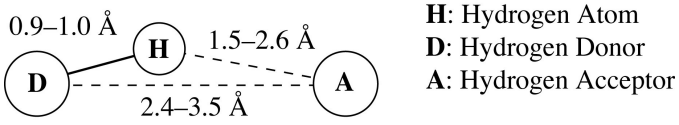
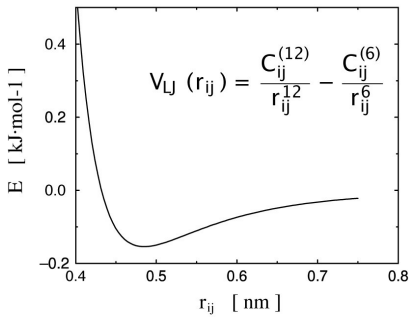


Figure 2: Illustration of a hydrogen bond. Hydrogen bonds are also somewhat longer than are covalent bonds; their bond distances (measured from the hydrogen atom) range from 1.5 to 2.6 Å; hence, distances ranging from 2.4 to 3.5 Å separate the two nonhydrogen atoms in a hydrogen bond. The strongest hydrogen bonds have a tendency to be approximately straight, such that the hydrogen-bond donor, the hydrogen atom, and the hydrogen-bond acceptor lie along a straight line (2).

Van der Waals interactions arise from induced dipoles, which result from time-fluctuations in the electronic-charge distribution around an atom. Therefore the charge distribution is not perfectly symmetric and this transient asymmetry in the electronic charge around the atoms acts through electrostatic interactions to induce a complementary asymmetry in the electron distribution around its neighbouring atoms. The resulting attraction between two atoms increases as they come too close from each other. Closer than a distance called the van der Waals contact distance, very strong repulsive forces become dominant. The physical origin of the repulsion is related to the Pauli principle: when the electronic clouds surrounding the atoms starts to overlap, the energy of the system increases abruptly.

The energies associated with van der Waals interactions are small. A typical van der Waals interaction contributes from 2 to 4 kJ·mol^{-1} per atom pair. When the surfaces of two large molecules come together, involving a large number of atom pairs interacting together, the net effect, becomes substantial (2). The van der Waals interactions and the repulsive forces between atoms are described by a semi-empirical law called the Lennard-Jones potential (see Fig. 3).



Figures 3: Illustration of van der Waals interactions and the Lennard-Jones potentials describing both the attractive (van der Waals) and repulsive behaviour of the atoms. The term $\propto 1/r_{ij}^{6}$, which dominates at large distance, constitutes the attractive part, while the term $\propto 1/r_{ij}^{12}$, which dominates at short distance, models the repulsion between atoms when they are brought very close to each other. The parameters $C_{ij}^{(12)}$ and $C_{ij}^{(6)}$ are chosen to fit the physical properties of the material.

AIM OF THE PROJECT

This D.Phil. work was designed to explore new potentials of oxygen-17 solid state NMR and molecular dynamics simulation for the application to non-covalent interactions in biomolecules, more specifically hydrogen bonding in biomolecules and ligand-protein interactions. The work has been split in two parts. Part I (Chapters 1.1–1.4) is focused on developing an approach for characterising hydrogen bonds within a biomolecule in a membrane environment by oxygen-17 solid state NMR. At first, crystallized amino acids were investigated, demonstrating the sensitivity potential of oxygen-17 as a probe extremely sensitive to the local environment of the label. The potential of this isotope to characterise hydrogen bonds in a membrane protein context was then demonstrated by applying successfully the method to a selectively labelled transmembrane peptide in hydrated multilamellar vesicles. Part II (Chapters 2.1–2.4) is focused on the use of molecular dynamic simulation to study ligand-protein interactions and hydrogen bonds within membrane peptides. This methodology was applied to several systems including bovine rhodopsin to investigate the early step of the activation mechanism of the protein, WALP peptides to complement experimental atomic-force microscopy studies, and vpu from HIV-1 to investigate drug-protein and protein-protein non-covalent interactions.

Each of the following chapter constitutes a complete and independent manuscript, as already published, in press or submitted for publication. Therefore the usual format of presenting a thesis has been slightly modified in order to accommodate the original versions of these written publications. References will be presented within each separate publication. Since two different topics will be presented in Parts I and II, both will be introduced by a separate overview and summarized by a separate conclusions section.

Chapter 1.1: Review of oxygen-17 solid state NMR applied to organic molecules

This chapter provides a snapshot of the current state of development of oxygen-17 solid state NMR applied to organic materials. The NMR techniques and enrichment protocols that are driving this progress are outlined. The ^{17}O parameters derived from solid state NMR experiments are

summarized and the structural sensitivity of the approach to effects such as hydrogen bonding highlighted. The prospects and challenges for ^{17}O solid state NMR of biomolecules are discussed.

Chapter 1.2: Resolution of multiple oxygen-17 signals in organic molecules

Study on ^{17}O-labelled glutamic acid and glutamate showing highly structure spectra. In the case of L- and D-glutamic acids, lines from all 4 oxygen sites are readily distinguished and partially assigned by MAS. Monosodium L-glutamate forms more complex crystals, with two molecular conformations per unit cells, which requires the application of more sophisticated methods, like double angle rotation and 3QMAS NMR to be resolved. This chapter shows that oxygen-17 solid state NMR provides enough resolution to study small ligands of the complexity of a small neurotransmitter uniformly labelled.

Chapter 1.3: Preparation and characterisation by ^{17}O solid state NMR of a transmembrane peptide.

Oxygen-17 solid state NMR was applied to a selectively labelled biomolecule, a selectively labelled transmembrane peptide, ^{17}O-[Ala12]-WALP23. The peptide has been studied as a lyophilized powder and incorporated in hydrated phospholipid vesicles. At high magnetic field, it is feasible to apply ^{17}O NMR to the study of membrane-incorporated peptides. Furthermore it is possible to estimate distances within oxygen-17 labelled peptides, based on empirical correlation built on amino acids.

Chapter 2.1: Review on conformational study on retinal bound to bovine rhodopsin

This review summarizes the information about the conformation of retinal in bovine rhodopsin, obtained by optical and solid state NMR spectroscopies, X-ray diffraction and *ab initio* theoretical methods, covering the dark-adapted state and various photo-activated intermediate states. An overview of the rhodopsin features relevant for the conformation of its chromophore retinal is also included, in particular the binding-pocket and the molecules of water present located near retinal.

Chapter 2.2: Molecular dynamic simulation (MD) focused on an accurate structural description of the chromophore and involving the parametrization of retinal in order to match experimentally derived information. The behavior of retinal in the dark-adapted state matches experimental information, including the last 2.2 Å crystal structure. It was also possible to study the early photo-

activation of bovine rhodopsin and identify retinal conformation related to photorhodopsin, bathorhodopsin. Known dynamic and structural information derived from experiments could also be matched, suggesting a succession of conformational events forming, common to the various paths followed by the chromophore during its photo-activation from dark-adapted state towards lumirhodopsin.

Chapter 2.3: WALP23 peptide represents a consensus sequence for transmembrane protein segments, and has special relevance for the packing of helices in proteins such as GPCRs. Based on the protocol for the previously described WALP23 peptide, a similar WALP including a cystein residue was designed for atomic-force microscopy experiments. MD simulation was used to complement Atomic force microscopy (AFM) for the imaging of bilayers containing the peptide and study single peptide unfolding and extraction from the membrane. MD simulations, at equilibrium and where additional forces were applied on selected atoms, have been used to simulate AFM force profile and suggest non-covalent molecular interactions involved.

Chapter 2.4: Simulation of $vpu_{1\text{-}32}$ pentamers interacting with amiloride derivatives

MD simulations were performed on a simplified model of the ion-channel in order to probe study the blocking action of cyclohexamethylene amiloride, a derivative of amiloride shown to inhibit ion channel activity, docked on a pentameric bundle. In the present 12 ns simulation study a putative binding site of blockers in a pentameric model bundle built of parallel aligned helices of the first 32 residues of Vpu was found near Ser-23. The blocker orientates along the channel axis with its alkyl ring pointing inside the pore, which leads to a blockage of the pore.

Appendix A: Systematic study of oxygen-17 labelled amino acids

^{17}O solid state NMR from 14 amino acids is reported greatly increasing the number investigated and showing the NMR interaction parameters for amino acids have a wide variation of χ_Q, η and δ_{iso}. In most cases well-separated resonances from carbonyl and hydroxyl oxygens with distinct second-order quadrupolar line shapes are observed with magic angle spinning (MAS). Double-angle rotation (DOR) produces a decrease in the line width by more than a factor of 40 for amino acids, providing very high resolution, ~1 ppm, spectra. A strong correlation (~–1200 ppm/Å) is found between δ_{iso} and the C–O bond length of the carbonyl oxygens.

Appendix B: Simulation of vpu_{1-52} in hydrated POPC bilayer

MD simulation has been used to refine the secondary structure of a model for the first 52 amino acids of Vpu (HV1H2). This part corresponds to the first two helices of the protein, which are thought to be associated with the membrane. The MD simulations also aim to study the connection of the transmembrane helix (helix-1) to the second helix (helix-2) residing at the surface of the membrane, more specifically, the flexible linker region, including a Glu-Tyr-Arg (EYR) motif. Repeated molecular dynamics simulations show that Glu-28 is involved in salt bridge formation with Lys-31 and Arg-34 establishing a kink between the two helices. Helix-2 remains in a helical conformation indicating its stability and function as a peptide "float", separating helix-1 from the rest of the protein known to interact with CD4 receptor.

REFERENCES

1. Atkins P, de Paula J. 2002. *Physical Chemistry*. New York: W.H.Freeman and Company
2. Berg JM, Tymoczko JL, Stryer L. 2002. *Biochemistry*. New York: W.H.Freeman and Company
3. Marchese A, George SR, Kolakowski LF, Jr., Lynch KR, O'Dowd BF. 1999. Novel GPCRs and their endogenous ligands: expanding the boundaries of physiology and pharmacology. *Trends Pharmacol Sci* 20: 370
4. Milligan G, Rees S. 1999. Chimaeric G alpha proteins: their potential use in drug discovery. *Trends Pharmacol Sci* 20: 118
5. Torres J, Stevens TJ, Samso M. 2003. Membrane proteins: the 'Wild West' of structural biology. *Trends Biochem Sci* 28: 137
6. Vollhardt KPC, Schore NE. 1998. *Organic Chemistry – Structure and Function*. New York: W.H.Freeman and Company

PART I

Introduction Part I – Solid State NMR

OVERVIEW

Nuclear magnetic resonance (NMR) spectroscopy can provide atomic-resolution structures of biological molecules, such as proteins. The nuclear spin is associated with a magnetic moment, and defines the basic resonance frequency. The exact resonance frequency depends on the chemical environment of each spins, as a result the NMR spectrum of a protein will show NMR signals with slightly different frequencies (called chemical shifts).

The use of nuclear magnetic resonance (NMR) spectroscopy in biological research has a long and successful history, primarily in early stages of drug discovery. Most of these NMR studies have been performed with liquid solutions and were conducted primarily for structure elucidation of biomacromolecules.

More recently, solid state NMR spectroscopy has been able to provide high-resolution structural information of selected part of biomolecules forming micro-crystaline or lyophilized powder (2, 3, 32), as well as to give detailed structural, dynamic and electronic information about drugs and ligands while constrained at their site of action in membrane-embedded receptors at near physiological condition (natural membrane fragment or reconstituted complexes) (14, 15, 19, 24, 25, 27, 28). Furthermore, solid state NMR is the only technique able to provide structural information about the conformation of membrane peptides (15 structures in the PDB, including Fd, Vpu, gramicidin and M2) (10, 17, 18, 31) and the oligomers they forms (6). Solid state NMR has also become a major player in the characterization of small molecules in the solid state. Worldwide regulatory authorities (FDA and ICH) now recognize the importance of solid state NMR spectroscopy and how the technique is intimately tied to the drug development process, where the technique has become used to set acceptance criteria for polymorphism in drug substances and drug products (approximately 80–90% of pharmaceutical products on the market exist in the solid form) (26).

Solid state NMR spectra are more complex than solution NMR spectra, because the full effect of anisotropic or orientation-dependent interactions are observed in the spectra, whenever it is averaged out in solution for small molecules due to rapid random tumbling. In the solid state, this motion is usually absent, and the anisotropic interactions present – chemical shift anisotropy (CSA), dipolar and quadrupolar couplings – lead to a broadening of the resonances (13, 23). These anisotropic interactions, on the one hand, have the significant disadvantage of hindering the resolution of distinct sites, but, on the other hand, contain valuable structural and dynamic information. Specifically, the CSA and quadrupolar interactions provide insight into electronic structure and bonding, while the dipolar coupling offers direct access to inter-nuclear distances.

Solid-state NMR can provide the same type of information available from corresponding solution NMR spectra, but a number of special techniques and / or equipment are needed.

NMR INTERACTION IN THE SOLID STATE

In the solid state, there are at least seven ways for a nuclear spin to communicate with its surroundings:

- ❶ Zeeman interaction of nuclear spins
- ❷ Direct dipolar spin interaction (through space dipole-dipole interaction)
- ❸ Various type of spin coupling, including indirect spin-spin coupling (J-coupling), nuclear-electron spin coupling (paramagnetic interaction), coupling of nuclear spins with molecular electric field gradients (quadrupolar interaction)
- ❹ Direct spin-lattice interactions
- ❺ Indirect spin-lattice interactions
- ❻ Chemical shielding and polarization of nuclear spins by electrons
- ❼ Coupling of nuclear spins to sound fields

Nuclear spin interactions are first distinguished on the basis of whether they are external or internal. One can define the total spin Hamiltonian operator:

$$H_T = H_{ext} + H_{\text{int}}$$

The Hamiltonian describing the external spins interaction:

$$H_{ext} = H_0 + H_1 \qquad \text{with } H_0 = \gamma\hbar \mathbf{B}_0 I_z$$

H_{ext} gathers the interactions with external fields $\mathbf{B}_0$ and $\mathbf{B}_1$. The interactions H_i refer to the Zeeman interaction of the spins with the magnetic fields applied on the system.

The Hamiltonian describing internal spins interactions:

$$H_{\text{int}} = H_{CS} + H_J + H_D + H_Q + H_L$$

H_{int} gathers the following interactions:

H_{CS}: chemical shielding interactions for spins *I* and *S*

H_J: indirect spin-spin coupling (*J* coupling)

H_D: direct dipolar interaction

H_Q: quadrupolar interactions for spins *I* and *S* (only on spin > 1/2)

H_L: interaction of spins I and S with the lattice

In the solid state, all these interactions can make a contribution and shift the spin state energies in the NMR spectra.

For most cases, one can assume the high-field approximation, which means the external interactions (Zeeman interaction and other external magnetic fields) are much higher than internal NMR interactions. Correspondingly, these internal interactions can be treated as perturbations on the Zeeman Hamiltonian.

BASIC SOLID STATE NMR TECHNIQUES

A number of methods have been developed in order to minimize large anistropic NMR interactions between nuclei and increase S/N in rare spin spectra:

Magic Angle Spinning

This technique has led to significant simplification of the solid state NMR spectra by reducing the powder pattern to an average isotropic value, where the observed signals are narrower (see Fig. 1). This is obtained by imposing the geometrical averaging of the magnetic anisotropic interactions through a rotation at an angle (so called magic angle) corresponding to the diagonal of a cube (54.7°). Chemical shielding, dipolar and quadrupolar interaction both contain a $(3\cdot\cos^2\theta - 1)$ dependence term with respect to the applied magnetic field. In solution, rapid anisotropic tumbling averages this spatial component to zero (due to the integration over $\sin\theta\, d\theta$). In the case of solids, another way to average this angular component is to average it mechanically, by spinning rapidly the sample with an angle relative to the static magnetic field so that the angular dependence $(3\cdot\cos^2\theta - 1)$ is average to zero. This is achieved by spinning the sample at an angle called magic angle (54.74°) (see Fig. 2).

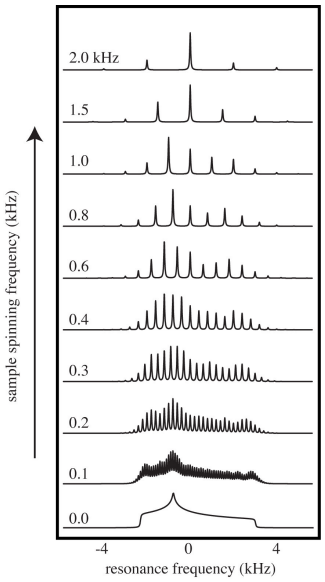


Figure 1: Simulated spectra showing the effect of MAS on the anisotropic lineshape due to a CSA interaction (Reproduced from (9)).

The rate of MAS must be greater than or equal to the magnitude of the anisotropic interaction to average it to zero. All anisotropic interactions with a magnitude below the frequency of spinning of the sample are averaged out. The technique is of limited use for “high-gamma” nuclei like protons and fluorine, which can have dipolar coupling in excess of 100 kHz. If the sample is spun at a rate less than the magnitude of the anisotropic interaction, a manifold sideband becomes visible, which

are separated by the rate of spinning (in Hz). Even with MAS slower than the breadth of the anisotropic interaction, S/N is increased, because the signal becomes localized under the spinning sidebands, rather than spread over the entire breadth as in case of the static NMR spectrum. In addition, analyses of the residual sidebands pattern can also provide valuable orientational information.

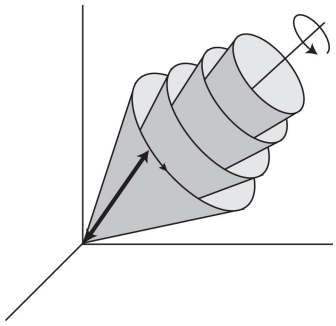


Figure 2: The effect of MAS for the specific example of a dipolar coupling between two spins. The four cones represent the range of positions adopted over the course of one rotor period for four different orientations of the internuclear vector relative to the rotor axis. The double-headed arrow represents an arbitrary position of one of the internuclear vectors (Reproduced from (1)).

Practically, the samples are finely powdered or dissolved in micelles and packed tightly into rotors, which are then spun at rates from 1 to 35 kHz, depending on the rotor size and type of experiment being conducted.

Cross Polarization

Polarization from abundant nuclei like ^{1}H and ^{31}P can be transferred to dilute or rare nuclei like ^{13}C, ^{15}N, in order to enhance signal to noise and reduce time between successive experiments.

The signal is enhanced due to the spins I having a high polarisation because of their high γ. The cross polarization enhances signal from dilute spins potentially by a factor of γ_I/γ_S, where I is the abundant spin and S is the dilute spin.

Since abundant spins are strongly dipolar coupled, they are therefore subject to large fluctuating magnetic fields resulting from motion. This induces rapid spin-lattice relaxation at the abundant nuclei. The end result is that one does not have to wait for slowly relaxing dilute nuclei to relax, rather, the recycle delay is dependent upon the T_1 of the abundant nuclei I (^{1}H, ^{19}F, etc.). This lead also to the improvement of the S/N since acquisitions can be made much faster.

Polarization is transferred during the spin locking period, called contact time, and a π/2 pulse is only made on protons to polarized them to the x-axis. They are then held there by a spin locking pulse:

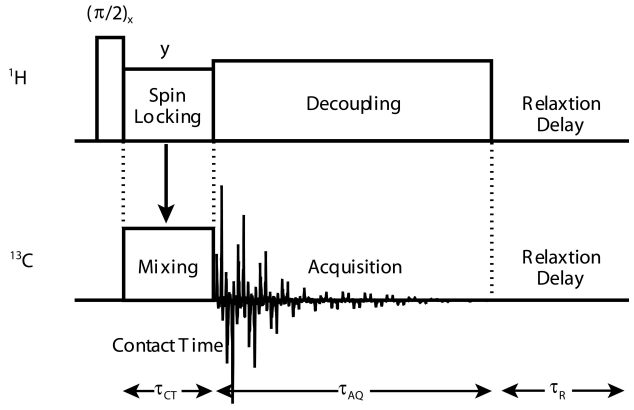


Figure 3: Pulse scheme for a typical cross-polarization (CP).

Cross polarization requires that nuclei are dipolar coupled to one another, but even works while samples are being spun rapidly at the magic angle (though not if the spinning rate is greater than the anisotropic interaction). Thus Cross polarization is often combined with MAS (CP-MAS).

The key to obtaining efficient polarization is the proper setting the Hartmann-Hahn match properly, which requires the r.f. fields of the dilute spin (e.g., $\geq {}^{1\text{-}13}C$) to be set equal to that of the abundant spin (e.g., $\geq {}^{1\text{-}1}H$) by adjusting the power on each of the channels:

$$\gamma_{^{13}C} \cdot B_{^{13}C} = \gamma_{^{1}H} \cdot B_{^{1}H}$$

$B_{^{13}C}$ and $B_{^{1}H}$ are known as the r.f. spin locking fields.

If these are set properly, the proton and carbon magnetization precess in the rotating frame at the same rate, allowing for transfer of the abundant spin polarization to the rare spin:

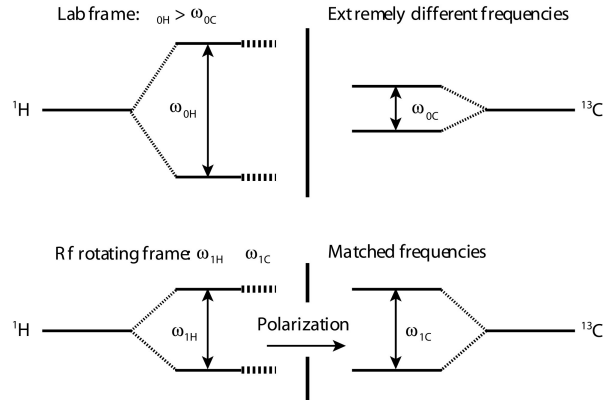


Figure 4: Illustration of the mechanism of Cross-Polarization (CP), where a proper setting of the Hartmann-Hahn match is required, by adjusting the power on each of the channels.

SOLID STATE NMR OF QUADRUPOLAR NUCLEI

Quadrupolar nuclei have a spin > 1/2 and an asymmetric distribution of nucleons giving rise to a non-spherical positive electric charge distribution. The asymmetric charge distribution in the nucleus is described by the nuclear electric quadrupole moment, eQ, which is an intrinsic property of the nucleus and is the same regardless of the environment.

The quadrupolar interaction, unlike all other anisotropic NMR interactions can be written as a sum of a first and second order interaction. The first order interaction is proportional to the nuclear quadrupole coupling constant C_Q, and the second-order interaction is proportional to $C_Q^2/\geq_0$ and is much smaller.

Table 1: NMR parameters of interest that can be extracted from quadrupolar nuclei spectra

Parameter	Symbols usually used	Unit
Isotropic Chemical Shift	δ_{iso}	[≥ ≥ ≥]
Chemical Shift Anisotropy Tensor (principal components in PAS)	σ_{11}, σ_{22}, σ_{33}	[ppm]
Orientation of CSA (Euler angles)	α, β, χ	[°]
Quadrupole Coupling Constant	$\chi_{\geq} = e^2qQ/h$	[Hz]
Nuclear Quadrupolar Electric Moment	eQ	[C · m^2]
Quadrupolar Asymmetry Parameter	$\eta_Q = (q_{xx} - q_{yy})/q_{zz}$	-
Electric field gradient tensor (main components)	V_{11}, V_{22}, V_{33} (= eq)	[V · m^2]

The Hamiltonian H_Q used to describe the quadrupolar coupling interaction is usually split in two terms: first order and second order quadrupolar Hamiltonian ($H_Q^{(1)}$ and $H_Q^{(2)}$). These two orders differ both in terms of symmetry and dependence relative to the magnetic field.

$$H_Q = H_Q^{(1)} + H_Q^{(2)}$$

Both $H_Q^{(1)}$ and $H_Q^{(2)}$ can be described in terms of spherical tensor operators. $H_Q^{(1)}$ is independent from the magnetic field and the symmetrical properties of the spherical tensors used to describe it

are such, that is averaged out to zero upon rotation along a single axis. On the other hand, $H_Q^{(2)}$ depends on the magnetic field strength and presents a more complex symmetrical behaviour: it is not averaged out to zero upon the rotation along a single axis (like a MAS experiment, see below). $H_Q^{(2)}$ is inversely proportional to the static external magnetic field B_0: $H_Q^{(2)} \propto \frac{-1}{\gamma B_0}$. This non-linear field dependence gives a significant decrease of the importance of $H_Q^{(2)}$ on the overall spectrum when the magnetic field used is stronger.

Quadrupolar nuclei interact with electric field gradients (EFGs) in the molecule, which are spatial changes in electric field around the nucleus. EFGs at the quadrupolar nuclei are described by a symmetric traceless tensor **V**. The magnitude of the first-order quadrupolar interaction is given by the nuclear quadrupole coupling constant, which accounts for both the intrinsic property of the nucleus and for its environment:

$$C_Q = eQ \cdot \frac{V_{33}}{h} \qquad \text{(in Hz)}$$

(V_{33} is the trace component with the largest absolute value, when **V** is put in its diagonalized form)

As the perturbing effect of the electric quadrupolar interaction is reduced at the higher magnetic fields, it is expected that the resolution of the signal from quadrupolar nuclei will increase, allowing the study of even more complex molecules. However, the Chemical shift anisotropy (CSA) of the nuclei is directly proportional to the static magnetic field. This means that for a particular spinning rate there is an optimal value for the static magnetic field!

Static spectra of quadrupolar nuclei

Static experiments are usually performed at several different fields or in conjunction with MAS in order to determine all the tensor properties of the label present in the sample. The limitation of the technique comes with the very little number of different sites it is able to resolve. For example, static ^{17}O NMR spectra of amino acid powder give very broad line, up to 100 kHz wide (8).

MAS of quadrupolar nuclei

High-resolution (or at least higher resolution) solid state NMR of quadrupolar nuclei like ^{17}O has been made possible through the application of magic angle spinning. However, the more complex nature of the quadrupolar interactions leads to an incomplete averaging of these interactions.

In a MAS NMR spectrum, the measured signal results from several interactions:

$$\nu = \nu_0 + \nu_{CS} + \nu_Q$$

ν_0 = Larmor frequency ν_{CS} = Chemical shift interaction
ν_Q = Residual first order and second order quadrupolar interaction, results from the coupling between the nuclear quadrupolar moment eQ, and the EFG (or eq) at the nucleus.

The resolution of MAS is limited by the incomplete averaging of the second order quadrupolar interactions. Very often MAS lead to 10-12 kHz wide bands showing a sizable second order quadrupolar coupling constant not averaged out by the rotation along a single axe. This is due to the intrinsic symmetry of the second order quadrupolar. Therefore more complex techniques are required to acquire quadrupolar nuclei spectra with a higher resolution.

To overcome the inadequacy of MAS experiments on quadrupolar nuclei, several techniques have been developed in order to average fully the quadrupolar coupling interactions. Two main avenues have been explored: an extended mechanical averaging of the sample (DOR, DAS) or a combination of multiple-quantum excitation and fast spinning (MQMAS) (5). Such techniques are intrinsically useful for signal assignment since they drastically increase the resolution by averaging out the second order quadrupolar interaction. DOR has one advantage on MQMAS is that no sensitivity is lost in the signal acquisition (in fact sensitivity is gained since the signal are much more narrow).

Double Orientation Resonance (DOR)

In experiments like Double Rotation (DOR) (21, 22, 30) or Dynamical Angle Spinning (DAS) (12, 16), the spatial averaging of the sample is achieved by spinning simultaneously around two different axes.
These methods achieve an almost complete averaging of the quadrupolar interaction (second order term) and allow the acquisition of solid state ^{17}O NMR spectra with resolution similar to half-integer spin spectra, but remain technologically very demanding, limited by probe technology. The spinning speed

that can be achieved around any of the two axes remains slower than the ones achieved by single angle MAS. In practice, line widths down to 1 ppm have been obtained (20).

DAS is similar to DOR, with the difference that the spinning angle of the sample changes at some point of the pulse program in a pulse synchronous way.

More complex pulse sequence can be combined with these methods. For example, Gann et al. have demonstrated a 2D DAS experiment (7), allowing an even better resolution, but at the cost of its sensitivity.

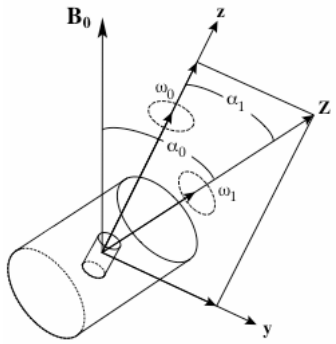


Figure 5: Principle of DOR, simultaneous rotation of the sample around two axis.

Multiple Quantum Magic Angle Spinning (MQMAS)

These experiments are a combination of multiple-quantum excitation and fast spinning called multiple-quantum MAS (MQMAS) (5). As an example, Triple-quantum MAS (3QMAS) is a two-dimensional (2D) MAS experiment designed to observe half-integer spin quadrupolar nuclei. The advantage of this technique is that it allows the use of a standard MAS probe. MQMAS has been mainly applied to inorganic sample and there are only two reports of this application on organic molecules (11, 29).

MQMAS can be used in two variants. Without rotor synchronization of the t_1 increment with the rotor period leading to strong spinning sidebands but the spectral width in the F1 dimension is not limited by the sample spinning frequency (side-bands result in loss of sensitivity and resolution). With rotor synchronization where $\Delta t_1 = \tau_R$ (rotor period), spectral width is equal to the spinning frequency and all spinning side-bands in the F1 dimension are folded onto the central bands, which increases the sensitivity and the resolution but limits the bandwidth to the maximum spinning frequency achievable.

When compared to 1D MAS, MQMAS allows much narrow line width (29) (Table 2).

Table 2: Comparison of 1D MAS and MQMAS (adapted from Wu et al. (2001)(29))

Technique	MQMAS	1D MAS
Line width [kHz]	0.1 – 0.2*	0.5 – 1.5*

*Providing that the ^{1}H-^{17}O H-bonding is weak.

MQMAS induces the reduction of the isotropic resonance in the F1 dimension by almost an order of magnitude as compared to a 1D spectrum, leading the main contribution in the residual line width to be the incomplete removal of the ^{1}H-^{17}O dipolar interactions (this can possibly be improved by using the Two-Pulse Phase Modulation (TPPM) for a better dipolar ^{1}H decoupling (29)). The other contributions to line broadening are poor crystalinity and inaccurate magic angle set-up.

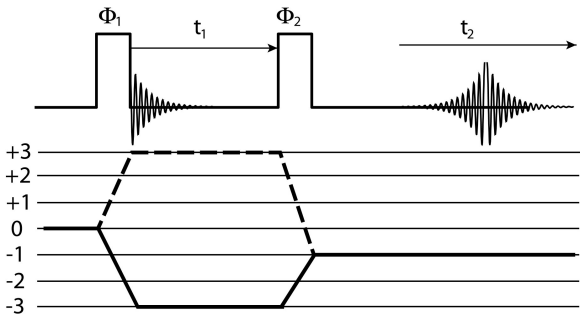


Figure 6: 3QMAS pulse sequence (adapted from (5)).

For inorganic sample, a comparison of the DOR and MQMAS techniques gave the ^{17}O DOR NMR spectra superior to ^{17}O 3QMAS NMR spectra with respect to the resolution by a factor of two (4), while the signal to noise ratio of DOR and 3QMAS NMR is comparable.

REFERENCES

1. Brown SP, Emsley L. 2002. Solid State NMR. In *Handbook of Spectroscopy*, ed. Vo-Dinh, Gauglitz: Wiley
2. Castellani F, van Rossum B, Diehl A, Schubert M, Rehbein K, Oschkinat H. 2002. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* 420: 98
3. Castellani F, van Rossum BJ, Diehl A, Rehbein K, Oschkinat H. 2003. Determination of solid-state NMR structures of proteins by means of three-dimensional 15N-13C-13C dipolar correlation spectroscopy and chemical shift analysis. *Biochemistry* 42: 11476
4. Freude D, Kärger J. 2002. NMR Techniques. In *Handbook of Porous Solids*, ed. F Schüth, KSW Sing, J Weitkamp, pp. 465. Weiheim: Wiley-VCH
5. Frydman L, Harwood JS. 1995. Isotropic Spectra of Half-Integer Quadrupolar Spins from Bidimensional Magic-Angle Spinning NMR. *J Am Chem Soc* 117: 5367
6. Fu R, Cotten M, Cross TA. 2000. Inter- and intramolecular distance measurements by solid-state MAS NMR: determination of gramicidin A channel dimer structure in hydrated phospholipid bilayers. *J Biomol NMR* 16: 261
7. Gann SL, Baltisberger JH, Wooten EW, Zimmermann H, Pines A. 1994. Cross Polarization and Dynamic-Angle Spinning of ^{17}O in L-Alanine. *Bull. Magn. Reson.* 16: 68–72

8. Goc R, Ponnusamy E, Tritt-Goc J, Fiat D. 1988. ^{17}O n.m.r. studies of amino acids in the solid state, in single- and polycrystalline forms. *Int. J. Pept. Protein Res.* 31: 130–136
9. Hodgkinson P, Emsley L. 1997. The reliability of the determination of tensor parameters by solid-state nuclear magnetic resonance. *J. Chem. Phys.* 107: 4808–4816
10. Kovacs F, Quine J, Cross TA. 1999. Validation of the single-stranded channel conformation of gramicidin A by solid-state NMR. *Proc Natl Acad Sci U S A* 96: 7910
11. Lemaitre V, Pike K, Watts A, Anupold A, Samoson A, et al. 2003. New Insights into the Bonding Arrangements of L- and D-Glutamates from Solid State 17O NMR. *Chemical Physics Letter* 371: 91
12. Llor A, Virlet J. 1988. Towards High-Resolution NMR Of More Nuclei In Solids - Sample Spinning With Time-Dependent Spinner Axis Angle. *Chemical Physics Letter* 152: 248
13. Mehring M. 1983. *Principles of High Resolution NMR in Solids*. New York: Springer
14. Middleton DA, Rankin S, Esmann M, Watts A. 2000. Structural insights into the binding of cardiac glycosides to the digitalis receptor revealed by solid-state NMR. *Proc Natl Acad Sci U S A* 97: 13602
15. Middleton DA, Robins R, Feng X, Levitt MH, Spiers ID, et al. 1997. The conformation of an inhibitor bound to the gastric proton pump. *FEBS Letters* 410: 269
16. Mueller KT, Sun BQ, Chingas GC, Zwanziger JW, Terao T, Pines A. 1990. Dynamic-Angle Spinning of Quadrupolar Nuclei. *Journal of Magnetic Resonance* 86: 470
17. Opella SJ, Marassi FM, Gesell JJ, Valente AP, Kim Y, et al. 1999. Structures of the M2 channel-lining segments from nicotinic acetylcholine and NMDA receptors by NMR spectroscopy. *Nat Struct Biol* 6: 374
18. Park SH, Mrse AA, Nevzorov AA, Mesleh MF, Oblatt-Montal M, et al. 2003. Three-dimensional structure of the channel-forming trans-membrane domain of virus protein "u" (Vpu) from HIV-1. *J Mol Biol* 333: 409
19. Patel AB, Crocker E, Eilers M, Hirshfeld A, Sheves M, Smith SO. 2004. Coupling of retinal isomerization to the activation of rhodopsin. *Proceedings of the National Academy of Science* 101: 10048–10053
20. Pike KJ, Lemaître V, Kukol A, Anupõld T, Samoson A, et al. 2004. Solid-State ^{17}O NMR of Amino Acids. *Journal of Physical Chemistry B* 108: 9256–9263
21. Samoson A, Lippmaa E. 1989. Synchronized Double-Rotation NMR-Spectroscopy. *Journal of Magnetic Resonance* 84: 410–416
22. Samoson A, Lippmaa E, Pines A. 1988. High-Resolution Solid-State NMR Averaging of 2nd-Order Effects by Means of a Double-Rotor. *Molecular Physics* 65: 1013
23. Schmidt-Rohr K, Spiess HW. 1999. *Multidimensional solid-state NMR and polymers*. London: Academic Press Limited. 478 pp.
24. Spooner PJ, Sharples JM, Goodall SC, Seedorf H, Verhoeven MA, et al. 2003. Conformational similarities in the beta-ionone ring region of the rhodopsin chromophore in its ground state and after photoactivation to the metarhodopsin-I intermediate. *Biochemistry* 42: 13371
25. Spooner PJ, Sharples JM, Verhoeven MA, Lugtenburg J, Glaubitz C, Watts A. 2002. Relative orientation between the beta-ionone ring and the polyene chain for the chromophore of rhodopsin in native membranes. *Biochemistry* 41: 7549
26. Tishmack PA, Bugay DE, Byrn SR. 2003. Solid-state nuclear magnetic resonance spectroscopy--pharmaceutical applications. *J Pharm Sci* 92: 441
27. Verdegem PJ, Bovee-Geurts PH, de Grip WJ, Lugtenburg J, de Groot HJ. 1999. Retinylidene ligand structure in bovine rhodopsin, metarhodopsin-I, and 10-methylrhodopsin from internuclear distance measurements using ^{13}C-labeling and 1-D rotational resonance MAS NMR. *Biochemistry* 38: 11316
28. Watts JA, Watts A, Middleton DA. 2001. A model of reversible inhinitors in gastric H+/K+-ATPase binding site determined by rotational echo double resonance NMR. *J. Biol. Chem.* 276: 43197
29. Wu G, Dong S. 2001. Two-dimensional ^{17}O multiple quantum magic-angle spinning NMR of organic

solids. *J. Am. Chem. Soc.* 123: 9119–9125

30. Wu Y, Sun BQ, Pines A, Samoson A, Lippmaa E. 1990. NMR Experiments with a New Double Rotor. *Journal of Magnetic Resonance* 89: 297
31. Zeri AC, Mesleh MF, Nevzorov AA, Opella SJ. 2003. Structure of the coat protein in fd filamentous bacteriophage particles determined by solid state NMR spectroscopy. *Proc. Natl. Acad. Sci.* 100: 6458
32. Zimmermann J, Jarchau T, Walter U, Oschkinat H, Ball LJ. 2004. 1H, 13C and 15N resonance assignment of the human Spred2 EVH1 domain. *J Biomol NMR* 29: 435

I.1 A Review of Oxygen-17 Solid State NMR of Organic Materials

Vincent Lemaître,[1,2] Mark E. Smith,[3] and Anthony Watts,[1]

[1] *Biomembrane Structure Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK ; email: anthony.watts@bioch.ox.ac.uk*

[2] *Nestec S.A., BioAnalytical Department, Vers-Chez-Les-Blanc, CH-1000 Lausanne 26, Switzerland*

[3] *Department of Physics, University of Warwick, Coventry, CV4 7AL, U.K.*

Keywords oxygen-17, solid state NMR, MAS, DOR, DAS, MQMAS, amino acids, organic molecules, peptide, hemeprotein, hydrogen-bond, biological applications, review.

■ **ABSTRACT** ^{17}O solid state NMR of organic materials is developing rapidly. This article provides a snapshot of the current state of development of this field. The NMR techniques and enrichment protocols that are driving this progress are outlined. The ^{17}O parameters derived from solid state NMR experiments are summarized and the structural sensitivity of the approach to effects such as hydrogen bonding highlighted. The prospects and challenges for ^{17}O solid state NMR of biomolecules are discussed.

INTRODUCTION

In the last ten years, the efforts to exploit oxygen-17 (^{17}O) in solid state NMR have greatly increased, with applications ranging from studies of amorphous materials (26, 27, 30, 51, 67, 73, 108), zeolites (16, 17) and minerals (7, 8, 99) to polymers (130, 131). Although oxygen-17 is still a relatively exotic nucleus for NMR study, the field has now matured to such a stage where applications of the techniques to problems of real biological significance are close to being realized.

The use of quadrupolar nuclei in general in biological systems was reviewed in 1998 (120). Here we review the variety of NMR approaches for studying ^{17}O in systems of biological interest, such as

organic molecules, amino acids and peptides. The different systems studied, the techniques applied and the results obtained will be discussed.

PERSPECTIVE AND OVERVIEW

One of the main reasons for studying oxygen is its ubiquity in biology. Indeed, oxygen controls or participates in nearly every biological process, especially those involving aerobic metabolism. Oxygen occupies a key position both at the structural and physiological level. In all macromolecules, including peptides, proteins, DNA and RNA (especially ribosome) and carbohydrates, oxygen has a major role in the molecular conformation observed, such as the secondary, tertiary and quaternary structure. In proteins that are stabilized through multiple hydrogen bonds, changes which occur between ground and transition states can also induce an important effect on enzyme catalysis (22). It is therefore not a surprise that oxygen atoms are involved in triggering, signalling and activation mechanisms. Many physiological processes, like substrate-enzyme, drug-receptor or protein-protein interactions, imply non-covalent bonding between the substrates. Taking into account the great importance of hydrogen bonding in biological systems, oxygen-17 has the potential of being a valuable probe of both structure and function.

Oxygen-17 is the only NMR-active oxygen isotope (nuclear spin = 5/2) and its application in NMR has been hindered by several of its intrinsic nuclear properties resulting in low sensitivity and complex spectra. Amongst these are (1) the small gyromagnetic ratio $\gamma = -3.6279 \times 10^{7} \cdot rad \cdot T^{-1} \cdot s^{-1}$ (comparable to ^{15}N), so that the resonance frequency is about $1/7^{th}$ that of protons (54.221 MHz at 9.39 T which is 400 MHz for ^{1}H), giving a relative sensitivity of 1.830 (relative to ^{13}C at constant field for equal numbers of nuclei (66)). (2) a low natural abundance of 0.037% leading to an absolute sensitivity of 0.0611 (compared to ^{13}C) so that isotopic enrichment is usually necessary; (3) a nuclear quadrupole moment $Q = 2.558 \times 10^{-30}$ m^2 (significant but much smaller than many other nuclei, e.g.^{27}Al); and (4) in a number of cases a large electric field gradient (EFG) is experienced. As a non-integer quadrupolar nucleus it is usually the central (1/2, −1/2) transition that is observed and the quadrupolar effects have to be considered beyond first order (73). This further reduces the

sensitivity compared to spin-1/2 nuclei and in solids it is the residual second-order quadrupolar broadening for half-integer spin quadrupolar nuclei that usually determines the line widths in magic angle spinning experiments (MAS) (see (107) and references therein). However compared to ^{27}Al in sites of comparable distortion second-quadrupole effects lead to only 0.064 as much broadening (107). Oxygen-17 exhibits a large NMR chemical shift range depending on the chemical function and the local environment (73), with the shift being distributed around ~2000 ppm (from water to dioxygen (41)).

As a quadrupolar nucleus, oxygen is sensitive to both the electric field gradient (EFG) and chemical shift anisotropy (CSA), which are anisotropic NMR interactions that can be used as probes for characterizing the local environment of the nucleus. The advantage of exploiting such anisotropic interactions (2^{nd} rank tensors) is that they are able to provide three-dimensional information about the local electronic structure. Indeed, both these parameters have been shown to be very sensitive to local structure and bonding (including hydrogen bonding). Anisotropic bonding of oxygen atoms in many solids produces strong electric field gradients at the oxygen nuclei and quadrupole coupling constants χ_Q of several MHz. The average for the quadrupole constant is 4.2 ± 1.5 MHz in inorganic material (average based on about 300 values (33, 73)). This implies, that for ^{17}O, the quadrupolar interaction is likely to be dominant, although other nuclear spin interactions such as dipolar interactions (~10^4 Hz) and chemical shift anisotropy (CSA) (~10-10^4 Hz) cannot be neglected. As even higher magnetic fields are applied the CSA will become increasingly important.

In 1981, it was already suggested that both EFG and CSA would be very sensitive to the electronic structure of the molecule and that they could potentially yield details and degrees of resolution (59) not available from X-ray or neutron diffraction. Amongst these structural details, are molecular dynamics, and inter- and intramolecular interactions such as hydrogen bonds. Small energy changes in the electronic structure cannot be detected by diffraction methods alone (1), because of the character of the irradiation inherent to these methods, and the inherent diffuse nature of electronic distributions which are not sufficiently resolved. It is more than twenty years since this statement was made, and since then very significant improvements in both experimental NMR (especially high magnetic fields with concomitant increases in sensitivity and resolution) and quantum

chemical simulation have made it possible to collect accurate data in order to relate NMR parameters to chemical structure.

^{17}O solution NMR is a useful technique to solve structural problems for small organic molecules. The fast reorientation of the molecules in solution (fast tumbling rate, $\tau_r > Q^{-1} \sim 10^{-6} - 10^{-7}$s (41)) results in narrow spectral lines due to the averaging of the quadrupolar interaction. Rapid quadrupolar relaxation of ^{17}O has, however, prevented successful applications of solution ^{17}O NMR to biological macromolecules (41). The large quadrupolar interaction of oxygen-containing functional groups can cause highly effective relaxation, which leads to strong broadening of the NMR signals, which can be severe for large molecules. ^{17}O solid state NMR has therefore be seen as an alternative for structural studies to ^{17}O solution NMR (29), since the relaxation times are much longer and the intrinsic spectral resolution is not limited by molecular weight.

GENERAL BACKGROUND OF EXPERIMENTAL SOLID STATE NMR TECHNIQUES APPLIED FOR HALF-INTEGER SPIN QUADRUPOLAR NUCLEI

A whole range of NMR techniques can be applied to half-integer spin quadrupolar nuclei (such as ^{17}O) in solid materials including static, magic angle spinning (MAS) and higher order averaging techniques. Static experiments are usually performed at several different magnetic fields, or in conjunction with MAS, to enable all the tensor properties of the ^{17}O sites present to be determined. The limitation of static work is the inherent resolution when dealing with resonance lines that can be very broad, up to 100 kHz wide from some amino acids (45). Hence in complex molecules with different oxygen sites there can be strong overlap between the different resonances.

High-resolution (or at least higher resolution) ^{17}O solid state NMR has been made possible through the application of MAS (Figure 1 for an illustration). However, the angular variation of the second-order quadrupolar interaction is more complex than $(3\cos^2\theta-1)$ so that this interaction is incompletely averaged under MAS. One of the most immediate consequences of this is that the

residual line widths under MAS (since spinning by MAS only narrows a factor of ~3.6 compared to a static sample (13)) means that the resolution improvement due to MAS is not as good as it would be for a spin-1/2 system, and often MAS of amino acids leads to resonances that are 10-12 kHz wide. Therefore more complex averaging techniques are required to acquire ^{17}O spectra with higher resolution.

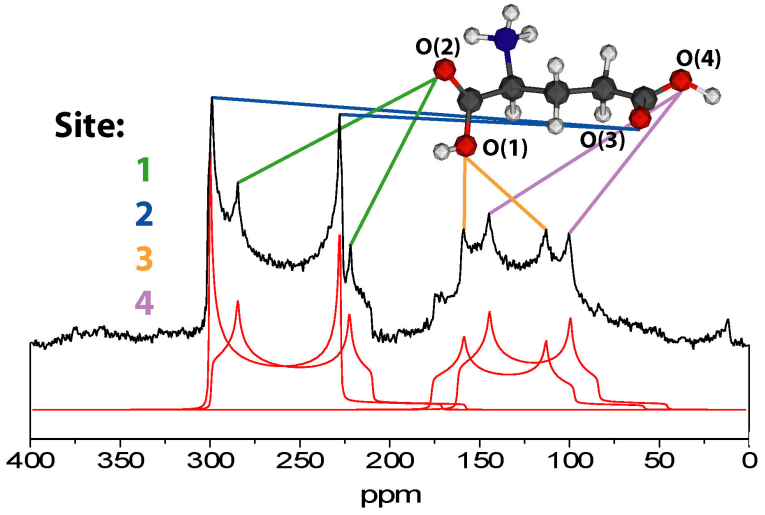


Figure 1 ^{17}O MAS NMR of L-glutamic acid·HCl. Resolution of the four distinct crystallographic oxygen sites. A highly structured spectrum is observed with 2 main resonances centered at ~260 and 125 ppm. Each resonance is composed of 2 strongly overlapping lines. The structure on the resonance at lower shift only became clear when high ^{1}H-decoupling powers were used (adapted from (70) with permission of the copyright owner).

Two main avenues have been explored to average the second-order quadrupolar broadening as well as first-order effects: extended mechanical averaging of the sample (DOR (95, 96, 128), DAS (72, 78)) or a combination of multiple-quantum excitation effectively averaging in spin space and fast spinning (MQMAS) (36). MQMAS is a two dimensional (2D) sequence that correlates multiple quantum signal with that of the central transition. In the 2D data set there exists directions where both the first and second-order quadrupole broadening are removed to produce the high resolution. An alternative high resolution approach for half-integer quadrupole nuclei is satellite transition MAS (10, 37). This correlates single quantum signal from the satellite and central transitions in a 2D data set, from which it is again possible to produce an isotropic high-resolution spectrum. The references given here provide more details about the physical background and implementation of these techniques, and they have been reviewed in detail previously (107).

The advantages of these techniques are illustrated by MQMAS which reduces the width in the isotropic F1 dimension, and DOR which reduces the width of the NMR spectra by almost an order of magnitude, when compared to a 1D MAS NMR spectrum (121). As an example L-monosodium glutamate (L-MSG) is shown (Figure 2). These are the least 5 signals but in MAS spectrum there is considerable overlap. However the additional narrowing provided by both DOR and MQ approaches gives complete resolution of the different signals (70). The main contribution to the residual line width is then probably incomplete removal of the ^{1}H-^{17}O dipolar interactions (this can possibly be improved by using the Two-Pulse Phase Modulation (TPPM) (14) for better dipolar ^{1}H decoupling (121)). Line broadening can be due to poor crystallinity and since the quadrupole interactions are large, it is required that the magic angle is accurately set-up so that this does not contribute to the line width. In MQMAS, the line width is not limited by the molecular weight of the system under observation (121). MQMAS can also suffer problems at high spinning rates and larger quadrupole interactions due to excitation effects of the MQ transitions. Consequences of this can be that sites that unequally excited so that MQ spectra cannot really be used for quantification ((107) and references therein).

A comparison of DOR and MQMAS techniques for an inorganic sample, shows the line widths of the ^{17}O DOR NMR spectra are narrower than for ^{17}O 3QMAS NMR by a factor of two (34).

The advent of higher magnetic fields, faster magic angle spinning (MAS) and the introduction of techniques to improve the spectral resolution has seen a significant increase in ^{17}O NMR reports from inorganic materials. The resonances are narrower at higher applied magnetic fields because of the inverse dependence of the second order quadrupolar broadening on the applied magnetic field ($\propto$ to B_o^2 in ppm). Progress in ^{17}O solid state NMR applied to biological sciences has been much more limited since ^{17}O NMR from organic materials presents an even greater challenge because typically much greater quadrupole interactions are experienced in such materials. Figure 3 summarizes some the milestones achieved during the last 20 years regarding mainly the study of organic and biologically-related molecules with ^{17}O solid state NMR, showing the relatively slow progresses until the last five years.

DOR has advantages over MQMAS in that no sensitivity is lost in the signal acquisition (in fact sensitivity is gained since the signal are narrower) and it is a 1D technique. The main disadvantage of DOR is that a highly specialised probehead is required. There have been continued improvements in the maximum rotation rates of the outer rotor. Current maximum rates are ~2 kHz, which means that with odd-order sideband suppression (95), sidebands are separated by ~4 kHz. Although still large numbers of sidebands can be produced they can be handled and the much-improved resolution would allow several different sites in the same sample to be resolved in the DOR spectrum.

^{17}O solid state NMR has already developed sufficiently that is it able to resolve simultaneously several oxygen sites, even if they are chemically very similar. For example, MAS experiment provides enough resolution to observe and characterize the four oxygen atoms from crystalline L-glutamic acid (Figure 1). DOR experiments provide even better resolution with all the 5 sites in L-MSG resolved (Figure 2). The increase in the resolution will allow the study of more complex molecules, perhaps even uniformly labeled peptides and more complex organic molecules.

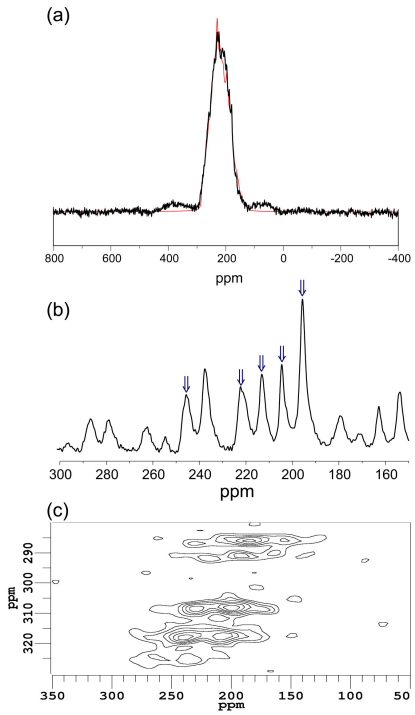


Figure 2 14.1 T ^{17}O NMR spectrum of L-Monosodium glutamate (a) MAS together with a simulation based on the 5 components deduced from the higher resolution data. (b) DOR (⇓ indicate the centrebands) and (c) the centreband of the 14.1 T 3Q MAS NMR data (adapted from (70) with permission of the copyright owner).

Recent reports of ^{17}O NMR from organic materials have included O_2 and CO interacting with hemeproteins, polymers, various crystalline organic molecules, amino acids and nucleic acid bases – these applications will be reviewed here.

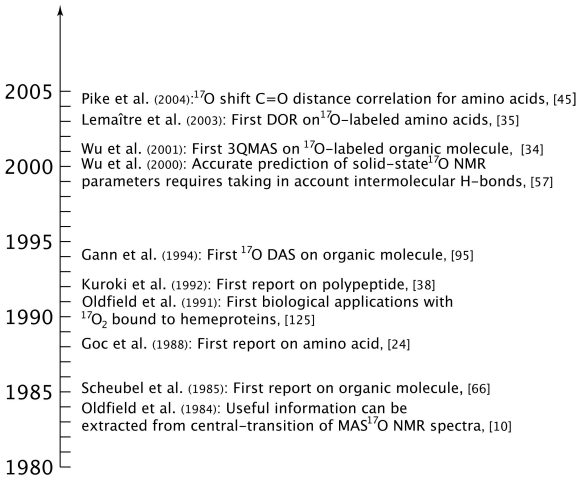


Figure 3 Important milestones in the application of solid state ^{17}O NMR to organic molecules and related materials.

SAMPLE PREPARATION

^{17}O solid state NMR requires selective isotopic enrichment of the sample. However, even low enrichment levels are enough to provide a sufficient increase in signal to determine the NMR parameters of the enriched sites. The lowest enrichment used for ^{17}O of organic molecules reported for NMR study is 6 atom% (6, 61, 63). Considering the current magnetic fields and probes available, and in order to ensure acquisition in a reasonable amount of time, a minimum enrichment of 10-15 atom% for crystals of organic molecules is recommended. For non-crystalline organic phases, higher enrichments are preferable (minimum 40 atom%).

Although it is not necessary to calculate very accurately the enrichment in ^{17}O, it is possible to monitor the enrichment using electron ionization mass spectrometry (EI-MS). As an example, Figure 4 shows the change in the mass spectrum upon enrichment in the case of L-glutamic acid, where enrichment was carried out using 18%-^{17}O-water. Upon enrichment, one observes an increase of the intensity of the peaks located at M+1, M+2, M+3 and M+4, where M is the molecular peak.

This indicates an overall isotopic enrichment of the molecule (in this case, L-Glutamic acid has four oxygen sites). Although the level of enrichment could be higher, this level of enrichment of the water is enough to give a sample suitable for solid state NMR studies.

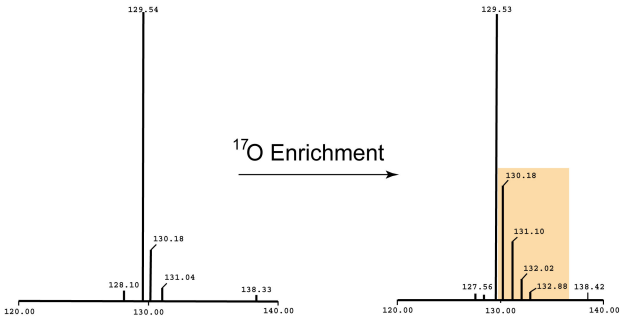


Figure 4 Electron ionisation mass spectrum of (a) L-glutamic acid and ^{17}O-enriched L-glutamic, starting with 18 atom-% ^{17}O-enriched water. The increase of the intensity of mass peak located at the left of the molecular peak shows that the ^{17}O enrichment was achieved.

The main precursors for ^{17}O enrichment are ^{17}O-water, ^{17}O-dioxygen (O_2 gas) or ^{17}O-carbon monoxide. For more information related to synthetic methods, see the review written by Butler (1991) (18) or refer directly to the publication related with the solid state NMR study where the individual enrichment schemes employed are usually discussed (Table 2 for a detailed list). Enrichment can be at specific sites or the sample uniformly labelled. All the evidence, for example from uniformly labelled amino acids, shows that the intensity ratio between the different sites agrees exactly with the expected value.

Synthesis of ^{17}O-enriched samples can be difficult, since the number of labeled precursors on the market remains fairly limited. It is interesting to note that in 2001, the precursors commercially available were mostly limited to H_2O, O_2 and CO. As a sign of the growing interest in ^{17}O labeling, it is now possible to buy ^{17}O-enriched organic molecules such as benzophenone, phenol or even amino acids. There are now two companies providing ^{17}O labeled organic molecules (Icon Isotope Service and Cambridge Isotopes).

Enrichment of amino-acids is possible and has been described more than 20 years ago (109, 110). Synthesis of selectively enriched peptides is possible, which takes place in solution for small di- and tri-peptides (109, 117). Furthermore, solid-phase peptide synthesis (with both Boc- (109) or Fmoc-protecting groups [Lemaitre et al., to be published]) can also be performed. The enrichment of the

protected amino acid can be either performed on the Boc- or MeO- protected amino acid (109) or the protection reaction (FMOC for example) can be conducted on the amino acid without any significant loss of label (86).

CHARACTERISTIC PARAMETERS

Generally, authors have determined as accurately and as completely as possible the NMR parameters (Table 1) characterizing the observed NMR signals. To obtain unambiguous simulations of the NMR spectra, it is usually necessary to record them at several fields (at least 2), or use a combination of methods e.g. static and MAS, DOR or MQ in order to generate enough constraints for the simulation procedure. Advances in both experimental NMR and computation of the NMR parameters have resulted in a much improved correlation between anisotropic NMR properties (EFG, CSA) and molecular structure.

Indirect spin coupling constants will not be discussed here, since there are no reports so far, for an organic molecule in ^{17}O solid state NMR. However, this could turn out to be another very useful parameter, which might be related to structure. Relaxation properties of ^{17}O-labelled solids also need to be investigated in more detail but very wide variations in T_1 have been determined from several seconds to a few ms.

Table 1: NMR parameters of interest that can be extracted from ^{17}O spectra

Parameter	**Symbols usually used**
Isotropic Chemical Shift	δ_{iso}
Chemical Shift Anisotropy Tensor (principal components in PAS)	σ_{11}, σ_{22}, σ_{33}
Orientation of CSA (Euler angles)	α, β, χ
Quadrupole Coupling Constant	$\chi = e^2qQ/h$
Nuclear Quadrupolar Electric Moment	eQ
Quadrupolar Asymmetry Parameter	$\eta_Q = (q_{xx} - q_{yy})/q_{zz}$
Electric field gradient tensor (main components)	V_{11}, V_{22}, V_{33} $(= eq)$

Chemical Shift

Changes in chemical shift can potentially be used to monitor conformational changes, substrate binding, changes in the hydrogen-bonding and protonation state of the molecule studied. ^{17}O chemical shifts are referenced against H_2O.

Since the first application of ^{17}O solid state NMR on MnO_2 by Jackson in 1963 (53), many other molecules have been characterized using this technique. The general trend of ^{17}O chemical shifts is shown in Figure 5, which is the result of the compilation of about 30 reports from the last 20 years or so. This figure is not exhaustive, since the number of organic molecules that have been studied by ^{17}O solid state NMR remains small and thus the examples given may not be representative of the entire chemical shift range of ^{17}O. The identity of the molecules and the individual values for the isotropic chemical shift (and also other parameters) can be found in Table 2. An important point to emphasize here is that the shift of the peak position in a solid state NMR spectrum of a second-order quadrupolar perturbed system does not correspond to the isotropic chemical shift (which is usually the case for solution spectra). If a second-order quadrupolar lineshape can be resolved then the isotropic chemical shift can be deduced from a single spectrum. An alternative approach is to carry out DOR and MQMAS combined with magnetic field variation, and again the isotropic chemical shift deduced. Care has to be taken to ascertain if papers when referring to the chemical shift, mean the peak position or the genuine isotropic value. In examining differences (e.g. from structural effects) it is isotropic chemical shifts that should be compared.

The values of the ^{17}O NMR chemical shifts for organic molecules in the solid state measured to date cover a range from 46 to 375 ppm. Comparison with similar chemical shift ranges measured in solution (75) shows significant differences, up to 100 ppm. This is mainly due to important modifications in the hydrogen-bond network experienced by the different chemical groups upon crystallization. Most of the changes tend to reduce the chemical shift of a specific chemical group from solution to solid state and can be interpreted as an increase of the hydrogen-bond length (vide infra). The correlation between the C-O π bond order (21), observed in solution is also seen in the solid state. Another feature mirroring that in solution, is the small chemical shift range observed for S-O bonds, which cannot be related to the bond order in a similar manner as for C-O bonds.

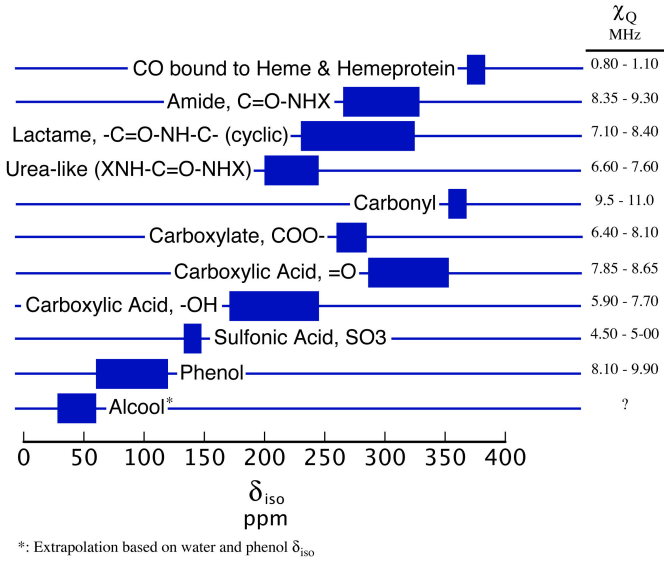


Figure 5 ^{17}O isotropic chemical shift and quadrupole coupling constant ranges for different X-O bonds of organic molecules in the solid state. There are some significant changes between solution (see (41)) and solid state.

Quadrupole Coupling Constant

Solid state NMR allows the determination of parameters like χ_Q and η_Q (Table 2) from powder or single-crystal static experiments, and MAS of polycrystalline samples. However, it is not the only method available to provide information on quadrupole interaction related parameters. A significant number of quadrupole coupling constants of organic molecules have been determined using nuclear quadrupole resonance (NQR) spectroscopy (15, 19, 41, 87, 88, 100-102) (Table 3).

Solid state NMR has shown that experimental χ_Q values cover the range 4.8 to 10.8 MHz for organic molecules in the solid state, while nuclear quadrupole resonance (NQR) suggests values for the components of the quadrupole tensor from –9 MHz to + 11.5 MHz for chemical groups involving C-O bonds (100). The sign of χ_Q gives information on the orientation of the field gradient. Figure 5 summarizes values for quadrupolar coupling constants that have been measured for different chemical functions of crystalline organic molecules. Similarly to the isotropic chemical shifts, significant differences for χ_Q have been observed between the solution and solid state (28, 86). The range over which χ_Q is distributed seems to be narrower in the solid state compared to solution (−8 to 17 MHz) (41), again highlighting some effects of the hydrogen-bond network.

The general rule observed for the solution state (41), that a decrease of the quadrupole coupling constant occurs with an increase of strength of hydrogen bonding, is also a general trend observed in the solid state (vide infra).

Ab initio Calculations

The detailed quantum chemical description of sites in organic molecules is rather complex. The most common approach to simplify the modelling of the nuclear spin interactions is to consider them independently from the rest of the system, using lattice relaxation terms. Other considerations can be used to simplify the modelling process of ^{17}O chemical shift and other relevant parameters. Reports of calculations of the ^{17}O NMR parameters in such materials have been presented to varying degrees of sophistication, including Finite Perturbation theory, Hartree-Fock and Density Functional Theory.

It is often found that inclusion of the hydrogen-bond network is necessary to achieve a good quality of fit, emphasising how sensitive oxygen-17 is regarding this type of interaction.

Wu et al. (2000) (127) used CP MAS to determine the amplitude and orientation of tensor interactions. They carried out DFT calculations with and without inclusion of an intermolecular hydrogen-bond network. Simulations using the software Gaussian (g98) with B3LYP as a level of theory, G-311+FG as a basis set and exchange functions Gauge including Atomic Orbital (GIAO) basis, have been performed for calculating the CS shielding.

There has been significant recent progress in first principles quantum mechanical calculation of NMR parameters that have greatly improved in theory, accuracy and speed with which they can be carried out. Rather than cluster-based calculations density functional theory using the Gauge including projector augmented wave (GIPAW) method (85) uses the true periodic nature of the crystal structure. The crystal structure is refined, by allowing the positions of the hydrogen atoms to relax. In the method, pseudopotentials are used to represent the electron cores. From these calculations both the chemical shielding and quadrupole interactions can be determined. Calculations of these parameters for ^{17}O using this approach have been used for zeolites (89) and glutamic acid polymorphs (132). For L-glutamic acid the quality of the calculations meant that the

peaks could confidently be assigned on the basis of these calculations and illustrates the complementarity between the solid state NMR measurements and the computational work. Computational work will also allow correlations with structural detail to be explored without having to prepare as many enriched samples. Calculations can also help determine the orientation of the chemical shift and quadrupole interactions relative to one another. In powder samples combining variable magnetic field, and static and MAS data allows both the interactions and their orientation to be estimated, especially if ^{1}H decoupling is used to remove dipolar interactions. It is however difficult from this data alone to deduce the orientation of these interactions relative to the molecular frame, whereas single crystal measurements can provide this information.

Assignment of ^{17}O NMR resonances

In most ^{17}O NMR studies of organic molecules to date the assignment of the different resonances has been relatively straightforward, with the number of different sites small - two different signals on average. Furthermore, the crystal structure of the molecule under study is almost always known. For this type of sample, an assignment based on the measurement of the chemical shift from a MAS experiment is usually sufficient, provided the two atoms belong to different chemical functions. As an example, for a molecule containing a single carboxylic group, the difference in NMR parameters between the hydroxyl (–OH, 170-190 ppm, χ_Q = 8.6-8.0 MHz) and the carbonyl (=O, 300-355 ppm, χ_Q = 7.7-7.35 MHz) is usually large enough to avoid any assignment confusion (86). In amino acid·HCl materials the two resonances are completely resolved and the carbonyl and hydroxyl can be immediately identified. However in L-alanine there is complete overlap between the two resonances as the isotropic chemical shifts differ by only 24 ppm so that it is difficult to directly uniquely identify the resonances.

To help assignment of the oxygen sites approaches using the strength of the coupling between the oxygen and surrounding protons have been applied. Typically, to understand the strength of the ^{1}H-^{17}O dipolar coupling, MAS or DOR experiments can be run where the strength of the proton decoupling is varied (86). The DOR experiment allows the best resolution of the different signals (line widths down to 1 ppm (86)). Then, the line broadening as a function of the decoupling strength

allows the assignment of a specific site from the crystal structure. The narrower the line, the less strongly is a particular oxygen atom coupled to the protons.

MQMAS can also be used for helping the attribution of signals to particular sites within the crystal structure (121). It has been shown that the frequency of a signal along F2 increases with a decrease in hydrogen-bond length (based on the carbonyl bond length that increases if the strength of the hydrogen-bond increases).

Another technique used previously in solid state ^{17}O NMR studies was to record spectra at different temperatures: the oxygen involved in the weakest hydrogen-bond network show the more significant changes in its NMR parameters. The method was demonstrated with L-leucine where two types of hydrogen-bonding in the unit cell can be distinguished (46).

All the methods developed so far are limited, in the sense they only allow a small number of different oxygen sites to be dealt with and their chemical environment should ideally differ significantly to allow ready assignment. Furthermore, X-ray or neutron diffraction structures are required in order to determine how each oxygen atom is involved in the hydrogen-bond network. Future work will have to include development of self-consistent NMR methods in order to assign ^{17}O signals in an independent way without relying on previously available structural information, and the computational work mentioned above will play an important role here.

SELECTED APPLICATIONS

Organic Molecules, Nucleic Acid & Amino Acids

Recent reports of ^{17}O NMR from organic materials have included hemeproteins, polypeptides, strongly hydrogen-bonded carboxylic acids, amino acids and nucleic acid bases.

Several classes of molecules have been more extensively studied, amongst them the individual components of DNA and proteins, including the nucleic acids (all of them) and a vast majority of the naturally occurring amino acids. It has to be noted that carbohydrates, an important class of

biologically important molecules with high oxygen content, have not been studied yet. The alcohol group is another chemical functionality that has not yet received much attention, with no solid state NMR report to date. The closest study to this functionality is that of a few phenolic molecules (2-nitrophenol, 4-nitrophenol and tyrosine) (29), characterized by wider signals than for amide groups.

The amide group has been investigated as a simplified model for the amide bond that mediates the link between the different amino acid residues, and is also present in several nucleic acid bases (T, G and U). Wu et al. explored several of them (benzanilide, N-methyl benzamide, acetanilide) (127, 129), determining the chemical shift tensor orientation.

Nucleic acid bases have been studied (121, 122, 124) (Figure 6). Uracil had both one (static and MAS) (122, 124) and two oxygen sites labeled (MAS and MQMAS) (121). From this work all the components of the EFG and CS tensor have been determined for all the nucleobases (amplitude and relative orientation). Furthermore, it has allowed comparison with the same molecules in solution, showing that for some oxygen sites, there are very large changes upon crystallization. The nature of those changes is related with the degree of involvement of each particular oxygen atom in the intermolecular hydrogen-bond network. The oxygen atoms that are able to form intramolecular hydrogen bonds show smaller changes in their NMR parameters upon crystallization. ^{17}O MAS gives line widths up to 10 kHz highlighting the sizable residual quadrupolar width. Static spectra show line widths from 30 to 60 kHz. Accuracy of the simulation is estimated to ± 10 ppm for chemical shift and ± 5° for Euler angles (122, 124).

Amino acids were amongst the first organic molecules to be studied by ^{17}O solid state NMR, beginning as early as 1988 (45, 46). At this time, the resolution was low and it took about 5 hours with a Bruker CXP 180 to collect a spectrum with a S/N of 20 for a sample with an enrichment of 20 atom-% in ^{17}O. In the case of the L-leucine where there are two inequivalent molecules per unit cell, only two inequivalent sites could be found. It was not possible to resolve the two oxygen sites of L-alanine before 1994 and the application of DAS. The EFG and CS tensors in the molecular frame could nevertheless be estimated in the earlier work (45).

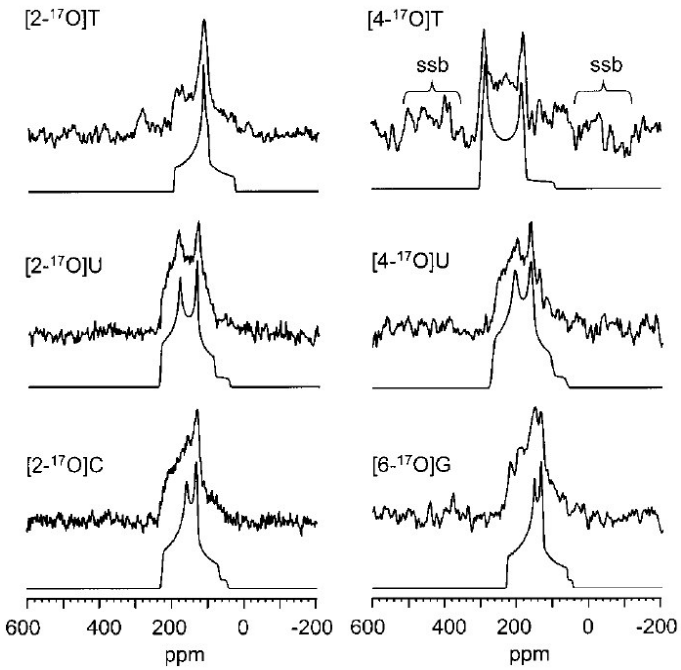


Figure 6 Experimental (upper trace) and simulated (lower trace) ^{17}O MAS NMR spectra of free nucleic acid bases at 11.75 T. Chemical shifts are referenced to liquid H_2O (adapted from (124) with permission of the copyright owner).

All the ^{17}O solid state NMR techniques available have now been applied to amino acids namely, static NMR on powders and single crystals as well as polycrystals, MAS, DAS, DOR and 3QMAS on polycrystalline materials.

Application of higher resolution techniques have highlighted how sensitive the ^{17}O NMR parameters of labeled amino acids are relative to crystallization conditions (pH) (70, 86) or even modification of the amino-group of the molecules (86) (e.g. Fmoc protection of the amino group). In those cases, major changes in the shape and the frequencies of the signals have been observed (Figure 7).

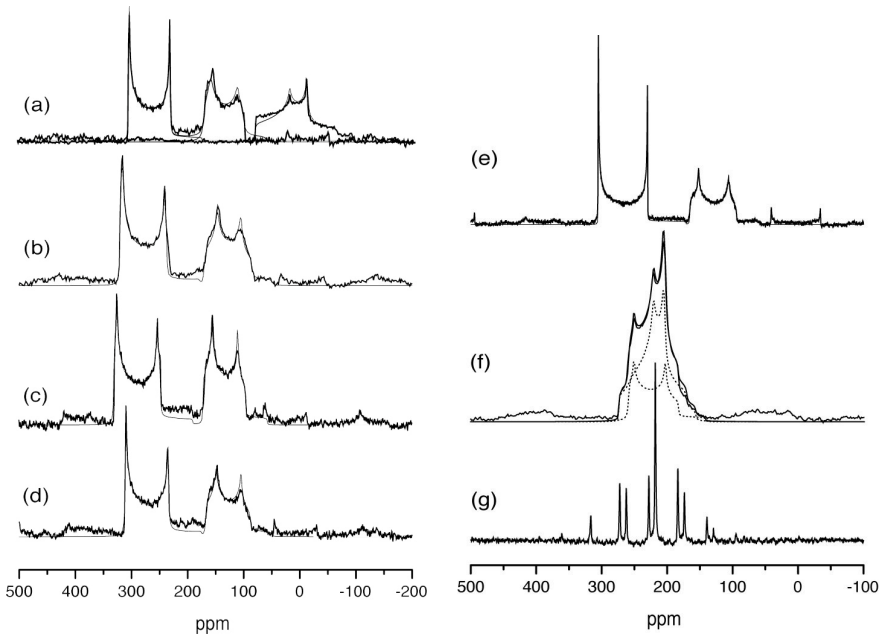


Figure 7 14.1 T ^{17}O MAS NMR spectra of (a) L-tyrosine·HCl, (b) L-asparagine·HCl, (c) L-valine·HCl, (d) glycine·HCl and (e) L-alanine·HCl together with simulations of the centrebands. (f) 14.1 T ^{17}O MAS NMR spectrum of L-alanine together with simulation, (g) DOR spectrum of L-alanine outer rotor speed 1800 Hz (adapted from (86)).

Small peptides and Polymers

^{17}O solid state NMR has been applied to study small peptides (glycylglycine and glycylglycine·HNO_3 (6, 63)) and uniformly labelled polypeptides (polyglycine I and II (6, 61, 63, 130), poly(L-alanine) I and II (114, 131)). Unfortunately, the low enrichment of the samples (6-10 atom-% for the different samples) has led to the observation of numerous artefacts in the observed spectra, many of which were recorded at lower applied magnetic fields.

Kuroki et al. (1994) (63) manufactured samples with a low enrichment, and the spectra accumulated at lower applied magnetic fields (9.39 T, 400 MHz ^{1}H) resulted in the authors reaching an erroneous conclusion about the carbonyl ^{17}O NMR parameters, especially regarding the orientation of the chemical shift tensor in the molecular frame (129). This was corrected in later papers where MAS at 25 kHz was applied (18.8 T, 800 MHz ^{1}H) and the quality of the data significantly improved (130, 131).

One of the most interesting applications of ^{17}O solid state NMR on polypeptides illustrates the ability of the technique to differentiate between parallel and antiparallel chains in the case of a 3_{10}-

helix, whereas ^{13}C and ^{15}N solid state NMR were unable to resolve any differences. Kuroki et al. reported that ^{17}O solid state NMR could be used to detect two inequivalent signals in polyglycine II, that would only be consistent with an antiparallel 3_{10}-helix and not a parallel 3_{10}-helix (61, 63). The two inequivalent oxygen sites were not observed in more recent publications by the same authors, where the same samples were studied at higher magnetic field and with higher MAS rate (131).

Nevertheless, it is possible to determine accurately the ^{17}O NMR parameters on labelled peptides, which constitute the most complex organic samples studied by ^{17}O solid state NMR so far. Indeed, the polypeptides studied are characterized by different degrees of polymerization (leading to a heterogeneous system) and different secondary structures can be obtained. Polyglycine I and II form respectively helix 3_{10}- and β-sheet, while poly(L-alanine) I and II forms are α-helix and β-sheet respectively.

There is no doubt that in future, more samples with this level of complexity, will be studied more routinely by ^{17}O NMR. Small oligopeptides and selectively labelled peptides in both crystalline or non-crystalline phases with a high enough enrichment level constitute an attainable goal for ^{17}O solid state NMR and important structural information could be extracted from such samples (130, 131).

Table 2: ^{17}O NMR parameters of organic molecules in the solid state.

Note: Usually, ^{17}O solid state NMR spectra are referenced relative to natural abundance water at room temperature, for which the chemical shift is arbitrarily set to 0 ppm. This gives a sharp resonance and there is a very small shift dependence on pH but this is not normally taken into account.

