Expression, Purification and Characterization of the HIV-1 coreceptor CCR5 and its ligand RANTES

and

High-pressure NMR investigation of hydrogen bonds in biomolecules

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Lydia Nisius

aus Deutschland

Basel, 2010

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Stephan Grzesiek
Prof. Dr. Andreas Engel

Basel, den 23.06.2009

Prof. Dr. Eberhard Parlow
(Dekan)

Summary

The first part of this thesis describes the expression, purification, and characterization of the membrane protein CCR5 and its ligand RANTES. The second part is concerned with the effect of pressure on the hydrogen bond network in proteins and nucleic acids.

The G protein-coupled receptor CCR5 is the major HIV-1 coreceptor and thus a very important target for the development of HIV-1 entry inhibitors. The large-scale expression and purification of CCR5, leading to 1 mg of pure protein per liter cell culture after purification, is described. The receptor's functionality is shown via recognition by a conformation-dependent antibody and by binding of the endogenous ligand RANTES. The possibility to obtain large amounts of functional CCR5 overcomes a major bottleneck for the structural characterization of this membrane protein (Nisius *et al.* 2008).

The receptor and its ligand were further characterized by NMR spectroscopy. First heteronuclear solution state NMR spectra of selectively labeled CCR5 were obtained. The interaction of the receptor with the chemokine RANTES was studied by solution NMR as well as in a lipid environment using solid state NMR experiments. In these experiments, binding of RANTES to CCR5 is observed.

Many chemokines undergo dimerization in the low micro- to millimolar range, but the biological relevance of this phenomenon is under debate. The temperature and pressure dependence of the RANTES monomer – dimer equilibrium was characterized by NMR spectroscopy. This study reveals the enthalpy and entropy contributions to RANTES dimerization. In addition, a pressure-induced unfolded state was observed for the first time.

A solely monomeric mutant of RANTES was designed and shown to inhibit HIV-1 infection *in vitro*. This is the basis for further studies of the complex between RANTES and CCR5 as well as CCR5-derived peptides (Nisius *et al.*, manuscript in preparation).

Hydrogen bonds are extremely important to stabilize the structures of proteins and nucleic acids. In this thesis, methods to measure the scalar couplings across hydrogen bonds in these biomolecules are described (Dingley *et al.* 2008, Cordier *et al.* 2008). They were applied to study pressure-dependent changes in the hydrogen bond scalar couplings in ubiquitin. The most pressure-labile regions of the protein were identified. These largely correspond to regions which are destabilized at elevated temperature (Nisius and Grzesiek, manuscript in preparation).

Pressure- and temperature-dependent hydrogen bond scalar couplings have also been measured in a small RNA hairpin. The data indicate that the average pressure-induced changes of hydrogen bond length are similar in nucleic acids and proteins.

Contents

Abbreviations and Symbols

2D7	monoclonal, conformation-dependent antibody against CCR5
Å	0.1 nm
AQP0	aquaporin 0
A-state	acidic state
BMRB	BioMagResBank, www.bmrb.wisc.edu
CCR5	CC chemokine receptor 5
CD	circular dichroism
CD4	cluster of differentiation
CMC	critical micelle concentration
COSY	correlation spectroscopy
CP	cross polarization
CRINEPT	cross-correlated relaxation-enhanced polarization transfer
CXCR4	CXC chemokine receptor 4
δ	chemical shift
D	diffusion coefficient
Da	dalton, gmol^{-1}
DDM	n-dodecyl-β-D-maltoside
DHPC	1,2-diheptanoyl-sn-glycero-3-phosphocholine
DMPA	1,2-dimyristoyl-sn-glycero-3-phosphatidic acid
DTT	dithiothreitol
η	viscosity
$E.\ coli$	$Escherichia\ coli$
EDTA	ethylenediaminetetraacetic acid
EL	extracellular loop
EM	electron microscopy
ESI-MS	electrospray ionization − mass spectrometry
γ	gyromagnetic ratio
ΔG_d	dissociation free energy
GDP	guanosine diphosphate
gp120	HIV envelope glycoprotein 120
GPCR	G protein-coupled receptor
GRK	GPCR kinase
GTP	guanosine triphosphate
ΔH	enthalpy of transition
{^{1}H}-^{15}N NOE	^{15}N steady state NOE upon ^{1}H saturation

HBC	hydrogen bond scalar coupling
HIV	human immunodeficiency virus
HRP	horse radish peroxidase
HSQC	heteronuclear single quantum coherence
ID_{50}	inhibitory dose (50%)
I_n	peak intensity of species n
I_r	relative peak intensity
I_{cross}	cross peak intensity
I_{ref}	reference peak intensity
IL	intracellular loop
INEPT	insensitive nuclei enhancement by polarization transfer
ITC	isothermal titration calorimetry
J	scalar coupling
$^{2h}J_{NN}$ $(^{h2}J_{NN})$	2-bond scalar coupling via hydrogen bond $(^{15}N - ^{15}N)$
$^{3h}J_{NC'}$ $(^{h3}J_{NC'})$	3-bond scalar coupling via hydrogen bond $(^{15}N - ^{13}C')$
k	$1.381 \ast 10^{-23}$ JK^{-1}, Boltzmann constant
K_d	dissociation constant
K_u	monomer/unfolded state equilibrium constant
LPR	lipid-to-protein ratio
MAS	magic angle spinning
MBCD	methyl-β-cyclodextrin
MIP-1α	macrophage inflammatory protein 1α
MOI	multiplicity of infection
MWCO	molecular weight cut-off
N	noise
NMR	nuclear magnetic resonance
NS	number of scans
NOE	nuclear Overhauser effect
OmpX, OmpF	outer membrane protein X, F
OG	n-octyl-β-glycopyranoside
p	pressure
p24	HIV capsid protein
PAGE	polyacrylamide gel electrophoresis
PBMC	peripheral blood mononuclear cell
PDB	RCSB Protein Data Bank, www.rcsb.org/pdb
PISEMA	polarization inversion spin-exchange at the magic angle
POPC	1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine

POPE	1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine
ppm	parts per million
R	8.314 JK^{-1}mol^{-1}, gas constant
R$_1$, R$_2$	longitudinal and transverse relaxation rate
R_H	hydrodynamic radius
RANTES	regulated upon activation, normal T cell expressed and secreted
RDC	residual dipolar coupling
RNA	ribonucleic acid
ΔS	entropy of transition
SDF-1	stromal cell-derived factor 1
Sf9	*Spodoptera frugiperda*
SDS	sodium dodecylsulfate
S/N	signal-to-noise ratio
T	absolute temperature
t$_1$, t$_2$, ...	chemical shift evolution times
T$_1$, T$_2$	longitudinal and transverse relaxation time
TBS	Tris-buffered saline
TM	transmembrane
TROSY	transverse relaxation-optimized spectroscopy
ΔV	volume change

1 The G protein-coupled receptor CCR5 and its interactions with the chemokine RANTES

1.1 Introduction

1.1.1 Membrane proteins

Membranes are crucial for all cells since they enclose their content and separate them from the environment. They maintain the essential differences between the intracellular and extracellular space and surround organelles such as mitochondria and the nucleus in eucaryotic cells. As important as the separation of organelles and cells is, so important is the exchange of matter and information.

Therefore, biological membranes are not only made up of lipid bilayers but also contain a substantial amount of proteins. These membrane proteins can serve as pores or transporters, as enzymes, adhesion molecules, receptors and much more. Approximately 30% of a eucaryotic genome encode the various membrane proteins, that can be classified as either peripheral or integral (von Heijne 2007). Peripheral membrane proteins are mainly water-soluble and tethered to the lipid bilayer via a fatty acid chain or other hydrophobic anchors. In contrast, integral membrane proteins are water-insoluble and span the membrane one or more times. Their transmembrane segments can be either α-helical or made up of β-sheets forming a β-barrel (von Heijne 1996). Despite their vast abundance, membrane proteins are underrepresented in the protein data bank (PDB) with less than 1% of the structures deposited there (Lacapere *et al.* 2007).

The first low-resolution model of a membrane protein, bacteriorhodopsin in purple membrane, was derived using electron microscopy (Henderson & Unwin 1975). Nowadays, data from electron microscopy of two-dimensional crystals can yield high-resolution structures (e.g. AQP0 has been resolved to 1.9 Å resolution (Gonen *et al.* 2005)). On the other end of the scale, cryo-electron microscopy is suited to study very large complexes in a single particle approach, although at relatively low resolution (Robinson *et al.* 2007). X-ray crystallography is the most successful technique for the determination of three-dimensional high-resolution structures, not only of proteins in general but also for membrane proteins. The introduction of new crystallization methods such as lipidic cubic phases (Landau & Rosenbusch 1996) allows resolution beyond 2 Å (Lacapere *et al.* 2007).

Nuclear magnetic resonance (NMR) spectroscopy is an alternative to x-ray crystallography for atomic resolution studies of proteins that cannot be crystallized. NMR spectroscopy can be performed in solution using detergent-solubilized membrane proteins with the major drawback being size limitation. Alternatively, solid state NMR spectroscopy can be applied to study lipid-embedded membrane proteins. While x-ray structures provide a static view on membrane proteins, NMR spectroscopy can also reveal their dynamics.

1.1.2 NMR spectroscopy of membrane proteins

Membrane proteins for NMR spectroscopy have to be heterologously overexpressed, since large amounts are needed and most experiments require the introduction of NMR-active isotopes (^{2}H, ^{13}C, ^{15}N). The proteins then need to be solubilized and purified before they can be prepared for solution or solid state NMR experiments.

NMR spectroscopy in solution

For liquid state NMR spectroscopy, membrane proteins have to be prepared in a soluble form. This can be achieved in various ways, such as detergent micelles, small bicelles (Park *et al.* 2006), nanodiscs (Lyukmanova *et al.* 2008), reversed micelles, or even organic solvents (Rastogi & Girvin 1999). The major obstacles in studying membrane proteins in solution is the large size of protein/detergent or protein/lipid complexes leading to long rotational correlation times and thus extensive line broadening (Opella & Marassi 2004). A second problem can arise from the protein's dynamics, since many membrane proteins have a less compact structure in micelles than they have in their native environment. This can lead to conformational exchange-based line broadening (Sanders & Sönnichsen 2006). With the resonances broadened by these factors, sensitivity and signal overlap become severe problems and it is more difficult to measure NMR parameters for structure determination such as NOEs (nuclear Overhauser effect), *J* couplings, and RDCs (residual dipolar couplings) (Sanders & Sönnichsen 2006).

In recent years, methods have been developed to extend the size limit in solution state NMR spectroscopy. The use of transverse relaxation-optimized spectroscopy (TROSY) (Pervushin *et al.* 1997) as a building block in pulse sequences can improve the spectral quality by selecting only the slowly relaxing components. This allows measurements of complexes with a molecular weight of up to 100 kDa. Combination of TROSY with cross-correlated relaxation-enhanced polarization transfer (CRINEPT) can push the limits of NMR spectroscopy in solution even further (Riek *et al.* 1999).

Another possibility to overcome the fast transverse relaxation rates of membrane proteins

in large complexes is their perdeuteration, thus reducing the dipole-dipole relaxation. The benefits of protein perdeuteration are also used in the ILV-labeling approach (Gardner & Kay 1998). Here, the methyl groups of isoleucine, leucine, and valine are protonated in otherwise deuterated proteins. This allows the measurement of methyl-methyl NOEs and was used in the structure refinement of OmpX in DHPC micelles (Fernández *et al.* 2004).

Problems arising from signal overlap especially in α-helical membrane proteins can be circumvented by segmental (Yagi *et al.* 2004) or selective labeling strategies. The latter has been applied to study rhodopsin with solution state NMR spectroscopy (Klein-Seetharaman *et al.* 1999, Werner *et al.* 2008).

The combined use of all these approaches has enabled the elucidation of nearly 20 structures of membrane proteins and even more fragments thereof by NMR spectroscopy in solution to date[1].

Solid state NMR spectroscopy

Solid state NMR spectroscopy allows the investigation of membrane proteins in a native-like environment. In this technique, the spectral resolution and the sensitivity are limited by the size- and orientation-dependence of the nuclear spin interactions, such as chemical shifts and dipolar couplings (Luca *et al.* 2005). Thus spectra of non-oriented, static samples suffer from extremely broad lines. This problem can be overcome by either using oriented samples or magic angle spinning (MAS) (Andrew *et al.* 1958).

Under MAS-conditions, a variety of spin-spin interactions that are directly related to internuclear distances can be measured (Luca *et al.* 2005), thus providing distance restraints for structure calculation. With respect to membrane proteins, bacteriorhodopsin and rhodopsin have been extensively studied using MAS-based methods (Herzfeld & Lansing 2002). Even if a membrane protein is not available in isotope-labeled form, solid state NMR spectroscopy offers the possibility to study binding of a labeled ligand. This was shown for neurotensin (Luca *et al.* 2003) and bradykinin (Lopez *et al.* 2008) binding to their respective G protein-coupled receptors.

For solid state NMR spectroscopy in oriented membranes, the proteins can either be embedded in bilayers on a glass support or in magnetically aligned bicelles. Using polarization inversion spin-exchange at the magic angle (PISEMA) high-resolution separated local field spectra of aligned proteins can be obtained. These spectra can be used to determine the tilt of helices with respect to the bilayer and for resonance assignment of the membrane protein (Opella & Marassi 2004). Several structures of membrane

---

[1](www.drorlist.com/nmr/MPNMR.html)

spanning peptides and proteins have been solved using oriented samples (Opella *et al.* 1999, Wang *et al.* 2001, Park *et al.* 2003).