		$\delta_{cs,\,iso}$ [ppm]	χ_Q [MHz]	η	σ_{11} [ppm]	σ_{22} [ppm]	σ_{33} [ppm]	α [°]	β [°]	γ [°]	Ref.	Structure Ref.
Organic Molecules												
Benzophenone	O	358±42	10.808	0.369	537±42	525±42	12±42	-	-	-	(98)	(32)
Benzamide	O	300	8.40	0.40	500±3	400±3	0±3	-	-	-	(29)	(39)
Benzanilide	O	320	8.97	0.15	580±5	450±5	-50±5	5	90	70	29, 129)	(56)
N-methylbenzamide	O	287	8.50	0.30	520	380	-40	5	90	78	(129)	(69)
Acetanilide	O	330	8.81	0.20	570	440	-20	0	89	69	(129)	(55)
2-nitro-[^{17}O]-phenol	OH	63	9.90	0.8	-	-	-	-	-	-	(29)	(52)
4-nitro-[^{17}O]-phenol	OH	80	9.70	0.85	-	-	-	-	-	-	(29)	(23)
D,L-Tyrosine·HCl	OH	117	8.10	1.0	-	-	-	-	-	-	(29)	(77)
α-Oxalic Acid	O1 (C=O)	301±14	8.30±0.23	0.07	476±18	413±11	14±13	22±12	26±12	20±1	(133)	(92)
	O2 (C-OH)	183±4	6.68±0.08	0.16	351±4	142±3	55±4	10±1	5±3	9±1		
Benzoic acid	O	230	5.70	1.0	-	-	-	-	-	-	(29)	(105)
Potassium Hydrogen benzoate	O1 (C=O)	290	8.50	0.20	-	-	-	-	-	-	(29)	(106)
	O2 (C-OH)	230	6.30	0.10	-	-	-	-	-	-		
Potassium Hydrogen benzoate	O1 (C=O)	287±2	8.30±0.02	0.23±0.05	470±5	380±5	10±5	0±5	90±5	30±5	(126)	(106)
	O2 (C-OH)	213±2	5.90±0.02	0.55±0.05	370±5	190±5	80±5	5±5	90±5	90±5		
Potassium Hydrogen benzoate	O1 (C=O)	285±5	8.30	0.15±0.01	-	-	-	-	-	-	(123)	(106)
	O2 (C-OH)	215±5	5.90±0.02	0.90±0.01	-	-	-	-	-	-		
Ammonium hydrogen benzoate	O1 (C-OH)	240	6.3	0.40	-	-	-	-	-	-	(29)	(82)
	O2 (C-OHN)	242	6.3	0.0	-	-	-	-	-	-		
Phthalic Acid	O1 (C=O)	312±5	7.2±0.1	0±0.05	-	-	-	-	-	-	(48)	(79)
	O2 (C-OH)	180±5	7.4±0.1	0±0.05	-	-	-	-		-		
Lithium hydrogen phthalate·2 H_2O	O	46±5	5.9±0.1	0.21±0.05	-	-	-	-	-	-	(48)	(60)
Dilithium phthalate hemihydrate	O	272±5	7.6±0.1	0.44±0.05	-	-	-	-	-	-	(48)	NA
Potassium hydrogen phthalate·2 H_2O	O	305±5	8.4±0.1	0.20±0.05	-	-	-	-	-	-	(48)	(80)
p-Toluenesulfonic acid·H_2O	O (S-O)	140	4.79	1.0	-	-	-	-	-	-	(29)	(5)
Urea	O	200±1	7.24±0.01	0.92	-	-	-	0	90	0/90	(28)	(90, 113)
Nucleic Acid Bases												
2-^{17}O-cytosine	O2	230	7.20	0.70	350	300	40	12	80	70	(124)	(12, 74)
6-^{17}O-guanine·H_2O	O6	230	7.10	0.80	395	285	10	5	87	67	(124)	(115)
2-^{17}O-thymine	O2	200±2	6.65±0.02	1.00±0.02	290±5	270±5	20±5	4	90	70	22, 124)	(83)
4-^{17}O- thymine	O4	325±2	8.40±0.02	0.10±0.02	570±5	360±5	20±5	0	84	84	22, 124)	(83)

[2-4-$^{17}O_2$]uracil	O2 O4	240±5 275±5	7.62±0.02 7.85±0.02	0.50±0.01 0.55±0.01	- -	- -	- -	- -	- -	- -	(121)	(111)
2-^{17}O-uracil	O2	245	7.61	0.50	400	330	10	0	89	82	(124)	(111)
4-^{17}O-uracil	O4	275	7.85	0.55	470	350	10	0	90	75	(124)	(111)
Amino Acids												
L-alanine	O	-	6.6	0.55	50	350	350	-	-	-	(45)	(68)
L-alanine	O1 O2	285±8 268±8	8.1±0.3 7.2±0.3	- -	- -	- -	- -	- -	- -	- -	(38)	(68)
L-alanine	O1 O2	284±0.5 260.5±0.5	7.86±0.05 6.53±0.05	0.28±0.02 0.70±0.02	- -	- -	- -	- -	- -	- -	(86)	(68)
L-alanine·HCl	O1 O2	327.8±0.5 176.7±0.5	8.31±0.05 7.29±0.05	0.00±0.02 0.20±0.02	- -	- -	- -	- -	- -	- -	(86)	(25)
Fmoc-L-alanine	O1 O2	303.3±0.5 175.7±0.5	7.89±0.05 6.95±0.05	0.16±0.02 0.12±0.02	- -	- -	- -	- -	- -	- -	(86)	NA
D-alanine	O1 O2	275±5 262±5	7.60±0.02 6.40±0.02	0.60±0.01 0.65±0.01	- -	- -	- -	- -	- -	- -	(121)	NA See (68)
L-cysteine·HCl	O1 O2	353.5±0.5 174.9±0.5	8.65±0.05 7.41±0.05	0.18±0.02 0.27±0.02	- -	- -	- -	- -	- -	- -	(86)	NA
L-glutamic acid·HCl	O1 or O4 O1 or O4 O2 O3	172.5±0.5 187.0±0.5 322.0±0.5 315.0±0.5	7.45±0.05 7.49±0.05 8.16±0.05 8.31±0.05	0.25±0.02 0.25±0.02 0.00±0.03 0.17±0.02	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -	(70)	(103)
D-glutamic acid·HCl	O1 or O4 O1 or O4 O3 O2	172.3±0.5 187.2±0.5 315.4±0.5 322.0±0.5	7.45±0.05 7.49±0.05 8.16±0.05 8.31±0.05	0.25±0.02 0.25±0.02 0.00±0.03 0.17±0.02	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -	(86)	NA See (103)
D,L-glutamic acid·HCl	O1 O2 & O3 O4	170±5 250±5 320±5	7.20±0.02 6.80±0.02 8.20±0.02	0.20±0.01 0.58±0.01 0.00±0.01	- - -	- - -	- - -	- - -	- - -	- - -	(121)	NA
Monosodium L-glutamate·H_2O	1 2 3 4 5	254±1.5 260±1.5 274±1.5 283±1.5 297±1.5	7.4±0.2 7.2±0.2 7.6±0.2 7.7±0.2 7.0±0.2	0.47±0.05 0.50±0.05 0.45±0.05 0.40±0.05 0.45±0.05	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	(70)	(97)
L-glutamine·HCl	O1 O2 O3	319.8±0.5 306±1 180±1	8.20±0.05 8.30±0.1 7.75±0.05	0.03±0.02 0.03±0.03 0.24±0.02	- - -	- - -	- - -	- - -	- - -	- - -	(86)	(104)

Glycine·HCl	O1 O2	336±0.5 185±0.5	8.40±0.05 7.60±0.05	0.00±0.02 0.25±0.02	- -	- -	- -	- -	- -	- -	(86)	(25)
L-isoleucine·HCl	O1 O2	347.1±0.5 182.6±0.5	8.52±0.05 7.40±0.05	0.06±0.02 0.22±0.02	- -	- -	- -	- -	- -	- -	(86)	NA
Leucine·HCl	O1 O2	342.7±0.5 182.6±0.5	8.39±0.05 7.50±0.05	0.05±0.02 0.20±0.02	- -	- -	- -	- -	- -	- -	(86)	NA
L-lysine·HCl	O1 O2	346.7±0.5 180.8±0.5	8.56±0.05 7.67±0.05	0.00±0.02 0.24±0.02	- -	- -	- -	- -	- -	- -	(86)	NA
L-phenylanaline·HCl	O1 O2	353.5±0.5 178.8±0.5	8.54±0.05 7.46±0.05	0.07±0.02 0.25±0.02	- -	- -	- -	- -	- -	- -	(86)	(2)
L-tyrosine·HCl	O1 O2 O3	327.0±0.5 183.0±0.5 83.0±0.5	8.22±0.05 7.35±0.05 8.56±0.05	0.00±0.02 0.19±0.02 0.65±0.02	- - -	- - -	- - -	- - -	- - -	- - -	(86)	(35)
D,L-tyrosine·HCl	O3	117±0.5	8.10±0.05	1.0±0.02	-	-	-	-	-	-	(29)	NA
L-valine·HCl	O1 O2	351±0.5 181±0.5	8.40±0.05 7.35±0.05	0.03±0.02 0.21±0.02	- -	- -	- -	- -	- -	- -	(86)	(58)
Fmoc-L-valine	O1 O2	324.1±0.5 167.3±0.5	8.42±0.05 7.48±0.05	0.08±0.02 0.27±0.02	- -	- -	- -	- -	- -	- -	(86)	NA
Peptides / Polypeptides												
GlyGly	O	265	8.55	0.45	546	382	-132	94	90	-87	(63)	(57)
GlyGly·HNO_3	O	280	8.75	0.47	559	408	-127	94	89	-81	(63)	(91)
Polyglycine I (3_{10}-helix)	O	299	8.55	0.26	574	425	-101	100	91	-79	(63)	(11)
Polyglycine I(3_{10}-helix) (3_{10}-helix)	O	304	8.36	0.30	-	-	-	-	-	-	(130)	(11)
Polyglycine II (β-sheet)	O	299	8.30	0.29	562	410	-108	92	89	-81	(63)	(24, 94)
Polyglycine II (β-sheet)	O	293	8.21	0.33	-	-	-	-	-	-	(130)	(24, 94)
Poly(L-alanine) (α-helix)	O	303	9.28	0.33	595	435	-121	-	-	-	(114)	(4)
Poly(L-alanine) (α-helix)	O	319	8.59	0.28	-	-	-	-	-	-	(131)	(4)
Poly(L-alanine) (β-sheet)	O	265	8.65	0.41	514	390	-110	-	-	-	(114)	(3)
Poly(L-alanine) (β-sheet)	O	286	8.04	0.28	-	-	-	-	-	-	(131)	(3)
Inorganic Molecules Complexed with Organic Molecules or Proteins												
p-toluenesulfonic acid·H_2O	O (H_2O)	30.0±0.5	7.05±0.02	0.0	88±2	1±2	1±2	-	-	-	(125)	(5)
Oxalic acid dihydrate	O (H_2O)	-	6.80±0.20	0.93±0.10	-	-	-	-	-	-	(134)	(93)
$C^{17}O$ bound to Fe(TPP)(CO)(NmeIm)	O (CO)	372	1.0	-	-	-	-	-	-	-	(76)	-
$C^{17}O$ bound to A_0 Myoglobin	O (CO)	372	1.1	-	-	-	-	-	-	-	(76)	-
$C^{17}O$ bound to A_1 Myoglobin	O (CO)	366	0.8	-	-	-	-	-	-	-	(76)	-

Table 3: ^{17}O quadrupole parameters of selected organic molecules in the solid state determined by NQR double resonance studies.

		χ_Q [MHz]	η	Ref
Organic Molecules				
Acetic acid (220K)	O1 (C=O) O2 (C-OH)	±8.108±0.005 −7.280±0.005	0.15±0.01 0.187±0.005	(15)
Acrylic acid (77K)	O1 (C=O) O2 (C-OH)	±7.565±0.005 −7.195±0.005	0.31±0.01 0.166±0.005	(15)
Aspirin (291K)	O1 (C=O) O2 (C-OH)	6.793±0.005 −6.683±0.005	0.551±0.005 0.224±0.005	(88)
Benzoic acid (291K)	O1 (C=O) O2 (C-OH)	5.780±0.007 −5.917±0.005	0.894±0.007 0.736±0.005	(88)
α-Chloroacetic acid (77K)	O1 (C=O) O2 (C-OH)	±8.207±0.010 −7.494±0.005	0.157±0.01 0.216±0.007	(15)
p-chloro-benzoic acid (291K)	O1 (C=O) O2 (C-OH)	6.113±0.007 −6.409±0.005	0.769±0.008 0.387±0.005	(88)
m-chloro-benzoic acid (291K)	O1 (C=O) O2 (C-OH)	6.440±0.020 −6.610±0.020	0.565±0.012 0.250±0.015	(88)
Cyanoacetic acid (77K)	O1 (C=O) O2 (C-OH)	±8.093±0.015 −7.427±0.005	0.028±0.01 0.083±0.005	(15)
Formic acid (77K)	O1 (C=O) O2 (C-OH)	±7.818±0.005 −.900±0.005	0.07±0.02 0.077±0.005	(15)
Fumaric acid (77K)	O1 (C=O) O2 (C-OH)	±8.226±0.010 −7.400±0.010	0.125±0.025 0.185±0.010	(15)
Hydroxyacetic acid (77K)	O1 (C=O) O2 (C-OH)	±8.068±0.015 −7.304±0.010	0.106±0.01 0.236±0.010	(15)
p-hydroxy-benzoic acid (291K)	O1 (C=O) O2 (C-OH)	5.873±0.007 −5.945±0.007	0.820±0.008 0.730±0.007	(88)
m-hydroxy-benzoic acid (291K)	O1 (C=O) O2 (C-OH)	5.756±0.008 −5.996±0.008	0.812±0.005 0.660±0.007	(88)
Imidazolium hydrogen maleate (291K)	O (O-H-O)	6.105±0.015	0.620±0.010	(87)
Isophthalic acid (291K)	O1 (C=O) O2 (C-OH)	7.198±0.010 −6.765±0.010	0.416±0.010 0.155±0.010	(88)
Maleic acid (77K)	O1 (C=O) O2 (C=O) O3 (C-OH) O4 (C-OH)	±7.485±0.010 ±8.593±0.010 −7.000±0.020 −7.233±0.010	0.28±0.01 0.105±0.015 0.04±0.03 0.04±0.02	(15)
Maleic acid (295K)	O1 (C=O) O2 (C=O) O3 (C-OH) O4 (C-OH)	±7.480±0.010 ±8.545±0.010 −6.935±0.010 −7.175±0.010	0.28±0.01 0.085±0.02 0.05±0.02 0.085±0.015	(15)
p-nitro-benzoic acid (291K)	O1 (C=O) O2 (C-OH)	5.985±0.005 −6.166±0.007	0.850±0.008 0.519±0.007	(88)
m-nitro-benzoic acid (291K)	O1 (C=O) O2 (C-OH)	5.950±0.017 −6.385±0.010	0.790±0.015 0.360±0.008	(88)
o-nitro-benzoic acid (291K)	O1 (C-OH) O2 (C-OH)	−5.461±0.020 −6.046±0.020	0.770±0.015 0.450±0.010	(88)
α-Oxalic acid (77K)	O1 (C=O) O2 (C-OH)	±8.455±0.003 −7.545±0.010	0.00±0.01 0.160±0.010	(15)
	O1 (C=O)	±8.471±0.003	0.00±0.01	(15)

α-Oxalic acid (295K)	O1 (C=O) O2 (C-OH)	±8.471±0.003 −7.540±0.010	0.00±0.01 0.175±0.010	(15)
β-Oxalic acid (77K)	O1 (C=O) O2 (C-OH)	±8.110±0.007 −7.350±0.010	0.08±0.02 0.135±0.010	(15)
β-Oxalic acid (291K)	O1 (C=O) O2 (C-OH)	±7.473±0.008 −6.933±0.008	0.237±0.005 0.000±0.020	(88)
β-Oxalic acid (295K)	O1 (C=O) O2 (C-OH)	±7.473±0.010 −6.930±0.010	0.24±0.01 0.00±0.02	(15)
Phthalic acid (291K)	O1 (C=O) O2 (C-OH)	7.078±0.007 −7.173±0.005	0.405±0.008 0.147±0.005	(88)
Potassium hydrogen benzoate (291K)	O (O-H-O)	6.165±0.007	0.591±0.007	(87)
Potassium hydrogen chloromaleate (291K)	O (O-H-O)	6.330±0.010	0.580±0.020	(87)
Potassium hydrogen diaspirinate (291K)	O (O-H-O)	6.214±0.015	0.551±0.015	(87)
Potassium hydrogen di-*m*-chlorobenzoate (291K)	O (O-H-O)	6.248±0.010	0.599±0.008	(87)
Potassium hydrogen di-*p*-chlorobenzoate (291K)	O (O-H-O)	6.213±0.008	0.569±0.007	(87)
Potassium hydrogen diformate (291K)	O1 (C=O)	5.462±0.008	0.833±0.008	(87)
	O2 (C-OH)	5.641±0.008	0.539±0.007	
Potassium hydrogen maleate (291K)	O (O-H-O)	6.074±0.010	0.589±0.020	(87)
Potassium hydrogen oxalate (291K)	O1 (C=O)	6.637±0.015	0.684±0.010	(87)
	O2 (C-OH)	6.911±0.010	0.086±0.010	
Potassium hydrogen phthalate (291K)	O1 (C=O)	6.611±0.015	0.758±0.010	(87)
	O2 (C-OH)	7.133±0.010	0.000±0.010	
Quinolic acid (291K)	O1 (C=O)	6.126±0.015	0.745±0.010	(87)
	O2 (C-OH)	6.277±0.015	0.633±0.010	
Salicylic acid (291K)	O1 (C=O) O2 (C-OH) O3 (C-OH)	−5.319±0.008 −5.825±0.008 9.197±0.020	0.860±0.008 0.624±0.007 0.982±0.010	(88)
Amino Acid				
DL-proline	A1 A2 A3 A4 B1 B2 B3 B4	6.79 6.45 6.17 6.08 7.72 7.78 8.28 8.37	0.170 0.436 0.663 0.703 0.242 0.150 < 0.1 < 0.05	(102)

Carbonyl group: The ^{17}O Chemical shift as a picometer ruler

One over-riding feature of NMR in structural studies is the sensitivity of chemical shift and dipolar coupling constants (D_{ab}) to internuclear distance ($r^3 \propto \frac{1}{D_{ab}}$). Thus for determination of local (over several hundreds of pm) distances, NMR surpasses virtually all other methods, both in accuracy and precision. For example, there is a good and strong (~1200 ppm / Å) linear correlation of the chemical shift (δ_{iso}) of the carbonyl oxygen with the C-O bond length for ^{17}O labelled sites (86) (Figure 8). The correlation is derived from the study of 9 amino acids co-crystallized with HCl and comparison of NMR δ and the distance R(C-O) between the two atoms in the carbonyl moiety of the carboxylic acid group of those amino acids.

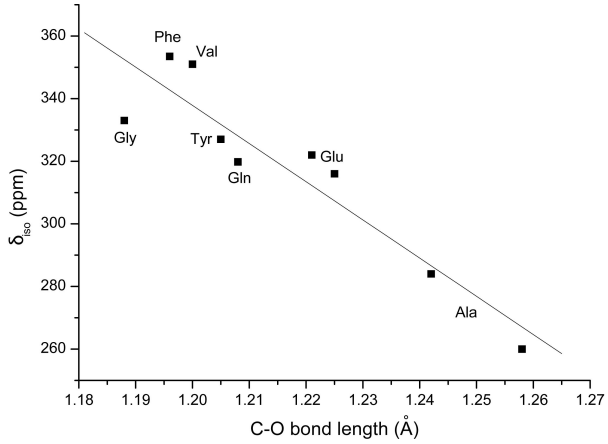


Figure 8 Correlation between shift of the C=O oxygens with the C-O bond length (adapted from (86)).

Study of Hydrogen Bonding

One of the core interests in applying ^{17}O solid state NMR is to probe hydrogen bonding. Oxygen atoms are usually optimally located to monitor any hydrogen-bonds, and due to its quadrupolar nature, it should be very sensitive to any changes of the strength of such interactions. X-ray crystallographers have stated that the N-H···O-C hydrogen bond is a universal feature of amino acid aggregation in the solid state (112). However, it has been now almost 10 years since several authors (63) have tried to correlate ^{17}O NMR parameters to molecular structure and hydrogen-bond length in particular.

So far, there is not a complete understanding between the hydrogen-bond properties (especially the bond-length) and the value of different NMR parameters. General trends can be observed as well as limited correlations between bond-length and ^{17}O NMR parameters. This highlights the complexity of the hydrogen-bond interactions and the importance of the relative orientations between the different partners involved in the hydrogen bond.

In the case of N-H···O=C hydrogen-bonds, it was shown that the quadrupole coupling tensor is a function of the R(N···O) distance between the nitrogen and the oxygen, both theoretically (62) and experimentally (63, 129). However, a simple general correlation has not yet been discerned. The following general trends have, nevertheless, been noted (28, 62, 63, 114, 124, 129-131):

1. e^2qQ/h decreases as the length of the hydrogen bond decreases (see ref. (54) for a definition of hydrogen-bond length) (in other words, if the strength of the hydrogen bond increases);
2. The principal value σ_{33} of ^{17}O chemical shift tensor moves upfield if the length of the hydrogen bond decreases.

In summary, with an increase in hydrogen-bond strength, the shielding of the nucleus increases, so that both chemical shift and χ_Q decrease. Furthermore, it has been shown that virtually all ^{17}O NMR parameters vary upon changes in the local hydrogen bond network (129). Wu and co-workers (28, 124, 127) have highlighted the importance of including the whole intermolecular hydrogen bond network in order to simulate accurately the components of the CS and EFG tensors.

Kuroki and co-workers have studied several dipeptides and polypeptides and tried to correlate e^2qQ/h with the R(N···O) distance between the nitrogen and the oxygen (Figure 9) (63, 130, 131). According to simulations performed by Wu et al. (129) on the amide chemical function, σ_{11} and σ_{22} increase while σ_{33} decreases with the hydrogen bond length. The behaviour of the first two components of the CSA is in agreement with observation from Kuroki and co-workers (63), with the exception of some discrepancies for σ_{33} that might have resulted from a simulation artefact (a small basis set was used – STO-6G).. . . . . (129) asymmetry parameters (η_Q) decrease with hydrogen-bond length increasing hydrogen-bond length, while the quadrupole coupling

constant increases (Figure 10). As a result, χ_Q, η_Q, σ_{11}, σ_{22} and σ_{33} vary with the hydrogen bond-length in systems comprised of very similar molecules crystallizing in slightly different ways (129).

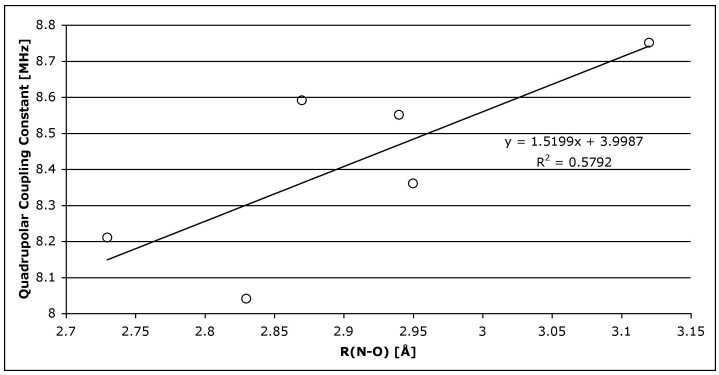


Figure 9 Plots of the e^2qQ/h against the N···O separation (R(N···O) (adapted from the latest data presented in (63, 130, 131)).

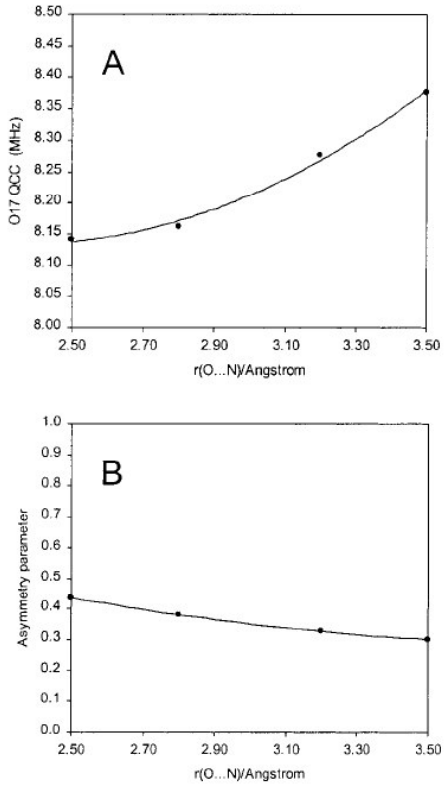


Figure 10 Calculated ^{17}O quadrupole coupling constant (A) and asymmetry parameter (B) versus r(O···N) for the N-methylacetamide/formamide model (adapted from (129) with permission of the copyright owner).

Hydrogen bonding for carbonyl groups often results in large upfield ^{17}O chemical shifts (63). Furthermore, it seems the whole network of the hydrogen bonds and the general geometry of the

local environment (i.e. secondary structure) has also to be taken into consideration. In poly(L-alanine), the α-helix and β-sheet forms show a difference in δ_{iso} of 33 ppm, which is related to the different hydrogen bonding of these two different secondary structures (131).

Solid-solid NMR has been used to study Short Strong Hydrogen Bonds (SSHB or LBHB for Low Barrier Hydrogen Bonds), which can contribute up to 20 kcal/mol stabilization to transition states. With a bond length of ~2.4 Å, lithium hydrogen phthalate monohydrate is a SSHB (O-O distance = 2.4 Å). Its MAS spectrum provides strong evidence for a SSHB due to a low isotropic chemical shift, with lowest reported for a carboxylate (46 ppm) (48). Therefore, ^{17}O solid state NMR can be easily used for identifying low barrier hydrogen bonds. This shows again that quadrupole interaction and chemical shift are strongly affected by hydrogen bonding.

Recently, Wu et al. (126) provided a comprehensive ^{17}O NMR study on potassium hydrogen benzoate. One of its interesting features is a single O-H-O symmetric hydrogen bond. Extensive simulation using both HF and DFT investigated the effect of such a symmetrical hydrogen bond on the ^{17}O NMR parameters. With the O-O distance 2.51Å here, the isotropic chemical shift observed is much larger than that for lithium hydrogen phthalate.

Nuclear quadrupole resonance (NQR) is an alternative technique that can be used to characterize ^{17}O-labelled organic molecules and study the effect of hydrogen-bonding on the ^{17}O electric field gradient tensor (100, 101) (Table 3). Poplett and co-workers studied hydrogen-bonding in carboxylic acids and were the first to notice a dependence between components of the quadrupole coupling tensor and the hydrogen-bond length (15, 87, 88). Seliger (100) has proposed using the principal value V_{33} of the electric field gradient tensor as a parameter reflecting the hydrogen bond strength. In the case of organic solids containing C-O-H···O=C hydrogen bonds, Seliger has showed that it is possible to build a simple model to correlate V_{33} with the length R(O…O) between the two oxygen atoms involved in the bond (Figure 11). This correlation is based on several assumptions that include (1) that the electric field gradient tensor at the oxygen sites is mainly of a relatively local nature, (2) that the electron distribution around the hydrogen nucleus is reflected in the principal value V_{33} of the EFG tensor along the out-of-plane principal direction which is perpendicular to the oxygen-hydrogen bond and (3) that no violent thermal motions are significantly

influencing the ^{17}O EFG tensor (librations, reorientations, hydrogen exchange, etc.). If the EFG tensor is partially averaged due to the thermal motions the correlation becomes invalid. Therefore, this empirical correlation works only at a low enough temperature so that the influence of thermal motions on the EFG tensor can be neglected. Interestingly, the parameters obtained from potassium hydrogen benzoate (126) fit Seliger's correlation for the hydrogen bond length.

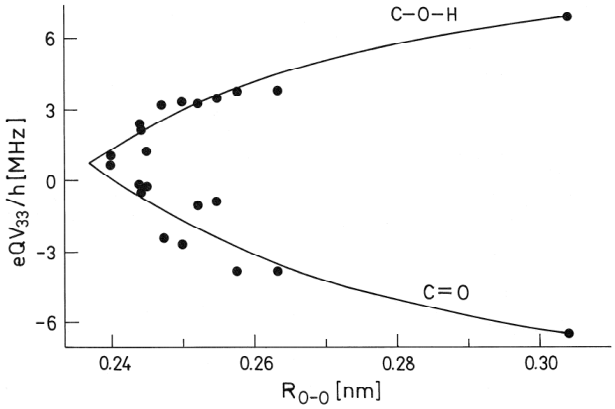


Figure 11: Empirical expression giving a reasonable description of the correlation V_{33} vs. R(O···O) in the range 0.24 nm < R(O···O) < 0.3 nm. The plot of eQV_{33}/h at the C–^{17}O–H and C=^{17}O···H oxygen sites versus the hydrogen bond length R(O···O). The full lines are calculated after: V_{33}(C-O-H) = 9.5 MHz – 27MHz $R_{O\cdots O}^{-3}$ - 1290MHz $R_{O\cdots O}^{-6}$, V_{33}(C=O) = -10.2 MHz + 56MHz $R_{O\cdots O}^{-3}$ + 1400MHz $R_{O\cdots O}^{-6}$, where the hydrogen bond distances is given in 0.1 nm (adapted from (100) with permission of the copyright owner).

With the exception of one report (31), the effects of the temperature on the quadrupole coupling constant has not been studied in organic molecules. By observing the effects of the temperature on the static ^{17}O solid state NMR spectrum of a single-crystal, Filsinger and co-workers characterized a motional process in hydrogen-bonded dimethylmalonic acid, suggesting a 180°-flips of the $(COOH)_2$ units. NQR studies have shown that the temperature has non-negligible effects on the values of the quadrupole coupling constants and the quadrupolar asymmetry parameters (15, 101). In principle variable temperature NMR and changes in the lineshape could be related to motion, although for many of the line shapes so far observed at room temperature there is little direct evidence of any effects of motion.

Studies of ^{17}O magnetic resonance in organic molecules have highlighted the complex dependence between the ^{17}O NMR parameters and molecular structure. The general conclusion is that EFG and chemical shift tensors are very sensitive to hydrogen bonding and can therefore be used to

characterize such interactions. Several empirical correlations (6, 63, 100, 129) matching some ^{17}O NMR parameters and the hydrogen bond length indicate that for some subset of molecules (or subset of hydrogen bonds) there is a relation between several of the ^{17}O NMR parameters and the hydrogen bond length, but these will need to be developed further.

Inorganic ^{17}O-labelled molecules bound to proteins and bilayers

O_2 and CO as ligands of Hemeproteins

Heme model compounds and hemeproteins have been extensively studied by ^{17}O solution NMR (40, 42-44, 64, 65, 84) and, only more recently, ^{17}O solid state NMR has also been used to provide structural data on such systems.

Since the first determination in 1991 of the ^{17}O NMR parameters from a $C^{17}O$ molecule bound to a model compound for heme protein (a "picket fence porphyrin") in the solid state (84) (a spinning speed of only 4.7 kHz - resulted in many spinning side-bands), the quality of the spectra has increased and solid state NMR has provided information about the conformation of not only bound-CO but also bound-O_2 in heme models and metalloproteins.

Oldfield et al. (1991) studied a model compound for oxyhemoglobin and oxymyoglobin, the iron-dioxygen complex of "picket fence porphyrin" using ^{17}O solid state NMR (81). It was possible to determine the ^{17}O NMR parameters (chemical shift tensor and nuclear quadrupole coupling constant) for all the oxygen sites from the porphyrin-bound O_2 molecule. Comparison with previous solution studies (44) showed observable changes in the NMR parameters. An estimation of the Fe-O-O bond angle (140°) for the model system was possible using the partial averaging of the shift tensor at room temperature. Temperature dependences indicate no changes in isotropic chemical shift for terminal oxygen while there is a low-frequency shift for the bridging oxygen – interpreted as the freezing of the system in one conformational substate. The nuclear quadrupole coupling constants are small when compared to O_3, suggesting a π-delocalization on the Fe-O-O fragment.

Oxymyoglobin and oxyhemoglobin themselves could also be observed in frozen solution, showing spectra similar to the model porphyrin.

Since 1998, two studies of metalloporphyrins and metalloproteins, using simultaneously ^{13}C and ^{17}O solid state NMR, ^{57}Fe Mössbauer and infrared vibrational spectroscopic techniques, and DFT quantum simulation to probe various Fe-O_2 or Fe-CO conformations, have provided accurate structural information on the conformation of the bound molecules relative to the Fe centre. McMahon et al. (1998) studied how a CO ligand binds to iron in metalloporphyrins and metalloproteins (76) (extending Park et al.'s work that was mainly focused on the use of ^{17}O solution NMR (84)). This approach using DFT to test several models against the spectroscopic data, was the first detailed quantum chemical analysis of metal-ligand geometries in metalloproteins, and has allowed the determination of the most probable ligand tilt and bend angles. This is a useful complement to X-Ray diffraction, where the resolution on the ligand (CO) could not sufficiently constrain its orientation (76) (resulting in a large distribution of angles from one crystal structure to another). This study shows that all spectroscopic methods taken together converge to a single conformation, very close to linear and untilted Fe-C-O geometry for all carbonmonoxyheme proteins. The same approach was applied to study several Fe-O_2 analogue metalloporphyrins (47). NMR results are well reproduced by DFT, enabling testing of various models of Fe-O_2 bonding in the porphyrin and metalloproteins. No evidence was found for two binding sites in an oxypicket fence porphyrin. Oxygen sites in complexes are more electronegative than that in the CO system, which strongly supports the idea that H-bonding to O_2 is a major contributor to O_2/CO discrimination in hemeproteins.

Molecules of water interacting with bilayers

Bilayers are by their very nature at the border between liquid and solids. Water molecules, which constitute an important part of such lyotropic phases, experience an anisotropic environment through the interactions with the bilayers and thus anisotropic molecular motions, resulting in the quadrupolar interaction not being averaged out to zero. Tricot et al. (1979) (116) compared ^{17}O and ^{2}H NMR spectra of water (both $^{2}H_2O$ and $H_2{}^{17}O$) in fully hydrated lamellar phase of DPL (dipalmitoyl-3sn-phosphatidylcrioline). Ordering of water molecule results in a quadrupolar

splitting which is related to the order parameter. In the case of ^{17}O, the quadrupolar splitting observed is larger than for ^{2}H.

FUTURE PROSPECTS

Use of ^{17}O as a probe for drug-receptor interaction

Several of the organic molecules studied as crystals also have other properties, such as interacting with one or several receptors or enzymes. Providing that ^{17}O solid state NMR has sufficient sensitivity (i.e. high magnetic field and high ^{17}O enrichment) to measure a signal, it should be possible to characterize the local electronic environment sensed by the ^{17}O labels of such molecules while interacting at the binding site / pocket of the protein. Information relative to the protonation-state of the bond-ligand could also be provided. Of course, information on the existence hydrogen-bonds between the ligand and a qualitative picture of the strength of the hydrogen-bonds involved could also be available.

Distance measurements: measuring distances between ^{1}H-^{17}O and ^{13}C-^{17}O nuclei

There has been very little work using ^{17}O NMR to determine distance. Linder *et al.* have indicated the possibility of correlating the quadrupole and dipole tensor by quadrupole separated local field experiments on powder samples (71). This has been applied to the ^{1}H-^{17}O system in $Mg(OH)_2$ and $Mg(OH)_x(OCH3)_{1-x}$ (118). Lee-Goldburg decoupling during the dipolar evolution provides better resolution in the dipolar dimension. The normal chemical shift dimension produces the second-order quadrupolar lineshape and the dipolar dimension allows the dipolar interaction to be calculated. The 2D data contain more information since the intensity distribution depends critically on the relative orientation of the two tensors. Since the dipolar interaction is axially symmetric, only the azimuthal and polar angles are necessary to describe the relative orientation, and the intensity distribution changes markedly with this orientation (Figure 12).

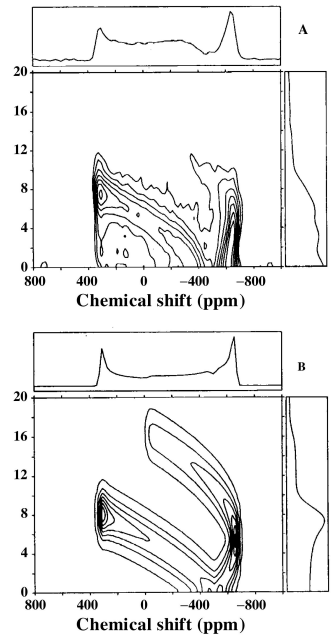


Figure 12: A quadrupole-dipole separated local field experiment for the OH group in $Mg(OH)_x(OCH_3)_{2-x}$ with A. the experimental data compared with B. the simulation of the intensity assuming the tensors are collinear from van Eck and Smith (1998) (118) with permission of the copyright owner.

In 2003, Hu et al. (49) presented the application of a 2D static NMR experiment (Figure 13) to provide a good approximation of the ^{1}H-^{17}O distance in ice at 90 K. To allow ready interpretation of the static experiment, it is best applied to systems with a single ^{1}H-^{17}O bond or hydrogen bond can be studied accurately. The method could in principle be applied to α-helix forming transmembrane peptides in which a single ^{17}O-labelled amino acid has been inserted and providing a single ^{1}H-^{17}O hydrogen bond (Figure 14). This approach might be able to focus on local deformation of the helix induced by the membrane.

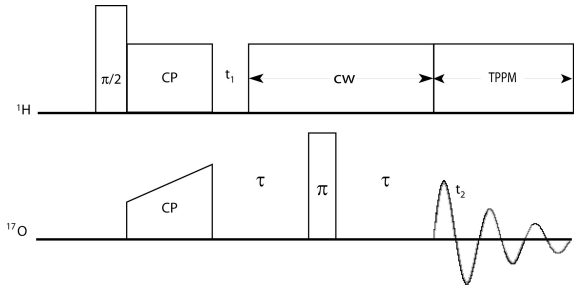


Figure 13: CP-Echo pulse sequence applied to a glycerol/^{17}O-water by Hu et al. (2003) (49) in order to estimate the average O-H distance in ice.

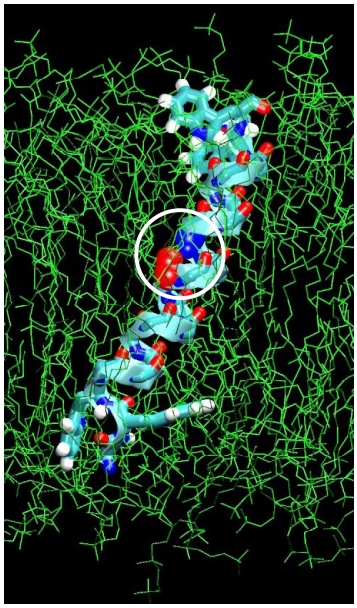


Figure 13: Example where a dipole-quadrupole correlation 2D static NMR experiment could be applied on a biological sample.

Developing robust MAS or DOR pulse sequences which allow the selective reintroduction of the dipolar interaction in order to measure distances between a ^{17}O nucleus and another nucleus of interest will be a very important step forward. REDOR and REAPDOR have already been demonstrated in the case of ^{13}C-^{17}O internuclear distances (20) for ^{17}O labelled water co-crystallized with L-asparagine and other S = 1, I = 1/2 I-S internuclear distances (50). Obviously more work in this direction should be carried out in order to have robust methodologies that can be applied to biological samples. Cross-polarisation (CP) can also be used to discriminate different sites as those with the protons in closer proximity cross-polarise more strongly at shorter contact times. This principle has been illustrated for ^{1}H-^{17}O for both conventional CP and generation of multiple quantum signal (9, 119).

With the demonstration that ^{17}O solid state NMR provides enough sensitivity to study selectively enriched peptides in a non-crystalline phase or inserted in vesicles, the development of efficient pulses sequences to measure internuclear distances between pairs of nuclei including an ^{17}O, should be the next logical step in order to realise the full potential of the technique.

Synthetic challenges

Not all the chemical functionalities involving oxygen are equally amenable to ^{17}O enrichment. For carboxylic groups, the exchange is relatively straightforward. Other chemical groups, like alcohol are more difficult to enrich and therefore much less common in the literature.

Synthesis of selectively enriched peptides is possible and is also compatible with solid-phase peptide synthesis (109). However, the techniques used for peptide synthesis prevent the insertion of ^{17}O labels in the side-chains of the amino acids, due to the chemical protection required for the synthesis to take place. Currently, the only way available to synthesize a peptide with ^{17}O-labelled side-chains would consist of adding a supplementary ^{17}O-enrichement step after purification of the peptide, increasing even more the cost of the sample. The preparation of such a sample has not yet been reported.

CONCLUSIONS

NMR of ^{17}O-labeled biomaterials is very sensitive to changes in the local bonding environment. Indeed, the chemical shift and quadrupole coupling constant change significantly in different protonation states of a same molecule or very similar molecules forming slightly different crystal forms.

The ability to resolve the different sites and to detect subtle bonding changes through ^{17}O NMR spectra suggests it has potential to become one of the most significant probe techniques for biochemical investigation of ligand-receptor interactions as it is a non-perturbing approach. Furthermore, molecular size is in principle not limiting (compare to solution) and crystallinity is not a requirement.

ACKNOWLEDGEMENTS

First and foremost, we would like to acknowledge Ray Dupree's stimulus, continued interest and experienced guidance in engaging us in studies of oxygen-17 as a useful and informative nucleus

for application in biology. VL thanks BBSRC for Industrial CASE Scholarship. The BBSRC, MRC, EPSRC and HEFCE are acknowledged for their support to AW for this work. AW was in receipt of a Professeral Research Fellowship, BBSRC. MES thanks EPSRC, HEFCE and the University of Warwick for supporting NMR equipment at Warwick and the Royal Society and Leverhulme Trust for a senior fellowship permitting him to start work on ^{17}O NMR in organic materials.

REFERENCES

1. Abragam A. 1961. *Principles of Nuclear Magnetism.* Oxford: Clarendon Press
2. Al-Karaghouli AR, Koetzle TF. 1975. Neutron diffraction study of L-phenylalanine hydrochloride. *Acta Crystallographica Section B* 31: 2461
3. Amott S, Dover SD, Elliot A. 1967. *Journal of Molecular Biology* 30: 201
4. Amott S, Wonacott AL. 1966. *Journal of Molecular Biology* 21: 371
5. Arora SK, Sundaralingam M. 1971. The crystal and molecular structure of 4-methyl sulfonic acid (p-toluenesulfonic acid) monohydrate, C7H8SO3-.H3O+, an oxonium salt. *Acta Crystallographica Section B* 27: 1293
6. Asakawa N, Kameda T, Kuroki S, Kurosu H, Ando S, et al. 1998. Structural Studies of Hydrogen-Bonded Peptides and Polypeptides by Solid-State NMR. *Annual Reports on NMR Spectroscopy* 35: 55
7. Ashbrook SE, Berry AJ, Wimperis S. 2001. ^{17}O multiple-quantum MAS NMR study of high-pressure hydrous magnesium silicates. *J. Am. Chem. Soc.* 123: 6360–6
8. Ashbrook SE, Berry AJ, Wimperis S. 2002. *Journal of Physical Chemistry B* 106: 773
9. Ashbrook SE, Wimperis S. 2000. Multiple-quantum cross-polarization and two-dimensional MQMAS NMR of quadrupolar nuclei. *Journal of Magnetic Resonance* 147: 238
10. Ashbrook SE, Wimperis S. 2002. Satellite-transition MAS NMR of spin I = 3/2, 5/2, 7/2, and 9/2 nuclei: sensitivity, resolution, and practical implementation. *Journal of Magnetic Resonance* 156: 269
11. Astbury WT, Dalgleish CH, Darmon SE, Sutherland GBBM. 1948. *Nature* 161: 780
12. Barker DL, Marsh RE. 1964. The crystal structure of cytosine. *Acta Crystallographica* 17: 1581
13. Behrens H, Schnabel B. 1982. *Physica B* 114: 185
14. Bennett AE, Rienstra CM, Auger M, Lakshmi KV, Griffin RG. 1995. Heteronuclear decoupling in rotating solids. *Journal of Chemical Physics* 103: 6951
15. Brosnan SGP, Edmonds DT, Poplett IJF. 1981. An 17O Nuclear Quadrupole Resonance Study of Some Carboxylic Acids. *Journal of Magnetic Resonance* 45: 451
16. Bull LM, Bussemer B, Anupõld T, Reinhold A, Samoson A, et al. 2000. A High-Resolution ^{17}O and ^{29}Si NMR Study of Zeolite Siliceous Ferrierite and ab Initio Calculations of NMR Parameters. *J Am Chem Soc* 122: 4948
17. Bull LM, Cheetham AK, Samoson A, Anupõld T, Reinhold A, et al. 1998. A high resolution 17O NMR study of siliceous zeolite faujasite. *J Am Chem Soc* 120: 3510
18. Butler LG. 1991. The NMR Parameters for Oxygen-17. In *17O NMR in Organic Chemistry*, ed. D Boykin, pp. 1. Boca Raton: CRC Press
19. Cheng CP, Brown TL. 1979. Oxygen-17 Nuclear Quadrupole Double Resonance Spectroscopy. 1. Introduction. Results for Organic Carbonyl Compounds. *J Am Chem Soc* 101: 2327
20. Chopin L, Vega S, Gullion T. 1998. A MAS NMR Method for Measuring 13C-17O Distances. *J Am Chem Soc* 120: 4406

21. Christ HA, Diehl P, Schneider HR, Dahn H. 1961. *Helvetica Chemica Acta* 44: 865
22. Cleland WW, Kreevoy MM. 1994. *Science* 264: 1887
23. Coppens P, Schmidt GM. 1965. The crystal structure of the -modification of p-nitrophenol near 90° K. *Acta Crystallographica* 18: 62
24. Crich FHC, Rich A. 1955. *Nature* 176: 780
25. di Blasio B, pavone V, Pedone C. 1977. *Cryst. Struct. Comm.* 6: 745
26. Dirken PJ, Kohn SC, Smith ME, van Eck ERH. 1997. Complete resolution of Si---O---Si and Si---O---Al fragments in an aluminosilicate glass by ^{17}O multiple quantum magic angle spinning NMR spectroscopy. *Chemical Physics Letters* 266: 568–574
27. Dirken PJ, Smith ME, Whitfield HJ. 1995. $^{17}0$ and ^{29}Si Solid State NMR Study of Atomic Scale Structure in Sol-Gel-Prepared TiO_2-$Si0_2$ Materials. *Journal of Physical Chemistry* 99: 395–401
28. Dong S, Ida R, Wu G. 2000. A combined Experimental and Theoretical 17O NMR Study of Crystalline Urea: An Example of Large Hydrogen-bonding Effects. *Journal of Physical Chemistry A* 104: 11194
29. Dong S, Yamada K, Wu G. 2000. Oxygen-17 Nuclear Magnetic Resonance of Organic Solids. *Zeitschrift für Naturforschung* 55a: 21–28
30. Du LS, Stebbins JF. 2003. *Journal of Physical Chemistry B* 107: 10063
31. Filsinger B, Zimmermann H, Haeberlen U. 1992. 17O NMR Evidence for 180°-Flips of (COOH)2 Units in Dimethylmalonic Acid. *Molecular Physics* 76: 157
32. Fleischer EB, Sung N, Hawkinson S. 1968. The Crystal Structure of Benzophenone. *Journal of Physical Chemistry* 72: 4311
33. Freude D. 2000. Quadrupole Nuclei in Solid-State NMR. In *Encyclopedia of Analytical Chemistry*, ed. R Meyers, pp. 12188. Chichester: John Wiley & Sons
34. Freude D, Kärger J. 2002. NMR Techniques. In *Handbook of Porous Solids*, ed. F Schüth, KSW Sing, J Weitkamp, pp. 465. Weiheim: Wiley-VCH
35. Frey MN, Koetzle TF, Lehmann MS, Hamilton WC. 1973. Precision neutron diffraction structure determination of protein and nucleic acid components. X. A comparison between the crystal and molecular structures of L-tyrosine and L-tyrosine hydrochloride. *Journal of Chemical Physics* 58: 2547
36. Frydman L, Harwood JS. 1995. Isotropic Spectra of Half-Integer Quadrupolar Spins from Bidimensional Magic-Angle Spinning NMR. *J Am Chem Soc* 117: 5367
37. Gan Z. 2000. Isotropic NMR Spectra of Half-Integer Quadrupolar Nuclei Using Satellite Transitions and Magic-Angle Spinning. *J Am Chem Soc* 122: 3242
38. Gann SL, Baltisberger JH, Wooten EW, Zimmermann H, Pines A. 1994. Cross Polarization and Dynamic-Angle Spinning of ^{17}O in L-Alanine. *Bull. Magn. Reson.* 16: 68–72
39. Gao Q, Jeffrey GA, Ruble JR, McMullan RK. 1991. A single-crystal neutron diffraction refinement of benzamide at 15 and 123 K. *Acta Crystallographica Section B* 47: 742
40. Gerothanassis IP. 1994. *Progress in NMR Spectroscopy* 26: 239
41. Gerothanassis IP. 1995. Oxygen-17. In *Encyclopedia of NMR*, ed. D Grant, K Harris, pp. 3430. New-York: John Wiley & Sons
42. Gerothanassis IP, Loock B, Momenteau M. 1992. *Chemical Communications*: 598
43. Gerothanassis IP, Momenteau M. 1987. *J Am Chem Soc* 109: 6944
44. Gerothanassis IP, Momenteau M, Loock B. 1989. Hydrogen-Bond Stabilization of Dioxygen, Conformation Excitation, and Autoxidation Mechanism in Hemoprotein Models As Revealed by 170 NMR Spectroscopy. *J Am Chem Soc* 111: 7006
45. Goc R, Ponnusamy E, Tritt-Goc J, Fiat D. 1988. ^{17}O n.m.r. studies of amino acids in the solid state, in single- and polycrystalline forms. *Int. J. Pept. Protein Res.* 31: 130–136
46. Goc R, Tritt-Goc J, Fiat D. 1989. ^{17}O NMR Study of Polycrystalline *L*-Leucine. *Bulletin of Magnetic Resonance* 11: 238–243

47. Godbout N, Sanders KL, Salzmann R, Havlin RH, Wojdelski M, Oldfield E. 1999. Solid-State NMR, Mössbauer, Crystallographic, and Density Functional Theory Investigation of Fe-O_2 and Fe-O_2 Analogue Metalloporphyrins and Metalloproteins. *J. Am. Chem. Soc.* 121: 3829–3844
48. Howes AP, Jenkins R, Smith ME, Crout DHG, Dupree R. 2001. Influence of low-barrier hydrogen bonds on solid state ^{17}O NMR spectra of labelled phthalate species. *Chemical Communications*: 1448–1449
49. Hu KN, Iuga D, Griffin RG. 2003. *Dynamic Nuclear Polarization Enhanced 17O Solid State NMR Spectroscopy on Biological Solids*. Presented at 44th Exprerimental NMR Conference, Savannah, USA
50. Hughes E, T. G, Goldbourt A, Vega S, Vega AJ. 2002. Internuclear Distance Determination of S = 1, I = 1/2 Spin Pair Using REAPDOR NMR. *Journal of Magnetic Resonance* 156: 230
51. Hussin R, Holland D, Dupree R. 1998. Does six-coordinate germanium exist in Na2O-GeO2 glasses? An oxygen-17 NMR investigation. *J. Non. Cryst. Sol.* 234: 440
52. Iwasaki F, Kawano Y. 1978. The crystal and molecular structure of o-nitrophenol. *Acta Crystallographica Section B* 34: 1286
53. Jackson JA. 1963. *Journal of Physical Chemistry* 24: 591
54. Jeffrey GA. 1997. *An Introduction to Hydrogen Bonding*. New York: Oxford University Press
55. Johnson SW, Eckert J, Barthes M, McMullan RK, Muller M. 1995. Crystal Structure of Acetanilide at 15 and 295 K by Neutron Diffraction. Lack of Evidence for Proton Transfer along the N-H***O Hydrogen Bond. *Journal of Physical Chemistry* 99: 16253
56. Kashio S, Ito K, Haisa M. 1979. *Bulletin of the Chemical Society of Japan* 52: 365
57. Kevick Å, Al-Karaghouli AR, Koetzle TF. 1977. Deformation electron density of -glycylglycine at 82 K. I. The neutron diffraction study. *Acta Crystallographica Section B* 33: 3796
58. Koetzle TF, Golic L, Lehmann MS, Verbist JJ, Hamilton WC. 1974. Precision neutron diffraction structure determination of protein and nucleic acid components. XV. Crystal and molecular structure of the amino acid L-valine hydrochloride. *Journal of Chemical Physics* 60: 4690
59. Kopple KD. 1981. On the Location of Beta-Turns. *Biopolymers* 20: 1913
60. Küppers H, Takusagawa F, Koetzle TF. 1985. *Journal of Physical Chemistry* 82: 5636
61. Kuroki R, Ando I, Shoji A, Ozaki T. 1992. A Structural Study of Polyglycine II in the Solid State by ^{17}O CP MAS NMR Spectroscopy. *Chemical Communications*: 433–434
62. Kuroki S, Ando S, Ando I. 1995. An MO study of nuclear quadrupolar coupling constant and nuclear shielding of the carbonyl oxygen in solid peptides with hydrogen bonds. *Chem. Phys.* 195: 107–116
63. Kuroki S, Takahashi A, Ando I, Shoji A, Ozaki T. 1994. Hydrogen-bonding Structural Study of Solid Peptides and Polypeptides containing a Glycine Residue by ^{17}O NMR Spectroscopy. *J. of Mol. Struct.* 323: 197–208
64. Lee HC, Cummings K, Hall K, Hager LP, Oldfield E. 1988. Oxygen-17 nuclear magnetic resonance spectroscopic studies of carbonmonoxyperoxidases. *Journal of Biological Chemistry* 263: 16118
65. Lee HC, Oldfield E. 1989. Oxygen-17 Nuclear Magnetic Resonance Spectroscopic Studies of Carbonmonoxy Hemoproteins. *J Am Chem Soc* 111: 1584
66. Lee K, Anderson WA. 1980. Nuclear Spins, Moments and Magnetic Resonance Frequencies. In *CRC Handbook of Chemistry and Physics*, ed. RC Weast, MJ Astle, pp. E70. Boca Raton: CRC Press
67. Lee SK, Stebbins JF. 2000. *Journal of Physical Chemistry B* 104: 4091
68. Lehmann M, Koetzle T, Hamilton W. 1972. Pecision neutron diffraction structure determination of protein and nucleic acid components. VIII: the crystal and molecular structure of the beta-forms of the amino acid L-glutamic acid. *J. Cryst. Mol. Struct.* 2: 225
69. Leiserowitz L, Tuval M. 1978. Molecular-packing modes. N-

Methylamides. *Acta Crystallographica Section B* 34: 1230

70. Lemaitre V, Pike KJ, Watts A, Anupõld T, Samoson A, et al. 2003. New Insights into the Bonding Arrangements of L- and D-Glutamates from Solid State 17O NMR. *Chemical Physics Letter* 371: 91
71. Lindner M, Höhener A, Ernst RR. 1980. *Journal of Physical Chemistry* 73: 4959
72. Llor A, Virlet J. 1988. Towards High-Resolution NMR Of More Nuclei In Solids - Sample Spinning With Time-Dependent Spinner Axis Angle. *Chemical Physics Letter* 152: 248
73. MacKenzie KJD, Smith ME. 2002. *Multinuclear Solid State NMR of Inorganic Materials*. Oxford: Pergamon Press
74. McClure RJ, Craven BM. 1973. New investigations of cytosine and its monohydrate. *Acta Crystallographica Section B* 29: 1234
75. McFarlane W, McFarlane CE. 1987. Chapter 14. In *Multinuclear NMR*, ed. J Mason: Plenum Press
76. McMahon MT, deBios AC, Godbout N, Salzmann R, Laws DD, et al. 1998. An Experimental and Quantum Chemical Investigation of CO Binding to Heme Proteins and Model Systems: A Unified Model Based on 13C, 17O, and 57Fe Nuclear Magnetic Resonance and 57Fe Mössbauer and Infrared Spectroscopies. *J Am Chem Soc* 120: 4784
77. Mostad A, Rumming C. 1973. *Chemica Acta Scandinavia* 27: 401
78. Mueller KT, Sun BQ, Chingas GC, Zwanziger JW, Terao T, Pines A. 1990. Dynamic-Angle Spinning of Quadrupolar Nuclei. *Journal of Magnetic Resonance* 86: 470
79. Nowacki W, Jaggi H. 1957. *Zeitschrift für Crystallography* 109: 272
80. Okaya Y. 1965. *Acta Crystallographica* 19: 879
81. Oldfield E, Lee HC, Coretsopoulos C, Adebodun F, Park KD, et al. 1991. Solid-State Oxygen-17 Nuclear Magnetic Resonance Spectroscopy Studies of [17O2] Picket Fence Porphyrin, Myoglobin, and hemoglobin. *J Am Chem Soc* 113: 8680
82. Oxton IA, Cameron TS, Knop O, McCulloch AW. 1977. *Canadian Journal of Chemistry* 55: 3831
83. Ozeki K, Sakabe N, Tanaka J. 1969. The crystal structure of thymine. *Acta Crystallographica Section B* 25: 1038
84. Park KD, Guo KM, Adebodun F, Chiu ML, Sligar SG, Oldfield E. 1991. Distal and proximal ligand interactions in heme proteins: correlations between C-O and Fe-C vibrational frequencies, oxygen-17 and carbon-13 nuclear magnetic resonance chemical shifts, and oxygen-17 nuclear quadrupole coupling constants in C17O- and 13CO-labeled species. *Biochemistry* 30: 2333
85. Pickard CJ, Mauri F. 2001. All-electron magnetic response with pseudopotentials: NMR chemical shifts. *Phys. Rev. B. Condensed Matter* 63: 245101
86. Pike KJ, Lemaître V, Kukol A, Anupõld T, Samoson A, et al. 2004. Solid-State ^{17}O NMR of Amino Acids. *Journal of Physical Chemistry B* 108: 9256–9263
87. Poplett IJF, Sabir M, Smith JAS. 1981. 17O and 2H Quadrupole Double Resonance in Some Acid Carboxylated. *Journal of Chemical Society, Faraday Transaction 2* 77: 1651
88. Poplett IJF, Smith JAS. 1981. 17O and 2H Quadrupole Double Resonance in Some Carboxylic Acid Dimers. *Journal of Chemical Society, Faraday Transaction 2* 77: 1473
89. Profeta M, Mauri F, Pickard CJ. 2003. Accurate first principles prediction of 17O NMR parameters in SiO2: assignment of the zeolite ferrierite spectrum. *J Am Chem Soc* 125: 541
90. Pryor AW, Sanger PL. 1970. Collection and interpretation of neutron diffraction measurements on urea. *Acta Crystallographica Section B* 26: 543
91. Rao NS, Parthasarathy R. 1973. Structure and conformational aspects of the nitrates of amino acids and peptides. I. Crystal structure of glycylglycine nitrate. *Acta Crystallographica Section B* 29: 2379
92. Sabine TM, Cox GW, Craven BM. 1969. *Acta Crystallographica Section B* B25: 2437

93. Sabine TM, Cox GW, Craven BM. 1969. A neutron diffraction study of -oxalic acid dihydrate. *Acta Crystallographica Section B* 25: 2437
94. Saito H, Ando I. 1989. *Annu. Rep. NMR Spectrosc.*: 210
95. Samoson A, Lippmaa E. 1989. Synchronized Double-Rotation NMR-Spectroscopy. *Journal of Magnetic Resonance* 84: 410–416
96. Samoson A, Lippmaa E, Pines A. 1988. High-Resolution Solid-State NMR Averaging of 2nd-Order Effects by Means of a Double-Rotor. *Molecular Physics* 65: 1013
97. Sano C, Nagashima N, Kawakita T, Iitaka Y. 1989. Crystal and molecular structures of monosodium L-glutamate monohydrate. *Analytical Sciences* 5: 121–122
98. Scheubel W, Zimmermann H, Haeberlen U. 1985. 17O Quadrupole Coupling and Nuclear Magnetic Shielding Tensors in Benzophenone. *Journal of Magnetic Resonance* 63: 544
99. Schramm S, Oldfield E. 1984. *J Am Chem Soc* 106: 2502
100. Seliger J. 1998. 17O Quadrupole Coupling in C-O-H…O=C Hydrogen Bonds. *Chemical Physics* 231: 81
101. Seliger J, Zagar V. 1998. 17O Nuclear Quadrupole Resonance Study of Proton Disorder in Solid Benzoic Acid, 4-Hydroxybenzoic Acid and 4-Nitrobenzoic Acid. *Chemical Physics* 234: 223
102. Seliger J, Zagar V, Blinc R, Kadaba PK. 1990. ^{1}H-^{17}O Nuclear Quadrupole Double Resonance in *D,L*-Proline. *Zeitschrift für Naturforschung* 45a: 733–735
103. Sequeira A, Rajagopal H, Chidambaram R. 1972. A neutron diffraction study of the structure of L-glutamic acid.HCl. *Acta Cryst.* B28: 2514–2519
104. Shamala N, Venkatesan K. 1972. *Cryst. Struct. Commun.* 1: 227
105. Sim GA, Robertson JM, Goodwin TH. 1955. The Crystal and Molecular Structure of Benzoic Acid. *Acta Crystallographica* 8: 157
106. Skinner JM, Stewart GMD, Speakman JC. 1954. The Crystal Structure of the Acid Salts of Some Monobasic Acids. Part III.* Potassium Hydrogen Dibenzoate. *Journal of the Chemical Society*: 180
107. Smith ME, van Eck ERH. 1999. Recent advances in experimental solid state NMR methodology for half-integer spin quadrupolar nuclei. *Progress in NMR Spectroscopy* 34: 159–201
108. Stebbins JF, Oglesby JV, Kroeker S. 2001. *American Mineralogist* 86: 1307
109. Steinschneider A, Burgar MI, Buku A, Fiat D. 1981. Labeling of amino acids and peptides with isotopic oxygen as followed by ^{17}O-N.M.R. *Int. J. Pept. Protein Res.* 18: 324–33
110. Steinschneider A, Fiat D. 1984. Carbonyl-17O n.m.r. of amino acid and peptide carboxamide and methyl ester derivatives. *Int J Pept Protein Res* 23: 591
111. Stewart RF, Jensen LH. 1967. Redetermination of the crystal structure of uracil. *Acta Crystallographica* 23: 1102
112. Suresh CG, Vijayan M. 1983. Occurrence and geometrical features of head-to-tail sequences involving amino acids in crystal structures. *Int J Pept Protein Res* 22: 129
113. Swaminathan S, Craven BM, McMullan RK. 1984. The crystal structure and molecular thermal motion of urea at 12, 60 and 123 K from neutron diffraction. *Acta Crystallographica Section B* 40: 300
114. Takahashi A, Kuroki S, Ando I, Ozaki T, Shoji A. 1998. Hydrogen-Bonded Structure and NMR Parameters of Oxygen-17 Labeled Poly(L-Alanine)s as Studied by Solid State Oxygen-17 NMR Spectroscopy. *Journal of Molecular Structure* 442: 195
115. Thewalt U, Bugg CE, Marsh RE. 1971. The crystal structure of guanine monohydrate. *Acta Crystallographica Section B* 27: 2358
116. Tricot Y, Niederberger W. 1979. Water Orientation and Motion in Phospholipid Bilayers: a Comparison between 17O- and 2H-NMR. *Biophysical Journal* 9: 195
117. Valentine B, Steinschneider A, Dhawan D, Burgar MI, St Amour T, Fiat D. 1985. Oxygen-17 n.m.r. of peptides. *Int. J. Pept. Protein Res.* 25: 56–68
118. van Eck ERH, Smith ME. 1998. Orientation of the quadrupole and

dipolar tensors of hydroxyl groups by 17O quadrupole separated local field NMR. *Journal of Chemical Physics* 108: 5904

119. Walter TH, Turner GL, Oldfield E. 1988. *Journal of Magnetic Resonance* 76: 106
120. Wu G. 1998. Recent developments in solid-state nuclear magnetic resonance of quadrupolar nuclei and applications to biological systems. *Biochem Cell Biol* 76: 429
121. Wu G, Dong S. 2001. Two-dimensional ^{17}O multiple quantum magic-angle spinning NMR of organic solids. *J. Am. Chem. Soc.* 123: 9119–9125
122. Wu G, Dong S, Ida R. 2001. Solid-State ^{17}O NMR of Thymine: a Potential New Probe to Nucleic Acid Base Pairing. *Chemical Communications*: 891–892
123. Wu G, Dong S, Ida R, Reen N. 2001. A Solid-State 17O Nuclear Magnetic Resonance Study of Nucleic Acid Bases. *Journal of the American Chemical Society* 124: 1768
124. Wu G, Dong S, Ida R, Reen N. 2002. A solid-state ^{17}O nuclear magnetic resonance study of nucleic acid bases. *J. Am. Chem. Soc.* 124: 1768–77
125. Wu G, Hook A, Dong S, Yamada K. 2000. A Solid-State NMR and Theoretical Study of the 17O Electric Field Gradient and Chemical Shielding Tensors of the Oxonium Ion in p-Toluenesulfonic Acid Monohydrate. *Journal of Physical Chemistry A* 104: 4102
126. Wu G, Yamada K. 2003. Determination of the 17O NMR Tensors in Potassium Hydrogen Dibenzoate: A Salt Containing a short O...H...O Hydrogen Bond. *Solid State Nuclear Magnetic Resonance* 24: 196
127. Wu G, Yamada K, Dong A, Grondey H. 2000. Intermolecular Hydrogen-Bonding Effects on the Amide Oxygen Electric-Field-Gradient and Chemical Shielding Tensors of Benzamide. *J Am Chem Soc* 122: 4215
128. Wu Y, Sun BQ, Pines A, Samoson A, Lippmaa E. 1990. NMR Experiments with a New Double Rotor. *Journal of Magnetic Resonance* 89: 297
129. Yamada K, Dong S, Wu G. 2000. Solid-State 17O NMR Investigation of the Carbonyl Oxygen Electric-Field-Gradient Tensor and Chemical Shielding Tensor in Amides. *J Am Chem Soc* 122: 11602
130. Yamauchi K, Kuroki S, Ando I. 2002. High-Resolution Solid-State ^{17}O NMR Studies of Polyglycines and their Hydrogen-Bonded Structures. *J. of Mol. Struct.* 602–603: 171–175
131. Yamauchi K, Kuroki S, Ando I, Ozaki T, Shoji A. 1999. ^{17}O NMR chemical shifts and quadrupole coupling constants in solid poly(*L*-alanine)s determined using a high-speed MAS technique. *Chem. Phys. Lett.* 302: 331–336
132. Yates JR, Pickard CJ, Payne MC, Dupree R, Profeta M, Mauri F. 2004. Theoretical Investigation of Oxygen-17 NMR Shielding and Electric Field Gradients in Glutamic Acid Polymorphs. *J. Phys. Chem. A* 108: 6032–6037
133. Zhang Q, Chekmenev EY, Wittebort RJ. 2003. 17O Quadrupole and Chemical Shielding Tensors in an H-bonded Carboxyl Group: alpha-Oxalic Acid. *J Am Chem Soc* 125: 9140
134. Zhang QW, Zhang HM, Usha MG, Wittebort RJ. 1996. 17O NMR and Crystalline Hydrates. *Solid State Nuclear Magnetic Resonance* 7: 147

I.2 New insights into the bonding arrangements of L- and D-glutamates from solid state ^{17}O NMR

Vincent Lemaître,[1,2] Kevin J. Pike,[3] Anthony Watts,[1] T Anupold,[4] A Samoson, [4] Mark E. Smith,[3] and Ray Dupree,[3]

[1] *Biomembrane Structure Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.*
[2] *Nestec S.A., BioAnalytical Department, Vers-Chez-Les-Blanc, CH-1000 Lausanne 26, Switzerland*
[3] *Department of Physics, University of Warwick, Coventry, CV4 7AL, U.K. ; email: r.dupree@warwick.ac.uk*
[4] *National Institute for Chemical Physics and Biophysics, Akadeemia Tee 23, Tallinn, Estonia*

Keywords Oxygen-17, amino acids, solid state NMR, hydrogen-bonds, 3Q MAS, DOR

■ **ABSTRACT** Magic angle spinning (MAS) from *L*- and *D*-glutamic acid·HCl at 14.1 T produces highly structured and very similar NMR spectra. Lines from all 4 oxygen sites are readily distinguished and assigned. These ^{17}O NMR spectra are very different from the previously reported ^{17}O spectrum of the *D,L*-form presumably because that was a racemic crystal. ^{17}O NMR from monosodium *L*-glutamate·HCl is very different again requiring the application of double angle rotation and triple quantum MAS NMR to provide resolution of five different sites. Hence high-resolution ^{17}O solid state NMR techniques other possible new insight into biochemical bonding processes.

INTRODUCTION

The importance of oxygen, as the most abundant element in the Earth's crust, and its ubiquity throughout living systems, might imply that it could play a central role in biological NMR studies. However, ^{17}O, the only NMR-active oxygen isotope, has a spin I = 5/2 so that it possesses a quadrupole moment (which often leads to significantly broadened lines from solids) which along with a low natural abundance (0.037%), makes its study by NMR still relatively uncommon. In

recent years, despite these difficulties, with the advent of higher magnetic fields, faster magic angle spinning (MAS) and techniques for improving resolution there has been a significant increase in ^{17}O NMR reports from inorganic materials such as gels, glasses, zeolites and mineral analogues (2-4, 6, 10-12, 14). However, there have been many fewer reports of ^{17}O NMR from organic materials as they present even more of a challenge due to the larger quadrupole interaction. Recent reports of ^{17}ONMR from organic materials have included hemeproteins (8), polypeptides (26), strongly hydrogen-bonded carboxylic acids (9), amino acids (23) and nucleic acid bases (24).

L-glutamate plays a significant role in many biochemical processes acting as one of the most important neurotransmitters and by activating several families of brain-receptors (18). Glutamate is very widely used within the food industry with ~1.5 million tons of monosodium glutamate (MSG) used each year (22). It is established that *L*-glutamate (but not *D*-) imparts a unique taste, termed "umami" (25), and is added to a wide variety of foodstuffs to improve their flavour (7). The obvious interest in this has led to several proposed molecular mechanisms to explain taste transduction but as yet none is fully accepted. A key experimental challenge is to provide high quality, detailed and unambiguous atomic scale information about the molecular bonding arrangement and changes that occur upon ligand-receptor interactions. This demands the development of experimental probe techniques to deliver this information. Solid state NMR is one non-perturbing approach, which can be used for the study of ligand-receptor interactions where molecular size is not limiting and crystallinity not a requirement (20, 21). Here it is shown that the ^{17}O NMR spectra from *D*- and *L*-forms of glutamic acid are very similar but very different from the *D,L*- form, and that five sites in MSG can be distinguished, opening up new possibilities for insight into molecular mechanisms responsible for the biological activity of glutamate.

EXPERIMENTAL DETAILS

L-glutamic acid·HCl ($C_5H_9NO_4$·HCl) was prepared following a literature procedure (19) using 20 at% ^{17}O H_2O. The *D*-form was purchased from Sigma-Aldrich, Gillingham (UK) and ^{17}O-enriched by exchange 20 at% ^{17}O H_2O. L-MSG monohydrate ($NaC_5H_8NO_4$·H_2O) was prepared by dissolving

L-glutamic acid in water and adding NaOH until neutral. The sample was recrystallised from water and the excess water was evaporated under vacuum. All samples were single phase to better than 99%.

Most of the NMR was carried out on a Chemagnetics Infinity 600 spectrometer at a frequency of 81.345 MHz. MAS and 3Q MAS experiments used a 4 mm probe spinning at ~16 kHz. For the MSG, where the direct MAS spectrum provided no distinctive spectral features from which to obtain the interaction parameters, variable field double angle rotation (DOR) was carried out. These experiments were undertaken using odd-order sideband suppression by triggering the acquisition between consecutive scans at two positions of the outer rotor differing by 180° (15). The outer rotor speed was varied between 1300 and 1700 Hz to determine the centrebands. ^{17}O DOR spectra were acquired at magnetic fields of 8.45 and 14.1 T at respective frequencies of 48.8 and 81.43 MHz. Spectra were referenced to water at 0 ppm. Spectral simulations were carried out using dmfit software (13).

Calculations of the NMR parameters were performed using GAUSSIAN 98 on different size clusters. Whilst calculations on a large enough cluster agree quite well (<1 ppm for the shift) with experiment for e.g., ^{29}Si in inorganic systems, for ^{17}O agreement is less good (3). Nevertheless with the relatively large differences observed in the experimental determined shifts it was felt that calculations would be a useful aid in site assignment. Calculations were performed using the GIAO method with different basis sets (up to 6.311G**) on different size clusters up to three molecules in extent. In the three molecule calculations the two neighbouring molecules of the specific oxygen site being considered were used.

RESULTS AND DISCUSSION

The ^{17}O MAS NMR spectrum from L-glutamic acid shows a highly featured spectrum with two main resonances centred at 260 and 125 ppm (Fig. 1a). Each of these resonances shows a number of singularities, and is composed of two strongly overlapping lines that are nevertheless readily separated in the spectral simulation. It should be noted that the features in the lower shift resonance

only became clear when high (100 kHz) ^{1}H-decoupling powers were used. The simulation parameters for these 4 sites are summarised in Table 1. For comparison DOR NMR data was obtained at 14.1 T (Fig. 1b) and 8.45 T.

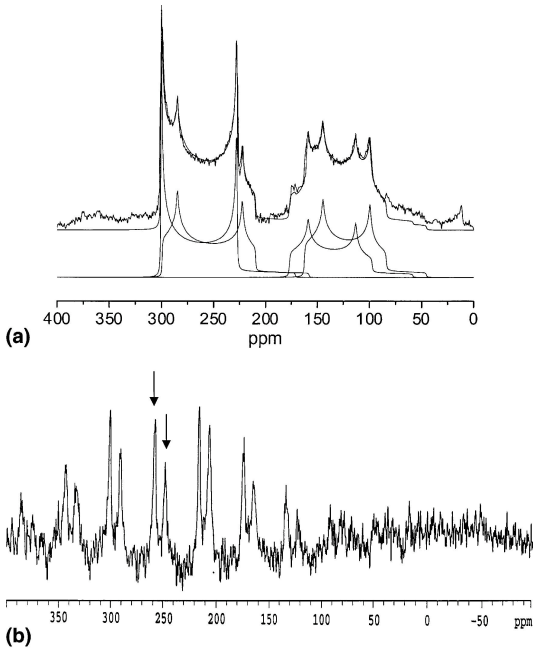


Figure 1: (a) 14.1 T ^{17}O MAS NMR spectrum of *L*-glutamic acid together with a simulation for the centreband only and (b) 14.1 T DOR data collected with the outer rotor spinning at 1700 Hz (↓ indicate the centrebands).

There is a large increase in the resolution of the spectra by employing DOR, which shows a series of narrow lines and associated spinning sidebands. By varying the spinning speed from 1300 to 1700 Hz the centrebands can be identified, with only two centrebands observed at both fields. The peak position peak in DOR spectra is given by

$$\delta_{\text{peak}} = \delta_{\text{cs,iso}} - \frac{3}{40} F(I) \frac{C_Q^2}{\nu_0^2} \left(1 + \frac{\eta^2}{3}\right) \qquad (1)$$

where $\delta_{\text{cs,iso}}$ is the isotropic chemical shift, C_Q is the quadrupole interaction parameter, ν_0 is the Larmor frequency, η is the quadrupole asymmetry parameter and *F(I)* is a spin-dependent factor which for $I = 5/2$ is 2/25. Analysis at multiple applied magnetic fields (1, 3, 12) allows the NMR interaction parameters to be deduced as a product termed the quadrupole product $P_Q = C_Q \sqrt{\left(1 + \frac{\eta^2}{3}\right)}$ which is dominated by C_Q. For the two observed peaks from *L*-glutamic

acid·HCl. P_Q and $\delta_{cs,iso}$ are 8.22 ± 0.10 MHz and 320.7 ± 0.5 ppm, and 8.41 ± 0.10 MHz and 314.6 ± 0.5 ppm. There is very good agreement between these parameters and the sites labelled 1 and 2 observed in MAS (Table 1).

Table 1 ^{17}O NMR interaction parameters for L-glutamic acid·HCl

Site	C_Q / [MHz] ± 0.03	η ± 0.02	δ_{iso} =(ppm) ± 0.5	Assignment
1	8.16	0.0	322.0	O3
2	8.31	0.17	315.0	O2
3	7.49	0.25	187.0	O1
4	7.45	0.25	172.5	O4

The likely assignment of this spectrum is that lines 1 and 2 can be identified with the two carbonyl oxygens (O2 and O3). This assignment agrees with the shifts observed for carbonyls in other amino acids as well as from carboxylic acids. The line with the largest shift has an asymmetry parameter of near zero indicating axial local symmetry, which means it most probably comes from the O3 site where the local geometry is more closely planar (Fig. 2) (17). This view is reinforced from the calculations, which always produce a slightly larger chemical shift for O3. The other two lines then come from OH oxygens (O1 and O4), as indicated by the distinctive shifts and the sensitivity of their spectrum to proton decoupling.

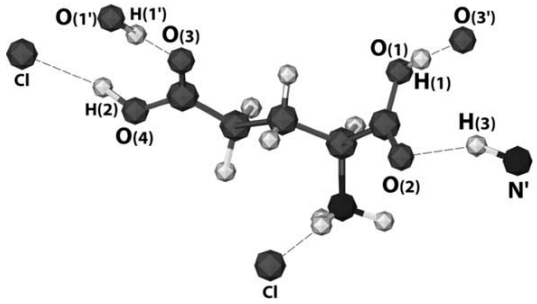


Figure 2: Structure of *L*-glutamic acid.

This is also the reason why these two signals are lost from the DOR spectrum as this was acquired on a single channel probe with no decoupling. Relatively strong dipolar coupling will not be averaged by spinning, as the outer rotor frequency is only 1.7 kHz, thus the magnetisation from these peaks will dephase comparatively rapidly and be lost from the spectrum. The electric field gradient for lines 3 and 4 is very similar with a ~15 ppm shift difference. Calculations over a range

of cluster sizes consistently show that O1 has a larger isotropic chemical shift by ~10–15 ppm than O4, thus line 3 is assigned to O1 and line 4 to O4. The richness of the NMR spectrum from the solid is lost in solution with only two Lorentzian-like lines are observed with shifts of 263 and 251 ppm. The details of the quadrupole interaction are lost because of the motional averaging so that the two carbonyls become indistinguishable as do the two hydroxyls. The very different shifts in solution indicate that the carbonyls and hydroxyls are much more similar here due to the different hydrogen bonding in the solid.

^{17}O MAS NMR from *D*-glutamic acid was also collected and, to within experimental error, the NMR parameters obtained from simulation were the same as those given in Table 1 for the *D*-form. This data from the *L*- and *D*-glutamic acids individually makes interesting comparison with that reported recently for *D, L*-glutamic acid (23) where the spectrum showed two resonances in approximately the same position as each of the pairs of lines that are observed here for both the *D*- and *L*-glutamic acids. Even under 3Q conditions there was no resolution into separate lines (23). There was, in addition, an extra peak of twice the intensity centred around ~200 ppm intermediate to the two peaks observed here in *D*- and *L*-glutamic acids. The assignment of the resonances in that report (23), where hydroxyls are assigned to the outer two peaks and carbonyls to the inner resonance, is at variance to that suggested here. The very unusual shift for the hydroxyl was ascribed to the O–H···Cl bond. However, the equivalent HCl compounds are used here so that a possible reinterpretation of the spectrum from the *D,L*-form can be made. Thus the outer two resonances are from carbonyls and hydroxyls that are closely similar to those present in the *L*- and *D*- structures since their parameters are in close agreement with the mean parameters observed for the two pairs of sites observed here. Then the intermediate peak comes from oxygens (appearing to have similar NMR parameters) that are very different to those in the chiral forms and result from the differing bonding in the *D,L*-form. The *D, L*-form of the glutamate could be either be a racemic conglomerate of chiral crystals or a racemic crystal. However, the presence of the additional resonance in the spectrum with 50% of the total intensity strongly implies the latter. The *D, L*-glutamic acid·HCl investigated by Wu et al. (23) was enriched from the *D, L*-monohydrate. This same material was used by Dunitz and Schweizer (5) to produce racemate anhydrous *D,L*-glutamic acid whose conformation was found to differ from the two chiral modifications of the monohydrate

(5). No detailed structure for racemic *D,L*-glutamic acid·HCl could be located in the literature. However, inspection of the structure of the racemic monohydrate (5) shows that whilst two oxygen sites have similar C–O bond lengths to the *L*-glutamate·HCl form the other two oxygens have bond lengths intermediate in length consistent with the intermediate shift observed. It is clear that, even from the limited numbers of oxygen observations that have already been made in the solid state, oxygen parameters can vary strongly with the details of the bonding arrangement. Hence differences in the packing conformation between the chiral forms and the *D,L*-racemate leads to changes in the oxygen bonding of the network so that in the racemate half the hydroxyls and carbonyls must be in very different (hydrogen) bonding environments. The ability to detect small non-covalent interactions is indicative of the sensitivity of this approach in opening up a great wealth of new detailed information.

The sensitivity of oxygen to changes in bonding is emphasised by comparison of the acid with the sodium salt where there are two distinct but very similar glutamate anions (16) doubling, in principle, the number of oxygen sites. The well-separated pairs of resonances with highly structured quadrupolar lineshapes for the acid are replaced by a single, almost featureless line with some minor structure in the MAS spectrum (Fig. 3a). This peak is at an intermediate shift to the signals from *L*-glutamic acid·HCl. To help resolve the different sites DOR and 3Q MAS were applied producing a large increase in the resolution of the spectra. DOR (Fig. 3b) shows a series of narrow lines and associated spinning sidebands. By varying the spinning speed from 1300 to 1700 Hz the centrebands can be identified. It appears that at least 5 isotropic lines can be resolved which is confirmed by the 3Q MAS NMR which also shows 5 lines in the centreband in Fig. 3c. The DOR and isotropic MQ spectra are differently affected by the quadrupole interaction [23] introducing an additional multiplicative factor of –19/12 in Eq. (1) for 3Q experiments for $I = 5/2$. Plotting the DOR and 3Q data allows an estimate of the combined quadrupole parameter, P_Q, and shift for each line (Fig. 3d) (1). These NMR parameters were consistent with simulations of the anisotropic slices of the 3Q data set where the asymmetry parameter, η, was found to be typically ~0.4–0.5 (Table 2) and a composite of the 5 lines produced a good simulation of the almost featureless MAS lineshape which could even reproduce some of the fine structure on the main peak (Fig. 3a).

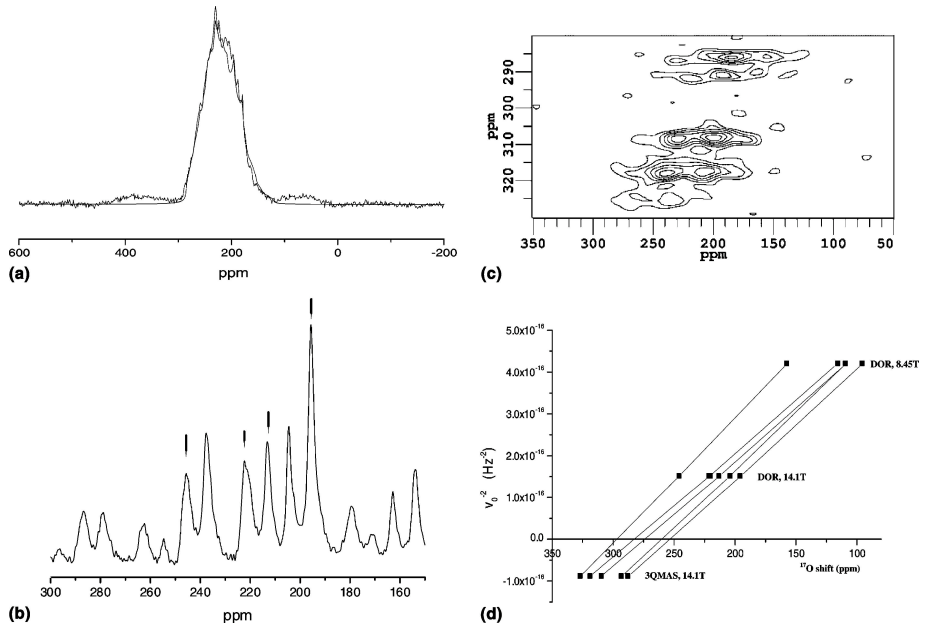


Figure 3: (a) 14.1 T ^{17}O MAS NMR spectrum of *L*-MSG together with a simulation based on the 5 components deduced from the higher resolution data. (b) 14.1 T DOR (↓ indicate the centrebands), (c) the centreband of the 14.1 T 3Q MQMAS NMR data and (d) the field dependence of the isotropic position from the DOR and 3Q data.

The shifts for all these lines cover a 43 ppm range centred on 274 ppm very different from *L*-glutamic acid. The electric field gradients are also very different, C_Q is within ~5% of 7.4 MHz, much less than O2 and O3 in *L*-glutamic acid and the asymmetry parameters of all sites are much larger. These changes must be due to the changed bonding in MSG. Although all 8 sites cannot be resolved, even with a resolution of better than 1 ppm, inspection of the structure (16) shows that the oxygens occupy 4 pairs of very similar sites. One of these pairs, possibly O5/O7, which show the greatest difference in the bond distances (0.10 Å), splits to give the 5 lines observed.

Table 2 ^{17}O NMR interaction parameters for *L*-MSG. P_Q and □$_{iso}$ are taken from the DOR results, □ and C_Q from the MQ slice simulations

Site	P_Q / (MHz) ± 0.05	δ_{iso} [ppm] ± 1.5	C_Q / (MHz) ± 0.2	η ± 0.05
1	7.93	254	7.4	0.47
2	7.76	260	7.2	0.50
3	8.11	274	7.6	0.45
4	8.15	283	7.7	0.40
5	7.45	297	7.0	0.45

CONCLUSION

The data presented here has further demonstrated the exquisite sensitivity of solid state ^{17}O NMR to changes in the bonding environment with specific application to biochemically significant molecules. The differences between the *D,L*- and *L*- and *D*-forms of glutamic acid·HCl must reflect changes in the (hydrogen) bonding within the compounds which determines the molecular packing. The chemical shift and electric field gradient of *L*-MSG formed from the acid indicate significant changes in the bonding at the oxygen sites occur. Higher resolution techniques than MAS are necessary to reveal these differences. 3Q MAS at a single field would be sufficient but the quality and accuracy of the data is greatly improved by the use of DOR. Neither ^{1}H, where the spectra begin to show signs of resolution for the hydroxyl protons only in glutamic acid at ultra high spinning speeds (>40 kHz), nor ^{13}C others anywhere near the degree of structural discrimination of ^{17}O. The ability to resolve the different sites and to detect changes from ^{17}O NMR spectra suggest that this will be a fruitful experimental approach to elucidate molecular pathways of biochemical recognition and could find widespread application.

ACKNOWLEDGEMENTS The EPSRC are thanked for supporting the NMR equipment at Warwick and the work on biological materials through grant GR/N29549. The Royal Society supports the collaboration with Tallinn and MES also thanks the Royal Society and Leverhulme Trust for the award of a Senior Research Fellowship 2001–2002. AW and VL thank BBSRC for support and Dr. E. Hughes for advice. The Estonian Science Foundation is also thanked for support.

REFERENCES

1. Anupõld T, Reinhold A, Sarv P, Samoson A. 1998. A comparison of double rotation and multi-quantum magic angle spinning spectra. *Solid State Nucl Magn Reson* 13: 87
2. Ashbrook SE, Berry AJ, Wimperis S. 1999. *Am. Mineral* 84: 1191
3. Bull LM, Bussemer B, Anupõld T, Reinhold A, Samoson A, et al. 2000. A High-Resolution ^{17}O and ^{29}Si NMR Study of Zeolite Siliceous Ferrierite and ab Initio Calculations of NMR Parameters. *J Am Chem Soc* 122: 4948
4. Dirken PJ, Kohn SC, Smith ME, van Eck ERH. 1997. Complete resolution of Si---O---Si and Si---O---Al fragments in an aluminosilicate glass by 17O multiple quantum magic angle spinning NMR spectroscopy. *Chemical Physics Letters* 266: 568–574
5. Dunitz J, Schweizer W. 1995. Anhydrous *D,L*-glutamic acid. *Acta Cryst.* C51: 1377
6. Gervais C, Babonneau F, Smith ME. 2001. Detection, quantification and magnetic field dependence of solid state 17O NMR X-O-Y (X,Y =Si,Ti)

linakges: Implications for characterising amorphous titania-silica based materials. *Journal of Physical Chemistry B* 105: 1971

7. Girardot NF, Peryam DR. 1954. MSG's power to perk up foods. *Food Engineering* 26: 71–72
8. Godbout N, Sanders KL, Salzmann R, Havlin RH, Wojdelski M, Oldfield E. 1999. Solid-State NMR, Mössbauer, Crystallographic, and Density Functional Theory Investigation of Fe-O2 and Fe-O2 Analogue Metalloporphyrins and Metalloproteins. *J Am Chem Soc* 121: 3829
9. Howes AP, Jenkins R, Smith ME, Crout DHG, Dupree R. 2001. Influence of low-barrier hydrogen bonds on solid state ^{17}O NMR spectra of labelled phthalate species. *Chemical Communications*: 1448
10. Hussin R, Dupree R, Holland D. 1999. *J . Non-Cryst. Solids* 246: 159
11. Kroeker S, Rice D, Stebbins JF. 2002. *Am. Mineral* 87: 572
12. MacKenzie KJD, Smith ME. 2002. *Multinuclear Solid State NMR of Inorganic Materials*. Oxford: Pergamon Press
13. Massiot D, Fayon F, Capron M, King I, Le Calvi S, et al. 2002. Modelling one- and two-dimensional solid-state NMR spectra. *Magnetic Resonance in Chemistry* 40: 70
14. Pickup DM, Mountjoy G, Wallidge GW, Anderson R, Cole JM, et al. 1999. *J. Mater. Chem.* 9: 1299
15. Samoson A, Lippmaa E. 1989. Synchronized Double-Rotation NMR-Spectroscopy. *Journal of Magnetic Resonance* 84: 410
16. Sano C, Nagashima N, Kawakita T, Iitaka Y. 1989. Crystal and molecular structures of monosodium L-glutamate monohydrate. *Analytical Sciences* 5: 121–122
17. Sequeira A, Rajagopal H, Chidambaram R. 1972. A neutron diffraction study of the structure of L-glutamic acid.HCl. *Acta Cryst.* B28: 2514–2519
18. Stanley BG. 1996. Glutamate and its receptors in lateral hypothalamic stimulation of eating. Chapter 14. In *Drug receptor subtypes and ingestive behaviour*, pp. 301–322. New York: Academic Press
19. Steinschneider A, Burgar MI, Buku A, Fiat D. 1981. Labeling of amino acids and peptides with isotopic oxygen as followed by ^{17}O-N.M.R. *Int J Pept Protein Res* 18: 324–33
20. Watts A. 1999. NMR of drugs and ligands bound to membrane receptors. *Current Opinion in Biotechnology* 10: 48
21. Watts A. 1999. Structural Resolution of Ligand–Receptor Interactions in Functional, Membrane-embedded Receptors and Proteins using Novel, Non-perturbing Solid-state NMR Methods. *Pharmacy and Pharmacology Communications* 7: 7
22. Whitfield J. 2002. Yum, amino acids. *Nature Science Update*
23. Wu G, Dong S. 2001. Two-dimensional ^{17}O multiple quantum magic-angle spinning NMR of organic solids. *J Am Chem Soc* 123: 9119–25
24. Wu G, Dong S, Ida R, Reen N. 2002. A solid-state ^{17}O nuclear magnetic resonance study of nucleic acid bases. *J Am Chem Soc* 124: 1768–77
25. Yamaguchi S, Kimizuka A. 1979. Psychometric studies on the taste of monosodium glutamate. In *Glutamic Acid: Advances in Biochemistry and Physiology*, ed. LJ Filer, S Garattini, MR Kare, WA Reynolds, RJ Wurtman, pp. 35–54. New York: Raven Press
26. Yamauchi K, Kuroki S, Ando I, Ozaki T, Shoji A. 1999. ^{17}O NMR chemical shifts and quadrupole coupling constants in solid poly(L-alanine)s determined using a high-speed MAS technique. *Chemical Physics Letters* 302: 331

I.3 Solid State ^{17}O NMR as a probe for structural studies of proteins in biomembranes

Vincent Lemaître,[1,2] Maurits R.R. de Planque,[2] Andy P. Howes,[3] Mark E. Smith,[3] Ray Dupree,[3] and Anthony Watts,[1]

[1] *Biomembrane Structure Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK ; email: anthony.watts@bioch.ox.ac.uk*
[2] *Nestec S.A., BioAnalytical Department, Vers-Chez-Les-Blanc, CH-1000 Lausanne 26, Switzerland*
[3] *Department of Physics, University of Warwick, Coventry, CV4 7AL, U.K.*

Key words oxygen-17, solid state NMR, MAS, peptide, vesicles, hydrogen-bond, biological applications.

■ **ABSTRACT** We report the first example of ^{17}O NMR spectra from a selectively labelled transmembrane peptide, ^{17}O-[Ala12]-WALP23, as a lyophilized powder and incorporated in hydrated phospholipid vesicles. It is shown that at high magnetic field it is feasible to apply ^{17}O NMR to the study of membrane-incorporated peptides. Furthermore, we were able to estimate distances within the selectively labelled WALP peptide, which represents a consensus transmembrane protein sequence. This work opens up new applications of ^{17}O solid state NMR on biological systems.

INTRODUCTION

A key experimental challenge to understand conformational details of membrane proteins is to provide unambiguous atomic scale information about the molecular bonding arrangement, and any changes that occur upon receptor activation. This demands the development of experimental probe techniques to deliver this information of biological and pharmaceutical importance. Solid state NMR is a non-perturbing approach which can be used to study ligand-protein interactions where molecular size is not limiting and crystallinity not a requirement.(16) As a first step in addressing this challenge by exploiting ^{17}O NMR in biomembrane applications, we report here ^{17}O solid state

NMR at high field of a ^{17}O selectively labelled transmembrane peptide in a biomimetic environment.

Oxygen plays a key role in the molecular conformation of biopolymers. Amongst the important information that can be obtained from ^{17}O, the only NMR-active oxygen isotope, is the isotropic chemical shift (δ_{iso}), the quadrupolar coupling constant (C_Q) and asymmetry parameter (η). These parameters are extremely sensitive to the electronic distribution around the nucleus, more specifically to its protonation state (13) and its involvement in hydrogen bonds. Furthermore, they contain structural information (1, 8, 13, 14), and several methods to determine internuclear distances between oxygen and other nuclei have been developed (4, 6, 13). This suggests that oxygen could play a central role in biological NMR studies. However, ^{17}O has a low natural abundance (0.037%), and a spin I=5/2 with a corresponding quadrupole interaction that is manifested as significantly broadened signals in NMR spectra. Consequently, ^{17}O solid state NMR studies are still relatively uncommon and selective labeling is essential. Despite these difficulties, in recent years with the advent of higher magnetic fields and techniques for improving resolution beyond magic angle spinning experiments (MAS), there has been a significant increase in ^{17}O NMR reports from inorganic materials such as glasses and zeolites (2, 10). There has been much less ^{17}O NMR reported from organic materials since ^{17}O presents even more of a challenge due to the larger quadrupole interaction and hence larger linewidths (9). Recent reports of ^{17}O NMR from biologically relevant materials have included inorganic molecules interacting with hemeproteins (12), polypeptides (18, 19), amino acids (13) and nucleic acid bases (17).

Here we report the first example of ^{17}O NMR spectra from a selectively labelled transmembrane peptide, ^{17}O-[Ala12]-WALP23, both as a lyophilized powder and incorporated in hydrated vesicles, opening up new possibilities for applications of ^{17}O solid state NMR on real biological systems. WALP23 is a synthetic peptide which represents a consensus sequence for transmembrane protein segments (5). This hydrophobic peptide forms well-defined and well-characterized transmembrane α-helices (5) and has special relevance for packing of α-helices in polytopic proteins such as G-protein coupled receptors.

MATERIAL AND METHODS

Fmoc-L-alanine was prepared starting from ^{17}O uniformly enriched L-alanine following a previously described procedure (15) using $H_2{}^{17}O$ (70% at ^{17}O). WALP23 was synthesized and transmembranously incorporated in multilamellar vesicles (MLV) of 1,2-distearoyl-*sn*-glycerol-3-phosphatidylcholine (DSPC) lipids (5). ^{17}O NMR experiments were carried out on Chemagnetics Infinity 600 and Varian/Magnex 800 wide-bore 89 mm spectrometers at frequencies of 81.345 and 108.419 MHz for ^{17}O, respectively. MAS was performed at a spin rate of 11-22 kHz. Spectral simulations were carried out using dmfit software (11).

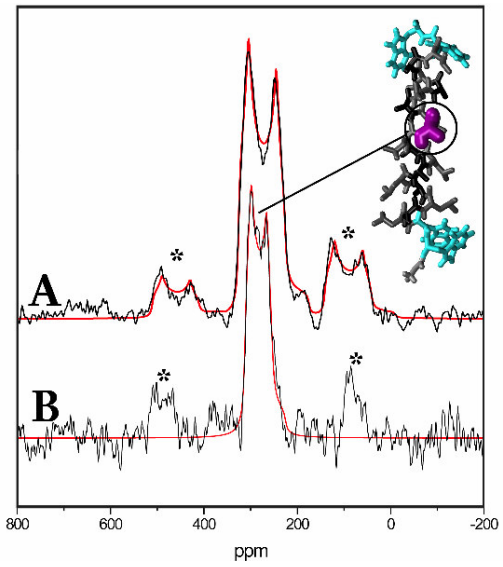


Figure 1. ^{17}O NMR spectra of lyophilized ^{17}O-[Ala12]-WALP23 at room temperature (a) at 14.1 T, 4 mm rotor spinning at ~15 kHz, with simulation, and (b) at 18.8 T, 2.5 mm rotor spinning at ~22 kHz, with simulation. Spinning sidebands are marked with an asterisk (*).

RESULTS AND DISCUSSION

The ^{17}O MAS NMR spectra of lyophilized ^{17}O-[Ala12]-WALP23 acquired at magnetic field strengths of 14.1 and 18.8 T show a single resonance, centered[17] at ~280 - 300 ppm, from the single label at alanine-12 (Fig. 1). Spinning side-bands of the central transition are also visible. The decrease of the second-order quadrupole broadening and associated increase in spectral resolution with increasing applied magnetic field is clearly evident. The simulation parameters for ^{17}O in this alanine environment are summarized in Table 1.

Since WALP23 is lyophilized from a solution of 2,2,2-trifluoroethanol, the α-helical conformation should be retained. The isotropic chemical shift for the label (317.5 ppm) is very close to that of another polymer, poly-L-alanine, which also forms ideal α-helices (319 ppm(19)). The difference between these two samples is only ~1.5 ppm. This suggests only small differences in local environment for the two peptides, although poly-L-Alanine is crystalline and ^{17}O-[Ala12]-WALP23 is not.

Table 1 ^{17}O NMR interaction parameters for L-alanine in two protonation states, [Ala12]-WALP23 lyophilized and in dry and hydrated MLVs (1:10, peptide:lipid molar ratio), and poly-L-alanine.

Site	C_Q [MHz]	η	δ_{iso} [ppm]	Assignment
L-Alanine (2)				
1	7.86±0.05	0.28±0.03	284.0±0.5	O1
2	6.53±0.05	0.70±0.03	260.5±0.5	O2
L-Alanine·HCl (2)				
1	8.31±0.02	0.00±0.02	327.8±0.5	C=O
2	7.29±0.02	0.25±0.02	176.7±0.5	C-OH
Fmoc alanine (2)				
1	7.89±0.05	0.16±0.02	303.3±0.5	C=O
2	6.95±0.05	0.12±0.02	175.7±0.5	C-OH
Lyophilized [Ala12]-WALP23				
1	8.45±0.10	0.21±0.03	317.5±0.5	C=O
[Ala12]-WALP23 in DSPC Vesicles				
1	8.42±0.10	0.21±0.03	311±1	C=O
[Ala12]-WALP23 in Hydrated DSPC Vesicles				
1	8.55±0.15	0.24±0.03	315±1	C=O
Poly-L-alanine (11)				
1	8.59	0.28	319	C=O

The ^{17}O MAS NMR spectrum of ^{17}O-[Ala12]-WALP23 in hydrated DSPC vesicles (Fig. 2) also shows a single resonance from the label inserted in alanine-12 centered around ~280 ppm. There is a partial overlap between the spinning side-bands and the central transition because of the lower spinning speed which was necessary for this sample to maintain vesicle integrity. The hydration water is also visible (0 ppm), but does not interfere with the main signal from the ^{17}O label. The lower spinning speed does not prevent determination of the ^{17}O NMR parameters which are given in Table 1. This experiment, which was designed as a test to evaluate the sensitivity of the method, shows that it is possible to study ^{17}O selectively labelled peptides in hydrated biomembranes.

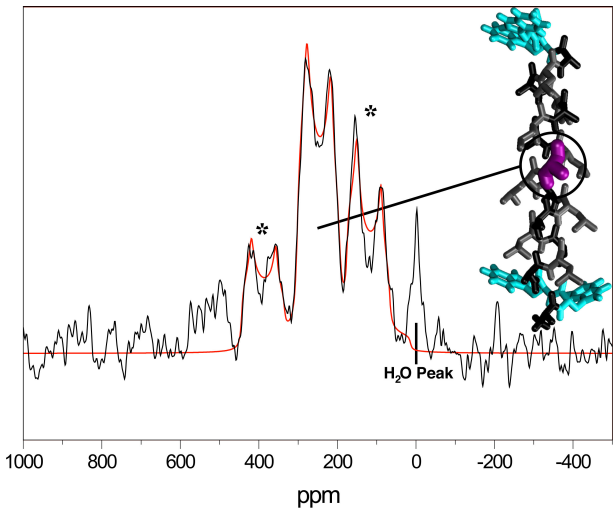


Figure 2: ^{17}O NMR spectra of ^{17}O-[Ala12]-WALP23 in hydrated DSPC vesicles (1:10 peptide/lipid molar ratio) at 14.1 T, 4 mm rotor spinning a ~11 kHz, with simulation. The lyophilized MLV sample was hydrated with one weight equivalent of water, and the spectrum was acquired at room temperature, with the lipids in the liquid-crystalline phase.

We recently obtained a correlation between the isotropic chemical shift and the C=O bond lengths determined by diffraction.(13) In addition, *ab initio* calculations on glutamic acids polymorphs also showed a correlation between these parameters. [18] These two correlations were very similar. Based on these data, the C=O bond length at the labelled site for lyophilized WALP is 1.217 Å, very close to the 1.220 Å for an ideal α-helix. For the WALP peptide incorporated in hydrated vesicles, the corresponding value is 1.223 Å, an increase in C=O bond length of ~0.006 Å. In order to further test this correlation, it was applied on other polymers studied previously by ^{17}O solid state NMR (polyglycine II and poly-L-alanine I & II, data not shown) and was shown to give bond lengths consistent with the known structures.

CONCLUSIONS

The data presented here are the first application of ^{17}O solid state NMR to characterize and estimate distances within a selectively labelled peptide of biological interest. The availability of high field NMR spectrometers (here up to 18.8 T with 89 mm bore magnet) allows useable S/N ratios to be readily achieved. It should be emphasized that this included a sample with <5% of the amino acid residues labelled, and which was further diluted by a factor of 10-20 because of the presence of

lipids and water. The ability to detect and characterize a signal from a selectively labelled peptide suggests that ^{17}O NMR will be a fruitful experimental approach to study hydrogen bonding in macromolecules. It could find widespread application in the study of selectively labelled peptides reconstituted in biological membranes, or selectively labelled organic molecules interacting with crystallized proteins or with membrane proteins in biological membranes. The high sensitivity of solid state ^{17}O NMR to the local bonding environment also allows very accurate determination of small changes in C=O bond length.

ACKNOWLEDGMENTS This work was supported by EPSRC BBSRC and EMBO. We thank Dr Dirk Rijkers and Dr Dahlia Fisher for their advice on the peptide synthesis. Magnex and Varian are acknowledged for their effort in building the first commercially available 800 MHz ^{1}H WB 89 mm NMR spectrometer.

SUPPORTING INFORMATION

Sample preparation

The enrichment of L-alanine was performed following an established protocol (15), and its characterization by solid state NMR reported in (13). The protection of the amino group of L-alanine was achieved through a standard Fmoc-protection procedure (3) without any significant loss of label, as reported in (13). For the synthesis and purification of ^{17}O-[Ala12]-WALP23 and the preparation of the multi-lamellar vesicles we followed the protocols described in (7) and (5) respectively.

Dry vesicles (60 mg), made of ^{17}O-[Ala12]-WALP23 and DSPC mixed in a 1:10 molar ratio were mixed with 60 μl of water (~1:1 solid-liquid mass ratio), leading to the formation of a dispersion. Most of the sample (> 85%) was loaded into a 4 mm Varian MAS rotor. The formation of the vesicles was monitored by ^{31}P static NMR before and after hydration, and before and after the ^{17}O-MAS.

(DSPC stands for 1,2-distearoyl-*sn*-glycero-3-phosphocholine)

NMR experiments

The ^{17}O NMR was carried out on Varian Infinity 600 or 800 spectrometers operating at frequencies of 81.345 MHz and 108.44 MHz respectively. MAS experiments used either a 4 mm probe spinning at ~15 kHz or a 2.5 mm probe (on the 800 MHZ spectrometer) spinning at ~ 22 kHz. A spin-echo using extended phase cycling (2) was used with the echo spacing set to the rotation period. Typical pulse lengths were ~ 1 μs and 2μs and spectral width was 400 kHz. Typical pulse delays were 25 ms with 200,000 - 400.000 acquisitions. All spectra were referenced to water at 0 ppm.

The hydrated sample was spun at 10kHz only to avoid damage to the sample. From our temperature tests on the 800 MHz spectrometer with a 4 mm probe, spinning at 10 kHz induces a temperature increase of about 10° in the probe, while the temperature for the phase transition for DSPC is 54.8 °C and is broadened by up to 10° by the presence of peptides. Thus the sample is a liquid-crystalline phase, as confirmed in other MAS NMR studies (20). The sample was checked by static ^{31}P NMR before and after MAS to confirm that it had not deteriorated whilst spinning.

REFERENCES

1. Asakawa N, Kameda T, Kuroki S, Kurosu H, Ando S, et al. 1998. Structural Studies of Hydrogen-Bonded Peptides and Polypeptides by Solid-State NMR. *Annual Reports on NMR Spectroscopy* 35: 55
2. Bull LM, Bussemer B, Anupõld T, Reinhold A, Samoson A, et al. 2000. A High-Resolution ^{17}O and ^{29}Si NMR Study of Zeolite Siliceous Ferrierite and ab Initio Calculations of NMR Parameters. *J Am Chem Soc* 122: 4948
3. Carpino LA, Han GY. 1972. The 9-Fluorenylmethoxycarbonyl Amino-Protecting Group. *Journal of Organic Chemistry* 37: 3404–3409
4. Chopin L, Vega S, Gullion T. 1998. A MAS NMR Method for Measuring 13C-17O Distances. *J Am Chem Soc* 120: 4406
5. de Planque MR, Kruijtzer JA, Liskamp RM, Marsh D, Greathouse DV, et al. 1999. Different membrane anchoring positions of tryptophan and lysine in synthetic transmembrane alpha-helical peptides. *J Biol Chem* 274: 20839
6. Hu KN, Iuga D, Griffin RG. 2003. *Dynamic Nuclear Polarization Enhanced 17O Solid State NMR Spectroscopy on Biological Solids*. Presented at 44th Exprerimental NMR Conference, Savannah, USA
7. Killian JA, Salemink I, de Planque MR, Lindblom G, Koeppe RE, 2nd, Greathouse DV. 1996. Induction of nonbilayer structures in diacylphosphatidylcholine model membranes by transmembrane alpha-helical peptides: importance of hydrophobic mismatch and proposed role of tryptophans. *Biochemistry* 35: 1037
8. Kuroki S, Ando S, Ando I. 1995. An MO study of nuclear quadrupolar coupling constant and nuclear shielding of the carbonyl oxygen in solid peptides with hydrogen bonds. *Chem. Phys.* 195: 107–116

9. Lemaître V, Smith ME, Watts A. 2004. A Review of Oxygen-17 Solid-State NMR of Organic Materials. *Solid State NMR* In Press
10. MacKenzie KJD, Smith ME. 2002. *Multinuclear Solid State NMR of Inorganic Materials*. Oxford: Pergamon Press
11. Massiot D, Fayon F, Capron M, King I, Le Calvi S, et al. 2002. *Magnetic Resonance in Chemistry* 40: 70
12. McMahon MT, deBios AC, Godbout N, Salzmann R, Laws DD, et al. 1998. An Experimental and Quantum Chemical Investigation of CO Binding to Heme Proteins and Model Systems: A Unified Model Based on 13C, 17O, and 57Fe Nuclear Magnetic Resonance and 57Fe Mössbauer and Infrared Spectroscopies. *J Am Chem Soc* 120: 4784
13. Pike KJ, Lemaître V, Kukol A, Anupõld T, Samoson A, et al. 2004. Solid-State ^{17}O NMR of Amino Acids. *Journal of Physical Chemistry B* 108: 9256–9263
14. Seliger J. 1998. 17O Quadrupole Coupling in C-O-H…O=C Hydrogen Bonds. *Chemical Physics* 231: 81
15. Steinschneider A, Burgar MI, Buku A, Fiat D. 1981. Labeling of amino acids and peptides with isotopic oxygen as followed by ^{17}O-N.M.R. *Int. J. Pept. Protein Res.* 18: 324–33
16. Watts A. 1999. NMR of drugs and ligands bound to membrane receptors. *Current Opinion in Biotechnology* 10: 48
17. Wu G, Dong S, Ida R, Reen N. 2002. A solid-state ^{17}O nuclear magnetic resonance study of nucleic acid bases. *J. Am. Chem. Soc.* 124: 1768–77
18. Yamauchi K, Kuroki S, Ando I. 2002. High-Resolution Solid-State ^{17}O NMR Studies of Polyglycines and their Hydrogen-Bonded Structures. *J. of Mol. Struct.* 602–603: 171–175
19. Yamauchi K, Kuroki S, Ando I, Ozaki T, Shoji A. 1999. ^{17}O NMR chemical shifts and quadrupole coupling constants in solid poly(*L*-alanine)s determined using a high-speed MAS technique. *Chem. Phys. Lett.* 302: 331–336
20. Zhou Z, Sayer BG, Hughes DW, Stark RE, Epand RM. 1999. Studies of phospholipid hydration by high-resolution magic-angle spinning nuclear magnetic resonance. *Biophys J* 76: 387

PART II

Introduction Part II – Molecular Dynamics Simulations on Membrane Peptides or Proteins

MOLECULAR DYNAMIC SIMULATION

Overview

Molecular dynamics simulation (MD) has become an important tool for understanding the physical basis of the structure, dynamics and function of biological molecules (26, 27). With continuously increasing computing power and the introduction of faster and more accurate algorithms, the study of systems with increasing size and complexity has been made possible. This method has a relatively broad field of application. One of the major contributions of MD has been the replacement of a protein viewed as relatively rigid structure, by a dynamic model where the function of the protein results from the conformational changes and the internal motions.

The next sections will be focused on the application of MD on biologically relevant protein/peptides in lipid bilayers, starting with a brief description of the MD technique, its potential for such systems and its main limitations. The systems studied in the next chapters include (1) bovine rhodopsin, (2) WALPC, a model for a transmembrane peptide, in order to complement biophysical studies conducted by AFM, where the bilayer surface was imaged and peptides extracted out of the bilayer using the AFM cantilever, and (3) Vpu from HIV-1, focusing on ligand-protein interactions. All the MD simulations were conducted in explicit hydrated lipid bilayers.

Molecular dynamics methodology

Like most theoretical methods applied on biomolecules (e.g. protein structure prediction, fold recognition, homology modeling), the MD relies mainly on statistical effective energy functions. Although the theoretical foundation of such functions is not clear, their usefulness has been demonstrated in many applications (31).

Interactions between atoms are divided in non-bonded interactions, using either a given cut-off radius between any pair of atoms or more complex methods (like the Particle Mesh Ewald used to compute long-range electrostatic interactions), and bonded interactions between atoms connected by chemical bonds. For the non-bonded interactions (electrostatic and van der Waals), a partial charge and parameters for van der Waals repulsion and attraction are assigned to each atom. The bonded interactions consist of bond, angle and dihedral terms. Bonds and angles are usually described as harmonic oscillators and dihedral angles are usually described by a suitable cosine expansion. A typical potential function has the form:

$$V = \sum_{i<j} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + \sum_{i<j} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \sum_{bonds} \frac{1}{2} k_{ij}^{b} (r_{ij} - b_{ij}^{0})^2 + \sum_{angles} \frac{1}{2} k_{ijk}^{\theta} (\theta_{ijk} - \theta_{ijk}^{0})^2 + \sum_{dih} k^{\varphi} (1 + \cos(n(\varphi - \varphi^{0})))$$

Figure 1: Typical potential energy function used for MD simulation (including simulated annealing). Red, electrostatic interactions; blue: van der Waals interactions; orange: bond interactions; purple: angle interactions; green: dihedral interaction. r_{ij} is the distance between atoms (or pseudo-atoms when CH_n groups are treated as one atom) i and j, q_i is the partial charge on atom i, A_{ij} and B_{ij} are Lennard–Jones parameters, k^b, k_{ijk}^{θ} and k^{φ} are force constants for bonds, angles and dihedrals, n is the dihedral multiplicity and b^0, θ_{ijk}^0, φ^0 are equilibrium values for the bond lengths, angles and dihedral angles, respectively.

This equation for the potential energy and the force field required for a proper parametrization of the system under study, imply several assumptions. The most obvious is the use of Newtonian mechanic, which results in a classical treatment of the system (the charge distribution on each atoms is fixed for the electrons in their ground state). Another important assumption lies in the pair-additivity of the interactions: non-bonded interactions involving three or more atoms are neglected. Other important assumptions are the point charge representation for atoms (electronic polarizability is neglected) and the use of simple quadratic forms for computational efficiency. For the same purpose, the bond-lengths are often constrained (SHAKE or LINCS algorithms).

These simplifications mean that MD has a number of important limitations. The force field required a large number of parameters, some of them are difficult to determine, and therefore does not guarantee the accuracy of the simulations. This uncertainty in the parameters makes necessary extensive testing of parameter sets on simple systems, which can be compared to experimental data.

In fact, often parameters are treated as empirical values that are obtained by fitting models to experimental data, e.g. a water model to experimental data on water. The other limitation is the calculation time-step has to be very short for the integration of the equations of motion to converge. A typical value in practice is 2 femtoseconds (10^{-15} s). This means that 500'000 computationally expensive integration steps are necessary to calculate the dynamics of a system during one nanosecond. The small time step combined with the practical limit on computer power limit both the lengths and the system size of current simulations. In practice, systems made of tens of thousands of particles, corresponding to a simulation box of about 5–10 nanometers, can be simulated to the nanosecond time scale (10-100 ns) (48).

Using the potential energy function described in Fig.1, the equations of motion can be solved for all atoms in the system by calculating the forces on all atoms and integrating in time. In principle, it is possible to calculate the complete dynamics of any system that can be described in terms of a simple interaction potential. The main result of such a calculation is a trajectory of all atoms in time: the coordinates and velocities of all atoms at any of the integration steps, potentially making MD simulation a powerful method to study the motions of atoms in a detailed manner.

AFM-pulling and Steered Molecular Dynamics

AFM-pulling is a subcategory of steered molecular dynamics (SMD) (22, 37), where external forces are voluntarily applied on arbitrary groups of atoms. AFM-pulling involves a virtual spring, where one side of the spring is connected to an atom or a group of atoms, while the other side is moved at constant speed in a chosen arbitrary direction The constant motion of the virtual spring has the effect of pulling the atoms away from their location following the spring to which they are bound. A time-dependent force is superimposed on the molecule bound to the virtual spring. In the case of AFM-pulling simulations of the extraction of a membrane peptide out of its bilayer, a force perpendicular-directed relative to the membrane plan is applied to a residue (supposed to interact with the cantilever of the AFM in the experiment), to initiate the move out of the bilayer. The applied force has the form:

$$F = K_d \, (Z_{Cantilever} - Z_{Cys})$$

Where $Z_{Cantilever}$ is the Z-coordinate of the side of the virtual cantilever moving at a constant speed along an arbitrary predefined vector; Z_{Cys} is the Z-coordinate the side of the virtual cantilever linked with the peptide. The two extremities of the virtual cantilever are related together with the equation of a harmonic spring with a spring constant K_d. The technique has been applied to unfold soluble protein, in complement of AFM experiments (Figure 2).

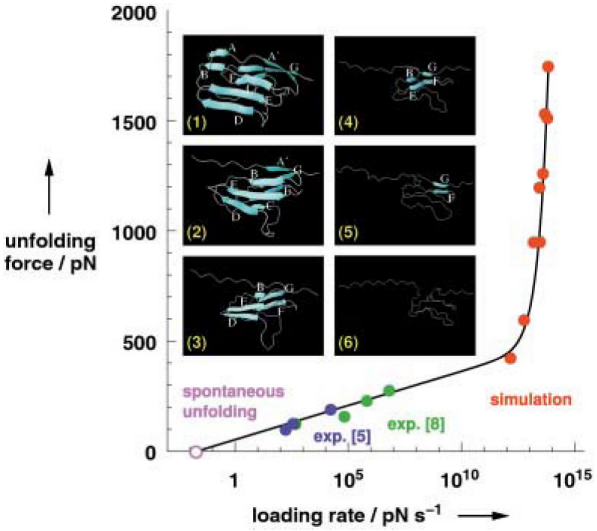


Figure 2: Example of the unfolding of a soluble protein – titin Ig – both measured (blue and green) and calculated (red). The loading rate is inversely proportional to the amount of time required for the unfolding (from (37)).

Molecular dynamics of lipid systems and membrane proteins

In the past few years, growing availability of structural information related to membrane proteins and technical improvements have enabled MD simulations of lipid bilayers and membrane proteins or peptides, which has become a very active subject of research as reported in two recent reviews (15, 20).

Among the methodological improvement for the simulation of lipids, one finds the optimization of force fields (31), the development of particle mesh Ewald summation to provide a better treatment of electrostatics than cut-off methods. It has also been shown that the simulation box should have flexible dimensions (NpγT ensemble) (46).

Whereas the simulation of globular proteins in vacuum may have produce reasonable results for the cases where the force fields parameters would implicitly include solvation effects, a similar

approach seems to be inapplicable for membranes and membranes proteins (15). Various methods have been developed for taking into account the membranes, including the use of an effective membrane-mimic potentials, simplified bilayers models and simulations in explicit lipid bilayers (15). The simulation of explicit lipid bilayers has proven to be the most realistic, with the advantages that the protein and the membrane are represented in microscopic details without assumption made regarding the dielectric properties of the whole system (15). Also, it allows simulating more complex lipid bilayers, where more than one type of lipids are included (20). The development of mixed lipids membrane models are in development, but will be soon available for membrane protein simulations.

Simulations of membrane proteins were pioneered with single peptides forming TM α-helices (reviewed in (20)), in order to refine sequence-based prediction of TM helix length (21) and to study the conformational dynamics and the role of proline-induced helix kink (47). Another example is the study of the voltage-dependent insertion of alamethicin at phospholipids/water interfaces was also studied, leading to a two-steps mechanism for the insertion (45).

The simulation of membrane peptides has been extended to the simulation of peptide bundles, including alamethicin (44), gramicidin (2, 10), various model for the TM segment of vpu from HIV-1 (7, 34), various models for the TM segments of the influenza A virus M2 protein (19, 49) and the M2δ of the nicotinic acetylcholine receptor (44). In the case where the exact three-dimensional structure was not available, models for various oligomeric states and with different protonation states were built and simulated. The criteria for the best models are the stability of the helical bundle in the membrane and the overall agreement with available experimental data (e.g. tilt angle of the helices, the spatial location and dynamical behaviour of functionally important residues).

The simulation of integral membrane proteins are yet among the most complex systems studied by MD simulations and can include in some cases the simulations of membrane protein oligomers (e.g. OmpF as a trimer in a hydrated lipid bilayer (43)), where the total number of atoms approaches 10^5. Throughout the literature, there are several examples of integral membrane proteins simulated in explicit lipid bilayers, including the bacterial potassium channel KcsA (41), mechanosensitive ion channel (MscL) (6), OmpF porin (24, 25, 43), bacteriorhodopsin (23) and bovine rhodopsin (8).

Due to the complexity of the conformational space for such molecule and the relatively short timescale available (10-100 ns), the choice of the initial conditions are important and requires an a priori knowledge of the protein conformation (usually derived from experiments). For the same reason, the range of questions that can be addressed is also limited to the stability/flexibility of the protein structure, the protein-lipid interactions, protein-induced perturbations to the bilayer, ligand-protein interactions provided a binding-pocket is identified. Other questions like folding and insertion in the membrane are beyond the available computational capabilities.

In summary, the main challenge for MD simulations remains the need of increasing the duration of the simulations in order to be able to follow long timescale motions (20), which would lead to an improved conformational space exploration and more accurate results.

SYSTEM STUDIED BY MD SIMULATIONS

The systems presented in this thesis reflect the diversity of membrane protein and peptides. They include bovine rhodopsin an integral membrane protein, Vpu from HIV-I a single membrane-spanning membrane protein which has the ability to form ion-channels and WALPC23 a synthetic peptide representing a consensus sequence for TM protein segments and which has special relevance for the packing of helices in membrane proteins.

Bovine rhodopsin

Bovine rhodopsin is a widespread photoreceptor, responsible for vision in the vertebrates. The protein contains 348 amino acids, has a molecular mass of ~40 kDa, and folds into seven TM helices varying in length from 19 to 34 residues and one cytoplasmic helix. It is the most intensively studied member of the GPCR family (18) and has been a model system for years (17) because it is the only member that is naturally present in high abundance in biological tissues (1).

Chapter 2.1 describe extensively the protein and its interactions with the chromophore, retinal, which is responsible for triggering the protein through conformational changes resulting after the absorption of a photon.

WALP peptides

WALP is an example of synthetic hydrophobic peptide which forms well-defined and well-characterized transmembrane α-helices(13). They are commonly used to study the interactions of α-helices with the surrounding lipid membrane (28, 29).

WALP peptides consist of an alternating alanine-leucine stretch of variable length and form a hydrophobic α-helix and are flanked on both ends by two tryptophans. These residues are often found in membrane-spanning proteins at the hydrophobic-hydrophilic interface of the bilayer (36). WALP peptides are usually associated with a number indicating the total number of residue in the peptidic chain. For example, WALP23 contains 23 residues in total.

WALP peptides have mainly been used to study the effect of mismatch between the hydrophobic transmembrane regions of integral membrane proteins and the hydrophobic core of the lipid bilayer. Such hydrophobic mismatch is thought to have an important effect on lateral segregation in membranes (35), on membrane protein activity and stability, and on protein sorting (14, 28). WALP can significant influence the organization of phosphatidylcholine model membranes, depending on the relative length of the peptide with respect to bilayer hydrophobic thickness (either too short or too long) (11, 13, 28, 29). It has been found that upon decreasing the relative hydrophobic length of the peptides, WALP peptides can induce the formation of the nonlamellar phases (H_{II} phase) in fluid PC bilayers (29) and alter the thickness of fluid PC bilayers (12). These changes seem to adequately relieve the hydrophobic mismatch, because the acyl chains of the lipids in these nonlamellar phases are more disordered, and, hence, their effective hydrophobic length is smaller.

WALP peptides have also been used to investigate the anchorage of transmembrane segments in bilayers, as a result, tryptophans residues have been shown to preferentially locate at the interface of the lipid headgroups and acyl chains (13).

Vpu from HIV–1

Vpu is an 81-residue phosphoprotein, of molecular mass 16 kDa, encoded human immunodeficiency virus type-1 (HIV-1) (5, 42). Vpu associates with the Golgi and ER membranes (16) and is composed of a hydrophobic N-terminal TM domain (residues 1-27) and a hydrophilic C-terminal cytoplasmic domain (54 residues) (33). Vpu has 2 major biological functions in the life cycle of the virus: (a) The N-terminal TM domain is responsible for enhancement of viron release from the cell surface, by formation of an ion channel altering the electrochemical gradient (9, 38). (b) The C-terminal cytoplasmic domain mediates in the ER the selective degradation of the CD4 protein through ubiquitination (9, 40).

Vpu is capable of self homo-oligomerization in the presence of lipid membranes both in vitro and vivo, with TM helices forming bundles (3, 33). Whilst the exact number of helices involved is uncertain, a pentamer or hexamer formation is favoured. Gel electrophoresis showed complexes of ~ 80 Kda, which correspond to the pentamer form (33). Previous MD simulations studies tend to favour the pentamer form as the most plausible oligomeric state. Lopez et al. (32) observe in a 3 ns simulation the disruption an hexameric bundle in to a pentamer and a monomer.

Vpu forms cation-selective channels in planar lipid bilayers and also in the plasma membrane of E. coli cells expressing Vpu. A synthetic peptide of residues 1-27 of Vpu has been shown to exhibit the same ion channel activity, indicating that the C-terminal does not have a direct role in generating a TM ion current (39). This peptide showed that Vpu does not discriminate Na^+ over K^+, shows marginal permeability to divalent cations and that conductance appears to be pH independent. H^+ permeability of membrane vesicles is unaffected by the presence of Vpu, suggesting that Vpu does not conduct H^+ (30).

The role of Vpu's ion channel is unclear, but it could contribute to HIV pathogenesis by destabilising host cell proteins. Expression of Vpu in Xenopus oocyte membranes decreases K^+ conductance through K^+ channels (4). The characterization of this inhibition suggests that Vpu has a negative effect on the expression of endogenous membrane proteins in host cells. The K^+ channel inhibition appears to dependent on the cytoplasmic domain of the protein, and thus associated with the CD4 degradative capacity of Vpu.

REFERENCES

1. Albert AD, Yeagle PL. 2002. Structural studies on rhodopsin. *Biochim Biophys Acta* 1565: 183
2. Allen TW, Bastug T, Kuyucak S, Chung SH. 2003. Gramicidin A channel as a test ground for molecular dynamics force fields. *Biophys J* 84: 2159
3. Bechinger B, Kinder R, Helmle M, Vogt TC, Harzer U, Schinzel S. 1999. Peptide structural analysis by solid-state NMR spectroscopy. *Biopolymers* 51: 174
4. Coady MJ, Daniel NG, Tiganos E, Allain B, Friborg J, et al. 1998. Effects of Vpu expression on Xenopus oocyte membrane conductance. *Virology* 244: 39
5. Cohen EA, Terwilliger EF, Sodroski JG, Haseltine WA. 1988. Identification of a protein encoded by the *vpu* gene of HIV-1. *Nature* 334: 532
6. Colombo G, Marrink SJ, Mark AE. 2003. Simulation of MscL gating in a bilayer under stress. *Biophys J* 84: 2331
7. Cordes FS, Kukol A, Forrest LR, Arkin IT, Sansom MSP, Fischer WB. 2001. The structure of the HIV-1 Vpu ion channel: modelling and simulation studies. *Biochimica et Biophysica Acta* 1512: 291
8. Crozier PS, Stevens MJ, Forrest LR, Woolf TB. 2003. Molecular dynamics simulation of dark-adapted rhodopsin in an explicit membrane bilayer: coupling between local retinal and larger scale conformational change. *J Mol Biol* 333: 493
9. Cullen BR. 1998. HIV-1 auxiliary proteins: Making connections in a dying cell. *Cell* 93: 685
10. de Groot BL, Tieleman DP, Pohl P, Grubmuller H. 2002. Water permeation through gramicidin A: desformylation and the double helix: a molecular dynamics study. *Biophys J* 82: 2934
11. de Planque MR, Goormaghtigh E, Greathouse DV, Koeppe RE, 2nd, Kruijtzer JA, et al. 2001. Sensitivity of single membrane-spanning alpha-helical peptides to hydrophobic mismatch with a lipid bilayer: effects on backbone structure, orientation, and extent of membrane incorporation. *Biochemistry* 40: 5000
12. de Planque MR, Greathouse DV, Koeppe RE, 2nd, Schafer H, Marsh D, Killian JA. 1998. Influence of lipid/peptide hydrophobic mismatch on the thickness of diacylphosphatidylcholine bilayers. A ^{2}H NMR and ESR study using designed transmembrane alpha-helical peptides and gramicidin A. *Biochemistry* 37: 9333
13. de Planque MR, Kruijtzer JA, Liskamp RM, Marsh D, Greathouse DV, et al. 1999. Different membrane anchoring positions of tryptophan and lysine in synthetic transmembrane alpha-helical peptides. *J Biol Chem* 274: 20839
14. Dumas F, Lebrun MC, Tocanne JF. 1999. Is the protein/lipid hydrophobic matching principle relevant to membrane organization and functions? *FEBS Lett* 458: 271
15. Efremov RG, Nolde DE, Konshina AG, Syrtcev NP, Arseniev AS. 2004. Peptides and proteins in membranes: what can we learn via computer simulations? *Curr Med Chem* 11: 2421
16. Ewart GD, Mills K, Cox GB, Gage PW. 2002. Amiloride derivatives block ion channel activity and enhancement of virus-like particle budding caused by HIV-1 protein Vpu. *Eur Biophys J* 31: 26
17. Ferretti L, Karnik SS, Khorana HG, Nassal M, Oprian DD. 1986. Total synthesis of a gene for bovine rhodopsin. *Proc Natl Acad Sci U S A* 83: 599
18. Filipek S, Teller DC, Palczewski K, Stenkamp R. 2003. The crystallographic model of rhodopsin and its use in studies of other G protein-coupled receptors. *Annu Rev Biophys Biomol Struct* 32: 375
19. Forrest LR, Kukol A, Arkin IT, Tieleman DP, Sansom MS. 2000. Exploring models of the influenza A M2 channel: MD simulations in a phospholipid bilayer. *Biophysical Journal* 78: 55

20. Forrest LR, Sansom MS. 2000. Membrane simulations: bigger and better? *Curr Opin Struct Biol* 10: 174
21. Forrest LR, Tieleman DP, Sansom MS. 1999. Defining the transmembrane helix of M2 protein from influenza A by molecular dynamics simulations in a lipid bilayer. *Biophys J* 76: 1886
22. Grubmüller H, Heymann B, Tavan P. 1996. Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science* 271: 997
23. Hayashi S, Tajkhorshid E, Schulten K. 2002. Structural changes during the formation of early intermediates in the bacteriorhodopsin photocycle. *Biophys J* 83: 1281
24. Im W, Roux B. 2002. Ion permeation and selectivity of OmpF porin: a theoretical study based on molecular dynamics, Brownian dynamics, and continuum electrodiffusion theory. *J. Mol. Biol.* 322: 851
25. Im W, Roux B. 2002. Ions and counterions in a biological channel: a molecular dynamics simulation of OmpF porin from *Escherichia coli* in an explicit membrane with 1 M KCl aqueous salt solution. *J. Mol. Biol.* 319: 1177
26. Karplus M. 2002. Molecular dynamics simulations of biomolecules. *Acc Chem Res* 35: 321
27. Karplus M, McCammon JA. 2002. Molecular dynamics simulations of biomolecules. *Nat Struct Biol* 9: 646
28. Killian JA. 1998. Hydrophobic mismatch between proteins and lipids in membranes. *Biochim Biophys Acta* 1376: 401
29. Killian JA, Salemink I, de Planque MR, Lindblom G, Koeppe RE, 2nd, Greathouse DV. 1996. Induction of nonbilayer structures in diacylphosphatidylcholine model membranes by transmembrane alpha-helical peptides: importance of hydrophobic mismatch and proposed role of tryptophans. *Biochemistry* 35: 1037
30. Lamb RA, Pinto LH. 1997. Do Vpu and Vpr of human immunodeficiency virus type 1 and NB of influenza B virus have ion channel activities in the viral life cycles? *Virology* 229: 1
31. Lazaridis T, Karplus M. 2000. Effective energy functions for protein structure prediction. *Curr Opin Struct Biol* 10: 139
32. Lopez CF, Montal M, Blasie JK, Klein ML, Moore PB. 2002. Molecular dynamics investigation of membrane-bound bundles of the channel-forming transmembrane domain of viral protein U from the human immunodeficiency virus HIV-1. *Biophys J* 83: 1259
33. Maldarelli F, Chen MY, Willey RL, Strebel K. 1993. Human-immunodeficiency-virus type-1 Vpu protein is an oligomeric type-I integral membrane protein. *J. Virol* 67: 5056
34. Moore PB, Zhong Q, Husslein T, Klein ML. 1998. Simulation of the HIV-1 Vpu transmembrane domain as a pentameric bundle. *FEBS Lett* 431: 143
35. Mouritsen OG, Bloom M. 1984. Mattress model of lipid-protein interactions in membranes. *Biophys J* 46: 141
36. Reithmeier RA. 1995. Characterization and modeling of membrane proteins using sequence analysis. *Curr Opin Struct Biol* 5: 491
37. Rief M, Grubmüller H. 2002. Force spectroscopy of single biomolecules. *Chemphyschem* 3: 255
38. Schubert U, Bour S, Ferrer-Montiel AV, Montal M, Maldarelli F, Strebel K. 1996. The two biological activities of human immunodeficiency virus type 1 Vpu protein involve two separable structural domains. *J. Virol.* 70: 809
39. Schubert U, Ferrer-Montiel AV, Oblatt-Montal M, Henklein P, Strebel K, Montal M. 1996. Identification of an ion channel activity of the Vpu transmembrane domain and its involment in the regulation of virus release from HIV-1-infected cells. *FEBS Lett.* 398: 12
40. Schubert U, Henklein P, Boldyreff B, Wingender E, Strebel K, Porstmann T. 1994. The human immunodeficiency virus type 1 encoded Vpu protein is phosphorylated by casein kinase-2 (CK-2) at positions Ser52 and Ser54 within a predicted alpha-helix-turn-alpha-helix-motif. *J. Mol. Biol.* 236: 16
41. Shrivastava IH, Sansom MS. 2002. Molecular dynamics simulations and KcsA channel gating. *Eur Biophys J* 31: 207
42. Strebel K, Klimkait T, Martin MA. 1988. Novel gene of HIV-1, *vpu*, and

its 16-kilodalton product. *Science* 241: 1221

43. Tieleman DP, Berendsen HJ. 1998. A molecular dynamics study of the pores formed by Escherichia coli OmpF porin in a fully hydrated palmitoyloleoylphosphatidylcholine bilayer. *Biophys J* 74: 2786
44. Tieleman DP, Berendsen HJ, Sansom MS. 1999. An alamethicin channel in a lipid bilayer: molecular dynamics simulations. *Biophys. J.* 76: 1757
45. Tieleman DP, Berendsen HJ, Sansom MS. 2001. Voltage-dependent insertion of alamethicin at phospholipid/water and octane/water interfaces. *Biophys J* 80: 331
46. Tieleman DP, Marrink SJ, Berendsen HJ. 1997. A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems. *Biochim Biophys Acta* 1331: 235
47. Tieleman DP, Shrivastava IH, Ulmschneider MR, Sansom MS. 2001. Proline-induced hinges in transmembrane helices: possible roles in ion channel gating. *Proteins* 44: 63
48. van der Spoel D, van Buuren AR, Apol E, Meulehhoff PJ, Tieleman DP, et al. 2002. *Gromacs User Manual, version 3.1.1*. Groningen: Department of Biophysical Chemistry, University of Groningen
49. Wu Y, Voth GA. 2003. Computational studies of proton transport through the M2 channel. *FEBS Lett* 552: 23

II.1 Conformational Studies of Retinal in Rhodopsin

Vincent Lemaître,[1,2] and Anthony Watts,[1]

[1] *Biomembrane Structure Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK ; email: anthony.watts@bioch.ox.ac.uk*

[2] *Nestec S.A., BioAnalytical Department, Vers-Chez-Les-Blanc, CH-1000 Lausanne 26, Switzerland*

Key words retinal, rhodopsin, photo-intermediate, photorhodopsin, bathorhodopsin, lumirhodopsin, metarhodopsin, NMR, RR, DARR, DQ, x-ray crystallography, circular dichroism, DFT, vision, phototransduction

■ **Abstract** The rhodopsin-bound retinal is a device for converting light into molecular conformational changes, leading first to the activation of rhodopsin itself and then the initiation of the signal transduction pathways leading to vision. The conformation of retinal and the changes resulting from light-absorption are crucial to an understanding of the mechanism of activation of rhodopsin, and this has been the subject of intense research where various experimental and theoretical techniques have been applied. This review summarizes the major results obtained by optical and solid state NMR spectroscopies, x-ray diffraction and *ab initio* theoretical methods, from the dark-adapted state through the various photo-activated intermediate states. An overview of the rhodopsin features relevant for the conformation of its chromophore retinal is also included, in particular the binding-pocket and the molecules of water present located near retinal.

INTRODUCTION

Vision is a complex process in which light is converted into a nervous impulse through a chain of biochemical reactions taking place along various specialized cells and tissues, from the retina of the eye to the brain. In the eye, specialized cells of the retina called rod cells contain a protein acting as a photoreceptor, rhodopsin and possess the unique ability to convert and generate the neuronal

signal. In order to become activated and induce the enzymatic cascade leading to vision (4, 117, 183, 197, 231, 282, 302), rhodopsin relies on its chromophore, retinal. Retinal is an organic molecule (Figure 1) covalently-bound to rhodopsin in anticipation of its photo-activation.

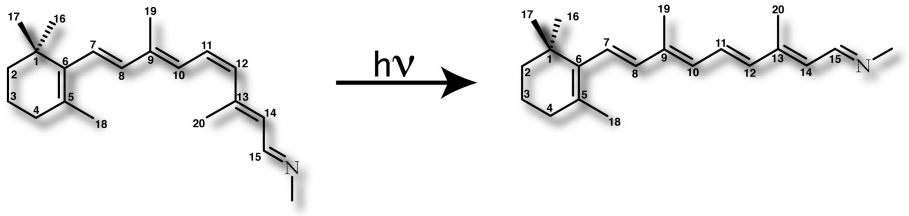


Figure 1: Chemical structure of 11-*cis*-retinal and all-*trans*-retinal.

The 11-*cis*-retinal isomer is a key photosensitive component of rhodopsin. It is a molecular trigger capable of converting the electromagnetic energy carried by a photon into mechanical energy to be transferred to the protein in order to induce slower conformational changes, eventually leading to the initialization of the chain of biochemical reactions leading and finally to neural signal. To perform its function, 11-*cis*-retinal undergoes a fast photoisomerization leading to all-*trans*-retinal, a process which triggers the conformational changes of the receptor to its active form (metarhodopsin II) (117, 231, 282, 302).

Considerable effort has led to a picture of the events taking place in rhodopsin at the molecular level over four decades. This review summarizes the important milestones in this process of elucidation of vision with a special emphasis on the initial trigger, retinal, with special emphasis on rationalization of the disparate bodies of data from all methods.

RHODOPSIN, A GPCR

Rhodopsin is a G protein-coupled receptor (GPCR), which, when activated, interacts with a heterotrimeric G protein (transducin in retina cells) to convert an extracellular signal into an intracellular signal. This superfamily is made of several families and rhodopsin has been classified as belonging to family A, the largest subfamily of GPCRs, gathering most of those identified so far. The family includes cone pigments, β-adrenergic and other ligand receptors (63, 74, 113, 148, 221, 231).

Most GPCRs respond to the binding of a ligand and undergo multiple conformational changes upon agonist binding (91). For rhodopsin, the signal is made up of two components: the covalently-bound chromophore, which undergoes *cis* to *trans* photoisomerization, and a photon (123, 182, 233).

Rhodopsin is the most intensively studied member of the GPCR family (31, 81, 82, 123, 233, 240) and has been a model system for years (80) because it is the only member that is naturally present in high abundance in biological tissues (4). Historically, rhodopsin has played a prominent role in the study of GPCRs, being the first GPCR to be sequenced (80), the first GPCR for which the role of individual amino acid was connected with the protein function (160) (with several examples: the role of the counter ion Glu113 (232) and other critical residues (85, 221)). Furthermore, rhodopsin triggered the study of other class of protein like the study of G proteins and transducin in particular (162, 194). Initial suggestions for spectral transduction have been formulated (159, 300, 304) with rhodopsin.

Regarding the expression, purification and reconstitution in membranes for the purpose of structural study, rhodopsin is an exception compared to other GPCRs. The protein is obtained from natural tissues where it is present in large quantities. Methods have been developed to isolate ROS disk membranes (246), and relatively large amounts of natural membrane containing predominantly one membrane protein can be obtained for structural study. Subsequently the purification of rhodopsin is achieved using an affinity column in presence of detergent (262). When extracted into detergent, rhodopsin maintains its chromophore (158). The reconstitution of purified rhodopsin into membranes of defined lipid is also achievable (131, 228). More recently, rhodopsin has been successfully expressed in heterologous system, HEK cells (65, 66) and baculovirus/sf9 (50), allowing the labelling of amino acids within the protein.

For the purpose of structural studies, rhodopsin has been studied either reconstituted into native membrane, in lipid bilayers, in micelles, lyophilized or as crystals. These various preparations have been shown to maintain the properties of retinal regarding the absorption of visible light and seem to be suitable to study the protein in its dark-adapted state. However, it has also been shown that the

ability of rhodopsin to become activated is very much dependent on the lipid composition of the membrane where it resides (see below).

The binding site for retinal is viewed as representative of the binding mode of an inverse agonist in other GPCRs (179), stabilizing the inactive conformation of rhodopsin (179), reducing the intrinsic receptor activity and leading to low basal activity. The activity against transducin (G_t) for the 11-*cis*-retinal-bound rhodopsin ground state is even lower that the one of the ligand-free opsin, which is itself 10^{-6} of the activity of the all-*trans*-bound metarhodopsin II state, indicating the magnitude of the inverse agonism imposed by the ligand (181). Such a large effect has to result in strong structural features (197). The photo-induced isomerization of the chromophore to the all-*trans* forms turns the chromophore into an agonist that triggers the transition of rhodopsin to the active state (229).

Bovine rhodopsin contains 348 amino acids, has a molecular mass of ~40 kDa, and folds into seven transmembrane helices varying in length from 19 to 34 residues and one cytoplasmic helix. In addition, rhodopsin exhibits other features common with many other GPCRs, like sites of asparagine-linked glycosylation in the amino-terminal region (110). There are two carbohydrate moieties at Asn2 and Asn15 that are oriented towards the lumen (extra-cellular) face (87, 110, 112). Two major species identified as $Man_3GlcNAc_3$ are attached to Ans2 and Asn15 in bovine rhodopsin (114) and later confirmed using mass spectrometry (7). Two palmitate-linked cysteines are close to the carboxyl-terminal region (145) at Cys322 and Cys323 in the cytoplasmic region (63, 145, 188, 195, 203, 211, 216, 272, 275). A disulfide bridge linking two extra-cellular loops is present between Cys110 and Cys187 (54, 127, 144, 146). Finally, the protein shows serine (Ser 334, 338 and 343) and threonine phosphorylation sites in the carboxyl-terminal region (112). The phosphorylation of these residues is a mechanism to desensitize signalling to G proteins (196).

Although the importance of dimerization or even the oligomerization has been emphasized important the function for many GPCRs (1, 5, 9, 37, 40, 99, 100, 151, 164, 184, 204, 256, 273), it is still controversial for rhodopsin. Recent Atomic Force Microscopy (AFM) studies have provided a direct observation of mice rhodopsin in native membrane, showing the proteins are organized as dimmers forming paracrystalline arrays (83, 84, 169). This finding contradicts many studies conducted over the last thirty years, where such behaviour had been set aside based on indirect methods of observation (21-23, 41, 43, 62, 170-172, 219, 230, 293). Based on the AFM data, a

molecular model as been proposed (169) and it has been postulated that the organization of rhodopsin could have strong implications for its mechanism of activation and regulation (83, 84, 169). The lipid-protein ratio in the membranes of rod cells outer segments is 60:1, where 95% of the proteins present in the membrane is rhodopsin (157, 212, 254). It is therefore expected to observe rhodopsin in close vicinity from each other and it cannot be precluded that these patterns are simply some packing arrangement allowing such a high-protein content to take place in the membrane, not necessarily directly related with the activation or regulation of the protein, like observed in previous AFM studies for synthetic peptide in bilayers (WALP) (224, 225). A functional study showing the implication of this supramolecular organization either on the activation or regulation of the protein is needed and would emphasize further the importance of this discovery.

STRUCTURAL DESCRIPTION OF BOVINE RHODOPSIN

Crystal structures

The crystal structure of bovine rhodopsin is the first and only detailed three-dimensional structural model for a full GPCR where its tertiary structure is resolved. The crystal structure of bovine rhodopsin was published in 2000 (205), resolving more than 93% of the amino acid side chains (Figure 2). It was subsequently refined (269), and further improvement in the resolution were achieved, leading to structures in which the position of water molecules within the protein were defined (168, 198). In one of the latest crystal structure, a resolution of 2.2 Å was obtained, finally resolving the accurately conformation of the chromophore (200).

Single-crystal micro-spectrophotometry has confirmed the relevance of three-dimensional models derived from x-ray crystallography (198), showing it is possible to follow the formation of the spectroscopic intermediary states of the chromophore upon the photo activation. From this point of view, the technique applied to crystal samples of rhodopsin showed the crystal to behave in a similar way as rhodopsin in suspension or membrane preparation regarding the absorption of light (198). It is possible to observe transition from bathorhodopsin to lumirhodopsin, as well as detecting intermediate metarhodopsin I and II. The spectra observed are consistent with previous

experiments on rhodopsin observed in either membrane or solution environment (198). The fact that rhodopsin constitutes 95% of the protein content of a membrane (157, 212) where protein-lipid ratio is 1:60 (254), i.e. an environment extremely concentrated, might account for the relevance of the x-ray structure, which also contains a few lipids.

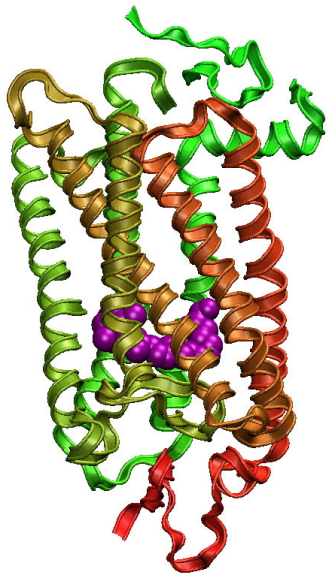


Figure 2: Molecular model of rhodopsin in its dark-state determined by X-ray crystallography with a resolution of 2.8 Å (205).

In an alternative approach, solution NMR applied on fragment peptides synthesized from rhodopsin amino-acid sequence allowed the determination of the secondary structure of the protein as well as short-range distance restraints (3, 147, 294, 296-299). These NMR data were combined with long-range distance constraints obtained by other experimental techniques (spin-label, solid state NMR, etc.) and using simulated annealing a three-dimensional model for rhodopsin in the dark-state was built (295). This approach is complementary to the crystal structure, bringing similar results to the crystal structure (295).

Organization of the transmembrane core

The x-ray crystal structure showed the expected seven transmembrane domains that had been predicted both computationally and from the low resolution data (198, 199, 205, 258, 269). The secondary structures, more specifically the definition of the transmembrane helices have shown to

be slightly affected (up to one helical turn for helices V and VI) by the space group of the crystal (168).

The helices differ in their length and are irregular. They tilt at various angles with respect to the expected membrane surface. Helices contain a mix of α- and 3_{10}-helices, possess a large number of kinks, twists (222). Furthermore, it has revealed that all transmembrane helices are bent to different extents, except for helix III, located in the middle of the helical bundle (205). Many of the bends and twists were spotted in the two-dimensional cryo-electron microscopy study of frog rhodopsin (236), preceding the three-dimensional x-ray crystal structure. The helical axes from that study can be readily superposed with the helices in the 1HZX, the refined model of 1F88 (269) (1HZX and 1F88 refers to the ID number of the structure in the Protein DataBank (12)). Further refinements of the x-ray crystal structure have not reduced the number of these conformational anomalies and therefore they are not seen as artefacts coming from the data set resolution or the method used for crystallographic refinement (269).

The kinks are all located near the retinal binding-pocket at proline and glycine sites. They are thought to serve as molecular hinges involved in the conformational changes associated with the chromophore isomerization (229) and to facilitate inter-helical polar interactions (168).

Helix II is kinked due to the presence of a GlyGly motif in the middle of the helix. Helix III is the most severely tilted helix of the bundle in the dark-adapted state. At the other extreme helices IV, VI and VII were suggested to be the closest to perpendicular to the membrane surface (236, 237). These features have survived into the most recent three-dimensional structure determinations for rhodopsin (4). Helix IV is the shortest helix and is significantly bent due to the presence of Pro170 and Pro171 (48). Helix V contains several residues that are part of the binding-pocket (Met207 and Phe212), it has been recently proposed to initiate the activation of the protein upon the isomerization of retinal (213). Helix VI seems to interact with other helices of the bundle only through van der Waals interactions. It shows considerable distortion due to the presence of Pro267, Pro291 and Pro303, a part of the highly conserved NPXXXY motif. Trp265 is a highly conserved residues among GPCRs (201). Helix VII contains the chromophore and interacts with all helix except helices IV and V. Helix VII is kinked near the chromophore, the helix showing some

distortion from ideality, where a few residues are folded in a 3_{10}-helix rather than an α-helix (168). The cytoplasmic helix, termed helix VIII, is located at the extension of helix VII. This helix in addition to the cytoplasmic loops II and III (reviewed in (116)) of rhodopsin, is part of the binding sites for transducin α-subunit ($G_t\alpha$) and plays a role in the regulation of the transducin γ-subunit ($G_t\gamma$) binding (69). Helix VIII is straight and amphipathic. The helix starts at Lys311 and contains Arg314, whose side-chains form hydrogen-bonds with the surrounding residues. Cys323 is the C-terminal residue of the helix, and is palmitoylated.

More recently extensive sequence search resulting in the introduction of a new concept, the helix packing moments (174), have highlighted residues of rhodopsin, which are conserved along the class A GPCR family, and involved in helix-helix interactions. Using the x-ray crystal structure, a possible role in the structure and function of the protein can be suggested for these small or weakly polar amino acids. They are thought to keep the helices bound in a rigid structure through hydrogen-bonding and small residue packing interaction. There is a large number of these residues conserved between helix I to helix IV, helices that are not thought to be mobile upon activation (174). In the case of helix VI, which plays an important role in the activation of the protein, does not have such highly conserved residues (174). Rather, its hydrophobic moment (67) is mainly oriented towards helix III and the interactions between these two helices are mainly conducted through highly conserved aromatic residues (Phe261, Trp265, etc.). The different nature of the interactions between helix III and VI might be directly related with their importance in the activation of rhodopsin. Interestingly, there are almost none of these small and weakly polar residues at the interface between helix V and VI. Furthermore helix V and VI are loosely packed in the x-ray structure (much more loosely than helices VI and VII), suggesting a potential role for helix V in the activation of the receptor (174), as shown in recent solid state NMR experiments (213).

Organization of the loops

The intracellular loops (cytoplasmic) are the less well defined by x-ray crystallography (some parts of these loops areas were not defined until recently) and are the regions with the highest B-factors. The last two crystal structures, where bovine rhodopsin was crystallized in two different space groups and where all residues were defined, highlight the flexibility of these loops, which is the area

of the protein where the most significant differences between the two crystal structures take place. Only a few interactions between these loops are observed. The commonly accepted explanation is that the loops probably mobile in solution. Almost certainly, the C-terminal residues of rhodopsin are also flexible and mobile in the physiological milieu, lacking a definite single conformation (36, 163).

The extracellular loops are more structured. The residues 173 to 198 from the second extra-cellular loop forms a twisted β-hairpin from residues 177 to 190 and contains Cys187, portion of a disulfide bridge. The extracellular loop 2 is part of the binding-pocket with residues Glu181 and Ser186 relatively close from the bound-retinal. Mutation of Cys187 gives rise to proteins that are abnormally glycosylated and cannot bind retinal (146). It also contains Glu181, a part of the β-sheet structure that comes in proximity of retinal, near atom C12. This residue is highly conserved in rhodopsin and short- and middle-wavelength visual pigments (63).

Organization of the binding-pocket

The chromophore is bound to Lys296 in transmembrane helix VII by a protonated Schiff base (122, 220). The counter ion Glu113 is located in the vicinity of the Nε atom of Lys296 (distance between N and the carboxylic oxygens are 3.3 Å and 3.5Å) (205).

Glu113 is highly conserved among all known vertebrate visual pigments (192, 232, 304). Upon studying how the mutation of charged residues on helix III and pH changes affect the maximum of absorption for the chromophore, Sakmar et al. concluded that Glu113 is the counter ion (232). This was later confirmed by x-ray crystallography (205). Glu113 has three important functions:

a. It stabilizes the protonated Schiff base by increasing the K_a for this group by up to 10^9. The pKa of the Schiff base in the binding-pocket is > 12-16, while 6-7 in solution (7.4 in alcohol (257)) and preventing its spontaneous deprotonation (64, 112, 257).
b. It is responsible for a major part of the bathochromic shift to 500 nm in the maximum absorption for visual pigments and to make them more sensitive to longer wavelengths as UV light is filtered by the front of the eye in most animals (304).

c. It stabilizes the connection between helix III and VII and is essential for locking the receptor in its inactive state (42). The disruption of this salt bridge, due to a mutation for example (e.g. E113N (226, 232)), results in a constitutive activation of rhodopsin in the absence of the chromophore (226). The addition of 11-*cis*-retinal, even not covalently-bound to rhodopsin is enough to suppress the constitutive activity (226). Indeed a mutant lacking Lys296 can bind 11-*cis*-retinal analogue (27, 111) and remain inactivated just through non-covalent interactions with the ligand (305). On the other hand, the proton carried by the Schiff base is not critical for the formation of metarhodopsin II, as it was shown that non-prototnated Schiff base can form the activated state and activate transducin (72, 73, 307). This neutralization process corresponds to binding of a cationic part in a diffusible ligand to other rhodopsin-like receptors (259).

As the disruption of this salt bridge (226, 306) upon proton transfer is thought to trigger conformational changes in rhodopsin which are necessary for G protein activation (76), the nature of the Schiff base – counter ion interaction in the ground state of rhodopsin has been a focus of many previous studies (64, 65, 118, 277).

The crystal structures of rhodopsin allow identifying the residues of the protein that are part of the binding-pocket, as well as their relative orientations (Figure 3) (205). Improvement in the resolution and the availability of crystal structures sharing different space groups have confirmed the organization of the binding-pocket, where twenty-seven amino acids are located within 4.5 Å of the retinal moiety (Tyr43, Met44, Leu47, Thr94, Glu113, Ala117, Thr118, Gly120, Gly121, Glu122, Tyr178, Glu181, Ser186, Cys187, Gly188, Ile189, Tyr191, Met207, His211, Phe212, Phe261, Trp265, Tyr268, Ala269, Phe293, Ala295 and Lys296) (Figure 3) (168, 198, 200, 205).

Details of importance for the mechanism of activation

Gly121 is a highly conserved residue near the middle of helix III, known to be involved in chromophore-protein interactions and containing the chromophore Schiff base counter ion at position Glu113. The replacement of Gly121 resulted in a relative reversal in the selectivity of the opsin apoprotein for reconstitution with 11-*cis*-retinal over all-*trans*-retinal in cone cell membranes. These results suggest that Gly121 is an important and specific component of the binding-pocket for

11-*cis*-retinal in rhodopsin (103). Mutation of Gly121 in rhodopsin causes 11-*cis*-retinal to act as a partial agonist rather than an inverse agonist, allowing the mutant to activate transducin in the dark (104). Glu122 has been reported to be protonated in the ground-state (71).

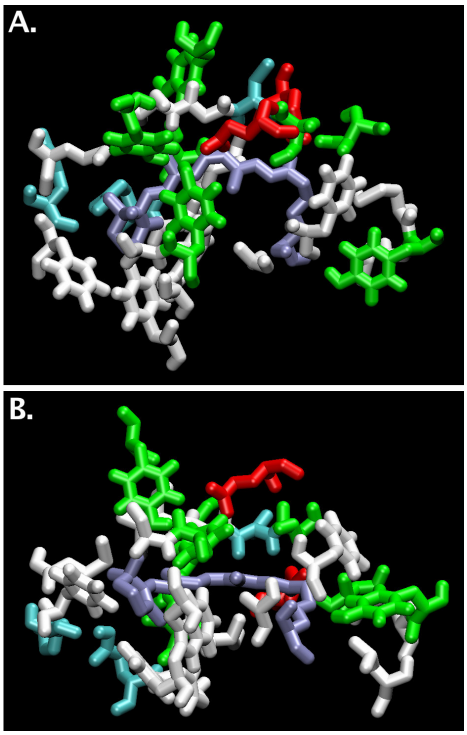


Figure 3: Binding-pocket of rhodopsin in its dark-state determined by X-ray crystallography with a resolution of 2.6 Å (198) (**A.** view from top, **B.** view from side). The conformation for the retinal is the 6-s-*cis* conformer proposed by (263), which fulfills most of the experimental restraints determined by NMR and circular dichroism. Different residues types are highlighted with different colors: lysine-bound retinal (dark blue), aromatic (white), polar (green), potentially charged (red) and hydrophobic (cyan). Hydrogen atoms where omitted for sake of clarity.

Glu181 is probably protonated in the ground-state of rhodopsin and therefore not serve as the primary counter ion (291). According results of two-photon spectroscopy, the binding-pocket for the dark-adapted state is electrically neutral (20). Furthermore, the maximum of absorption λ_{max} for the Raman vibrational spectrum for the E181Q mutant is nearly identical to the wild-type rhodopsin (292). On the other hand, there is a significant shift in the UV/Vis maximum of absorption (10 nm red-shift) (270), the vibrational spectra fail to show an appropriate band of a protonated carboxylic acid (141, 190) and repeated NMR measurements (49, 105, 187, 251, 277) suggest the presence of a negative charge close to the isomerising double bond. The question has also been addressed by applying static Poisson-Boltzmann calculation and MD simulations in explicit solvent, strongly suggesting the presence of a charge on Glu181 (227). It has also been reported that a corresponding

Glutamic acid residue at the same position serves as a counter ion in retinochrome, a member of the rhodopsin family (270). In long-wavelength visual pigments, such as red visual pigments, a histidine residue occupies this position. The histidine residue and a lysine residue present three residues toward the C-terminus form a chloride ion-binding site (284). Chloride ion, when bound, further causes a bathochromic shift by ~40 nm for these cone visual pigments (51). Since Glu181 is part of a hydrogen-bond network (198), it cannot be excluded that the proton is partly delocalized, which would be an explanation the discrepancies between these various theoretical and experimental methods.

WATER MOLECULES NEAR THE BINDING POCKET

There are molecules of water present in rhodopsin and they have been shown to be essential for a proper photo-cascade. Various level of dehydration can block the photo-cascade at various stage from bathorhodopsin to metarhodopsin I (179).

The calculation of the accessible surface reveals several cavities within bovine rhodopsin. The resolution of the first x-ray crystallographic structures was not high enough to assign any molecules in those holes, especially that these features are not expected for proteins containing helices, since the divergence of the helices will provide space between them that might not be fully occupied by amino acid side-chains (269). Two cavities are present near Glu113, and on the other side of the Schiff base nitrogen atom away from Glu113. Both cavities are buried in the protein with no access to the surface of the protein, but it has been postulated that dynamical motions of the structure should be sufficient to allow water to diffuse into these sites (269).

Various crystallographic structures of rhodopsin with a resolution of 2.65 Å or higher (168, 198, 200) have reported the presence seven water molecules consistently found in the two rhodopsin molecules A and B of the rhodopsin crystal (198). The number of water molecules defined in the crystal varies with the resolution of the crystal structure and the space group of the crystal. More water molecules are present, mainly located near the loop regions, for the better-resolved structures

(168, 200) and are therefore unlikely to influence the binding-pocket and the conformation of chromophore (200).

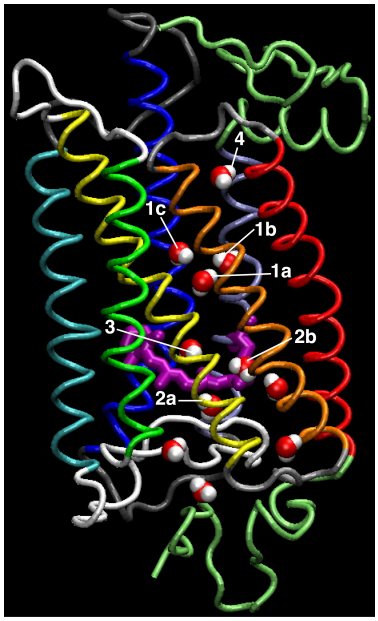


Figure 4: Molecular model of rhodopsin in its dark-state determined by X-ray crystallography with a resolution of 2.6 Å, showing the presence of water molecules present in the protein (198).
From a functional point of view, it has been demonstrated that the activation of rhodopsin is sensitive to the content of bound-water in the protein (186) and that upon transition from metarhodopsin I to metarhodopsin II, the release of bound-water is expected (186).

Interestingly, many of the highly conserved residues among rhodopsin-like GPCRs are found to form binding sites for these molecules of water. For example, Asn73 (2.40), Asp83 (2.50), Cys264 (6.47), Asn302 (7.49) and Tyr306 (7.53) (between bracket, the residue ID according Ballesteros notation (8)) were found to be involved in non-bonded interactions with molecules of water trapped inside the protein (198). Other residues participating in the water binding-sites are particularly conserved in a variety of visual pigments, regardless of the distance from the binding-pocket (198).

Table 1 gathers the protein residues – water interactions found in bovine rhodopsin in the dark-adapted state. These water molecules are displayed with rhodopsin (198) in Figure 4. More water molecules have been identified (168, 200), as suggested by a MD simulation study (124), however none of these newly and consistently found molecules seems to play an important functional role (200).

Table 1: Molecules of water present in rhodopsin according to x-ray crystal structure (1L9H) and H-bond network (198). See Figure 7 for the water ID number.

Water ID	1a	1b	1c	2a	2b	3	4
	Asp83 Gly120	Asp83 Ser298 Asn302	Met257 Tyr301 Asn302	Glu181 Ser186 Cys187	Phe91 Glu113	Cys264 Tyr268 Pro291	Thr62 Asn73 Tyr306

Some of the molecules of water present in rhodopsin appear to be of primordial importance for a proper understanding of the mechanism of action for the protein (198). Indeed, the existence of a hydrogen-bond network involving the side-chains from amino-acid belonging to different helices, can account for two important features of the activation of the protein: (1) the concerted motions during the activation; the water acting as some pieces of glue, maintaining some residues together, while the isomerization of retinal results in the destruction of this network of interactions, leading rhodopsin to its activated conformation(s) and (2) transfer of proton between residues – the presence of water make it possible to transfer proton between residues remotely located, through reorientation of the O-H bond of the water molecules.

The molecules of water present in rhodopsin are shared in a few clusters of water molecules. Some of them are localized around strategic residues of the protein (198).

Water located near Asp83, Gly120 and Asn302 (Cluster 1)

A cluster of three water molecules (referred to as cluster 1) is found near residues Asp83, Gly120 and Asn302. The water cluster mediates interactions with residues in helices II, III, VI and VII by the mean of a hydrogen-bond network, resulting in inter-helical constraints (198) (with the participation of residues Met257 helix VI) and Asp83 (helix II)). Disruption of these constraints is likely to be involved in the activation process (90, 239).

Furthermore, Asp83 is directly hydrogen-bonded to Asn55 (helix I). Thus three highly conserved residues, Ans55, Asp83 and Asn302, constitute a hydrogen-bond chain in the ground state structure of bovine rhodopsin (198). A structural link between Asp83 and Tyr301 (helix VII) by a water molecule is also consistent with IR spectroscopy study on a Y301F mutant (198).

Cluster of water located near the Schiff base (cluster 2)

The second site (referred as cluster 2) contains two molecules of water and the Schiff base. One is in the vicinity of Glu113, the counter ion of the protonated Schiff base and the second is located between the side-chains of Glu181 and Ser 186 (198).

Water molecules are expected to be present close to the protonated Schiff base (50, 65). Upon very thorough dehydration, the absorbance spectrum of rhodopsin is shifted to about 390 nm, suggesting the deprotonation of the Schiff base proton (55). Furthermore, the H-D exchange occurs rapidly at the Schiff base proton upon addition of D_2O, indicating the Schiff base is accessible from the bulk. The exchange takes place even in the dark-state as another evidence of a water molecule close to the Schiff base (88, 178, 202).

Glu181 is located near three other polar residues (Tyr191, Tyr192 and Tyr268 whose hydroxyl groups cover part of the retinal binding-pocket near C9 (198)) therefore there is a continuous hydrogen-bond network lying in the extra-cellular side of the chromophore binding-site from Glu113 to Glu181 (198).

The second molecule of the cluster does not dwell between Glu113 and the Schiff base, but rather is stabilizing the salt bridge by lowering the pKa of Glu113. It also fills a small but prominent spatial gap between helices II and III at the retinal binding-pocket. There is a discrepancy between the interpretation of IR and NMR data at the Schiff base (65, 277) suggesting a “complex counter ion” with a molecule of water between the Schiff base and the counter ion and other experimental techniques. It is not consistent with x-ray (198), other IR spectroscopy studies (189, 190) and with the activity of a E113D mutant (D can still work as a counter ion, resulting in a red-shifted absorption) (132, 137).

Other water molecules

The water molecule located near Glu181 and Ser186 is significant in terms of mechanism of colour regulation in the related retinal proteins. Mutation of Glu181 to His or Gln has been reported to induce shift in the maximum of absorption of the chromophore in the UV-VIS (270), but not for the vibrational spectra (292). This supports the idea that a rearrangement of the residues and water molecules near this cluster of water molecules is most likely to play a crucial role in the spectral tuning of the different type of Rhodopsin (198). For rhodopsin crystallized in the trigonal form, a water molecule between Trp265 and the NpxxY motif was identified (168).

PHOTO-INTERMEDIATES OF RETINAL: KINETICS & ENERGETICS

The chromophore is bound in the hydrophobic core of the molecule, causing its absorption maximum at approximately 380 nm to be shifted bathochromically to its characteristic 500 nm absorption maximum if intact in rhodopsin (155).

The retinal is photosensitive and confers this property to the whole receptor. The irradiation of rhodopsin isomerizes the 11-*cis*-retinal to all-*trans*, which triggers a chain of conformational changes in the opsin that induces an enzymatic cascade leading to vision (117, 183, 197, 231, 282, 302). After photo-activation of rhodopsin, the all-*trans*-retinal is expelled from the receptor and is cycled, through the adjacent pigmented epithelium, to 11-*cis*-retinal, ready to recombine with a bleached opsin (rhodopsin without retinal) to regenerate rhodopsin (4).

It takes about a few milliseconds to form the active species of rhodopsin, the so-called metarhodopsin II. The transition to metarhodopsin II stimulates the binding of the G protein transducin (G_t) (75, 101, 182, 218). Transducin becomes activated (116, 197) and the α subunit of transducin initiates the strongly amplifying cGMP cascade (11) by binding to the phosphodiesterase, and culminating in the hydrolysis of cGMP. The reduction in cGMP levels leads to closure of millions of the plasma membrane Na^+ channels (cGMP cation channels) (93, 171, 260), which results in a hyperpolarization of the plasma membrane that is transmitted to the synapse at the base of the rod cell, followed by synaptic signalling (10, 75, 112, 156, 218, 261).

Energetics

Rhodopsin provides an environment in which its 11-*cis*-retinal chromophore can undergo a *cis* to *trans* conformational switch in response to absorption of a photon with a very high quantum yield of 0.65 (25)–0.67 (17).

The retinal in its binding-pocket is a molecule for which thermal isomerization is very difficult to obtain (i.e. slow: $k \sim 10^{-11}\ s^{-1}$), while the photo-isomerization is extremely fast ($k \sim 10^{12}\ s^{-1}$). The ratio between the two is 10^{23}, compare to 10^7 for the Schiff base in solution. Therefore, rhodopsin

provides an environment where thermal isomerization of retinal is very unlikely compare to the photo-isomerization.

This process is completed in less than 200 fs (217, 238, 283), which is the fastest photochemical reaction known (238). The initial movements of the chromophore are thought to be tightly constrained by the surrounding protein, due to the very short time scale of the photoisomerization (244).