1.1.3 G protein-coupled receptors

G protein-coupled receptors form the largest group of cell surface receptors involved in signal transduction, accounting for more than 2% of the proteins encoded by the human genome (Takeda *et al.* 2002). The large abundance of these membrane proteins is also refelcted by the wide range of ligands, which include neurotransmitters, hormones, small proteins and peptides, and odorants. Even photons can trigger siganling via GPCRs. Thus it is not surprising that GPCRs are involved in many physiological processes as well as numerous diseases. In fact, more than 30% of the therapeutics currently in the market target these receptors (Wise *et al.* 2002).
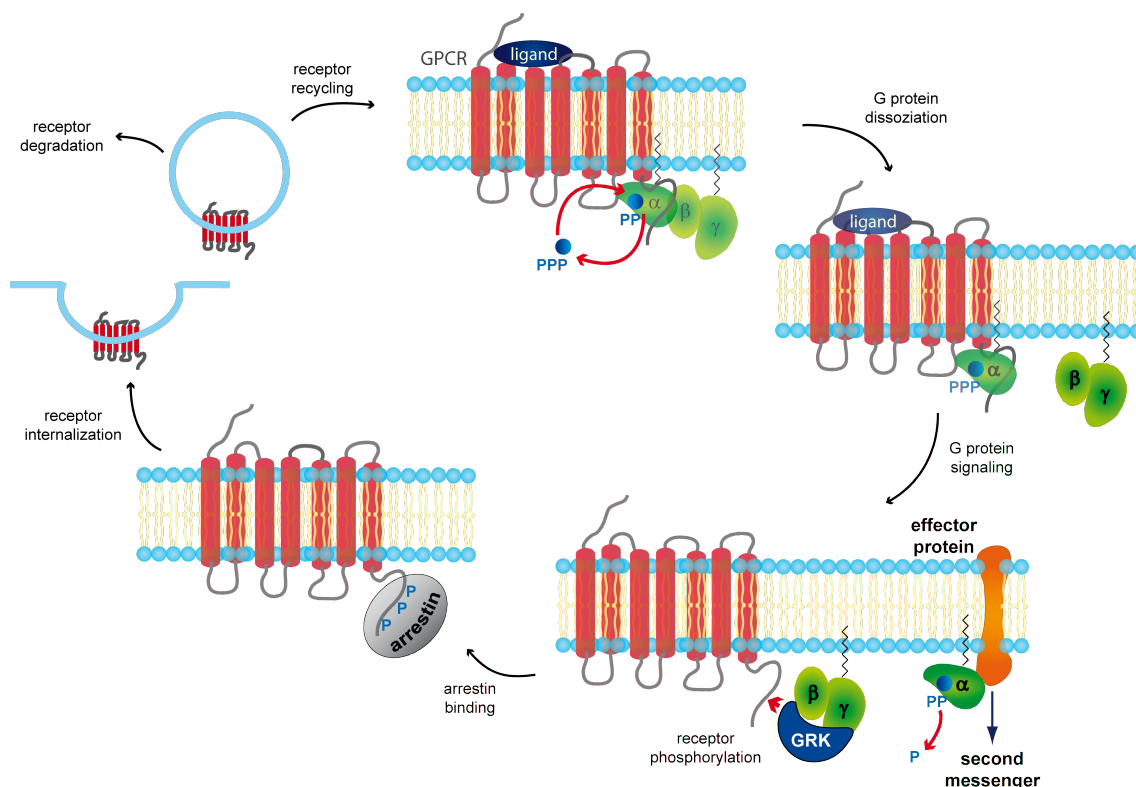
Figure 1.1 Activation, signaling, and down-regulation of G protein-coupled receptors.

All GPCRs contain the common structural feature of seven transmembrane helices connected by loops of varying length. GPCRs are situated in the cell membrane with the N-terminus at the extracellular side and the C-terminus in the cytosol. With this architecture, they are well suited to propagate signals from the extracellular side across the membrane to the interior. Ligand binding at the outside induces structural changes

in the receptors cytosolic parts, where they are coupled to heterotrimeric G proteins. This interaction gives rise to the name of the whole receptor family.

The G protein consists of a GDP/GTP binding α subunit and a β/γ subunit. Receptor activation triggers the exchange of GDP for GTP in the α subunit and its subsequent dissociation from Gβ/γ. Both, Gα and Gβ/γ then signal through activation or deactivation of downstream effectors that produce second messengers. After hydrolysis of GTP to GDP the heterotrimer reassociates and the activation cycle is terminated (De Vries *et al.* 2000). The downstream events that can occur after GPCR activation are as diverse as the processes the receptors are involved in, ranging from regulation of ion channels to the activation or suppression of gene expression. Which of these cascades is activated depends on the nature of the G protein α, β, and γ subunits, that belong to gene families with 16, 5, and 12 known members, respectively (Pierce *et al.* 2002). As a negative feedback, the receptors can also be down-regulated. One such mechanism involves phosphorylation of the GPCR's C-terminus and subsequent binding of β-arrestin. Since β-arrestin also interacts with clathrin, the receptor can be internalized in a clathrin-coated vesicle and thus is not available for signaling until it is recycled to the cell surface (Goodman *et al.* 1996). In recent years, also other, G protein-independent signaling mechanism have been found (McDonald *et al.* 2000).

GPCRs can be classified into three groups according to specific motifs, relationship to a reference receptor, etc., with class A being the largest (Mustafi & Palczewski 2009). This class contains rhodopsin-like proteins such as the biogenic amine receptors or the olfactoric receptors. From a structural point of view, GPCRs of class A are best characterized. All high resolution GPCR structures elucidated so far belong to this class. The first of those was the structure of bovine rhodopsin determined in the year 2000 (Palczewski *et al.* 2000) and up to now it remains the only one crystallized in its native state (Lodowski & Palczewski 2009). All other GPCRs with known structure required substantial mutations to form high quality crystals. These receptors are the human β_2-adrenergic and A$_{2A}$ adenosine receptors as well as the turkey β_1-adrenergic receptor (Cherezov *et al.* 2007, Rasmussen *et al.* 2007, Warne *et al.* 2008, Jaakola *et al.* 2008). Structurally, these receptors show a high degree of overall similarity with very distinct differences in the ligand binding site. The fact that even small differences in the helix arrangement can have a large effect on the ligand binding interface makes it very difficult to model the structure of other GPCRs and predict the topology of the binding sites (Schlyer & Horuk 2006).

1.1.4 The CC chemokine receptor 5

The chemokine receptors are class A GPCRs that occur on the surface of leukocytes. Their endogenous ligands are chemotactic cytokines (chemokines), which are involved in leukocyte activation and migration. According to conserved cystein residues in their N-terminus, chemokines can be divided into four classes, α (CXC), β (CC), γ (C), and δ (CX$_3$C) (Sodhi *et al.* 2004). Chemokines can also be classified by function into homeostatic and inflammatory subgroups. Most chemokines can bind to and activate more than one type of receptor and conversely, one chemokine receptor usually has more than one ligand.
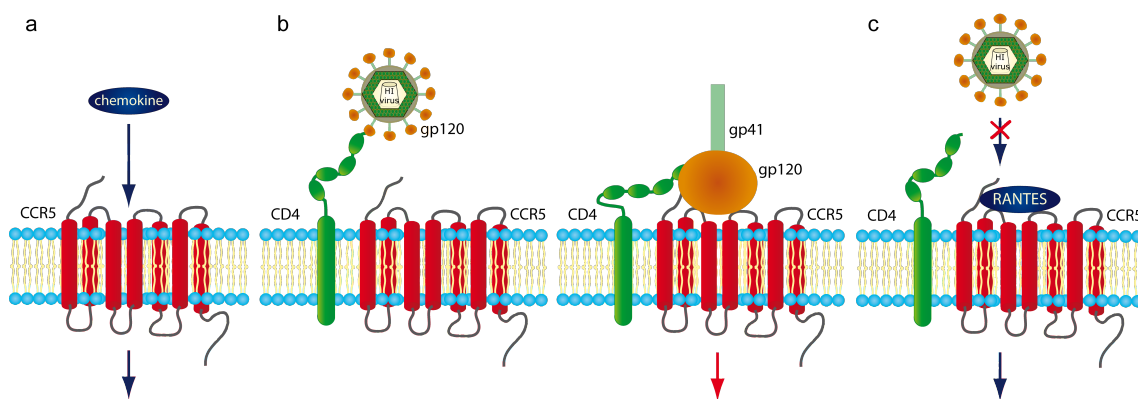

Figure 1.2   CCR5's role in physiology and disease. **a** Chemokine binding to CCR5 activates G protein-linked signaling pathways, leading for example to activation of macrophages. **b** The HI virus uses CCR5 as a coreceptor for entry into the host cell. This process requires the ternary complex of the viral envelope glycoprotein gp120, CD4 and CCR5. **c** Binding of the endogenous ligands such as RANTES inhibits the gp120/CCR5 interaction.

Some chemokine receptors can even bind non-chemokine ligands, for example the receptors CCR5 and CXCR4 can interact with the HIV-encoded glycoprotein gp120. In fact, these two chemokine receptors function as HIV-1 coreceptors and are necessary for the viral entry into host cells. Different expression of these two receptors forms the basis of HIV-1 – tropism, dividing virus subtypes into R5- or macrophage-tropic and X4- or T-cell-tropic strains. Normally, in initial stages of infection, CCR5 is utilized as a coreceptor whereas CXCR4 is associated with progression of the disease (Oberlin *et al.* 1996). The fundamental role of CCR5 in HIV infections is emphasized by a truncated form of the receptor, Δ32-CCR5, that leads to slower progression (heterozygous case) or even to resistance against infection (homozygous case) (Liu *et al.* 1996).

Prior to infection, the viral and host cell membranes have to come in contact to eventually fuse. This is possible by the initial interaction of the viral envelope glycoprotein

gp120 with CD4 on the host cell that leads to a conformational change in the glycoprotein. Thereby the CCR5 interaction side becomes accessible and gp120 can also bind to its coreceptor. This induces another structural rearrangement that also affects the glycoprotein gp41, which is associated with gp120 and connects it to the viral membrane. After conformational change, gp41 can be inserted into the host cell membrane. In turn, gp41 forms a 6-helix bundle, bringing the two membranes so close together to allow their fusion (Lederman *et al.* 2006).

The β chemokines MIP-1α, MIP-1β, and RANTES are able to prevent this process when bound to CCR5. Thus, a number of chemically modified analogues, especially of RANTES, have been developed and shown to be very effective in inhibiting HIV infections, e.g. PSC-RANTES (Lederman *et al.* 2004). A drawback of such molecules is that they cannot be cheaply produced in large amounts. Therefore, also fully recombinant chemokines have been engineered with the aim to separate the anti-HIV activity from the the ability to signal via CCR5 (Gaertner *et al.* 2008).

In addition to these variants of CCR5-natural ligands also a number of small molecule anti-HIV CCR5 antagonists has been developed. The first of those was TAK-779 (Baba *et al.* 1999), which is used only in research. The only compound approved for clinical use so far is maraviroc, showing a high anti-viral potency and good pharmacological properties (Dorr *et al.* 2005, Wood & Armour 2005). Maraviroc is predicted to form a strong salt bridge with Glu283 in CCR5. Further interactions are assumed with the residues Trp86, Tyr108, Tyr251, and Ile198. These predictions are based on a homology model of CCR5 using the known structure of bovine rhodopsin (Kondru *et al.* 2008). However, the antagonist binding pocket of the modeled structure was too small for any known ligand and therefore could only be refined using the size and shape of known small molecule ligands. This shows the difficulties in rational drug design if only homology models but no structures of a GPCR are available and emphasizes the need of more high-resolution structures as well as information on receptor dynamics.

While the antagonist binding site seems to be situated between the transmembrane helices, for the interaction with the chemokines and gp120 also the N-terminus and the extracellular loops are important. Binding of these ligands to the N-terminus requires the unusual posttranslational *O*-sulfation of tyrosines (Farzan *et al.* 1999). Although in CCR5 the tyrosine residues 3, 10, 14, and 15 could be sulfated, it was shown that modification of Y10 and Y14 is sufficient for the binding of gp120. A structure of gp120 in complex with CD4 and a tyrosine-sulfated antibody has been determined and used for docking of a NMR structure of a double-sulfated CCR5 N-terminal peptide (Huang *et al.* 2007). With respect to RANTES binding it was also shown that the interaction
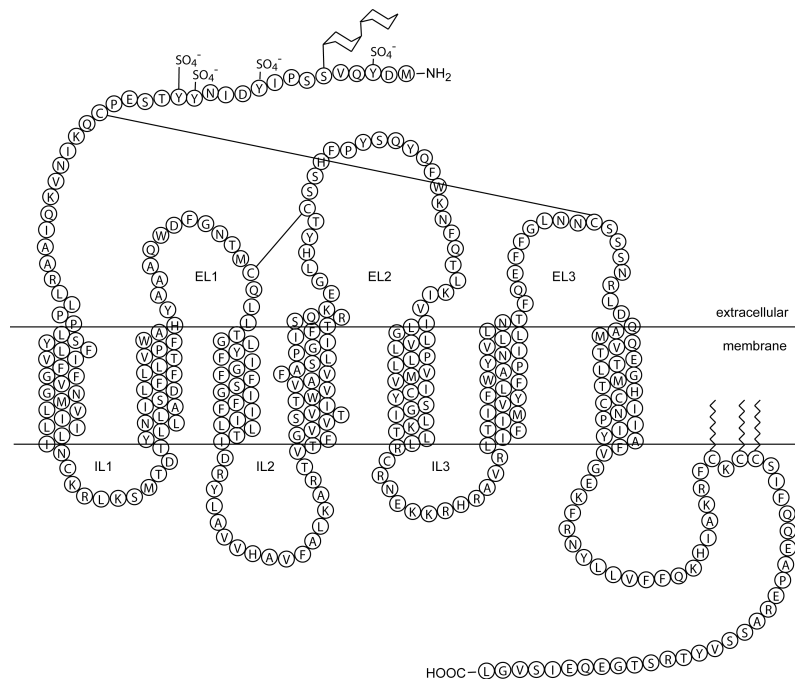
Figure 1.3   Schematic representation of CCR5's topology depicting the posttranslational modifications. CCR5 is sulfated at several tyrosines in the N-terminus, *O*-glycosylated at serine 6, and palmitoylated at three cysteins in the C-terminus. The structure is stabilized by two disulfide bonds in the extracellular part. EL1 − EL3 marks the extracellular loops, IL1 − IL3 are the intracellular loops. Helix positions are drawn according to (Oppermann 2004).

between the chemokine and a sulfated, CCR5-derived peptide is two orders of magnitude stronger compared to the unsulfated peptide (Duma *et al.* 2007). In addition to sulfation, the CCR5 N-terminus is also glycosylated at serine 6 with *O*-glycans containing negatively charged sialic acid residues (Farzan *et al.* 1999). The negatively charged surface formed by both, the sialic acid and the sulfotyrosines, is necessary for binding of the chemokines which have a positively charged surface area. A further characteristic of CCR5 are two disulfide bonds connecting the extracellular loops EL1 and EL2 as well as the N-terminus and EL3. These disulfide bridges stabilize CCR5 and are mandatory for chemokine binding, but not absolutely necessary for HIV infection (Blanpain *et al.* 1999). As many other GPCRs CCR5 is palmitoylated at cysteins 321, 323, and 324 in its C-terminus (Blanpain *et al.* 2001) which can function as an additional membrane anchor. In the receptors C-terminus there are also several serines that can be phosphorylated by G protein-coupled receptor kinases (GRK) in order to regulate cell signaling.