In 1972, it was estimated that the *cis*-to-*trans* transition had was shorter than 6 ps (32). More recently, the availability of lasers able to generated 6 fs long pulses allows to follow the formation of the photoproduct in its very early stage (283). The *cis*-to-*trans* conversion starts within 50-60 fs after the absorption of a photon and is achieved within a few hundreds femtoseconds (~200 fs) (238, 283).

Due to this fast photochemical process, it is believed that a large fraction of the energy of a photon, 32 kcal/mol, is first stored in the chromophore–rhodopsin complex (45, 118). The energy conversion is remarkably efficient since ~60% of the energy of the absorbed photon is stored in the primary product.

Several mechanism for the energy storage in bathorhodopsin have been proposed: (1) charge separation between the positively charged Schiff base and its counter ion (Glu113) (118, 232, 304), (2) distortion of the conjugated backbone of retinal (286) or (3) a combination of the two (17).

One argument in favour of the charge separation mechanism was gained from a simple electrostatic calculation showing that the translation of the protonated Schiff-base by 2.5 Å relative to the counter ion result in a change in free energy of ~33 kcal/mol (in a low dielectric environment, $\varepsilon = 2$) (118).

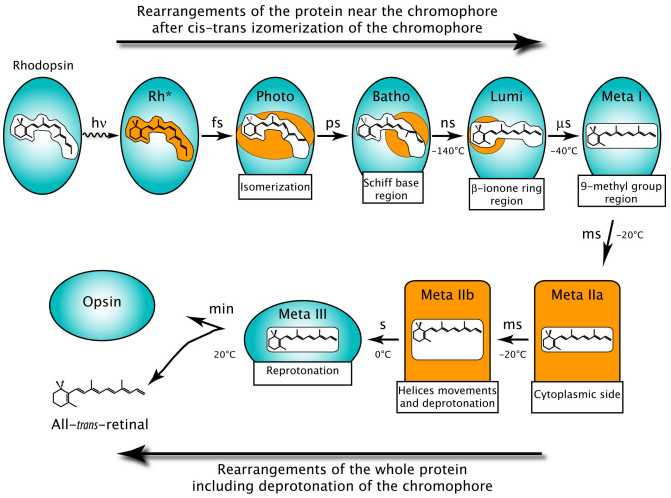
Experimental data gained by ^{13}C solid state NMR has shown that there was no substantial difference between the electronic charge distribution in the chromophores of rhodopsin, isorhodopsin and bathorhodopsin, excluding charge separation between the chromophore and the counter ion as a significant mechanism for energy storage (248), and argues in favour of the energy being stored in

the structural distortions of the chromophore. These conclusions are also supported by Raman and FTIR experiments (6, 60, 206, 208).

The finding was rationalized in a quantum simulation approach showing that the isomerization leads to a highly strained all-*trans* form in bathorhodopsin where the energy is stored in torsional distortion of the conjugated backbone that are compatible with the very short time of formation of the primary photoproduct. Furthermore, the role of the counter ion is important for the energy storage mechanism, by locking the positive charge borne by the protonated Schiff base in the terminal region of the molecule and accounting for the stiffness of the backbone (13).

Photo-cycle and spectroscopic intermediates

Several of these intermediates in the visual transduction process can be characterized by spectroscopic methods including flash photolysis and various low-temperature spectroscopic measurements (15, 61, 70, 77, 125, 134, 135, 137, 149, 150, 154, 167, 176, 177, 207, 209, 242, 243, 271, 278-280, 301) (Figure 5A). This relates back to early work on the relation between the energetic of the activating light and the coupling to protein and retinal conformations (44, 118) (reviewed in (197, 241)) (Figure 5B). The intermediate states that have been isolated are summarized in Figure 5. Low-temperature and time-resolved UV/VIS spectroscopic measurements have shown that photo-bleaching of rhodopsin involves several intermediates and most likely has more than one pathway (142, 153, 210).



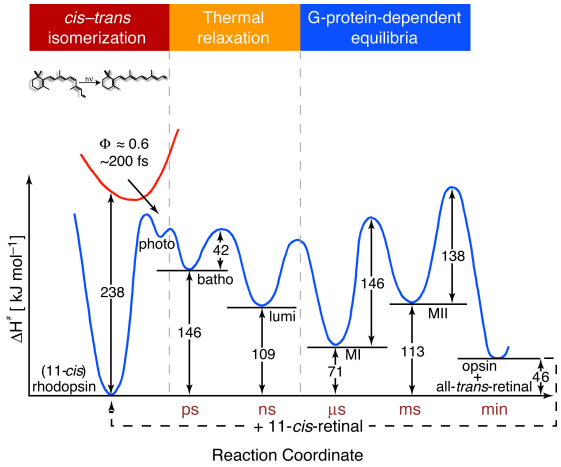


Figure 5: (A) Schematic illustration of the photobleaching process of rhodopsin (adapted from (242)). (B) Reaction coordinate of rhodopsin activation after absorption of a photon and electronic excitation (figure adapted from (241), data from (44-46, 128)).

CONFORMATIONAL STUDY OF RETINAL IN THE DARK-ADAPTED STATE

The study of the conformation of retinal in the binding-pocket of rhodopsin has been an intense subject of study in the past twenty-five years. Several analytical techniques have been applied, including x-ray crystallography, IR, Raman, UV-VIS (see previous sections) and solid state NMR spectroscopy.

In this section we review the contributions that have lead to the determination of molecular parameters – e.g. torsion angles, inter-atomic distances and orientation restraints for retinal in its binding sites, with a special emphasis on the contribution of solid state NMR. Table 2 summarizes the list of all the NMR reports where retinal was studied bound to rhodopsin. Information brought by circular dichroism (CD) is also summarized.

The x-ray structure of rhodopsin and its further refinements (198, 205, 269) have led to a rather accurate picture of the retinal in its dark-adapted state. Concerning the β-ionone ring, the diffraction data at 2.8 Å resolution is sufficient to identify and orient the ring, although there is a possibility of partial occupancy of rings rotated by 180° about C6-C7 bond (205, 269). However, until recently, uncertainties regarding the ground-state structure of the chromophore remain, due to the rather low

atomic resolution available (2.6–2.8 Å). Furthermore, retinal explores various short-lived photo-activated states, which will be even more difficult to be resolved by x-ray crystallography.

Various contributions have tried to constrain the conformation and retinal in its binding-pocket with a special attention on the following features:

a. The conformation of the chromophore chain;
b. The absolute twist of various bonds within the bound-chromophore, with a special emphasis on the C6–C7, C10–C11, C12–C13 bonds;
c. The conformation of the β-ionone ring relative to the chromophore;
d. The absolute configuration of the Schiff-base and its protonation state;
e. The orientation of the retinal in the binding-pocket and its electronic environment.

These details are very important for a proper understanding of the protein. To illustrate the point, previous NMR studies on bacteriorhodopsin have demonstrated that several factors contribute to red-shifting the visible absorption band to 568 nm. These include (1) a 6-s-*trans*-retinal conformation (109, 274), (2) a weak interaction between the Schiff base nitrogen and its associated counter ion (56, 107), and (3) protein perturbations near the β-ionone ring (119).

Rhodopsin is stable in its dark-adapted state. Consequently, it is the photo-state for which the more information has been gathered.

From the x-ray diffraction, the chromophore is represented by a well-defined density profile, and its conformation appear to be 6-s-*cis*, 11-*cis*, 12-s-*trans*, 15-16-*anti* (7, 200), as proposed in early solid state NMR work (187, 250). The protonation of the Schiff base is not discernible on the x-ray data. Another features not visible on the x-ray are the conformational out-of-plane distortions of a 12-s-*trans*-retinal chromophore due to non-bonded steric intra-ligand and ligand-protein interactions (79, 276).

Circular dichroism

The conformational of retinal by circular dichroism has focused mainly on the study of the polyene chain and the twisting of the double bonds in the dark-adapted state.

There are indications that the chromophore chain is strongly twisted in the central region, around C10 to C13 (79, 173, 180, 276). These twists may be responsible for the initial ultra-fast isomerization reaction since it induces a significant slope with respect to C11–C12 torsion in the Frank-Condon region of the excited-state potential energy surface (89, 283, 286). The resolution of the diffraction data is not sufficient for determining out-of-plane torsional angles and is not accurate enough to base high-quality excited-state calculations. An important aspect for the structure of retinal in the dark-adapted state, related with very rapid photoisomerization, is whether the chromophore is initially deformed to a state of higher energy state, thus potentially adding to the rate of the isomerization process.

The absolute configuration of the conformer in the dark-state around the C12–C13 bond is relevant because it is expected to influence the sense of rotation during the isomerization. According circular dichroism, there is even a preferred sense of twist for the chromophore (86, 130). Different experimental and theoretical studies provide controversial evidence about the absolute conformation of the C12–C13 bond, reporting either positive (33, 34, 140, 227) or a negative helicity (106, 175, 268). The most recent CD report favours a positive helicity (33) as confirmed in a recent 2.2 Å crystallographic structure (200) and in theoretical studies (263).

Solid state NMR ^{13}C & ^{15}N chemical-shifts studies

The first conformational studies of rhodopsin-bound retinal by MAS solid state NMR were published in 1987 (187, 250), either for the protein studied as a lyophilized powder (187), in detergent (250) or reconstituted in phospholipids (187). These studies, like many others to follow, used a selectively labelled chromophore, in order to avoid spectral overlap and to easy the assignment of the observed signals.

It was concluded from the comparison of the ^{13}C5 chemical shift of the bound-retinal with several retinal model compounds, that the C6-C7 bond in retinal bound to rhodopsin adopts a *cis* conformation (so-called 6-s-*cis* conformation). Based on the retinal model compounds, it was shown that the chemical shift of C5 depends strongly on the torsion angle between the β-ionone ring and the conjugated chain (108).

This was found in agreement with a model explaining the light-absorption properties of retinal (139) and later confirmed in several other studies, amongst them some of the more recent (49, 198, 205, 251, 255, 269).

The chemical shift of ^{13}C5 was investigated more in details, looking at the different components of the chemical shift tensor (251), concluding that the elements of the tensor were closer to a 6-s-*cis* than a 6-s-*trans* conformation of the C6-C7 bond.

The conformation of the C6-C7 bond is extremely important, as it influences the visible absorption band of retinal. Previous calculations indicate that the absorbance of the bound-retinal is sensitive to C6-C7 twist (16, 119). Indeed, the torsion angle around C6-C7 can either extend the π-electron conjugation to the C5=C6 double bond in planar conformations or by breaking the conjugated π-system in twisted conformations (251). From the crystal structure of retinal model compounds, it has been shown that a strictly planar 6-s-*cis* in retinal is prevented by steric interaction between the C18-methyl group and the hydrogen atom on C8 (120).

The chemical shift of ^{13}C14 was in agreement with a protonated Schiff base model in with the C=N bond in an anti-configuration (250). The chemical shift of ^{13}C12 presented the evidence of a negative charge in its proximity (187). In a third ^{13}C-MAS NMR study where ^{13}C-singly-labelled retinal molecules were investigated, this particular shift was confirmed. Even further, it was shown that the chemical shifts observed for carbon atoms 8 to 13 were shifted with the largest deshielding (6.2 ppm) localized at carbon 13 (251), consistent with a protein perturbation in the vicinity of C13

Isorhodopsin (9-*cis*-retinal), a by-product resulting a retinal that failed to activate the receptor, was also inspected using the same methodology (the primary photo-product can follow several paths: back to 11-*cis*, to lumirhodopsin, or to 9-*cis* rhodopsin). Isorhodopsin exhibit large perturbations of the ^{13}C chemical shift for C7 and C13, consistent with a protein perturbation near C13 (in comparison with polycrystalline 9-*cis*-retinal). Furthermore, the study shows similar ^{13}C chemical shift with 11-*cis*-retinal rhodopsin, with the exception of carbon C7 to C12 that are not analogous. This suggests that the interactions between isorhodopsin and the binding-pocket are different for this part of the molecule (251).

The ^{13}C shifts of 11-*cis*-[8–20-$^{13}C_{10}$]-retinal prepared by total synthesis and incorporated into opsin to form rhodopsin in the natural lipid membrane environment (277), were assigned by MAS-NMR dipolar correlation spectroscopy in a single experiment. The comparison with data of singly labelled retinylidene ligands in detergent-solubilized rhodopsin shows an agreement within an error margin of 1 ppm. A deviation of 1.5 ppm on C13 and the systematic deviation of the other carbons indicate a slightly less positive charge polarization in the natural lipid environment compared to rhodopsin in detergent (277).

The use of multispin labelling in combination with 2D correlation spectroscopy improves the relative accuracy of the shift measurements. It further allows looking at the electronic structure of the ligand. The protonation of the Schiff base involves the delocalization of the positive charge into the polyene, it results in the increase of the charge density alternating between the odd- and even-numbered atoms at the Schiff base end of the chromophore: positive charge at odd-numbered lattice positions induce correlated negative charge polarization on the even-numbered positions via Coulomb interaction (277).

The chemical shift were used to analyze the electronic structure of the retinylidene ligand at three different levels: (i) by specifying interactions between the ^{13}C-labeled ligand and the G-protein-coupled receptor target, (ii) by making a charge assessment of the protonation of the Schiff base in rhodopsin, and (iii) by evaluating the total charge on the carbons of the retinylidene chromophore (277).

A conjugation defect is the predominant ground state. This property governs the molecular electronics of retinal in rhodopsin. The cumulative chemical shifts at the odd-numbered carbons ($\Delta\sigma_{odd}$) of 11-*cis*-protonated Schiff base models relative to the unprotonated Schiff base can be used to measure the extent of delocalization of positive charge into the polyene (277). The total amount of positive charge delocalized in the polyene can be divided into three contributions: (1) polarization by the electronegative Schiff base nitrogen, (2) protonation of the Schiff base nitrogen and (3) mutual polarization effects between the counter ion and the protonated Schiff base nitrogen (277).

The complete ^{1}H and ^{13}C assignments of the 11-*cis*-retinylidene chromophore in its ligand-binding site determined with MAS-NMR (49) was obtained in 2002 and indicate the presence of non-bonded interactions between the protons of the methyl groups of the retinylidene β-ionone ring and the protein. These non-bonded interactions are attributed to nearby aromatic acid residues Phe208, Phe212, and Trp265 that are in close contact with, respectively, H16/H17 and H18. Furthermore, this study confirmed that C16 and C17, although chemically equivalent, are not spectrocopically equivalent (255). This can be accounted for solely on the basis of intra-molecular steric effects, as the same in equivalence is observed in the crystalline analogue of retinal. The shift difference between the two methyl groups is compatible with an axial and equatorial methyl (255). Furthermore, the binding of the chromophore involves a chiral selection of the ring conformation, resulting in equatorial and axial positions for methyl C16 and C17. Altogether, these data also suggest a 6-s-*cis* orientation for the ring (49, 255). While it was suggested that C19 interact strongly with the protein, NMR shows the absence of conformational shift for C19 (255).

The linewidth for C10 and C11 is ~70 Hz, close to 45 Hz for pure polycrystalline retinal, showing that for this part of the retinal the heterogeneity in the protein is comparable to the one in the crystal and considered as ordered on the time-scale of NMR (55, 276). The single rotation resonance patterns also provide strong evidence that there is a unique conformation for the retinal polyene-chain in the protein (55, 276). The hypothesis of the protein with two possible conformations for retinal, where only one is active is excluded (55, 276). This was further confirmed in 2002 (49), as the average linewidth observed for the signals from the uniformly-labelled ^{13}C-retinal bound to rhodopsin, suggests that the chromophore is trapped in the tightly packed binding-pocket in a single conformation. More specifically, the interpretation of the chemical shift leads to conclusion that the conformation of the β-ionone ring is unique and not a mix of the two conformations 6-s-*cis* and 6-s-*trans* (49). The first clue for a tightly packed retinal came in 1963, when it was showed that all-*trans*-retinal could not be accommodated in the binding pocked of opsin (302).

As mentioned earlier, it is possible to express rhodopsin in heterologous system at the level of mg / litre, opening up the possibility to introduce isotopic labels for specific amino acid types. This was

achieved for two different expression systems, opening up the possibility to carry ^{15}N NMR on the Schiff base itself (50, 65) (reviewed in (66)). This might offer the possibility to study the interactions between retinal and its binding-pocket and starts to be applied (213) (see below).

The introduction of double quantum filtering techniques has been mentioned as a possibility to select only the most closely packed residues (165, 223), where the strong dipolar couplings between the labels on the ligand and the protein are used to "filter out" signals from amino acid that are not in van der Waals contact with the ligand (66).

The first examples of heterologous expression of rhodopsin contain ^{15}N-labelled L-lysine residues, allowing studying the Schiff base using ^{15}N NMR. There are 11 L-lysine residues in rhodopsin. However, Lys296, which is involved in the Schiff base, has a different characteristic chemical shift (50, 65).

The ^{15}N resonance corresponding to the protonated retinylidene Schiff base nitrogen was observed at 156.8 ppm in the MAS spectrum of in [ε-$^{15}N_2$]-L-lysine-rhodopsin (65) and 155.9 ppm for ^{15}N in [α, ε-$^{15}N_2$]-L-lysine-rhodopsin (50). This suggests that the Schiff base in rhodopsin is protonated, in agreement with numerous previous reports.

The comparison of the ^{15}N chemical shift of the lysine-bound retinal in rhodopsin with protonated Schiff base model compounds (for which the distance between the protonated Schiff base and the counter ion differs) allows the estimation of the distance between the Schiff base and its counter ion (65). The chemical shift corresponds to an effective Schiff base-counter ion distance of greater than 4 Å (65).

This has suggested the presence of a "complex counter ion", consistent with structural water in the binding-site hydrogen-bonded with the Schiff base nitrogen and the Glu113 counter ion (50, 65). The complex counter ion was proposed by (57, 193) and would explain IR spectroscopy results, where the C=N frequency is shifted, also suggesting a molecule of water to be tightly associated with the counter ion (59, 105, 189, 257). X-ray crystallography confirmed the presence of a molecule of water tightly associated with the counter ion and close to the Schiff base, but not located between the Schiff base and Glu113 (198).

Table 2 : Summary of NMR-derived studies focused on the conformation of retinal in bovine rhodopsin (SR = Spinning Rate [kHz]).

Ref.	Method	Chromophore	Sample
(250)	^{13}C-CP-MAS-NMR SR: 2.1–2.5 kHz $\nu(^{13}C)$ = 75.0 MHz	11-*cis*-[5, 14-$^{13}C_2$] retinal dark-adapted state	Rhodopsin in detergent (Ammonyx-LO) (T = 238.15–258.15 K)
(187)	^{13}C-CP-MAS-NMR SR: 2.5 kHz $\nu(^{13}C)$ = 79.9 MHz	11-*cis*-[5, 12-$^{13}C_2$] retinal dark-adapted state	Rhodopsin as a lyophilized powder
(187)	^{13}C-CP-MAS-NMR SR: 4.5 kHz $\nu(^{13}C)$ = 79.9 MHz	11-*cis*-[5, 12-$^{13}C_2$] retinal dark-adapted state	Rhodopsin reconstituted in diphytanoyl glycerophosphocholine
(251)	^{13}C-CP-MAS-NMR SR: 3.0–4.0 kHz $\nu(^{13}C)$ = 79.0 MHz	11-*cis*-[5,6,7,8,910,11,12,13,14,15,19,20-$^{13}C_1$] retinal dark-adapted state, isorhodopsin	Rhodopsin in detergent (Ammonyx-LO) (T = 238.15–258.15 K)
(248)	^{13}C-CP-MAS-ss-NMR SR: 2.5 kHz $\nu(^{13}C)$ = 50.3 MHz	11-*cis*-[5,6,7,8,910,11,12,13,14,15,19,20-$^{13}C_1$] retinal bathorhodopsin	Rhodopsin in detergent (Ammonyx-LO) (T < 130 K)
(249)	^{13}C-CP-MAS-ss-NMR SR: ??? $\nu(^{13}C)$ = 90.4 MHz	11-*cis*-[13, 15-$^{13}C_2$] retinal metarhodopsin II	Rhodopsin reconstituted into DPPC bilayers, 1:100 molar ratio (T = 223.15 K). Illumination at –5 – -15°C for ~30s
(79)	MAS-ss-NMR: 2Q-HLF SR: 4.4 kHz $\nu(^{1}H)$ = 200 MHz	11-*cis*-[10, 11-$^{13}C_2$] retinal dark-adapted state	Rhodopsin reconstituted into retina native membrane and then washed β-cyclodextrine to remove excess retinal.
(96)	^{2}H-static-ss-NMR Oriented sample $\nu(^{2}H)$ = 61 MHz	11-*cis*-[19, 19, 19-$^{2}H_3$] retinal 11-*cis*-[20, 20, 20-$^{2}H_3$] retinal dark-adapted state	Rhodopsin reconstituted into DMPC bilayers, 1:50 molar ratio (T = 277.15 K, h_r = 52%), ^{2}H depleted H_2O. Oriented sample on glass plates.
(94)	^{2}H-static-ss-NMR Oriented sample $\nu(^{2}H)$ = 61.4 MHz	11-*cis*-[20, 20, 20-$^{2}H_3$] retinal dark-adapted state, metarhodopsin I	Rhodopsin reconstituted into DMPC bilayers, 1:50 molar ratio (T ???), ^{2}H depleted H_2O. Oriented sample on glass plates.
(276)	^{13}C-CP-MAS-ss-NMR SR: ~3Hz from RR resonance conditions $\nu(^{13}C)$ = 100.6 MHz	11-*cis*-[10, 20-$^{13}C_2$] retinal 11-*cis*-[11, 20-$^{13}C_2$] retinal other retinal analogues dark-adapted state, metarhodopsin I	Rhodopsin reconstituted into retina native membrane and then washed β-cyclodextrine to remove excess retinal.
(65)	^{13}C,^{15}N-CP-MAS-ss-NMR SR: 4-5.8 kHz $\nu(^{15}N)$ = 36.4 MHz	11-*cis*-retinal	Isotopically labelled rhodopsin (6-^{15}N-lysine and 2-^{13}C-glycine) reconstituted in DOPC (T = 213.15 K).
(50)	^{15}N-CP-MAS-NMR SR: ??? $^{15}\nu$ = 40.561 MHz	11-*cis*-retinal	Isotopically labelled rhodopsin ([α, ε-$^{15}N_2$]-L-lysine) reconstituted into native bovine retina lipids (T = 210 K)
(78)	MAS-ss-NMR: 2Q-HLF SR: 4.4 kHz $\nu(^{1}H)$ = 200 MHz	11-*cis*-[10, 11-$^{13}C_2$] retinal dark-adapted state	Rhodopsin reconstituted into retina native membrane and then washed β-cyclodextrine to remove excess retinal.

(95)	^{2}H-MAOSS-NMR SR : 2.860±0.003 kHz $\nu(^{2}H)$ = 61.4 MHz	11-*cis*-[18, 18, 18-$^{2}H_3$] retinal 11-*cis*-[19, 19, 19-$^{2}H_3$] retinal 11-*cis*-[20, 20, 20-$^{2}H_3$] retinal dark-adapted state, metarhodopsin I	Rhodopsin reconstituted into DMPC bilayers, 1:50 molar ratio (T 213 K), ^{2}H depleted H_2O. Oriented sample on glass plates.
(277)	^{13}C-CP-MAS-ss-NMR SR: 12 kHz $\nu(^{13}C)$ = 188 MHz	11-*cis*-[8,9,10,11,12,13,14,15,19,20-$^{13}C_{10}$] retinal dark-adapted state	Rhodopsin reconstituted into retina native membrane and then washed β-cyclodextrine to remove excess retinal.
(49)	^{1}H, ^{13}C-CP-MAS-NMR SR : 12 kHz ^{13}C = 188 MHz Radio frequency-driven dipolar recoupling corr. 2D heteronuclear corr.	u-^{13}C-retinal dark-adapted state	Rhodopsin reconstituted into retina native membrane.
(255)	^{13}C-CP-MAS-ss-NMR $\nu(^{13}C)$ = 125.8 MHz	11-*cis*-[8, 18-$^{13}C_2$] retinal 11-*cis*-[8, 16/17-$^{13}C_2$] retinal	Rhodopsin regenerated in native bovine retina, washed with β-cyclodextrin and reconstituted in native bovine retina membrane (T= 213.15 K).
(254)	^{13}C-CP-MAS-ss-NMR SR: 10.0 kHz $\nu(^{13}C)$ = 125.8 MHz	11-*cis*-[8, 18-$^{13}C_2$] retinal 11-*cis*-[8, 16/17-$^{13}C_2$] retinal	Rhodopsin regenerated in native bovine retina, washed with β-cyclodextrin and reconstituted in native bovine retina membrane (T= 213.15 K).
(213)	^{13}C-CP-MAS-ss-NMR DARR SR: 10.0 kHz $\nu(^{1}H)$ = 600 MHz	11-*cis*-[12,14,15,19,20-$^{13}C_5$] retinal metarhodopsin II	Isotopically labelled rhodopsin (^{13}C-selectively-labelled Tyr, Gly, Ser, Thr) in detergent (T = 193.15 K).
(39)	^{13}C-CP-MAS-ss-NMR DQ filtered SR: 5.5 kHz $\nu(^{1}H)$ = 400 MHz	11-*cis*-[10, 11-$^{13}C_2$] retinal 11-*cis*-[11, 12-$^{13}C_2$] retinal 11-*cis*-[12, 13-$^{13}C_2$] retinal 11-*cis*-[13, 14-$^{13}C_2$] retinal 11-*cis*-[14, 15-$^{13}C_2$] retinal dark-adapted state	Rhodopsin reconstituted into retina native membrane and then washed β-cyclodextrine to remove excess retinal (T = 173 K).
(234)	^{2}H-static-ss-NMR Oriented sample $\nu(^{2}H)$ = 76.75 MHz	11-*cis*-[18, 18, 18-$^{2}H_3$] retinal 11-*cis*-[19, 19, 19-$^{2}H_3$] retinal 11-*cis*-[20, 20, 20-$^{2}H_3$] retinal dark-adapted state	Regenerated rhodopsin reconstituted into POPC bilayer (1:50). Purified by hydroxyapatite chromatography before reconstitution.

Structural restraints by solid state NMR

These studies are characterized by the use of more complex NMR experiments, either allowing the estimation of molecular parameters using spectral information from the labels, like inter-nuclear distances, angles and even bond-lengths of the chromophore.

It has been shown that the selective reintroduction of dipole interactions in MAS experiments allows the measurements of internuclear distances with resolution in the order of 0.2 Å (215), reviewed in (247, 252). The accuracy of rotational resonance can be improved by the inclusion of double-quantum filter in the experiment (143). Rotational resonance has been applied on various ^{13}C selectively labelled retinal to determine distances between labels (254, 255, 276). These distances can provide rather long-distance restraints allowing constraining accurately the conformation of the retinal or at least significant segments of it.

Verdegem et al. (1999), in the first distance determinations in a ligand at the active site of a membrane protein, measured the C10–C20 and C11–C20 distances (276) (Table 4). From these distances and under the assumption of a similar geometry as that determined for the 11-*cis*-retinal in its crystal structure (92), it was confirmed that the retinylidene is 11-*cis*-12-s-*trans* (121, 282) and the experiment also validates the induced-fit theory for ligand-protein interaction. It was also shown that the C10–C13 unit is conformationally twisted (276).

Table 4: Distance restraints within retinal for rhodopsin in the dark-adapted state determined by ^{13}C–^{13}C rotational resonance (sample temperature ~ 210 K).

Atom Pair	C10–C20	C11–C20	C8–C16	C8–C17	C8–C18
Distance [Å]	3.04 ± 0.15	2.93 ± 0.15	4.05 ± 0.25	4.05 ± 0.25	2.95 ± 0.25
Ref.	(276)	(276)	(255)	(255)	(255)

The spatial structure of the C10-C11=C12-C13-C20 motif for a 10-methyl retinal analogue were also examined by rotational resonance (276). This retinal analogue was shown to display an increased non-planarity in the chromophore ($r_{10, 20}$ = 3.47 ± 0.15 Å and $r_{(10\text{-}CH3), 13}$ = 3.14 ± 0.15 Å), while being shown to be less efficient than retinal to isomerizes in rhodopsin, invalidating the theory where the more efficient isomerization rate of retinal in its binding-pocket compare to retinal in solution is due only to the conformational twist of retinal induced by the rhodopsin binding-pocket (58). Therefore, the effect of methyl substituents in the polyene region is more complex than

the simple mechanism of increasing the slope of the excited-state surface in the Franck-Condon region (276).

Rotational resonance solid state NMR was also used to provide distance restraints (Table 4) regarding the relative orientation between the β-ionone ring and the polyene chain of the chromophore in rhodopsin (255) (a key parameters for the visible absorption band of retinal, see before). In this paper, two components corresponding to two different conformations were observed. A minor component of the spectrum (approximately 26% of the chromophore), which was later shown to be due to an impurity (49), and a major spectral component (approximately 74%) has the chemical shift resolution required for measuring inter-nuclear distances to ^{13}C in the retinal chain (C8) separately from each of these methyl groups. The distances measured between C8–C16, C8–C17 and C8–C18, show that the major portion of retinylidene in rhodopsin has a twisted 6-s-*cis* conformation (255).

From the measured distance, molecular modelling based on the crystal structure of polycrystalline retinal, suggests that the distance between C8 and C18 (2.95 Å, also considered by the authors as the more precise distance measured) that the chain is twisted out-of-plane with respect to the ring by a modest amount (C5-C6-C7-C8 torsion angle = –28 ± 7°) and C8 is equidistant between the ring methyl groups at C1 with C17 turned into *axial* orientation, in agreement chemical shift data.

Solid state NMR experiment called double-quantum heteronuclear local field spectroscopy (2Q-HLF), allowing a direct determination of molecular torsional angles without estimating inter-nuclear distances, were also applied to retinal in rhodopsin. The experiment needs to be calibrated on a sample where the torsion angle is known, a polycrystalline sample of retinal (79). Using 2Q-HLF, the determination of the H–C10–C11–H torsional angle of retinal bound to rhodopsin was possible and an angle of 160 ± 10° was determined (79). Another angle was reported for the H–C14–C15–H angle using the same method (55), whose value is 165 ± 5°. These two angles indicate a significant deviation from the planar s-*trans* conformation.

In very recent and novel approach, the carbon-carbon bond-lengths of an important section of the retinal polyene chain have been determined for non-crystalline sample of rhodopsin in its dark-state

(39). Using Double quantum solid state NMR (with a recoupling sequence $R14_2^6$) (38), the measurement give an accuracy of less than 3 picometer, highlighting localized perturbation on the length of the retinal C-C bonds (Table 5). The bond-lengths were corrected to take into consideration vibrational effect to facilitate the comparison the x-ray crystal structure. These experiments confirms a significant drift from the x-ray crystal structure (~1-5% of the bond-length on average), which with a resolution of 2.6–2.8 Å (198, 205, 269) is not able to resolve single C–C bond (Table 9). Interestingly, it also confirms a previous theoretical study in which DFT was used to simulate Retinal in the binding-pocket (see below) (263). As shown in Table 5, all the bond-lengths of the DFT model published in 2002 are within the error margin of the DQ bond-length measurements, although there are some discrepancies on the bond-length alternation.

Table 5: Bond Distance within retinal for rhodopsin in the dark-adapted state determined by ^{13}C-^{13}C DQ filtered (sample temperature ~ 173 K) (39) compared with x-ray crystal structure (198) and quantum DFT simulations of retinal for rhodopsin in the dark-adapted (263).

Atom Pair	C10–C11	C11–C12	C12–C13	C13–C14	C14–C15	Ref.
Distance [Å]	1.469±0.021	1.393±0.020	1.440±0.023	1.398±0.019	1.458±0.022	NMR, (39), *
	1.439±0.025	1.363±0.024	1.410±0.027	1.368±0.024	1.428±0.026	NMR, (39), **
	1.486	1.371	1.480	1.355	1.501	X-ray, (198), #
	1.426	*1.378*	*1.436*	*1.391*	*1.420*	DFT, (263)

(*) before vibrational correction, (**) after vibrational correction, (#) average on chain A and B of the x-ray crystal structure.

The perturbation in the bond-length alternation in retinal is consistent with a model where the positive charges from the protonated Schiff base is partially delocalized on the polyene chain, with a partial concentration at the C13 position. The effect is rather strong, since the bond-length alternation is almost suppressed in the polyene chain around C11-C12. This coincides with the proximity of a molecule of water located between Glu181 and Ser186. It is therefore suggested that these polar residues associated with the water molecule assist the rapid isomerization by perturbating the conjugation in the C11-13 region and stabilizing a positive partial charge near the C11-C12 bond (39).

This also brings some more information regarding the chemical shift of C13 attributed to the simultaneous presence of Glu181, Ser186 and a molecule of water (39). Another residue might also influence the particular shift observed at C13, since Tyr268 is located within 4.5 Å and might also stabilize the delocalization of charged induce by the conjugaison (277).

^{2}H solid state NMR has also been applied to oriented rhodopsin samples containing retinal with a CD_3 group on one the C18, C19 or C20 methyl groups. From these samples, the orientations of the C8–C18, C9–C19 and C13–C20 bond vectors relative to the membrane normal have been determined (94, 95, 98). Two variants of ^{2}H NMR have been tested: static samples (94, 98, 234) and MAOSS (95).

The analysis of the ^{2}H NMR spectra provided angles for the individual labelled chemical bond vectors providing orientation constraints relevant for the three-dimensional structure of the polyene chain of the chromophore in the binding-pocket. The structural interpretation of the ^{2}H spectra relies on the computation of the mosaic spread with the simulation of the labels at various orientations in order to restrict the orientation of the vector bond in the rim of a double cone, that has a symmetrical distribution of the membrane normal. The measurement consists of an average orientation of the retinal in its binding-pocket combined with the average orientation of the protein, with the spread of the angles measured to be 10°.

Rhodopsin was studied by static ^{2}H NMR reconstituted in both DMPC and POPC (Table 3) (95, 98, 234). The samples in DMPC were studied by ^{2}H-MAOSS-NMR, providing an improved definition of the deuterium signals compare to static NMR on the same sample (95) (Table 3). The lipid composition of the two sample differ from the native membrane, especially for the sample reconstituted in DMPC, known to Rhodopsin was to prevent the full activation of the protein, blocking it in metarhodopsin I state (185). A consequence is that the hydrophobic thickness of the bilayer might differ from native membrane and may influence the anchoring of the protein in the bilayer.

The orientation provided by the three labelled methyl groups, do not provide by itself enough restraints retinal, unless assumptions of an ideal geometry for subparts of retinal are made. Gröbner

et al. (2000) (95) also suggested a model for the retinal in the binding-pocket, using from combination of MAOSS and static ^{2}H NMR, plus previously determined restraints determine by (78, 276). The modelling strategy appears to be flawed due to a lack of information. The most striking example is the 6-s-*trans* orientation of the ring against most of the experimental data available at that time (^{13}C chemical shifts (49, 187, 250) and circular dichroism (136)). The orientation of the methyl group C18 is restrained and orientated along a double-cone symmetrical with respect to the bilayer plane, where the angle with the normal to the bilayer could be either 21 ± 5° (6-s-*trans*) or 159 ± 5° (6-s-*cis*). Retrospectively, the main two arguments for the 6-*s-trans* orientation, was the similarity with bacteriorhodopsin and the way it was combined with a model derived from the low-resolution (5Å) electron-microscopy structure oriented so that helix VII is parallel to the membrane plane. Consequently, the model had to be revised when further experimental evidence that the ring was in fact 6-s-*cis* came out (33, 49, 198, 205, 255, 269).

Table 3: Orientation restraint relative to the membrane normal for retinal methyl group in the dark-adapted state by static ^{2}H solid state NMR (sample temperature ~ 277 K) and ^{2}H MAOSS (sample temperature ~ 140 K).

Bond	C5–C18	C9–C19	C13–C20	Ref.
Orientation relative to Z [°]	N.A.	42 ± 5 or 138 ± 5	30 ± 5 or 150 ± 5	static ^{2}H solid state NMR (94, 95, 98)
	70 ± 3 or 110 ± 3	52 ± 3 or 128 ± 3	68 ± 2 or 112 ± 5	static ^{2}H solid state NMR (234)
	21 ± 5 or 159 ± 5	44 ± 5 or 136 ± 5	30 ± 5 or 150 ± 5	^{2}H MAOSS (95)

CONFORMATIONAL STUDY OF RETINAL IN PHOTO-ACTIVATED STATES

Retinal in the bathorhodopsin state

The ^{13}C MAS-NMR study of retinal in the bathorhodopsin photo-intermediate trapped at low temperature has shown very little difference in the chemical shift for the retinal (carbons at positions 8, 10, 11, 12, 13, 14, 15) compare to the retinal in the dark-adapted state (248). Relative to the all-trans-retinal protonated Schiff base chloride, a large perturbation is observed for C13. It was already observed in the dark-adapted state and does not change position upon activation. Smaller shift for C10, C11 and C12, similar to the one observed for the dark-adapted state (248).

The interpretation of these data is that the energy stored in the primary photoproduct does not raise any substantial changes in the average electron density at the labelled positions, and is similar to dark-adapted state and isorhodopsin (248). The implications for the mechanism of energy storage in retinal have been discussed in a previous section.

In bathorhodopsin, the chemical shift for $^{13}C14$ remains in agreement with a protonated Schiff base model with an anti C=N bond (248). The average linewidth for the ^{13}C signals was 75-80 Hz in this photo-state (248).

Another conclusion of the study is that the protein-chromophore interactions that affect the charge polarization in the conjugated system in bathorhodopsin are similar to those in rhodopsin and isorhodopsin (248).

Raman spectroscopy showed that retinal in the bathorhodopsin state is s-*trans* about the C10-C11 single bond (206)

Retinal in the metarhodopsin I state

The first conformational study on metarhodopsin I state was focused on the polyene tail in the region where the isomerization takes place (276). Similarly to the dark-state distances C10–C20 and

C11–C20 were measured by ^{13}C-^{13}C rotational resonance (Table 7), showing an increase of the C10–C20 distance by more than 1.3 Å upon activation, whereas the C11–C20 remains almost unchanged.

Table 7: Distance restraints within retinal in the metarhodopsin I state determined by ^{13}C-^{13}C rotational-resonance (sample temperature ~ 210 K).

Atom Pair	C10–C20	C11–C20	C8–C17	C8–C18
Distance [Å]	≥ 4.35 ± 0.15	2.83 ± 0.15	4.05 ± 0.25	2.95 ± 0.25
Ref.	(276)	(276)	(254)	(254)

A second rotational resonance study of retinal in the metarhodopsin I state (254), focused on the observation of ^{13}C nuclei introduced into the β-ionone ring (at the C16, C17, and C18 methyl groups) and into the adjoining segment of the polyene chain (at C8) and allowed distances to be measured between several of the labels using ^{13}C rotational resonance (Table 7).

For the rotational resonance study, an indirect photo-conversion via the primary intermediate, bathorhodopin, was adopted as the preferred method since 44% conversion to the metarhodopsin I component could be achieved, with only low levels (18%) of ground-state rhodopsin remaining. The additional photoproduct, isorhodopsin, was resolved in ^{13}C spectra from C8 in the chain, at levels of 38%, and was shown using rotational resonance NMR to adopt a 6-s-*cis* conformation between the ring and the polyene chain. The C8 resonance was not shifted in the metarhodopsin I spectral component but was strongly broadened, revealing that the local conformation had become less well defined in this segment of the chain. This line broadening slowed rotational resonance exchange with the C17 and C18 ring methyl groups but was accounted to show that, despite the chain being more relaxed in metarhodopsin I, its average conformation with respect to the ring was similar to that in the ground state protein. Conformational restraints are also retained for the C16 and C17 methyl groups on photo-activation, which, together with the largely preserved conformation in the chain, suggest that the ring remains with strong contacts in its binding-pocket prior to activation of the receptor.

The only significant chemical shift change that could be detected for the ring methyl groups on photo-activation was in the C18 resonance, which increased from 22.1 to 22.5 ppm (254). The large splitting between C16 and C17 (-4.3 ppm) describes the unique orientation of these geminal methyl

groups in rhodopsin (255), and its retention on photo-activation is important for the determination of the mechanism of activation (254).

It has been predicted (166) that on relaxation from bathorhodopsin, C8 must move past the C18 methyl group thereby reversing the sense of rotation between the ring and the chain. Despite it was not possible to measure exchange to both methyl groups at C1. The information gained from rotational resonance concludes from the data for C17 alone that the chain of the chromophore in metarhodopsin I also retains the same sense of rotation around the C6-C7 bond as in the ground state protein. This is based on the observation that the asymmetric orientation of C16 and C17 about the ring plane remains unchanged in metarhodopsin I. A reversal in the sense of twist in the chain (positive torsion) would significantly lengthen the distance between C8 and C17 since all positive torsion angles place C8 at least 4.5 Å from C17.

The first ^{2}H NMR study on metarhodopsin I using solid state NMR was reported in 1998 (94), where static ^{2}H NMR was applied on 11-*cis*-[20, 20, 20-$^2H_3$]-retinal in sample oriented on glass plate. Rhodopsin reconstituted in DMPC bilayers by ISDU (Isopotential Spin Dry Ultracentrifugation) (97). The choice of the lipid relies the propention of DMPC to prevent the formation of metarhodopsin II, as it has been shown that in DMPC bilayers, the equilibrium between metarhodopsin I and II depends on the temperature (185). Below T_m, the amount of metarhodopsin II formed is almost negligible amount and strongly shifted toward metarhodopsin I relative to the native rod outer segment disk membrane (185).

In a second ^{2}H NMR study, the orientations relative to the membrane normal for two other methyl groups, C18 and C19, were determined using ^{2}H-MAOSS-NMR (95). The ^{2}H NMR spectrum of methyle C20 (CD_3) suggests an orientation relative to the membrane normal to be 60 ± 10° or 120 ± 10° (Table 6), compared to 30 ± 5° or 150 ± 5° for the dark-state (Table 3), this suggest a reorientation of the C20 methyle by either 30 ± 15° or 90 ± 15° upon activation. Recent work on metarhodopsin II (213) favours the first possibility as the most plausible with a large reorientation of the C20 upon activation (≥ 90°).

Table 6: Orientation restraint relative to the membrane normal for retinal methyl group in the metarhodopsin I state by ^{2}H MAOSS (95) (sample temperature ~ 140 K).

Bond	C5-C18	C9-C19	C13-C20
Orientation relative to Z [°]	62±7 or 118±7	65±10 or 115±10	60±10 or 120±10

The H-C10-C11-H torsional angle of the chromophore in the metarhodopsin I state has been estimated using 2Q-HLF. The torsional angles was estimated to be 180 ± 25° for metarhodopsin-I. The result is consistent with current models of the photo-induced conformational transitions in the chromophore, in which the 11-*cis*-retinal ground state is twisted, while the later photo-intermediates have a planar all-*trans* conformation (78).

Retinal in the metarhodopsin II state

Two solid state NMR studies have covered metarhodopsin II, the so-called signalling state of rhodopsin (214, 249). The earliest study aimed to study the protonation state of the retinal Schiff base linkage (249) and the second aimed defining the position of the chromophore relative to some residues belonging to the binding-pocket (214).

In the early 1990, there was a controversy concerning retinal in the metarhodopsin II state. On one side it was argued that metarhodopsin II contains an unprotonated Schiff base on the basis of the photo-reversibility of metarhodopsin II to rhodopsin (302). While on the other side, it proposed that the retinal was not bound anymore to the binding site or least similar to a carbinolamine (46, 47). Neither Raman nor FTIR could solve the problem, since it was not possible to assign corresponding vibration bands.

As the ^{13}C chemical shift of both the ^{13}C15 and ^{13}C13 of retinal in the metarhodopsin II state are characteristic of an unprotonated Schiff base, at the same time as the ^{13}C15 shift is significantly different from that of retinal or a tetrahedral carbinolamine group (a group that was previously proposed as an intermediate in the hydrolysis of the Schiff base at the metarhodopsin II state), it was concluded that the retinal remained covalently-bound to metarhodopsin II forming a deprotonated Schiff base, and that carbinolamine is not a good model for retinal in the activated state (249). Therefore the ^{13}C chemical shift provide convincing evidence that Schiff base deprotonation occurs in the metarhodopsin I – metarhodopsin II transition (249).

A comparison with the ^{13}C chemical shift of retinal in ground-state and bathorhodopsin shows a shift for C13 in agreement with a negative charge in its surrounding, while at the metarhodopsin II

state the chemical shift for this particular carbon is not significantly perturbed anymore from unprotonated Schiff base model compounds (249). This suggests a change of the conformation of the protein around C13.

Patel et al. (2004) (214) provide more information on how the retinal isomerization would disrupt helix interactions locking the receptor off in the dark state and suggest a mechanism of activation for rhodopsin. In this study, a variant of rotational resonance called 2D dipolar-assisted rotational resonance NMR (DARR) (52, 266) was applied to obtain distance restraints between between ^{13}C-labels on the retinal chromophore and specific ^{13}C-labels on tyrosine, glycine, serine, and threonine in the retinal binding-site of rhodopsin.

DARR is a 2D solid state NMR experiment that reintroduces the ^{1}H-^{13}C heteronuclear dipolar coupling between labels that have van der Waals contact (52, 266). In the case of ^{13}C nuclei, this corresponds to distances up to 5.5 Å between the nuclei, in which case a cross-peak appears in the 2D spectrum between the ^{13}C two labels (214).

The strategy followed consisted in determining ^{13}C–^{13}C retinal-protein contact in the dark-adapted state of rhodopsin, then confirming these contacts using the x-ray crystal structure, before looking for ^{13}C–^{13}C retinal-protein contacts in the metarhodopsin II intermediate state. This strategy allows to determine both the location of the retinal in the binding-pocket and the direction of the isomerization (214).

From these restraints and with the help of the crystal structure of the protein in its dark-adapted state (198), it was possible to establish a list of ligand-protein contact for the dark-adapted and (198) metarhodopsin II states. The distance restraints gained with this method are listed in Table 8.

The collected data were interpreted under the assumption that the retinal is moving rather than the protein transforming its structure during the activation. The essential aspects of the isomerization trajectory suggested from the retinal-protein contacts observed in the active metarhodopsin II intermediate (Table 8) imply a large rotation of the C20 methyl group ($\geq 90°$) toward extra-cellular loop 2 and a 4 – 5 Å translation of the retinal chromophore toward transmembrane helix V.

Furthermore, the photo-activation of rhodopsin involves a coupled motion of transmembrane helices V, VI, and VII (214).

The rotation of C20 is reasonable, as the x-ray crystal structure shows that C20 is less densely packed than C19. The fact that C19 is largely blocked is interpreted by the author as the fact C19 direct the translation of retinal toward helix V. C19 is also seen as an element preventing the rotation of the β-ionone ring. Although these NMR experiments do not provide any information regarding the orientation of the β-ionone ring, they are consistent with no motion of the ring as suggested for the metarhodopsin I state (254).

The interactions of the β-ionone ring with the protein control the activation mechanism (133). This statement is reinforced with the DARR experiments on retinal in the metarhodopsin II, where the presence of the ring is required to allow the translation motion of the retinal resulting in the increased contacts between the ring and helix V leading to the displacement of helix V, the disruption of helix-helix interactions and the activation of the receptor (214).

In the same paper, Patel et al. optimized the full conversion of rhodopsin to metarhodopsin II and trap it at low temperature as a stable and homogeneous specie. Various sample preparations were tried, including rhodopsin reconstituted in rod outer segment membrane, saturated lipids or solubilized in DM detergent micelles. DM micelles is the only environment where full metarhodopsin II conversion occur for samples with the optical density required for NMR, probably due to the more flexible environment provided by the micelles (214). However, a loss of activity of ~10% was observed upon illumination, probably due to the formation of isorhodopsin and other side-reactions.

Table 8: Summary of distance restraints gained by DARR (213) (for the dark-adapted state, the distances are measured from the 1L9H crystal structure (198)). **V** = cross-peak observed, **X** = cross-peak not observed. Between bracket, distance as measured from the crystal-structure (dark-adapted state) or distance estimation from DARR (MetaII).

DARR Contact	C19		C20	
Amino Acid (position)	Dark (Dist. in crystal)	Meta II (DARR)	Dark (Dist. in crystal)	Meta II (DARR)
4'-^{13}C-Tyr178 (loop e2)	X (6.1 Å)	X (≥ 5.5 Å)	X (9.8 Å)	V (≤ 5.5 Å)
4'-^{13}C-Tyr191 (loop e2)	V (5.1 Å)	X (≥ 5.5 Å)	X (8.4 Å)	X (≥ 5.5 Å)
4'-^{13}C-Tyr268 (helix VI)	V (4.8 Å)	X (≥ 5.5 Å)	V (4.4 Å)	X (≥ 5.5 Å)
2-^{13}C-Gly114 (helix III)	X (6.8 Å)	X (≥ 5.5 Å)	X (7.0 Å)	V (≤ 5.5 Å)
2-^{13}C-Gly121 (helix III)	X (8.0 Å)	X (≥ 5.5 Å)	X (8.0 Å)	X (≥ 5.5 Å)
2-^{13}C-Gly188 (loop e2)	V (4.8 Å)	X (≥ 5.5 Å)	X (6.3 Å)	X (≥ 5.5 Å)
			()	
1-^{13}C-Thr118 (helix III)	N.A. (4.0 Å)	N.A. (≥ 5.5 Å)	X (7.4 Å)	X (≥ 5.5 Å)

DARR Contact	C14		C15	
Amino Acid (position)	Dark (Dist. in crystal)	Meta II	Dark (Dist. in crystal)	Meta II
4'-^{13}C-Tyr178 (loop e2)	X (9.1 Å)	V (≤ 5.5 Å)	X (10.5 Å)	V (≤ 5.5 Å)
4'-^{13}C-Tyr191 (loop e2)	X (9.3 Å)	X (≥ 5.5 Å)	X (10.3 Å)	X (≥ 5.5 Å)
4'-^{13}C-Tyr268 (helix VI)	X (5.9 Å)	X (≥ 5.5 Å)	X (6.2 Å)	X (≥ 5.5 Å)
3-^{13}C-Ser186 (loop e2)	V (4.0 Å)	X (≥ 5.5 Å)	V (3.4 Å)	X (≥ 5.5 Å)

PHOTOLYSIS AND CROSS-LINKING OF RETINAL DERIVATIVES

The photolysis of retinal derivatives leading to cross-linking can be achieved for two classes of derivatives. The photolysis of a non-isomerizable 11-*cis*-locked retinal analogue results in a cross-linking with Trp265 and Leu266, showing that C-3 of the β-ionone ring is close to helix VI in the dark state (303), while the photolysis of a photo-reactive retinal analogue with an unlock 11-ene results in a cross-linking with Trp265, Ala117, Glu122, Phe115, Ser127 and Trp128, showing that C3 of the β-ionone ring is close to both helix III and VI (191).

In another photolysis experiment, an analogue very similar to retinal, the ring-modified 3-diazo-4-keto-11-Z-retinal, was used to regenerate rhodopsin. The results obtained with previous photolysis experiment with 11-cis-locked analogue were confirmed for the dark-adapted state, with the ring only found to cross-link with Trp265 (a residue of the binding-pocket) in its dark-adapted and primary photo-intermediate (bathorhodopsin) states. The cross-linking was also achieved with other thermally trapped photo-intermediates, in an attempt to track the movement the C3 atom from the β-ionone ring within the protein. However, at lumirhodopsin and the later intermediates, the cross-linking was only observed with Ala169 in helix IV, remote from the chromophore binding-pocket, suggesting different residues than suggested in a previous study (191), and leading to the conclusion that the β-ionone ring of the chromophore is ejected from the binding-pocket prior to formation of metarhodopsin I (24, 253).

A motion of the β-ionone ring towards Ala169 suggests a remarkable conformational change at the early lumirhodopsin stage – a translation of almost 13 Å according the crystal structure (205) – disrupting major contacts within the binding-pocket and then no major adjustment in the location of the attached chromophore on progression to metarhodopsin I and II. It results in a model (26) where it is difficult to implicate the chromophore in any concerted conformational or electronic transformation leading to activation after the lumirhodopsin state, other than to assume that these

events are initiated on releasing the β-ionone ring from the binding-pocket (68, 138). Another surprising feature of this study is that the modified chromophore seemed to drive the early photo-equilibrium toward bathorhodopsin, since at the lumirhodopsin and later stages no cross-linking within any remaining ground-state rhodopsin was detected. Recent molecular dynamics calculations (129) predicted that the ring modified for photo-affinity cross-linking can induce deformations in the protein once flipped from its binding-pocket, but did not consider whether modifications interfere with stabilizing interactions within the binding-pocket. Other studies show the implication of helix III and VI upon activation of GPCRs (76, 103, 104, 239), but do not account for retinal-helix contact.

The β-ionone ring of the chromophore is in contact with residues from the binding-pocket belonging to helix V as shown in Figure 6. The binding-pocket is derived from the latest crystallographic structure for rhodopsin in the ground state (198) and the chromophore in A (shown in yellow) was modelled with the NMR restraints available for the ring segment (255) and aligned with the ring location provided in the crystallographic structure. The binding-pocket is viewed perpendicular to the membrane normal (and to the face of the chromophore ring) and with the extra-cellular side of the protein above. Phe212 appears within a flexible region of helix V above Pro215 and contacts the C2-C3 edge of the ring. Ala169, the site of photo-labelling in the early photo-intermediate, lumirhodopsin, is shown behind the chromophore (and the extracellular loop 2), is situated at the top of the short helix IV and is around 15 Å from C3 on the ring. The primary restraints to the ring rotating out of the binding-pocket toward Ala196 are from Met207 and Phe208. In B, the chromophore ring is shown in the same location within the binding-pocket but has been modified with a diazo group (N_2) attached to C3 as in the photo cross-linking analogue. The diazo group cannot be accommodated in this position since it penetrates through the van de Waals space at the bottom of Phe212. This structural modification of the ring required that the C2-C3-C4 segment be made approximately coplanar with the C3-N2 bond to reflect the changes in hybridization at C3 and also C4 (keto group not shown) in the photo-affinity analogue.

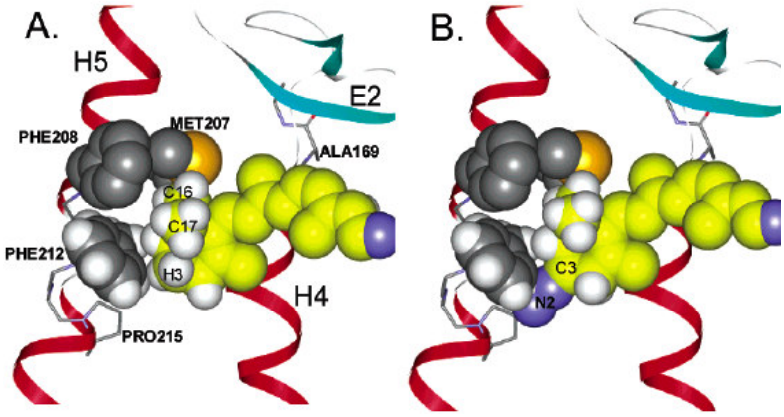


Figure 6: In A, Binding-pocket contacts for the chromophore ring with residues on helix V of rhodopsin. In B, the chromophore ring is shown in the same location within the binding-pocket but has been modified with a diazo group (N_2) attached to C3 as in the photo cross-linking analogue. The diazo group cannot be accommodated in this position since it penetrates through the van de Waals space at the bottom of Phe212 (from (254)).

MODELLING OF RETINAL IN THE BOVINE RHODOPSIN BINDING-POCKET

Interestingly, *ab initio* simulations of the retinal-Schiff base in vacuum have shown that the deprotonated state is the most stable, suggesting the binding-pocket of the protein has to stabilize the protonated specie (235, 265). General theoretical considerations suggest that a highly polar environment stabilized the ionized site in a protein matrix (288), highlighting the importance of the binding-pocket, its amino acids composition and geometry.

In situ modelling of retinal isomerization in rhodopsin and other retinal proteins has been the object of several theoretical studies (reviewed in (287)). Several modelling approaches have been used to study *ab initio* the conformation of the lysine-bound retinal in the dark-state. Many early studies used position constraints on different atoms of retinal to represents the restraining effect of the binding-pocket (14, 18, 19, 285). In some studies, the binding-pocket was modelled using a deformable environment (286), which can react in some extend to conformational relaxation of retinal after isomerization. The electrostatic effect of the binding-pocket has been taken in account

by dielectric models (285) or by including an explicit negative charge into the calculations (18, 19, 267).

Other studies have included *ab initio* molecular dynamics simulations (also known as Carr-Parrinello molecular dynamics) (13, 28, 30), MNDO method (245), MO (289, 290) and more recently DFT including the geometry of the binding-pocket (263). The more recent investigations of retinal in rhodopsin have included a full representation of the protein (263), but so far bacteriorhodopsin has been studied more extensively (115, 126). The photoisomerization of a simplified model for the retinal Schiff base has also be simulated using semi-empirical method showing a similar time-scale for the isomerization of the *cis*-bond as experimentally observed (281).

Some early semi-empirical molecular calculations (102, 105, 106), using the ^{13}C chemical shift experimentally measured as restraints have been used to retrain the position of the counter ion relative to the retinal. One of the main results of these calculations was that a single negative charged in the vicinity of C12 is enough to account for the anomalous changes in the charge density reflected by the ^{13}C chemical shift. Unfortunately, this has led to an erroneous conclusion that the counterion was located within 3 Å of carbon C12 (another model including two charges was also considered in this study, but led to poor fit of the chemical fit (102)). The main reason for this failure resides in the lack of information to model the binding-pocket, which has been proven to be more complex than considered in the model, where the binding-pocket only included a single or two point charge and a molecule of water. As it seems now, there are two glutamic acids in the binding-pocket, Glu113 and 181, which have the ability to carry a negative charge and both close enough to be located within 3 Å from the retinal polyene chain depending on the orientation of their side-chain (198, 205, 269).

Ab initio molecular dynamics has been focusing on the delocalization of the positive charge along the conjugated carbon chain of retinal and the storage of energy in the chromophore upon activation. A hybrid approach is used to induce the isomerization in a model of the chromophore of rhodopsin, by applying an external classical force field to the C10-C11-C12-C13 dihedral angle, in addition to the *ab initio* forces. One of the main conclusion is that the energy is mainly stored in torsional strain in the primary photoproduct (13), in agreement with other simulation studies (13,

286). It is found that the *trans*ition state is characterized by a well-defined bond pattern along the chromophore and by a positive charge displacement. Furthermore, the relaxation of the chromophore is associated with a coherent propagation of a conjugation defect that transports the excess positive charge along the chromophore-conjugated backbone (161). This soliton-like charge propagation is strongly coupled to the vibrational degrees of freedom due to the non-planarity of the molecule and its dampin is attributed to coupling with out-of-plane vibrational degrees due to non-planarity of the chromophore (28).

Ab initio molecular dynamics simulation of the chromophore starting in a conformation where distances between C10–C20 and C11–C20 were fixed accordingly solid state NMR (276), brings 3.08 Å and 3.05 Å for these distances after optimization of the structure (30). It also predicts significant deviations from planarity in the region C9–C14. The deformation is distributed along the backbone along the whole C9–C14 region rather than concentrated in a single specific bond (30), with a total out-of-plane deformation of approximately 30 degrees (30), in agreement with solid state NMR measurement for some of these torsion angles (79).

From solid state NMR, the C10-C11–C12-C13 bond has an estimated torsion angle of about -44 ± 10° (under the assumption 11-*cis*-12-s-cis-retinal adopt a similar conformation to its crystal structure but the torsion angle around C11-C12). The angle correlates well with the angle of 39° found for the C12–C13 bond in crystalline 11-*cis*-retinal (92). Previous Carr-Parrinello molecular dynamics simulations (13) suggest a distribution of the out-of-plane distortion over three bonds including C12–C13. These angles are thought to result from steric hindrance between hydrogen atoms on C10 and C14. The absolute sense of rotations agrees with chiroptical data obtained by (33, 34): C13 is on one side of the molecular plane and the C10 proton on the other. Previous estimations based on Raman spectroscopy (208) and the interpretation of NMR ^{13}C chemical shifts restrains lead to the same value (105).

Ab initio molecular dynamics simulations on retinal have shown that the relative position with the chromophore of a counterion or its omission has a strong influence the ^{13}C chemical shifts (29). The method has been used to compute ^{13}C chemical shift on retinal-protonated Schiff base models in a rather good agreement, highlighting the need of a detailed description of the binding-pocket for more accurate results (29). Nevertheless, the study shows qualitative features such as the reduction

in the amplitude of the bond-length alternation between single and double carbon bonds in the vicinity of the protonated Schiff base, due to the positive charge delocalization (29).

Indeed, the availability of the crystal structure and a more detailed description of the binding-pocket allowed the refinement of the *ab initio* simulation of the bound-retinal.

The reaction path for the retinal in rhodopsin has been considered using MD simulations (129, 227, 229). So far the model used was the refined model 1HZX (2.8Å, where the molecules of water were either missing or added using computational methods). Using this model, all studies suggest a hula-twist conformational change occurring during the photo-cycle, followed by the rotation of the ring. The confirmation of the reaction mechanism has been aided by consideration of designed chemical mimics of retinal that alter the photo-cycle in controlled ways.

Yamada et al. (2002) presented the application of *ab initio* MO calculations at the RHF/6-31g* level on a model system consisting of Schiff base-retinal, Glu113 (counter ion) and up to eight residues, for which the Schiff base-retinal was considered in its protonated and deprotonated states (289). Unfortunately, no molecules of water were considered. The result of these computations shows that the carboxyl group of Glu113 (counterion) has the most significant contributions to the stability of the protonated Schiff base (289). The individual contribution of residue belonging to the binding-pocket, was studied showing the addition of any residue to the protonated Schiff base in vacuum stabilizes the protonated state of the Schiff base (289). Furthermore, an additive rule roughly holds for the contributions to the stabilization energy from individual amino acid residues (289). In terms of stabilization, charged residues like Glu113 have the more important contribution, followed by polar residues like Ser186 and Cys187, which have the most important contribution amongst neutral amino acids. The MO calculations also show that the charge separation state is energetically stable for the molecular model system consisting of retinal, Glu113 and some neutral amino acids, while in the absence of any molecule of water (289).

In a recent study, Density Function Theory (DFT) and Molecular Dynamics Simulation (MD) were used to investigate several simplified binding-pocket models, including Glu113 and a combination of the following residues: a molecule of water bound to Glu113, Thr 94 and Cys187 (35, 264). It

was concluded that Thr94 was the main factor for stabilizing the protonated Schiff-base. When Glu113 was the only residue present, the chromophore deprotonates spontaneously and the proton is transferred to Glu113. The addition of Thr94 in the calculation was enough to stabilize the protonated Schiff-base and prevent the proton transfer. The further addition of more elements (Cys187 and a water molecule) would stabilize further the interacting ions, but to a weaker extend.

Density functional theory (DFT) was also used to derived *ab initio* the conformation of retinal in the dark-adapted state. The method allows the inclusion of a large fraction of the protein (the 27 next-nearest amino acids) and a molecule of water present near Glu113 (the water molecule near Glu181 is missing), Sugihara et al. (2002) could study the influence of the protein pocket on the three-dimensional structure of the 11-*cis*-retinal Schiff base (SB) chromophore (263) and address the following questions in their simulations: (i) The conformation of the β-ionone ring. The protein pocket tolerates both conformations, 6-s-*cis* and 6-s-*trans*, with a total energy difference of 3.6 kcal/mol in favour of the former (263). Of the two possible 6-s-*cis* conformations, the one with a negative twist angle (optimized value: -35°) is strongly favoured, by 3.6 kcal/mol, relative to the one in which the dihedral is positive (263) (Figure 7). (ii) Out-of-plane twist of the chromophore. The environment induces a non-planar helical deformation of the chromophore, with the distortions concentrated in the central region of the chromophore, from C10 to C13. The dihedral angle between the planes formed by the bonds from C7 to C10 and from C13 to C15 is 42° (263). (iii) The absolute configuration of the chromophore. The dihedral angle about the C12-C13 bond is +170° from planar s-*cis*, which imparts a positive felicity on the chromophore (263). The internuclear distances are in close agreement with most of the information gained from circular dichroism and NMR studies (the twist around the C11-C12 bond (33), inter-nuclear distances measured by rotational resonance (255, 276) and bond-length measurements by DQ (39)), preceding the publication of two major NMR studies of retinal bound to rhodopsin in the dark-adapted state (39). The methyl groups on the β-ionone ring are pointing towards the side-chain of aromatic residues, in agreement with the ^{13}C chemical shift (255), making this model the most credible model generated so far of the retinal in its binding-pocket.

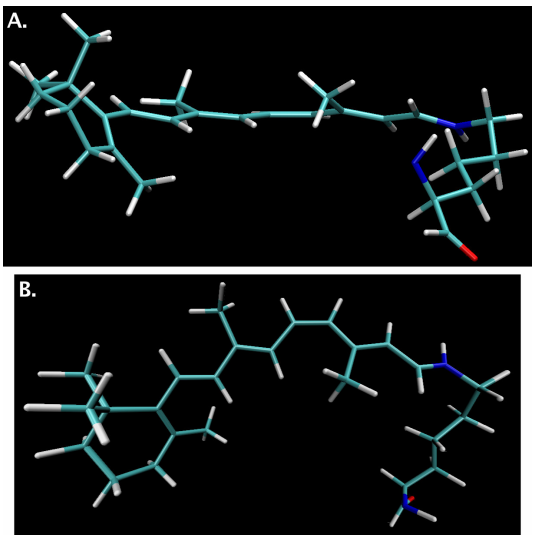


Figure 7: Deformation of the C11-C12 bond according the more recent circular dichroism study (presented here 6-s-*cis* conformer proposed by (263), which fulfills most of the experimental restraints determined by NMR and CD). **A,** view from side (orthographic projection). **B,** view from top (perspective projection).

In a recent publication (39), the accuracy of DFT simulation was questioned on the basis that the technique could not reproduce the bond-length alternation observed for crystallized protonated Schiff base models. However, the conclusion relies on an unhomogeneous comparison between experimental data for solid state crystallized retinal protonated Schiff base and DFT simulation of an isolated retinal protonated Schiff base in vacuum (closer to a gas- than solid state). Therefore this comparison is flawed.

The general conclusion from these *ab initio* studies is double. On one hand, quantum methods, like DFT, are mature enough to provide high-resolution information, provided that a relatively well-defined geometry for the binding-pocket (2.8 Å resolution) and a reasonable starting point for the orientation of the ligand are available. On the other hand, this restrains drastically the number of other potential candidates due to the current shortage of structural information regarding membrane proteins in general and GPCRs in particular. Overcoming this difficulty is undoubtedly a major challenge to face for *ab initio* simulation methods.

Quantum methods, and DFT in particular have shown it was possible to predict the conformation of retinal in rhodopsin provided the binding-pocket was accurately defined. Unfortunately, the method might not be readily applicable to further photo-activated states (especially the late photo-

intermediate where conformational changes have already taken place in rhodopsin), because of the requirement of an accurately defined binding-pocket.

MD SIMULATIONS OF THE PHOTOACTIVATION OF RETINAL IN BOVINE RHODOPSIN

The x-ray crystal structure of rhodopsin is a major step toward understanding the mechanistic details of the visual process and the experimentally determined geometry of the protein constitutes a starting point for the structures of further intermediates and pathways to be investigated and tested using theoretical and computational methods.

One of the challenge introduced by rhodopsin for *in silico* study is the large separation in time-scales between the steps in photo-cycle of the protein with respect to the current limits in computer simulations of large proteins. While it takes a few ms for the protein to reach the activated-state and much more before the retinal to be regenerated (152), the state-of-the-art computer simulation of large protein with explicit solvent and membrane description are still limited to a 10-100 ns (53). Therefore MD is not able to follow the activation of the receptor from the dark-adapted state to metarhodopsin II yet.

So far, all the MD simulations reported used models built-up from the refined 2.8 Å x-ray crystal structure 1HZX, for which no information regarding the position of water molecules inside rhodopsin is available. In order to take this critical parameter for the protein structure and dynamics into account, several molecules of water were added to some of the cavities inside the protein using electrostatic-based predictions. With the availability of an x-ray crystal structure with an improved resolution to 2.6 Å (1L9H) showing the exact location of water molecules, the current models used by MD simulations might be outdated. As no report presenting MD simulations with a model for rhodopsin built from 1L9H has been presented yet, the results yield by current MD simulations should be interpreted carefully, until further simulations using 1L9H as a starting point for the model building confirms that the slight differences in the positioning of the molecules of water do

not affect the overall picture. Furthermore, an additional difficulty for MD simulation relies in the requirement of a good parametrization for the retinal for which accurate data have not been made available until 2002 (263), but their use has not been reported yet.

Two extensive molecular dynamics calculations accounting for the modelled POPC/water environment (229) or octane/water (227) in an explicit description and starting from rhodopsin X-ray structure 1HZX, have explored the first nanoseconds after the isomerization of the retinal, trying to connect the *cis-trans* isomerization of the retinal with the protein changes (227, 229). The 11-*cis* to all-*trans* isomerization, induced by transiently switching the potential energy function governing the C11–C12 dihedral angle of retinal, leads to an internally strained retinal, which relaxes after a few nanoseconds by a switching of the β-ionone ring to an essentially planar all-*trans* conformation (227, 229). Transition of 6-s-*cis*- to 6-s-*trans*-retinal (227, 229) is observed, as well as changes in the hydrogen-bond network with the counter ion Glu113. Furthermore, the energy initially stored internally in the distorted retinal is transformed into non-bonded interactions of retinal with its environment (229). The most significant conformational changes in the binding-pocket are the straightening of retinal polyene chain and the separation of its β-ionone ring from Trp265 in a motion of the β-ionone ring from Tyr268 to Ala169 (227, 229) in accordance with previous photo-cross-linking data (24) but against the more recent results gained by solid state NMR (213, 254).

Comparison of simulations with Glu181 in different protonation states strongly suggests that this loop residue located close to the 11-*cis* bond bears a negative charge (227).

Molecular dynamics simulations also provide evidence that the protein tightly confines the absolute conformation of the retinal around the C12-C13 bond to a positive helicity (227).

This structural transition of the retinal induces in turn significant conformational changes of the protein backbone, especially in helix VI (227, 229), and the cytoplasmic loop 2 where transient structural changes involving the conserved salt bridge between Glu-134 and Arg-135 are observed (229). One of the MD studies suggests also a folding of the cytoplasmic loop 3 connecting helices V and VI, inward toward the bundle core (229), while experimental data suggest the creation of the binding site for transducin in the metarhodopsin II state by motion of the cytoplasmic loops 2 and 3

outward (2, 5). These features prepare the protein for major structural transformations achieved later in the photocycle and suggest a possible molecular mechanism for the early steps of intra-molecular signal transduction in a prototypical G protein-coupled receptor (227, 229). The rotation of helix VI, which was postulated by site-directed spin labelling (76) was not observed, but rather generated by steered molecular dynamics (229).

CONCLUSIONS & PERSPECTIVES

Conformation of retinal in the ground state is accurately defined, where both experiment and theory provide the same picture.

The main contributors for this picture are solid state NMR and rotational resonance for providing the first accurate distance restraints on the chromophore and x-ray crystallography for defining the binding pocket and later confirming the geometry of the chromophore. Quantum simulations and DFT in particular has considerably matured, being able to provide an accurate picture of the chromophore based solely on the geometry of the binding pocket, ahead of the publication of the experiment on the chromophore.

It is expected that solid state NMR will become a major player for future structural studies of retinal photo-intermediates, as it is the only experimental method which so far has been able to provide accurate structural (distance restraints) for retinal photo-intermediates, using rotational resonance based methods. Furthermore, crystallization of rhodopsin in one of active-states appears challenging due to both the short lifetime of the activated-states and the complex kinetic of activation resulting in the coexistence of several different intermediate states during illumination.

ACKNOWLEDGEMENTS

VL would like to thank Dr Paul Spooner for extremely valuable discussions about retinal and rhodopsin, and Dr Minoru Sugihara for providing the coordinates of the 6-s-cis conformer of retinal in its binding-pocket as computed with DFT.