Organization of the chapter

In chapter 1.2, the large-scale expression and purification of CCR5 and its characterization is described. The following section (chapter 1.3) shows initial results that have been obtained using solution and solid state NMR spectroscopy to investigate the receptor and its interaction with maraviroc and the chemokine RANTES. Finally, in chapter 1.4 the design of a solely monomeric RANTES mutant is described and the temperature- and pressure-dependence of the RANTES monomer/dimer equilibrium is discussed.

1.2 Large-scale expression and purification of the major HIV-1 coreceptor CCR5 and characterization of its interaction with RANTES

Original Publication

Nisius, L., Rogowski, M., Vangelista, L., and Grzesiek, S.

Large-scale expression and purification of the major HIV-1 coreceptor CCR5 and characterization of its interaction with RANTES
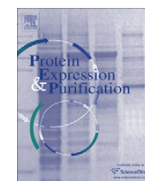
Large-scale expression and purification of the major HIV-1 coreceptor CCR5 and characterization of its interaction with RANTES

Lydia Nisius [a], Marco Rogowski [a], Luca Vangelista [b], Stephan Grzesiek [a,*]

[a] Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland
[b] Unit of Human Virology, DIBIT, San Raffaele Scientific Institute, Via Olgettina 58, Milan, Italy

ARTICLE INFO

ABSTRACT

The G protein-coupled receptor CCR5 is a human chemokine receptor involved in the activation and migration of leukocytes. CCR5 is also the major HIV-1 coreceptor that, together with human CD4 and the viral glycoprotein gp120, promotes virus entry into host cells. Thus inhibition of the CCR5-gp120 interaction presents a promising route to prevent HIV infections. Atomic structural details of the interaction between CCR5 and its cognate chemokines or gp120 are presently unknown due to the general difficulties of membrane protein structure determination. Here, we report the high-yield expression of human CCR5 in baculovirus-infected Sf9 insect cells. Highly purified (>90%) CCR5 is obtained in detergent-solubilized form at yields of about 1 mg/l cell culture. The conformational integrity of recombinant CCR5 after purification is shown by immunoprecipitation with the conformation-dependent monoclonal antibody 2D7, CD and NMR spectroscopy. The detergent micelles contain CCR5 in monomeric and dimeric forms, which can be separated by size exclusion chromatography and characterized individually. Further functional characterization by isothermal titration calorimetry indicates that the recombinant receptor interacts with its cognate chemokine RANTES. This interaction is strongly suppressed when sulfation of CCR5 is inhibited in the insect cells.

© 2008 Elsevier Inc. All rights reserved.

G protein-coupled receptors (GPCRs)[1] form the largest group of cell-surface receptors involved in signal transduction. They account for about one third of all membrane proteins in humans and share a conserved global structure comprising seven transmembrane helices [1]. In response to interaction with agonists at their extracellular surface, GPCRs provoke extremely diverse physiological reactions by modulation of the intracellular G-protein signaling cascade. Their prominent role in signaling malfunction and disease is emphasized by the fact that over 30% of the total drugs in the market target GPCRs [2]. Despite their importance, only two high-resolution structures of GPCRs have been solved so far: the crystal structures of bovine rhodopsin [3] and of the β_2-adrenergic receptor [4,5]. This shortage of data results from the well-known difficulties impeding quantitative studies of membrane proteins, i.e. low yields in expression, low stability, and the amphiphilic nature of the membrane protein's surface, which leads to problems in solubilization, purification, and reconstitution into a functional state amenable to structural and biophysical analysis.

The GPCR CCR5 (CC chemokine receptor 5, where CC indicates a conserved double cysteine motif in the cognate chemokines) is involved in the regulation of immune cell trafficking. It is activated after binding the endogenous chemokines MIP-1α (macrophage inflammatory protein 1α), MIP-1β, and RANTES (regulated on activation, normal T-cell expressed and secreted) [6]. CCR5 also plays a crucial role in human immunodeficiency virus-1 (HIV-1) infections [7,8]. The entry of HIV-1 into host cells usually requires the sequential interaction of the viral envelope glycoprotein 120 (gp120) with the host-cell factor CD4 and with either CCR5 or CXCR4 (CXC chemokine receptor 4), leading to fusion of the viral and host-cell membranes. During the early stages of HIV infection, viral isolates tend to use exclusively CCR5 for cell entry, while later in the disease, isolates that use CXCR4 emerge in part of the patients [9,10]. High-resolution structural information of CCR5 is currently lacking, but a model structure has been derived based on the rhodopsin X-ray structure and biochemical interaction data [11]. Recently, it has been shown that a peptide derived from the N-terminus of CCR5 becomes α-helical upon binding to gp120 [12].

Binding of the endogenous chemokines obstructs interaction of CCR5 with the viral envelope glycoprotein gp120 and thereby hinders infection [13]. Of these chemokines RANTES is the most potent HIV-blocking agent [14–16].

The N-terminus of CCR5 plays a crucial role in chemokine and gp120 recognition, and its post-translational sulfation at tyrosines 3, 10, 14, and 15 [17] is a prerequisite for high affinity binding to both chemokines [18] and gp120 [17]. Nuclear magnetic resonance

(NMR) chemical shift mapping has revealed the interaction surface of the CCR5 N-terminus on the RANTES structure and corroborates the requirement of sulfation for strong binding [19]. Further post-translational modifications include *O*-glycosylation at serine 6 [18] and palmitoylation at cysteines 321, 323, and 324 [20]. Whereas palmitoylation presumably adds conformational stability to the C-terminus by anchoring to the lipid membrane, the functional role of glycosylation is under debate [18].

Both chemokine receptors as well as their cognate chemokines can form stable dimers. In particular, CCR5 has been shown to form homodimers as well as heterodimers with CCR2 [21]. Similarly, RANTES, MIP-1α and MIP-1β occur in monomeric and dimeric form, depending on concentration and pH. The functional role or consequences of dimerization on signaling or HIV-1 coreceptor function are currently not understood. However, there is evidence that only the chemokine monomers interact with the receptor [19,22].

A more detailed biophysical and structural characterization of CCR5 and its interactions will be important for understanding chemokine signaling and virus entry as well as for the design of HIV entry inhibitors. Such studies require the reliable production of functional CCR5 on the tens of milligram scale. Different expression systems have been reviewed with respect to GPCR production [23,24]. A very promising route is the expression via the baculovirus/insect cell system, which allows the production of milligram quantities of membrane proteins per liter of cell culture and is easy to handle and scale up. In addition, insect cells are able to perform post-translational modifications similar to those introduced by mammalian cells. The successful expression of several GPCRs has been reported in insect cells, including among others [25] the second HIV-1 coreceptor, CXCR4 [26]. For the latter, functional expression could be achieved in *Trichoplusia ni* (Hi5) insect cells with yields of about one milligram of pure protein per 8 l of cell culture.

Here, we describe the large-scale recombinant expression of human CCR5 in baculovirus-infected *Spodoptera frugiperda* (Sf9) insect cells. Several detergents were tested for the solubilization of the receptor with the detergent FosCholine-12 yielding the best results. Highly purified protein (>90%) was obtained using a double tag strategy, combining immobilized metal affinity and *Strep*-Tactin affinity chromatography. Using this procedure routinely about one milligram of the receptor can be obtained in purified and detergent-solubilized form from one liter of cell culture. This material is recognized by several antibodies including the conformation-dependent monoclonal antibody 2D7, it shows the characteristic circular dichroism (CD) and ^{1}H NMR spectra of an α-helical protein, and interacts with RANTES at micromolar affinity. In addition, monomeric and dimeric forms of the detergent-solubilized receptor could be separated and characterized individually.

Experimental procedures

Generation of expression constructs and preparation of recombinant baculovirus

A pcDNA3.1 vector (Invitrogen) containing the cDNA encoding for human CCR5 C-terminally fused to an additional 6-histidine tag (CCR5-His) was kindly provided by Paolo Lusso (San Raffaele Scientific Institute, Milan, Italy). This construct was digested with XhoI and NsiI (New England BioLabs). Subsequently, the isolated CCR5-His DNA fragment was cloned into the pFastBac™Dual vector (Invitrogen) under the control of the P10 promoter. After transformation into *Escherichia coli*, clones with the insertion were selected by ampicillin. The corresponding plasmid was then used as a template to introduce the second C-terminal tag (*Strep*-tag II) [27] via two subsequent PCR steps with the following oligonucleotides as primers: 1. 5′-CACCATCACC ATCACCATTG GTCCCACCCT TGAATGCATA

GCATGCGG-3′ and 5′-CCGCATGCTA TGCATTCAAG GGTGGGACCA ATGGTGATGG TGATGGTG-3′; 2. 5′-CACCATTGGT CCCACCCTCA GTTCGAGAAG TGAATGCATA GCATGCGG-3′ and 5′-CCGCATGCTA TGCATTCACT TCTCGAACTG AGGGTGGGAC CAATGGTG-3′. After this step the complete C-terminal double-tag amino acid sequence was HHHHHHWSHPQFEK. The resulting double-tagged CCR5 was therefore named CCR5HS. Purified plasmid DNA with the confirmed incorporation of the 24 bases encoding the *Strep*-tag II was transformed into DH10Bac™ chemically competent *E. coli* cells (Invitrogen). Positive clones were selected according to the manufacturer's protocol and used for bacmid DNA isolation. Sf9 insect cells were transfected with this bacmid using BaculoPORTER (GeneTherapySystems) as a transfection agent. P0 viral stock was isolated after 5 days, plaque-purified and used to produce high-titer baculoviral stock as described in the manufacturer's protocol.

Sf9 cell culture

Spodoptera frugiperda (Sf9) cells were routinely maintained in suspension culture at levels between 0.3 and 4×10^6 cells/ml in serum-free Sf900II insect cell culture medium (Gibco) supplemented with antibiotics (25000 U/l penicillin G and 25 mg/l streptomycin). Erlenmeyer flasks (0.1–5 l) were filled to about 10–20% and shaken at 80 rpm and 27 °C.

Small-scale expression and analysis

For the optimization of expression conditions, several parameters were varied. Thus 10-ml cell cultures at a density of 2×10^6 cells/ml were infected with different amounts of virus at a multiplicity of infection (MOI) of 0.2–5. Furthermore, the addition of 0, 10, or 25% of fresh medium at 24 h post infection was tested. From each test expression, samples were taken every 2 h from 36 to 48 h and at 62 h post infection. Cell pellets of these samples were resolubilized in RIPA buffer (50 mM Tris, pH 7.4, 300 mM NaCl, 1% NP-40 substitute, 0.25% sodium deoxycholate, 0.1% SDS, 10 mM β-mercaptoethanol) and subjected to Western blot analysis.

Large-scale expression

For large-scale expression, cells were grown to a density of $2.0–2.5 \times 10^6$ cells/ml and inoculated with the high-titer baculovirus stock using an MOI of 1. Typically aliquots of 1–1.5 l of cell culture were kept in 5-l flasks covered with gaze. Twenty-four hours post infection, 10% of fresh medium was added to the cell culture. Cells were harvested 44 h post infection by centrifugation at 500*g* and washed with PBS once. The resulting cell pellets were stored at −70 °C. At this temperature, they remained stable for at least several months without any obvious sign of degradation such as loss of recognition of CCR5 by the conformation-specific monoclonal antibody 2D7.

Membrane preparation

All purification steps were carried out at 4 °C. Typically, the cell pellet obtained from about 1 l of cell culture was resuspended in 30 ml buffer A (50 mM NaH$_2$PO$_4$, pH 7.5, 300 mM NaCl, 10% glycerol, 1× Complete Protein Inhibitor, EDTA-free (Roche)). Cells were broken by Dounce homogenization (Wheaton, tight pestle) and centrifuged at 3000*g* for 15 min to separate homogenate and unbroken cells. The resulting pellet was resuspended in 10 ml buffer A, homogenized, and centrifuged again at 3000*g*. Supernatants were pooled and centrifuged at 30,000*g* for 1 h. The supernatant was discarded, and the pellet was resuspended in 20 ml buffer A and centrifuged at 100,000*g* for 30 min. The supernatant was again discarded, and the pellet containing the membrane fraction was

stored at −70 °C. Similar to the frozen cell pellets, these membrane preparations could be kept at this temperature for several months without any obvious signs of degradation.