LITERATURE CITED

1. Abdalla S, Lother H, Quitterer U. 2000. AT1-receptor heterodimers show enhanced G-protein activation and altered receptor sequestration. *Nature* 407: 94
2. Acharya S, Saad Y, Karnik SS. 1997. Transducin-alpha C-terminal peptide binding site consists of C-D and E-F loops of rhodopsin. *J Biol Chem* 272: 6519
3. Albert AD, Yeagle PL. 2000. Domain approach to three-dimensional structure of rhodopsin using high-resolution nuclear magnetic resonance. *Methods Enzymol* 315: 107
4. Albert AD, Yeagle PL. 2002. Structural studies on rhodopsin. *Biochim Biophys Acta* 1565: 183
5. Arimoto R, Kisselev OG, Makara GM, Marshall GR. 2001. Rhodopsin-transducin interface: studies with conformationally constrained peptides. *Biophys J* 81: 3285
6. Bagley KA, Balogh-Nair V, Croteau AA, Dollinger G, Ebrey TG, et al. 1985. Fourier-transform infrared difference spectroscopy of rhodopsin and its photoproducts at low temperature. *Biochemistry* 24: 6055
7. Ball LE, Oatis JE, Jr., Dharmasiri K, Busman M, Wang J, et al. 1998. Mass spectrometric analysis of integral membrane proteins: application to complete mapping of bacteriorhodopsins and rhodopsin. *Protein Sci* 7: 758
8. Ballesteros JA, Weinstein H. 1992. Analysis and refinement of criteria for predicting the structure and relative orientations of transmembranal helical domains. *Biophys J* 62: 107
9. Baneres JL, Parello J. 2003. Structure-based analysis of GPCR function: evidence for a novel pentameric assembly between the dimeric leukotriene B4 receptor BLT1 and the G-protein. *J Mol Biol* 329: 815
10. Baylor DA, Burns ME. 1998. Control of rhodopsin activity in vision. *Eye* 12 (Pt 3b): 521
11. Bennett N, Michel-Villaz M, Kuhn H. 1982. Light-induced interaction between rhodopsin and the GTP-binding protein. Metarhodopsin II is the major photoproduct involved. *Eur J Biochem* 127: 97
12. Berman H, Westbrook J, Feng Z,Gilliland G, Bhat TN, Weissig H,Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28: 235
13. Bifone A, de Groot HJ, Buda F. 1997. Energy Storage in the Primary Photoproduct of Vision. *Journal of Physical Chemistry B* 101: 2954–2958
14. Birge RR. 1981. Photophysics of light transduction in rhodopsin and bacteriorhodopsin. *Annu Rev Biophys Bioeng* 10: 315
15. Birge RR. 1990. Nature of the primary photochemical events in rhodopsin and bacteriorhodopsin. *Biochim Biophys Acta* 1016: 293
16. Birge RR, Bocian DF, Hubbard LM. 1982. Origins of Inhomogeneous Broadening in the Vibronic Spectra of Visual Chromophores and Visual Pigments. *J Am Chem Soc* 104: 1196–1207
17. Birge RR, Einterz CM, Knapp HM, Murray LP. 1988. The nature of the primary photochemical events in rhodopsin and isorhodopsin. *Biophys J* 53: 367
18. Birge RR, Hubbard LM. 1980. Molecular dynamics of *cis-trans* isomerization in rhodopsin. *J Am Chem Soc* 102: 2195–2205
19. Birge RR, Hubbard LM. 1981. Molecular dynamics of *trans-cis* isomerization in bathorhodopsin. *Biophys J* 34: 517
20. Birge RR, Murray LP, Pierce BM, Akita H, Balogh-Nair V, et al. 1985. Two-photon spectroscopy of locked-11-*cis*-rhodopsin: evidence for a protonated Schiff base in a neutral protein binding site. *Proc Natl Acad Sci U S A* 82: 4117
21. Blasie JK, Dewey MM, Blaurock AE, Worthington CR. 1965. Electron microscope and low-angle x-ray diffraction studies on outer segment membranes from the retina of the frog. *J Mol Biol* 14: 143
22. Blasie JK, Worthington CR. 1969. Molecular localization of frog retinal receptor photopigment by electron

microscopy and low-angle X-ray diffraction. *J Mol Biol* 39: 407

23. Blasie JK, Worthington CR. 1969. Planar liquid-like arrangement of photopigment molecules in frog retinal receptor disk membranes. *J Mol Biol* 39: 417
24. Borhan B, Souto ML, Imai H, Shichida Y, Nakanishi K. 2000. Movement of retinal along the visual transduction path. *Science* 288: 2209
25. Boucher F, Leblanc RM. 1985. Energy storage in the primary photoreaction of bovine rhodopsin. A photoacoustic study. *Photochem Photobiol* 41: 459
26. Bourne HR, Meng EC. 2000. Structure. Rhodopsin sees the light. *Science* 289: 733
27. Bownds D. 1967. Site of attachment of retinal in rhodopsin. *Nature* 216: 1178
28. Buda F, de Groot HJ, Bifone A. 1996. Charge Localization and Dynamics in Rhodopsin. *Physical Review Letters* 77: 4474–4477
29. Buda F, Giannozzi P, Mauri FJ. 2000. Density Functional Theory Study of the Structure and ^{13}C Chemical Shifts of Retinylidene Iminium Salts. *Journal of Physical Chemistry B* 104: 9048–9053
30. Buda F, Touw SI, de Groot HJ. 2001. An *ab initio* molecular dynamics modeling of the primary photochemical event in vision. In *Perspective on Solid State NMR in Biology*, ed. SR Kiihne, HJ de Groot, pp. 111–122. Dordrecht: Kluwer Academic Publishers
31. Burns ME, Baylor DA. 2001. Activation, deactivation, and adaptation in vertebrate photoreceptor cells. *Annu Rev Neurosci* 24: 779
32. Busch GE, Applebury ML, Lamola AA, Rentzepis PM. 1972. Formation and decay of prelumirhodopsin at room temperatures. *Proc Natl Acad Sci U S A* 69: 2802
33. Buss V. 2001. Inherent chirality of the retinal chromophore in rhodopsin-A nonempirical theoretical analysis of chiroptical data. *Chirality* 13: 13
34. Buss V, Kolster K, Terstegen F, Vahrenhorst R. 1998. Absolute Sense of Twist of the C12=C13 Bond of the Retinal Chromophore in Rhodopsin–Semiempirical and Nonempirical Calculations of Chiroptical Data. *Angew. Chem. Int. Ed.* 37: 1893–1895
35. Buss V, Sugihara M, Entel P, Hafner J. 2003. Thr94 and Wat2b Effect Protonation of the Retinal Chromophore in Rhodopsin. *Angew. Chem. Int. Ed.* 42: 3245–3247
36. Cai K, Langen R, Hubbell WL, Khorana HG. 1997. Structure and function in rhodopsin: topology of the C-terminal polypeptide chain in relation to the cytoplasmic loops. *Proc Natl Acad Sci U S A* 94: 14267
37. Canals M, Marcellino D, Fanelli F, Ciruela F, de Benedetti P, et al. 2003. Adenosine A2A-dopamine D2 receptor-receptor heteromerization: qualitative and quantitative assessment by fluorescence and bioluminescence energy transfer. *J Biol Chem* 278: 46741
38. Carravetta M, Edén M, Johannessen OG, Luthman H, Verdegem P, et al. 2001. Estimation of Carbon-Carbon Bond Lengths and Medium-Range Internuclear Distances by Solid-State Nuclear Magnetic Resonance. *J Am Chem Soc* 123: 10628–10638
39. Carravetta M, Zhao X, Johannessen OG, Lai WC, Verhoeven MA, et al. 2004. Protein-induced bonding perturbation of the rhodopsin chromophore detected by double-quantum solid-state NMR. *J Am Chem Soc* 126: 3948
40. Carrillo JJ, Pediani J, Milligan G. 2003. Dimers of class A G protein-coupled receptors function via agonist-mediated trans-activation of associated G proteins. *J Biol Chem* 278: 42578
41. Chabre M, Saibil H, Worcester DL. 1976. Neutron diffraction studies of oriented retinal rods. *Brookhaven Symp Biol*: III77
42. Cohen GB, Oprian DD, Robinson PR. 1992. Mechanism of activation and inactivation of opsin: role of Glu113 and Lys296. *Biochemistry* 31: 12592
43. Cone RA. 1972. Rotational diffusion of rhodopsin in the visual receptor membrane. *Nat New Biol* 236: 39
44. Cooper A. 1979. Energy uptake in the first step of visual excitation. *Nature* 282: 531
45. Cooper A. 1981. Rhodopsin photoenergetics: lumirhodopsin and the complete energy profile. *FEBS Lett* 123: 324
46. Cooper A, Converse CA. 1976. Energetics of primary processes in

visula escitation: photocalorimetry of rhodopsin in rod outer segment membranes. *Biochemistry* 15: 2970

47. Cooper A, Dixon SF, Nutley MA, Robb JL. 1987. Mechanism of Retinal Schiff Base Formation and Hydrolysis in Relation to Visual Pigment Photolysis and Regeneration: Resonance Raman Spectroscopy of a Tetrahedral Carbinolamine Intermediate and Oxygen-18 Labeling of Retinal at the Metarhodopsin Stage in Photoreceptor Membranes. *J Am Chem Soc* 109: 7254–7263
48. Cordes FS, Bright JN, Sansom MS. 2002. Proline-induced distortions of transmembrane helices. *J Mol Biol* 323: 951
49. Creemers AF, Kiihne S, Bovee-Geurts PH, DeGrip WJ, Lugtenburg J, de Groot HJ. 2002. ^{1}H and ^{13}C MAS NMR evidence for pronounced ligand-protein interactions involving the ionone ring of the retinylidene chromophore in rhodopsin. *Proc Natl Acad Sci U S A* 99: 9101
50. Creemers AF, Klaassen CH, Bovee-Geurts PH, Kelle R, Kragl U, et al. 1999. Solid state ^{15}N NMR evidence for a complex Schiff base counterion in the visual G-protein-coupled receptor rhodopsin. *Biochemistry* 38: 7195
51. Crescitelli F. 1977. In *Handbook of Sensory Physiology*. Berlin: Springer-Verlag
52. Crocker E, Patel AB, Eilers M, Jayaraman S, Getmanova E, et al. 2004. Dipolar assisted rotational resonance NMR of tryptophan and tyrosine in rhodopsin. *J Biomol NMR* 29: 11
53. Crozier PS, Stevens MJ, Forrest LR, Woolf TB. 2003. Molecular dynamics simulation of dark-adapted rhodopsin in an explicit membrane bilayer: coupling between local retinal and larger scale conformational change. *J Mol Biol* 333: 493
54. Davidson FF, Loewen PC, Khorana HG. 1994. Structure and function in rhodopsin: replacement by alanine of cysteine residues 110 and 187, components of a conserved disulfide bond in rhodopsin, affects the light-activated metarhodopsin II state. *Proc Natl Acad Sci U S A* 91: 4029
55. de Grip WJ, DeLange F, Klaassen CH, Verdegem PJ, Wallace-Williams S, et al. 1999. Photoactivation of rhodopsin: interplay between protein and chromophore. *Novartis Found Symp* 224: 102
56. de Groot HJ, Harbison GS, Herzfeld J, Griffin RG. 1989. Nuclear magnetic resonance study of the Schiff base in bacteriorhodopsin: counterion effects on the ^{15}N shift anisotropy. *Biochemistry* 28: 3346
57. DeGrip WJ, DeCaluwe GLJ, Pistorius AMA, van Aalten DMF, Janssen JJM, et al. 1993. In *Molecular mechanism of generation of electric signals in sensory cells*, ed. F Tokunaga, pp. 26–40. Osaka, Japan: The Taniguchi Foundation
58. DeLange F, Bovee-Geurts PH, VanOostrum J, Portier MD, Verdegem PJ, et al. 1998. An additional methyl group at the 10-position of retinal dramatically slows down the kinetics of the rhodopsin photocascade. *Biochemistry* 37: 1411
59. Deng H, Huang L, Callender R, Ebrey T. 1994. Evidence for a bound water molecule next to the retinal Schiff base in bacteriorhodopsin and rhodopsin: a resonance Raman study of the Schiff base hydrogen/deuterium exchange. *Biophys J* 66: 1129
60. Deng H, Manor D, Weng G, Rath P, Koutalos Y, et al. 1991. A resonance Raman study of octopus bathorhodopsin with deuterium labeled retinal chromophores. *Photochem Photobiol* 54: 1001
61. Doukas AG, Aton B, Callender RH, Ebrey TG. 1978. Resonance Raman studies of bovine metarhodopsin I and metarhodopsin II. *Biochemistry* 17: 2430
62. Downer NW, Cone RA. 1985. Transient dichroism in photoreceptor membranes indicates that stable oligomers of rhodopsin do not form during excitation. *Biophys J* 47: 277
63. Ebrey T, Koutalos Y. 2001. Vertebrate photoreceptors. *Prog Retin Eye Res* 20: 49
64. Ebrey TG. 2000. pKa of the protonated Schiff base of visual pigments. *Methods Enzymol* 315: 196
65. Eilers M, Reeves PJ, Ying W, Khorana HG, Smith SO. 1999. Magic angle spinning NMR of the protonated

retinylidene Schiff base nitrogen in rhodopsin: expression of 15N-lysine- and ^{13}C-glycine-labeled opsin in a stable cell line. *Proc Natl Acad Sci U S A* 96: 487

66. Eilers M, Ying W, Reeves PJ, Khorana HG, Smith SO. 2002. Magic angle spinning nuclear magnetic resonance of isotopically labeled rhodopsin. *Methods Enzymol* 343: 212
67. Engelman DM, Steitz TA, Goldman A. 1986. Identifing nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Ann. Rev. Biophys. Biophys. Chem.* 15: 321
68. Ernst OP, Bartl FJ. 2002. Active states of rhodopsin. *Chembiochem* 3: 968
69. Ernst OP, Meyer CK, Marin EP, Henklein P, Fu WY, et al. 2000. Mutation of the fourth cytoplasmic loop of rhodopsin affects binding of transducin and peptides derived from the carboxyl-terminal sequences of transducin alpha and gamma subunits. *J Biol Chem* 275: 1937
70. Esquerra RM, Che D, Shapiro DB, Lewis JW, Bogomolni RA, et al. 1996. Chromophore reorientations in the early photolysis intermediates of bacteriorhodopsin. *Biophys J* 70: 962
71. Fahmy K, Jager F, Beck M, Zvyaga TA, Sakmar TP, Siebert F. 1993. Protonation states of membrane-embedded carboxylic acid groups in rhodopsin and metarhodopsin II: a Fourier-transform infrared spectroscopy study of site-directed mutants. *Proc Natl Acad Sci U S A* 90: 10206
72. Fahmy K, Sakmar TP. 1993. Light-dependent transducin activation by an ultraviolet-absorbing rhodopsin mutant. *Biochemistry* 32: 9165
73. Fahmy K, Siebert F, Sakmar TP. 1994. A mutant rhodopsin photoproduct with a protonated Schiff base displays an active-state conformation: a Fourier-transform infrared spectroscopy study. *Biochemistry* 33: 13700
74. Fahmy K, Siebert F, Sakmar TP. 1995. Photoactivated state of rhodopsin and how it can form. *Biophys Chem* 56: 171
75. Fain GL, Matthews HR, Cornwall MC, Koutalos Y. 2001. Adaptation in vertebrate photoreceptors. *Physiol Rev* 81: 117
76. Farrens DL, Altenbach C, Yang K, Hubbell WL, Khorana HG. 1996. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* 274: 768
77. Farrens DL, Khorana HG. 1995. Structure and function in rhodopsin. Measurement of the rate of metarhodopsin II decay by fluorescence spectroscopy. *J Biol Chem* 270: 5073
78. Feng X, Verdegem PJ, Eden M, Sandstrom D, Lee YK, et al. 2000. Determination of a molecular torsional angle in the metarhodopsin-I photointermediate of rhodopsin by double-quantum solid-state NMR. *J Biomol NMR* 16: 1
79. Feng X, Verdegem PJ, Lee YK, Sandström D, Edén M, et al. 1997. Direct determination of a molecular torsional angle in the membrane protein rhodopsin by solid-state NMR. *J Am Chem Soc* 119: 6853–6857
80. Ferretti L, Karnik SS, Khorana HG, Nassal M, Oprian DD. 1986. Total synthesis of a gene for bovine rhodopsin. *Proc Natl Acad Sci U S A* 83: 599
81. Filipek S, Stenkamp RE, Teller DC, Palczewski K. 2003. G protein-coupled receptor rhodopsin: a prospectus. *Annu Rev Physiol* 65: 851
82. Filipek S, Teller DC, Palczewski K, Stenkamp R. 2003. The crystallographic model of rhodopsin and its use in studies of other G protein-coupled receptors. *Annu Rev Biophys Biomol Struct* 32: 375
83. Fotiadis D, Liang Y, Filipek S, Saperstein DA, Engel A, Palczewski K. 2003. Atomic-force microscopy: Rhodopsin dimers in native disc membranes. *Nature* 421: 127
84. Fotiadis D, Liang Y, Filipek S, Saperstein DA, Engel A, Palczewski K. 2004. The G protein-coupled receptor rhodopsin in the native membrane. *FEBS Lett* 564: 281
85. Franke RR, Konig B, Sakmar TP, Khorana HG, Hofmann KP. 1990. Rhodopsin mutants that bind but fail to activate transducin. *Science* 250: 123
86. Fukada Y, Shichida Y, Yoshizawa T, Ito M, Kodama A, Tsukida K. 1984. Studies on structure and function of rhodopsin by use of

cyclopentatrienylidene 11-cis-locked-rhodopsin. *Biochemistry* 23: 5826

87. Fukuda MN, Papermaster DS, Hargrave PA. 1979. Rhodopsin carbohydrate. Structure of small oligosaccharides attached at two sites near the NH2 terminus. *J Biol Chem* 254: 8201
88. Ganter UM, Gartner W, Siebert F. 1988. Rhodopsin-lumirhodopsin phototransition of bovine rhodopsin investigated by Fourier transform infrared difference spectroscopy. *Biochemistry* 27: 7480
89. Garavelli M, Bernadi F, Olivucci M, Vreven T, Klein S, et al. 1998. Potential-energy surfaces for ultrafast photochemistry static and dynamics aspects. *Faraday Discuss* 110: 51–70
90. Gether U, Kobilka BK. 1998. G Protein-coupled Receptors. II. Mechanism of agonist activation. *Journal of Biological Chemistry* 273: 17979
91. Gether U, Lin S, Ghanouni P, Ballesteros JA, Weinstein H, Kobilka BK. 1997. Agonists induce conformational changes in transmembrane domains III and VI of the beta2 adrenoceptor. *Embo J* 16: 6737
92. Gilardi RD, Karle J. 1972. The crystal and molecular structure of 11-cis-retinal. *Acta Crystallographica Section B* 28: 2605–2612
93. Gorczyca WA, Gray-Keller MP, Detwiler PB, Palczewski K. 1994. Purification and physiological evaluation of a guanylate cyclase activating protein from retinal rods. *Proc Natl Acad Sci U S A* 91: 4014
94. Grobner G, Burnett I, Choi G, Glaubitz C, Mason J, Watts A. 1998. Structure and conformation of the retinal chromophore in bovine rhodopsin. *Biochem Soc Trans* 26: S308
95. Grobner G, Burnett IJ, Glaubitz C, Choi G, Mason AJ, Watts A. 2000. Observations of light-induced structural changes of retinal within rhodopsin. *Nature* 405: 810
96. Grobner G, Choi G, Burnett IJ, Glaubitz C, Verdegem PJ, et al. 1998. Photoreceptor rhodopsin: structural and conformational study of its chromophore 11-*cis* retinal in oriented membranes by deuterium solid state NMR. *FEBS Letters* 422: 201
97. Grobner G, Taylor A, Williamson PT, Choi G, Glaubitz C, et al. 1997. Macroscopic orientation of natural and model membranes for structural studies. *Anal Biochem* 254: 132
98. Groebner G, Burnett I, Choi G, Glaubitz C, Mason J, Watts A. 1998. Structure and conformation of the retinal chromophore in bovine rhodopsin. *Biochem Soc Trans* 26: S308
99. Guo W, Shi L, Javitch JA. 2003. The fourth transmembrane segment forms the interface of the dopamine D2 receptor homodimer. *J Biol Chem* 278: 4385
100. Hamm HE. 2001. How Activated Receptors Couple to G Proteins. *Proc Natl Acad Sci U S A* 98: 4819
101. Hamm HE. 2001. How activated receptors couple to G proteins. *Proc Natl Acad Sci U S A* 98: 4819
102. Han M, DeDecker BS, Smith SO. 1993. Localization of the retinal protonated Schiff base counterion in rhodopsin. *Biophys J* 65: 899
103. Han M, Lin SW, Smith SO, Sakmar TP. 1996. The effects of amino acid replacements of glycine 121 on transmembrane helix 3 of rhodopsin. *J Biol Chem* 271: 32330
104. Han M, Lou J, Nakanishi K, Sakmar TP, Smith SO. 1997. Partial agonist activity of 11-*cis*-retinal in rhodopsin mutants. *J Biol Chem* 272: 23081
105. Han M, Smith SO. 1995. High-resolution structural studies of the retinal--Glu113 interaction in rhodopsin. *Biophys Chem* 56: 23
106. Han M, Smith SO. 1995. NMR constraints on the location of the retinal chromophore in rhodopsin and bathorhodopsin. *Biochemistry* 34: 1425
107. Harbison GS, Herzfeld J, Griffin RG. 1983. Solid-state nitrogen-15 nuclear magnetic resonance study of the Schiff base in bacteriorhodopsin. *Biochemistry* 22: 1
108. Harbison GS, Mulder FM, Pardoen JA, Lugtenburg J, Herzfeld J, Griffin RG. 1985. High-Resolution Carbon-13 NMR of Retinal Derivatives in the Solid State. *J Am Chem Soc* 107: 4809–4816
109. Harbison GS, Smith SO, Pardoen JA, Courtin JM, Lugtenburg J, et al. 1985. Solid-state ^{13}C NMR detection of a perturbed 6-s-trans chromophore in

bacteriorhodopsin. *Biochemistry* 24: 6955

110. Hargrave PA. 1977. The amino-terminal tryptic peptide of bovine rhodopsin. A glycopeptide containing two sites of oligosaccharide attachment. *Biochim Biophys Acta* 492: 83
111. Hargrave PA, Bownds D, Wang JK, McDowell JH. 1982. Retinyl peptide isolation and characterization. *Methods Enzymol* 81: 211
112. Hargrave PA, McDowell JH. 1992. Rhodopsin and phototransduction. *Int Rev Cytol* 137B: 49
113. Hargrave PA, McDowell JH. 1992. Rhodopsin and phototransduction: a model system for G protein-linked receptors. *Faseb J* 6: 2323
114. Hargrave PA, McDowell JH, Feldmann RJ, Atkinson PH, Rao JK, Argos P. 1984. Rhodopsin's protein and carbohydrate structure: selected aspects. *Vision Res* 24: 1487
115. Hayashi S, Tajkhorshid E, Schulten K. 2002. Structural changes during the formation of early intermediates in the bacteriorhodopsin photocycle. *Biophys J* 83: 1281
116. Hofmann KP. 1999. Signalling states of photoactivated rhodopsin. *Novartis Found Symp* 224: 158
117. Hofmann KP. 2000. In *Handbook of Biological Physics: Molecular Mechanisms in Visual Transduction*, ed. DG Stavenga, WJ DeGrip, EN Pugh, pp. 91–142. New-York: Elsevier/North-Holland
118. Honig B, Ebrey T, Callender RH, Dinur U, Ottolenghi M. 1979. Photoisomerization, energy storage, and charge separation: a model for light energy transduction in visual pigments and bacteriorhodopsin. *Proc Natl Acad Sci U S A* 76: 2503
119. Honig B, Greenberg AD, Dinur U, Ebrey TG. 1976. Visual-pigment spectra: implications of the protonation of the retinal Schiff base. *Biochemistry* 15: 4593
120. Honig B, Hudson B, Sykes BD, Karplus M. 1971. Ring orientation in -ionone and retinals. *Proc Natl Acad Sci U S A* 68: 1289
121. Hubbard R. 1966. The stereoisomerization of 11-*cis*-retinal. *J Biol Chem* 241: 1814
122. Hubbard R. 1969. Absorption spectrum of rhodopsin: 500 nm absorption band. *Nature* 221: 432
123. Hubbell WL, Altenbach C, Hubbell CM, Khorana HG. 2003. Rhodopsin structure, dynamics, and activation: a perspective from crystallography, site-directed spin labeling, sulfhydryl reactivity, and disulfide cross-linking. *Adv Protein Chem* 63: 243
124. Huber T, Botelho AV, Beyer K, Brown MF. 2004. Membrane model for the G-protein-coupled receptor rhodopsin: hydrophobic interface and dynamical structure. *Biophys J* 86: 2078
125. Hug SJ, Lewis JW, Einterz CM, Thorgeirsson TE, Kliger DS. 1990. Nanosecond photolysis of rhodopsin: evidence for a new, blue-shifted intermediate. *Biochemistry* 29: 1475
126. Humphrey W, Lu H, Logunov I, Werner HJ, Schulten K. 1998. Three electronic state model of the primary phototransformation of bacteriorhodopsin. *Biophys J* 75: 1689
127. Hwa J, Reeves PJ, Klein-Seetharaman J, Davidson F, Khorana HG. 1999. Structure and function in rhodopsin: further elucidation of the role of the intradiscal cysteines, Cys-110, -185, and -187, in rhodopsin folding and function. *Proc Natl Acad Sci U S A* 96: 1932
128. Imai H, Mizukami T, Imamoto Y, Shichida Y. 1994. Direct observation of the thermal equilibria among lumirhodopsin, metarhodopsin I, and metarhodopsin II in chicken rhodopsin. *Biochemistry* 33: 14351
129. Ishiguro M, Hirano T, Oyama Y. 2003. Modelling of photointermediates suggests a mechanism of the flip of the beta-ionone moiety of the retinylidene chromophore in the rhodopsin photocascade. *Chembiochem* 4: 228
130. Ito M, Katuta Y, Imamoto Y, Shichida Y, Yoshizawa T. 1992. Conformational analysis of the rhodopsin chromophore using bicyclic retinal analogues. *Photochem Photobiol* 56: 915–919
131. Jackson ML, Litman BJ. 1982. Rhodopsin-phospholipid reconstitution by dialysis removal of octyl glucoside. *Biochemistry* 21: 5601
132. Jager F, Fahmy K, Sakmar TP, Siebert F. 1994. Identification of glutamic acid

113 as the Schiff base proton acceptor in the metarhodopsin II photointermediate of rhodopsin. *Biochemistry* 33: 10878

133. Jager F, Jager S, Krutle O, Friedman N, Sheves M, et al. 1994. Interactions of the beta-ionone ring with the protein in the visual pigment rhodopsin control the activation mechanism. An FTIR and fluorescence study on artificial vertebrate rhodopsins. *Biochemistry* 33: 7389
134. Jäger F, Ujj L, Atkinson GH. 1997. Vibrational spectrum of bathorhodopsin in the room-temperature rhodopsin photoreaction. *J Am Chem Soc* 119: 12610–12618
135. Jager S, Han M, Lewis JW, Szundi I, Sakmar TP, Kliger DS. 1997. Properties of early photolysis intermediates of rhodopsin are affected by glycine 121 and phenylalanine 261. *Biochemistry* 36: 11804
136. Jager S, Lewis JW, Zvyaga TA, Szundi I, Sakmar TP, Kliger DS. 1997. Chromophore structural changes in rhodopsin from nanoseconds to microseconds following pigment photolysis. *Proc Natl Acad Sci U S A* 94: 8557
137. Jager S, Lewis JW, Zvyaga TA, Szundi I, Sakmar TP, Kliger DS. 1997. Time-resolved spectroscopy of the early photolysis intermediates of rhodopsin Schiff base counterion mutants. *Biochemistry* 36: 1999
138. Jang GF, Kuksa V, Filipek S, Bartl F, Ritter E, et al. 2001. Mechanism of rhodopsin activation as examined with ring-constrained retinal analogs and the crystal structure of the ground state protein. *J Biol Chem* 276: 26148
139. Kakitani H, Kakitani T, Rodman H, Honig B. 1985. On the mechanism of wavelength regulation in visual pigments. *Photochem Photobiol* 41: 471
140. Kampermann H, Kolster K, Buss V. 2001. Conformation and absolute configuration of 11,12-cyclopropyl retinal analogs - an RHF/DFT/CIS study. *Journal of Molecular Modeling* 7: 132–139
141. Kandori H, Kinoshita N, Yamazaki Y, Maeda A, Shichida Y, et al. 1999. Structural change of threonine 89 upon photoisomerization in bacteriorhodopsin as revealed by polarized FTIR spectroscopy. *Biochemistry* 38: 9676
142. Kandori H, Shichida Y, Yoshizawa T. 2001. Photoisomerization in rhodopsin. *Biochemistry (Mosc)* 66: 1197
143. Karlsson T, Edén M, Luthman H, Levitt MH. 2000. Efficient Double-Quantum Excitation in Rotational Resonance NMR. *Journal of Magnetic Resonance* 145: 95
144. Karnik SS, Khorana HG. 1990. Assembly of functional rhodopsin requires a disulfide bond between cysteine residues 110 and 187. *J Biol Chem* 265: 17520
145. Karnik SS, Ridge KD, Bhattacharya S, Khorana HG. 1993. Palmitoylation of bovine opsin and its cysteine mutants in COS cells. *Proc Natl Acad Sci U S A* 90: 40
146. Karnik SS, Sakmar TP, Chen HB, Khorana HG. 1988. Cysteine residues 110 and 187 are essential for the formation of correct structure in bovine rhodopsin. *Proc Natl Acad Sci U S A* 85: 8459
147. Katragadda M, Chopra A, Bennett M, Alderfer JL, Yeagle PL, Albert AD. 2001. Structures of the transmembrane helices of the G-protein coupled receptor, rhodopsin. *J Pept Res* 58: 79
148. Khorana HG. 1992. Rhodopsin, photoreceptor of the rod cell. An emerging pattern for structure and function. *J Biol Chem* 267: 1
149. Kim JE, McCamant DW, Zhu L, Mathies RA. 2001. Resonance Raman structural evidence that the *cis*-to-*trans* isomerization in rhodopsin occurs in femtoseoncds. *Journal of Physical Chemistry B* 105: 1240–1249
150. Kim JE, Pan D, Mathies RA. 2003. Picosecond dynamics of G-protein coupled receptor activation in rhodopsin from time-resolved UV resonance Raman spectroscopy. *Biochemistry* 42: 5169
151. Klco JM, Lassere TB, Baranski TJ. 2003. C5a receptor oligomerization. I. Disulfide trapping reveals oligomers and potential contact surfaces in a G protein-coupled receptor. *J Biol Chem* 278: 35345
152. Klein-Seetharaman J. 2002. Dynamics in rhodopsin. *Chembiochem* 3: 981
153. Kliger DS, Lewis JW. 1995. Spectral and kinetic characterization of visual

pigment photointermediates. *Israel Journal of Chemistry* 35: 289–307

154. Klinger AL, Braiman MS. 1992. Structural comparison of metarhodopsin II, metarhodopsin III, and opsin based on kinetic analysis of Fourier transform infrared difference spectra. *Biophys J* 63: 1244
155. Kosower EM. 1988. Assignment of groups responsible for the "opsin shift" and light absorptions of rhodopsin and red, green, and blue iodopsins (cone pigments). *Proc Natl Acad Sci U S A* 85: 1076
156. Koutalos Y, Yau KW. 1993. A rich complexity emerges in phototransduction. *Curr Opin Neurobiol* 3: 513
157. Krebs W, Kuhn H. 1977. Structure of isolated bovine rod outer segment membranes. *Exp Eye Res* 25: 511
158. Kühne W. 1977. *Vision Res* 17: 1273–1316
159. Kuwayama S, Imai H, Hirano T, Terakita A, Shichida Y. 2002. Conserved proline residue at position 189 in cone visual pigments as a determinant of molecular properties different from rhodopsins. *Biochemistry* 41: 15245
160. Kwok-Keung Fung B, Stryer L. 1980. Photolyzed rhodopsin catalyzes the exchange of GTP for bound GDP in retinal rod outer segments. *Proc Natl Acad Sci U S A* 77: 2500
161. La Penna G, Buda F, Bifone A, de Groot HJ. 1998. The transition state in the isomerization of rhodopsin. *Chemical Physics Letters* 294: 447–453
162. Lambright DG, Noel JP, Hamm HE, Sigler PB. 1994. Structural determinants for activation of the alpha-subunit of a heterotrimeric G protein. *Nature* 369: 621
163. Langen R, Cai K, Altenbach C, Khorana HG, Hubbell W. 1999. Structural Features of the C-Terminal Domain of Bovine Rhodopsin: A Site-Directed Spin-Labeling Study. *Biochemistry* 38: 7918
164. Lee SP, O'Dowd BF, Rajaram RD, Nguyen T, George SR. 2003. D2 dopamine receptor homodimerization is mediated by multiple sites of interaction, including an intermolecular interaction involving transmembrane domain 4. *Biochemistry* 42: 11023
165. Lee YK, Kurur ND, Helmle M, Johannessen OG, Nielsen NC, Levitt MH. 1995. Efficient dipolar recoupling in the NMR of rotating solids. A sevenfold symmetric radiofrequency pulse sequence. *Chem Phys Lett* 242: 304–309
166. Lewis JW, Fan GB, Sheves M, Szundi I, Kliger DS. 2001. Steric barrier to bathorhodopsin decay in 5-demethyl and mesityl analogues of rhodopsin. *J Am Chem Soc* 123: 10024
167. Lewis JW, van Kuijk FJ, Thorgeirsson TE, Kliger DS. 1991. Photolysis intermediates of human rhodopsin. *Biochemistry* 30: 11372
168. Li J, Edwards PC, Burghammer M, Villa C, Schertler GF. 2004. Structure of Bovine Rhodopsin in a Trigonal Crystal Form. *J. Mol. Biol.* (In Press)
169. Liang Y, Fotiadis D, Filipek S, Saperstein DA, Palczewski K, Engel A. 2003. Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes. *J Biol Chem* 278: 21655
170. Liebman PA, Entine G. 1974. Lateral diffusion of visual pigment in photorecptor disk membranes. *Science* 185: 457
171. Liebman PA, Parker KR, Dratz EA. 1987. The molecular mechanism of visual excitation and its relation to the structure and composition of the rod outer segment. *Annu. Rev. Physiol.* 49: 765
172. Liebman PA, Weiner HL, Drzymala RE. 1982. Lateral diffusion of visual pigment in rod disk membranes. *Methods Enzymol* 81: 660
173. Lin SW, Groesbeek M, van der Hoef I, Verdegem P, Lugtenburg J, Mathies RA. 1998. Vibrational assignment of torsional normal modes of rhodopsin: probing excited-state isomerization dynamics along the reactive C11=C12 torsion coordinate. *Journal of Physical Chemistry B* 102: 2787–2806
174. Liu W, Eilers M, Patel AB, Smith SO. 2004. Helix packing moments reveal diversity and conservation in membrane protein structure. *J Mol Biol* 337: 713
175. Lou J, Hashimoto K, Berova N, Nakanishi K. 1999. Enantioselective Binding of an 11-cis-Locked

Cyclopropyl Retinal. The Conformation of Retinal in Bovine Rhodopsin. *Organic Letters* 1: 51–54

176. Lugtenburg J, Mathies RA, Griffin RG, Herzfeld J. 1988. Structure and function of rhodopsins from solid state NMR and resonance Raman spectroscopy of isotopic retinal derivatives. *Trends Biochem Sci* 13: 388
177. Mah TL, Lewis JW, Sheves M, Ottolenghi M, Kliger DS. 1995. Low-temperature trapping of early photointermediates of alpha-isorhodopsin. *Photochem Photobiol* 62: 356
178. Mathies R, Oseroff AR, Stryer L. 1976. Rapid-flow resonance Raman spectroscopy of photolabile molecules: rhodopsin and isorhodopsin. *Proc Natl Acad Sci U S A* 73: 1
179. Mathies RA. 1999. Photons, femtoseconds and dipolar interactions: a molecular picture of the primary events in vision. *Novartis Found Symp* 224: 70
180. Mathies RA, Lugtenburg J. 2000. Chapter 2: The primary photoreaction of rhodopsin. In *Handbook of Biological Physics*, ed. DG Stavenga, WJ de Grip, EN Pugh, Jr., pp. 55–90. Amsterdam: North-Holland
181. Melia TJ, Jr., Cowan CW, Angleson JK, Wensel TG. 1997. A comparison of the efficiency of G protein activation by ligand-free and light-activated forms of rhodopsin. *Biophys J* 73: 3182
182. Meng EC, Bourne HR. 2001. Receptor activation: what does the rhodopsin structure tell us? *Trends Pharmacol Sci* 22: 587
183. Menon ST, Han M, Sakmar TP. 2001. Rhodopsin: structural basis of molecular physiology. *Physiological reviews* 81: 1659
184. Milligan G, Ramsay D, Pascal G, Carrillo JJ. 2003. GPCR dimerisation. *Life Sci* 74: 181
185. Mitchell DC, Kibelbek J, Litman BJ. 1991. Rhodopsin in dimyristoylphosphatidylcholine-reconstituted bilayers forms metarhodopsin II and activates Gt. *Biochemistry* 30: 37
186. Mitchell DC, Litman BJ. 1999. Effect of protein hydration on receptor conformation: decreased levels of bound water promote metarhodopsin II formation. *Biochemistry* 38: 7617
187. Mollevanger LC, Kentgens AP, Pardoen JA, Courtin JM, Veeman WS, et al. 1987. High-resolution solid-state ^{13}C-NMR study of carbons C-5 and C-12 of the chromophore of bovine rhodopsin. Evidence for a 6-s-*cis* conformation with negative-charge perturbation near C-12. *Eur J Biochem* 163: 9
188. Morrison DF, O'Brien PJ, Pepperberg DR. 1991. Depalmitylation with hydroxylamine alters the functional properties of rhodopsin. *J Biol Chem* 266: 20118
189. Nagata T, Terakita A, Kandori H, Kojima D, Shichida Y, Maeda A. 1997. Water and peptide backbone structure in the active center of bovine rhodopsin. *Biochemistry* 36: 6164
190. Nagata T, Terakita A, Kandori H, Shichida Y, Maeda A. 1998. The hydrogen-bonding network of water molecules and the peptide backbone in the region connecting Asp83, Gly120, and Glu113 in bovine rhodopsin. *Biochemistry* 37: 17216
191. Nakayama TA, Khorana HG. 1990. Orientation of retinal in bovine rhodopsin determined by cross-linking using a photoactivatable analog of 11-cis-retinal. *J Biol Chem* 265: 15762
192. Nathans J. 1990. Determinants of visual pigment absorbance: identification of the retinylidene Schiff's base counterion in bovine rhodopsin. *Biochemistry* 29: 9746
193. Nishimura S, Kandori H, Maeda A. 1997. Transmembrane signaling mediated by water in bovine rhodopsin. *Photochem. Photobiol.* 66: 796
194. Noel JP, Hamm HE, Sigler PB. 1993. The 2.2 Å crystal structure of transducin-alpha complexed with GTP gamma S. *Nature* 366: 654
195. O'Brien PJ, Zatz M. 1984. Acylation of bovine rhodopsin by ^{3}H palmitic acid. *J Biol Chem* 259: 5054
196. Ohguro H, Johnson RS, Ericsson LH, Walsh KA, Palczewski K. 1994. Control of rhodopsin multiple phosphorylation. *Biochemistry* 33: 1023
197. Okada T, Ernst OP, Palczewski K, Hofmann KP. 2001. Activation of rhodopsin: new insights from structural

and biochemical studies. *Trends Biochem Sci* 26: 318

198. Okada T, Fujiyoshi Y, Silow M, Navarro J, Landau EM, Shichida Y. 2002. Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc Natl Acad Sci U S A* 99: 5982
199. Okada T, Palczewski K. 2001. Crystal structure of rhodopsin: implications for vision and beyond. *Curr Opin Struct Biol* 11: 420
200. Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. 2004. The Retinal Conformation and its Environment in Rhodopsin in Light of a new 2.2 Å Crsystal Structure. *J. Mol. Biol.* (In Press)
201. Oliveira L, Paiva AC, Sander C, Vriend G. 1994. A common step for signal transduction in G protein-coupled receptors. *Trends Pharmacol Sci* 15: 170
202. Oseroff AR, Callender RH. 1974. Resonance Raman spectroscopy of rhodopsin in retinal disk membranes. *Biochemistry* 13: 4243
203. Ovchinnikov Yu A, Abdulaev NG, Bogachuk AS. 1988. Two adjacent cysteine residues in the C-terminal cytoplasmic fragment of bovine rhodopsin are palmitylated. *FEBS Lett* 230: 1
204. Overton MC, Chinault SL, Blumer KJ. 2003. Oligomerization, biogenesis, and signaling is promoted by a glycophorin A-like dimerization motif in transmembrane domain 1 of a yeast G protein-coupled receptor. *J Biol Chem* 278: 49369
205. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, et al. 2000. Crystal Structure of Rhodopsin: A G protein-coupled receptor. *Science* 277: 687
206. Palings I, Pardoen JA, van den Berg E, Winkel C, Lugtenburg J, Mathies RA. 1987. Assignment of fingerprint vibrations in the resonance Raman spectra of rhodopsin, isorhodopsin, and bathorhodopsin: implications for chromophore structure and environment. *Biochemistry* 26: 2544
207. Palings I, Pardoen JA, van den Berg E, Winkel C, Lugtenburg J, Mathies RA. 1987. Assignment of fingerprint vibrations in the resonance Raman spectra of rhodopsin, isorhodopsin, and bathorhodopsin: implications for chromophore structure and environment. *Biochemistry* 26: 2544–56
208. Palings I, van den Berg EM, Lugtenburg J, Mathies RA. 1989. Complete assignment of the hydrogen out-of-plane wagging vibrations of bathorhodopsin: chromophore structure and energy storage in the primary photoproduct of vision. *Biochemistry* 28: 1498
209. Pan D, Ganim Z, Kim JE, Verhoeven MA, Lugtenburg J, Mathies RA. 2002. Time-resolved resonance Raman analysis of chromophore structural changes in the formation and decay of rhodopsin's BSI intermediate. *J Am Chem Soc* 124: 4857
210. Pan D, Mathies RA. 2001. Chromophore structure in lumirhodopsin and metarhodopsin I by time-resolved resonance Raman microchip spectroscopy. *Biochemistry* 40: 7929
211. Papac DI, Thornburg KR, Bullesbach EE, Crouch RK, Knapp DR. 1992. Palmitylation of a G-protein coupled receptor. Direct analysis by tandem mass spectrometry. *J Biol Chem* 267: 16889
212. Papermaster DS, Dreyer WJ. 1974. Rhodopsin content in the outer segment membranes of bovine and frog retinal rods. *Biochemistry* 13: 2438
213. Patel AB, Crocker E, Eilers M, Hirshfeld A, Sheves M, Smith SO. 2004. Coupling of retinal isomerization to the activation of rhodopsin. *Proceedings of the National Academy of Science* 101: 10048–10053
214. Patel AB, Crocker E, Eilers M, Hirshfeld A, Sheves M, Smith SO. 2004. Coupling of retinal isomerization to the activation of rhodopsin. *Proc Natl Acad Sci U S A*
215. Peersen O, Groesbeek M, Aimoto S, Smith SO. 1995. Analysis of Rotational Resonance Magnetization Exchange Curves from Crystalline Peptides. *J Am Chem Soc* 117: 7228–7237
216. Pepperberg DR, Morrison DF, O'Brien PJ. 1995. Depalmitoylation of rhodopsin with hydroxylamine. *Methods Enzymol* 250: 348

217. Peteanu LA, Schoenlein RW, Wang Q, Mathies RA, Shank CV. 1993. The first step in vision occurs in femtoseconds: complete blue and red spectral studies. *Proc Natl Acad Sci U S A* 90: 11762

218. Polans A, Baehr W, Palczewski K. 1996. Turned on by Ca^{2+}! The physiology and pathology of Ca^{2+}-binding proteins in the retina. *Trends Neurosci* 19: 547

219. Poo M, Cone RA. 1974. Lateral diffusion of rhodopsin in the photoreceptor membrane. *Nature* 247: 438

220. Rando RR. 1996. Polyenes and vision. *Chem Biol* 3: 255

221. Rao VR, Oprian DD. 1996. Activating mutations of rhodopsin and other G protein-coupled receptors. *Annu Rev Biophys Biomol Struct* 25: 287

222. Riek RP, Rigoutsos I, Novotny J, Graham RM. 2001. Non-alpha-helical elements modulate polytopic membrane protein architecture. *J Mol Biol* 306: 349

223. Rienstra CM, Hatcher ME, Mueller LJ, Sun B, Fesik SW, Griffin RG. 1998. Efficient Multispin Homonuclear Double-Quantum Recoupling for Magic-Angle Spinning NMR: ^{13}C–^{13}C Correlation Spectroscopy of U-^{13}C-Erythromycin A. *J Am Chem Soc* 120: 10602–10612

224. Rinia HA, Boots JW, Rijkers DT, Kik RA, Snel MM, et al. 2002. Domain formation in phosphatidylcholine bilayers containing transmembrane peptides: specific effects of flanking residues. *Biochemistry* 41: 2814

225. Rinia HA, Kik RA, Demel RA, Snel MM, Killian JA, et al. 2000. Visualization of highly ordered striated domains induced by transmembrane peptides in supported phosphatidylcholine bilayers. *Biochemistry* 39: 5852

226. Robinson PR, Cohen GB, Zhukovsky EA, Oprian DD. 1992. Constitutively active mutants of rhodopsin. *Neuron* 9: 719

227. Rohrig UF, Guidoni L, Rothlisberger U. 2002. Early steps of the intramolecular signal transduction in rhodopsin explored by molecular dynamics simulations. *Biochemistry* 41: 10799

228. Rothschild KJ, DeGrip WJ, Sanches R. 1980. Fourier transform infrared study of photoreceptor membrane. I. Group assignments based on rhodopsin delipidation and reconstitution. *Biochim Biophys Acta* 596: 338

229. Saam J, Tajkhorshid E, Hayashi S, Schulten K. 2002. Molecular dynamics investigation of primary photoinduced events in the activation of rhodopsin. *Biophys J* 83: 3097

230. Saibil H, Chabre M, Worcester D. 1976. Neutron diffraction studies of retinal rod outer segment membranes. *Nature* 262: 266

231. Sakmar TP. 1998. Rhodopsin: a prototypical G protein-coupled receptor. *Prog Nucleic Acid Res Mol Biol* 59: 1

232. Sakmar TP, Franke RR, Khorana HG. 1989. Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. *Proc Natl Acad Sci U S A* 86: 8309

233. Sakmar TP, Menon ST, Marin EP, Awad ES. 2002. Rhodopsin: insights from recent structural studies. *Annu Rev Biophys Biomol Struct* 31: 443

234. Salgado GF, Struts AV, Tanaka K, Fujioka N, Nakanishi K, Brown MF. 2004. Deuterium NMR Structure of Retinal in the Ground State of Rhodopsin. *Biochem.* 43: 12819–12828

235. Scheiner S, Duan X. 1991. Effect of intermolecular orientation upon proton transfer within a polarizable medium. *Biophys J* 60: 874

236. Schertler GF, Hargrave PA. 1995. Projection structure of frog rhodopsin in two crystal forms. *Proc Natl Acad Sci U S A* 92: 11578

237. Schertler GF, Villa C, Henderson R. 1993. Projection structure of rhodopsin. *Nature* 362: 770

238. Schoenlein RW, Peteanu LA, Mathies RA, Shank CV. 1991. The first step in vision: femtosecond isomerization of rhodopsin. *Science* 254: 412

239. Sheikh SP, Zvyaga TA, Lichtarge O, Sakmar TP, Bourne HR. 1996. Rhodopsin activation blocked by metal-ion-binding sites linking transmembrane helices C and F. *Nature* 383: 347

240. Shi L, Javitch JA. 2002. The binding site of aminergic G protein-coupled receptors: the transmembrane

segments and second extracellular loop. *Annu Rev Pharmacol Toxicol* 42: 437

241. Shichida Y, Imai H. 1998. Visual pigment: G-protein-coupled receptor for light signals. *Cell Mol Life Sci* 54: 1299
242. Shichida Y, Imai H. 1999. Amino acid residues controlling the properties and functions of rod and cone visual pigments. *Novartis Found Symp* 224: 142
243. Shichida Y, Tachibanaki S, Mizukami T, Imai H, Terakita A. 2000. Heterogeneity of rhodopsin intermediate state interacting with transducin. *Methods Enzymol* 315: 347
244. Shieh T, Han M, Sakmar TP, Smith SO. 1997. The steric trigger in rhodopsin activation. *J Mol Biol* 269: 373
245. Singh D, Hudson BS, Middleton C, Birge RR. 2001. Conformation and orientation of the retinyl chromophore in rhodopsin: a critical evaluation of recent NMR data on the basis of theoretical calculations results in a minimum energy structure consistent with all experimental data. *Biochemistry* 40: 4201
246. Smith HG, Jr., Stubbs GW, Litman BJ. 1975. The isolation and purification of osmotically intact discs from retinal rod outer segments. *Exp Eye Res* 20: 211
247. Smith SO, Aschheim K, Groesbeek M. 1996. Magic angle spinning NMR spectroscopy of membrane proteins. *Quat. Rev. Biophys.* 29: 395
248. Smith SO, Courtin J, de Groot H, Gebhard R, Lugtenburg J. 1991. ^{13}C magic-angle spinning NMR studies of bathorhodopsin, the primary photoproduct of rhodopsin. *Biochemistry* 30: 7409–15
249. Smith SO, de Groot H, Gebhard R, Lugtenburg J. 1992. Magic angle spinning NMR studies on the metarhodopsin II intermediate of bovine rhodopsin: evidence for an unprotonated Schiff base. *Photochem Photobiol* 56: 1035–1039
250. Smith SO, Palings I, Copie V, Raleigh DP, Courtin J, et al. 1987. Low-temperature solid-state ^{13}C NMR studies of the retinal chromophore in rhodopsin. *Biochemistry* 26: 1606–11
251. Smith SO, Palings I, Miley ME, Courtin J, de Groot H, et al. 1990. Solid-state NMR studies of the mechanism of the opsin shift in the visual pigment rhodopsin. *Biochemistry* 29: 8158
252. Smith SO, Peersen OB. 1992. Solid-State NMR Approaches for Studying Membrane Protein Structure. *Annual Review of Biophysics and Biomolecular Structure* 21: 25
253. Souto ML, Um J, Borhan B, Nakanishi K. 2000. *Helvetica Chemica Acta* 83: 2617–2628
254. Spooner PJ, Sharples JM, Goodall SC, Seedorf H, Verhoeven MA, et al. 2003. Conformational similarities in the beta-ionone ring region of the rhodopsin chromophore in its ground state and after photoactivation to the metarhodopsin-I intermediate. *Biochemistry* 42: 13371
255. Spooner PJ, Sharples JM, Verhoeven MA, Lugtenburg J, Glaubitz C, Watts A. 2002. Relative orientation between the beta-ionone ring and the polyene chain for the chromophore of rhodopsin in native membranes. *Biochemistry* 41: 7549
256. Stanasila L, Perez JB, Vogel H, Cotecchia S. 2003. Oligomerization of the alpha 1a- and alpha 1b-adrenergic receptor subtypes. Potential implications in receptor internalization. *J Biol Chem* 278: 40239
257. Steinberg G, Ottolenghi M, Sheves M. 1993. pKa of the protonated Schiff base of bovine rhodopsin. A study with artificial pigments. *Biophys J* 64: 1499
258. Stenkamp RE, Filipek S, Driessen CA, Teller DC, Palczewski K. 2002. Crystal structure of rhodopsin: a template for cone visual pigments and other G protein-coupled receptors. *Biochim Biophys Acta* 1565: 168
259. Strader CD, Sigal IS, Candelore MR, Rands E, Hill WS, Dixon RA. 1988. Conserved aspartic acid residues 79 and 113 of the beta-adrenergic receptor have different roles in receptor function. *J Biol Chem* 263: 10267
260. Stryer L. 1986. Cyclic GMP Cascade of Vision. *Annu. Rev. Neurosci.* 9: 87
261. Stryer L. 1996. Vision: from photon to perception. *Proc Natl Acad Sci U S A* 93: 557
262. Stubbs GW, Smith HG, Jr., Litman BJ. 1976. Alkyl glucosides as effective

solubilizing agents for bovine rhodopsin. A comparison with several commonly used detergents. *Biochim Biophys Acta* 426: 46

263. Sugihara M, Buss V, Entel P, Elstner M, Frauenheim T. 2002. 11-*cis*-retinal protonated Schiff base: influence of the protein environment on the geometry of the rhodopsin chromophore. *Biochemistry* 41: 15259
264. Sugihara M, Buss V, Entel P, Hafner J. 2004. The Nature of the Complex Counterion of the Chromophore in Rhodopsin. *J. Phys. Chem. B* 108: 3673–3680
265. Tajkhorshid E, Suhai S. 2000. The dielectric effect of the environment on the pKa of the retinal Schiff base and on the stabilization of the ion pair in bacteriorhodopsin. *J Mol Struct (THEOCHEM)* 501–502: 297–313
266. Takegoshi K, Nakamura K, Terao T. 2001. ^{13}C–^{1}H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem Phys Lett* 344: 631–637
267. Tallent JR, Hyde EW, Findsen LA, Fox GC, Birge RR. 1992. Molecular dynamics study of primary photochemical event in rhodopsin. *J Am Chem Soc* 114: 1581–1592
268. Tan Q, Lou J, Borhan B, Karnaukhova E, Berova N, Nakanishi K. 1997. Absolute Sense of Twist of the C12-CI3 Bond of the Retinal Chromophore in Bovine Rhodopsin Based on Exciton-Coupled CD Spectra of 11,12-DihydroretinaI Analogues. *Angew. Chem. Int. Ed. Engl.* 36: 2089–2093
269. Teller DC, Okada T, Behnke CA, Palczewski K, Stenkamp RE. 2001. Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). *Biochemistry* 40: 7761
270. Terakita A, Yamashita T, Shichida Y. 2000. Highly conserved glutamic acid in the extracellular IV-V loop in rhodopsins acts as the counterion in retinochrome, a member of the rhodopsin family. *Proc Natl Acad Sci U S A* 97: 14263
271. Thorgeirsson TE, Lewis JW, Wallace-Williams SE, Kliger DS. 1992. Photolysis of rhodopsin results in deprotonation of its retinal Schiff's base prior to formation of metarhodopsin II. *Photochem Photobiol* 56: 1135
272. Traxler KW, Dewey TG. 1994. Effects of depalmitoylation on physicochemical properties of rhodopsin. *Biochemistry* 33: 1718
273. Trettel F, Di Bartolomeo S, Lauro C, Catalano M, Ciotti MT, Limatola C. 2003. Ligand-independent CXCR2 dimerization. *J Biol Chem* 278: 40980
274. van der Steen R, Biesheuvel PL, Mathies RA, Lugtenburg J. 1986. Retinal Analogues with Locked 6-7 Conformations Show That Bacteriorhodopsin Requires the 6-s-*trans* Conformation of the Chromophore. *J Am Chem Soc* 108: 6410–6411
275. Veit M, Sachs K, Heckelmann M, Maretzki D, Hofmann KP, Schmidt MF. 1998. Palmitoylation of rhodopsin with S-protein acyltransferase: enzyme catalyzed reaction versus autocatalytic acylation. *Biochim Biophys Acta* 1394: 90
276. Verdegem PJ, Bovee-Geurts PH, de Grip WJ, Lugtenburg J, de Groot HJ. 1999. Retinylidene ligand structure in bovine rhodopsin, metarhodopsin-I, and 10-methylrhodopsin from internuclear distance measurements using ^{13}C-labeling and 1-D rotational resonance MAS NMR. *Biochemistry* 38: 11316
277. Verhoeven MA, Creemers AF, Bovee-Geurts PH, De Grip WJ, Lugtenburg J, de Groot HJ. 2001. Ultra-high-field MAS NMR assay of a multispin labeled ligand bound to its G-protein receptor target in the natural membrane environment: electronic structure of the retinylidene chromophore in rhodopsin. *Biochemistry* 40: 3282
278. Vogel R, Siebert F. 2001. Conformations of the active and inactive states of opsin. *J Biol Chem* 276: 38487
279. Vogel R, Siebert F. 2003. Fourier transform IR spectroscopy study for new insights into molecular properties and activation mechanisms of visual pigment rhodopsin. *Biopolymers* 72: 133
280. Vogel R, Siebert F, Mathias G, Tavan P, Fan G, Sheves M. 2003. Deactivation of rhodopsin in the transition from the signaling state meta

II to meta III involves a thermal isomerization of the retinal chromophore C[double bond]D. *Biochemistry* 42: 9863

281. Vreven T, Bernardi F, Garavelli M, Olivucci M, Robb MA, Schlegel HB. 1997. Ab initio photoisomerization dynamics of a simple retinal chromophore model. *J Am Chem Soc* 119: 12687–12688
282. Wald G. 1968. The molecular basis of visual excitation. *Nature* 219: 800
283. Wang Q, Schoenlein RW, Peteanu LA, Mathies RA, Shank CV. 1994. Vibrationally coherent photochemistry in the femtosecond primary event of vision. *Science* 266: 422
284. Wang Z, Asenjo AB, Oprian DD. 1993. Identification of the Cl^--binding site in the human red and green color vision pigments. *Biochemistry* 32: 2125
285. Warshel A. 1976. Bicycle-pedal model for the first step in the vision process. *Nature* 260: 679
286. Warshel A, Barboy N. 1982. Energy storage and reaction pathway in the first step of the vision process. *J Am Chem Soc* 104: 1457
287. Warshel A, Parson WW. 2001. Dynamics of biochemical and biophysical reactions: insight from computer simulations. *Q Rev Biophys* 34: 563
288. Warshel A, Russell ST. 1984. Calculations of electrostatic interactions in biological systems and in solutions. *Q Rev Biophys* 17: 283
289. Yamada A, Kakitani T, Yamamoto S, Yamamoto T. 2002. A computational study on the stability of the protonated Schiff base of retinal in rhodopsin. *Chemical Physics Letters* 366: 670–675
290. Yamamoto S, Wasada H, Kakitani T. 1999. Ab initio MO study on the potential energy surfaces for twisting around the C11=C12 bond of the protonated Schiff base of retinal. *Theochem* 461-462: 463–471
291. Yan EC, Kazmi MA, De S, Chang BS, Seibert C, et al. 2002. Function of extracellular loop 2 in rhodopsin: glutamic acid 181 modulates stability and absorption wavelength of metarhodopsin II. *Biochemistry* 41: 3620
292. Yan EC, Kazmi MA, Ganim Z, Hou JM, Pan D, et al. 2003. Retinal counterion switch in the photoactivation of the G protein-coupled receptor rhodopsin. *Proc Natl Acad Sci U S A* 100: 9262
293. Yeager MJ. 1976. Neutron diffraction analysis of the structure of retinal photoreceptor membranes and rhodopsin. *Brookhaven Symp Biol*: III3
294. Yeagle PL, Albert AD. 1998. Structure of the G-protein-coupled receptor, rhodopsin: a domain approach. *Biochem Soc Trans* 26: 520
295. Yeagle PL, Albert AD. 2002. Use of nuclear magnetic resonance to study the three-dimensional structure of rhodopsin. *Methods Enzymol* 343: 223
296. Yeagle PL, Alderfer JL, Albert AD. 1997. Three-Dimensional Structure of the Cytoplasmic Face of the G Protein Receptor Rhodopsin. *Biochemistry* 36: 9649
297. Yeagle PL, Alderfer JL, Salloum AC, Ali L, Albert AD. 1997. The first and second cytoplasmic loops of the G-protein receptor, rhodopsin, independently form beta-turns. *Biochemistry* 36: 3864
298. Yeagle PL, Alderfer JL, Salloum AC, Ali L, Albert AD. 1997. The First and Second Cytoplasmic Loops of the G-Protein Receptor, Rhodopsin, Independently Form Turns. *Biochemistry* 36: 3864
299. Yeagle PL, Danis C, Choi G, Alderfer JL, Albert AD. 2000. Three dimensional structure of the seventh transmembrane helical domain of the G-protein receptor, rhodopsin. *Mol Vis* 6: 125
300. Yokoyama S. 2002. Molecular evolution of color vision in vertebrates. *Gene* 300: 69
301. Yoshizawa T, Horiuchi S. 1973. *Biochemistry and Physiology of Visual Pigments*. New York: Springer-Verlag
302. Yoshizawa T, Wald G. 1963. Pre-lumirhodopsin and the bleaching of visual pigments. *Nature* 197: 1279
303. Zhang HZ, Lerro KA, Yamamoto T, Lien TH, Sastry L, et al. 1994. The localization of the chromophore in rhodopsin: a photoaffinity study. *J Am Chem Soc* 116: 10165–173
304. Zhukovsky EA, Oprian DD. 1989. Effect of carboxylic acid side chains

on the absorption maximum of visual pigments. *Science* 246: 928

305. Zhukovsky EA, Robinson PR, Oprian DD. 1991. Transducin activation by rhodopsin without a covalent bond to the 11-*cis*-retinal chromophore. *Science* 251: 558

306. Zhukovsky EA, Robinson PR, Oprian DD. 1992. Changing the location of the Schiff base counterion in rhodopsin. *Biochemistry* 31: 10400

307. Zvyaga TA, Fahmy K, Sakmar TP. 1994. Characterization of rhodopsin-transducin interaction: a mutant rhodopsin photoproduct with a protonated Schiff base activates transducin. *Biochemistry* 33: 9753

II.2 Molecular Dynamic Simulations Of Retinal In Rhodopsin: From The Dark-Adapted State Towards Lumirhodopsin

Vincent Lemaître,[1,2] Philip Yeagle,[3] and Anthony Watts,[1]

[1] *Biomembrane Structure Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK ; email: anthony.watts@bioch.ox.ac.uk*

[2] *Nestec S.A., BioAnalytical Department, Vers-Chez-Les-Blanc, CH-1000 Lausanne 26, Switzerland*

[3] *Biology Department, University of Connecticut, 91 North Eagleville Road, USA.*

Key words retinal, rhodopsin, photo-intermediate, photorhodopsin, bathorhodopsin, lumirhodopsin, vision, phototransduction

■ **Abstract** This study aims at linking the formation of photo-intermediates and the conformational changes observed in the chromophore of rhodopsin during the early steps of the protein activation. The lysine-bound retinal in rhodopsin has been studied by molecular dynamic simulation, focusing on its conformation in the dark adapted-state (10ns) and in the early steps after the isomerization of the 11-*cis* bond to *trans* (up to 10ns). The parametrization for the chromophore is based on a previous quantum study [Sugihara M, Buss V, Entel P, Elstner M, Frauenheim T. 2002. Biochemistry 41: 15259] and shows good conformational agreement with the most recent experimental contributions. The isomerization, induced by switching the function governing the dihedral angle for the C11=C12 bond, was repeated with slightly different starting conformations. From the repeated simulations of the isomerization, it can be shown that the retinal-model undergoes a conserved activation pattern. The conformational changes are sequential, spreading away from the C11=C12 bond: isomerization of the C11=C12 bond, rotation of methyl group C20, increased fluctuations at the β-ionone ring. The dynamic of these changes suggests they are linked with photo-intermediates observed by spectroscopy. The exact moment when these events take place after the isomerization, is modulated by the starting conformation, suggesting retinal isomerizes through multiple activation paths that are slightly different. Interestingly, the amplitudes of the structural fluctuations observed for the protein in the dark-adapted state and after the

isomerization are similar, suggesting a subtle mechanism for the transmission of the information from the chromophore to the protein.

INTRODUCTION

The rhodopsin-bound retinal is a molecular device for converting light into molecular conformational changes. The irradiation of rhodopsin isomerizes the 11-*cis*-retinal to all-*trans*, triggering a chain of conformational changes in the opsin that induces an enzymatic cascade leading to vision (42, 56, 75). The conformation of retinal and its evolution resulting from light-absorption are crucial to an understanding of the mechanism of activation of rhodopsin, and this has been the subject of intense research (6, 26, 57). Rhodopsin is the most intensively studied member of the G protein-coupled receptor (GPCR), mainly because it is the only GPCR that is naturally present in high abundance in biological tissues (1).

Here, we present computational studies of bovine rhodopsin inserted in an explicit membrane environment (Fig. 1). We address structural and dynamical properties of the chromophore in the protein, for both the dark-adapted state and the early stage after the induction of the retinal isomerization. Specifically, we endeavour to refine the mechanism for signal propagation within the chromophore and its transmission to the protein.

The experimental data relevant for this study are either related to the dynamic or the conformation of retinal bound to bovine rhodopsin in the dark-adapted state or occurring before the lumirhodopsin state, photo-intermediate which MD simulation cannot be reached due to the time required and current computational power limitations.

Rhodopsin provides an environment in which its 11-*cis*-retinal chromophore can undergo a *cis* to *trans* conformational switch in response to absorption of a photon with a very high quantum yield of 0.67 (5). When light strikes the retina and is absorbed by the photo-pigment, the resulting structural changes induced by the retinal upon photo-activation lead to a series of defined intermediates and are referred as the photo-cycle of the protein (reviewed in (42, 60)) Low-

temperature and time-resolved UV/VIS spectroscopic measurements have shown that photo-bleaching of rhodopsin involves several intermediates and most likely has more than one pathway (32, 34, 48).

The primary photoproduct, photorhodopsin, is formed within a very short time (200 fs). This state is very transient and cannot be isolated (31, 59). A model for the potential energy surface of the excited state, based on UV-VIS femto-spectroscopy data, shows that the torsion angle of the bound C11=C12 is changed by 75° wihin 30 fs, while the full isomerization takes place in 200 fs (40). The initial movements of the chromophore are thought to be tightly constrained by the surrounding protein, due to the very short time scale of the photo-isomerization (61).

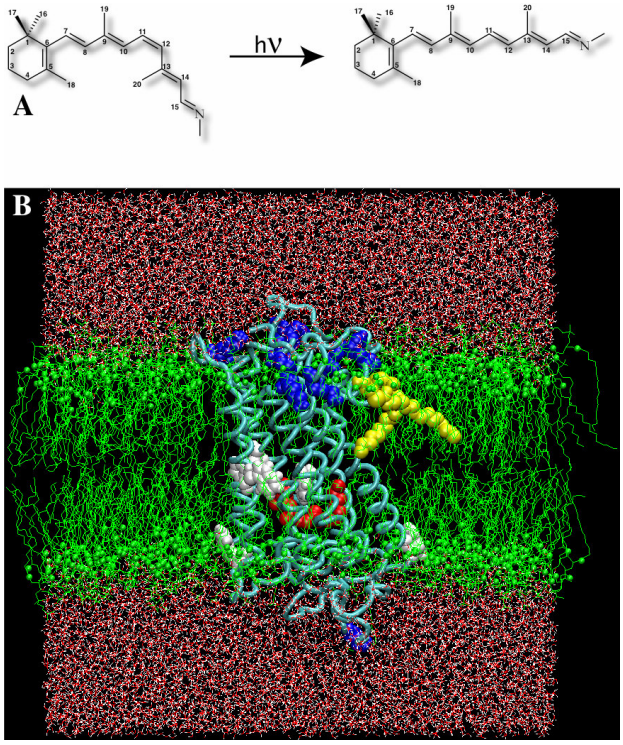


Figure 1: (**A**) Chemical structure of 11-*cis*-retinal and all-*trans*-retinal. (**B**) Illustration of the simulation box (blue: lysine, green: POPC, red: lysine-bound retinal, white: tryptophan, yellow: palmitoylated-cystein, red&white: water).

Subsequently, photorhodopsin thermally relaxes within a few picoseconds to a distorted *all-trans* configuration, bathorhodopsin (47). In bathorhodopsin, the chromophore is expected to carry the energy transmitted by the photon, not having had the time to relax. Models for the energy storage in bathorhodopsin-retinal models suggest that absorbed energy can be stored via (i) structural changes

in the retinal chromophore itself, (ii) alterations in the interactions between the retinal and its protein environment, or (iii) a combination of both. Many studies, including calorimetric, NMR, infrared, and Raman data for bathodopsin studied at low temperature, suggest that the distortion in the conformation of the chromophore account for an important part in the energy storage (2, 11, 16, 17, 19, 25, 47, 62). An early NMR study on bathorhodopsin suggest the energy stored in the primary photoproduct does not raise any substantial changes in the average electron density of the polyene chain, and is similar to dark-adapted state and isorhodopsin (62), where the largest perturbation is observed for C13. Vibrational spectroscopy studies at low and room temperature confirm that the vibrational modes in the Schiff base region of the retinal chromophore in the dark-adapted state of rhodopsin remain the same upon bathorhodopsin formation (17, 52). The charge separation between the Schiff base and its counter-ion, therefore, does not make a major contribution to this part of the energy storage/transduction mechanism. Furthermore, significant changes between the dark-adapted state and bathorhodopsin occur in the polyene chain, indicating substantial twisting of the retinal backbone (17, 19, 39, 47, 52). These conclusions were drawn for bathorhodopsin at low temperature (T< 100K) or at room temperature (17, 52).

On a nanosecond time scale (~120 ns (30)), bathorhodopsin establishes an equilibrium with a blue-shifted intermediate (BSI) before the mixture decays to form lumirhodopsin (in ~150 ns (30). The Blue-Shifted Intermediate (BSI) is not observed under cryogenic conditions, but rather observed by flash photolysis at ambient temperature (28) The equilibrium between bathorhodopsin and BSI is independent of the protein environment, but can be influenced by chemical modification of the retinal, suggesting that the bathorhodopsin to BSI only involves conformational changes of the chromophore (34). The proposed barrier for the formation of BSI is the steric interaction between H8 (hydrogen atom on carbon C8 of the retinal) and the C5-methyl group of retinal (36, 37).

The decay from BSI to reach lumirhodopsin in the nanosecond timescale can be significantly affected by the amino composition of the binding-pocket, but not by the chromophore (34, 53). Therefore it seems this step involves the adjustment of the protein matrix to the isomerised chromophore (34). In the lumirhodopsin state, the salt bridge between the protonated Schiff base and Glu113 is broken and retinal is believed to be fully relaxed (48). Breaking the salt bridge

decreases the pKa of the protonated Schiff base, facilitating the transfer of the proton to Glu113 at a later stage (10, 35).

Lumirhodopsin is then transformed into metarhodopsin I and subsequently metarhodopsin II, where important conformational changes are taking place in the protein and the active conformation for G-protein coupling reached (34). The ^{13}C NMR chemical shifts provide convincing evidence that Schiff base deprotonation occurs in the transition between metarhodopsin I and metarhodopsin II (63), as suggested by the pH dependence of UV/VIS spectrum observed for the metarhodopsin I state (77).

Rhodopsin is also the only GPCRs for which accurate structural information is available. Several crystal structures are available (44, 46), even indicating the position of hydration water within the protein for the most recent ones (38, 43, 45). The availability of this structural information has allowed the various theoretical studies to take place, including Molecular Dynamic simulation (MD) (14, 27, 54, 55) and Quantum simulations on fragment of the protein (using Density Function Theory for example) (8, 45, 68, 69, 76).

Whereas the crystallographic data reveal a detailed description of the receptor in the inactive dark state and until recently low-resolution information on the chromophore, solid state NMR was able to provide accurate structural and electronic information on the chromophore, either in the dark-adapted state or in one of its photo-intermediate states. Rotational resonance has been used successfully for providing internuclear distances or restraints in the dark-adapted state, metarhodopsin I and II (49, 66, 67, 73). The validity of the method was confirmed for the dark-adapted state with the improvement in the x-ray crystal structure (45). Furthermore, the analysis of the chemical shift at various photo-activated states and their comparison with simplify Schiff-base model or rhodopsin-bound retinal in the dark state can be used to probe the electronic environment around the isotopic labels selectively inserted in the chromophore (12, 41, 62-65, 74).

The information gathered by vibrational spectroscopy provide information on the chromophore difficult to translate into an accurate molecular description, i.e. meaning interatomic distances, while structural approaches like x-ray and NMR provide a few accurate snapshots on the molecular

conformation of the chromophore in various states of the protein activation pattern. Molecular dynamic simulation is a tool with the potential of linking these two important experimental contributions, at least on a short timescale (~10ns), with the ability to provide accurate models, which can then be tested by designing adequate experiences.

The results presented in this contribution provide a continuation of the previous MD studies on retinal in rhodopsin (54, 55) in the sense that the same method is applied, but also some refinement in the chromophore model. The conformation and the molecular description for the chromophore are based on a quantum study which describes most of the structural information available for the dark-adapted state (68). The experimental data regarding the conformation of retinal also were only made available recently. Furthermore, the protein was modelled from a crystal structure with a resolution of 2.6 Å (1L9H) including functionally important water molecules and is simulated in an explicit lipid bilayer where the minimum distance between the protein and its mirror images is at least 25 Å.

Anticipating our results, we find that although several slightly different routes for possible structural relaxations are available, all of them lead to similar conformational changes and that the order, in which they take place define a fixed pattern. Immediately after isomerization the chromophore is forced into a highly strained configuration where the deformation is localized to the isomerized bond, the other part of retinal being unchanged. The deformation is then transmitted further away to methyl group C20, which undergo a large amplitude rotation, finally pointing towards the extracellular loops. Finally the perturbation reaches the ring, where increased vibrational fluctuations are observed upon isomerization of the 11-*cis* bond. In the combined 30ns of MD simulation, the ring flipping or moving away from its aromatic binding-pocket, was not observed. Within a multi-nanosecond simulation, the structurally distorted retinal relaxes towards a more planar geometry. The geometric strain is mainly released through a swinging motion of a fraction of the polyene tail, resulting in the rotation of methyl group C20. The protonated Schiff-base interaction with Glu113 maintains this side of the chromophore rather fixed. Therefore, the perturbation generated upon isomerization of the 11-*cis* bond seems to be transferred from the

retinal to the protein by a subtle mechanism mediated the β-ionone ring, where its vibrational activity rather than its location is being perturbed.

MATERIAL AND METHODS

Molecular Dynamics Calculations

MD simulations are performed with the GROMACS v3.1.5 package, using gromos43A2, extended to improve the simulation the lipid components for the force field (4). Simulations are run at a temperature of 300 K and a pressure of 1 bar in an isothermal-isobaric ensemble (NPT) with periodic boundaries present. Both a Berendsen temperature and pressure couplers are chosen to keep these parameters constant. The time step for the simulations is 1 fs in the case of the short 5 ps MD simulations and 2 fs for the others. A LINCS algorithm is used to maintain the geometry of the molecules. Long-range electrostatic interactions are calculated with the particle-mesh Ewald (PME) method. PME tends to slow down the computation but increase its quality since it removes any cut-off electrostatic interactions. Lennard-Jones interactions are cut off at 14 Å. The single point charge (SPC) water model (3) is used to describe the water in the simulation box.

Lipids

The bilayer consists in a POPC patch including initially 288 lipids and 16337 water molecules. The topology file for the POPC molecule has been previously described and is available from http://moose.bio.ucalgary.ca/Downloads.

Protein

The crystal structure of the dark-adapted state rhodopsin (1L9H from PDB) was used as a starting point for building the molecular model of the protein. Missing residues in the intracellular loops have been added using Modeller (58). The simulation of rhodopsin involves two functionalized amino acid residues, palmytoylated-cysteine and retinal-bound lysine, for which it was necessary to write a molecular description. The topologies for these residues were optimized and tested in

preliminary runs. The parameters used to simulate the palmytoyl chain are based on the parameters used to simulate the lipids.

Retinal

Retinal is carefully parametrized in order to reproduce experimental data previously gathered for retinal in the dark-state (7, 9, 67, 72) in the gromos43a2 force field. The parametrization is based on a previous DFT theoretical study on 6-*s-cis* retinal (68), confirmed by a recent 2.2 Å resolution crystal structure (45). The bond-lengths, bond angles and dihedral angles from this model have been used as parameters for the retinal topology. The partial charge on the nuclei were approximated based on the Mulliken charges of the DFT model. The 6-*s-cis* conformer proposed by Sugihara et al. was used as the starting conformation for retinal in rhodopsin and used to replace the chromophore present in the 1L9H crystal structure.

Insertion of protein into bilayer and equilibration

The protein is inserted in the bilayers using a method aiming to generate a suitable cavity in the interior of the lipid-bilayer (21), using the solvent-accessible surface of the protein as a template, during the course of a short steered molecular dynamics simulation (SMD) of a solvated lipid membrane (500 ps SMD run). The protein-lipid equilibration is achieved through a three-stage process. First, overlapping water molecules are removed, then lipid molecules whose headgroups are located within 1.5Å from the protein surface are removed; the protein-lipid interface is optimized by applying repulsive forces perpendicular to the protein surface, to the remaining lipid atoms inside the volume occupied by the protein surface until it is emptied. The protein itself may then be inserted into the bilayer. Counter-ions are added. The system is being energy-minimized at each step involving the addition or the removal of any molecular species (steepest descent algorithm). Finally, the system is being equilibrated in successive short MD simulations (5 runs of 200 ps each) were positions restraints are applied on the protein and progressively decreased.

MD runs

One 10 ns long MD simulation was run for the dark-adapted state. The isomerization of the chromophore was studied in a series of six MD simulations: Three 5ps-simulations with a 1fs time

step where frames were collected for every time step of the calculation, aiming at following the first picoseconds of the isomerization and three 10ns-simulations with a 2fs time step, aiming at following the interactions between retinal and its binding-pocket.

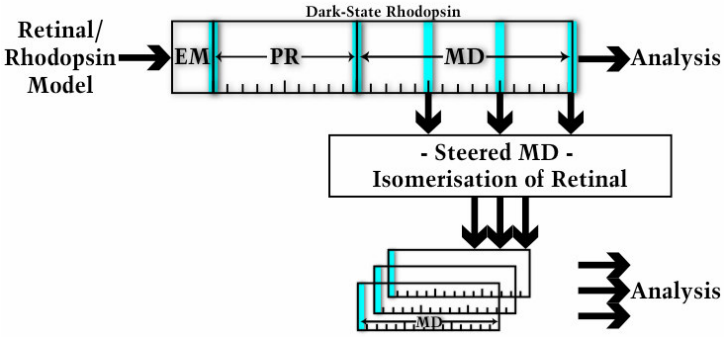


Figure 2: Strategy to study the isomerization of retinal. A 10ns-MD of bovine rhodopsin is run to generate frames used as starting conformation for the isomerization (including starting velocities). Frames after 500ps, 1000ps and 1500ps were used as seeds for SMD simulations where retinal is isomerized from 11-*cis* to all-*trans*. The interpretation is based on the comparison between the SMDs where retinal is isomerized with the MD of bovine rhodopsin in its dark-adapted state.

Photo-activation of retinal using steered MD techniques previously described (54, 55), where the dihedral parameters for the 11-*cis* bond are switched to describe a *trans* double bond. The same technique was used to induce the isomerization of the chromophore.

Computers

MD simulations were run on various computers including a PowerBook G4 running under Mac OS X (1.25 GHz 512MB RAM) for the shortest MD and the data analysis, or an 8 processors Beowulf-class cluster made of 4 dual processor X-serve G4 rack (2 x 1.25 GHz, 1 GB RAM) units with GHz switches and UPS, running under a Mac OS X (10.2.8). MD calculations were run using simultaneously several processors (2 or 4) with the MPI middleware distributing the processes over all units.

Data Analysis

For the analysis (tilt, kink angle, etc.), only the core residues of the TM helices were included: (H1) 35-64, (H2) 72-98, (H3) 109-138, (H4) 151-172, (H5) 201-225, (H6) 245-277, (H7) 286-308. The retinal binding pocket was defined as the amino acids present within 4.5 Å from the retinal as described by Palczewski et al. (46). All molecular structures were drawn using VMD (29).

RESULTS AND DISCUSSION

The MD simulations presented here have two goals. First present a MD model for retinal based on quantum simulations in close agreement with the structural data available and in a second time extrapolate this model to study the early stage of the isomerization of retinal. The 10ns-MD simulation of bovine rhodopsin in the dark-state shows a distorted retinal in a relatively rigid binding-pocket. Most of the motions taking place in the protein are located in the intra-cellular loops. The isomerization takes place in a 100fs-scale and is followed by several conformational changes within retinal at the picosecond and nanosecond time scale. At a 10ns-timescale, most of the changes deriving from the isomerization take place in the retinal itself. The comparison of the motion of the protein in general and the motion of the residues involved in the binding-pocket on a 10ns-timescale does not show large-amplitude conformational changes that can be directly linked with the isomerization of retinal and interpreted as the protein response to the isomerization. Large amplitude motions of the intracellular loops happen both in the dark-adapted state and after the isomerization. Since this study focused on retinal and its binding pocket, the nature of these motions was not investigated further.

Rhodopsin in the dark-adapted state

During the 10ns-MD in the dark adapted state, the RMSD of rhodopsin increases quickly to reach an average value of about 2 Å, around which it oscillate during most of the simulation (see Fig. 3). The RMSD computed for the residues from the binding pocket show a different behaviour, close to 1.5 Å. The comparison between the two RMSD curves indicates that most of the motion in the protein takes place in the loops. This is confirmed by the RMSF computed for the protein (Fig.3),

where the highest values are obtained for the intracellular loop and the C-terminal domain. This behaviour has been confirmed experimentally by x-ray crystallography where it is well-known that this part of the protein is the most mobile and was actually missing in the earliest crystal structures (38, 46, 71).

Modelling 11-*cis*-retinal in the dark-adapted state

The retinal model is parametrized based on a theoretical contribution where the conformation of 11-*cis* retinal in the binding pocket was determined using DFT for the value at equilibrium of all parameters and using standard values from the force field (68). The dihedral interactions for all the bonds in the polyene chain were set with the standard values used for double-bond in order to give the model the rigidity required to reproduce accurately experimentally derived information (i.e. interatomic distances measured by solid state NMR (67)). This parametrization leads to a model for the chromophore that reproduces reasonably well the structural information available.

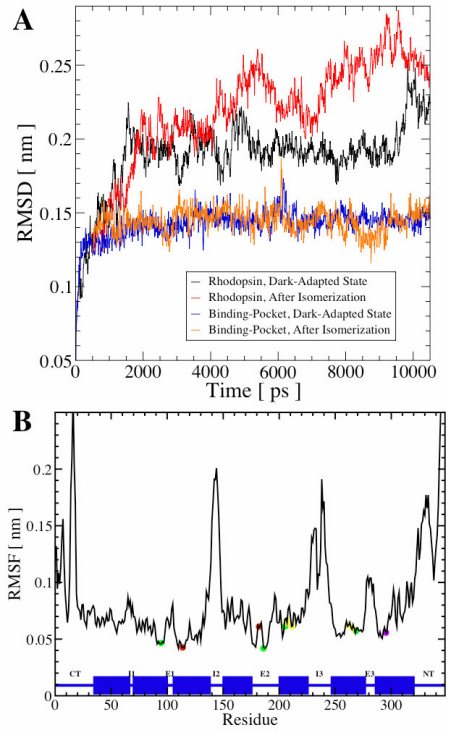


Figure 3: (**A**) Backbone RMSD of bovine rhodopsin in the dark adapted state (black) and after isomerization of the retinal (red, starting point = dark-adapted state trajectory, after 500ps). RMSD curves of the residues belonging to the binding-pocket are also displayed for the dark-adapted state (blue) and after isomerization (orange). (**B**) Root-Mean Square Fluctuation of rhodopsin in the dark-adapted state. The helices and several residues from the binding-pocket are highlighted (purple: retinal-bound lysine, red: glutamic acid, green: serine or methionine, yellow: phenylalanine or

tryptophan). The secondary structure is highlighted in blue at the bottom to highlight the position of the transmembrane helices.

The RMSD computed for the lysine-bound retinal (Figure 4A, black) fluctuates around 0.8 Å, even lower than the average RMSD for the binding pocket (Figure 3A). RMSD computed for various fragment shows that most of the deviation from the structure of reference (retinal conformer derived by DFT (68)) takes place in the polyene chain (Figure 4A, blue), the deformation of the ring geometry less affected, although it is the part of the molecules which fluctuates the more with the methyl groups (RMSF, Figure 5, black).

Rotational resonance solid-state NMR was used for distance determinations within the rhodopsin-bound chromophore in membrane for the dark-adapted state and metarhodopsin I (67, 72) (Table 1 and 2). From these distances, it was shown that the C10–C13 unit is conformationally twisted (72). Rotational resonance solid-state NMR was also used to provide distance restraints regarding the relative orientation between the β-ionone ring and the polyene chain of the chromophore in rhodopsin (67). The distances measured between C8–C16, C8–C17 and C8–C18, show that the major portion of retinylidene in rhodopsin has a twisted 6-s-*cis* conformation (67).

Table 1: Distance restraints within retinal for rhodopsin in the dark-adapted state determined by MD at 300K, DFT and ^{13}C–^{13}C rotational resonance NMR at 210 K.

Atom Pair	C10–C20	C11–C20	C8–C16	C8–C17	C8–C18
MD-Distance [Å]	3.09 ± 0.14 *	3.06 ± 0.01 *	4.42 ± 0.10 *	3.88 ± 0.17 *	3.08 ± 0.13 *
DFT-Distance [Å]	3.07	3.09	4.35	3.98	3.00
Ref.	(68)	(68)	(68)	(68)	(68)
NMR-Distance [Å]	3.04 ± 0.15 #	2.93 ± 0.15 #	4.05 ± 0.25 #	4.05 ± 0.25 #	2.95 ± 0.25 #
Ref.	(72)	(72)	(67)	(67)	(67)

* = Standard deviation of the fuctuations during the MD : # = Experimental error.

Table 2: Distance restraints within retinal for rhodopsin determined by MD within 10ns after the isomerization of retinal at 300K, and ^{13}C–^{13}C rotational resonance NMR for metarhodopsin I at 210 K.

Atom Pair	C10–C20	C11–C20	C8–C16	C8–C17	C8–C18
MD-Distance [Å]	4.44 ± 0.14 * 4.44 ± 0.11 * 4.48 ± 0.11 * (4.45 ± 0.13)	3.05 ± 0.09 * 3.04 ± 0.09 * 3.08 ± 0.09 * (3.06 ± 0.09)	4.39 ± 0.11 * 4.39 ± 0.10 * 4.42 ± 0.12 * (4.40 ± 0.11)	3.96 ± 0.17 * 3.98 ± 0.17 * 3.90 ± 0.16 * (3.95 ± 0.17)	3.32 ± 0.17 * 3.15 ± 0.15 * 3.41 ± 0.13 * (3.29 ± 0.15)
NMR-Distance [Å]	4.35 ± 0.15 #	2.83 ± 0.15 #	N.A.	4.05 ± 0.25 #	2.95 ± 0.25 #
Ref.	(72)	(72)		(66)	(66)

- = Standard deviation of the fuctuations during the 10ns-MD : # = Experimental error.

One of the goals of this study is to implement an accurate retinal model to be used for MD simulations of bovine rhodopsin in the dark-adapted state. For this purpose, the distance-restraints defined by solid state NMR constitute an important block of available experimentally-derived structural information. Table 1 include the averages for the corresponding distances as computed from our theoretical model, while Figure 10A displays the instantaneous distance as a function of the simulation time. The comparison between shows a fairly good agreement between the model and the experiments, with only one distance out by 0.12 Å from the error range of the NMR estimation. Table 3 and Figure 6 aimed at the same comparison for selected dihedral angles, showing a decent agreement between the model and NMR-derived distances (note that the temperature between experiment and simulation differs by 70 K). The averaged-distances are also fluctuating close to the values defined by the DFT-derived 6-s-*cis* conformer (68).

While setting-up a model for retinal aims at describing accurately as possible a crucial fragment of rhodopsin by MD, it also allows investigating more accurately the dynamic of the chromophore in its binding-pocket and can be apply for the further investigation of rhodopsin-mutants in the dark-adapted state.

Table 3: Dihedral angles for selected bonds for retinal in the dark-adapted state, determined by MD, DFT and NMR at 210K.

Dihedral Angle	MD Average	DFT	NMR
C5-C6–C7-C8	-39 ± 9° *	-35°	-28 ± 7° #, • (67)
C10-C11–C12-C13	-20 ± 11° *	-12°	± 44 ± 10° #, • (72)
C8-C9–C10-C11	166 ± 11° *	175°	160 ± 10° # (22)
C13-C14–C15-NZ	180 ± 15° *	173°	165 ± 5° # (15)

* = Standard deviation of the fluctuations during the MD : # = Experimental error ; • = NMR-based estimation are based on the 11-cis-retinal crystal structure for which the bond angle was rotated in order to match ^{13}C-^{13}C rotational resonance distances.

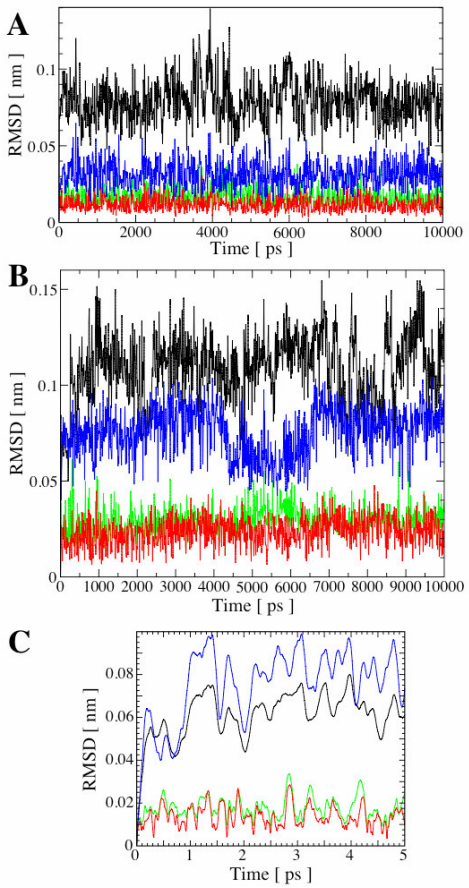


Figure 4: RMSD for the lysine-bound retinal (black) and fragments of the retinal – β-ionone ring (red), β-ionone ring + C7–C8 (green) and polyene tail (blue). (A) Dark-adapted state, (B) 10ns after isomerization of retinal after 500ps MD in Dark adapted state, (C) First 5ps after isomerization of retinal, after 500ps MD in Dark-adapted state.

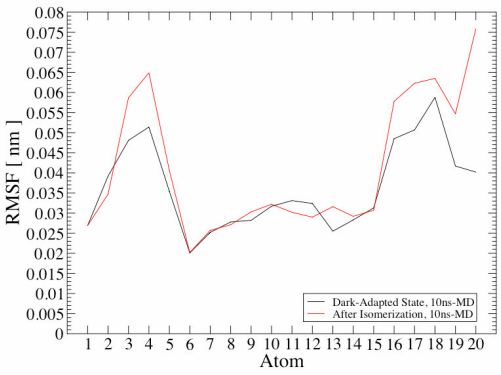


Figure 5: Root-Mean Square Fluctuation (RMSF) for retinal, illustrating the increased vibrations at the level of selected parts of retinal after isomerization. Atoms 3 to 5 from the β-ionone ring (in

contact with Helix V) and all methyl groups (atoms 16 to 20) show a visible increase of their corresponding RMSF upon isomerization.

Various parameters are useful to characterize the dynamic of the retinal model in the dark-adapted state. The RMSF is one of them (Figure 5), highlighting amongst other things that the methyl groups are the most mobile parts of the chromophore, and that both C11 and C12 are the atoms within the polyene chain fluctuating the more around their average location. The standard deviations of various averaged interatomic distances and dihedral angles are also good indicators. In the dark-state, the standard deviation between non-directly bonded atoms was below 0.2 Å. At the same time dihedral angles exhibit standard deviation around 10°, which can affect significantly the local geometry of the molecules. This model sets the picture of retinal as a rather twisted and flexible molecule, while maintaining some interatomic distances relatively constant (fluctuation are less than 5% of the average value).

Transition from 11-*cis* to all-*trans*-retinal

The isomerization of retinal has been studied in two groups of short 5ps-MDs and long 10ns-MDs, each containing three simulations starting from a different starting point. The starting point for each of the MDs is taken from the 10ns-MD of rhodopsin in the dark adapted-state, respectively at time 500ps, 1000ps and 1500ps (Figure 2 for an illustration).

The choice of two time-scales results from a technical compromise: in order to follow accurately the isomerization of retinal at the femto- to picosecond level, the output of a large number of frame and a very short stimestep is required (time-step 1 fs, frame saved every calculation step). This approach is not practicable for larger simulations, the space for storing the data and the RAM memory required for the analysis being too prohibitive. As a result, these MDs use a slightly different set-up (time-step is 2 fs, frame saved every 500 calculation steps) to explore longer simulation-time at the expense of a lower time-resolution.

The isomerization is obtained by switching the potential function describing the dihedral interaction for the 11-*cis* bond. This approach has been described in details in previous reports (24, 55). The force field used can be expected to properly include the essential steric and electrostatic effects. However, it has some limitations. One of them is the use of united pseudo-atoms for aliphatic apolar

hydrogen atoms, which are not explicitly described. A consequence of this approach is to squash the distribution of partial charges on the atoms. As previous studies have shown that the changes in the partial charges upon isomerization are rather small, to with a maximum of 0.1 e between the 11-*cis* to the all-*trans* (55), the same partial charges were kept to describe retinal in the dark-adapted and photo-activated states.

Concerning the presentation of the data, the averages presented are performed on the set of three MDs of the same duration. Due to space limitations, the figures present only the MD simulations starting from 500ps in the dark-adapted MD (the other traces are available in the supporting information). When compared with the simulation for the dark-adapted state, the curves are displayed so that they share the same timeframe, i.e. if the isomerization was initiated after 500 ps MD in the dark-state, the curve for the "after isomerization" trajectory starts at 500ps.

The comparison of the RMSD curves for rhodopsin (Figure 3A) shows that the isomerization of retinal leads the protein to follow a different path during the MD than the one followed in the dark-adapted state (note that the conformation of reference for all RMSD curves is the same: the conformation of the protein at time = 0 of the MD simulation in the dark-adapted state). From the RMSF plot, most of the motions taking place in the protein are located near the loops (data not shown), in particular the intracellular loops and the C-terminal part of the protein, in a scheme similar to the dark-adapted state simulation.

The results presented here illustrate the path followed by the molecular model of retinal upon isomerization. The results strongly suggest that the path has a Brownian diffusion and depends on both starting coordinates and velocities of the retinal. Since only a few of the many available paths have been followed for a relatively short time, this analysis cannot be exhaustive, but should rather interpreted in semi-quantitative way.

Dynamics of the isomerization and pattern of activation

The time isomerization of the chromophore from 11-*cis* to all-trans starts within 50 fs after the isomerization is initiated and takes up to 200 fs to be completed depending on the exact starting

conformation of the retinal. The next important conformational change within retinal is the flip of methyl C20, which takes place within 500 fs to 1.5 ps after the isomerization starts.

The isomerization of the model is followed by monitoring the dihedral angle of the C10-C11–C12-C13 bond as a function of time (Figure 6C for a detailed description of the first 5ps). From the three 5ps-MDs, it takes about 150-200 fs for retinal to isomerize around the C11-C12 bond. In addition, the different conformations generated by each of the MD simulations were clustered using a full linkage algorithm (and a cut-off of 0.07 Å, Figure 8). This type of analysis allows to simplify the visualization of the conformation changes within retinal and to identify in all three simulations short-lived cluster with a lifetime of about 50-800fs (range derived from the three MDs), where the C11-C12 bond has been isomerized without the rest of the polyene chain having the time to rearrange itself. Due to the time-scale and the nature of the conformational disorders, it is tempting to associate these clusters with the photorhodopsin intermediate. Further the very short lifetime of these clusters is an explanation in itself of why it is not possible to isolate this photo-intermediate.

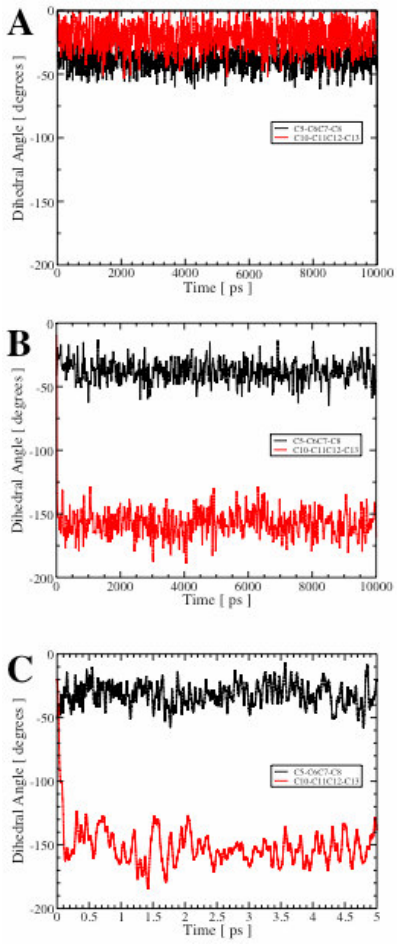


Figure 6: Dihedral angle C5-C6–C7-C8 and C10-C11–C12-C13 for retinal bound to rhodopsin in the dark-adapted state (**A**), 10ns-MD describing the isomerization of retinal after 500ps-MD in dark-adapted state (**B**), 5ps-MD describing the isomerization of retinal after 500ps-MD in dark-adapted state (**C**).

Another important conformational changes happening within the first picoseconds after the isomerization is initiated, is the rotation of the methyl group C20 bound to the polyene chain at carbon C13. As the rotation of the methyl group takes place along an axis, which is approximately the normal to the membrane plane, the angle defined by the bond vector C13-C20 and the Z-axis of the simulation box is adequate to follow the orientation of the bond inside the binding-pocket. Figure 7B presents an example of the rotation undergone by the methyl group, where it takes about 1ps for the rotation to be complete. The average-time needed for the rotation to take place over the three 5ps-MD, is 1.33 ps (1.5ps/1ps/1.5ps) and the average amplitude of the rotation is 95° (100°/93°/93°). Further cluster-analysis allows identifying clusters (Figure 8) whose lifetime starts shortly after the methyl group has flipped over and that have a lifetime of several picoseconds,

strongly suggests it is related with bathorhodopsin because of the timescale at which this conformational change takes place and its amplitude.

Table 4: Dihedral angles for selected bonds for retinal after the isomerization, determined by 10ns-MD within 2ns after the isomerization of retinal at 300K, and NMR for retinal in metarhodopsin I at 210K.

Dihedral Angle	MD Average	NMR
C5-C6–C7-C8	-38 ± 9° * -36 ± 9° * -38 ± 9° * (37 ± 9°)	-28 ± 7° #, •
C10-C11–C12-C13	-157 ± 12° * -162 ± 12° * -159 ± 12 * (159 ± 12°) *	N.A.
C9-C10–C11-C12	176 ± 13° * 171 ± 16° * 175 ± 11° * (174 ± 14°)	180 ± 15° #

* = Standard deviation of the fluctuations during the MD : # = Experimental error ; • = NMR-based estimation are based on the 11-cis-retinal crystal structure for which the bond angle was rotated in order to match ^{13}C-^{13}C rotational resonance distances.

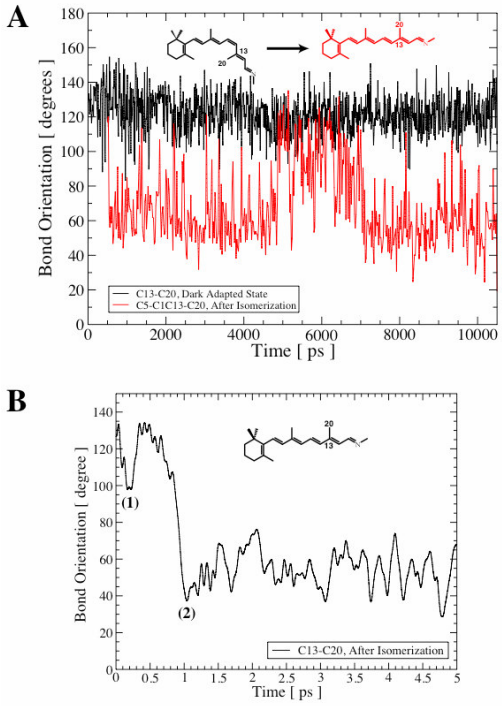


Figure 7: Angle of bond C13-C20 with the Z-axis (normal to the membrane plane), showing a large-scale rotation of methyl group C20 upon isomerization (70-90°, depending on the starting conformation). (**A**) 10ns-MD of bovine rhodopsin in the dark-adapted state (black) and after isomerization of retinal (red), after 500ps-MD in the dark. (**B**) 5ps-MD of bovine rhodopsin after isomerization of retinal (isomerization initiated after 500ps in the dark). (**1**) Point corresponding to the end of the isomerization of the C11-C12 double bond, interpreted as being "photorhodopsin".

(**2**) Point corresponding to the flip of methyl-C20, interpreted as being "bathorhodopsin". The bond orientation is given related to the normal to the membrane plane, 0° corresponding to a vector pointing towards the extracellular-side and 180° correspond to a vector pointing towards the cytoplasmic side of the membrane.

Altogether, the timescale and the amplitude of the various conformational changes observed within the chromophore during the first 5ps after isomerization, suggest they may be related with previously identified photo-intermediate photorhodopsin and bathorhodopsin. In addition, the isomerization from slightly different conformation for the 11-*cis* retinal leads to a distribution of the time required for transition from one photo-intermediate to another (i.e. and the conformational changes they imply). This is interpreted as the existence of multiple pathways leading to the activation of the protein and might explain the coexistence of several photo-intermediate states. Nevertheless, the sequence of the observed conformational changes, remain the same in all the simulations.

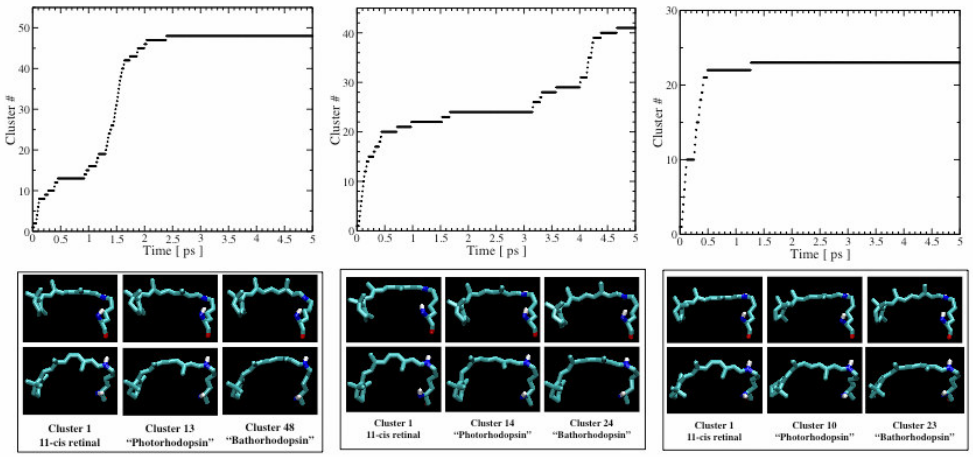


Figure 8: Cluster analysis on several 5ps-MD simulations where retinal is being isomerized from different starting-points taken from the trajectory of bovine rhodopsin in the dark-adapted state. Conformations are sampled every 1fs from the trajectory and a full-linkage algorithm with a 0.07 Å cut-off is used.

In order to extend the three short 5ps-MD described in the previous paragraphs, three longer 10ns-MD simulations, starting from the same points of the trajectory where rhodopsin is simulated during 10ns in the dark-adapted state have been run, allowing to study the conformational changes taking place in retinal on a much longer timescale.

In parallel to these very localized conformational changes at the picosecond scale, other conformational changes take place at the nanosecond-scale. These changes are not localized as the

earliest changes within the chromophore, but rather involve the entire retinal molecules and are therefore subtler to identify. A careful comparison with the 10ns-MD of rhodopsin in the dark-adapted state is required to ensure it results from the isomerization and not from a thermal fluctuation.

Among these changes, there is an extension of the retinal length by about 1.0-1.4 Å, and can be tracked by following the interatomic distance between either C13 or C15 and C4 (Table 5). Another conformational changes is the increase of the interatomic distance between C8 and C18, by a modest 0.2 Å, as a result from a slight change in the orientation of the ring (Figure 10B). The methyl groups undergo increased fluctuation (Figure 5), which can affect their average orientation (Figure 7).

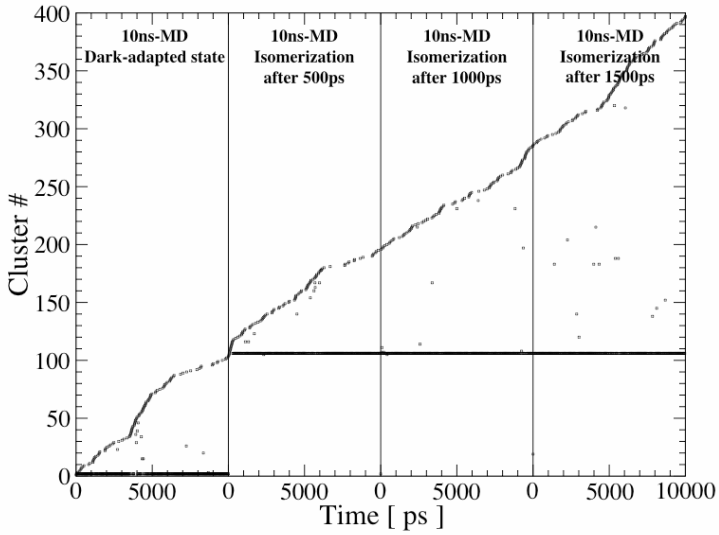


Figure 9: Cluster analysis on several all the 10ns-MD simulations. Conformations are sampled every 20ps from the trajectory (a total of 2000 conformations analyzed) and a full-linkage algorithm with a 0.21 Å cut-off is used. Two main clusters can be identified, cluster #2 gathering 78% of the retinal conformers in the dark-adapted state and cluster #106 gathering also 78% of the retinal conformers after isomerization.

Further cluster analysis performed simultaneously on all 10ns-MD (Figure 9) shows that both retinal in the dark-adapted state and after isomerization spend most of their time in two distinct clusters. For both states, 78% of conformation can be gathered in one cluster with a cut-off of 0.021 Å (no cluster visible with a cut-off lower than 0.018 Å). This means that within 10ns, retinal does not fluctuate more than ~0.2 Å-RMSD away from its average conformation. Furthermore, the

amplitude of the fluctuations seems to be the same for retinal in the dark and retinal after isomerization, i.e. on the nanosecond time scale, conformational changes are driven mainly through thermal fluctuations or perturbation that have a similar amplitude.

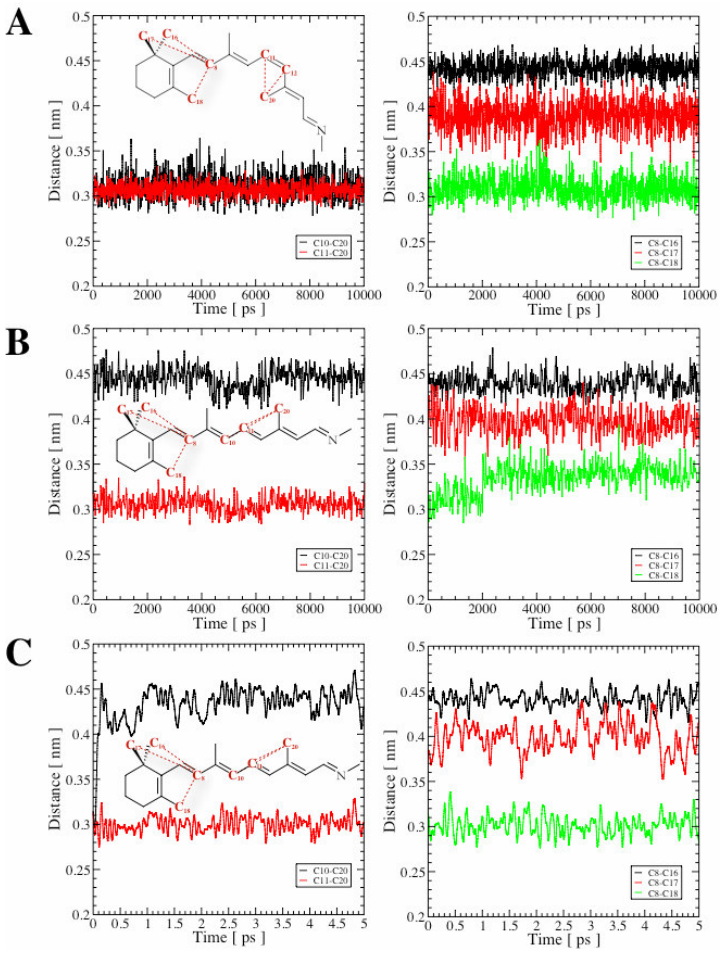


Figure 10: Illustration of intramolecular distances within retinal, experimentally-measured by solid state NMR in the dark-adapted state and in metarhodopsin I. (**A**) 10ns-MD in the dark-adapted state, (**B**) example of a 10ns-MD where the isomerization of retinal is initiated at time = 0ps (500ps-MD in the dark-adapted state was run previously). (**C**) same as (**B**) with a simulation time of 5ps.

Table 5: Intramolecular distances illustrating the length extension of retinal upon isomerization. The corresponding distances are longer in the crystallized all-trans retinal, suggesting the chromophore can extend more.

Intramolecular Distances	Dark-Adapted State	After Isomerization	Crystallized All-trans-retinal
C13-C4 [Å]	8.81 ± 0.31 *	10.16 ± 0.33 # 10.16 ± 0.34 # 9.92 ± 0.42 # (10.08 ± 0.36)	11.48
C15-C4	10.44 ± 0.41 *	11.87 ± 0.48 #	13.55

[Å]		11.92 ± 0.44 # 11.50 ± 0.55 # (11.76 ± 0.51)	

*: average on a single 10ns-MD of bovine rhodopsin in the dark-adapted state**: average on the three 10ns-MD of bovine rhodopsin after isomerizaiton of retinal.

Relaxation of the Binding Pocket

A general feature of the retinal model after isomerization is an increase in the vibrational fluctuations of some fragments of the molecule. These fluctuations are highlighted in Figure 4B, where the amplitude of the RMSD fluctuations has increased compare to the dark-state (Figure 4A). Furthermore the RMSF plot for retinal after isomerization (Figure 5) highlights the methyl groups (C16-C20) and atoms C3-C4 from the β-ionone ring, as the positions in the chromophore where increased vibrational activity takes place. This activity results from the process of reciprocal adaptation between the chromophore and its binding-pocket after the isomerization. A transfer of this vibrational activity certainly takes place between the retinal and its binding pocket. Figure 11B illustrates this at the level of the β-ionone ring, where the time-fluctuations for distance between the ring and the surrounding residues increase and can possibly act as a way to transfer the signal from the retinal to the protein. The amplitude of these fluctuations – 0.1 to 0.3 Å – are small and may explain why it takes milliseconds for the protein to reach its activated-state, the perturbation induced by the chromophore on the protein being relatively small. Furthermore, the β-ionone ring moves closer to helix V and away from helix VI by a distance of 0.1-0.2 Å. The amplitude of this displacement is rather tiny and simply means that the contact between the β-ionone ring and helix V is reinforced in the first nanoseconds of the MDs describing the isomerization of retinal. It might result either from the extension of the retinal polyene tail or from the modest increase of the vibrational level near the ring.

In this model, not all the parts of the retinal are affected by the isomerization. The most striking example is the salt-bridge between the protonated Schiff-base and Glu113, an important part of the molecule, does not seem to be affected at all in the early stage of the isomerization, since both the average distance between the two groups and the fluctuation around the average are changed by less than 0.1 Å upon activation (Figure 11A). This is in agreement with experimental data since the salt-bridge is not expected to break before metarhodopsin I (~ms after isomerization).

Protein conformational change

At this stage, it is difficult to identify alterations in the conformation of the protein, which would undoubtedly results from the isomerization of retinal. The RMSD and RMSF plots point at various small alterations in the conformation of the protein during the MD simulations after initiation of the isomerization. Nevertheless, these changes, which can account for changes in the tilt angles of the transmembrane helices by a few degrees (data not shown), are difficult to interpret, since they are not consistently observed along the three simulations and not extremely different from the fluctuations observed for MD simulations in the dark-adapted state. Longer simulation with a timescale of a few 100ns would be required to investigate these conformational changes more thoroughly.

Simulating all-trans retinal in the early stages of the activation of rhodopsin

In the metarhodopsin I state, the polyene tail in the region where the isomerization takes place, distances conform with the *cis-trans* isomerization where measured, showing an increase of the C10–C20 distance by more than 1.3 Å upon activation, whereas the C11–C20 remains almost unchanged (72). In the same photo-intermediate, distance measured by rotational resonance and the observation of ^{13}C chemical shift introduced into the β-ionone ring (66) (at the C16, C17, and C18 methyl groups) and into the adjoining segment of the polyene chain (at C8) suggest the orientation of the ring remains unchanged upon isomerization. The only significant chemical shift change that could be detected for the ring methyl groups on photo-activation was in the C18 resonance, which increased from 22.1 to 22.5 ppm (66). The large splitting between C16 and C17 (-4.3 ppm) describes the unique orientation of these geminal methyl groups in rhodopsin (67), and its retention on photo-activation is important for the determination of the mechanism of activation (66).

The comparison between computed and experimentally measured data is for the less inhomogeneous since they are related to different states of the protein and at least 250 ns separate them. Nonetheless, these values constitute a point of reference to compare to the distorted all-trans retinal simulated with.

The comparison highlights two points. First, the distances C11-C20 is very close to the value taken in the metarhodopsin I, as a result of the rotation of the C20 methyl group following the isomerization. Second, the agreement with the interatomic distances measure near the ring is less striking, mainly due to the 0.2-0.3 Å deviation from the C8-C18 distance as measured by NMR. The fluctuations result directly from the isomerization of the retinal and the stress it induces on the chromophore, and are likely to persist until the binding-pocket has accommodated the all-*trans* retinal (i.e. lumirhodopsin).

The slight increase of the C8-C18 distance is also accompanied by ring-fluctuations resulting momentaneous reorientation of methyl C18 away from Trp265. While the ring oscillates between its initial orientation and its new oritentation where methyl C18 is rotated away from Trp265, it cannot be interpreted as the β-ionone ring moving away from the binding-pocket, as shown in Figure 11, where the average distance between the ring and the aromatic binding pocket remain constant, only their standard deviations increasing. The amplitude of the fluctuation seems too small to prevent the ring to recover its initial orientation later once the retinal has relaxed.

Mechanism of Activation

This theoretical study suggests that the early stage of the activation of rhodopsin are localized mainly on retinal, where a serie of conformational changes happen in a sequential manner, spreading away from the isomerized bond. First the isomerization of the C11=C12 bond within a few hundreds femtoseconds, then the rotation of methyl group C20 moving up from Trp265 within 1-2 ps. These highly localized changes are then followed by increased fluctuations at the level of the whole molecule, mainly localized on the methyl groups and the β-ionone ring. These fluctuations disturb the orientation of the ring and its relative position relative to helix V and VI (Table 6), as well with the aromatic residues in the surrounding (Figure 11). The amplitude of these change are tenuous and involves changes in the ring that are in terms of a few tenths of Ångtröms, and cannot be interpreted as the ring moving away from the aromatic binding pocket.

It suggests the 11-*cis* bond relaxes quickly after isomerization (C11 and C12 are slightly less mobile after isomerization according RMSF – Figure 5), transferring the mechanical stress to another part

of the retinal in a sequential way, starting at the 11-*cis* bond and spreading away from the epicentre of the isomerization.

Table 6: Distance between the β-ionone ring and neighbour helices. Within 10ns after isomerization, the β-ionone ring only shows a slight move toward helix V, away from helix VI.

	Dark-Adapted State	After Isomerization
Ring – Helix V [Å]	3.22 ± 0.26	3.12 ± 0.28 3.11 ± 0.26 3.17 ± 0.29
Ring – Helix VI [Å]	2.93 ± 0.21	3.08 ± 0.28 3.18 ± 0.32 3.08 ± 0.29

The simulations suggests the β-ionone ring is to play an important role in the isomerization, through the increased fluctuations in its orientation compared to the dark-adapted state, and the slight increase of the contact made with helix V (aligned with the polyene chain).

Recently, helix V has been suggested to play an important role in the activation of rhodopsin (50). The simulations seem to show the premises of the increased interactions between retinal and helix V as part of a mechanism through which the retinal transfers the signal to the protein.

The salt bridge on the other extremity of the chromophore anchors solidly the ring to helix III, at least during the first 10ns of the activation (Figure 11). As a consequence, the only part of the retinal allowed to move is its ring, which is also take advantage of a slightly less tight packing (it is known that various retinal conformers can be docked in this part of the binding pocket).

The extrapolation of the 10ns-MDs suggests that the small fluctuation of the ring could play a part in the process initiating the activation of the receptor. The view, far from describing retinal as a bottle-opener forcing its way through the protein, suggest rather a small-oscillator whose tiny action, repeated on several 100ns could trigger a conformation changes in the protein.

This concept follow the evolution in the description of weak noncovalent interactions, mainly driven by atomic-force microscopy, where these interactions are described as having limited lifetimes and so would fail under any level of force if pulled on for the right length of time (18). With this view in mind, retinal would just apply a localized perturbation with the right amplitude and timescale leading to the protein activation.

In such a model, retinal does not need to move away from its binding pocket to achieve its task, and can therefore be associated with recent findings that the orientation of the ring is unchanged in the metarhodopsin I (66) and the mechanism suggested by a variant of rotational resonance (DARR) (13, 70) was applied to obtain distance restraints between between ^{13}C-labels on the retinal chromophore and specific ^{13}C-labels on tyrosine, glycine, serine, and threonine in the retinal binding-site of rhodopsin (50).

This recent study provides more information on how the retinal isomerization would disrupt helix interactions locking the receptor off in the dark state and suggest a mechanism of activation for rhodopsin and highlights two essential aspects of the isomerization trajectory suggested from the retinal-protein contacts observed in the active metarhodopsin II intermediate imply a large rotation of the C20 methyl group (≥ 90°) toward extra-cellular loop 2 and a 4–5 Å translation of the retinal chromophore toward transmembrane helix V.

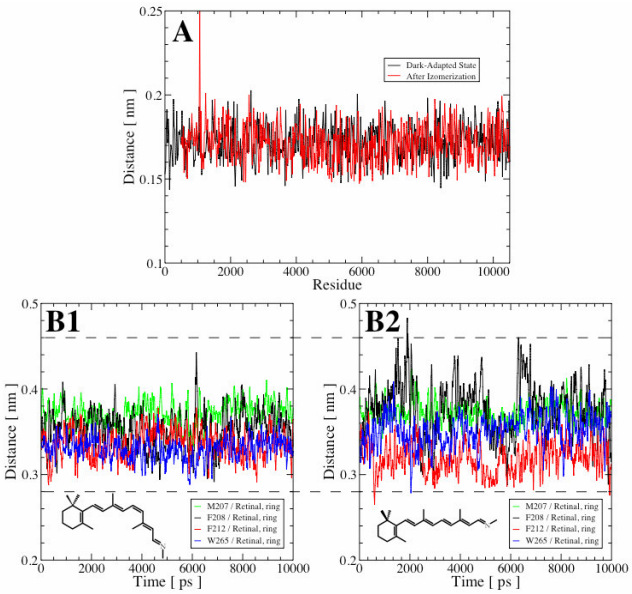


Figure 11: (**A**) Illustration of the stability of the interaction between the protonated Schiff-base and the counter-ion (Glu113), both in the dark-adapted state (1.72 ± 0.12 Å) and in the early stage of the activation (1.72 ± 0.14 Å). (**B**) Distance between β-ionone ring and selected residues from the binding pocket (carbon-carbon distance only). The increased fluctuation in these distances upon isomerization (B2) illustrates the increased level of vibration of the ring after isomerization (left = dark-adapted state, right = after isomerization).

Ring flip

In this MD study, the β-ionone ring remained in the binding-pocket, in a conformation very similar to the one observed in the dark-state, in agreement with recent NMR structural study on bovine rhodopsin (dark-state (67) and metarhodopsin I (66)).

The flip of the ring upon activation proposed in a previous study was not observed with the model used for retinal. The rigidity of the ring compare to previous studies, where the rotation of the ring was observed results from two factors: the conformation of the retinal in the dark-adapted state differs slightly and the treatment of the dihedrals is different. Clearly, the possibility for the ring to flip is governed by the choice of one parameter for the potential energy function describing the C5=C6-C7=C8 (Lemaître, Data not shown). In order for the chromophore model to fit experimental information available for the orientation of the ring resulting from ^{13}C-^{13}C distance measurement by solid state NMR, all the dihedral interactions within polyene chain of retinal were treated as double-bonds, thus increasing the rigidity of the molecule and preventing the ring to flip. This choice is also justified by the bond order as determined by solid state NMR, which appears to be larger than 1 for almost all the bonds in the polyene chain (Malcolm Levitt, 2004, personal communication). This illustrates the difficulties to simulate a system in which the electronic delocalization plays an important role.

Potential pitfalls

We recognize that there are limitations to this computational approach. There are intrinsic limitations to MD simulations, which include the classical treatment of the system and a static attribution of the electronic charges to the nuclei. The other challenge introduced by rhodopsin for computational study is the large separation in time-scales between the steps in photo-cycle of the protein with respect to the current limits in computer simulations of large proteins. While it takes a few ms for the protein to reach the activated-state and much more before the retinal to be regenerated (33), the state-of-the-art computer simulation of large protein with explicit solvent and membrane description are still limited to a 10-100 ns (14). Therefore MD is not yet able to follow the activation of the receptor from the dark-adapted state to the active species of rhodopsin, the so-called metarhodopsin II, where the binding and activation of the G protein transducin (G_t) occurs

(20, 23, 51). Rather it is limited to the first photo-intermediates, where most conformational changes take place on the chromophore or its binding-pocket.

CONCLUSION

This contribution presents molecular dynamic simulations of retinal in its binding-pocket, with bovine rhodopsin inserted in a hydrated POPC bilayer. The retinal model based on previous DFT calculations allows matches well the structural information available. The extrapolation of this model to study the dynamics and conformation of the earliest photointermediates, provides a reasonable model for the activation of retinal and suggest a possible path for the transfer of the signal to the protein. Clear structural features are suggested for both photorhodopsin and bathorhodopsin. The evolution of the chromophore in the multiple 10ns-MD suggests how the motion of retinal can be coupled to the protein, i.e. through slightly increased vibrational activity of the ring and a slight extension of retinal moving maintaining close contact with helix V. To the opposite, the repetition of the isomerization from different starting retinal conformation does not allow to identify a clear and defined pattern for the activation of the protein, in agreement with the timescale required for the protein activation (ms).

ACKNOWLEDGEMENTS

VL thanks Dr Minoru Sugihara for providing us with the coordinate file describing the DFT-derived 11-*cis*,6-s-*cis* retinal conformer published in 2002 and Dr Paul Spooner for the many discussions we had about rhodopsin-bound retinal. Prof. Peter Tieleman is thanked for making available pre-equilibrated POPC bilayer. VL thanks BBSRC for support.

LITERATURE CITED

1. Albert AD, Yeagle PL. 2002. Structural studies on rhodopsin. *Biochim Biophys Acta* 1565: 183
2. Bagley KA, Balogh-Nair V, Croteau AA, Dollinger G, Ebrey TG, et al. 1985. Fourier-transform infrared difference spectroscopy of rhodopsin and its photoproducts at low temperature. *Biochemistry* 24: 6055
3. Berendsen HJC, Griegera JR, Straatsma TP. 1987. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* 91: 6269

4. Berger O, Edholm O, Jahnig F. 1997. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys J* 72: 2002
5. Birge RR, Einterz CM, Knapp HM, Murray LP. 1988. The nature of the primary photochemical events in rhodopsin and isorhodopsin. *Biophys J* 53: 367
6. Burns ME, Baylor DA. 2001. Activation, deactivation, and adaptation in vertebrate photoreceptor cells. *Annu Rev Neurosci* 24: 779
7. Buss V. 2001. Inherent chirality of the retinal chromophore in rhodopsin-A nonempirical theoretical analysis of chiroptical data. *Chirality* 13: 13
8. Buss V, Sugihara M, Entel P, Hafner J. 2003. Thr94 and Wat2b Effect Protonation of the Retinal Chromophore in Rhodopsin. *Angew. Chem. Int. Ed.* 42: 3245–3247
9. Carravetta M, Zhao X, Johannessen OG, Lai WC, Verhoeven MA, et al. 2004. Protein-induced bonding perturbation of the rhodopsin chromophore detected by double-quantum solid-state NMR. *J Am Chem Soc* 126: 3948
10. Cohen GB, Oprian DD, Robinson PR. 1992. Mechanism of activation and inactivation of opsin: role of Glu113 and Lys296. *Biochemistry* 31: 12592
11. Cooper A. 1979. Energy uptake in the first step of visual excitation. *Nature* 282: 531
12. Creemers AF, Kiihne S, Bovee-Geurts PH, DeGrip WJ, Lugtenburg J, de Groot HJ. 2002. ^{1}H and ^{13}C MAS NMR evidence for pronounced ligand-protein interactions involving the ionone ring of the retinylidene chromophore in rhodopsin. *Proc Natl Acad Sci U S A* 99: 9101
13. Crocker E, Patel AB, Eilers M, Jayaraman S, Getmanova E, et al. 2004. Dipolar assisted rotational resonance NMR of tryptophan and tyrosine in rhodopsin. *J Biomol NMR* 29: 11
14. Crozier PS, Stevens MJ, Forrest LR, Woolf TB. 2003. Molecular dynamics simulation of dark-adapted rhodopsin in an explicit membrane bilayer: coupling between local retinal and larger scale conformational change. *J Mol Biol* 333: 493
15. de Grip WJ, DeLange F, Klaassen CH, Verdegem PJ, Wallace-Williams S, et al. 1999. Photoactivation of rhodopsin: interplay between protein and chromophore. *Novartis Found Symp* 224: 102
16. de Grip WJ, Gray D, Gillespie J, Bovee PH, Van den Berg EM, et al. 1988. Photoexcitation of rhodopsin: conformation changes in the chromophore, protein and associated lipids as determined by FTIR difference spectroscopy. *Photochem Photobiol* 48: 497
17. Deng H, Callender RH. 1987. A study of the Schiff base mode in bovine rhodopsin and bathorhodopsin. *Biochemistry* 26: 7418
18. Evans E. 1998. Energy landscapes of biomolecular adhesion and receptor anchoring at interfaces explored with dynamic force spectroscopy. *Faraday Discuss*: 1
19. Eyring G, Mathies R. 1979. Resonance Raman studies of bathorhodopsin: evidence for a protonated Schiff base linkage. *Proc Natl Acad Sci U S A* 76: 33
20. Fain GL, Matthews HR, Cornwall MC, Koutalos Y. 2001. Adaptation in vertebrate photoreceptors. *Physiol Rev* 81: 117
21. Faraldo-Gomez JD, Smith GR, Sansom MS. 2002. Setting up and optimization of membrane protein simulations. *Eur Biophys J* 31: 217
22. Feng X, Verdegem PJ, Lee YK, Sandström D, Edén M, et al. 1997. Direct determination of a molecular torsional angle in the membrane protein rhodopsin by solid-state NMR. *J Am Chem Soc* 119: 6853–6857
23. Hamm HE. 2001. How activated receptors couple to G proteins. *Proc Natl Acad Sci U S A* 98: 4819
24. Hayashi S, Tajkhorshid E, Schulten K. 2002. Structural changes during the formation of early intermediates in the bacteriorhodopsin photocycle. *Biophys J* 83: 1281
25. Honig B, Ebrey T, Callender RH, Dinur U, Ottolenghi M. 1979. Photoisomerization, energy storage, and charge separation: a model for light energy transduction in visual

pigments and bacteriorhodopsin. *Proc Natl Acad Sci U S A* 76: 2503

26. Hubbell WL, Altenbach C, Hubbell CM, Khorana HG. 2003. Rhodopsin structure, dynamics, and activation: a perspective from crystallography, site-directed spin labeling, sulfhydryl reactivity, and disulfide cross-linking. *Adv Protein Chem* 63: 243
27. Huber T, Botelho AV, Beyer K, Brown MF. 2004. Membrane model for the G-protein-coupled receptor rhodopsin: hydrophobic interface and dynamical structure. *Biophys J* 86: 2078
28. Hug SJ, Lewis JW, Einterz CM, Thorgeirsson TE, Kliger DS. 1990. Nanosecond photolysis of rhodopsin: evidence for a new, blue-shifted intermediate. *Biochemistry* 29: 1475
29. Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. *J Mol Graph* 14: 33
30. Jager S, Lewis JW, Zvyaga TA, Szundi I, Sakmar TP, Kliger DS. 1997. Chromophore structural changes in rhodopsin from nanoseconds to microseconds following pigment photolysis. *Proc Natl Acad Sci U S A* 94: 8557
31. Kandori H, Shichida Y, Yoshizawa T. 1989. Absolute absorption spectra of batho- and photorhodopsins at room temperature. Picosecond laser photolysis of rhodopsin in polyacrylamide. *Biophys J* 56: 453
32. Kandori H, Shichida Y, Yoshizawa T. 2001. Photoisomerization in rhodopsin. *Biochemistry (Mosc)* 66: 1197
33. Klein-Seetharaman J. 2002. Dynamics in rhodopsin. *Chembiochem* 3: 981
34. Kliger DS, Lewis JW. 1995. Spectral and kinetic characterization of visual pigment photointermediates. *Israel Journal of Chemistry* 35: 289–307
35. Kuwata O, Yuan C, Misra S, Govindjee R, Ebrey TG. 2001. Kinetics and pH dependence of light-induced deprotonation of the Schiff base of rhodopsin: possible coupling to proton uptake and formation of the active form of Meta II. *Biochemistry (Mosc)* 66: 1283
36. Lewis JW, Fan GB, Sheves M, Szundi I, Kliger DS. 2001. Steric barrier to bathorhodopsin decay in 5-demethyl and mesityl analogues of rhodopsin. *J Am Chem Soc* 123: 10024
37. Lewis JW, Pinkas M, Sheves M, Ottolenghi M, Kliger DS. 1995. Structural changes in early photolysis intermediates of rhodopsin from time-resolved spectral measurements of artificial pigments sterically hindered along the chromophores chain. *J Am Chem Soc* 117: 918–923
38. Li J, Edwards PC, Burghammer M, Villa C, Schertler GF. 2004. Structure of Bovine Rhodopsin in a Trigonal Crystal Form. *J. Mol. Biol.* (In Press)
39. Marcus MA, Lewis A. 1979. Assigning the resonance Raman spectral features of rhodopsin, isorhodopsin and bathorhodopsin in bovine photostationary state spectra. *Photochem Photobiol* 29: 699
40. Mathies RA. 1999. Photons, femtoseconds and dipolar interactions: a molecular picture of the primary events in vision. *Novartis Found Symp* 224: 70
41. Mollevanger LC, Kentgens AP, Pardoen JA, Courtin JM, Veeman WS, et al. 1987. High-resolution solid-state ^{13}C-NMR study of carbons C-5 and C-12 of the chromophore of bovine rhodopsin. Evidence for a 6-s-*cis* conformation with negative-charge perturbation near C-12. *Eur J Biochem* 163: 9
42. Okada T, Ernst OP, Palczewski K, Hofmann KP. 2001. Activation of rhodopsin: new insights from structural and biochemical studies. *Trends Biochem Sci* 26: 318
43. Okada T, Fujiyoshi Y, Silow M, Navarro J, Landau EM, Shichida Y. 2002. Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc Natl Acad Sci U S A* 99: 5982
44. Okada T, Palczewski K. 2001. Crystal structure of rhodopsin: implications for vision and beyond. *Curr Opin Struct Biol* 11: 420
45. Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. 2004. The Retinal Conformation and its Environment in Rhodopsin in Light of a new 2.2 Å Crsystal Structure. *J. Mol. Biol.* (In Press)
46. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, et al. 2000. Crystal Structure of Rhodopsin: A G

protein-coupled receptor. *Science* 277: 687

47. Palings I, van den Berg EM, Lugtenburg J, Mathies RA. 1989. Complete assignment of the hydrogen out-of-plane wagging vibrations of bathorhodopsin: chromophore structure and energy storage in the primary photoproduct of vision. *Biochemistry* 28: 1498
48. Pan D, Mathies RA. 2001. Chromophore structure in lumirhodopsin and metarhodopsin I by time-resolved resonance Raman microchip spectroscopy. *Biochemistry* 40: 7929
49. Patel AB, Crocker E, Eilers M, Hirshfeld A, Sheves M, Smith SO. 2004. Coupling of retinal isomerization to the activation of rhodopsin. *Proceedings of the National Academy of Science* 101: 10048–10053
50. Patel AB, Crocker E, Eilers M, Hirshfeld A, Sheves M, Smith SO. 2004. Coupling of retinal isomerization to the activation of rhodopsin. *Proc Natl Acad Sci U S A*
51. Polans A, Baehr W, Palczewski K. 1996. Turned on by Ca^{2+}! The physiology and pathology of Ca^{2+}-binding proteins in the retina. *Trends Neurosci* 19: 547
52. Popp A, Ujj L, Atkinson GH. 1996. Bathorhodopsin structure in the room-temperature rhodopsin photosequence: picosecond time-resolved coherent anti-Stokes Raman scattering. *Proc Natl Acad Sci U S A* 93: 372
53. Randall CE, Lewis JW, Hug SJ, Björling SC, Eisner-Shanas I, et al. 1991. A new photolysis intermediate in artificial and native visual pigments. *J Am Chem Soc* 113: 3473–3485
54. Röhrig UF, Guidoni L, Rothlisberger U. 2002. Early steps of the intramolecular signal transduction in rhodopsin explored by molecular dynamics simulations. *Biochemistry* 41: 10799
55. Saam J, Tajkhorshid E, Hayashi S, Schulten K. 2002. Molecular dynamics investigation of primary photoinduced events in the activation of rhodopsin. *Biophys J* 83: 3097
56. Sakmar TP. 1998. Rhodopsin: a prototypical G protein-coupled receptor. *Prog Nucleic Acid Res Mol Biol* 59: 1
57. Sakmar TP, Menon ST, Marin EP, Awad ES. 2002. Rhodopsin: insights from recent structural studies. *Annu Rev Biophys Biomol Struct* 31: 443
58. Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M. 1995. Evaluation of comparative protein modeling by MODELLER. *Proteins* 23: 318
59. Schoenlein RW, Peteanu LA, Mathies RA, Shank CV. 1991. The first step in vision: femtosecond isomerization of rhodopsin. *Science* 254: 412
60. Shichida Y, Imai H. 1998. Visual pigment: G-protein-coupled receptor for light signals. *Cell Mol Life Sci* 54: 1299
61. Shieh T, Han M, Sakmar TP, Smith SO. 1997. The steric trigger in rhodopsin activation. *J Mol Biol* 269: 373
62. Smith SO, Courtin J, de Groot H, Gebhard R, Lugtenburg J. 1991. ^{13}C magic-angle spinning NMR studies of bathorhodopsin, the primary photoproduct of rhodopsin. *Biochemistry* 30: 7409–15
63. Smith SO, de Groot H, Gebhard R, Lugtenburg J. 1992. Magic angle spinning NMR studies on the metarhodopsin II intermediate of bovine rhodopsin: evidence for an unprotonated Schiff base. *Photochem Photobiol* 56: 1035–1039
64. Smith SO, Palings I, Copie V, Raleigh DP, Courtin J, et al. 1987. Low-temperature solid-state ^{13}C NMR studies of the retinal chromophore in rhodopsin. *Biochemistry* 26: 1606–11
65. Smith SO, Palings I, Miley ME, Courtin J, de Groot H, et al. 1990. Solid-state NMR studies of the mechanism of the opsin shift in the visual pigment rhodopsin. *Biochemistry* 29: 8158
66. Spooner PJ, Sharples JM, Goodall SC, Seedorf H, Verhoeven MA, et al. 2003. Conformational similarities in the beta-ionone ring region of the rhodopsin chromophore in its ground state and after photoactivation to the metarhodopsin-I intermediate. *Biochemistry* 42: 13371
67. Spooner PJ, Sharples JM, Verhoeven MA, Lugtenburg J, Glaubitz C, Watts A. 2002. Relative orientation between the beta-ionone ring and the polyene chain for the chromophore of

rhodopsin in native membranes. *Biochemistry* 41: 7549

68. Sugihara M, Buss V, Entel P, Elstner M, Frauenheim T. 2002. 11-*cis*-retinal protonated Schiff base: influence of the protein environment on the geometry of the rhodopsin chromophore. *Biochemistry* 41: 15259
69. Sugihara M, Buss V, Entel P, Hafner J. 2004. The Nature of the Complex Counterion of the Chromophore in Rhodopsin. *J. Phys. Chem. B* 108: 3673–3680
70. Takegoshi K, Nakamura K, Terao T. 2001. ^{13}C–^{1}H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem Phys Lett* 344: 631–637
71. Teller DC, Okada T, Behnke CA, Palczewski K, Stenkamp RE. 2001. Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). *Biochemistry* 40: 7761
72. Verdegem PJ, Bovee-Geurts PH, de Grip WJ, Lugtenburg J, de Groot HJ. 1999. Retinylidene ligand structure in bovine rhodopsin, metarhodopsin-I, and 10-methylrhodopsin from internuclear distance measurements using ^{13}C-labeling and 1-D rotational resonance MAS NMR. *Biochemistry* 38: 11316
73. Verdegem PJE, Helmle M, Lugtenburg J, de Groot HJM. 1997. Internuclear Distance Measurements up to 0.44 nm for Retinals in the Solid State with 1-D Rotational Resonance 13C MAS NMR Spectroscopy. *J Am Chem Soc* 119: 169
74. Verhoeven MA, Creemers AF, Bovee-Geurts PH, De Grip WJ, Lugtenburg J, de Groot HJ. 2001. Ultra-high-field MAS NMR assay of a multispin labeled ligand bound to its G-protein receptor target in the natural membrane environment: electronic structure of the retinylidene chromophore in rhodopsin. *Biochemistry* 40: 3282
75. Wald G. 1968. The molecular basis of visual excitation. *Nature* 219: 800
76. Yamada A, Kakitani T, Yamamoto S, Yamamoto T. 2002. A computational study on the stability of the protonated Schiff base of retinal in rhodopsin. *Chemical Physics Letters* 366: 670–675
77. Yan EC, Kazmi MA, Ganim Z, Hou JM, Pan D, et al. 2003. Retinal counterion switch in the photoactivation of the G protein-coupled receptor rhodopsin. *Proc Natl Acad Sci U S A* 100: 9262

II.3 Membrane – Peptide Interactions Studied by Molecular Dynamic Simulations of Atomic Force Microscopy Experiments

Vincent Lemaître,[1,2] Maurits R. de Planque,[1] and Anthony Watts,[1]

[1] *Biomembrane Structure Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK ;*

[2] *Nestec S.A., BioAnalytical Department, Vers-Chez-Les-Blanc, CH-1000 Lausanne 26, Switzerland*

Key words Atomic Force Microscopy, Molecular Dynamic Simulation, alpha-helical, transmembrane peptide, Membrane, DPPC.

■ **Abstract** WALP23 peptide represents a consensus sequence for transmembrane protein segments, and has special relevance for the packing of helices in proteins such as GPCRs. MD simulation was used to complement Atomic force microscopy (AFM) for the imaging of bilayers containing the peptide and study single peptide unfolding and extraction from the membrane. MD simulations, at equilibrium and where additional forces were applied on selected atoms, have been used to simulate AFM force profile.

INTRODUCTION

Mechanism of protein folding and unfolding are fundamental questions of structural biology (12, 16). With the availability of atomic force microscope and laser tweezer, the question of unfolding protein molecules by external force can be probed (6, 22, 31, 32, 40, 43). Atomic-force microscopy (AFM) and force spectroscopy of single biomolecules allows the measurement of force-extension profile resulting from the application of a stretching force to an individual molecule of biopolymer (33, 40, 41) (reviewed in (42)).

This contribution presents Molecular Dynamic simulations (MD) of WALP membrane peptide in order to mimic AFM experiments within a timescale computationally accessible, in an attempt to

link AFM to molecular models. It aims at describing how stable are membrane-spanning proteins and what are the factors contributing to their stable integration

WALP peptides are synthetic hydrophobic peptides which represent a consensus sequence for transmembrane (TM) protein segments (11) (reviewed in (10)) and have special relevance for packing of α-helices in polytopic proteins such as G-protein coupled receptors. The specific interest of such peptides for AFM studies relies in their ability to form peptides-aggregates in lipid bilayers (44, 45), which are relevant as simplified models for multiple-segments membrane-spanning proteins, since the peptides experience both protein-protein interactions (interactions with other peptides present in the aggregates) and lipid-protein interactions. They are commonly used to study the interactions of α-helices with the surrounding lipid membrane (23, 24) and the anchorage of transmembrane protein segments at the hydrophobic-hydrophilic interface of the bilayer (39), showing the tryptophans residues to preferentially locate at the interface of the lipid headgroups and acyl chains (11).

The surface bilayers containing WALP has been imaged by AFM (44, 45), showing the peptides tend to form aggregates in the membrane, forming characteristic aggregation patterns interpreted as the peptide forming linear assemblies. Force spectroscopy experiments, where the cantilever of the AFM is used to apply force on the object studied, have been run on WALP23 derivatives. Compare to other biological systems studied by AFM, like bacteriorhodopsin (21), WALP offers a simpler topology, and thus easier to interpret. These experiments have shown it is possible to unfold / extract the peptide out of the membrane, suggesting the peptide unfold upon being extracted (17).

In this study, the peptide considered is a 23-meres peptide having a hydrophobic length equivalent to a WALP17 peptide and including at one extremity a cystein residue, located before the tryptophans. The peptide is called WALPC23 and the sequence is:

$^{+}H_3N$–ACAGAWWLALALALALALALWWA–COO^{-}

The extremities are charged and not capped like in previous AFM studies on this type of peptides (44, 45). WALP have already been studied by MD simulation, focusing on the determination of the tilt angle for WALP peptides with various length in various lipid bilayers (38) (DMPC and DPPC).

Molecular Dynamics simulation (MD) is a technique that can be applied complementarily to single-molecule studies and AFM in particular. It has been applied in order to help the interpretation of AFM data regarding the adsorption of proteins on a gold surface (4, 5), as well as to complement protein unfolding studies through the use of a variant of MD, called Steered Molecular Dynamic simulation (SMD), a computational method for studying force-induced reaction in biopolymers (18, 20, 28). This theoretical method mimics atomic force microscopy and optical tweezer experiments, at a time scale accessible with the currently available computational power.

SMD allows identifying structural features of the forced unfolding pathways, which are not provided by presently available measurements. Molecular dynamic simulation describing force-unfolding of protein has been described for protein and peptides in solution (25, 27, 28, 30, 34, 35) and have produced a useful complement to experimental studies (42). Although SMD has been applied to pulling lipids out of a lipid bilayer (29), it has not been applied yet to the study of peptides or proteins embedded in lipid bilayer. Such systems are more complex to study, since not only the contributions of the protein-protein interactions but also the contributions of the lipid-protein interactions have to be considered.

Summarizing our coming results, the MD and SMD simulations presented here were performed with the purpose of providing an atomistic interpretation of AFM-pulling data. Three directions were followed, including (1) providing a molecular model to based on AFM imaging data of a bilayer containing WALP23 (44, 45), (2) using SMD to simulate as accurately as possible a single pulling event of a membrane-spanning peptide from its bilayer environment and (3) providing a systematic study of the AFM-pulling parameters and a qualitative interpretation of their effects on the AFM spectra, based on molecular models.

MATERIAL AND METHODS

MD simulations

MD and SMD simulations are performed using GROMACS v3.1.4 (2, 26), which allows applying forces along Z-axis of the simulation box. The force field used is gromos43A2 for the peptides, extended to improve the simulation of the lipid components (3). Simulations are run at 300 K in an isothermal-isobaric ensemble (NPT). Periodic boundaries are present and a Berendsen temperature- and pressure-coupling scheme is chosen to keep these parameters constant. The time step for the simulation is 2 fs. A LINCS algorithm is used to maintain the geometry of the molecules. Long-range electrostatic interactions are calculated with the particle-mesh Ewald (PME) method (electrostatic interaction up to 9 Å are treated in direct-space and beyond in the Fourier-space). Lennard-Jones interactions are cut off at 14 Å. The simple point charge (SPC) water model (1) is used to describe the water in the simulation box.

Peptides

The peptide coordinates were generated using Swisspdbviewer (47) under the hyptothesis it forms an ideal α-helix. The side-chains of the tryptophan residues were oriented so that the N-H group points away from the hydrophobic core of the membrane. The monomer was subjected to steepest-descent energy minimization. The model consists in a row of five peptides (cf AFM imaging (44, 45)) aligned in alternating orientation (anti-parallel orientation (10, 48)) was chosen as to model the WALPC peptide in a bilayer. The peptide-row is then subjected to steepest-descent energy minimization before the insertion into a bilayer takes place. The peptides were positioned manually so that their initial orientation relative to the membrane is distributed beween 0° and 10° (10).

Lipid bilayer

We have started with a lipid bilayer based on previous published work (36, 37). The bilayer consists in a DPPC patch in the gel-phase including initially 128 lipids and 3655 water molecules and previously equilibrated in a 100 ns MD simulation (area per lipid = 0.645 ± 0.010 nm^2) (available at http://www.lce.hut.fi/research/polymer/downloads.shtml).

Insertion of protein into bilayer

The peptides are being inserted in the bilayer using a method (15) aiming to generate a suitable cavity in the interior of the lipid-bilayer, using the solvent-accessible surface of the protein as a template during the course of a short steered molecular dynamics simulation (SMD) of a solvated lipid membrane (1000 ps SMD run). This is achieved by a three-stages process: first, overlapping water molecules are removed; then lipid molecules whose headgroups are located within 1.5 Å from the peptide-row surface are removed; the peptide-lipid interface is optimized by applying repulsive forces perpendicular to the peptide-row surface, to the remaining lipid atoms inside the volume occupied by the protein surface until it is emptied, leaving 100 DPPC lipids and 3646 water molecules in the box (17'053 atoms; Figure 1). After this, the peptide-row itself may then be inserted, leading to a 1:20 peptide-lipid system. The system is being energy-minimized at each step involving the addition or the removal of any molecular species (steepest descent algorithm).
Finally, the system is being equilibrated in successive short MD simulations (6 runs of 200 ps each) were positions restraints are applied on the peptides and progressively decreased to zero.

The system was then either run for a 20ns-MD simulation in order to study the behaviour of the protperties of the peptide-row model, or subjected to shorter steered MD simulations where forces were applied on the cystein residues of one of the peptides.

Steered MD simulation

Forces are applied on the peptide through a virtual cantilever moving at a constant velocity along an axis perpendicular to the membrane plane and passing through the peptide being pulled. One of the extremities of the cantilever is bound to the peptide and the other is moved away from the peptide along the pulling-axis. The virtual cantilever is a harmonic spring characterized by a spring constant defining its stiffness. It does not have any shape nor undergoes any hydrodynamic drag in the simulation. The set-up is described in Figure 2, which provides a schematic representation of the set-up used for the AFM-pulling simulation leading to the computation of a force-extension profile similar to the one experimentally measured.

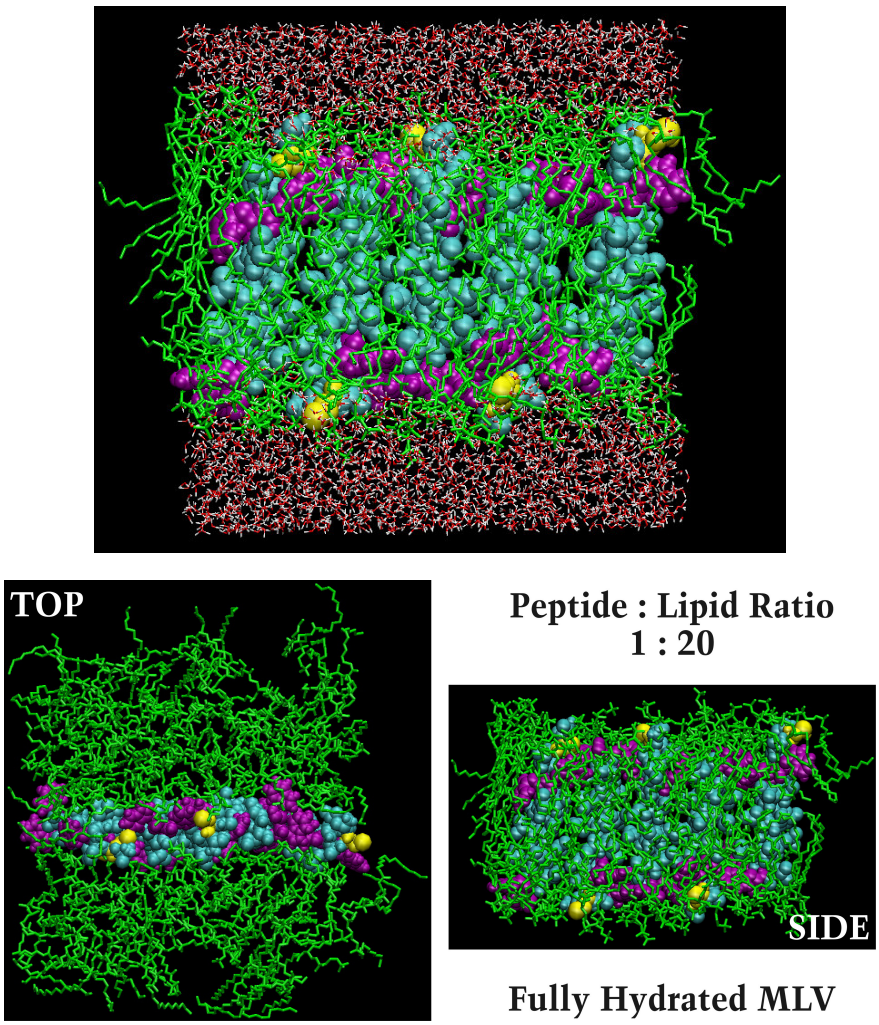


Figure 1: Illustration of the simulation box used for the equilibration of the peptide row and the 20ns-MD simulation. The peptide row is shown to highlight the presence of 5 WALPC peptides. DPPC lipids: green; Peptide: Cyan; Tryptophan residues: Purple; Cystein residues: Yellow; Water: Red & White.

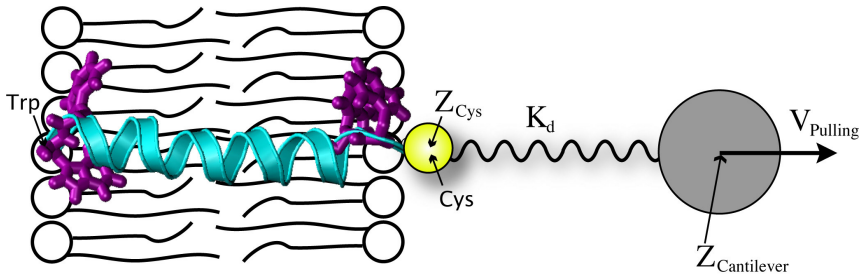


Figure 2: Schematic representation of a single peptide molecule subjected to the force generated by a virtual spring to which it is linked and where the other extremity of the spring is moving at a constant speed. The molecule undergoes an increasing pulling force until it readjusts its position relative to the virtual spring. This set-up is used to simulate AFM-experiments in which a single peptide is being extracted from a lipid bilayer.

Various cantilever stiffness – 16.67 N/m, 1.67 N/m, 0.17 N/m and 0.07 N/m – and pulling-rate – 1.25 Å/ps, 0.125 Å/ps, 0.0125 Å and 0.0063 Å/ps – were simulated, as well as up to three different peptides were pulled in order to gain some insight for the intrinsic variability of the model. This lead to steered MD simulations ranging from 70 ps to 17ns in which the AFM-pulling of one peptide is simulated. For these particular simulations, a larger box was required to simulate the extraction of the peptide out of the membrane. The lipid-peptide equilibrated bilayer was transferred in a larger simulation box containing 12'070 molecules of water (42'325 atoms; Figure 3). This new simulation box was subjected to steepest-descent energy minimization before the AFM-pulling simulation takes place.

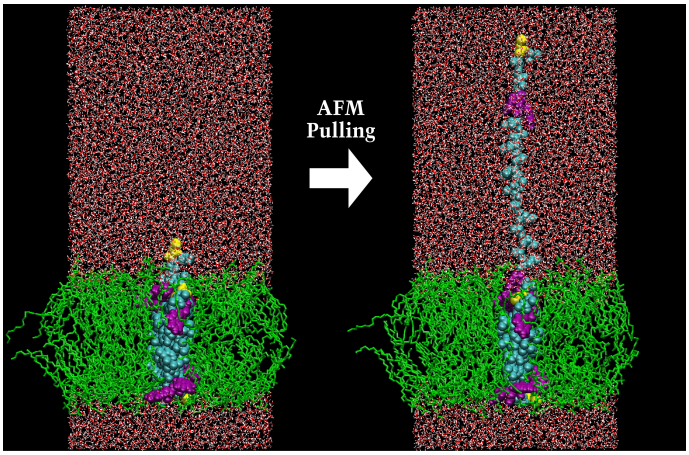


Figure 3: Illustration of the simulation box used for AFM-pulling simulation, before and after the pulling. The peptide row is shown from profile. DPPC lipids: green, Peptide: Cyan, Tryptophan residues: Purple, Cystein residues: Yellow, Water: Red & White.

The figures were generated either with VMD (19) for the molecular views or xmgrace for the graphics.

RESULTS AND DISCUSSION

Simulation of WALPC in a gel-phase DPPC bilayer

Assessment of the molecular model

The molecular model for the aggregation of the peptides was based on all the experimental data available, including the composition of the samples used for AFM study (one peptide for 20 DPPC

lipids), as well as spectroscopic data regarding the relative orientation of interacting peptides, where it was shown that WALP peptide dimmers aggregate forming anti-parallel dimmers (10, 48). It also includes information from AFM imaging studies (44, 45) of WALP peptides, which have shown the presence of line patterns suggesting the formation of peptide-lines separated by a few lipids. Therefore, a row of five peptides aligned in alternating orientation was chosen as to model the WALPC peptide in a bilayer.

In order to test the model was properly equilibrated, a 20 ns MD simulation was set-up. A visual inspection of the simulation box (Figure 4) indicates that the peptide row remains intact during the simulation and no distortion in its shape. The bilayer remains intact, neither being distorted. The temperature, the root-mean-square deviation (RMSD) of the system and the secondary structure of the peptides were monitored during the simulation as an indication of convergence of the hydrated lipid-protein system. The temperature average was constant at 299.85 ± 1.64 K and the RMSD (Figure 5) shows an increase from 0 to 0.3 nm before oscillating between 0.25 and 0.35 nm. This range of value, which could be considered high for the crystal structure of a globular protein, is reasonable for a system made of five peptides free to move and change their relative orientation within the bilayer. Another features indicating the system is stable is that the secondary structure of all the peptides remained unchanged on the 20 ns time-scale of the calculation (Figure 6).

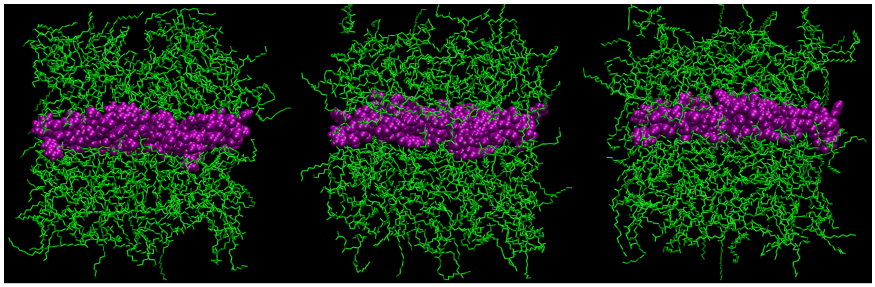


Figure 4: Snapshots of the 20ns-MD simulation of WALPC, illustrating the stability of the peptide-row model on a 20 ns time-scale.

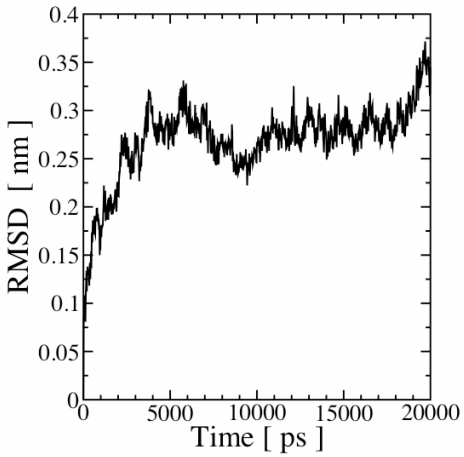


Figure 5: RMSD of the WALPC peptide-row. The relatively high values are due to the intrinsic flexibility of the peptide aggregates, where the peptides are free to move relatively to each other, resulting in the frequent reorientation of the peptides relative to the membrane plane.

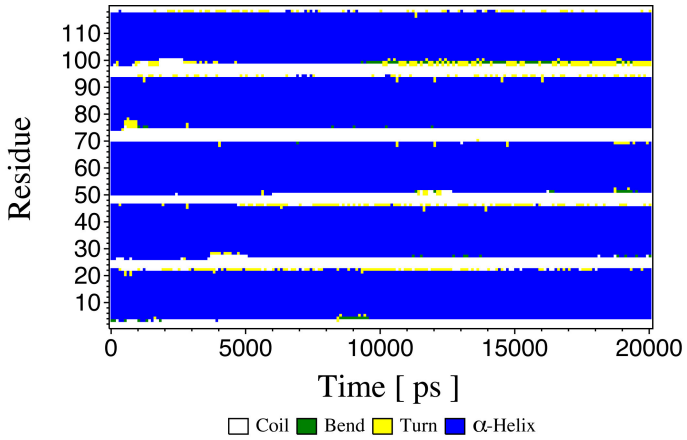


Figure 6: Secondary structure of the five WALPC peptides simulated over 20ns, showing five stable α-helical structures, one for each peptide, which is 23 residues long. The extremities of the helices tend to be more mobile and not showing any secondary structure.

Orientation of the transmembrane peptide in the bilayer

In order to illustrate the mobility of the peptides in the aggregate, the tilt angle of the peptide located in the middle of the row (peptide on which most the AFM-pulling are simulated from) is displayed as a function of time in Figure 7, showing the α-helix can change its orientation up to 20° relative to an axis perpendicular to the bilayer plane within 20 ns. It also that the average tilt angle is 10.7 ± 3.8°, which is a reasonable value compare to previous NMR studies on similar peptide in fluid bilayers (8, 9, 46).

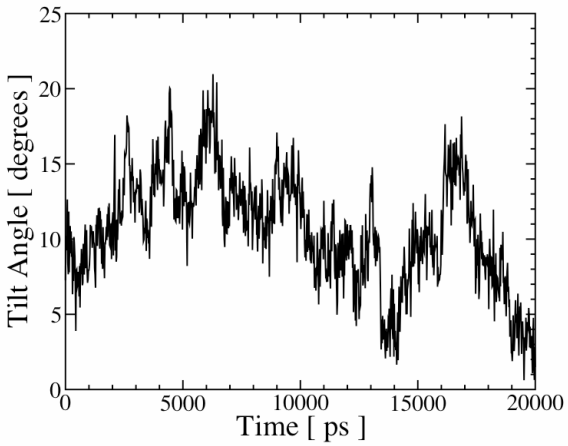


Figure 7: Tilt angle of the peptide located in the middle of the peptide row, illustrating of the flexibility of the peptide-row model (Average: 10.7 ± 3.8°). Other peptides show similar behaviours.

Surface analysis of the bilayer

From previous AFM imaging study of WALP peptide in gel-phase DPPC bilayer (45), a deep comprised between 1 to 3 Å compare to the average membrane surface was observed for WALP23 (slightly deeper for WALP16), suggesting the peptides aggregate or induce changes in the packing of the lipids in the vicinity of the peptides. Similar behaviour is observed for WALPC peptide in DPPC bilayer (Contera et al., personal communication).

Although both the time-scale and size of the system used in this MD would not allow spotting changes in the lipid-phase, it is possible to compare the distances between the average surface of the membrane and the extremities of the peptides. Two different approaches for estimating the averaged-distance were used. The first method is based on a comparison between the coordinates of the lipid phosphate group and the coordinates of the peptides extremities (centre of mass of the last residues). The average distance between the two groups on each side of the membrane is 1.67 ± 0.71 Å. Another way to estimate the distance between the membrane surface and the peptide is to compute the membrane density profile. The density profile shows the distance between the maximum of density for the phosphate groups from the lipids and the maximum of density for the peptide to be ~8 Å, which probably is over-estimating the distance, considered that the maxima of the peptide density profile correspond the average position of the tryptophan residues (the heaviest side-chain in the peptide). On the other hand, the distance between the points where the density

curves for both the phopsphate group and the peptide increase and where it become possible to encounter a member of these group is ~3 Å. Therefore based on the density profile, the distance between the membrane surface and the peptides extremities is less than 8 Å but more than 2 Å, and subjected to the local condition of the sample. Altogether, this leads to a picture compatible with AFM imaging data.

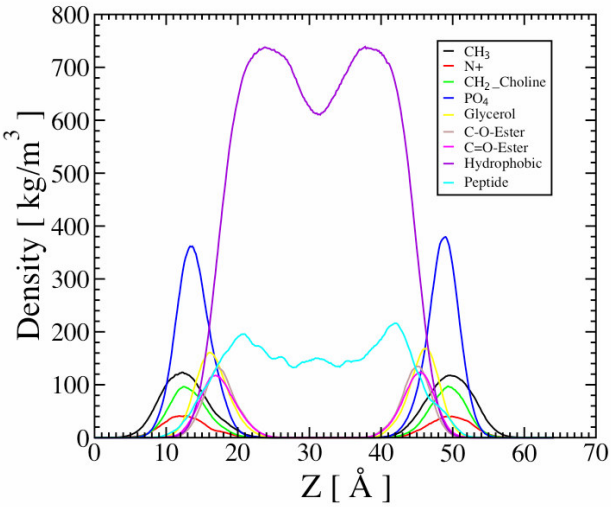


Figure 8: Density profile of the membrane. From this profile, it is expected that the peptides form ridge comprised between 3–8 Å (from the distances between the maximums of the phosphate groups (PO_4) and the peptide distribution curves).

Simulation of a single AFM-pulling curve

Choice of the pulling parameters

The loading rate (in pN/s) is the speed by which the cantilever moves away from the sample multiplied by the spring constant of the cantilever and is commonly used as the parameter use to compare pulling event conducted with different pulling-rate or cantilever stiffness.

In the example presented here, the simulation the unfolding/extraction of a single WALP peptide was achieved using a spring constant of 0.167 N/m for the virtual cantilever stiffness and a pulling rate of 0.0125 Å/ps (or 1.25 m/s or 10^5-10^6 times faster than experimentally), resulting in a loading rate of loading rate = $2.1 \cdot 10^{11}$ pN/s. The cantilever stiffness chosen is similar to the stiffness available experimentally. Furthermore, the pulling-rate ensures the simulation can be completed in a reasonable computation-time. The temperature during the SMD was constant at 299.94 ± 1.06 K.

Experimentally AFM-pulling of WALP23 out of a DPPC bilayer has been studied. The experimental conditions the closest to the simulation, in term of loading rate, are $2.5 \cdot 10^5$ pN/s (cantilever spring constant = ~0.025 N/m and pulling rate = 10 µm/s (17)), giving a force distribution range of 75-1000 pN, with the most likely value at 100 pN. The 7.9ns-MD simulation presented here ($2.1 \cdot 10^{11}$ pN/s) exhibits a maximum force at ~700 pN (Figure 11). The comparison between the maximum forces obtained in the simulations and the experiments shows that, although the timescale between simulation and experiment are very different, they share a similar order of magnitude. This suggests that the 5 to 6 orders of magnitude difference in pulling-rate does not affect the unfolding/extraction pathway of the peptide as much as it could have been anticipated.

Structural analysis

The unfolding and pulling of the peptide out of the membrane was followed using various type of analysis in order to identify the structural events taking place during the AFM-pulling and which influence the force-extension profile and the order in which they take place. A force is applied on the cystein residue of the third peptide of the model (Cys48), leading to the extraction of the peptide.

A visual inspection of the MD trajectory of the pulled-peptide allows identifying various structural events, 15 of them were gathered in Table 1 with a short description and highlighted in the figures illustrating the simulated AFM-pulling.

The root-mean-square deviation (RMSD) (Figure 9) indicates four stages in the pulling of the peptide. From 0 to ~800 ps (RMSD is ~0.1 nm), the force is building and nothing happen to the peptide. Then, from ~800 ps until ~5000 ps, the peptide unfold rather slowly, with its residues on the side which is not directly pulled out of the membrane, remaining immobile (RMSD increases from ~0.1 nm to ~0.6 nm). The slow unfolding is followed by a rapid collapse of the helix, followed by the extension of the peptiditic-chain. The unfolding kinetic is highlighted by the secondary-structure plot (Figure 10). Simultaneously the peptide starts moving slowly through the bilayer from ~5000 ps to ~6000 ps and is characterized by a rise in the RMSD, moving from ~0.6 to ~1.3 nm. The last stage consists in the extended peptidic-chain moving an increasing velocity through the

bilayer. This stage is characterized by a constant RMSD at ~1.3 nm indicating the peptide only undergoes translational or rotational motions leaving its overall conformation rather unaffected (From 6 ns to the end of the simulation).

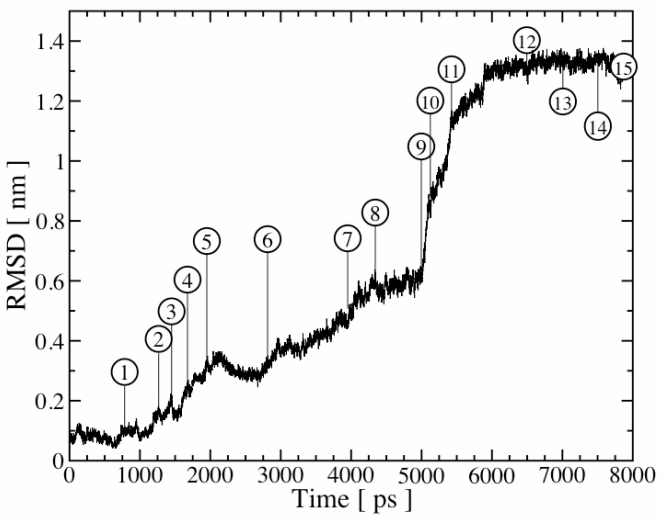


Figure 9: RMSD of the peptide extracted from the peptide bilayer during the simulation of an AFM-pulling with a 0.167 N/m cantilever moving at 0.0125 Å/ps resulting in a 7870ps MD simulations. Some of the structural events are highlighted and described in Table 1.

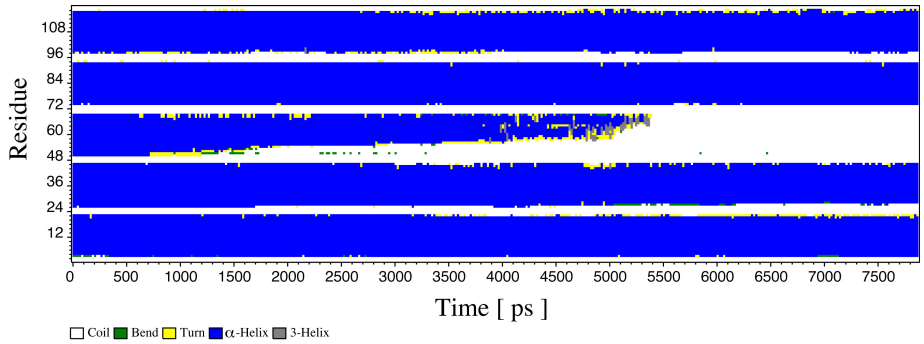


Figure 10: Secondary structure of the peptides during the AFM-pulling simulations, showing four stable α-helical peptides and the progressive unfolding of the pulled-peptide (the peptide in the middle of the peptide-row, residue 47 to 69 on this plot), while its neighbours remain unaffected. The figure shows the acceleration of the unfolding of the helix with time, suggesting that the shorter α-helix is destabilized compare to a larger one.

Table 1: Summary of collision, unfolding and stretching events taking place during the AFM-Pulling. The residue ID are indicated both in reference to the simulation box (5 peptides and 115 residues) and between bracket in reference to the peptide sequence (1 peptide, 23 residues).

#	Time [ps]	Description
1	720	Start of unfolding the helix: hydrogen-bond of residue Ala 49 breaks. From 720 to 800 the unwound residues extend (*Ala3*).
2	1200	Unfolding a second residue: Gly50 (*Gly4*).
3	1450	Unfolding a third residue: Ala51 (*Ala5*).
4	1700	From 1600 ps to 1700 ps, the unwound residues are being extended.
5	1950	Unfolding a fourth residue: Trp52 (*Trp6*).
6	2950	From 2000 ps to 2940 ps, Trp52 (*Trp6*) collides with Trp residue of a neighboured helix, preventing unfolding of the helix. At 2950 ps, the helix unfolds at the level of Trp53 (*Trp7*), whose side-chain still interact with the lipids.
7	4050	Leu54 (*Leu8*) residue unfolds and the side-chain of Trp52 (*Trp6*) residue moves out of bilayer.
8	4800	From 4070 ps to 4800 ps, four more residues unfold from Ala55 (*Ala9*) to Leu58 (*Leu12*).
9	5050	The side-chain of Trp53 (*Trp7*) residue moves out of bilayer and the unfolding rate increases, several residues simultaneously unfolding at the same time.
10	5200	From 5200 ps to 5350 ps, the remaining helix collapses, and the peptide extends quickly.
11	5600	The peptidic chain moves slowly through the membrane, with the Trp67 (*Trp21*) and Trp68 (*Trp22*) residues colliding with the side-chains of the leucine residues of the neighbour helices, at the same time the peptide chain is further extended.
12	7000	The side-chains of Trp67 (*Trp21*) reaches the middle of bilayer and the peptidic chain is fully extended - almost straight.
13	7500	Collision between the side-chains of Trp68 (*Trp22*) from the peptide being pulled and a Trp residue from one of the neighbourhood peptides.
14	7700	After Trp68 (*Trp22*) overtakes the Trp residues of the other peptide, the pulled-peptide starts to move quickly out of the membrane.
15	7870	Peptide moves out of the lipid bilayer.

Figure 11 shows the force-extension profile obtained from the SMD simulations. The x-axis is either the position of the virtual cantilever along the pulling axis (black) or the position of Cys48 (second residue - Cys2 - in the third peptide in the model) on which the pulling-force is applied (red). The y-axis is the pulling-force itself. The force-extension profile shows the presence of two successive maxima. From the trajectory, they can easily be interpreted as the peptide unfolding (z-coordinate 0 to 4 nm for Cys48, or 0 to 7.2 nm for the cantilever, until marker no.11) followed by the extension and the extraction of the peptidic-chain out of the bilayer (z-coordinate 4 to 6.3 nm for Cys48, or 7.2 to 10.3 nm for the cantilever, marker no.14). Smaller peaks can be interpreted as either unfolding event or side-chain collision between the pulled-peptide and its neighbours (Table 1).

The position of the peptide can also be followed as a function of time. Figure 12 illustrates the trajectory followed by both the virtual cantilever (black) and Cys48 (red) along the pulling-axis. The distance between the two curves is directly proportional to the measured force and related by the equation $F = k \cdot \Delta x$, where Δx is the distance between Cys48 and the cantilever along the pulling-axis.

Experimentally, the pulling-length measured for the cantilever is close to 8 nm (7.5 ± 1.5 nm) (17), closer to the force-profile as determined for Cys48 than for the virtual cantilever. The longer path for the cantilever comes from the distortions in its length induced by the fast pulling-rate a systematic study of the influence of the cantilever stiffness shows that the pulling length indicated by the cantilever depends on both its stiffness and the pulling-rate applied (see next section). However, it cannot be shorter than the distance covered by Cys48 during its extraction from the bilayer.

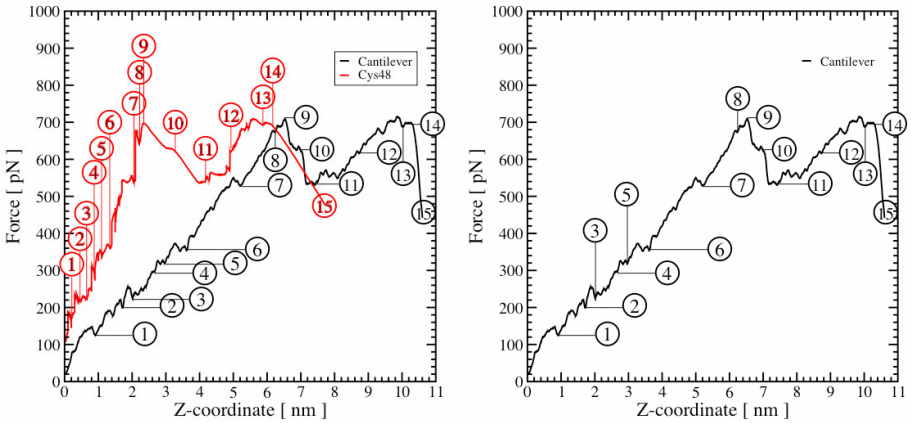


Figure 11: AFM Force-extension profile simulated with a 0.167 N/m cantilever moving at 0.0125 Å/ps resulting in a 7870 ps MD simulations. Some of the structural events are highlighted and describe in Table 1. To eliminate the fast fluctuations, the pulling force was smoothed through averaging (3'9350'000 points were computed during the SMD, finally averaged to 7870 points).

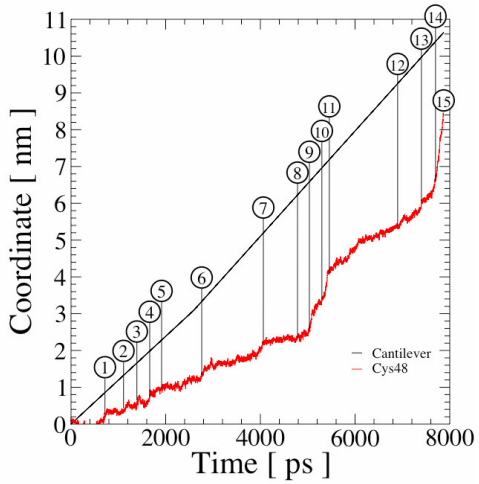


Figure 12: Positions along the pulling axis for Cys48, and the cantilever, during the peptide-extraction. Some of the structural events are highlighted and described in Table 1. The plot shows that a distortion between the position of the cantilever and the cystein-residue. The difference in the position of these two results in the applied force, since the two position are linked through the equation of a harmonic spring.