Two step purification of CCR5HS

Several detergents, i.e. Cymal-5, dodecylmaltoside (DDM), dodecylphosphocholine (FosCholine-12), C12E9 (all Anatrace), Brij 58, Brij 78, and Digitonin (Sigma–Aldrich) were tested for solubilization of CCR5HS membrane fractions. Best results (see text) were obtained for FosCholine-12. Using this detergent, solubilization was carried out in the following way: typically membranes obtained from 1 l of cell culture were thawed and resuspended in 30 ml buffer B (50 mM NaH$_2$PO$_4$, pH 7.5, 300 mM NaCl, 10% glycerol, 1% FosCholine-12). The mixture was gently shaken for 2 h and centrifuged at 100,000g for 30 min. The supernatant was supplemented with imidazole to a final concentration of 15 mM and loaded onto an equilibrated Ni-column (5 ml, Chelating Sepharose fast flow (Amersham Biosciences)). The column was washed consecutively with 10 ml buffer C (50 mM NaH$_2$PO$_4$, pH 7.5, 300 mM NaCl, 10% glycerol, 1 M NaCl, 0.1% FosCholine-12), 50 ml buffer D (50 mM NaH$_2$PO$_4$, pH 7.5, 300 mM NaCl, 10% glycerol, 50 mM imidazole, 0.1% FosCholine-12), and 20 ml buffer E (50 mM NaH$_2$PO$_4$, pH 7.5, 300 mM NaCl, 10% glycerol, 80 mM imidazole, 0.1% FosCholine-12). The protein was eluted with 15 ml buffer F (50 mM NaH$_2$PO$_4$, pH 7.5, 300 mM NaCl, 10% glycerol, 500 mM imidazole, 0.1% Fos Choline-12). The eluate was mixed with the same volume of buffer G (100 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, 0.1% FosCholine-12) and loaded onto an equilibrated *Strep*-Tactin column (5 ml). The column was washed with 50 ml buffer G and the protein eluted with 15 ml buffer H (buffer G + 2.5 mM desthiobiotin). Protein concentration was determined by UV absorbance at 280 nm and by a BCA (bicinchoninic acid) assay.

Immunoprecipitation with 2D7 antibody

Approximately 1 µg purified CCR5HS was incubated with 1.25 µg antibody 2D7 (Becton Dickinson) for 1 h at 4 °C. The mixture was centrifuged for 3 min at 14,000g to separate any possible precipitate. The supernatant was added to 50 µl of 50% equilibrated protein A/protein G sepharose beads (Amersham Bioscience). The beads were shaken at 1250 rpm for 1 h at 4 °C. Subsequently, the supernatant was discarded and the beads were washed three times with 1 ml buffer G. Fifty microliters 2× SDS sample buffer (125 mM Tris, pH 6.8, 10% glycerol, 5.5% β-mercaptoethanol, 3% sodium dodecyl sulfate, 0.4% bromphenolblue) were added to the beads and shaken at 1250 rpm for 1 h at room temperature. The resulting supernatant was subjected to Western blot analysis.

Gel electrophoresis and Western blotting

Protein electrophoresis was performed on precast 4–20% gradient gels (Pierce) according to the manufacturer's protocol. Samples were mixed with 2× SDS sample buffer before loading, but not heated, since this leads to aggregation of the protein. For high-sensitivity protein detection, gels were silver-stained.

For Western blots, the protein was transferred to nitrocellulose membranes using a semidry blotting system. The transfer time was 1 h at 0.1 A for a 6 × 8 cm gel in transfer buffer (50 mM Tris, 40 mM glycine, 1% SDS, 20% methanol, pH 9.0). The membrane was blocked for 1 h in 3% BSA in TBST (Tris-buffered saline with 0.05% Tween-20). Subsequently the membrane was incubated (1 h) with 2.5 µg/ml primary antibody (either penta-His antibody (Qiagen) or polyclonal anti-human CCR5 N-terminus antibody (eBioscience)) in TBST supplemented with 1% BSA. This was followed by incubation (1 h) with 4 µg/ml HRP- (horse raddish peroxidase-) con-

jugated secondary antibody (Qiagen) in TBST supplemented with 5% non-fat dry milk. Washing was performed for 3 × 2 min after incubation with the primary and secondary antibody using TBST. Protein bands were visualized using a chemiluminescent HRP substrate (Roche) according to the manufacturer's instructions.

Size exclusion chromatography and static light scattering

Purified human CCR5HS was concentrated to 3 mg/ml using a centrifugal filter device (AmiconUltra4 10 kDa MWCO, Millipore). The concentrated protein was applied to an equilibrated Superdex200 10/300 GL size exclusion chromatography column. Buffer G was used for equilibration of the column as well as for the separation of CCR5HS monomer and dimer. Fractions were analyzed by UV and static light scattering (Minidawn; Wyatt Technology Corporation).

ITC analysis of RANTES-E66S binding

Mutant RANTES-E66S where glutamic acid 66 is replaced by serine was expressed and purified as described previously [19]. Binding of the chemokine RANTES-E66S to CCR5HS was tested by isothermal titration calorimetry (ITC) using a VP ITC instrument (Microcal). A 5.25 µM solution of the purified receptor (2 ml in buffer J (50 mM NaH$_2$PO$_4$, pH 7.5, 300 mM NaCl, 0.5% FosCholine-12)) was placed into the ITC sample cell at 25 °C. RANTES-E66S was prepared in identical buffer at a concentration of 50 µM and titrated into the sample cell in steps of 10 µl. As a control, the experiment was repeated under identical conditions using buffer J devoid of CCR5. ITC data were analyzed using the programs Origin (OriginLab) and Matlab (The MathWorks).

CD spectroscopy

CD spectra of purified CCR5HS were recorded on a Chirascan CD spectrometer (Applied Photophysics Ltd.). The protein (0.375 µM) was in phosphate buffered saline (PBS) with 0.1% Cymal-5. Measurements were performed in a 10 mm quartz cuvette at 20 °C. Four spectra were averaged and corrected for the buffer contribution.

NMR spectroscopy

Purified human CCR5HS was separated via gel filtration, and the fractions containing monomeric or dimeric CCR5HS were pooled and concentrated to a final volume of 200 µl in buffer J supplemented with 5% D$_2$O for locking. Concentrations of the NMR samples were 30 and 90 µM for the CCR5HS monomer and dimer, respectively. All experiments were performed at 25 °C on a Bruker DRX800 spectrometer, equipped with a triple-resonance, Z-gradient cryo-probe. NMR data were processed using xwinnmr 3.5 and NMRPipe [28]. One-dimensional proton NMR spectra were recorded using the 1-1 echo sequence [29]. Pulsed field-gradient (PFG) NMR spectroscopy [30] was used to determine the diffusion coefficient of the CCR5HS monomer and dimer in solution. Six 1D spectra with different gradient strengths were recorded and the relative intensities I analyzed according to $I/I_0 = \exp(-D\gamma^2 G^2\delta^2(\varDelta - \delta/3))$ where I_0 is the intensity in the absence of a gradient, D is the diffusion coefficient, G and δ are the magnitude and duration of the magnetic field-gradient pulses, respectively, $\varDelta$ is the time between field-gradient pulses, and γ is the gyromagnetic ratio of the observed nucleus. The hydrodynamic radius R_H was calculated from the diffusion coefficient via the Stokes–Einstein equation $R_H = kT/6\pi\eta D$, assuming a spherical diffusor and using the viscosity η obtained from a reference experiment on a soluble protein (ovalbumin) of similar size under equal conditions. The hydrodynamic radius was then used to calculate the apparent molecular weight m

of the detergent-solubilized protein according to $m = 4\pi R_H^3 \rho N_A / 3$, where ρ is the density and N_A Avogadro's number.

Results

Large-scale expression and purification of CCR5

The cDNA sequence corresponding to the human CCR5 gene followed at the C-terminus by the double tag sequence (6× His/*Strep*-tag II, HHHHHHWSHPQFEK) was cloned into the pFastBac™Dual expression vector under the control of the strong P10 promotor. In the following, this double-tagged protein is referred to as CCR5HS. The expression of CCR5HS in Sf9 cells was optimized in terms of the multiplicity of infection (MOI) and the expression time using 10-ml test cultures. An MOI of 1 and addition of 10% fresh medium after 24 h gave the best expression results with the optimal harvesting time being 44 h post infection. Small-scale membrane preparation, solubilization with Cymal-5 [31] and subsequent Western blotting revealed that CCR5HS occurs in monomeric as well as dimeric form, even in SDS sample buffer. This indicates a very high stability of the dimer. A further characterization of the monomeric and dimeric forms is given below. On a 4–20% gradient gel (Fig. 1), the monomer migrates as a 30 kDa band, while the theoretical molecular weight is 43 kDa. The dimer band appears at an apparent molecular weight of approximately 60 kDa. Expression of the full-length protein was confirmed by antibody binding to the CCR5HS termini (polyclonal anti-human CCR5 N-terminus antibody for the N-terminus and penta-His antibody for the C-terminus). Both detections revealed identical bands corresponding to the CCR5HS monomer and dimer (*vide infra*).

Large-scale expression was successfully performed in 5-l shaking flasks filled to about 30% as described in the Experimental procedures section. Cell pellets were stable at −70 °C for at least six months. Membranes containing CCR5HS were isolated from these cell pellets by several homogenization and centrifugation steps. The stability of the resulting membrane preparation stored at −70 °C was checked by recognition with the antibody 2D7, which binds to a conformation-dependent motif in the CCR5 second extracellular loop [32]. No differences were observed over a period of at least three months. For solubilization of CCR5HS, different detergents were tested (Cymal-5, DDM, FosCholine-12, C12E9, Brij 58, Brij 78, Digitonin). In agreement with earlier studies [33], Cymal-5 and DDM were able to solubilize and stabilize the receptor. In addition, FosCholine-12 gave similarly good results for the solubilization and stabilization, but also showed better homogeneity in an electron microscopy screen. Therefore, FosCholine-12 was used for all further experiments. Purification of CCR5HS was performed by subsequent Ni- and *Strep*-Tactin-affinity chromatography. After this two-step procedure, the protein was >90% pure as analyzed by silver staining (Fig. 1).

Antibody recognition and secondary structure

For routine analysis, the presence of CCR5HS was detected by Western blots with a monoclonal anti-His antibody (Fig. 2a, lane 1). The insertion of the second purification tag (*Strep*-tag II) was confirmed by binding of a monoclonal antibody (anti-*Strep*-tag II) directed against the eight amino acids (WSHPQFEK) of this tag (data not shown). The protein was also tested for binding of a polyclonal antibody raised against the N-terminal amino acids 6–20 (SSPIYDINYYTSEPC) of human CCR5. The recognition by this antibody (Fig. 2a, lane 2) confirmed the expression and integrity of the entire protein including its N-terminus.

The conformation of CCR5HS was probed by immunoprecipitation with 2D7. This antibody recognizes a conformation-dependent epitope on CCR5, which includes the second extracellular loop, and is frequently used to assay the structural integrity of CCR5 [33]. CCR5HS was also recognized by 2D7 after solubilization and purification, suggesting that it retained its proper conformation during this process (Fig. 2a, lane 3). The stability of the purified protein was assayed by immunoprecipitation with 2D7 after incubation for different times at 4 °C and at room temperature. CCR5HS was recognized to about 50% after two weeks incubation at 4 °C (Fig. 2b), whereas the stability at room temperature was much lower. This loss of antibody recognition also establishes that the detection of CCR5 by 2D7 is specific.

The secondary structure of the purified CCR5HS was analyzed by CD spectroscopy (Fig. 2c). Although data collection was only possible down to 200 nm due to the absorption of buffer and detergent, the CD spectrum shows typical minima at about 223 and 210 nm clearly indicating the high content of α-helical structure, which is expected for a seven transmembrane helix GPCR.

Separation of monomeric and dimeric forms of CCR5

The micelles containing the monomeric and dimeric forms of CCR5HS could be separated by size exclusion chromatography. Fig. 3a shows the UV absorbance of the gel filtration eluate (solid line) as well as the molecular masses of the protein-detergent complexes (dashed lines) determined by an in-line static light scattering detector. The light scattering signal corresponds to a molecular weight of about 75 and 130 kDa for the CCR5HS monomer and the dimer micelles, respectively. Samples of the corresponding fractions after gel filtration were analyzed by SDS–PAGE, which confirmed the separation into monomeric and dimeric forms of CCR5HS (Fig. 3b). After separation, these monomeric and dimeric forms did not interconvert even after one week storage at 4 °C.

The separated CCR5HS monomer and dimer micelles were also characterized by NMR spectroscopy. Fig. 3c shows the amide proton regions of two one-dimensional ^{1}H spectra recorded on the CCR5HS monomer and dimer micelles. Both spectra are very similar and correspond to the expected low dispersion of an α-helical protein. From the similarity of both spectra it may be concluded that there are no major structural differences between the monomeric and dimeric forms of CCR5. However, minor conformational changes cannot be excluded. In particular, small conformational changes are likely at the monomer–monomer interface and may also be possible at the intracellular and/or extracellular interface, thus leading to a differential signaling capacity of the two forms.
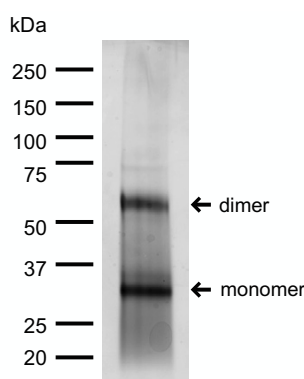

Fig. 1. SDS–PAGE of purified recombinant human CCR5HS. The purified receptor was analyzed by SDS–PAGE and subsequent silver staining. The band at about 30 kDa corresponds to the CCR5HS monomer, whereas the receptor dimer migrates at about 60 kDa. The purity of the protein preparation is estimated to be more than 90%.

The size of the monomeric and dimeric micelles was also determined by NMR diffusion measurements. For this purpose, one-dimensional ^{1}H spectra were recorded at different magnetic field-gradient strengths on FosCholine-12-solubilized monomeric and dimeric CCR5HS. Fig. 3d shows the logarithm of the relative spectral intensities plotted against the square of the relative gradient strength. The slope of the linear fits yields diffusion coefficients of 5.3×10^{-11} and 4.4×10^{-11} m^2 s^{-1} for the monomeric and dimeric micelles at 25 °C, respectively. An apparent viscosity of 1.27 cP of the detergent buffer solution (2% FosCholine-12) was obtained from a reference experiment comparing the diffusion of ovalbumin (45 kDa) in water and in the detergent buffer solution. Using this viscosity value and a density of 1.2 g/cm^3 for the micelles, molecular weights of 102 and 175 kDa are obtained for the monomeric and dimeric micelles, respectively. These molecular weights include the mass of the water bound to the micelles. Typical hydration of proteins amounts to about 30–40% (w/w) [30]. Subtracting e.g. 30% hydration water from the total micellar masses yields values of 78 and 135 kDa (protein+detergent) for the CCR5HS monomer and dimer micelles, which are in good agreement with the results from static light scattering.

Ligand binding

As a functional test, the binding of the chemokine RANTES to detergent-solubilized CCR5HS was assayed by isothermal titration calorimetry (ITC). Due to the decreased stability of CCR5 at low pH values, and in order to simulate a physiological environment, the binding studies were carried out at pH 7.5. At this pH, native RANTES has a high tendency to aggregate. Therefore, the binding was assayed by the mutant RANTES-E66S (hereafter referred to as RANTES), which has a much lower tendency to aggregate at elevated (7.5) pH values, while retaining the biological activity and CCR5 binding capacity of the native chemokine [19,34,35].

The ITC data show that the titration of RANTES (50 μM, 10-μl aliquots) into a buffer solution of FosCholine-12-solubilized CCR5HS (5 μM, ~1.4 ml) releases heat (Fig. 4a). This indicates that the association of CCR5HS and RANTES is exothermic. In contrast, the control titration of RANTES into an identical buffer/detergent solution without CCR5HS causes a small absorption of heat (Fig. 4c). This effect appears to result from the dilution of RANTES and the associated shift of the monomer–dimer equilibrium towards the monomer. Indeed, a pH-dependent dimer dissociation constant in the micromolar range has been observed in NMR studies [19]. The analysis of the ITC data is complicated by the fact that both RANTES and the receptor are in monomeric and dimeric forms. Using the simplifying assumption that all the available receptor is able to bind the ligand and that RANTES is predominantly monomeric at the used concentrations (<7 μM), the integrated ITC data can be fitted to a binding isotherm with an apparent dissociation constant of approximately 1 μM (Fig. 4b).

Post-translational modifications

In mammalian cells, the CCR5 N-terminus undergoes post-translational modifications by the addition of sulfate groups to tyrosines 3, 10, 14, and 15 [17] as well as by glycosylation of serine 6 [18]. In particular, the sulfation of CCR5 is a requirement for high affinity interactions with its ligands [17,19]. To probe whether sulfation of CCR5HS takes place in Sf9 cells, CCR5HS was expressed with the sulfation inhibitor sodium chlorate [36] added to the cell culture medium. The expression level was very similar to that obtained without the inhibitor, and the protein was also recognized by the conformation-dependent antibody 2D7 (data not shown). The purified protein was then assayed for binding of RANTES by ITC experiments under the same conditions as before. Fig. 4d shows that the released heat does not vary significantly over the course of the titration (ligand to protein ratio 0–1.8) and is considerably smaller than for normal CCR5HS (Fig. 4a). This indicates that RANTES binding is strongly inhibited when CCR5 is expressed in the presence of the sulfation inhibitor. Thus, it is concluded that CCR5HS is sulfated when expressed in normal Sf900II medium, but that sulfation is inhibited in the presence of sodium chlorate. This is in agreement with the reported sulfation of CXCR4 in Hi5 cells that also is inhibited by sodium chlorate [26].
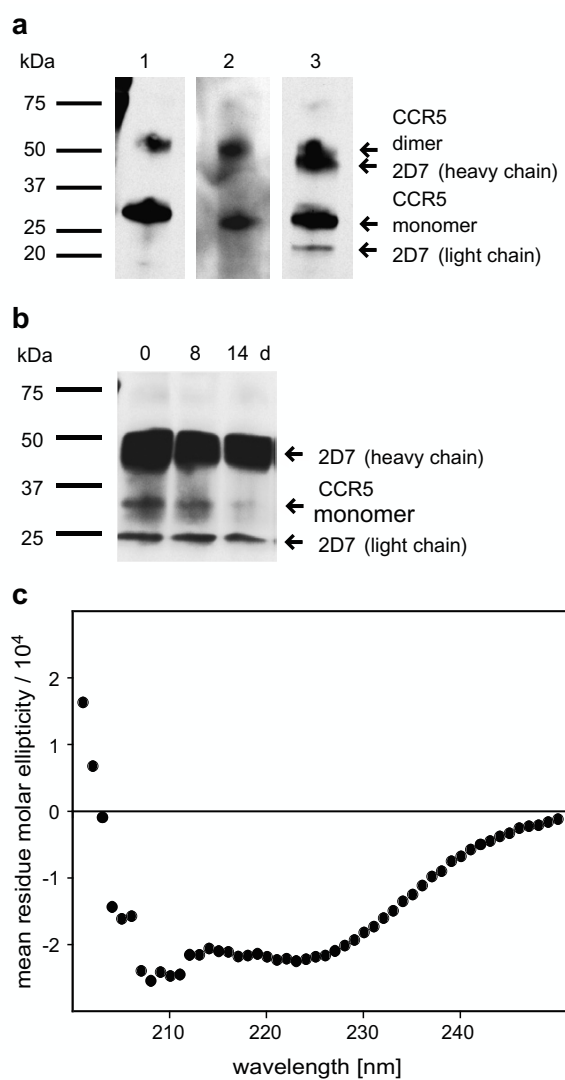

Fig. 2. Structural integrity of recombinant human CCR5HS. (a) Antibody binding and immunoprecipitation of purified CCR5HS. Binding of a C-terminal antibody (monoclonal) directed against the His-tag (lane 1) and of an N-terminal antibody (polyclonal) directed against the amino acids 6–20 (lane 2) was used to confirm the expression of the full-length protein. The structural integrity of purified CCR5HS was probed by immunoprecipitation with the conformation-dependent antibody 2D7 and subsequent Western blot analysis with the anti-His antibody (lane 3). All lanes show the monomer as well as the dimer of CCR5HS. In lane 3, the heavy and the light chain of the 2D7 antibody can be seen in addition. (b) Decrease of the amount of CCR5 recognised by the antibody 2D7 over the course of 14 days at 4 °C. (c) CD spectrum of detergent-solubilized CCR5HS. The high content of α-helical structure is evident from the characteristic minima at 223 and 210 nm.
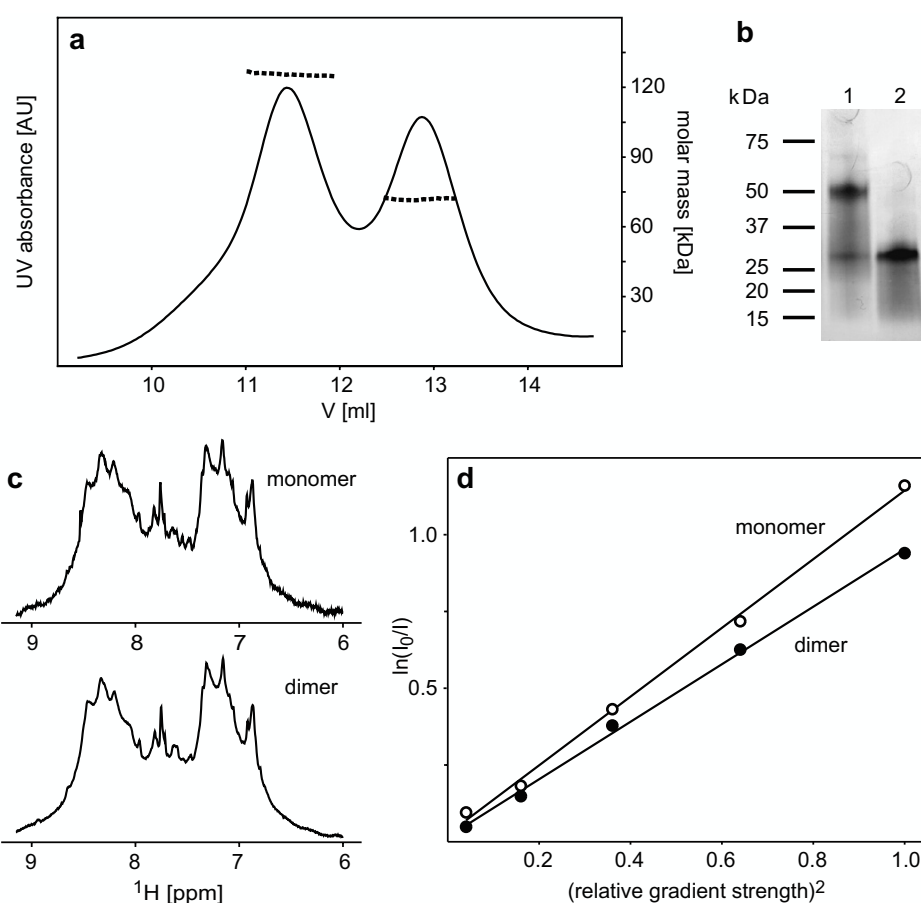
Fig. 3. Separation and characterization of CCR5HS monomer and dimer. (a) Purified CCR5HS was separated by size-exclusion chromatography in order to obtain the monomeric and the dimeric form. The eluate was analyzed by its UV absorbance (solid line) as well as by static light scattering (dashed lines). The latter correspond to molecular weights of the protein-detergent complexes of about 75 and 130 kDa for the monomer and the dimer, respectively. (b) Silver-stained SDS–PAGE of the separated CCR5HS dimer (lane 1) and monomer (lane 2) fractions. (c) 1D ^{1}H NMR spectra of the separated CCR5HS monomer and dimer fractions. Both, the monomer and the dimer show the typical pattern of an α-helical protein. Spectra are scaled to the same size. (d) Results of ^{1}H NMR experiments to determine the diffusion coefficient of the CCR5HS detergent micelles in the separated monomer and dimer fractions. The straight-line fits correspond to a diffusion coefficient of $5.3 \times 10^{-11}\,\mathrm{m^2\,s^{-1}}$ for the monomer (open circles) and $4.4 \times 10^{-11}\,\mathrm{m^2\,s^{-1}}$ for the dimer (filled circles).

Discussion

As the primary HIV-1 coreceptor, CCR5 is of major importance for the development of HIV entry inhibitors and microbicides. Despite this fact, only very limited structural information is available about CCR5 and its interactions with ligands. One of the major obstacles in such studies is the general lack of efficient procedures for high-yield functional expression and purification for membrane proteins and GPCRs in particular. In recent years, it has been shown that baculovirus-infected insect cells are capable of producing relatively large amounts of functional receptors [25]. Thus, the expression of the second HIV-1 coreceptor, the CXC chemokine receptor CXCR4, has been reported with yields of about one milligram pure protein per 8 liter of Hi5 cell culture [26].

Here, we have shown that human CCR5HS can be expressed with high yields in Sf9 insect cells using recombinant baculovirus as a vector for the gene. Pure protein was obtained by a double tag strategy consisting of a large-capacity separation by His-tag/Ni affinity chromatography followed by a highly specific separation via the *Strep*-tag II/*Strep*-Tactin interaction. Elution from both these matrices can be performed using small and readily available molecules (imidazole and desthiobiotin, respectively). Thus the

scale-up of this purification procedure is very easy and convenient. The purity of the receptor after these two chromatography steps is higher than 90% as judged from silver-stained gels (Fig. 1). An overall yield of about one milligram of purified, detergent-solubilized human CCR5HS is obtained per liter of cell culture.

The solubilized receptor is recognized by the conformation-dependent antibody 2D7, which binds to a structural motif in the second extracellular loop [32]. This suggests that the recombinant CCR5HS is correctly folded even in detergent-solubilized form. The expected high content of α-helical secondary structure was confirmed by CD measurements and one-dimensional ^{1}H NMR spectroscopy.

Using isothermal titration calorimetry we have been able to show that the recombinant human CCR5HS expressed in Sf9 cells interacts with one of its ligands, the chemokine RANTES. Using the crude assumption of a monomer-to-monomer interaction, a dissociation constant of about 1 μM is obtained. This magnitude is similar to the value of 1.2 μM derived by NMR spectroscopy for binding of the N-terminal (residues 1–25, Y10,Y14-sulfated) CCR5 peptide to RANTES monomer [19]. The agreement may be fortuitous, since the latter data were obtained at pH 3.8 in contrast to the pH 7.5 of the ITC titration, considering also that the electrostatic interaction
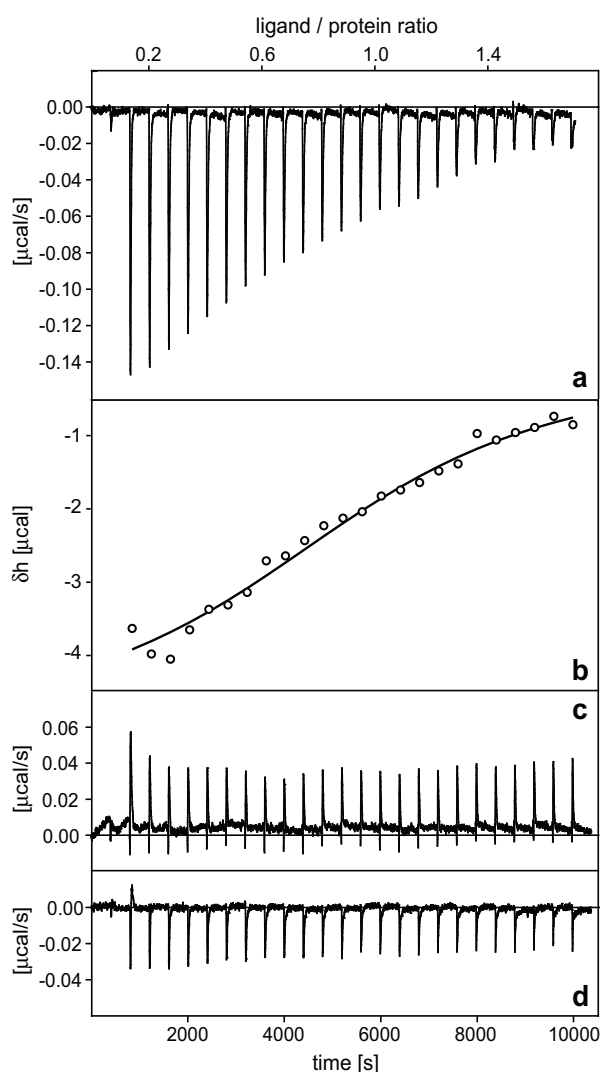
Fig. 4. Isothermal titration calorimetry of CCR5HS with RANTES. (a) Titration of 50 μM RANTES (10 μl steps) into 5.25 μM detergent-solubilized CCR5HS (1.4 ml). (b) Enthalpy per mole of CCR5HS derived from the raw data in (a) versus ligand to protein ratio. Individual titration steps were integrated using Origin, the resulting data points (open circles) were fitted assuming a one-to-one stoichiometry (solid line) by using Matlab. (c) Control titration of 50 μM RANTES (10 μl steps) into 1.4 ml buffer without CCR5. (d) Titration of CCR5HS expressed in the presence of the sulfation inhibitor sodium chlorate. Titration conditions are identical to (a).

tion with RANTES as probed by ITC is strongly reduced. This leads to the conclusion, that the receptor is sulfated when expressed in the absence of the inhibitor.