The trajectory can also be described from the point of view of the peptide extension, which first increase at the rate the peptide is unfolded (up to marker no.9), then extended (from marker no.9 to no.11), before the peptidic chain moves through the bilayer and the stress on the peptide partially relaxed (from marker no.11 to no.15). Interestingly, the maximum extension of the peptide is not reached when the maximum force is applied on Cys48, but a little bit later (between marker no.10 and no.11), indicating a short reaction-delay for the peptide.

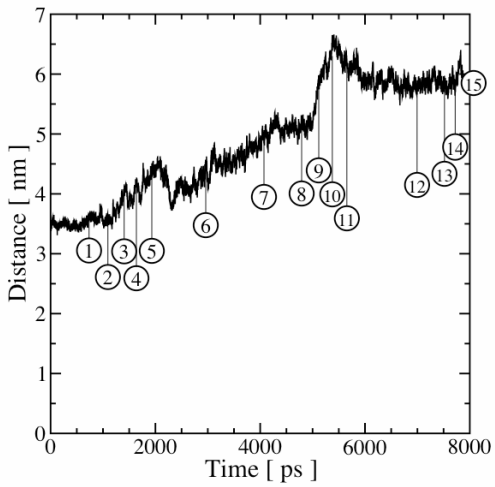


Figure 13: Peptide end-to-end distance during the simulated AFM-pulling. Molecular events are highlighted and described in Table 1. While the coordinate of Cys48 continuously increase during the pulling (see Figure 9), the peptide length stays constant during several time intervals and can even retract (between event 5 and 6). The length of the peptide is ~33 Å long when folded in an a-helix and up to ~67 Å when unfolded and dragged through the bilayer.

Importance of tryptophan residues

In the process of pulling the peptide out, some residues play a more outstanding role than others. This is the case of the tryptophan residues at two stages of the process. First Trp52 (*Trp6*) and Trp53 (*Trp7*) are able to prevent or slow down the unfolding, as described for the observed structural events no.6, no.7 and no.9. The effect is mainly carried out through interaction with the side-chains of other Tryptophan residues from neighbouring peptides and by interacting with polar group of the lipids. In a second stage, Trp67 (*Trp21*) and Trp68 (*Trp22*) play an important role anchoring the unfolded peptidic-chain in the bilayer, slowing it down from moving across the bilayer (see marker no.12, no.13 and no.14). This role is carried again through the interactions with other bulky tryptophan side-chains from the neighbouring peptides and through collision with the leucine side-chains from the neighbouring peptides. This results into either in the need of increasing the time necessary to extract the peptide and thus allowing time for the bulky side-chains to rearrange and find another pathway (as observed in event no.6), or in the need of increasing the applied pulling-force to apply, and thus forcing its way through (as observed in shorted simulations where stronger forces are applied).

Limitation of the technique

As shown in the previous sections, SMD has the ability to simulate a single force-extension profile curve, provided a reasonable molecular model can be provided. The limitation of the technique lays for one part in the fast pulling rate applied and the sample variability, leading to the necessity of repeating the extraction 1000 times or more in order to gain a correct estimation of the most-likely pulling-force – the relevant thermodynamic quantity. At the moment, such a task would require more computational power than it can reasonably be provided in a research laboratory.

Systematic study of AFM-pulling parameters

This section is about the systematic study of the pulling parameters and how they do affect the pulling of the WALPC peptide out of the bilayer. A series of 12 SMDs ranging from 70 ps up to 16.5 ns and covering more than three orders of magnitude for the pulling-rate and the spring constant of the virtual cantilever. The various SMD conditions are summarized in Table 2.

From Table 2, it can be seen that, at constant spring constant, the time needed to simulate the AFM-pulling of the peptide does not increase linearly with the pulling-rate, but slower. This behaviour is observed for the four spring constant (K_d) use in this study.

Table 2: Pulling-time as a function of the spring constant K_d and pulling-rate.

Pulling Time [ps]		K_d [N / m]			
		0.06	**0.167**	**1.667**	**16.67**
Pulling Rate [Å / ps]	**1.25**	310	182	80	70
	0.125	1960	1135	680	670
	0.0125	N.A.	7870	6000	5630
	0.00625	N.A.	N.A.	16465	N.A.

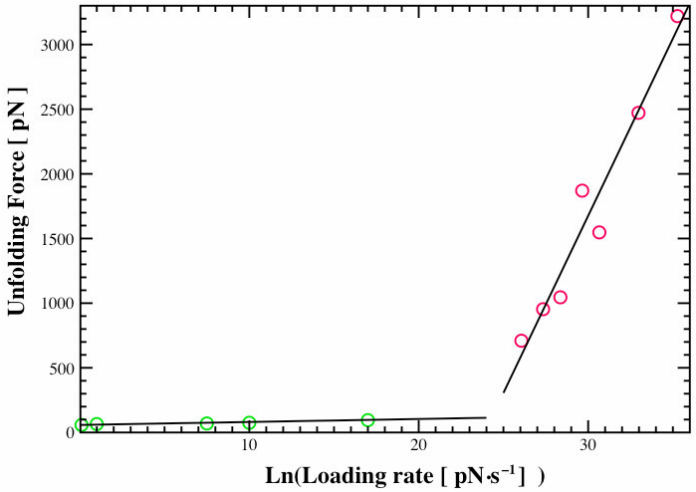


Figure 14: Effect of the loading-rate on the maximum force observed during the peptide unfolding and extraction out of its bilayer. Green: measured on WALP23 in gel-phase DPPC as published in (17) ; Red: calculated on WALPC23 in gel-phase DPPC. The same asymptotic behaviour with a rapid increase of the unfolding force with the loading rate is observed, as previously described for the unfolding of soluble protein (42). The standard deviation for the different points are omitted in the figure, but can range up to ~200 pN for simulations and up to ~12 pN for the experiments. Interestingly the Standard deviation seems to increase together with the loading rate and the unfolding force.

From various combinations of parameters used in the twelve simulations, seven different loading rate can be extracted. Figure 14 (red) displays the maximum force observed in the different simulations (or averaged maximum force when several simulations were conducted for a same loading rate) as a function of the loading rate. For a comparison, the values obtained experimentally for WALP23 with the same gel-phase DPPC (17) are displayed in green on the same figure. The figure suggests two different regimes, with a sharp increase in the maximum force when a certain loading rate level is obtained. The figure shows lot of similarity with a similar figure representing the unfolding of a soluble protein titin-Ig (42). The main difference with the soluble protein is the slope of the asymptote, which seems to results from the properties of the systems themselves.

The results from the systematic study of the effect of the spring constant and the pulling-rate on the calculated force profile are displayed on Figure 15, showing the maximum force observed during the unfolding of the peptide increases with the loading rate applied as reported previously, both experimentally (13, 14, 42) and calculated for soluble proteins (7).

Important distortions between the position of the pulled residue (Cys48) and the virtual cantilever arise are when the amplitude of the applied force increase. For obvious reason related with its k_d, a stiffer cantilever induces less distortion in the force-extension profile than a soft cantilever. This has some importance relative to the information carried through the cantilever, which at high pulling-rate and low spring constant appear to be an almost featureless straight line. Transposed to slower pulling-rates used experimentally, such distortions should be weaker and most probably negligible. Nevertheless, a stiffer cantilever might still provide force-profiles where with more features than a very soft cantilever.

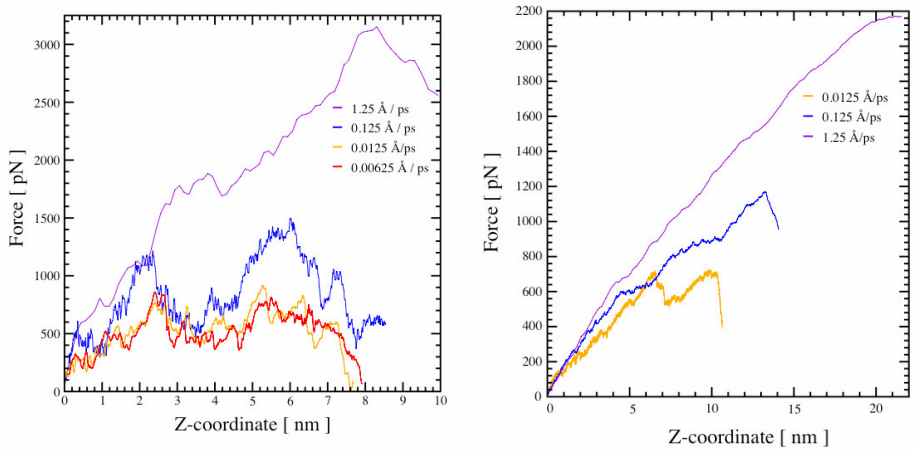


Figure 15: Effect of the pulling-rate on the force-extension profile for two K_d (1.667 N/m and 0.167 N/m).

The intrinsic variability within the molecular model used for the various SMD has been probed for three different peptides out of five (helix II, III and IV) of the peptide-row at three different pulling rate (1.25, 0.125 and 0.0125 Å/ps) with a spring constant of 1.667 N/m, showing a variability of about 10 to 15% for the maximum observed force. Shifts in the position and the force-intensity maxima are observed and due to small differences between the conformation and relative disposition of the peptides, leading to different protein–protein interactions, more specifically, the way the side-chains collide with each other. The positions where the maximum force is observed is also shifted by up to ~1nm. Altogether, the variability obtained by pulling three neighbour-peptides, sharing a similar environment suggests the randomness observed in the force profile comes directly from the randomness of the distribution for the side-chains orientation within the peptides.

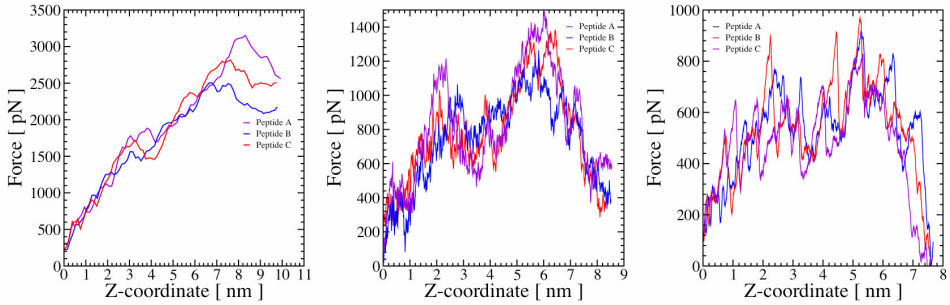


Figure 16: Estimation of the intrinsic sample variability for three different pulling-rates (1.25, 0.125 and 0.0125 Å/ps from left to right) for three out of five peptides present in the peptide-row. K_d is 1.667 N/m.

CONCLUSION

MD and SMD provide a useful insight into the interpretation of atomic force microscopy, both for imaging and pulling. In the case of AFM-pulling experiments, SMD can provide both a qualitative insight in the effect of the pulling-parameters on the force-extension profile measured with the AFM. The limitation of the technique resides in the statistical nature of the non-bonded interactions, which requires the pulling experiments to be repeated by a large number (more than 100), in order to provide a correct estimation of the most probable force necessary to break the interaction. The simulation of such a property would require more computational power than reasonably available at this time. Nevertheless, it provides a useful interpretation of the force-extension profile gathered by atomic force microscopy, suggesting which part of the peptides or process molecular process lead to the force-extension profile as observed.

This contribution supports the conclusion from Ganchev that the bilayer interface region plays an important role in stable integration of transmembrane α-helices in membranes. In addition we suggest that Trp – Trp interactions (side-chain collision and electrostatic interactions) as well as Trp lipid interaction (electrostatic interactions) play a critical role in the anchoring of the peptide in the membrane.

ACKNOWLEDGEMENT Dr Mikko Karttunen and Dr Michael Patra are thanked for making the coordinates of the 100ns-equilibrated DPPC bilayer available. VL thanks BBSRC for support.

LITERATURE CITED

1. Berendsen HJC, Griegera JR, Straatsma TP. 1987. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* 91: 6269
2. Berendsen HJC, van der Spoel D, van Drunen R. 1995. GROMACS: a message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* 91: 43–56
3. Berger O, Edholm O, Jahnig F. 1997. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys J* 72: 2002
4. Bizzarri AR, Bonanni B, Costantini G, Cannistraro S. 2003. A combined atomic force microscopy and molecular dynamics simulation study on a plastocyanin mutant chemisorbed on a gold surface. *Chemphyschem* 4: 1189
5. Bizzarri AR, Costantini G, Cannistraro S. 2003. MD simulation of a plastocyanin mutant adsorbed onto a gold surface. *Biophys Chem* 106: 111
6. Carrion-Vazquez M, Marszalek PE, Oberhauser AF, Fernandez JM. 1999. Atomic force microscopy captures length phenotypes in single proteins. *Proc Natl Acad Sci U S A* 96: 11288
7. Cheng F, Shen J, Luo X, Jiang H, Chen K. 2002. Steered molecular dynamics simulations on the "tail helix latch" hypothesis in the gelsolin activation process. *Biophys J* 83: 753
8. de Planque MR, Bonev BB, Demmers JA, Greathouse DV, Koeppe RE, 2nd, et al. 2003. Interfacial anchor properties of tryptophan residues in transmembrane peptides can dominate over hydrophobic matching effects in peptide-lipid interactions. *Biochemistry* 42: 5341
9. de Planque MR, Goormaghtigh E, Greathouse DV, Koeppe RE, 2nd, Kruijtzer JA, et al. 2001. Sensitivity of single membrane-spanning alpha-helical peptides to hydrophobic mismatch with a lipid bilayer: effects on backbone structure, orientation, and extent of membrane incorporation. *Biochemistry* 40: 5000
10. de Planque MR, Killian JA. 2003. Protein-lipid interactions studied with designed transmembrane peptides: role of hydrophobic matching and interfacial anchoring. *Mol Membr Biol* 20: 271
11. de Planque MR, Kruijtzer JA, Liskamp RM, Marsh D, Greathouse DV, et al. 1999. Different membrane anchoring positions of tryptophan and lysine in synthetic transmembrane alpha-helical peptides. *J Biol Chem* 274: 20839
12. Dobson CM, Sali A, Karplus M. 1998. Protein Folding: A Perspective from Theory and Experiment. *Angew. Chem. Int. Ed. Eng.* 37: 868–893
13. Evans E. 1998. Energy landscapes of biomolecular adhesion and receptor anchoring at interfaces explored with dynamic force spectroscopy. *Faraday Discuss*: 1
14. Evans E, Ludwig F. 2000. Dynamic strength of molecular anchoring and material cohesion in fluid biomembranes. *J. Phys.: Condens. Matter* 12: A315–A320
15. Faraldo-Gomez JD, Smith GR, Sansom MS. 2002. Setting up and optimization of membrane protein simulations. *Eur Biophys J* 31: 217
16. Fersht AR. 1999. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. New York: Freeman
17. Ganchev D, Rijkers DTS, Snel MME, Killian JA, de Kruijff B. 2004. The strength of integration of transmembrane α-helical peptides in lipid bilayers as determined by Atomic

Force Spectroscopy. *Biochemistry* In Press
18. Grubmüller H, Heymann B, Tavan P. 1996. Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science* 271: 997
19. Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. *J Mol Graph* 14: 33
20. Izrailev S, Crofts AR, Berry EA, Schulten K. 1999. Steered molecular dynamics simulation of the Rieske subunit motion in the cytochrome bc(1) complex. *Biophys J* 77: 1753
21. Janovjak H, Kessler M, Oesterhelt D, Gaub H, Muller DJ. 2003. Unfolding pathways of native bacteriorhodopsin depend on temperature. *Embo J* 22: 5220
22. Kellermayer MS, Smith SB, Granzier HL, Bustamante C. 1997. Folding-unfolding transitions in single titin molecules characterized with laser tweezers. *Science* 276: 1112
23. Killian JA. 1998. Hydrophobic mismatch between proteins and lipids in membranes. *Biochim Biophys Acta* 1376: 401
24. Killian JA, Salemink I, de Planque MR, Lindblom G, Koeppe RE, 2nd, Greathouse DV. 1996. Induction of nonbilayer structures in diacylphosphatidylcholine model membranes by transmembrane alpha-helical peptides: importance of hydrophobic mismatch and proposed role of tryptophans. *Biochemistry* 35: 1037
25. Krammer A, Lu H, Isralewitz B, Schulten K, Vogel V. 1999. Forced unfolding of the fibronectin type III module reveals a tensile molecular recognition switch. *Proc Natl Acad Sci U S A* 96: 1351
26. Lindahl E, Hess B, Van der Spoel D. 2001. GROMACS 3.0: A package for molecular simulation and trajectrory analysis. *J.Mol.Mod.* 7: 306–317
27. Lu H, Isralewitz B, Krammer A, Vogel V, Schulten K. 1998. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys J* 75: 662
28. Lu H, Schulten K. 1999. Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins* 35: 453
29. Marrink SJ, Berger O, Tieleman P, Jahnig F. 1998. Adhesion forces of lipids in a phospholipid membrane studied by molecular dynamics simulations. *Biophys J* 74: 931
30. Marszalek PE, Lu H, Li H, Carrion-Vazquez M, Oberhauser AF, et al. 1999. Mechanical unfolding intermediates in titin modules. *Nature* 402: 100
31. Mehta AD, Rief M, Spudich JA. 1999. Biomechanics, one molecule at a time. *J Biol Chem* 274: 14517
32. Mehta AD, Rief M, Spudich JA, Smith DA, Simmons RM. 1999. Single-molecule biomechanics with optical methods. *Science* 283: 1689
33. Oberhauser AF, Marszalek PE, Erickson HP, Fernandez JM. 1998. The molecular elasticity of the extracellular matrix protein tenascin. *Nature* 393: 181
34. Paci E, Karplus M. 1999. Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations. *J Mol Biol* 288: 441
35. Paci E, Karplus M. 2000. Unfolding proteins by external forces and temperature: the importance of topology and energetics. *Proc Natl Acad Sci U S A* 97: 6521
36. Patra M, Karttunen M, Hyvonen MT, Falck E, Lindqvist P, Vattulainen I. 2003. Molecular dynamics simulations of lipid bilayers: major artifacts due to truncating electrostatic interactions. *Biophys J* 84: 3636
37. Patra M, Karttunen M, Hyvonen MT, Falck E, Vattulainen I. 2004. Lipid Bilayers Driven to a Wrong Lane in Molecular Dynamics Simulations by Subtle Changes in Long-Range Electrostatic Interactions. *J. Phys. Chem. B* 108: 4485–4494
38. Petrache HI, Zuckerman DM, Sachs JN, Killian JA, Koeppe II RE, Woolf TB. 2002. Hydrophobic Matching Mechanism Investigated by Molecular Dynamics Simulations. *Langmuir* 18: 1340–1351
39. Reithmeier RA. 1995. Characterization and modeling of membrane proteins using sequence analysis. *Curr Opin Struct Biol* 5: 491
40. Rief M, Gautel M, Oesterhelt F, Fernandez JM, Gaub HE. 1997. Reversible unfolding of individual titin

immunoglobulin domains by AFM. *Science* 276: 1109

41. Rief M, Gautel M, Schemmel A, Gaub HE. 1998. The mechanical stability of immunoglobulin and fibronectin III domains in the muscle protein titin measured by atomic force microscopy. *Biophys J* 75: 3008
42. Rief M, Grubmüller H. 2002. Force spectroscopy of single biomolecules. *Chemphyschem* 3: 255
43. Rief M, Oesterhelt F, Heymann B, Gaub HE. 1997. Single Molecule Force Spectroscopy on Polysaccharides by Atomic Force Microscopy. *Science* 275: 1295
44. Rinia HA, Boots JW, Rijkers DT, Kik RA, Snel MM, et al. 2002. Domain formation in phosphatidylcholine bilayers containing transmembrane peptides: specific effects of flanking residues. *Biochemistry* 41: 2814
45. Rinia HA, Kik RA, Demel RA, Snel MM, Killian JA, et al. 2000. Visualization of highly ordered striated domains induced by transmembrane peptides in supported phosphatidylcholine bilayers. *Biochemistry* 39: 5852
46. Strandberg E, Ozdirekcan S, Rijkers DT, Van Der Wel PC, Koeppe RE, 2nd, et al. 2004. Tilt Angles of Transmembrane Model Peptides in Oriented and Non-Oriented Lipid Bilayers as Determined by ^{2}H Solid-State NMR. *Biophys J* 86: 3709
47. SwissPDB-Viewer. http://ca.expasy.org/spdbv.
48. Yano Y, Takemoto T, Kobayashi S, Yasui H, Sakurai H, et al. 2002. Topological stability and self-association of a completely hydrophobic model transmembrane helix in lipid bilayers. *Biochemistry* 41: 3073

II.4 Interaction of amiloride and its derivatives with Vpu from HIV-1: a MD simulation study

Vincent Lemaître,[1,2] Rehan Ali, [1] Chris G. Kim,[1] Anthony Watts,[1] and Wolfgang B. Fischer,[1,3]

[1] *Biomembrane Structure Unit, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, UK*
[2] *Nestec S.A., BioAnalytical Science Department, Vers-Chez-Les-Blanc, CH-1000 Lausanne 26, Switzerland*
[3] *Bionanotechnology Interdisciplinary Research Consortium, Clarendon Laboratory, Department of Physics, Oxford University, Parks Road, Oxford OX1 3SU, UK,* e-mail: wolfgang.fischer@bioch.ox.ac.uk

Keywords Viral ion channel, Vpu, HIV-1, molecular dynamic simulation, drug-protein interaction, amiloride, hexamethyleneamioride.

■ **Abstract** Vpu is an 81-residue membrane-protein with a single transmembrane segment that is encoded by HIV-1 and is involved in the enhancement of virion release via formation of an ion channel. Cyclohexamethylene amiloride (Hma) has been shown to inhibit ion channel activity. In the present 12 ns simulation study a putative binding site of Hma blockers in a pentameric model bundle built of parallel aligned helices of the first 32 residues of Vpu was found near Ser-23. Hma orientates along the channel axis with its alkyl ring pointing inside the pore, which leads to a blockage of the pore.

INTRODUCTION

Vpu is an 81 amino acid protein encoded by HIV-1 with a high degree of sequence conservation (47). Its function in the life cycle of HIV-1 is twofold: (i) to interact with CD4 in the endoplasmic reticulum to initiate the ubiquinine-mediated degradation of the CD4–Vpu complex and (ii) to enhance particle release at the site of the plasma membrane altering the electrochemical gradient via ion channel formation by homo–oligomerisation (33) in the lipid membrane. Whilst the first function is fairly established, the second is still open to debate (28). It is still not unambiguously established whether ion channel activity is an intrinsic property of Vpu. If so, then a selective

blocker needs to be found to give proof, which has been reported using the derivatives of amiloride (Am) (16). There are reasonably good structural data available to permit generation of molecular models of Vpu (for an overview see (8, 18)), in which there is a helical transmembrane (TM) domain followed by a larger cytoplasmic domain with a helix–loop–helix–helix/turn motif. Such models can be used in molecular dynamics simulations to investigate the mechanism of function and also to initiate computational blocker (leading possibly to drug) screening and development.

This study is based on recent findings cyclohexamethylene amiloride (Hma) but not amiloride (Am) itself, blocks channel activity of both a peptide corresponding to the TM segment of Vpu and full-length Vpu when reconstituted into lipid bilayers (16). A homo-pentameric bundle based on the first 32 amino acids of Vpu in a helical motif has been generated. Simulations in the presence of the blocker aim to shed light on: (i) what is the site of action for the blocker with the protein, (ii) whether the blocker affect the integrity of the bundle, and (iii) what is the most favourable conformation adopted by the blocker. Calculations have been performed with Am as the non-potent blocker for comparison. Two different protonation states for each blocker have been simulated.

MATERIALS AND METHODS

Model building

A pentameric Vpu TM bundle based on the first 32 amino acids of Vpu (HV1H2), QPIPIVAIVA^{10} LVVAIIIAIV^{20} VWSIVIIEYR^{30} KI, was generated with the program X-Plor 3.1 (9) based upon a protocol described in detail elsewhere (11, 26). This protocol generates symmetrical arrangements of five α-helices and was based on a simulated annealing (SA) procedure with short molecular dynamics (MD) simulations. The initial inter-helical distance was set to 0.94 nm. During a first stage, which involves the simulated annealing, five structures were generated based on van der Waals interactions. In the next stage (stage 2) each of these five structures was used as a starting point for the consecutive steps, which take electrostatic interactions into account. At the end 25 structures were obtained. Inter-helical root mean square deviation (RMSD) values were calculated

for each bundle and the bundle with the lowest average RMSD (i.e. the most symmetrical) was selected.

The structures of Am (3,5-diamino-6-chloro-N-(diaminomethylene)-pyrazinecarboxamide) and Hma (5-(N,N-hexamethylene) amiloride) in both a deprotonated (AM, HMA) and a protonated state (proton at the guanidinium group, AM^+, HMA^+), were generated by applying DS Viewer Pro 5.0 (Accelerys) and were consequently used in the AutoDock software (34) to locate the putative binding site of the blockers with Vpu. In total 10 runs were performed with 30'000 energy minimization steps. Thereby the conformation of the bundle was kept rigid, whilst the blocker was allowed to be flexible. All structures located the energetically most favourable site within the pore close to the serine residues. From a total of 200 putative docking conformations, the one closest to one of the serines was chosen for MD simulations. The respective blockers were then placed inside the pore. Overlapping water molecules were removed. The system was then subject to a 300-ps equilibration keeping only the position restraints on the backbone of the protein. The topology of each of the blockers was determined by manually adapting the output of PRODRG (45) for the force field GROMOS43a2.

MD simulations

The *in-silico* bundle was placed in a lipid bilayer consisting of 96 lipid molecules (1-palmitoyl-2-oleoyl-*sn*-glycerol-3-phosphatidyl-choline, POPC) in which a hole, already generated by removing enough lipids to avoid any overlap with the bundle, was present (19). The bilayer–bundle system was consequently solvated with 40 water molecules per lipid to give a system of around 20 000 atoms per experiment. As each helix has a net charge of +1, five chloride ions (six in the presence of the protonated blocker) were also added to compensate for any extra charge within the simulation box. The solvated system was equilibrated for 300 ps with the peptide restrained, followed by the production phase of fully unrestrained 12-ns MD simulation.

For MD Simulations GROMACS v3.1.4 software was used. Simulations were run at 300 K in an isothermal-isobaric ensemble (NPT). Periodic boundaries were present and a Berendsen temperature and pressure coupling were chosen to keep these parameters constant. The time step for the simulation is 2 fs. A LINCS algorithm is used to maintain the geometry of the molecules. Long-

range electrostatics were calculated with the particle-mesh Ewald (PME) method using standard GROMACS parameters: grid dimensions 0.12 nm, interpolation order 4. Double counting correction for short-range forces was applied. Lennard-Jones and short-range coulomb interactions were cut off at 1.1 and 0.9 nm, respectively. The simple point charge water model (7) was used to describe the water in the simulation box. Simulations were performed on a Dell Precision 420 with a dual Pentium III 1GHz processor or on a Dell Precision 330 with a Pentium IV 1.5 GHz. The pore radii were calculated using HOLE2 (41) with radii not being considered if larger than 1.0 nm.

RESULTS

The site of Am–Vpu interaction and that of its derivative Hma with Vpu has been elucidated by a docking approach. Independent of the number of TM helices the most likely binding site proposed for both blockers is around Ser-23 (Fig. 1, top pictures). At the starting structure, the guanidinium group of each of the blockers is pointing towards the serine residues of the bundle. At the end of the simulations the blockers remain at the site of Ser-23. A view down the pore from the C- to the N-terminus reveals that both Am (Fig. 1A) and Hma (Fig. 1B) stay in contact with two helices most of the time (Fig. 1, bottom pictures). Independent of the protonation state, Am seems to allow more space within the pore for water or ions to pass through than Hma.

The protein

The RMSD of the backbone atoms for each of the simulated pentamers, without blocker and in the presence of either AM^+ (Fig. 2A) or HMA^+ (Fig. 2B), remains below 0.35 nm with respect to their initial conformation. The presence of AM and AM^+ does not have any impact on the RMSD. In both cases the curve can hardly be distinguished from the bundle without blocker. In the presence of HMA the bundles shows some deviation from the bundle without blocker. Over a time span of the first 6 ns the values for the bundle with either of the blockers are slightly higher than the bundle without blocker. After another 6 ns the bundle with the HMA^+ remains approximately 0.05 nm

higher than the values for the bundle without blocker. The RMSD for the bundle with the HMA approaches the values of the plain bundle.

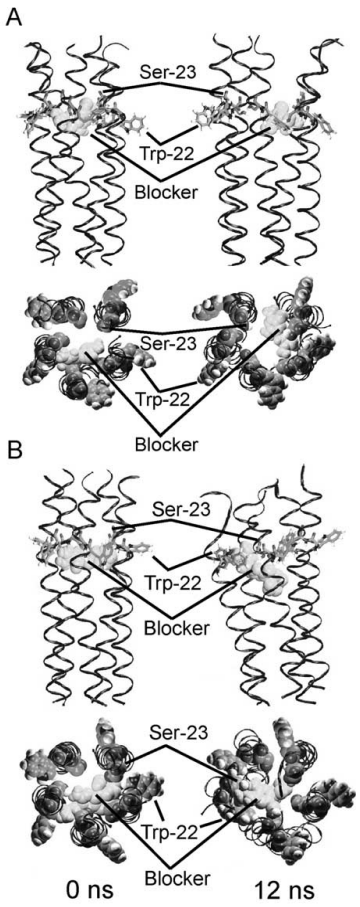


Figure 1 Representation of the pentameric bundle of the TM part of Vpu in the presence of AM^+ (light grey, van der Waals representa-tion) in a side view (upper panel) and view from the C-terminus into the pore (lower panel) (A). The peptide bundles are shown as ribbons (black). Tryptophans and serines are highlighted as sticks (grey scale is used). Lipid and water molecules are omitted for clarity. HMA^+ (light grey) is shown in the same way (B).

The averaged root mean square fluctuation (RMSF) per residue (data not shown) of the bundle is indicative of a larger deviation of more than 0.3 nm for residues on the C-terminal end than on the N-terminal end. This is in common for all the simulations independent of the presence of the blocker and its protonation state and indicative of a loss of helicity (DSSP analysis, data not shown) (14). For the membrane spanning amino acids, the fluctuation is below 0.15 nm.

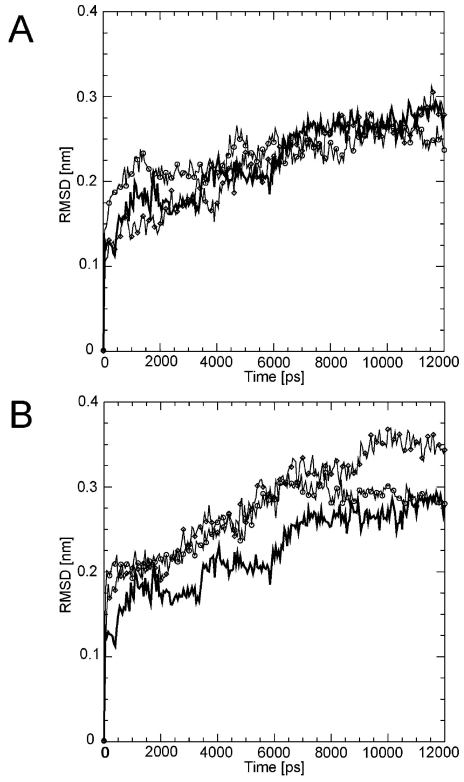


Figure 2 RMSD of the Cα traces for the pentameric bundle without blocker (black line), and in the presence of protonated (diamonds) and deprotonated (circles) blockers. Amiloride (A) and HMA (B).

The presence of either of the blockers in the deprotonated state does not have a major impact on the overall structure of the bundle. The averaged tilt angle for the bundle without blocker is about 5.8 ± 2.7°, whilst the values for the bundles in the presence of the deprotonated blockers do not exceed 11.0 ± 4.1°. AM^+ and HMA^+ induce larger averaged tilt angles of 19.6 ± 3.5° and 14.9 ± 7.5°, respectively. Changes for the average kink-angle and cross-angle are small but significant (with a confidence interval of 99.99 %) indicating that the blockers do affect the properties of the bundle. The average inter-helical distance of the helices does not appear to change in the presence of any of the blockers.

The pore is divided into three areas to refine the analysis of the effect of the blockers on the pore, the N-terminal (residues 1–11), middle (residues 12–22) and the C-terminal sections (residues 23–32). Time-averages for both the averaged minimum radii have been computed for each of the sections and are displayed in Table 1. Simulation of the bundle without any blocker reveals a

minimum pore radius of 0.12 ± 0.06 nm for the C-terminal section. The simulation with AM^+ shows an opening of the bundle, with an average radius of 0.21 ± 0.05 nm. The same calculation, taking into account the presence of the blocker and the consequent obstruction of the pore resulting from it, allows computation of the apparent pore radius and reveals a slightly smaller radius of 0.20 ± 0.06 nm. HMA^+ also induces a larger pore radius at this section of the bundle with values of 0.19 ± 0.06 nm (and for the apparent pore radius 0.18 ± 0.06 nm). The middle section remains almost unaffected by the blockers. Average radii of 0.18 ± 0.03 nm (AM^+) and 0.17 ± 0.03 nm (HMA^+) are close to the value for the bundle without any blocker at 0.18 ± 0.03 nm. However, the presence of the blockers decreases the apparent radii to 0.14 ± 0.05 nm (for AM^+) and 0.06 ± 0.04 nm (for HMA^+). The pore radius in the presence of HMA^+ would not even allow for a single-file water passage as found in gramicidin A (15, 39). Thus, HMA^+ completely blocks the pore. At the N-terminal section, an average pore radius of 0.11 ± 0.04 nm is calculated without the presence of the blocker in the pore. The presence of AM^+ at the binding site is accompanied by a widening of up to 0.17 ± 0.04 nm for the N-terminal section. In the presence of HMA^+ the pore narrows (0.13 ± 0.05 nm) compared to the simulations with an empty bundle.

Simulations with the blockers in a neutral state are also performed. Although this state is unlikely to happen in nature, it allows us to gain more insight in the blocker–protein interactions by performing 'impossible' virtual experiments. Removing the electrical charge on the guanidinium group (deprotonated blockers) leads to an overall narrowing of the pore radii in all three sections (Table 1). The decrease of the pore opening is more pronounced for the simulations with AM, particularly in the middle section with a radius: 0.14 ± 0.03 nm. For HMA, there is a widening of the pore radius in the middle section (0.06 ± 0.04 nm C-terminal section, 0.16 ± 0.04 nm middle section, 0.09 ± 0.04 nm N-terminal section). In the middle section the apparent radii decrease to 0.08 ± 0.03 nm and 0.07 ± 0.03 nm for AM and HMA, respectively. The two deprotonated blockers result in a blocking of the C-terminal part with almost identical apparent radii of ~ 0.05 nm.

To summarize the results, AM^+ and to a lesser extent HMA^+, lead to a widening of both the C- and N-terminal sections. In the presence of HMA^+, the diameter of the N-terminal section is reduced

compared to the simulations of Vpu bundle in the presence of AM^+. HMA^+ induces a 'funnel' like shape in the pore, in addition to an almost complete obstruction of the middle section of the pore. Simulating the unrealistic situation, where charges on the blockers are removed, leads to an overall narrowing of the pore. The blocking resulting from the presence of HMA seems to occur mainly by occlusion of the pore. These changes, although relatively small, are found to be significant at a confidence level of 99.99%.

Water molecules are found only within the pore at the C-terminal end approaching the blockers during the entire length of the simulations. Towards the N-terminal end, for which the pore is lined by highly hydrophobic residues, no water molecules can be detected. The water molecules, which are present when each of the models is built, escape from the pore during the 300-ps equilibration step and do not re-enter the pore during the simulation.

Table 1 Averaged minimum pore radii of the pentameric bundle when divided into three sections: the C-terminal (C), middle (M) and N-terminal (N) sections.

	C(nm)	M (nm)	N (nm)
Vpu	0.12 ± 0.06	0.18 ±0.03	0.11 ±0.04
Bundle (+AMþ)	0.21 ±0.05	0.18 ±0.02	0.17 ±0.04
Bundle+AMþ	0.20 ±0.06	0.14 ±0.05	0.17 ±0.04
Bundle (+HMAþ)	0.19 ±0.06	0.17 ±0.03	0.13 ±0.05
Bundle+HMAþ	0.18 ±0.06	0.06 ±0.04	0.13 ±0.06
Bundle (+AM)	0.05 ±0.03	0.14 ±0.03	0.09 ±0.03
Bundle+AM	0.05 ±0.03	0.08 ±0.03	0.09 ±0.03
Bundle (+HMA)	0.06 ±0.04	0.16 ±0.04	0.09 ±0.04
Bundle+HMA	0.05 ±0.04	0.07 ±0.03	0.08 ±0.04

Each segment is of equal length (approximately 12 amino acids). Minimum pore radii were averaged over the entire length of each simulation. The apparent radii are calculated with (e.g. bundle+HMA^+) and without (e.g. bundle (+HMA^+)) considering the blockers in the pore.

The blocker

The RMSD values for AM and AM^+ (data not shown) do not exceed 0.15 nm within the duration of the simulation. The RMSD values for the HMA and HMA^+ remain around 0.15 nm, with a larger fluctuation for the latter (0.1 nm < RMSD < 0.2 nm). The larger values and spread are due to the flexible hexamethylene ring of the AM derivative. These values are within the range of values obtained for K^+ channel blocker toxins simulated in aqueous solution (23). In addition, the RMSFs of the individual atoms of the blockers support this result. The central body of Am (a 3,4,6-substituted pyrazine ring) and Hma remain fairly rigid (RMSF < 0.1 nm) independent of the protonation state of the blockers. Only the hydrogen atoms of the amino groups of the pyrazine ring

show an RMSF slightly above 0.1 nm. For both blockers the hydrogen atoms of the guanidinium group show the largest fluctuations (> 0.1 nm). Overall, the curves are very similar for both blockers when compared to similar atoms with lower values for all atoms in the protonated blocker. The atoms of the hexamethylene ring in Hma are seen to fluctuate around 0.1 nm. The protonation state does not effect the fluctuation of this part of the molecule.

The different conformations generated during the MDs for all the four blockers tested were clustered using a full linkage algorithm, with a cut-off of 0.03 nm for amiloride and 0.04 nm for Hma. In the case of AM, this means that molecules with an RMSD smaller than 0.03 nm relative to all the existing members of a cluster will belong to this cluster. Cluster analysis reveals the most populated conformation for each of the blockers (Fig. 3). AM^+ (Fig. 3A) and AM (Fig. 3B) spend 96.0 % and 87.9 %, respectively, of their time in this conformation. For HMA^+ (Fig. 3C) and HMA (Fig. 3D) the values are 95.1 % and 91.5 %, respectively. For the protonated blockers, the carbonyl group linked to the guanidinium group points towards the primary amine group of the pyrazine ring. Deprotonation of the blockers reveals a conformational change in the blocker so that the amine part of the guanidinium group is now pointing to the primary amine group of the pyrazine ring.

The AM^+ forms on average 1.3 hydrogen bonds with water molecules during the simulation. The threshold for hydrogen bonding was A–H < 2.5 Å and an angle A–H–D < 60°, with A defining the acceptor, D the donor. The number of hydrogen bonds indicates that, on average, more than one hydrogen bond is formed between the blocker and the solvent. About 0.82 hydrogen bonds are formed with the side chains of the serine residues and about 0.19 hydrogen bonds with the backbone of the peptide. Barely any hydrogen bonds are formed with the tryptophan side chains. For the AM, the total number of hydrogen bonds is reduced in the following order: serine side chain (0.42), solvent (0.14), almost no interactions with the backbone (0.08) and tryptophans (0.00). For HMA^+, the water molecules are also the most frequent hydrogen-bond contacts with the blocker (1.7), followed by hydrogen bonding to the serine side chains (0.77) and tryptophan (0.26). For HMA, the overall averaged number of hydrogen bonds is also lower than for HMA^+. Hydrogen bonding with the side chains of the serines (0.33), the solvent (0.25) and the backbone (0.22) are all less frequent than for the protonated blocker. Tryptophans are no longer significantly involved in hydrogen-

bonds with HMA (0.02). As a result, AM^+ interacts mainly with the serine side-chains, whilst HMA^+ also interacts with the tryptophans.

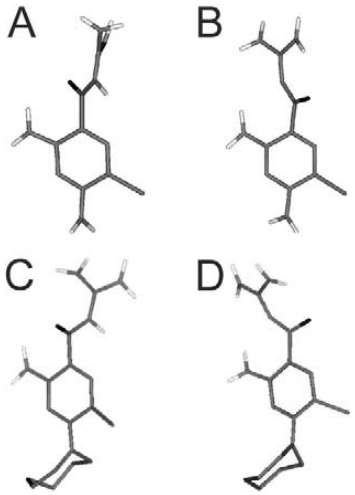


Figure 3 Most frequently occupied conformations of AM (A: protonated, B: deprotonated) and HMA (C: protonated, D: deprotonated).

Am remains with its long axis almost parallel to the membrane plane independent of its protonation state: 5.3 ± 23.3° protonated, 8.6 ± 51.8° deprotonated. The large standard deviation is due to the frequent reorientation of the blocker. Rotational motion around the membrane normal (z-axis), in the plane of the membrane (x-y plane) is indicated by oscillations in the range of ± 40° for AM^+. Hma, however is more likely to orientate itself out of the membrane plane towards the membrane normal: 43.0 ± 11.8° protonated, 11.3 ± 12.7° deprotonated. The hexamethylene ring orientates itself towards the hydrophobic N terminal end of the bundle. The lower standard deviation is indicative for a relatively narrow band of possible orientations held for a relatively long time during the simulations. HMA^+ screens for the hydrophobic area of the pore with its hydrophobic methylene ring. Rotation around the membrane normal is restricted to slight fluctuations around -15° for HMA^+.

DISCUSSION

In a recent publication (16) the blocking of either full-length Vpu and a peptide corresponding to the TM segment of Vpu reconstituted into lipid membranes by Am derivatives has been demonstrated.

Also the viral ion channel p7 from hepatitis C virus can be blocked by the Am derivative, Hma (38). With the present investigations we aim to shed light on the mechanism of interaction between these blockers and a Vpu bundle. The computational model of the bundle is based on the available structural and functional information in the literature (for an overview see (8, 18).

Plausibility of the bundle model

There is computational and experimental evidence which supports a model where Vpu adopts a pentameric oligomer-isation state. Early computational studies based on the TM part of Vpu were performed with constrained pentameric and hexameric bundles (22). The potential energy profiles of the interactions of a cation and an anion with the pore are in favour of the cation for the pentameric bundle, which sup-ports the idea of a pentameric oligomerisation state of the bundle. In an experimentally driven approach using Fourier transform infrared (FTIR) spectroscopy, a pentameric bundle showed the lowest potential energy profile and matched most closely the experimentally derived data, such as the tilt angle of the peptide inserted into lipid membranes (27). Comparison of conductance data with model calculations excludes a tetrameric assembly and supports the pentameric model (13). Furthermore, recent MD simulations show instabilities of hexameric bundles of Vpu compared to pentameric bundles, which is interpreted as another indication of the pentameric model as the working model (32).

The consideration of how to build bundles of Vpu has been discussed previously in detail elsewhere (12). In brief, the serines point into the pore, which moves the tryptophans to the outside of the bundle. The orientation of hydrophilic residues, such as serines, towards the pore is also established for the nicotinic acetylcholine receptor [23,24]. The consequent position of the tryptophans in the bundle model of Vpu has its experimental verification in the fact that helical membrane proteins, such as bacteriorhodopsin (37, 40), as well as β-barrel pores, such as porins (36), have these amino acids anchoring the protein within the membrane. The macroscopic analysis of the computational bundles, such as the average tilt angles of the bundle, falls within the range found experimentally (for an overview see (18)). The average tilt angle for the bundle without blocker (5.8 ± 2.7°) seems to be lower than the most recent experimentally derived value of 13° from nuclear magnetic resonance spectroscopy (35) and closer to the value derived from FTIR spectroscopy (6.5 ± 1.6°)

(27). The differences may arise from the use of TM segments, which differ in length. The different lipids used in the individual approaches and the con-sequent phase conditions they impose may also account for the differences. The kink angle calculated from the present bundle structure, taking the Cα-atoms of the residues Ile-16 to Ala-18 as centre points, is slightly lower compared to the experimentally derived values (35). However, calculations of the kink angle for the extended model $Vpu_{1–52}$ as a single entity in the lipid bilayer (43) fit well with the experimentally derived data (35). In the current model, the arginine residues point into the pore in such a way as to compensate the selectivity for cations. However, the arginine side chain is very flexible and is able to occupy a large conformational space by swinging in and out of the pore. In a kinked model, the arginine shows large movements during the simulations (43). This flexibility might have an impact on the electrostatic profile of the C-terminal mouth of the channel and its ion selectivity. As a consequence anions could pass without being trapped by the positive charges. Glutamate and lysine, the other charged residues present at the C-terminal end, are positioned at the outside of the bundle and are proposed to be involved in clamping of the bundle (18). The clamp might be weak and permit a dynamic behaviour of the bundle on a biologically relevant time scale. These dynamics could also allow the flux of larger ions, such as phosphate ions, through the pore (17).

Consequently, the present model is a reasonably good model for the simulations on a Vpu bundle. The missing water molecules within the pore towards the N-terminal end reflect the strong hydrophobic character of this part of the TM segment combined with the small pore radius. Earlier calculations using cut-o¡ radii for the treatment of electrostatic interactions revealed continuous water columns within the pentameric bundle (13). Recent investigations of theoretical hydrophobic pore models using PME calculations show that the presence of water in these pores is strongly dependent on the pore radius and on membrane polarisation effects (1, 2, 5). In addition the Lennard–Jones parameters of the water molecules (25) and the hydrophobic/hydrophilic character of the walls with which the waters can interact affect the presence of water molecules (42). Regarding the pore radius, below a certain threshold diameter water molecules would not pass through a pore (2, 5). With increasing pore size, water will eventually pass in an intermittent way (1, 2) or “avalanche-like fashion” (5). Simulations on longer time scales are needed to assess the proper behaviour of the water molecules in confined geometries. A protein model from the

mechanosensitive channel MscS, with a narrow hydrophobic part (hydrophobic lock) within the pore similar to Vpu, also shows the escape or "dewetting" of this part of the pore (4). The Vpu bundle model used might represent a proper bundle with a temporary state of low water content. It is possible that the hydrophobic part contributes to the low conductance found for the peptide corresponding to the TM segment of Vpu and reconstituted into lipid membranes (17, 35).

Plausibility of the blocker-protein binding site

Am is a molecule with a considerable number of hydrogen bond donor groups (amino functions) capable of forming hydrogen bonds with hydrogen bond acceptors, such as the serines. Consequently the C-terminal end of the Vpu bundle with the serines as putative hydrogen bond acceptors should be a plausible binding site. In addition, the tryptophans, in an equilibrated bundle model located at the helix/lipid interface, may also act as hydrogen bond acceptors. Glu-28, which is located at the outside of the bundle and could be a hydrogen bond acceptor, has been found not to interact with either blocker in a pure docking approach (C. Kim, V. Lemaitre, A. Watts, W. B. Fischer, to be published). The positive charge of Arg-30 is pointing into the pore (at pH 7) and seems to have only the chloride atom and the carbonyl oxygen atom of the blockers as putative electrostatic counterparts.

With respect to the Am body, the same conclusion applies for Hma. The hydrophobic ring adds an amphipathic character to the blocker. This hydrophobic part is involved in hydrophobic interactions with hydrophobic residues of the bundle within the pore towards the N-terminal end. It should be noted that molecules containing several methyl groups have been found to block voltage-gated sodium channels from frog nerve axons independent of the overall size of the blockers (24). This suggests that molecules carrying methyl groups or Hma with its hexamethylene ring share a common blocking mechanism.

Experimentally and computationally derived evidence for a binding site and binding geometry of amiloride and other blockers in ion channels

Hydrophobic interactions with parts of the channel adjacent to the AM binding site have been found to increase the binding affinity of the blocker (21). For Am, a putative binding site in the Na^+

channel has been proposed to be approximately 20% within the transmembrane electrical field (20, 31) on the extracellular side (6). The orientation allows the guanidinium group to penetrate into the pore (reviewed in (3)). A point mutation within the TM region replacing a crucial serine residue also abolishes affinity for Am (46). Consequently, the binding site proposed, with the guanidinium group in the vicinity of the serines of the putative pore of Vpu in the present study, is in agreement with findings on the epithelial sodium channel.

Effect of hydrogen bonding

The protonated blockers are the physiologically relevant forms in aqueous solution. The positive charge on the blockers has been removed in this study to evaluate the role of charge in the blocking mechanism. The trend of the data for the protonated blockers parallel the data of the protonated blockers with respect to their effect on the protein. Comparison of the blocker data shows differences in conformation, larger fluctuation of the guanidinium group and less hydrogen bonding for the deprotonated blockers. Deprotonation of the blocker results in a loss of hydrogen bond formation with side chains (serines, tryptophan), backbone and water.

HMA^+ also forms hydrogen bonds with tryptophan, with the consequence that this residue is moved towards the blocker into the helix–helix interface and almost into the pore (Fig. 1B). AM^+ and also AM form fewer hydrogen bonds on suggesting a weaker interaction with the bundle. Hydrogen bonding and electrostatic interactions play a major role in the structural conformation of helical TM proteins/peptides in the vicinity of hydrophobic areas and also within them (10, 29). The low dielectric slab of a bilayer imposes a driving force to saturate any free hydrogen bond with a proton acceptor (30). The location of the guanidinium group of the blockers and the serines are at the hydrophilic (headgroup region)/hydrophobic (hydrophobic slab of the bilayer) interface of the peptide bundle. Consequently the guanidinium group might anchor the blocker at this particular position.

Mode of action

The averaged radius of the pore depends on the presence of the blockers and in particular of HMA^+. A pronounced narrowing of the apparent pore size is observed when HMA^+ is in the pore. A pronounced narrowing of the apparent aperture of the middle of the bundle is observed when HMA^+

is in the pore. For AM^+ the reduction in pore diameter is less and would therefore allow waters, and possibly ions with them, to pass the pore under an electrochemical gradient. The proposed mechanism of HMA+ blocking might also involve interaction with the tryptohans. The increase in the tilt angle of the bundles in the presence of AM+ and HMA+ can be best explained by considering the helices as sliding along each other. This mechanistic view is supported by the observed increase in the cross angle.

In this study the situation is presented when the blocker has already reached its binding site. In our current bundle model, arginine side chains are flexible and free to change their orientation, whilst the backbone is embedded within a helical environment. On the *in vivo* time scale these side chains may fluctuate, allowing the blocker to reach its binding site. Like other titratable residues, if close enough to each other (44), all arginines might not be simultaneously protonated. Temporary deprotonation of one arginine would allow for an attraction of the guanidinium group towards the arginines and the mouth of the pore. This would lead to the embedding of the blocker in the pore as outlined.

CONCLUSIONS

The hydrophilic residue Ser-23 is a favourable binding site of the blockers. The mode of blocking of HMA^+, once it is in the pore, is that of steric hindrance. This steric hindrance is supported by a combination of factors such as hydrogen bond formation with tryptophan and implementation of slightly altered pore geometry. AM^+, although favouring the site within the bundle, is still able to allow for space for water molecules and ions to pass through. Regarding the pore radius, its impact on the pore is effectively opposite that of the derivative.

ACKNOWLEDGEMENT V.L. thanks BBSRC for Industrial CASE Scholarship. WBF thanks the E. P. Abraham Research Fund for financial support of this work. The Engineering and Physical Science Research Council (EPSRC), Medical Research Council (MRC), the Biological Science Research Council, the Bionanotechnology IRC and CJ Corporation (for financial support to CGK) are all acknowledged for grant support to AW.

REFERENCES

1. Allen R, Hansen J-P, Melchionna S. 2003. Molecular dynamics investigation of water permeation through nanopores. *J. Chem. Phys.* 119: 3905–3919
2. Allen R, Melchionna S, Hansen J-P. 2002. Intermittent Permeation of Cylindrical Nanopores by Water. *Phys. Rev. Lett.* 89: 175502.1–175502.4
3. Alvarez de la Rosa D, Canessa CM, Fyfe GK, Zhang P. 2000. Structure and regulation of amiloride-sensitive sodium channels. *Annu. Rev. Physiol.* 62: 573
4. Anishkin A, Sukharev S. 2004. Water dynamics and dewetting transitions in the small mechanosensitive channel MscS. *Biophys J* 86: 2883
5. Beckstein O, Sansom MSP. 2003. Liquid-vapor oscillations of water in hydrophobic nanopores. *Proc. Natl. Acad. Sci. USA* 100: 7063–7068
6. Benos DJ. 1982. Amiloride: a molecular probe of sodium transport in tissues and cells. *Am J Physiol* 242: C131
7. Berendsen HJC, Griegera JR, Straatsma TP. 1987. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* 91: 6269
8. Bour S, Strebel K. 2003. The HIV-1 Vpu protein: a multifunctional enhancer of viral particle release. *Microbes and Infections*
9. Brünger AT. 1992. *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*. New Haven, Ct.: Yale University Press
10. Chin CN, von H. 2000. Charge pair interactions in a model transmembrane helix in the ER membrane. *J Mol Biol* 303: 1
11. Cordes F, Kukol A, Forrest LR, Arkin IT, Sansom MSP, Fischer WB. 2001. The structure of the HIV-1 Vpu ion channel: modelling and simulation studies. *Biochim. Biophys. Acta* 1512: 291
12. Cordes FS, Kukol A, Forrest LR, Arkin IT, Sansom MSP, Fischer WB. 2001. The structure of the HIV-1 Vpu ion channel: modelling and simulation studies. *Biochimica et Biophysica Acta* 1512: 291
13. Cordes FS, Tustian A, Sansom MSP, Watts A, Fischer WB. 2002. Bundles consisting of extended transmembrane segments of Vpu from HIV-1: computer simulations and conductance measurements. *Biochemistry* 41: 7359–7365
14. Cordes FS, Tustian AD, Sansom MS, Watts A, Fischer WB. 2002. Bundles consisting of extended transmembrane segments of Vpu from HIV-1: computer simulations and conductance measurements. *Biochemistry* 41: 7359
15. de Groot BL, Tieleman DP, Pohl P, Grubmüller H. 2002. Water permeation through gramicidin A: desformylation and the double helix: a molecular dynamics study. *Biophys. J.* 82: 2934
16. Ewart GD, Mills K, Cox GB, Gage PW. 2002. Amiloride derivatives block ion channel activity and enhancement of virus-like particle budding caused by HIV-1 protein Vpu. *Eur Biophys J* 31: 26
17. Ewart GD, Sutherland T, Gage PW, Cox GB. 1996. The Vpu protein of human immunodeficiency virus type 1 forms cation-selective ion channels. *J. Virol.* 70: 7108–7115
18. Fischer WB. 2003. Vpu from HIV-1 on an atomic scale: experiments and computer simulations. *FEBS Lett.* 552: 39–46
19. Fischer WB, Forrest LR, Smith GR, Sansom MS. 2000. Transmembrane domains of viral ion channel proteins: a molecular dynamics simulation study. *Biopolymers* 53: 529
20. Fyfe GK, Canessa CM. 1998. Subunit composition determines the single channel kinetics of the epithelial sodium channel. *J. Gen. Physiol.* 112: 423
21. Garty H, Palmer LG. 1997. Epithelial sodium channels: function, structure, and regulation. *Physiol. Rev.* 77: 359
22. Grice AL, Kerr ID, Sansom MS. 1997. Ion channels formed by HIV-1 Vpu: a modelling and simulation study. *FEBS Lett* 405: 299–304
23. Grottesi A, Sansom MSP. 2003. Molecular dynamics simulations of a

K^+ channel blocker: Tc1 toxin from *Tityus cambridgei*. *FEBS Lett.* 535: 29

24. Hille B. 1971. The permeability of the sodium channel to organic cations in myelinated nerve. *J. Gen. Physiol.* 58: 599
25. Hummer G, Rasaiah JC, Noworyta JP. 2001. Water conducting through the hydrophobic channel of a carbon nanotube. *Nature* 414: 188
26. Kerr ID, Sankararamakrishnan R, Smart OS, Sansom MSP. 1994. Parallel helix bundles and ion channels: Molecular modelling via simulated annealing and restrained molecular dynamics. *Biophys. J.* 67: 1501
27. Kukol A, Arkin IT. 1999. Vpu transmembrane peptide structure obtained by site-specific fourier transform infrared dichroism and global molecular dynamics searching. *Biophys. J.* 77: 1594–1601
28. Lamb RA, Pinto LH. 1997. Do Vpu and Vpr of human immunodeficiency virus type 1 and NB of influenza B virus have ion channel activities in the viral life cycles? *Virology* 229: 1
29. Landolt-Marticorena C, Williams KA, Deber CM, Reithmeier AF. 1993. Non-random distribution of amino acids in the transmembrane segments of human type-I single span membrane proteins. *J. Mol. Biol.* 229: 602
30. Lew S, Ren J, London E. 2000. The effect of polar and/or ionizable residues in the core and flanking regions of hydrophobic helices on transmembrane conformation and oligomerization. *Biochemistry* 39: 9632
31. Li JH, Cragoe EJ, Jr., Lindemann B. 1987. Structure-activity relationship of amiloride analogs as blockers of epithelial Na channels: II. side-chain mkodifications. *J. Membr. Biol.* 95: 171
32. Lopez CF, Montal M, Blasie JK, Klein ML, Moore PB. 2002. Molecular dynamics investigation of membrane-bound bundles of the channel-forming transmembrane domain of viral protein U from the human immunodeficiency virus HIV-1. *Biophys J* 83: 1259
33. Maldarelli F, Chen MY, Willey RL, Strebel K. 1993. Human immunodeficiency virus type 1 Vpu protein is an oligomeric type I integral membrane protein. *J Virol* 67: 5056
34. Morris GM, Goodsell DS, Huey R, Hart WE, Halliday S, et al. 1998. Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function. *Journal of Computational Chemistry* 19: 1639
35. Park SH, Mrse AA, Nevzorov AA, Mesleh MF, Oblatt-Montal M, et al. 2003. Three-dimensional structure of the channel-forming trans-membrane domain of virus protein "u" (Vpu) from HIV-1. *J Mol Biol* 333: 409
36. Pautsch A, Schulz GE. 2000. High-resolution structure of the OmpA membrane domain. *J Mol Biol* 298: 273–82
37. Pebay-Peyroula E, Rummerl G, Rosenbusch JP, Landau EM. 1997. X-ray structure of bacteriorhodopsin at 2.5 Angstroms from microcrystals grown in lipid cubic phases. *Science* 277: 1676–1681
38. Premkumar A, Wilson L, Ewart GD, Gage PW. 2004. Cation-selective ion channels formed by p7 of hepatitis C virus are blocked by hexamethylene amiloride. *FEBS Lett* 557: 99
39. Roux B, Karplus M. 1991. Ion transport in a model gramicidin channel: structure and thermodynamics. *Biophys. J.* 59: 961
40. Sass H, Buldt G, Gessenich R, Hehn D, Neff D, et al. 2000. Structural alterations for proton translocation in the M state of wild-type bacteriorhodopsin. *Nature* 40: 649–653
41. Smart OS, Neduvelil JG, Wang X, Wallace BA, Sansom MSP. 1996. Hole: A program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph.* 14: 354
42. Spohr E, Trokhymchuk A, Henderson D. 1998. Adsorption of water molecules in slit pores. *Electroanal. Chem.* 450: 281–287
43. Sramala I, Lemaitre V, Faraldo-Gomez JD, Vincent S, Watts A, Fischer WB. 2003. Molecular dynamics simulations on the first two helices of Vpu from HIV-1. *Biophys J* 84: 3276
44. Tieleman DP, Breed J, Berendsen HJ, Sansom MS. 1998. Alamethicin channels in a membrane: molecular

dynamics simulations. *Faraday Discuss.*: 209

45. van Aalten DMF, Bywater R, Findlay JB, Hendlich M, Hooft RW, Vriend G. 1996. PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J Comput Aided Mol Des* 10: 255
46. Waldmann R, Champigny G, Lazdunski M. 1995. Functional degenerin-containing chimeras identify residues essential for amiloride-sensitive Na+ channel function. *J. Biol. Chem.* 270: 11735
47. Willbold D, Hoffmann S, Rösch P. 1997. Secondary structure and tertiary fold of the human immunodeficiency virus protein U (Vpu) cytoplasmatic domain in solution. *Eur. J. Biochem.* 245: 581

Conclusions

Conclusions

Results from this thesis can be split into three categories, which pertain to each part of the system, 0xygen-17 solid state NMR and MD simulation on membrane protein and peptides. The remaining conclusions concern those drawn from the use of these two tools to study noncovalent interactions within biomolecules. These three categories will be discussed below.

1. OXYGEN-17 SOLID STATE NMR

Through the application of recent advances in solid state NMR methodologies and equipment, we have demonstrated the feasibility of studying oxygen-17 labelled molecules in a biomembrane environment.

2. MOLECULAR DYNAMICS SIMULATIONS

3. GENERAL CONCLUSIONS

Appendices

A.1 Solid State ^{17}O NMR of Amino Acids

Kevin J. Pike,[1] Vincent Lemaître,[2,3] A Kukol,[4] T Anupõld,[4] A Samoson,[4] Andrew P. Howes,[1] Anthony Watts,[2] Mark E. Smith,[1] and Ray Dupree,[1]

[1] *Departments of Physics and Biological Sciences, University of Warwick, Coventry CV4 7AL, UK ; e-mail: r.dupree@warwick.ac.uk*

[2] *Biomembrane Structure Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK*
[3] *Nestec S.A., BioAnalytical Department, Vers-Chez-Les-Blanc, CH-1000 Lausanne 26, Switzerland*
[4] *Institute for Chemical Physics and Biophysics, Akadeemia Tee 23, Tallinn, Estonia.*

Received: January 5, 2004; In Final Form: April 20, 2004

Keywords Oxygen-17, amino acids, solid state NMR, hydrogen-bonds, 3Q MAS, DOR

■ **ABSTRACT** ^{17}O solid state NMR from 14 amino acids is reported here, greatly increasing the number investigated. In most cases well-separated resonances from carbonyl and hydroxyl oxygens with distinct second-order quadrupolar line shapes are observed using a 600 MHz spectrometer with fast magic angle spinning (MAS). This is in contrast to the motionally averaged resonances usually seen from amino acids in solution. For amino acids double-angle rotation (DOR) produces a decrease in the line width by more than a factor of 40, providing very high resolution, ~1 ppm, spectra. The oxygen lines in alanine and the carbonyl oxygens in *L*-glutamic acid hydrochloride are assigned using ^{1}H-decoupled DOR. The NMR interaction parameters for amino acids show a wide variation of χ_Q, from 6.4 to 8.6 MHz, η, from 0.0 to 0.9, and δ_{iso}, from 83 to 353 ppm. The high quality of the MAS NMR line shapes obtained at 14.1 T means that even small changes in parameters can be very accurately deduced, offering the possibility of ^{17}O NMR as a sensitive probe of structural changes in these and related compounds. The *D*-and *L*-forms of glutamic acid hydrochloride are shown to have the same NMR parameters to within error, which are very different from those reported in the literature for the *D,L*-form. A strong correlation (~–1200 ppm/Å) is found between δ_{iso} and the C–O bond length of the carbonyl oxygens. On the basis of

these data, enriching specific amino acids in more complex polypeptides and proteins could provide site-selective information about the bonding and functionality of different sites in biomolecules. An estimate is made of the possible detection limit for such species.

INTRODUCTION

The widespread occurrence of oxygen in many technologically and scientifically interesting materials would suggest that NMR studies of this nucleus could provide much crucial information about the structure and bonding in a wide range of materials. However, the spin $I = 5/2$ nucleus ^{17}O with a natural abundance of 0.037% is the only NMR-active isotope of oxygen, and as it possesses a quadrupole moment, ^{17}O resonances are often significantly broadened in solids (38). It is the combination of low sensitivity and sometimes large line widths that make ^{17}O NMR studies still relatively uncommon. However, there is high sensitivity of the NMR parameters such as the isotropic chemical shift (covering ~1000 ppm in organic molecules) and the quadrupole parameters to structural detail. Study of oxygen has been further encouraged by the advent of higher magnetic fields, faster magic angle spinning (MAS), and techniques for improving resolution (38) (e.g., dynamic angle spinning (DAS), double-angle rotation (DOR), multiple-quantum (MQ) MAS). This has led to a significant increase in ^{17}O NMR reports from inorganic materials (27) with examples of ^{17}O NMR studies of amorphous materials to determine nanoscale phase separation in silica-based gels (8, 13, 31) and the distribution of bonding species in a range of glasses (7, 18, 39, 53) . There have also been studies of crystalline materials such as zeolites and mineral analogues (3, 4, 20, 33), but with much less work on organic materials since the powder line shapes are typically much broader because of the large electric field gradient as a result of the increased covalency of the M-O bonds in these materials.

A key experimental challenge for biomolecular chemistry is to provide high-quality, detailed, and unambiguous atomic-scale information about the molecular bonding arrangement and changes that occur upon ligand-receptor interaction. Solid state NMR is one nonperturbing approach which can be used to study such interactions where molecular size is not limiting and crystallinity not a

requirement (44, 45) , and the ubiquity of oxygen throughout living systems should imply that ^{17}O is an important nucleus for such studies. Oxygen plays a key role in intra-and intermolecular interactions, with hydrogen-bonding important in biological processes so that ^{17}O could provide detailed information about the dynamics and structure of amino acids both in the solid state and in solution. This demands the development of experimental probe techniques to deliver this information. ^{17}O NMR from inorganic materials such as that cited above has shown that the sensitivity of the NMR parameters to structural detail gives great encouragement for the NMR study of organic materials. Recent reports of high-field ^{17}O NMR from organic materials have included heme proteins (16), strongly hydrogen-bonded carboxylic acids (17), nucleic acid bases (48), and other organic materials (9).

In proteins it is important to understand processes such as folding, which is related to functional properties and disease, in terms of the conformation through the secondary structure. It is the amino acid sequence and the interresidue bonding that influence this structure. The more common nuclei for the NMR study of such problems are ^{1}H, ^{13}C, and ^{15}N. ^{17}O would be another possible NMR approach, with the potential importance of ^{17}O being recognized through attempts more than 20 years ago to obtain ^{17}O NMR spectra from amino acids and peptides in solution. The work of Fiat and co-workers presented detailed ^{17}O enrichment protocols, most often specifically labeling the carbonyl site. The solution spectra from a series of [^{17}O]-carbonyl-exchanged amino acids gave single intense carbonyl resonances typically between 249 and 265 ppm (40, 43). Spectra from solids are potentially more informative than those from solution as both the shift and the quadrupole coupling constant $\chi_Q = \dfrac{e^2Qq}{\left(2I(2I-1)\eta\right)}$; eQ is the quadrupole moment and eq is the maximum component of the electric field gradient) and the asymmetry parameter η can be determined. Furthermore, different sites are not averaged to single resonances by motion and/or exchange that is present in solution. Often the solid state is more representative of the natural bonding state (e.g., in a membrane) than a solution. Also knowledge of the interaction parameters aids interpretation of the solution spectra (e.g., if χ_Q is known, NMR provides the correlation time more unambiguously). Hence, extension to the solid quickly followed the initial solution-state studies (14, 15). However, there were a number of factors that counted against the ability to obtain the high-quality NMR

spectra which would allow NMR parameters to be unambiguously determined including low applied magnetic field (e.g., 6.35 T), initially no MAS, and then only modest MAS rates. The second-order quadrupole powder patterns that were obtained often showed significant distortion from the true line shapes with, for example, narrow spectral features such as the singularities often disproportionately intense. Static NMR had the added disadvantage that the line shape contains both second-order quadrupole and chemical shift anisotropy (CSA) effects so that it was difficult to unambiguously determine the interaction parameters. The problems with simulation and excitation of this early ^{17}O NMR work can be gauged by values of χ_Q as low as 4 MHz being reported for polyglycine II compared with a true value of 8.2 MHz (22, 23).

In recent years, the situation has rapidly improved with the more widespread availability for solid state NMR studies of magnetic fields of ≥ 14.1 T and MAS rates of >25 kHz. Work compared four [^{17}O]carbonyl-labeled polyglycines which show very different hydrogen-bonding with variations of 8.30–8.55 MHz in χ_Q, 0.26-0.47 in η, and 288-407 ppm in the isotropic chemical shift ($\delta_{cs,iso}$) (23). The richness of the solid state data over those from solution where there is much less variation in the bonding and there are averaging effects is illustrated by a shift variation of ~120 ppm from these compounds compared to the report of a single shift of 267 ppm in solution (23). This shift variation is also much greater than those reported for ^{13}C and ^{15}N from the same compounds. The variation in these parameters between polyglycines I and II is related to differences in the molecular packing between the β-sheet and 3_{10}-helix forms (50). A ^{1}H-^{17}O double-resonance NQR study of polycrystalline *D,L*-proline showed eight different signals that were found to fall into two groups that could be assigned to hydroxyl (χ_Q = 6.08–6.79 MHz) and carbonyl (χ_Q = 7.72-8.73 MHz) oxygens (35). This assignment could be made on the basis of the relative values of χ_Q and because in the ^{1}H-^{17}O double-resonance approach the signal assigned to the hydroxyls shows a much stronger enhancement. Two poly(*L*-alanines) which adopted the α-helix and β-sheet forms were readily distinguished through the different χ_Q and $\delta_{cs,iso}$ values of the enriched carbonyl sites (51). A two-field DAS study of a uniformly ^{17}O labeled sample of the parent amino acid *L*-alanine showed two completely resolved ^{17}O signals, with the position of the lines at 11.7 T differing by 29 ppm. High-power ^{1}H decoupling substantially narrowed the DAS lines (12). The two signals arise from

distinct sites in the structure, with O(1) the conventional hydrogen bond with O···N, whereas the other oxygen, O(2), site is bifurcated as the attached proton is also connected to two other nitrogens. In *D*-alanine a 3Q MAS study was carried out with again two sites observed (47). These two studies gave reasonable agreement between the shifts of the two sites, and for one site $P_Q = \chi_Q(1 + \eta^2/3)^{1/2}$ also agreed within error. However, for the other site the values for P_Q from the two studies did not agree within error, and this will be examined in this paper. ^{17}O NMR of tyrosine hydrochloride labeled at the phenyl OH (i.e., not at the COOH) position gave a shift more than 150 ppm smaller than that of either site in alanine with a much larger χ_Q. Comparison of static and MAS NMR spectra indicated that there was a CSA of ~100 ppm (9). Recently, we reported a detailed ^{17}O NMR study of the *L*-form of glutamic acid hydrochloride (26) which gave NMR parameters very different from those of some of the sites reported for the *D,L*-form (47). The *D,L*-form of glutamic acid showed a resonance additional to those of the pure L-form with ~50% of the total intensity, which implies that *D,L*-glutamic acid is a racemic crystal as opposed to a racemic conglomerate of chiral crystals (47).

Despite this initial work on ^{17}O in amino acids, there is still no detailed knowledge of the range of values the NMR interaction parameters take (especially of C-OH) nor the relationship these parameters have to structure. It has been suggested that χ_Q has a linear relationship to the N···O bond length (50). However, a recent computational study using density functional theory calculated the NMR interaction parameters from α -helix and β-sheet conformations and suggested there was no dependence on the bond length (41). Distinct quadrupolar parameters from each of these conformations were calculated with χ_Q changing by 0.53 MHz, much larger than the accuracy with which it can be determined. The variations come from differences in the hydrogen bond and backbone dihedral angles, thus solid state ^{17}O NMR approach offers much potential for understanding subtle variations in the local bonding. In this paper we seek to extend the understanding of the utility of ^{17}O by greatly increasing the number of ^{17}O parameters from amino acids, reporting the NMR parameters of 14 amino acid samples.

EXPERIMENTAL DETAILS

Amino acids were ^{17}O-enriched using either a slight variant of the procedure described previously for ^{18}O enrichment (42) or that described previously by us for enriching *L*-glutamic acid (26, 40). In brief, the amino acid was suspended in a mixture of ^{17}O-enriched water and dioxane (1:3, v/v). Water was used in 25 fold molar excess over the amino acid. The mixture was treated with a continuous stream of HCl gas obtained from $MgCl_2$/ H_2SO_4 for 2 h while being kept at 90°C. The sample was lyophilized and the $H_2{}^{17}O$/dioxane recovered. Water enriched to 10-20 atom % $H_2{}^{17}O$ was found to be sufficient to achieve an acceptable signal-to-noise (S/N) ratio for the solid state NMR spectroscopy of these simple amino acids. The HCl form of the amino acid is produced using this enrichment route. [$^{17}O_2$]-*L*-Alanine hydrochloride was prepared by acid-catalyzed exchange with $H_2{}^{17}O$ following literature procedures (40, 47). A 900 mg sample of *L*-alanine was dissolved in 1 g of $H_2{}^{17}O$ in a 3.0 mL glass vial. The solution was saturated with dry HCl gas, sealed, and then heated to 100°C for approximately 24 h. The sample was then cooled to room temperature and allowed to crystallize, followed by several successive recrystallizations in the $H_2{}^{17}O$ used for the exchange. Finally, the sample was cooled to 4°C, and crystals were recovered. $H_2{}^{17}O$ was recovered using microdistillation under a N_2 atmosphere. [$^{17}O_2$]-*L*-Alanine was prepared using a similar procedure with less alanine (800 mg), more water (2 g), and only partial saturation with HCl. The initial amino acids were all supplied at >99.8% purity followed by straightforward processing by placing them in water and passing HCl gas, which should not change the form of the amino acid present other than producing ^{17}O enrichment and the HCl salt. The identity of the product was confirmed by solid state ^{13}C NMR and powder XRD. ^{13}C NMR was performed on a Chemagnetics Infinity 600 or 360 spectrometers operating at 150.92 and 90.55 MHz, respectively. A 4 mm probe was used spinning at 7.5-10 kHz with contact times of 1 ms. Powder XRD was run on a Bruker D5005 diffractometer using Cu KRI radiation at a wavelength of 1.5405 Å. The patterns were taken from $2\theta = 10°$ to $2\theta = 90°$ in 0.2° steps, spending 4–8 s at each point. Both ^{13}C and powder XRD confirmed the samples were as expected and were single phase. On some samples ^{35}Cl NMR was also used as further confirmation.

Most of the ^{17}O NMR was carried out on a Chemagnetics Infinity 600 spectrometer at a frequency of 81.345 MHz. MAS and 3Q MAS experiments used either a 4 mm probe spinning at ~16 kHz or a 3.2 mm MAS probe spinning at 17–22 kHz. The recycle delay was typically 3 s, which was sufficient to prevent saturation. Typically, ~20K scans were sufficient to give an adequate S/N ratio with the 3.2 mm probe for samples produced from 20% ^{17}O enriched water. A spin echo using extended phase cycling (21) was used with the echo spacing set to the rotation period. High-power ^{1}H decoupling using the XiX scheme (5) was employed where necessary. It was very important to be exactly on angle as the combination of the echo and ^{1}H decoupling produced very high quality second-order quadrupolar line shapes. Even small deviations from the angle significantly affected the NMR parameters deduced from the line shape. Experiments using DOR NMR were carried out using odd-order sideband suppression where the acquisition of successive scans is triggered at orientations of the outer rotor differing by 180° (34).The outer rotor speed was varied between 1300 and 1800 Hz to determine the centerbands. Again a recycle delay of typically 3 s was required, but the very much narrower resonance lines meant that only ~2000 scans were needed. The DOR probe, which was constructed in one of our laboratories, also has the facility for simultaneous ^{1}H irradiation, and in some cases spectra were acquired with a range of decoupling fields up to 34 kHz. All spectra were referenced to water at 0 ppm. Some additional spectra were acquired at a magnetic field of 8.45 T and a frequency of 48.8 MHz. Spectral simulations were carried out using the dmfit software (29).

RESULTS

Examples of typical ^{17}O MAS NMR spectra from amino acids are shown in Figure 1 from *L*-tyrosine hydrochloride (a), *L*-asparagine hydrochloride (b), *L*-valine hydrochloride (c), and *L*-glycine hydrochloride (d). In these samples which are uniformly labeled in the COOH group, there are two well-separated resonances, both with distinct second-order quadrupole line shapes. Also shown in Figure 1 are simulations which are of high accuracy because of the clear separation of the lines and the well-delineated line shapes, with the parameters deduced summarized in Table 1. The two lines are from the carbonyl and hydroxyl oxygens, with the higher shift line from the carbonyl.

This assignment is justified by comparison with the shifts observed from carbonyls and hydroxyls in solution. Also the ^{17}O resonances from solid poly(amino) acids selectively enriched at the carbonyl site typically have $\chi_Q \approx 8.2$ MHz and $\delta_{cs,iso} \approx 300$ ppm, further confirming this assignment. The spectrum of a *L*-tyrosine sample selectively enriched at the hydroxyl group attached to the phenol group (Cambridge Isotope Laboratories) is also shown in Figure 1a; it is much sharper than that previously reported (47), allowing a more accurate determination of the parameters. The values of χ_Q (8.56 MHz) and η (0.65) are much higher than those of the hydroxyl in the COOH group (7.35 MHz, 0.19), and the shift is 100 ppm less.

The spectra of the other two oxygen site amino acid hydrochlorides are similar to those in Figure 1, and their NMR parameters are given in Table 1, which has representatives of all the main classes of amino acids, nonpolar, polar, acidic, and basic. The sensitivity of the ^{17}O NMR parameters is well illustrated by a sample of valine that was initially ^{17}O-enriched and then subsequently reacted with 9-fluorenylmethoxycarbonyl (fmoc) to protect the oxygen-labeled sites during further synthesis. The ^{17}O MAS NMR spectrum again shows two clearly separated second-order quadrupolar line shapes, from the carbonyl and hydroxyl oxygens, but their parameters are significantly different from those of pure valine (Table 1).

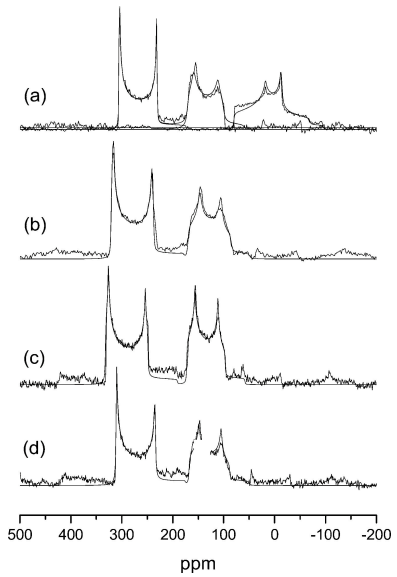


Figure 1 MAS NMR (14.1 T) spectra of (a) *L*-tyrosine hydrochloride, (b) *L*-asparagine hydrochloride, (c) *L*-valine hydrochloride, and (d) glycine hydrochloride together with simulations of the centerbands. (Note the *L*-tyrosine spectrum is a composite of two samples, one ^{17}O-enriched in the two oxygens of the carboxylate group and the other in the phenol position.)

Figure 2a shows the spectrum of *L*-alanine hydrochloride, together with a simulation, which is similar to those of the other acid hydrochlorides with clearly resolved CdO and OH lines. The spectrum of *L*-alanine shown in Figure 2b is very different from that obtained for *L*-alanine hydrochloride and from the typical spectra of the other amino acid hydrochlorides. The two strongly overlapping lines lie between those of *L*-alanine hydrochloride, but sufficient features can be resolved to allow a good simulation. To confirm and further refine the parameters in the simulation both DOR, shown in Figure 2c, and 3Q MAS NMR were carried out. As can be seen DOR produces a large increase in the resolution, and two narrow lines at 228 and 218 ppm, with associated spinning sidebands, can be identified by varying the spinning speed from 1300 to 1800 Hz. The peak position (δ_{peak}) in DOR spectra is given by

$$\delta_{peak} = \delta_{cs,iso} - \frac{3}{40} F(I) \frac{C_Q^2}{\nu_0^2} \left(1 + \frac{\eta^2}{3}\right) \qquad (1)$$

where $\delta_{cs,iso}$ is the isotropic chemical shift, ν_0 is the Larmor frequency, and $F(I)$ is a spin-dependent factor, which for $I = 5/2$ is 2/25.

Table 1 ^{17}O NMR Interaction Parameters for Amino Acids structure

Compd, line	$\delta_{cs,iso}$ ± 0.5 ppm	χ_Q ± 0.05 MHz	η ± 0.02	assignment	ref	structure ref
L-glutamic acid hydrochloride						
1	322.0	8.16	0.0 ± 0.03	O2	(26)	(36)
2	315.0	8.31	0.17	O3		
3	187.0	7.49	0.25	O1 or O4		
4	172.5	7.45	0.25	O1 or O4		
D-glutamic acid hydrochloride						
1	322.0	8.22	0.03 ± 0,03	O2	This study	na
2	315.4	8.35	0.17	O3		
3	187.2	7.49	0.25	O1 or O4		
4	172.3	7.45	0.25	O1 or O4		
D,L-glutamic acid hydrochloride						
1	320	8.2	0.0		(47)	na[b]
2	250	6.8	0.58			
3	250	6.8	0.58			
4	170	7.2	0.20			
D-alanine						
1	275 ± 5	7.60 ± 0.02	0.60 ± 0.01	O1	(47)	na
2	262 ± 5	6.40 ± 0.02	0.65 ± 0.01	O2		
L-alanine						
1	285 ± 8	8.1 ± 0.3[a]		O1	(12)	(25)
2	268 ± 8	7.2 ± 0.3[a]		O2		
L-alanine						
1	284	7.86	0.28	O1	This study	(25)
2	260.5	6.53	0.70	O2		
L-alanine hydrochloride						
1	327.8	8.31	0.0	C=O	This study	(6)
2	176.7	7.29	0.20	OH		

fmoc-protected *L*-alanine						
1	303.3	7.89	0.16		This study	na
2	175.7	6.95	0.12			
glycine hydrochloride						
1	336	8.40	0.0	C=O	This study	(6)
2	185	7.60	0.25	OH		
L-tyrosine hydrochloride						
1	327.0	8.22	0.0	C=O	This study	(11)
2	183.0	7.35	0.19	OH		
3	83	8.56	0.65	O3 (OH)		
D,L-tyrosine hydrochloride						
3	117	8.1	1.0	OH	(47)	na
L-asparagine hydrochloride						
1	342.5	8.55	0.0	C=O	This study	na
2	178.5	7.49	0.30	OH		
L-valine hydrochloride						
1	351	8.40	0.03	C=O	This study	(19)
2	181	7.35	0.21	OH		
fmoc-protected *L*-valine						
1	324.1	8.42	0.08		This study	na
2	167.3	7.48	0.27			
L-leucine hydrochloride						
1	342.7	8.39	0.05	C=O	This study	na
2	183.1	7.50	0.20	OH		
L-isoleucine hydrochloride						
1	347.1	8.52	0.06	C=O	This study	na
2	182.6	7.40	0.22	OH		
L-lysine hydrochloride						
1	346.7	8.56	0.0	C=O	This Study	Na
2	180.8	7.67	0.24	OH		
L-phenylanaline hydrochloride						
1	353.5	8.54	0.07	C=O	This study	(1)
2	178.8	7.46	0.25	OH		
L-cysteine hydrochloride						
1	353.5	8.65	0.18	C=O	This study	na
2	174.9	7.41	0.27	OH		
L-glutamine hydrochloride						
1	319.8	8.20	0.03	C=O	This study	(37)
2	306 ± 1	8.30 ± 0.1	0.03 ± 0.03			
3	180 ± 1	7.75	0.24	OH		

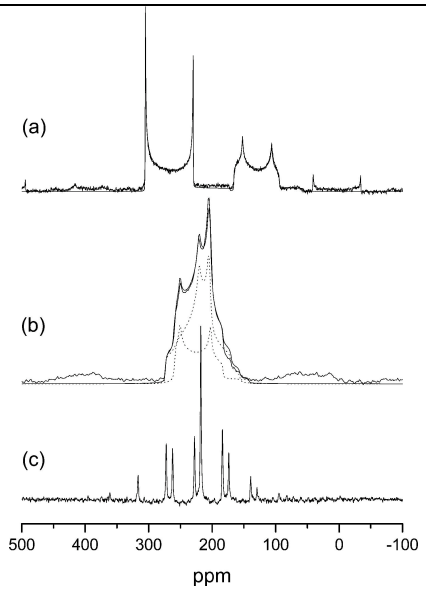


Figure 2. (a) ^{17}O MAS NMR spectrum of l-alanine hydrochloride together with simulation, (b) MAS NMR spectrum of *L*-alanine together with simulation, and (c) DOR spectrum of *L*-alanine, outer rotor speed 1800 Hz.