Using size exclusion chromatography, we have been able to separate micelles containing the receptor monomer and dimer. The dimeric form of CCR5HS can be observed even under SDS–PAGE conditions, suggesting that its assembly is very stable. The functional role of GPCR dimerization is controversial [38]. It has been shown that coexpression of full-length CCR5 together with the deletion mutant CCR5Δ32 inhibits the transport of the full-length receptor to the cell-surface, presumably due to receptor dimerization [39]. This process may be responsible for the delayed onset of symptoms in CCR5/CCR5Δ32 heterozygote HIV patients. However, there is currently no conclusive evidence whether CCR5 monomers or dimers or both are functionally important. The possibility to separate monomeric and dimeric detergent-solubilized CCR5 at high yields opens the way for detailed biophysical studies on this topic. Efforts in this direction are currently underway in our lab.

In conclusion, we have shown that highly purified, detergent-solubilized, functional CCR5 can be obtained in high yields from the baculovirus/insect cell system. The availability of large amounts of CCR5 should enable further structural studies as well as a thorough biophysical characterization of the interactions of CCR5 with its various ligands.

Acknowledgments

References

[1] K.L. Pierce, R.T. Premont, R.J. Lefkowitz, Seven-transmembrane receptors, Nat. Rev. Mol. Cell Biol. 3 (2002) 639–650.
[2] A. Wise, K. Gearing, S. Rees, Target validation of G-protein coupled receptors, Drug Discov. Today 7 (2002) 235–246.
[3] K. Palczewski, T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I. Le Trong, D.C. Teller, T. Okada, R.E. Stenkamp, M. Yamamoto, M. Miyano, Crystal structure of rhodopsin: a G protein-coupled receptor, Science 289 (2000) 739–745.
[4] V. Cherezov, D.M. Rosenbaum, M.A. Hanson, S.G. Rasmussen, F.S. Thian, T.S. Kobilka, H.J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka, R.C. Stevens, High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor, Science 318 (2007) 1258–1265.
[5] S.G. Rasmussen, H.J. Choi, D.M. Rosenbaum, T.S. Kobilka, F.S. Thian, P.C. Edwards, M. Burghammer, V.R. Ratnala, R. Sanishvili, R.F. Fischetti, G.F. Schertler, W.I. Weis, B.K. Kobilka, Crystal structure of the human beta2 adrenergic G-protein-coupled receptor, Nature 450 (2007) 383–387.
[6] C.J. Raport, J. Gosling, V.L. Schweickart, P.W. Gray, I.F. Charo, Molecular cloning and functional characterization of a novel human CC chemokine receptor (CCR5) for RANTES, MIP-1beta, and MIP-1alpha, J. Biol. Chem. 271 (1996) 17161–17166.
[7] G. Alkhatib, C. Combadiere, C.C. Broder, Y. Feng, P.E. Kennedy, P.M. Murphy, E.A. Berger, CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1, Science 272 (1996) 1955–1958.
[8] A. Sodhi, S. Montaner, J.S. Gutkind, Viral hijacking of G-protein-coupled-receptor signalling networks, Nat. Rev. Mol. Cell Biol. 5 (2004) 998–1012.
[9] G. Scarlatti, E. Tresoldi, A. Bjorndal, R. Fredriksson, C. Colognesi, H.K. Deng, M.S. Malnati, A. Plebani, A.G. Siccardi, D.R. Littman, E.M. Fenyo, P. Lusso, In vivo evolution of HIV-1 co-receptor usage and sensitivity to chemokine-mediated suppression, Nat. Med. 3 (1997) 1259–1265.
[10] D.R. Littman, Chemokine receptors: keys to AIDS pathogenesis?, Cell 93 (1998) 677–680.

between RANTES and CCR5 [19] is expected to vary significantly with pH. Compared to cell binding assays (e.g. $K_D = 0.4$ nM [18]), the observed binding of RANTES to the detergent-solubilized receptor is rather weak. This may be caused by physico-chemical effects such as a destabilization of the receptor structure in the detergent micelles. However, it may also be related to recent findings that intracellular binding of the G protein to CCR5 switches the receptor from low to high affinity for the chemokine [37]. Thus, the absence of the G protein in our system may significantly decrease the chemokine affinity.

The interaction with RANTES was also used to probe for the sulfation of CCR5HS. For this, the receptor was expressed in the presence of a sulfation inhibitor. Under these conditions, the CCR5HS expression level and binding to 2D7 is similar to that obtained under normal cell culture conditions. However, the CCR5HS interac-

[11] M.G. Paterlini, Structure modeling of the chemokine receptor CCR5: implications for ligand binding and selectivity, Biophys. J. 83 (2002) 3012–3031.

[12] C.C. Huang, S.N. Lam, P. Acharya, M. Tang, S.H. Xiang, S.S. Hussan, R.L. Stanfield, J. Robinson, J. Sodroski, I.A. Wilson, R. Wyatt, C.A. Bewley, P.D. Kwong, Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4, Science 317 (2007) 1930–1934.

[13] F. Cocchi, A. DeVico, A. Garzinodemo, S.K. Arya, R.C. Gallo, P. Lusso, Identification of Rantes, Mip-1-Alpha, and Mip-1-Beta as the Major Hiv-Suppressive Factors Produced by Cd8(+) T-Cells, Science 270 (1995) 1811–1815.

[14] S. Polo, V. Nardese, C. De Santis, C. Arcelloni, R. Paroni, F. Sironi, A. Verani, M. Rizzi, M. Bolognesi, P. Lusso, Enhancement of the HIV-1 inhibitory activity of RANTES by modification of the N-terminal region: dissociation from CCR5 activation, Eur. J. Immunol. 30 (2000) 3190–3198.

[15] V. Nardese, R. Longhi, S. Polo, F. Sironi, C. Arcelloni, R. Paroni, C. DeSantis, P. Sarmientos, M. Rizzi, M. Bolognesi, V. Pavone, P. Lusso, Structural determinants of CCR5 recognition and HIV-1 blockade in RANTES, Nat. Struct. Biol. 8 (2001) 611–615.

[16] O. Hartley, H. Gaertner, J. Wilken, D. Thompson, R. Fish, A. Ramos, C. Pastore, B. Dufour, F. Cerini, A. Melotti, N. Heveker, L. Picard, M. Alizon, D. Mosier, S. Kent, R. Offord, Medicinal chemistry applied to a synthetic protein: development of highly potent HIV entry inhibitors, Proc. Natl. Acad. Sci. USA 101 (2004) 16460–16465.

[17] M. Farzan, T. Mirzabekov, P. Kolchinsky, R. Wyatt, M. Cayabyab, N.P. Gerard, C. Gerard, J. Sodroski, H. Choe, Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry, Cell 96 (1999) 667–676.

[18] N. Bannert, S. Craig, M. Farzan, D. Sogah, N.V. Santo, H. Choe, J. Sodroski, Sialylated *O*-glycans and sulfated tyrosines in the NH2-terminal domain of CC chemokine receptor 5 contribute to high affinity binding of chemokines, J. Exp. Med. 194 (2001) 1661–1673.

[19] L. Duma, D. Haussinger, M. Rogowski, P. Lusso, S. Grzesiek, Recognition of RANTES by extracellular parts of the CCR5 receptor, J. Mol. Biol. 365 (2007) 1063–1075.

[20] C. Blanpain, V. Wittamer, J.M. Vanderwinden, A. Boom, B. Renneboog, B. Lee, E. Le Poul, L. El Asmar, C. Govaerts, G. Vassart, R.W. Doms, M. Parmentier, Palmitoylation of CCR5 is critical for receptor trafficking and efficient activation of intracellular signaling pathways, J. Biol. Chem. 276 (2001) 23795–23804.

[21] J.Y. Springael, E. Urizar, M. Parmentier, Dimerization of chemokine receptors and its functional consequences, Cytokine Growth Factor Rev. 16 (2005) 611–623.

[22] H. Jin, X. Shen, B.R. Baggett, X. Kong, P.J. LiWang, The human CC chemokine MIP-1beta dimer is not competent to bind to the CCR5 receptor, J. Biol. Chem. 282 (2007) 27976–27983.

[23] C.G. Tate, R. Grisshammer, Heterologous expression of G-protein-coupled receptors, Trends Biotechnol. 14 (1996) 426–430.

[24] V. Sarramegna, F. Talmont, P. Demange, A. Milon, Heterologous expression of G-protein-coupled receptors: comparison of expression systems from the standpoint of large-scale production and purification, Cell. Mol. Life Sci. 60 (2003) 1529–1546.

[25] M. Akermoun, M. Koglin, D. Zvalova-Iooss, N. Folschweiller, S.J. Dowell, K.L. Gearing, Characterization of 16 human G protein-coupled receptors expressed in baculovirus-infected insect cells, Protein Expr. Purif. 44 (2005) 65–74.

[26] A. Dukkipati, J. Vaclavikova, D. Waghray, K.C. Garcia, In vitro reconstitution and preparative purification of complexes between the chemokine receptor CXCR4 and its ligands SDF-1alpha, gp120-CD4 and AMD3100, Protein Expr. Purif. 50 (2006) 203–214.

[27] A. Skerra, T.G. Schmidt, Applications of a peptide ligand for streptavidin: the Strep-tag, Biomol. Eng. 16 (1999) 79–86.

[28] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (1995) 277–293.

[29] V. Sklenar, A. Bax, Spin-echo water suppression for the generation of purephase two-dimensional NMR-spectra, J. Magn. Reson. 74 (1987) 469–479.

[30] A.J. Dingley, J.P. Mackay, B.E. Chapman, M.B. Morris, P.W. Kuchel, B.D. Hambly, G.F. King, Measuring protein self-association using pulsed-field-gradient NMR spectroscopy: application to myosin light chain 2, J. Biomol. NMR 6 (1995) 321–328.

[31] T. Mirzabekov, H. Kontos, M. Farzan, W. Marasco, J. Sodroski, Paramagnetic proteoliposomes containing a pure, native, and oriented seven-transmembrane segment protein, CCR5, Nat. Biotechnol. 18 (2000) 649–654.

[32] L. Wu, G. LaRosa, N. Kassam, C.J. Gordon, H. Heath, N. Ruffing, H. Chen, J. Humblias, M. Samson, M. Parmentier, J.P. Moore, C.R. Mackay, Interaction of chemokine receptor CCR5 with its ligands: multiple domains for HIV-1 gp120 binding and a single domain for chemokine binding, J. Exp. Med. 186 (1997) 1373–1381.

[33] T. Mirzabekov, N. Bannert, M. Farzan, W. Hofmann, P. Kolchinsky, L. Wu, R. Wyatt, J. Sodroski, Enhanced expression, native purification, and characterization of CCR5, a principal HIV-1 coreceptor, J. Biol. Chem. 274 (1999) 28745–28750.

[34] V. Appay, A. Brown, S. Cribbes, E. Randle, L.G. Czaplewski, Aggregation of RANTES is responsible for its inflammatory properties. Characterization of nonaggregating, noninflammatory RANTES mutants, J. Biol. Chem. 274 (1999) 27505–27512.

[35] L.G. Czaplewski, J. McKeating, C.J. Craven, L.D. Higgins, V. Appay, A. Brown, T. Dudgeon, L.A. Howard, T. Meyers, J. Owen, S.R. Palan, P. Tan, G. Wilson, N.R. Woods, C.M. Heyworth, B.I. Lord, D. Brotherton, R. Christison, S. Craig, S. Cribbes, R.M. Edwards, S.J. Evans, R. Gilbert, P. Morgan, E. Randle, N. Schofield, P.G. Varley, J. Fisher, J.P. Waltho, M.G. Hunter, Identification of amino acid residues critical for aggregation of human CC chemokines macrophage inflammatory protein (MIP)-1alpha, MIP-1beta, and RANTES. Characterization of active disaggregated chemokine variants, J. Biol. Chem. 274 (1999) 16077–16084.

[36] Y. Zhao, D.C. Sane, Expression of a recombinant baculovirus for vitronectin in insect cells: purification, characterization of post-translational modifications and functional studies of the recombinant protein, Arch. Biochem. Biophys. 304 (1993) 434–442.

[37] J.Y. Springael, P.N. Le Minh, E. Urizar, S. Costagliola, G. Vassart, M. Parmentier, Allosteric modulation of binding properties between units of chemokine receptor homo- and hetero-oligomers, Mol. Pharmacol. 69 (2006) 1652–1661.

[38] S. Terrillon, M. Bouvier, Roles of G-protein-coupled receptor dimerization, EMBO Rep. 5 (2004) 30–34.

[39] M. Benkirane, D.Y. Jin, R.F. Chun, R.A. Koup, K.T. Jeang, Mechanism of transdominant inhibition of CCR5-mediated HIV-1 infection by ccr5delta32, J. Biol. Chem. 272 (1997) 30603–30606.

1.3 NMR spectroscopic study of the CC chemokine receptor CCR5 and its interaction with maraviroc and RANTES

Abstract

The β chemokine receptor CCR5 is a member of the G protein-coupled receptor (GPCR) family class A. It shares the typical structural motif of these proteins with seven membrane spanning helices. CCR5 is involved in both, crucial processes of the immune system as well as infections by the human immunodeficiency virus 1 (HIV-1). Thus the receptor and its endogenous ligands are very important targets but, as GPCRs in general, difficult to study. Purified CCR5 has been used to investigate the receptor – ligand interaction by both, solution and solid state nuclear magnetic resonance (NMR) spectroscopy. Preparations of selectively ^{15}N-isoleucine and ^{15}N-tyrosine labeled CCR5, solubilized by the detergent FosCholine-12, allowed observation of this membrane protein in two dimensional ^{1}H/^{15}N HSQC spectra for the first time. Unlabeled, detergent-solubilized CCR5 was used to investigate the binding of the small molecule antagonist maraviroc as well as the interaction with the chemokine RANTES. Here, initial proofs for the binding of maraviroc to the recombinant human receptor are shown. NMR spectroscopic interaction studies with RANTES proved to be more difficult. In NMR titration experiments, ligand resonances are weakened but not shifted. Solid state NMR spectroscopy was used to investigate the interaction between RANTES and CCR5 in a membrane-like environment. The successful reconstitution of CCR5 into lipids and initial RANTES-binding studies using solid state NMR spectroscopy are reported.

1.3.1 Introduction

Despite enormous technical progress, the fraction of membrane protein structures deposited in the protein data bank is still very small. Furthermore, among this small number of three dimensional structures not all classes of membrane proteins are equally well represented. While there is a number of β-barrel membrane protein and α-helical channel structures deposited, e.g. (S.Hiller *et al.* 2008, Hilf & Dutzler 2008), so far there are only few structures of G protein-coupled receptors available (Palczewski *et al.* 2000, Cherezov *et al.* 2007, Rasmussen *et al.* 2007, Warne *et al.* 2008, Jaakola *et al.* 2008). More structural information about this protein family would be very important, since GPCRs account for one third of all mammalian membrane proteins (Pierce *et al.* 2002) and many pharmaceutical compounds target these receptors. The lack of structural information is due to problems at all levels of the structure determination process, starting with low yield expression, difficult solubilization and purification as well as poor stability in detergents. Although all GPCR structures elucidated so far were solved by X-ray crystallography, this technique usually requires protein engineering and antagonist binding to reduce the inherent flexibility of these receptors (Kobilka 2007). Furthermore, extreme efforts are necessary to find mutants that are stabilized but still functional (Rosenbaum *et al.* 2007, Magnani *et al.* 2008). Therefore, at the moment X-ray crystallography is limited to receptors in one, mostly inactivated state. If insights into the mechanism of function are to be gained from the structures, additional information obtained with other techniques is necessary.

In recent years, NMR spectroscopy has gained more and more importance in the study of membrane associated and membrane integrated proteins in terms of both, structure and dynamics (Opella & Marassi 2004). With respect to GPCRs, the tryptophan backbone/side chain dynamics of the vertebrate photoreceptor rhodopsin have been studied by NMR spectroscopy (Klein-Seetharaman *et al.* 2004, Werner *et al.* 2007). Furthermore, its C-terminus was characterized by comparison of the partially labeled receptor with a 19 amino acid peptide (Werner *et al.* 2008). Progress was also made in studying the rhodopsin analogue sensory rhodopsin (pSRII) from *Natronomonas pharaonis*. This GPCR-like, seven transmembrane helix protein was investigated using both, solid state NMR (Etzkorn *et al.* 2007) as well as solution state NMR spectroscopy (Gautier *et al.* 2008). The investigation in the solid state concentrated on the proteins secondary structure, dynamics, and membrane topology. Using solution state NMR spectroscopy, Gautier *et al.* could achieve nearly complete backbone assignment, that allowed analysis of secondary structure and backbone dynamics. This study revealed the importance of the detergent choice for successful NMR investigations. The experimental time for example was drastically reduced by changing the detergent from dodecyl maltoside to

diheptanoyl phosphatidylcholine. The good data quality in this detergent even allowed structure calculation (D. Nietlispach, personal communication). These studies are encouraging and show that NMR spectroscopy, both in the liquid and solid state, is suitable to study GPCRs and similar proteins. Easier than complete three-dimensional structures, NMR spectroscopy can provide valuable information about the dynamics of those receptors already characterized by X-ray crystallography.

Here we describe initial investigations of the human CC chemokine receptor 5 (CCR5) and its interacting ligand RANTES using both, solution and solid state NMR spectroscopy. First feasibility studies of the receptor alone were done with a selectively ^{15}N-Ile labeled sample. Since antagonist binding was helpful in many crystallographic studies of GPCRs, binding of the antagonist maraviroc was tested using selectively ^{15}N-Tyr labeled CCR5. Also the isotope-labeled ligand RANTES was used and titrated with unlabeled CCR5 in order to study the ligand receptor interaction. With the same goal, the RANTES/CCR5 complex was reconstituted into lipids and used for solid state NMR spectroscopy.

1.3.2 Results and Discussion

Choice of the detergent for solution state NMR spectroscopy

The detergent which is used to keep a membrane protein in solution is essential for NMR spectroscopy. Therefore, several detergents were tested for their ability to solubilize and stabilize CCR5. The conformation of the receptor was assayed by immunoprecipitation with the conformation-dependent monoclonal antibody 2D7 (see chapter 1.2). Some detergents such as octyl glucoside are not able to keep CCR5 in a state that is recognized by this antibody and were therefore not further tested (Figure 1.4a).

Since the amount of protein required for NMR spectroscopy is rather large, initially electron microscopy was used to check the quality of preparations in different detergents (Cymal-5, DDM, FosCholine-12, C12E9, Brij 58, Brij 78, Digitonin). Out of the tested detergents FosCholine-12 showed the best homogeneity in the electron microscopy images (Figure 1.4b). Characterization of CCR5 in this detergent is described in chapter 1.2 and revealed that FosCholine-12-solubilized receptor is not only recognized by the 2D7 antibody but also interacts with its ligand RANTES. Therefore, all solution state NMR experiments described here were performed with CCR5 solubilized by FosCholine-12.
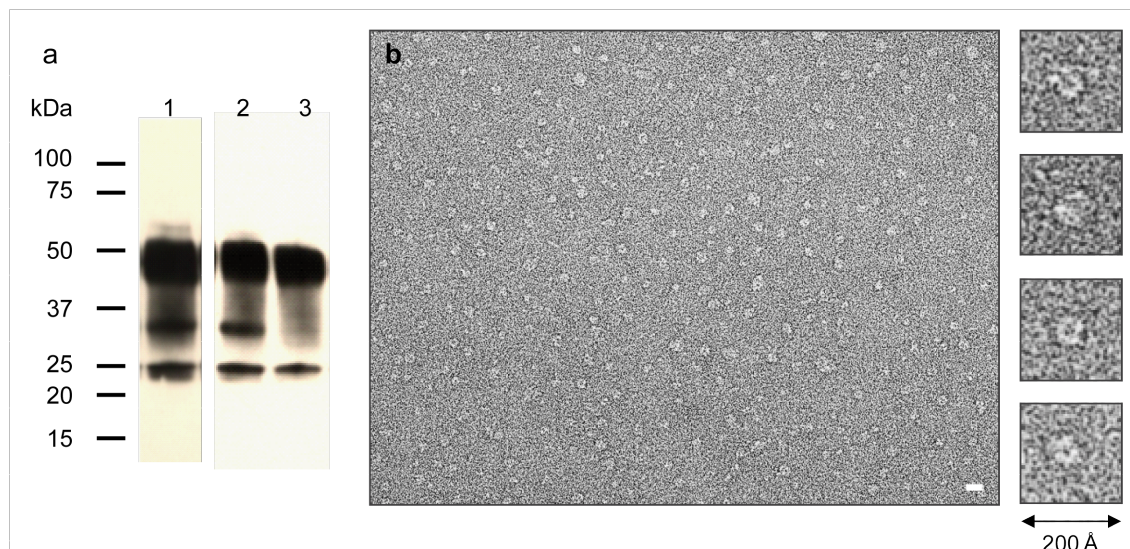
Figure 1.4 **a** Western blot after immunoprecipitation with the antibody 2D7, showing bands for the heavy (50 kDa) and light (25 kDa) chain of this antibody and the CCR5 monomer (30 kDa). In the detergents FocCholine-12 (lane 1) and Cymal-5 (lane 2) the protein is recognized by the antibody whereas octyl glucoside (lane 3) is not suitable. **b** Electron micrograph of CCR5 in FosCholine-12 (scale bar: 100 Å) with some particles enlarged on the right. Electron microscopy images for the detergent screen were taken by Petr Broz.

Amino acid-type selective labeling of CCR5 in baculovirus-infected insect cells

Large-scale recombinant expression of functional proteins is necessary but not sufficient for NMR spectroscopic investigations. For many experiments, uniform or selective labeling of the proteins with NMR active isotopes is required.

CCR5 was expressed in *Sf9* insect cells as described in chapter 1.2. For these cells, most amino acids are essential and have to be added to the cell culture medium. Thus production of isotope-labeled proteins requires a medium without amino acids and any undefined source of those (Strauss *et al.* 2003). For amino acid-type selective labeling of CCR5, the commercially available insect culture medium SF-4 Baculo Express ICM was custom made by BioConcept (Allschwil, Switzerland) without yeast extract and depleted of the amino acid that was to be isotope-labeled.

The three amino acids serine, isoleucine, and tyrosine were chosen to test the expression of selectively labeled CCR5 and for initial NMR spectroscopic studies. Serine is very abundant in the termini and loop regions while isoleucine is the major amino acid within the transmembrane helices. In addition, tyrosine was chosen because it represents a very important point of interaction with RANTES, which binds to sulfated tyrosine residues in the CCR5 N-terminus (Y3, Y10, Y14, Y15). Figure 1.5 shows the topology of CCR5

with all serine residues colored yellow, isoleucine residues marked in red and all tyrosine residues in blue. CCR5 here and in the following discussion refers to the human protein with an additional C-terminal 6His- and *Strep*-tag. These purification tags are not shown in Figure 1.5.
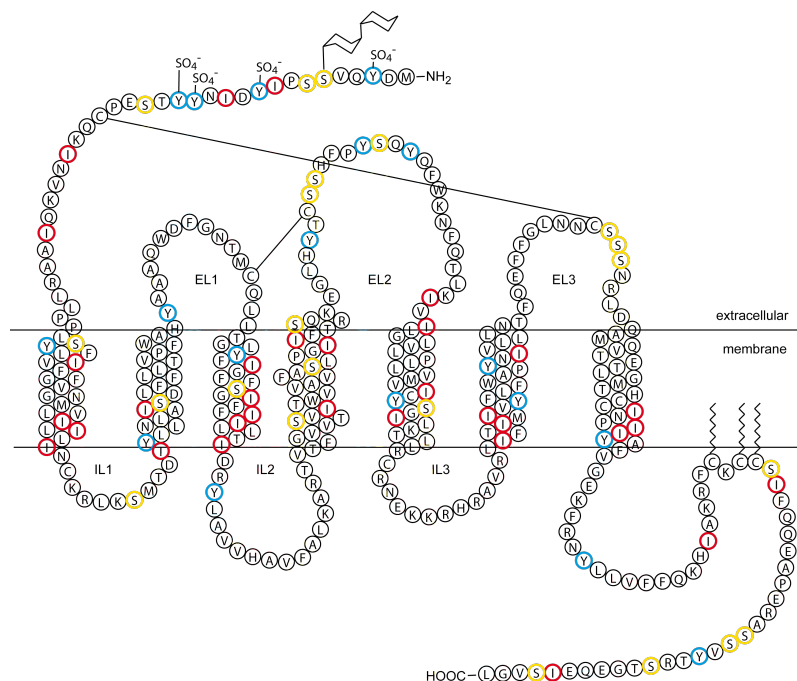

Figure 1.5 Topology of CCR5 highlighting those amino acids, that were chosen for selective labeling of the receptor (serine (yellow), isoleucine (red), and tyrosine (blue)). Helix positions are drawn according to (Oppermann 2004).

For the large-scale expression of selectively labeled CCR5, cells were pre-cultured in full medium (SF-4 Baculo Express ICM). Because it was shown before, that the cell internal pool of unlabeled amino acids is small (Strauss *et al.* 2003), the medium was changed to amino acid-depleted SF-4 medium only prior to infection as described in the Materials and Methods section. Expression was performed in 5-l shaking flasks filled with a maximum of 1 l cell culture. Cell harvest, membrane preparation, and purification was performed as described (see chapter 1.2). The expression levels of CCR5 in amino acid depleted medium are slightly lower compared to full medium.

One substantial problem during amino acid-type selective labeling is cross-labeling of other amino acids (scrambling). For *Sf9* insect cells, cross-labeling occurs for glycine (Strauss *et al.* 2003), alanine, glutamate, and aspartate (Adam *et al.* 2005). We also observed scrambling of ^{15}N-serine when incorporated into CCR5 or conversion of other amino acids into serine. Although normal signal intensities in 1D proton NMR spectra were obtained, no signals could be observed in 2D ^{1}H-^{15}N HSQC spectra. For labeling

with isoleucine and tyrosine, scrambling is not problematic, although in the case of selectively ^{15}N-tyrosine labeled CCR5 some peaks occur in the region where normally side chain NH$_2$-groups appear. These signals might arise from glutamine, that can be synthesised from glutamate by the used insect cells after degradation of ^{15}N-tyrosine (Drews *et al.* 2000).