Analysis at multiple applied magnetic fields (2, 38) allows P_Q and $\delta_{cs,iso}$ to be deduced. Alternately, if an MQ MAS experiment is carried out, one field is sufficient since in a 3Q MAS experiment the quadrupolar shift from the second term in eq 1 changes sign and is multiplied by 10/17 (27, 32). Thus, the peak position in the isotropic dimension will be that corresponding to a negative $1/\nu_0^2$, and plotting the center of gravity against the inverse of the square of the Larmor frequency allows both the isotropic chemical shift and, from the gradient, P_Q to be deduced. Such a plot is shown in Figure 3. Also shown on the plot are the isotropic shifts determined from the MAS simulation. The anisotropic dimension in the 3Q data clearly shows that the line with an isotropic shift of 260.5 ppm had a high η and that with a shift of 284 ppm a low η, in agreement with the MAS simulation. The values of $\delta_{cs,iso}$, χ_Q, and η for both forms of alanine are given in Table 1. The P_Q calculated from our χ_Q and η agrees with the DAS measurements (12) (to within their stated error), as do the shifts. The agreement is not as good with the measurements of Wu et al. on *D*-alanine (47). Wu et al. find both sites have a high η (0.65 and 0.60 ± 0.01) compared with the low η (0.28 ± 0.03) found here for the line at 284 ppm that is clearly needed to simulate the MAS spectrum. The shift of this line is also somewhat removed from the value, 275 ± 5 ppm, obtained by Wu. Assignment of the two lines is less clear from the NMR parameters for *L*-alanine than for *L*-alanine hydrochloride since the chemical shift differs by only 25 ppm compared with the ~150 ppm shift difference typical of the amino acid hydrochlorides (Table 1). The undecoupled DOR line width of the line at 218 ppm in the DOR spectrum is a little larger than that of the line at 228 ppm, indicating perhaps some residual proton coupling contribution to the line width, and in fact the line widths of both lines were found to depend strongly on the ^{1}H decoupling field, initially *increasing* as the decoupling field is increased.

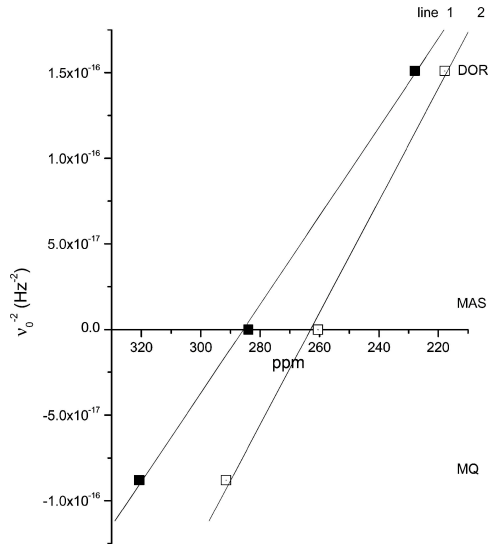


Figure 3. Peak position (ppm) for the two *L*-alanine lines as a function of the inverse square of the apparent Larmor frequency. $\delta_{cs,iso}$ determined from the MAS simulation is also shown.

Figure 4 shows the DOR spectrum with no decoupling (middle), with a decoupling field of 17 kHz (upper), and with 34 kHz decoupling (lower), and the line width as a function of decoupling field is shown in Figure 5. A maximum is seen in the line width of the 218 ppm peak which occurs at ~17 kHz decoupling.

Although the details of this are, as yet, not fully understood, it is almost certainly a rotary resonance effect (46) since the decoupling field is approximately twice the spinning frequency of the inner rotor (i.e., $\nu_{1H} = 2\nu_r$), which is 8.5 kHz for the outer rotor speed of 1800 Hz used here (49). There is also a smaller less well delineated peak in the line width of both lines for decoupling fields of 5-10 kHz. The line at 218 ppm is approximately 50% broader at 17 kHz decoupling than that at 228 ppm, whereas at high decoupling powers their widths are similar; thus, the 218 ppm resonance experiences a significantly greater ^{1}H dipolar coupling. The local structure around the oxygens is shown in Figure 6, and the resonances can be readily assigned, since O1 is close to one proton at 1.86 Å, whereas O2, unusually for an amino acid, has strong interaction with two protons at 1.78 and 1.83 Å (25). Hence, O2 undoubtedly has the stronger dipolar coupling and corresponds to the peak at 218 ppm with an isotropic shift of 260.5 ppm. Further confirmation of the assignment is the high η, which is consistent with the unusual interaction with two protons. In a fashion similar to that of fmoc-protected valine the NMR parameters of fmoc-protected alanine are significantly different

from those of the unprotected acid (Table 1); in particular for both protected acids the carbonyl shift is ~25 ppm smaller than for the unprotected form.

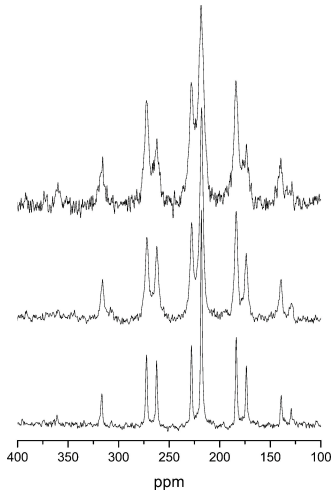


Figure 4. DOR spectrum of *L*-alanine at different ^{1}H decoupling fields: top, decoupling 17 kHz; middle, no decoupling; bottom, decoupling 34 kHz.

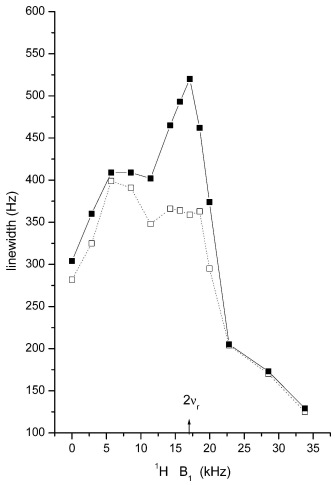


Figure 5 Variation of the line widths for the two lines in the DOR spectrum of *L*-alanine as a function of the ^{1}H decoupling field (kHz).

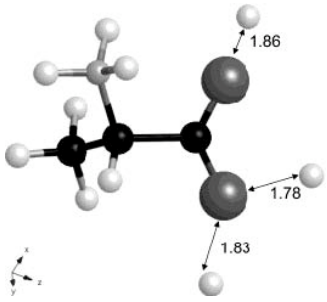


Figure 6 *L*-Alanine structure (25) showing the O-H distances.

The ^{17}O MAS (Figure 7a) NMR spectrum from *D*-glutamic acid is more complex since the molecule has two inequivalent COOH groups so that there are two inequivalent carbonyl and hydroxyl groups in each unit cell. Hence, there are two resonances in both the carbonyl and hydroxyl regions.

Despite the extensive overlap in the MAS NMR spectrum resulting from an isotropic chemical shift difference of only ~7 ppm for the carbonyl oxygens, a detailed simulation can be carried out that separates them completely, allowing the NMR parameters from each to be deduced (Table 1) as was discussed in detail recently for *L*-glutamic acid (26). The two lines at lower shift come from OH oxygens (O1 and O4), and to obtain the resolution shown in Figure 7, high-power (>100 kHz) ^{1}H decoupling was necessary. The DOR spectrum (Figure 7b) shows only two lines spaced by ~10 ppm. There is no evidence of narrowed DOR resonances from the hydroxyls that would appear at isotropic shifts of ~135 and 120 ppm. The maximum ^{1}H decoupling in these DOR experiments is only 34 kHz, and the absence of the signals from the hydroxyl oxygens is attributed to the much larger ^{1}H dipolar coupling of these oxygen sites where the protons are ~1.0 Å away. Consequently, the lines are not narrowed by the outer rotation rate of up to 1800 Hz with the decoupling power available. As with *L*-alanine the differing responses of the ^{17}O DOR line widths to ^{1}H decoupling allows assignment of these two lines. The line widths of the two carbonyl oxygens with no decoupling were 290 ± 15 and 360 ± 20 Hz, and this already implies that the line with the larger line width is for an oxygen closer to protons. Inspection of the structure (36), given in Figure 8, shows that the nearest proton to O2 is at 1.96 Å whereas the closest H to O3 is at 1.62 Å. The narrower line at ~258 ppm in the DOR spectrum narrows by around 30% on the application of 28 kHz ^{1}H decoupling, whereas the line width of the broader line at ~249 ppm is still greater than the undecoupled line width. Thus, we can assign these to O2 and O3, respectively.

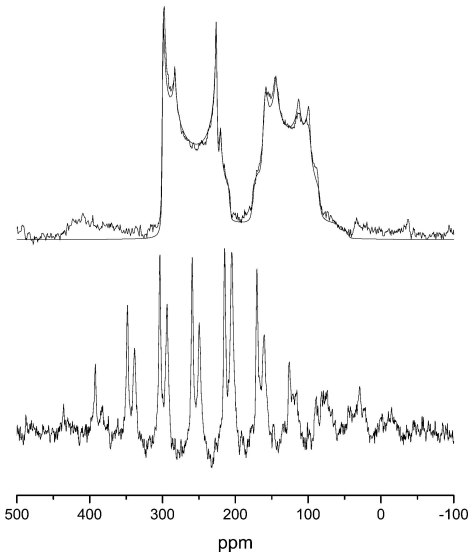


Figure 7. ^{17}O NMR (14.1 T) spectra from *D*-glutamic acid hydrochloride: (a) MAS together with simulation of the centerband and (b) DOR, outer rotor speed 1800 Hz.

DISCUSSION

The data on the *D*-glutamic acid allow a direct comparison with those of the *L*-glutamic acid that was recently analyzed in detail (26). Within error, the parameters deduced from the normal MAS spectra (which can be accurately determined because of the high degree of detail within the spectra) appear similar. Although DOR averages the second-order interactions, and hence some of the information that would distinguish sites is lost, the production of narrow well-resolved resonances provides further confirmation of the similarity of the NMR parameters of these samples. Despite being able to determine the position to an accuracy of ~1 ppm, it appears that the shifts of the two carbonyl oxygens are the same in the two forms.

A rather different ^{17}O NMR spectrum of glutamic acid hydrochloride was presented by Wu et al (47). The spectrum was analyzed as having a line at 170 ppm with χ_Q = 7.2 MHz and η = 0.20 and another at 320 ppm with χ_Q = 8.2 MHz and η = 0, with the other two sites having very similar shifts of ~250 ppm with χ_Q = 6.8 MHz and η = 0.58. Although the two outer lines are at positions similar to those found for the OH sites, O1 and O4, and CdO sites, O2 and O3, respectively, the lines at 250 ppm are not observed here. Lemaitre et al. (26) suggested that, since the sample of Wu was enriched

from *D,L*-glutamic acid monohydrate, it may have been racemic *D,L*-glutamic acid hydrochloride. (*D,L*-Glutamic acid monohydrate was used by Dunitz et al. (10) to produce racemic anhydrous *D,L*-glutamic acid.) Unfortunately, there is no structure in the literature for racemic *D,L*-glutamic acid hydrochloride, but interestingly, a recent calculation of the NMR parameters for anhydrous *D,L*-glutamic acid (52) does predict that one site, O2, will have parameters, 256 ppm, 6.8 MHz, and η = 0.72, similar to those of the intermediate line observed by Wu et al.

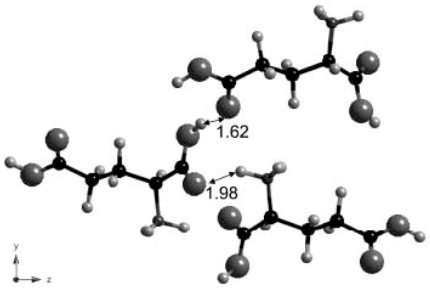


Figure 8 Structure of *L*-glutamic acid hydrochloride (36) showing the distance from the carbonyl atoms to the nearest proton.

The range of ^{17}O NMR parameters from the amino acid hydrochlorides shows some very interesting features (Table 1), with the OH and carbonyl oxygens showing quite clearly defined ranges of the parameters. For hydroxyls χ_Q lies in the range 7.29-7.67 MHz and $\delta_{cs,iso}$ varies from 172.5 to 187 ppm, while for the carbonyls χ_Q lies in the range 8.16-8.65 MHz and $\delta_{cs,iso}$ varies from 315 to 353.5 ppm. The asymmetry parameters are also closely defined, and are significantly different for the hydroxyl and carbonyl oxygens, indicating the differences in the bonding between these two sites. Generally, the OH sites tend to exhibit asymmetry parameters significantly removed from axial symmetry in the range 0.19-0.30, indicating that the proton is a little displaced from the line joining the oxygens. The carbonyl sites are generally close to axial symmetry, typically in the range 0.00-0.06, showing that the CdO bond is strongly defined. The O3 carboxyl site in glutamic acid has η = 0.17, probably because the relatively close H on the adjoining molecule (at 1.62 Å) is affecting the bonding.

Despite the quite tightly defined ranges of these parameters, the high-quality spectra that are obtained mean that the different amino acids can be readily distinguished. So, for example, mixtures of two amino acids could certainly be separated in both the MAS spectra, from the features of the second-order quadrupole line shape, and from the different net shifts (eq 1) in the DOR, producing

separate resonances. As the variation of the NMR parameters between the different amino acids is much greater than the accuracy with which they can be measured, the variations can be used as a sensitive probe of the structure. One of the key steps now necessary is to further develop this probe to establish links between the NMR parameters and structural motifs. For the carbonyl oxygens there is a strong correlation (r = –0.93) between $\delta_{cs,iso}$ and the C-O bond length, which is shown in Figure 9. As the C-O bond length increases the shift decreases with a slope of ~–1200 ppm/Å. There is also a correlation that is not as good between the electric field gradient and C-O bond length. It is interesting that although the asymmetry parameter is close to zero for C-O distances ≤ 1.22 Å, with longer bond lengths it seems to increase rapidly, indicating increased hydrogen bonding.

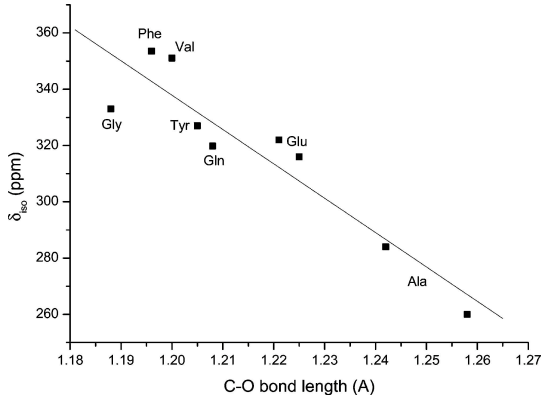


Figure 9 Correlation of $\delta_{cs,iso}$ with the C-O bond length for the carbonyl oxygen in amino acids for which structural data are available.

An exciting possibility for this type of work lies in recent improvements in full electronic calculations of the NMR parameters using the gauge-included projector augmented wave method that uses pseudopotentials in a periodic structure rather than having to use a cluster approximation (30). As these calculations are based on the crystal structures, they provide a direct link between the structural features and the NMR parameters. It will be shown in a separate paper that calculations of the NMR parameters are in very good agreement with the measured values (52).

Another question that these experimental data allow to be addressed is the likely capability of ^{17}O NMR to probe more complex biomolecules containing amino acids as building blocks. Given the highly defined line shapes, spectra made up of several different amino acids could be deconvolved under MAS. DOR also provides an important alternative since, with sufficient ^{1}H decoupling,

highly resolved spectra are obtained with line widths of only ~1-2 ppm. Thus, samples with many inequivalent oxygen sites could be resolved under these conditions. Also, as is common practice for ^{13}C and ^{15}N studies of biomolecules, site-specific ^{17}O labeling could be carried out. For most samples studied here ≤ 20 atom % enriched H_2O was used; however, this could be increased to ≥ 75 atom % enrichment. This, in combination with higher magnetic fields and application of techniques that enhance the signal either by manipulation of the satellite transition populations (28) or possibly through the application of a train of echoes (e.g., QCPMG) (24), means that structurally dilute oxygen sites could be observed. It is entirely feasible that biomolecules comprising 50-100 amino acid residues enriched at one position could be observed by a combination of MAS and DOR. This capability would bring many molecules of biological importance into the compass of solid state ^{17}O NMR.

CONCLUSION

^{17}O NMR data can be readily collected at reasonably high fields (14.1 T) from amino acids with good signal-to-noise ratios from samples where only 10-20 atom % ^{17}O enriched water was used in the sample preparation. Typical ^{17}O MAS NMR spectra consist of well-separated second-order quadrupole line shapes from the carbonyl and hydroxyl oxygens. The highly detailed MAS centerbands obtained mean that the NMR parameters can be accurately deduced and variations between the different amino acids readily determined. DOR produces significant additional narrowing over the MAS spectra, typically a factor ~40, so that high-resolution spectra with line widths of ~1 ppm are obtained. A combination of MAS, DOR, and 3Q MAS can provide a high accuracy of the NMR parameters, even for samples for which there is significant overlap between the different resonances in the MAS NMR spectra. *L*-Alanine shows a spectrum very different from those of the amino acid hydrochlorides because of the very different bonding arrangements at the different oxygen sites in the two forms. Variation of the DOR line width with ^{1}H decoupling power allows the assignment of different carbonyls because of their differing spatial proximity to the protons. Decoupling powers significantly greater than the 34 kHz used here will be necessary to narrow the ^{17}O resonances from hydroxyl oxygens under DOR. A strong correlation is observed

between the chemical shift $\delta_{cs,iso}$ and the C-O bond length so that ^{17}O has significant potential to provide new information about bonding in these systems. The sensitivity and resolution obtained provide much optimism that ^{17}O NMR data could usefully be collected from much larger biomolecules, for example, to provide information about site-specific binding and receptor-ligand interactions.

ACKNOWLEDGEMENTS The EPSRC is thanked for supporting the NMR equipment at Warwick and the work on biological materials through Grant GR/N29549. The Royal Society has supported the collaboration with Tallinn, Estonia, and M.E.S. also thanks the Royal Society and Leverhulme Trust for the award of a Senior Research Fellowship 2001-2002 under which some of this work was carried out. A.W. and V.L. thank BBSRC for support and Dr. E. Hughes for advice. The Estonian Science Foundation is also thanked for support, and Dr. S. P. Brown is thanked for collecting the ^{13}C CP MAS NMR data on the alanine samples.

REFERENCES

1. Al-Karaghouli AR, Koetzle TF. 1975. Neutron diffraction study of L-phenylalanine hydrochloride. *Acta Crystallographica Section B* 31: 2461
2. Anupõld T, Reinhold A, Sarv P, Samoson A. 1998. A comparison of double rotation and multi-quantum magic angle spinning spectra. *Solid State Nucl Magn Reson* 13: 87
3. Ashbrook SE, Berry AJ, Wimperis S. 2001. ^{17}O multiple-quantum MAS NMR study of high-pressure hydrous magnesium silicates. *J. Am. Chem. Soc.* 123: 6360–6
4. Bull LM, Bussemer B, Anupõld T, Reinhold A, Samoson A, et al. 2000. A High-Resolution ^{17}O and ^{29}Si NMR Study of Zeolite Siliceous Ferrierite and ab Initio Calculations of NMR Parameters. *J Am Chem Soc* 122: 4948
5. Detken A, Hardy EH, Ernst M, Meier BH. 2002. *Chem. Phys. Lett.* 356: 298
6. di Blasio B, pavone V, Pedone C. 1977. *Cryst. Struct. Comm.* 6: 745
7. Dirken PJ, Kohn SC, Smith ME, van Eck ERH. 1997. Complete resolution of Si---O---Si and Si---O---Al fragments in an aluminosilicate glass by ^{17}O multiple quantum magic angle spinning NMR spectroscopy. *Chemical Physics Letters* 266: 568–574
8. Dirken PJ, Smith ME, Whitfield HJ. 1995. ^{17}0 and ^{29}Si Solid State NMR Study of Atomic Scale Structure in Sol-Gel-Prepared TiO_2-$Si0_2$ Materials. *Journal of Physical Chemistry* 99: 395–401
9. Dong S, Yamada K, Wu G. 2000. Oxygen-17 Nuclear Magnetic Resonance of Organic Solids. *Zeitschrift für Naturforschung* 55a: 21–28
10. Dunitz J, Schweizer W. 1995. Anhydrous *D,L*-glutamic acid. *Acta Cryst.* C51: 1377
11. Frey MN, Koetzle TF, Lehmann MS, Hamilton WC. 1973. Precision neutron diffraction structure determination of protein and nucleic acid components. X. A comparison between the crystal and molecular structures of L-tyrosine and L-tyrosine

hydrochloride. *Journal of Chemical Physics* 58: 2547

12. Gann SL, Baltisberger JH, Wooten EW, Zimmermann H, Pines A. 1994. Cross Polarization and Dynamic-Angle Spinning of ^{17}O in L-Alanine. *Bull. Magn. Reson.* 16: 68–72
13. Gervais C, Babonneau F, Smith ME. 2001. Detection, quantification and magnetic field dependence of solid state ^{17}O NMR X-O-Y (X,Y =Si,Ti) linakges: Implications for characterising amorphous titania-silica based materials. *Journal of Physical Chemistry B* 105: 1971–1977
14. Goc R, Ponnusamy E, Tritt-Goc J, Fiat D. 1988. ^{17}O n.m.r. studies of amino acids in the solid state, in single- and polycrystalline forms. *Int. J. Pept. Protein Res.* 31: 130–136
15. Goc R, Tritt-Goc J, Fiat D. 1989. ^{17}O NMR Study of Polycrystalline *L*-Leucine. *Bulletin of Magnetic Resonance* 11: 238–243
16. Godbout N, Sanders KL, Salzmann R, Havlin RH, Wojdelski M, Oldfield E. 1999. Solid-State NMR, Mössbauer, Crystallographic, and Density Functional Theory Investigation of Fe-O_2 and Fe-O_2 Analogue Metalloporphyrins and Metalloproteins. *J. Am. Chem. Soc.* 121: 3829–3844
17. Howes AP, Jenkins R, Smith ME, Crout DHG, Dupree R. 2001. Influence of low-barrier hydrogen bonds on solid state ^{17}O NMR spectra of labelled phthalate species. *Chemical Communications*: 1448–1449
18. Hussin R, Dupree R, Holland D. 1999. *J . Non-Cryst. Solids* 246: 159
19. Koetzle TF, Golic L, Lehmann MS, Verbist JJ, Hamilton WC. 1974. Precision neutron diffraction structure determination of protein and nucleic acid components. XV. Crystal and molecular structure of the amino acid L-valine hydrochloride. *Journal of Chemical Physics* 60: 4690
20. Kroeker S, Rice D, Stebbins JF. 2002. *Am. Mineral* 87: 572
21. Kunwar AC, Turner GL, Oldfield E. 1986. *J. Magn. Reson.* 69: 124
22. Kuroki R, Ando I, Shoji A, Ozaki T. 1992. A Structural Study of Polyglycine II in the Solid State by ^{17}O CP MAS NMR Spectroscopy. *Chemical Communications*: 433–434
23. Kuroki S, Takahashi A, Ando I, Shoji A, Ozaki T. 1994. Hydrogen-bonding Structural Study of Solid Peptides and Polypeptides containing a Glycine Residue by ^{17}O NMR Spectroscopy. *J. of Mol. Struct.* 323: 197–208
24. Larsen FH, Farnan I. 2002. *Chem. Phys. Lett.* 357: 403
25. Lehmann M, Koetzle T, Hamilton W. 1972. Pecision neutron diffraction structure determination of protein and nucleic acid components. VIII: the crystal and molecular structure of the beta-forms of the amino acid L-glutamic acid. *J. Cryst. Mol. Struct.* 2: 225
26. Lemaître V, Pike KJ, Watts A, Anupõld T, Samoson A, et al. 2003. New Insights into the Bonding Arrangements of *L*- and *D*-Glutamates from Solid State ^{17}O NMR. *Chem. Phys. Lett.* 371: 91–97
27. MacKenzie KJD, Smith ME. 2002. *Multinuclear Solid State NMR of Inorganic Materials*. Oxford: Pergamon Press
28. Madhu PK, Pike KJ, Dupree R, Levitt MH, Smith ME. 2003. *Chem. Phys. Lett.* 367: 150
29. Massiot D, Fayon F, Capron M, King I, Le Calvi S, et al. 2002. Modelling one- and two-dimensional solid-state NMR spectra. *Magnetic Resonance in Chemistry* 40: 70
30. Pickard CJ, Mauri F. 2001. All-electron magnetic response with pseudopotentials: NMR chemical shifts. *Phys. Rev. B. Condensed Matter* 63: 245101
31. Pickup DM, Mountjoy G, Wallidge GW, Anderson R, Cole JM, et al. 1999. *J. Mater. Chem.* 9: 1299
32. Pike KJ, Malde RP, Ashbrook SE, McManus J, Wimperis S. 2000. Multiple-quantum MAS NMR of quadrupolar nuclei. Do five-, seven- and nine-quantum experiments yield higher resolution than the three-quantum experiment? *Solid State Nucl Magn Reson* 16: 203
33. Pingel UT, Amoureux J-P, Anupold T, Bauer F, Ernst H, et al. 1998. *Chem. Phys. Lett.* 294: 345
34. Samoson A, Lippmaa E. 1989. Synchronized Double-Rotation NMR-

Spectroscopy. *Journal of Magnetic Resonance* 84: 410–416
35. Seliger J, Zagar V, Blinc R, Kadaba PK. 1990. ^{1}H-^{17}O Nuclear Quadrupole Double Resonance in *D,L*-Proline. *Zeitschrift für Naturforschung* 45a: 733–735
36. Sequeira A, Rajagopal H, Chidambaram R. 1972. A neutron diffraction study of the structure of L-glutamic acid.HCl. *Acta Cryst.* B28: 2514–2519
37. Shamala N, Venkatesan K. 1972. *Cryst. Struct. Commun.* 1: 227
38. Smith ME, van Eck ERH. 1999. Recent advances in experimental solid state NMR methodology for half-integer spin quadrupolar nuclei. *Progress in NMR Spectroscopy* 34: 159–201
39. Stebbins JF, Zhao PD, Lee SK, Oglesby JV. 2001. *J. Non-Cryst. Solids* 293: 67
40. Steinschneider A, Burgar MI, Buku A, Fiat D. 1981. Labeling of amino acids and peptides with isotopic oxygen as followed by ^{17}O-N.M.R. *Int. J. Pept. Protein Res.* 18: 324–33
41. Torrent M, Mansour D, Day EP, Morokuma K. 2001. *J. Phys. Chem. A* 105: 4546
42. Torres J, Kukol A, Arkin IT. 2000. Use of a single glycine residue to determine the tilt and orientation of a transmembrane helix. A new structural label for infrared spectroscopy. *Biophys. J.* 79: 3139–3143
43. Valentine B, Steinschneider A, Dhawan D, Burgar MI, St Amour T, Fiat D. 1985. Oxygen-17 n.m.r. of peptides. *Int. J. Pept. Protein Res.* 25: 56–68
44. Watts A. 1999. NMR of drugs and ligands bound to membrane receptors. *Current Opinion in Biotechnology* 10: 48
45. Watts A. 1999. Structural Resolution of Ligand–Receptor Interactions in Functional, Membrane-embedded Receptors and Proteins using Novel, Non-perturbing Solid-state NMR Methods. *Pharmacy and Pharmacology Communications* 7: 7
46. Wi S, Logan JW, Sakellariou D, Walls JD, Pines A. 2002. *J. Chem. Phys.* 117: 7024
47. Wu G, Dong S. 2001. Two-dimensional ^{17}O multiple quantum magic-angle spinning NMR of organic solids. *J. Am. Chem. Soc.* 123: 9119–9125
48. Wu G, Dong S, Ida R, Reen N. 2002. A solid-state ^{17}O nuclear magnetic resonance study of nucleic acid bases. *J. Am. Chem. Soc.* 124: 1768–77
49. Wu Y, Sun BQ, Pines A, Samoson A, Lippmaa E. 1990. NMR Experiments with a New Double Rotor. *Journal of Magnetic Resonance* 89: 297
50. Yamauchi K, Kuroki S, Ando I. 2002. High-Resolution Solid-State ^{17}O NMR Studies of Polyglycines and their Hydrogen-Bonded Structures. *J. of Mol. Struct.* 602–603: 171–175
51. Yamauchi K, Kuroki S, Ando I, Ozaki T, Shoji A. 1999. ^{17}O NMR chemical shifts and quadrupole coupling constants in solid poly(*L*-alanine)s determined using a high-speed MAS technique. *Chem. Phys. Lett.* 302: 331–336
52. Yates JR, Pickard CJ, Payne MC, Dupree R, Profeta M, Mauri F. 2004. Theoretical Investigation of Oxygen-17 NMR Shielding and Electric Field Gradients in Glutamic Acid Polymorphs. *J. Phys. Chem. A* 108: 6032–6037
53. Zhao PD, Kroeker S, Stebbins JF. 2000. *J. Non-Cryst. Solids* 276: 122

A.2 Molecular dynamics simulations on the first two helices of Vpu from HIV-1

Issara Sramala,[1,2] Vincent Lemaître,[1,3] José D. Faraldo-Gómez,[4] Sébastien Vincent,[3] Anthony Watts,[1] Wolfgang B. Fischer,[1]

[1] *Biomembrane Structure Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK ; email: wolfgang.fischer@bioch.ox.ac.uk*

[2] *Present address: Institute of Molecular Biology and Genetics, Mahidol University, Salaya Campus, Nakornpathom, Thailand*

[3] *Nestec S.A., BioAnalytical Department, Vers-Chez-Les-Blanc, CH-1000 Lausanne 26, Switzerland*

[4] *Laboratory of Molecular Biophysics, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, UK*

Keywords HIV-1, Vpu, viral membrane protein, molecular dynamics simulations, protein structure.

■ **Abstract** Vpu is an 81 amino acid protein of HIV-1 with two phosphorylation sites. It consists of a short N-terminal end traversing the bilayer and a longer cytoplasmic part. The dual functional role of Vpu is attributed to these topological distinct regions of the protein. The first 52 amino acids of Vpu (HV1H2) have been simulated, which are thought to be embedded in a fully hydrated lipid bilayer and to consist of a transmembrane helix (helix-1) connected via a flexible linker region, including a Glu-Tyr-Arg (EYR) motif, with a second helix (helix-2) residing with its helix long axis on the bilayer surface. Repeated molecular dynamics simulations show that Glu-28 is involved in salt bridge formation with Lys-31 and Arg-34 establishing a kink between the two helices. Helix-2 remains in a helical conformation indicating its stability and function as a peptide "float", separating helix-1 from the rest of the protein. This leads to the conclusion that Vpu consists of three functional modules, helix-1, helix-2 and the remaining residues towards the C-terminal end.

INTRODUCTION

Viral genomes encode for a series of short membrane proteins with an average length of ~ 100 amino acids (10, 50). Most of these proteins fulfil their role to enable and/or improve the efficiency of viral replication (M2, NB, CM2, Vpu et al.) (21), whereas others are involved in the construction of, for example, protein shells to wrap the viral genome (Pf1, M13, and others). A common theme of all of these small proteins is that some of them have at least two functional roles and can therefore be seen as multi-functional tools.

Vpu is encoded only in the genome of HIV-1 (10, 50) together with accessory proteins such as, Vif, Vpr, and Nef. Vpu is found especially in the endoplasmic reticulum and is neither found in the virion nor in the cell membrane of the infected cell (49). It has a dual role in the life cycle of HIV-1, which are (i) to enhance viral particle release (15, 41, 46) and (ii) to be involved in the degradation of the HIV-1 receptor protein CD4 (5, 45, 49, 59). Although it is found that for the latter role the cytoplasmic part of Vpu is responsible, enhanced particle release is thought to be due to channel activity initiated by the assembly of a few Vpu proteins, which is dependent on the transmembrane (TM) part (17, 34, 46). To substantiate this, mutations in the TM part show a lower rate of particle secretion (51).

There is a significant amount of structural information available for the 81 amino acid protein Vpu. This information is based of investigations of peptides analogous of Vpu, and full length Vpu derived from NMR (9, 19, 25, 31, 33, 58, 61, 62), CD- (61), and FTIR-spectroscopy (29). In summary, it has been concluded that Vpu has a short N-terminal extramembraneous end and a TM spanning helix that is tilted by ~ 10º - 20º. A loop connects the TM helix with a second helix (helix-2) that resides with its helix long axis parallel to the membrane surface. Another flexible part, including the two phosphorylation sites (Ser-52 and Ser-56), connects helix-2 with a third helix. Towards the C-terminal end another short helix (19) or a turn (58) exists, depending on the conditions under which the data were recorded.

MD simulations play an important role in linking structural information from experiments with the mechanisms of the protein on an atomic scale. Investigations on K+ channels may stand asone of the most comprehensive examples in this respect (reviewed in (44)), In the case of Vpu, MD simulations have dealt so far with the analysis of the TM part of Vpu (11, 12, 23, 30, 38). Up to 6 TM helices have been aligned in parallel to form an assembly of homo-oligomers embedded in a low dielectric slab (23), an octane slab as a bilayer mimic (30, 38), and a fully hydrated lipid bilayer (11, 12, 30). All studies have in common a fairly good agreement with structural data from experiment. They also suggest that a pentameric bundle should be the minimum assembly for Vpu to form an ion conducting pore (12, 23, 30).

In an earlier computational study the TM helix (helix-1) of Vpu embedded in a fully hydrated lipid bilayer (20) has been investigated. In this study the focus is on an extended model $Vpu_{1–52}$ including helix-2 laying on top of the lipid membrane. It has been proposed that amino acids Glu-28, Tyr-29, and Arg-30 (EYR-motif) play a key role in the bend of the strand. Up to now detailed structural data on this region are not available. Driven by the good agreement of the recent computational data with those from spectroscopy, the role of these particular amino acids on the stability of the protein solely based on the simulations can be proposed. MD simulations have been repeated to assess the reliability of the data acquisition (56). The data will be discussed against the background of computational investigations on other viral membrane proteins such as the Pf1 coat protein from filamentous bacteriophage (37, 43, 52), which adopt a similar kinked like structure. In addition, the idea that Vpu protein consists of “functional modules” is addressed.

MATERIALS AND METHODS

A helical model of the first 52 amino acids of VPU WAS GENERATED, Vpu_{1-52}:

QPIPIVAIVA[10] LVVAIIIAIV[20] VWSIVIIEYR[30] KILRQRKIDR[40] LIDRLIERAE[50] DS

using a combined simulated annealing and molecular dynamics simulations (SA/MD) approach applying the program Xplor (7). An extended description of the procedure is given in detail elsewhere (28). The procedure is summarized in brief. The SA/MD protocol comprises two stages.

In Stage 1 an idealized α-helix is constructed on the basis of the backbone Cα-atoms of the peptide. All other atoms of the individual amino acids are superimposed on their particular Cα atoms. During Stage 1 the side chain atoms "evolve" from the Cα-atoms whilst retaining the Cα-atoms in fixed positions. Beginning the annealing at 1000 K, the weights for bond length, bond angles, planarity and chirality are gradually increased. A repulsive van der Waals term is slowly introduced after an initial delay up to 80 % of the original values. This allows the atoms to pass each other. 5 structures for each helix were obtained. Each helix from Stage 1 is used for 5 molecular dynamics runs in Stage 2. In Stage 2, initial velocities correspond to 500 K. Harmonic restraints hold the Cα-atoms, but are relaxed as the temperature drops from 500 to 300 K. At this stage distance restraints are introduced and the Cα-atoms are allowed to move. In Stage 2 electrostatic interactions are introduced into the potential energy function. The main-chain atoms obtain their charges corresponding to the PARAM19 parameter set. Partial charges on side-chain atoms of polar side chains are gradually scaled up (from 0.05 to 0.4 times their full value) during the temperature reduction from 500 to 300 K. The scaling factor 0.4 is also applied during the 5-ps dynamics and energy minimization. A distance dependent dielectric function is used (7), with a switching function smoothly truncating distant electrostatic interactions. In Stage 2, 5 x 5 = 25 helices are obtained and the most straight helix is used to run the simulations.

The chosen helix was manually bent around residues Glu-28 to Ile-32 using the program Swiss PDB-Viewer to adopt the helix-loop-helix motif. The Φ- and Ψ-values (Φ/Ψ in °) were intended to be held in a helical conformation: Glu-28: -70.4/0.2; Tyr-29: -65.2/-42.5; Arg-30: -84.7/-13.8; Lys-31: -86.2/10.0; Ile-32: -57.6/-18.0 (all values after the bend). This procedure allowed the Cα-atom of residue Asp-39 to point towards the bilayer surface and the Cα-atom of residue Arg-48 to point away from the surface (see also (25)).

A hole in a lipid bilayer consisting of 288 POPC (1-palmitoyl-2-oleoyl-*sn*-glycerol-3-phosphatidyl-choline) molecules, was created by overlaying a bent peptide model with the lipid bilayer. Lipid molecules for which the phosphorous atoms were overlapping with the GRASP-representation of the peptide were removed (18). This leads to a removal of an unequal number of lipids on both sides of the bilayer resulting in 125 lipids in the leaflet with helix-2 and 136 lipids in the other leaflet (261 lipids in total) with 13'572 lipid atoms. After insertion of Vpu_{1-52} (529 peptide atoms) the

peptide / lipid system (Figure 1A) was hydrated with 16'169 water molecules (48507 water atoms). The overall number of atoms in the simulation at this stage is 62'610, including 2 Cl^- ions to neutralize the simulation box. To allow the lipid molecules next to helix-2 to fill the "hole" beneath it, 600 ps of equilibration at 300 K with the protein restrained is performed (Figure 1B). Thereby the peptide coordinates were restrained. The production phase was run at 300 K for up to 9 ns and twice up to 3 ns, each starting with randomly chosen initial velocities.

Gromacs 2.0 and 3.0 software (http://www.gromacs.org) was used including the SPC water model (3). Simulations were run on either a VALinux Beowulf cluster of four dual Pentium III 700-MHz processors or a DELL OptiPlex GX1 with a Pentium III 450-MHz processor. The simulations used a time step of 2 fs and a LINCS algorithm to keep the geometry of the molecules. A isothermal-isobaric ensemble (NPT) was used with periodic boundaries and anisotropic pressure coupling (Gromacs 2.0) or Berendsen temperature and pressure coupling (Gromacs 3.0). Long-range electrostatic interactions have been calculated with the particle-mesh Ewald (PME) method. Lennard–Jones and short-range Coulombic interactions were cut off at 1.0 and 0.9 nm, respectively.

RESULTS

The evolution of the adaptation of $Vpu_{1–52}$ to the lipid membrane environment is shown in Figure 1A. At the beginning, the angle between the TM helix (helix-1) and helix-2 is ~ 100°, with helix-2 almost parallel to the membrane plane. In this configuration, Ser-23 is not buried under helix-2. Helix-2 resides on top of the lipid head-group region (grey spheres in Fig. 1A). The snapshots taken at 3 and 6 ns imply a wavelike motion of $Vpu_{1–52}$ of being partially ejected from the bilayer and driven back into it. At 9 ns helix 2 is found to be embedded in the lipid head group region, as would be expected for its partially hydrophobic character along the membrane facing side of the helix. Helix-1 adopts a considerable kink in certain time steps. Throughout the simulation the lipids remain fairly packed beneath helix-2 (Fig. 1B) allowing the hydrophobic part of the lipid molecules to be in contact with the bilayer facing hydrophobic side of helix-2. The lipid head groups do not manoeuvre underneath helix-2 but rather stay away from helix-2.

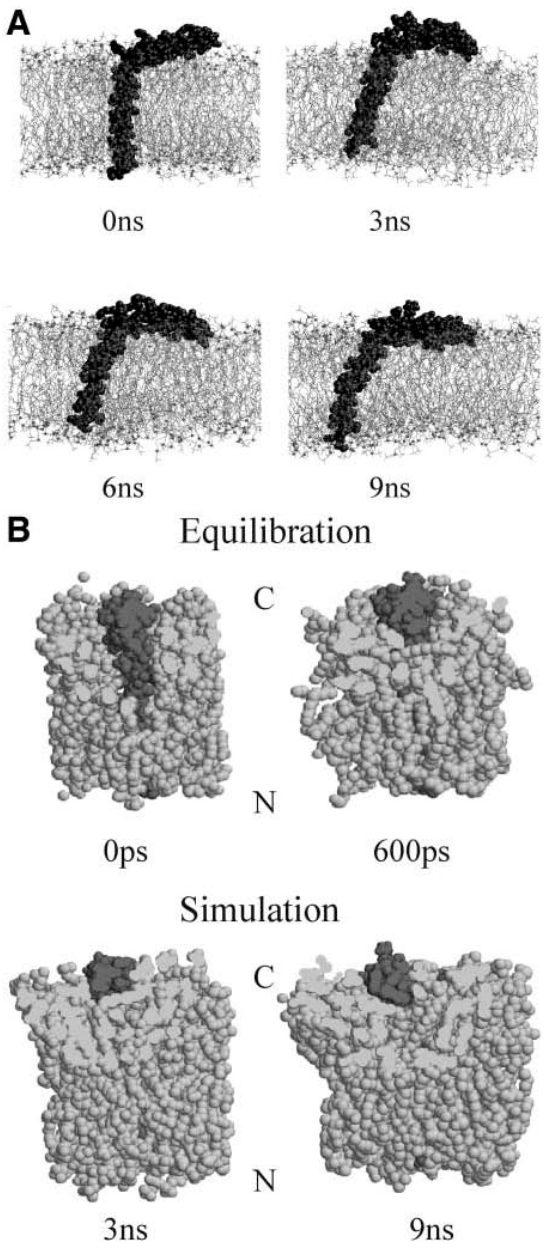


Figure 1 Snapshots of the protein (black) embedded in a lipid bilayer (thin grey chains, grey spheres represent the phosphorous headgroups) at 0, 3, 6, and 9 ns (A). View from the C- to the N-terminal end of Vpu along the axis of helix-2 (dark) lying on top of the bilayer (grey) before (0 ps) and after 600 ps of equilibration with the protein (dark) restraint (B). Snaphots taken from the simulations at 3 ns and 0 ns are shown in the lower half. The figures are created with the software Rasmol in the "slab" mode showing the atoms with their van der Waals radii. Plain grey spots are sliced atoms due to the "slab" mode representation.

The RMSD for the Cα atoms for the simulation over 9 ns levels off after ~ 0.5 ns reaching values between 0.2 and 0.25 nm (black trace in Fig. 2). The trace shows three maxima at around 2.5, 5 and

6 ps. Values for a second and third run over 3 ns behave in a similar way, levelling off after ~ 0.5 ns but at slightly higher and lower values (grey traces in Fig. 2) compared to the run over 9ns.

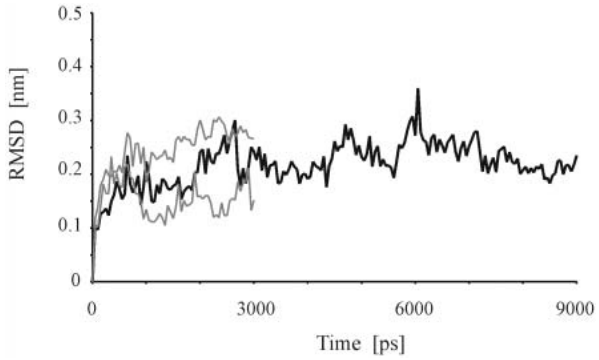


Figure 2 Root mean square deviation (RMSD) of the Cα-atoms of the simulation over 9 ns (black trace, thick line) and two simulations each over 3 ns (grey and black thin lines).

The interconnecting region of the two helices is highlighted in Figure 3. At the beginning of the simulation Asp-39 is pointing towards the hydrophobic region of the bilayer, whereas Arg-44 stretches into the aqueous environment (Fig. 3A). After 3 ns both residues reverse their position with Asp-39 now facing the aqueous phase and Arg-44 the lipid bilayer. This revolution is accompanied by the unwinding of parts of the region interconnecting the two helices (Fig. 3 B). Although residues Glu-28 (light grey, Fig. 3B) and Tyr-29 (grey, Fig. 3B) remain within a helical conformation of the backbone, Arg-30 (black, Fig. 3B) and Lys-31 (black, Fig. 3B) seem to be involved in the unwinding process. Arg-34 (light grey, Fig. 3B) is back in a helical conformation. A more thorough analysis of the ϕ and ψ values of these residues indicate that Glu-28 and Tyr-29 do not deviate from standard values for helices (6) during the entire duration of the 9-ns simulation (Fig. 4). Arg-30 is involved in the unwinding process since its ϕ and ψ values undergo large deviations especially within the first half of the simulation. The values for Lys-31 rapidly adopt values of around –100° for ϕ and ~ 150° for ψ within the first 200 ps. The ϕ and ψ values for Ile-32 fluctuated around ~ –100° (for ϕ) and ~ –50° (for ψ) and Leu-33 returned almost completely to a helical conformation. Averaged data for 0, 3, and 9 ns are given in Table 1. In the second and third run (both 3 ns) all traces match those shown in Figure 4 for the first 3 ns. At the C- and the N-terminal end no significant unwinding of the helices is observed.

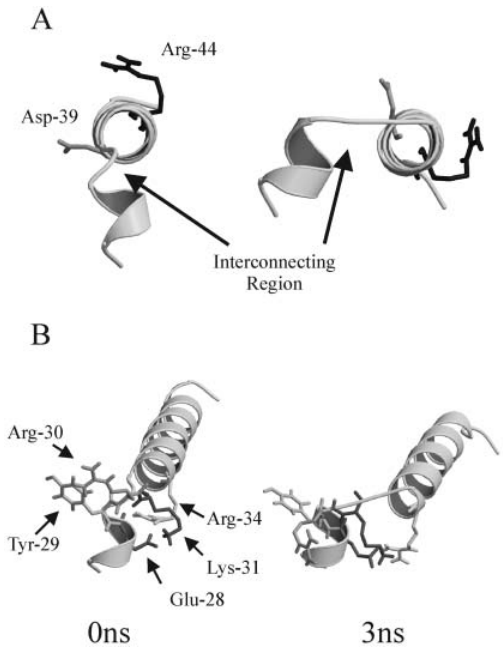


Figure 3 View from the N- to the C-terminal end of helix-2 along its long axis (A). Highlighted are residues Asp-39 (grey) and Arg-44 (black). View from top onto the kink between hlix-1 and helix-2 (B). Residues shown are Glu-28 (light grey), Tyr-29 (grey), Arg-30(black), Arg-34 (light grey). In both panels, lipids and water molecules are omitted for clarity. For data representation the Molscript software is used.

The helical environment around Glu-28 to Arg-30 is supported by a salt bridge between Glu-28 and Lys-31, shown in Figure 3B, with a distance of around 0.15 nm during the entire duration of the 9-ns simulation (black trace in Fig. 5). Arg-34 also forms a salt bridge with Glu-28 (grey trace in Fig. 5) after 2 ns forming a "complex salt bridge" (39). In repeated simulations, the salt bridge between Glu-28 and Lys-31 undergoes a more rapid adaptation of an average distance of 0.15 nm compared to the salt bridge formation by Glu-28 with Arg-34 (Fig. 5, B and C). It seems that the formation of the salt bridge between Glu-28 and Arg-34 triggers or at least supports the revolution of helix-2, as mentioned, accompanied with the unwinding of the helical structure between Arg-30 and Ile-32.

As a result, Glu-28 and Tyr-29 are unlikely to be the key residues forming the bend between helix-1 and helix-2. Arg-30 is not in a helical conformation for short time steps during the simulation, which supports earlier findings (12) in that it is most likely fulfilling its role as a flexible residue. However Lys-31 and Ile-32 seem to be involved in forming the kink, since they are largely diverting from the normal values for an ideal helix over the entire duration of the simulation.

Table 1: Simulated structural data of Vpu_{1-52}.

		Helix 1				Helix 2			
Runs		1st	2nd	3rd	Average	1st	2nd	3rd	Average
Tilt angle [°]	0ns	4.3 (3.8)	8.6 (3.6)	7.0 (3.2)	6.6 (3.8)	7.5 (2.1)	-2.3 (4.8)	4.3 (2.3)	3.1 (5.2)
	3ns	16.3 (2.3)	21.1 (1.7)	18.0 (1.9)	18.5 (2.8)	-15.4 (2.0)	-4.3 (1.5)	2.3 (3.1)	-5.8 (7.8)
	9ns	23.3 (2.1)	-	-	-	-10.2 (2.0)	-	-	-
Kink angle [°]	0ns	5.8 (3.8)	13.0 (1.2)	8.6 (4.3)	9.1 (4.4)	4.8 (3.2)	13.0 (1.5)	3.1 (1.5)	6.9 (4.9)
	3ns	7.1 (2.9)	15.4 (1.3)	4.2 (1.6)	8.9 (5.3)	13.0 (3.3)	11.0 (0.7)	2.6 (0.5)	8.9 (5.0)
	9ns	7.2 (2.6)	-	-	-	7.0 (3.9)	-	-	-

Data results from an average of five data sets recorded at 0-400, 2600-3000, and 8600-9000 ps in steps of 100 ps including the standard deviation shown in brackets. The average values derive from 15 data in the particular time frame and their standard deviation. For the 2nd and 3rd runs data are only available up to 3 ns.

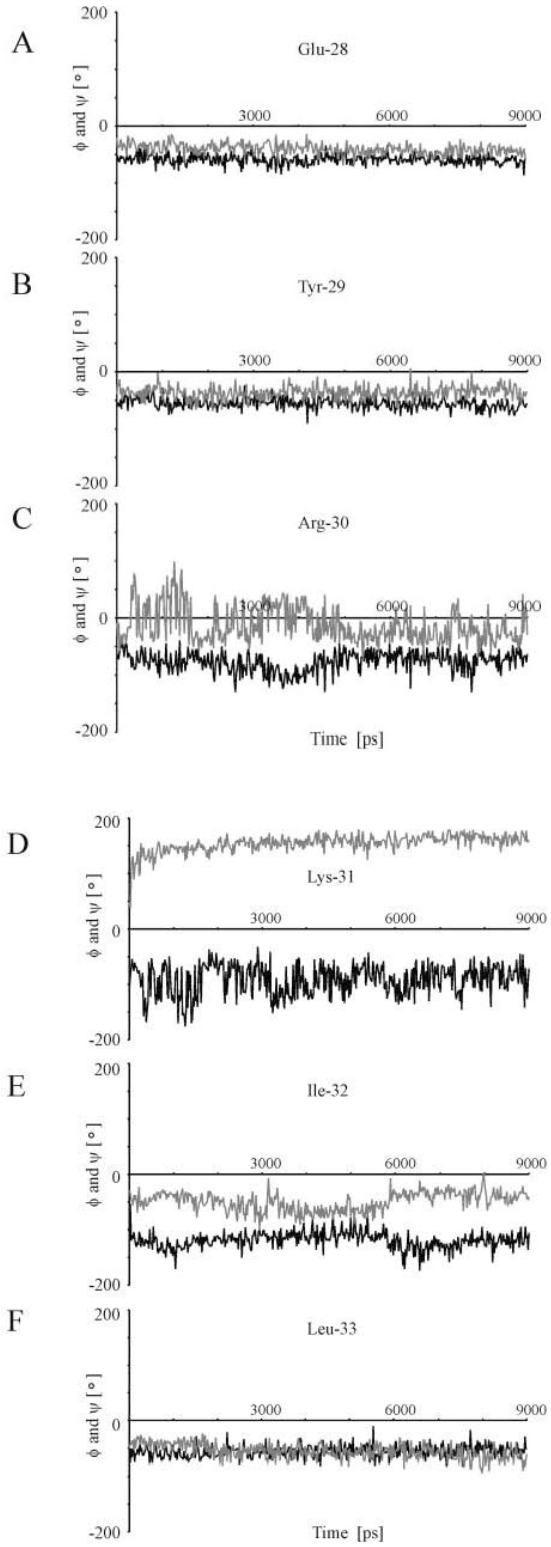


Figure 4 Time-dependent representation of the ϕ- (black traces) and ψ- values (grey traces) for the residues Glu-28 (A), Tyr-29 (B), Arg-30 (C), Lys-31 (D), Ile-32 (E), and Leu-33 (F).

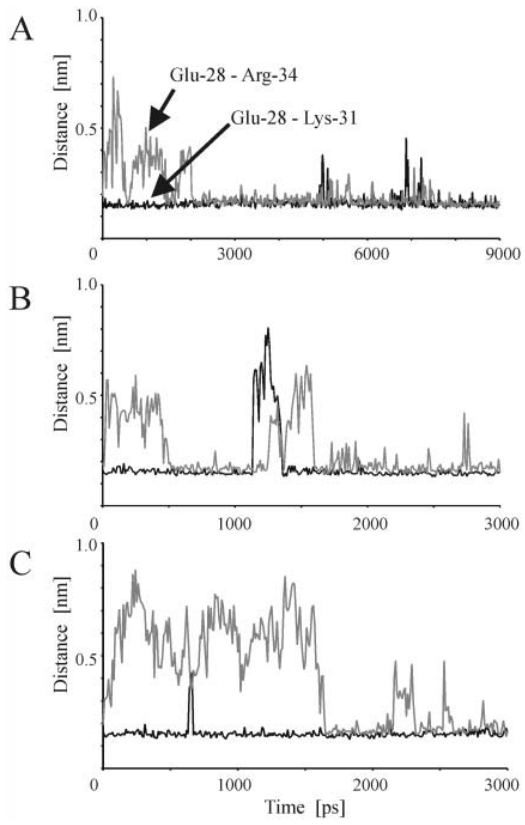


Figure 5 Time-dependent representation of the distances between Glu-28 – Lys-32 (black trace) and Glu-28 – Arg-34 (grey trace). Panel A shows the simulation over 9 ns, whereas panels B and C show the data for the two 3-ns simulations.

There is a sequence of alternating negatively and positively charged residues in Vpu: Asp-39, Arg-40, Asp-43, Arg-44, Glu-47, Arg-48, Asp-51, which forms another "complex salt bridge" (data not shown). The sequence is held together within distances of less than 0.2 nm between each of the residues. Since two hydrophobic residues separate each of the DR/ER repeats (3.6 residues per turn for an ideal helix) the hydrophilic residues involved in the salt bridge perform a slightly left handed twist. The complex salt bridge is a combination of (i,i+1)D/E,R and (i,i+3)R,D/E motives. This is different to the (i,i+4)E,K motif which has been found to promote helix conformation (35). The fact that the salt bridge is not buried (24, 27, 57) and the negatively charged residue starts at the amino terminal end (35), suggests that the motif has a stabilizing effect on the helical conformation.

After 0.5 ns of simulation, helix-1 approaches an average tilt angle of ~ (23.3 ± 2.1)º (Table 1 and black trace in Figure 6 A). Helix-2 penetrates with its C-terminal end into the lipid bilayer confining an average angle of its helix long axis with the plane of the bilayer at ~ (10.2 ± 2.0)º. This final

conformation is reached within the first 3 ns (grey trace in Figure 6 A). The resulting angle between these two helices is shown in Figure 6 B. During the simulation the angle between the two helices oscillates between 90º and 100º. Helix-1 is seen to be considerably kinked at ~ (7.2 ± 2.6)º with occasionally larger values of up to 15º. The latter value is still lower than the angle of ~ 25° found for prolin induced kinked TM helices in membrane proteins (2, 4, 14). Helix-2 adopts kink angles in the same range as helix-1 (Table 1). In Table 1, averaged values for the two repeated runs over 3 ns are listed. They show the spread ranging from 4º to 15º for both helices. The kink of helix-1 is in the same range a found for the same helix in an assembly of homooligomers forming a pore (12). It is suggested that helix-2 lies along the membrane surface is not a "ideal straight" helix.

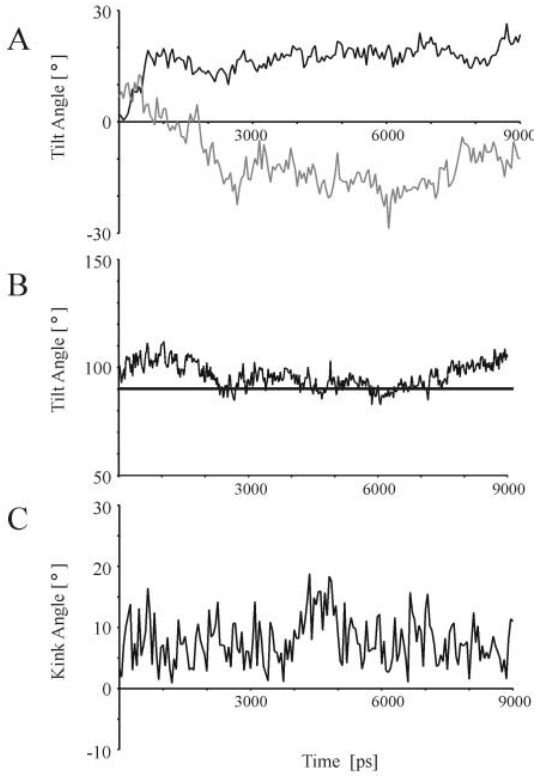


Figure 6: Evolution of the tilt angles of helix-1 (black trace) and helix-2 (grey trace) (A). Tilt angle between helix-1 and helix-2 is calculated from the cross product of the two vectors defined by the two helices: $\cos \alpha(t) = \mathbf{u}_1(t) \cdot \mathbf{u}_2(t)$; with $\mathbf{u}_1$ and $\mathbf{u}_2$ defining the vectors along the helix long axis of helix-1 and helix-2 (B). Kink angle for helix-1 (C).

DISCUSSION

Generation of the model

The peptide

The starting model has been chosen with some arbitrary constraints, firstly the choice of the residues involved in the kink connecting helix-1 and helix-2 and the adopted angles of the resulting amide bonds for these residues. This bent structure has not been seen as the unique structural motif for the kink, but will hold as a plausible first trial on the investigations of the kink. Secondly, the helix–helix bend is undertaken in a direction so that the Ser-23 is not buried under helix-2. This arrangement is covered by the fact that Asp-39 and Arg-44 has been found experimentally (25) to be aligned as shown in Fig. 3, graph on the left. Another plausible reason for the direction of the kink of helix-2 is that in a pentameric or hexameric assembly of the whole protein, serines-23 would point into the putative pore. Helix-2 consits of a hydrophilic and hydrophobic ridge. In the current model the orientation of helix-2 allows both ridges to face their favoured environment: the hydrophilic side faces the hydrophobic slab of the bilayer. Thirdly, a tilt angle found experimentally for both helices in the starting structure is not adopted. It is expected that the experimentally proposed conformation represents a local minimum and that the simulation protocol would allow finding this minimum if the starting structure would represent an energetically unfavourable conformation. Also tilt angles vary with method (CD, solid state NMR), laboratories and may be determined by lipid thickness (42).

The repeated simulations with different initial velocities result in almost identical structural features (e.g. Fig. 6). Thus, with the simulation protocol a relatively stable local minimum can be detected that allows then an interpretation of the data with more certainty. It also indicates to which extend the interpretation of events in a single simulation can be analysed. A single simulation presents trends rather than absolute figures.

Another essential factor required for a proper analysis of the data is the length of the simulation. Interpretation of the data obtained after 3 ns would result in the conclusion that the starting structure

might be essential for the outcome. Because the peptide is moving out of the bilayer (Fig. 1A), the appropriate starting structure might not have been chosen. Taking the full data set into account, the MD protocol adjusts for an unfavourable starting position. After 9 ns, independently of the starting structure used, the peptide embeds itself properly in the bilayer according to its hydrophobic shape. It cannot be ruled out that this overall translation parallel to the bilayer normal is as an oscillating movement on a longer time scale or a feature of the short time scales used (Fig. 1A).

Although different starting velocities and one long simulation (9 ns) have been chosen, the outcome might not cover all conformational space involved in a stable local minimum characteristic for the bent. Future work also needs to address different geometrical starting structures with respect to the insertion of the peptide into the lipid bilayer and the kink, especially because MD simulations do not detect large conformational changes.

The lipid bilayer

A lipid bilayer with an uneven number of lipids has been used recently to characterize the properties of the KcsA ion channel (48). An uneven number of lipids on both sides of the bilayer, reflects a frustrated system. In the current model the ”vacuum” underneath helix-2 of Vpu causes the lipid tails to fill this space, which means there is a frustration in this section of the bilayer. In a dynamic system like a lipid bilayer it is also possible to a certain extend that the lipid tails of the opposing leaflet participate in filling-up the space. Leaving the number of lipids equal on both sides would cause strong frustration on the level of the phospholipids headgroup region on the side of the leaflet containing helix-2. In both cases the frustrated system will possibly bias the peptide structure according to the different lateral pressure implemented (63). In this study the first type of frustration has been chosen because such a system would allow the amphiphatic helix-2 to settle at the lipid headgroup/lipid tail interface.

Calculation of the number of lipids per area ($Å^2$) reveals similar values for both leaflets over the entire range of the simulation indicative of identical conditions on both sides within the time frame of the simulations.

A pk_A calculation of the starting structure at 0 ns reveals all titrable residues to be in their default ionization (charged) state (Sameer, Lemaitre, Jakobsson, and Fischer, data not shown) except for Glu-51, which should be partially charged. Calculations on the structure after 0.5 ns indicate that Glu-51 has reached a position stabilizing its ionized state. It needs to be stated that the ionization state might change throughout a complete simulation and would need a more frequent assessment in a simulation.

Comparison of experimental and computational data on the shape of Vpu and with other viral membrane proteins

To a certain extent, the data derived from the simulations here match with those from NMR experiments. The average tilt angle for the TM helix obtained in this study is (23.3 ± 2.1)º. This is within the range of data obtained in other simulation studies on helix-1 assembled parallel in homo-oligomeric bundles (~15º (12)) and also of those from NMR spectroscopy on truncated / full length Vpu (33, 62). Consequently, the tilt angle seems to be independent of the rest of the protein and an intrinsic property of Vpu. Helix-2 does not hold its starting position. It rather undergoes an almost 90º revolution within the first 3 ns and remains in this position. It might speculated at this stage whether this discrepancy with experimental data (25) is due to a long lasting oscillating movement of helix 2 in respect to the time frame of the simulation and to short to be detected by NMR spectroscopy, which is anyway carried out in gel phase bilayer at low hydration (25, 32). A reason for this could also be that the experimental data were taken with a truncated cytoplasmic peptide rather than with a peptide analogous to the one used in this study. According to the present data the TM helix is stable up to Glu-29 / Tyr-30. This matches the most recent results on an extended TM segment of Vpu reconstituted into micelles using H/D exchange studies in combination with solution NMR spectroscopy ($Vpu_{2–37}$ in (31)), and is within the minimum range of a stable TM helix up to Val-25 proposed by MD simulations (20). It is expected that the extend of helix-2 up to Ser-52 is due to the missing phosphorous group at this residue. According to results from NMR spectroscopy this helix should be slightly shorter (31).

The shape of $Vpu_{1–52}$ with a TM helix and a second helix lying parallel to the lipid bilayer plane, is similar to that found for the viral coat proteins Pf1 (1, 40, 47), M13 (22, 26, 55, 60) of filamentous bacteriophages, and most recently for phospholamban (36). Molecular dynamics simulations of Pf1 indicate that the helix residing on the membrane surface is rather strongly pointing with its N-terminal end into the bilayer (43) compared to a more moderate tilt of ~ 10º for helix-2 in Vpu_{1-52} pointing with the corresponding C-terminus towards the lipid bilayer. In contrast to the studies on Pf1, the C-terminal end of helix-2 in $Vpu_{1–52}$ is not as mobile as the corresponding N-terminal end

in Pf1. Regarding the tilt angle of the TM helix, for Pf1 an angle of 30º is suggested in a Monte Carlo simulation upon the insertion process of Pf1 into the membrane (37) applying a mean-field potential to approximate the effect of the membrane. The outcome of the simulations on $Vpu_{1–52}$ represents a common motif of helix alignment in a monotopic membrane protein with one helix traversing the lipid bilayer and the other helix lying parallel to the membrane surface (53). Further simulations on Vpu with phosphorylated serines (Ser-52 and Ser-56) will be needed to address the effect of the phosphates on the alignment of helix-1 and the rest of the cytoplasmic part of Vpu in the simulations.

The data support the idea that the two helices are interconnected via a flexible linker, allowing the two helices to some extent to move independent from each other (see also (31)). It might even suggested that helix-1 has enough flexibility to form a putative channel via oligomerisation with other proteins, and helix-2 fulfils its role as a "protein float" riding on the bilayer separating helix-1 on the N-terminal side from the third helix (helix-3) and the rest of the protein towards the C-terminal side. The C-terminal side would then be able to dock to CD4, involving probably large mechanical movements, but without affecting helix-1. The results further implement the existence of "functional" modules present in Vpu. The idea of "functional" modules is an extension of the findings about proteins made up from modular repeated "building" blocks (see (8, 16)). No larger repeat units of this type are found and would be necessary for constructing the whole protein (building modules). The protein is simply too short, but each structural unit has its essential functional role (functional modules).

The EYR-motif

From the simulations it can be concluded that the first residues involved in the EYR motif retain a reasonable helical motif. According to the conformation, (i) Glu-28 is further able to support the loop via the formation of salt bridges with Lys-31 and Arg-34; (ii) Tyr-29 can anchor the protein within the lipid bilayer (13, 54), and (iii) Arg-30 is free to point into a pore and function as a putative selectivity filter. Residues involved in the connection of the two helices by forming a loop are the following residues Lys-31 to Arg-34. Salt bridge formation of Lys-31 with Glu-38 and Arg-

34 stabilise the loop. A similar role of causing a kink in connection with a loop is attributed to a close location of a Lys-20 to Asp-18 in the viral coat protein Pf1 (52).

Modeled structures demand the validation of the data and to which extent information can be deduced from them. With a single MD simulation, structural features can be resolved, however, even with repeated simulations the data have to be ”smoothed” by the interpreter. The results regarding the salt bridges (Fig. 5) would indicate from Fig. 5A, that the salt bridge Glu-28 – Lys-31 is constantly formed over the first 3ns. Fig. 5,B and C, show that, also within the first 3ns, some rupture might occur. Compared to the results from salt bridge Glu-28 – Arg-34, and in light of the repeated simulations, the statement should be such that the formation of the salt bridge settles faster for Glu-28 – Lys-31 within the first 3 ns.

Future studies need to address more conformational space regarding initial starting structures to suggest structural motifs involved in protein function. Also lipid frustration and its effect on protein structure in an all-atom simulation have to be explored in more detail.

CONCLUSIONS

The study supports that flexible parts are essential to connect units with specific functional roles and that Vpu is consequently built of “functional modules”. This allows for structural features such as e.g. the tilt angle of helix-1 to be independent of the rest of the protein. Helix-2 might form a “peptide float” riding on top of the bilayer. It is also concluded, as a consequence of the simulations on the individual modules, that it is possible to yield meaningful results representative for the whole protein.

Vpu uses an internal salt bridge formation involving three residues to connect helix-1 with helix-2 and to orient the second helix on top of the bilayer. The EYR-motif seems to be an important motif involved in confining the conformation of Vpu.

Acknowledgement We thank P.Tieleman (Calgary,Canada) for providing us with an equilibrated lipid membrane and thanks to S.Grage (Oxford) for interesting discussions. The Engineering and Physical Sciences Research Council (EPSRC) is acknowledged for financial support (to A.W.). The Biotechnology and Biological Sciences Research Council (BBSRC) is acknowledged for a Professional Research Fellowship (to A.W.) and grant support.

REFERENCES

1. Azpiazu I, Gomez-Fernandez JC, Chapman D. 1993. Biophysical studies of the Pf1 coat protein in the filamentous phage, in detergent micelles, and in a membrane environment. *Biochemistry* 32: 10720
2. Barlow D, Thornton J. 1988. Helix geometry in proteins. *J. Mol. Biol.* 201: 601
3. Berendsen HJC, Griegera JR, Straatsma TP. 1987. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* 91: 6269
4. Boncheva M, Vogel H. 1997. Formation of stable polypeptide monolayers at interfaces: controlling molecular conformation and orientation. *Biophys. J.* 73: 1056
5. Bour S, Schubert U, Strebel K. 1995. The human immunodeficiency virus type 1 Vpu protein specifically binds to the cytoplasmic domain of CD4: implications for the mechanism of degradation. *J. Virol.* 69: 1510
6. Branden C, Tooze J. 1991. *Introduction to Protein Structure. Garland, New York.*
7. Brünger AT. 1992. *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*. New Haven, Ct.: Yale University Press
8. Campbell ID, Downing AK. 1998. NMR of modular proteins. *Nat Struct Biol* 5 Suppl: 496
9. Coadou G, Evrard-Todeschi N, Gharbi-Benarous J, Benarous R, Girault J-P. 2001. Conformational analysis by NMR and molecular modelling of the 41-62 hydrophilic region of HIV-1 encoded virus protein U (Vpu). Effect of the phosphorylation on sites 52 and 56. *C. R. Acad. Sci. Paris, Chemie / Chemistry* 4: 751
10. Cohen EA, Terwilliger EF, Sodroski JG, Haseltine WA. 1988. Identification of a protein encoded by the *vpu* gene of HIV-1. *Nature* 334: 532
11. Cordes FS, Kukol A, Forrest LR, Arkin IT, Sansom MSP, Fischer WB. 2001. The structure of the HIV-1 Vpu ion channel: modelling and simulation studies. *Biochimica et Biophysica Acta* 1512: 291
12. Cordes FS, Tustian A, Sansom MSP, Watts A, Fischer WB. 2002. Bundles consisting of extended transmembrane segments of Vpu from HIV-1: computer simulations and conductance measurements. *Biochemistry* 41: 7359–7365
13. de Planque MR, Boots JW, Rijkers DT, Liskamp RM, Greathouse DV, Killian JA. 2002. The effects of hydrophobic mismatch between phosphatidylcholine bilayers and transmembrane alpha-helical peptides depend on the nature of interfacially exposed aromatic and charged residues. *Biochemistry* 41: 8396
14. Deisenhofer J, Epp O, Miki K, Huber R, Michel H. 1985. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3A resolution. *Nature* 318: 618
15. Deora A, Spearman P, Ratner L. 2000. The N-terminal matrix domain of HIV-1 Gag is sufficient but not necessary for viral protein U-mediated enhancement of particle release through a membrane-targeting mechanism. *Virol.* 269: 305

16. Doolittle RF. 1995. The multiplicity of domains in proteins. *Annu. Rev. Biochem.* 64: 287
17. Ewart GD, Sutherland T, Gage PW, Cox GB. 1996. The Vpu protein of human immunodeficiency virus type 1 forms cation-selective ion channels. *J. Virol.* 70: 7108–7115
18. Faraldo-Gomez JD, Smith GR, Sansom MSP. 2002. Setting up and optimization of membrane protein simulations. *Eur. Biophys. J.* 31: 217
19. Federau T, Schubert U, Flossdorf J, Henklein P, Schomburg D, Wray V. 1996. Solution structure of the cytoplasmic domain of the human immunodeficiency virus type 1 encoded virus protein U (Vpu). *Int J Pept Protein Res* 47: 297
20. Fischer WB, Forrest LR, Smith GR, Sansom MSP. 2000. Transmembrane domains of viral ion channel proteins: a molecular dynamics simulation study. *Biopolymers* 53: 529
21. Fischer WB, Sansom MS. 2002. Viral ion channels: structure and function. *Biochim Biophys Acta* 1561: 27
22. Glaubitz C. 2000. An Introduction to MAS NMR Spectroscopy on Oriented Membrane Proteins. *Concepts in Magnetic Resonance* 12: 137
23. Grice AL, Kerr ID, Sansom MS. 1997. Ion channels formed by HIV-1 Vpu: a modelling and simulation study. *FEBS Lett* 405: 299–304
24. Hendsch ZS, Tidor B. 1994. Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Prot. Sci.* 3: 211
25. Henklein P, Kinder R, Schubert U, Bechinger B. 2000. Membrane interactions and alignment of structures within the HIV-1 Vpu cytoplasmic domain: effect of phosphorylation of serines 52 and 56. *FEBS Lett.* 482: 220
26. Henry GD, Sykes BD. 1992. Assignment of amide ^{1}H and ^{15}N NMR resonances in detergent-solubilized M13 coat protein: a model for the coat protein dimer. *Biochemistry* 31: 5284
27. Horowitz A, Serrano L, Avron B, Bycroft M, Fersht AR. 1990. Strength and co-operativity of contributions of salt bridges to protein stability. *J. Mol. Biol.* 216: 1031
28. Kerr ID, Sankararamakrishnan R, Smart OS, Sansom MSP. 1994. Parallel helix bundles and ion channels: Molecular modelling via simulated annealing and restrained molecular dynamics. *Biophys. J.* 67: 1501
29. Kukol A, Adams PD, Rice LM, T. BA, Arkin IT. 1999. Experimentally based orientational refinement of membrane protein models: a structure for the influenza A M2 H+ channel. *J. Mol. Biol.* 286: 951
30. Lopez CF, Montal M, Blasie JK, Klein ML, Moore PB. 2002. Molecular dynamics investigation of membrane-bound bundles of the channel-forming transmembrane domain of viral protein U from the human immunodeficiency virus HIV-1. *Biophys J* 83: 1259
31. Ma C, Marassi FM, Jones DH, Straus SK, Bour S, et al. 2002. Expression, purification, and activities of full-length and truncated versions of the integral membrane protein Vpu from HIV-1. *Protein Sci* 11: 546
32. Marassi FM, Ma C, Gesell JJ, Opella SJ. 2000. Three-dimensional solid-state NMR spectroscopy is essential for resolution of resonances from in-plane residues in uniformly (15)N-labeled helical membrane proteins in oriented lipid bilayers. *J Magn Reson* 144: 156
33. Marassi FM, Ma C, Gratkowski H, Straus SK, Strebel K, et al. 1999. Correlation of the structural and functional domains in the membrane protein Vpu from HIV-1. *Proc Natl Acad Sci U S A* 96: 14336
34. Marassi FM, Opella SJ, Juvvadi P, Merrifield RB. 1999. Orientation of cecropin A helices in phospholipid bilayers determined by solid-state NMR spectroscopy. *Biophys J* 77: 3152
35. Marqusee S, Baldwin RL. 1987. Helix stabilization by $Glu^{-}...Lys^{+}$ salt bridges in short peptides of *de novo* design. *Proc. Natl. Acad. Sci. USA* 84: 8898
36. Mascioni A, Karim C, Zamoon J, Thomas DD, Veglia G. 2002. Solid-state NMR and rigid body molecular dynamics to determine domain orientations of monomeric

phospholamban. *J. Am. Chem. Soc.* 124: 9392

37. Milik M, Skolnick J. 1993. Insertion of peptide chains into lipid membranes: an off-lattice Monte Carlo dynamics model. *Prot. Struc. Func. Gen.* 15: 10
38. Moore PB, Zhong Q, Husslein T, Klein ML. 1998. Simulation of the HIV-1 Vpu transmembrane domain as a pentameric bundle. *FEBS Lett* 431: 143
39. Musafia B, Buchner V, Arad D. 1995. Complex salt bridges in proteins: statistical analysis of structure and function. *J. Mol. Biol.* 254: 761
40. Nambudripad R, Stark W, Opella SJ, Makowski L. 1991. Membrane-mediated assembly of filamentous bacteriophage Pf1 coat protein. *Science* 252: 1305
41. Paul M, Mazumder S, Raja N, Jabbar MA. 1998. Mutational analysis of the human immunodeficiency virus type 1 Vpu transmembrane domain that promotes the enhanced release of virus-like particles from the plasma membrane of mammalian cells. *J. Virol.* 72: 1270
42. Ridder ANJA, van de Hoef W, Stam J, Kuhn A, de Kruijff B, Killian JA. 2002. Importance of hydrophobic matching for spontaneous insertion of a single-spanning membrane protein. *Biochemistry* 41: 4946
43. Roux B, Woolf TB. 1996. Molecular dynamics of Pf1 coat protein in a phospholipid bilayer. In *Biological Membranes.*, ed. KM Merz Jr., B Roux, pp. 555. Boston: Birkhäuser
44. Sansom MSP, Shrivastava IH, Bright JN, Tate J, Capener CE, Biggin PC. 2002. Potassium channels: structures, models, simulations. *Biochim. Biophys. Acta* 1565: 294–307
45. Schubert U, Bour S, Ferrer-Montiel AV, Montal M, Maldarelli F, Strebel K. 1996. The two biological activities of human immunodeficiency virus type 1 Vpu protein involve two separable structural domains. *J. Virol.* 70: 809
46. Schubert U, Ferrer-Montiel AV, Oblatt-Montal M, Henklein P, Strebel K, Montal M. 1996. Identification of an ion channel activity of the Vpu transmembrane domain and its involment in the regulation of virus release from HIV-1-infected cells. *FEBS Lett.* 398: 12
47. Shon K-J, Kim Y, Colnago LA, Opella SL. 1991. NMR studies of the structure and dynamics of membrane-bound bacteriophage Pf1 coat protein. *Science* 252: 1303
48. Shrivastava IH, Sansom MSP. 2000. Simulations of ion permeation through a potassium channel: molecular dynamics of KcsA in a phospholipid bilayer. *Biophys. J.* 78: 557–570
49. Strebel K, Klimkait T, Maldarelli F, Martin MA. 1989. Molecular and biochemical analysis of human deficiency virus type 1 *vpu* protein. *J. Virol.* 63: 3784
50. Strebel K, Klimkait T, Martin MA. 1988. Novel gene of HIV-1, *vpu*, and its 16-kilodalton product. *Science* 241: 1221
51. Tiganos E, Friborg J, Allain B, Daniel NG, Yao X-J, Cohen EA. 1998. Structural and functional analyis of the membrane-spanning domain of the Human Immunodeficiency Virus Type 1 Vpu protein. *Virology* 251: 96
52. Tobias DJ, Klein ML, Opella SJ. 1993. Molecular dynamics simulation of Pf1 coat protein. *Biophys. J.* 64: 670
53. Turner RJ, Weiner JH. 1993. Evaluation of transmembrane helix prediction methods using the recently defined NMR structures of the coat proteins from bacteriophages M13 and Pf1. *Biochim. Biophys. Acta* 1202: 161
54. Ulmschneider MB, Sansom MS. 2001. Aminoacid distributions in integral membrane protein structures. *Biochimica et Biophysica Acta* 1512: 1
55. van der Ven FJM, van Os JWM, Aelen JMA, Wymenga SS, Remerowski ML, et al. 1993. Assignment of ^{1}H, ^{15}N, and backbone ^{13}C resonances in detergent-solubilized M13 coat protein via multinuclear multidimensional NMR: a model for the coat protein monomer. *Biochemistry* 32: 8322
56. van Gunsteren WF, Mark AE. 1998. Validation of molecular dynamics simulation. *J. Chem. Phys.* 108: 6109
57. Waldburger CD, Schildbach JF, Sauer RT. 1995. Are buried salt bridges important for protein stability and

conformational specifity? *Struc. Biol.* 2: 122

58. Willbold D, Hoffmann S, Rösch P. 1997. Secondary structure and tertiary fold of the human immunodeficiency virus protein U (Vpu) cytoplasmatic domain in solution. *Eur. J. Biochem.* 245: 581
59. Willey RL, Maldarelli F, Martin MA, Strebel K. 1992. Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. *J. Virol.* 66: 7193
60. Wolkers WF, Haris PI, Pistorius AMA, Chapman D, Hemminga MA. 1995. FT-IR spectroscopy of the major coat protein of M13 and Pf1 in the phage and reconstituted into phospholipid systems. *Biochemistry* 34: 7825
61. Wray V, Federau T, Henklein P, Klabunde S, Kunert O, et al. 1995. Solution structure of the hydrophilic region of HIV-1 encoded virus protein U (Vpu) by CD and 1H NMR spectroscopy. *Int J Pept Protein Res* 45: 35
62. Wray V, Kinder R, Federau T, Henklein P, Bechinger B, Schubert U. 1999. Solution structure and orientation of the transmembrane anchor domain of the HIV-1-encoded virus protein U by high-resolution and solid-state NMR spectroscopy. *Biochemistry* 38: 5272
63. Zheng S, Strzalka J, Ma C, Opella SJ, Ocko BM, Blasie JK. 2001. Structural studies of the HIV-1 accessory protein Vpu in langmuir monolayers: synchrotron X-ray reflectivity. *Biophys. J.* 80: 1837–50