Solution state NMR spectroscopy of selectively-labeled CCR5 in detergent micelles

Solution state NMR spectroscopy of integral membrane proteins is a challenge due to the slow tumbling of the protein/detergent complex. In GPCRs, an additional difficulty arises from the low range of signal dispersion resulting from their mainly α-helical structures. We have shown before that one dimensional ^{1}H NMR spectra of CCR5 can be obtained and used for diffusion measurements (see Chapter 1.2). From 1D spectra the amide proton T$_2$ can be estimated to $\approx$ 4.5 ms at 15 °C.
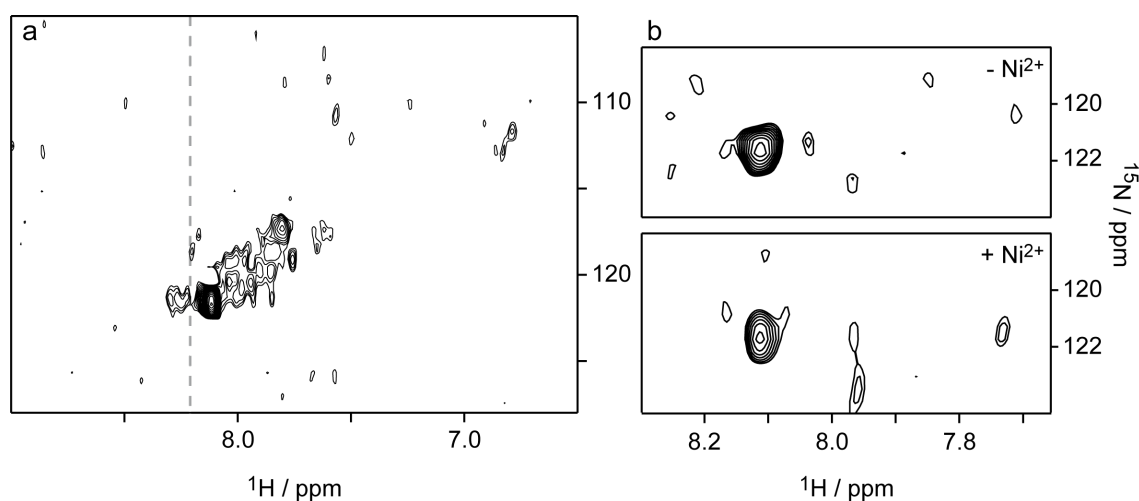

Figure 1.6   **a** Selected region of the ^{1}H-^{15}N HSQC spectrum of 100 μM ^{15}N-isoleucine-labeled CCR5 in FosCholine-12 micelles. The spectrum was measured at 40°C with a total measurement time of 85 h. Within that time not all of the 32 expected resonances can be observed. The grey dashed line represents the random coil amide proton shift for isoleucine. **b** The strongest observed peak is shown before and after the addition of 1 mM NiSO$_4$ to the sample. Due to Ni^{2+}-binding to the CCR5 His-tag, the signal is bleached out.

To assess the feasibility of further NMR spectroscopic investigations of detergent-solubilized CCR5 in the liquid state, the receptor was selectively labeled with either ^{15}N-isoleucine or ^{15}N-tyrosine. These preparations allowed observation of CCR5 in heteronuclear NMR experiments for the first time. Figure 1.6a shows the ^{1}H-^{15}N HSQC spectrum of selectively labeled ^{15}N-Ile CCR5 at 40 °C. The spectrum illustrates that several, even

though not all resonances of the labeled amino acid can be seen. Taking the strongest peak as an indicator, it is observed that the intensity decreases by 30% within 48 h, suggesting that the protein aggregates over time at the elevated temperature.

Out of the observed isoleucine signals, the very strong resonance most probably corresponds to isoleucine 348 in the very C-terminus of CCR5. This amino acid is separated from the His-tag by only three other residues. This assignment was verified by a Ni-bleaching experiment using the paramagnetic properties of Ni^{2+} which lead to very fast relaxation of nearby nuclei. The ^{15}N-Ile CCR5 sample was titrated with $NiSO_4$ and HSQC spectra were recorded before and after the addition. With increasing Ni^{2+} concentrations the peak intensity decreases, since Ni^{2+} is bound by the His-tag and bleaches out resonances of amino acids in close vicinity. Figure 1.6b shows the strongest peak of ^{15}N-Ile CCR5 before and after addition of $NiSO_4$. If the Ni^{2+} is complexed by EDTA and washed out, the peak is recovered. The remaining signals cannot be assigned to certain residues, but most likely they represent isoleucines in the more flexible termini and loop regions. Nevertheless, compared to the random coil amide proton chemical shift for isoleucine (8.19 ppm at pH 7.0 and 35 °C (Wüthrich 1986)), most observed signals are shifted upfield, which corresponds to an α-helical conformation.
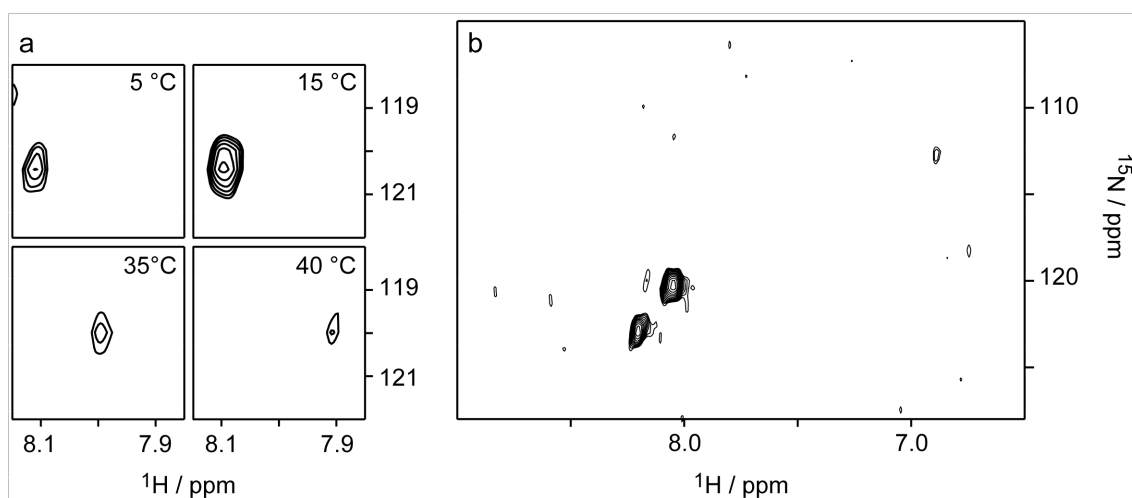

Figure 1.7   **a** Selected region of the ^{1}H-^{15}N HSQC spectrum of ^{15}N-tyrosine labeled CCR5 depicting the strongest peak. The results of four short experiments (1 h) at different temperatures (5, 15, 35, and 40 °C) are shown. The signal intensity is highest at 15 °C. **b** ^{1}H-^{15}N HSQC spectrum of 85 μM ^{15}N-tyrosine labeled CCR5 in FosCholine-12 micelles. The spectrum was acquired for 20 h at 15 °C.

In Figure 1.7a the strongest peak occurring in the ^{1}H-^{15}N HSQC spectrum of ^{15}N-Tyr CCR5 is shown at different temperatures. This peak can already be seen after short

measurement times (1 h). It is obvious that the signal intensity is extremely dependent on the temperature, with the best spectrum obtained at 15 °C. At lower temperature (5 °C) the resonance is weaker due to slower tumbling. In contrast, at higher temperatures even stronger signals would be expected. The observed decrease in signal intensity above 15 °C can be explained by motions that are in the intermediate range ($\mu s - ms$) and lead to line broadening. Thus the peaks disappear at higher temperatures. The different optimal temperature for measurements of ^{15}N-Tyr CCR5 compared to the ^{15}N-Ile labeled protein can be due to the different nature of these residues. Isoleucine is more hydrophobic than tyrosine (especially if the latter is sulfated) and its side chain is likely to stick to the protein rather than point into the solution. Therefore, higher temperatures are needed to observe these residues in the large complex.

Figure 1.7b shows a selected region of the ^{1}H-^{15}N HSQC spectrum of ^{15}N-Tyr CCR5 after a 20 h experiment at 15 °C. Two major resonances appear at proton chemical shifts between 8.0 and 8.25 ppm.

The two spectra of selectively labeled CCR5 show that heteronuclear NMR experiments with this receptor are in principle possible. Nevertheless, the spectral quality has to be significantly improved. The spectra obtained from ^{15}N-Ile and ^{15}N-Tyr CCR5 show less peaks than would be expected from the receptor's amino acid sequence (32 isoleucine and 18 tyrosine). One reason for the low number of visible peaks could be aggregation of the receptor under the conditions used. Although this is contradictory to our results from size-exclusion chromatography-coupled static light scattering and 1D ^{1}H NMR diffusion measurements, there are factors that could lead to aggregation of CCR5 during the NMR sample preparation. The selectively labeled protein was concentrated to yield ≈ 100 μM samples, a procedure which could lead to aggregation. Possibly, FosCholine-12, which was used for all NMR experiments is not able to prevent aggregation of the receptor completely, especially at high protein concentrations.

Since the use of highly-concentrated DHPC as detergent was very successful in the study of sensory rhodopsin pSRII (Gautier *et al.* 2008) this detergent was also tested for CCR5. Although the 1D proton NMR spectra do not change very much upon addition of 7% DHPC, slight improvement in 2D spectra can be observed when DHPC is added to the old (16 months) ^{15}N-Ile labeled sample (data not shown).

Binding of the antagonist maraviroc to CCR5

The ^{15}N-Tyr labeled CCR5 sample was further used to investigate the effect of the antagonist maraviroc. A five times excess of maraviroc was added from a highly concentrated H$_2$O stock solution. One dimensional ^{1}H NMR spectra were used to estimate the amide

proton T_2 for both molecules (Figure 1.8a-c). While $T_2(H^N)$ of the protein remains unchanged, the resonances of maraviroc show transverse proton relaxation times in the order of 28 ms (15 °C) after addition to CCR5, which is much less than observed for maraviroc alone under the same conditions (102 ms). This indicates that maraviroc at least partially binds to CCR5.
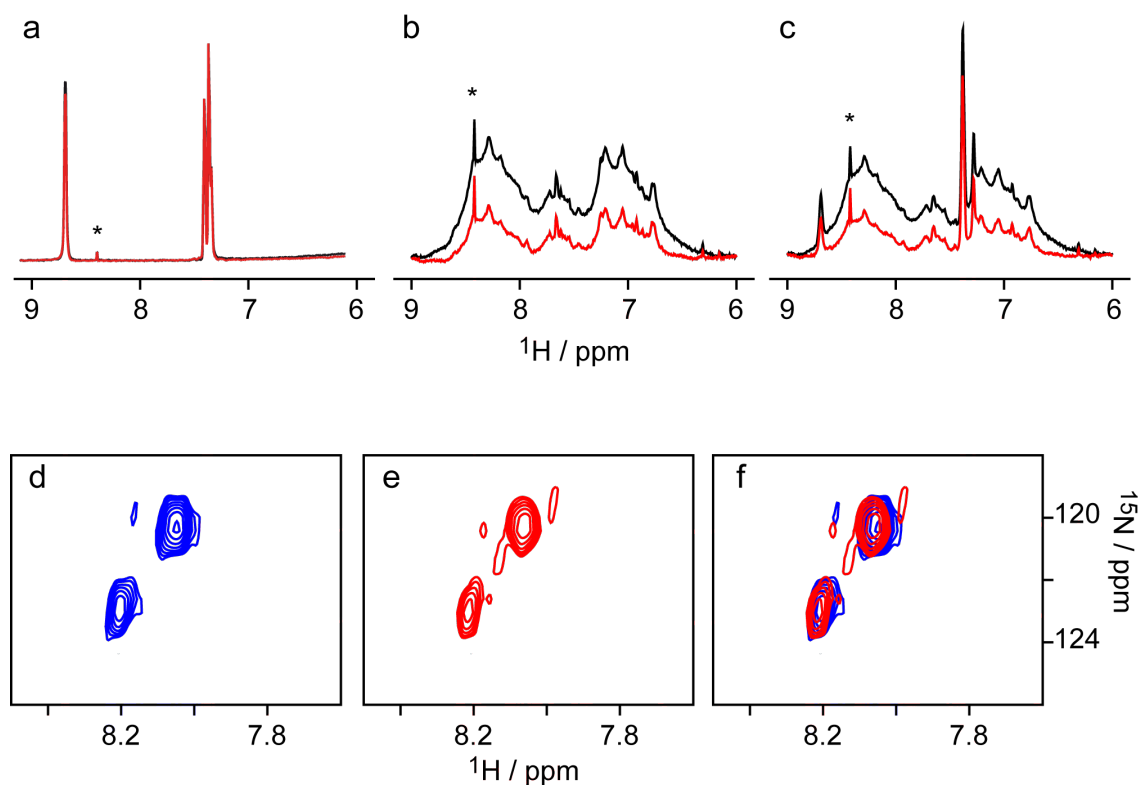

Figure 1.8 One and two dimensional NMR spectra of selectively ^{15}N-tyrosine labeled CCR5 and its antagonist maraviroc. **a – c** 1D ^{1}H NMR spectra of **a** 300 μM maraviroc, **b** 85 μM CCR5, and **c** the receptor/antagonist complex (85 μM CCR5, 425 μM maraviroc). The buffer conditions for all samples were the same (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 5% D$_2$O, pH 7.5, $\approx$ 2% FosCholine-12, 15 °C). Spectra were recorded with the 1-1 sequence with either a short delay (0.1 ms, black) or a longer delay (2.1 ms, red). The asterisk marks a small molecular weight impurity that does not relax significantly within the 2 × 2 ms delay difference. **d – f** 2D ^{1}H-^{15}N HSQC spectra of **d** ^{15}N-Tyr labeled CCR5, **e** the CCR5/maraviroc complex, and **f** the overlay of the two.

Figure 1.8d-f shows the comparison of the ^{1}H-^{15}N HSQC spectra of tyrosine labeled CCR5 with and without maraviroc as well as the overlay of the two. Very small chemical shift changes in the proton dimension can be observed. Nevertheless, since only two peaks are observed and they shift to a similar extent in the same direction, this result should be considered with caution.

The stability of CCR5 in complex with maraviroc can be estimated from the peak intensities in three consecutive ^{1}H-^{15}N HSQC experiments. The signals decrease in intensity only marginally within 48 h, with the intensity difference being about 0.5 times the top-level noise. Thus the sample seems to be more stable compared to ^{15}N-Ile CCR5. Whether this effect is due to the lower temperature (15 °C *versus* 40 °C) or binding of the antagonist maraviroc is currently unclear. However, it was observed before in antibody binding experiments (see chapter 1.2) that CCR5 is longer recognized if kept at 4 °C compared to 25 °C.
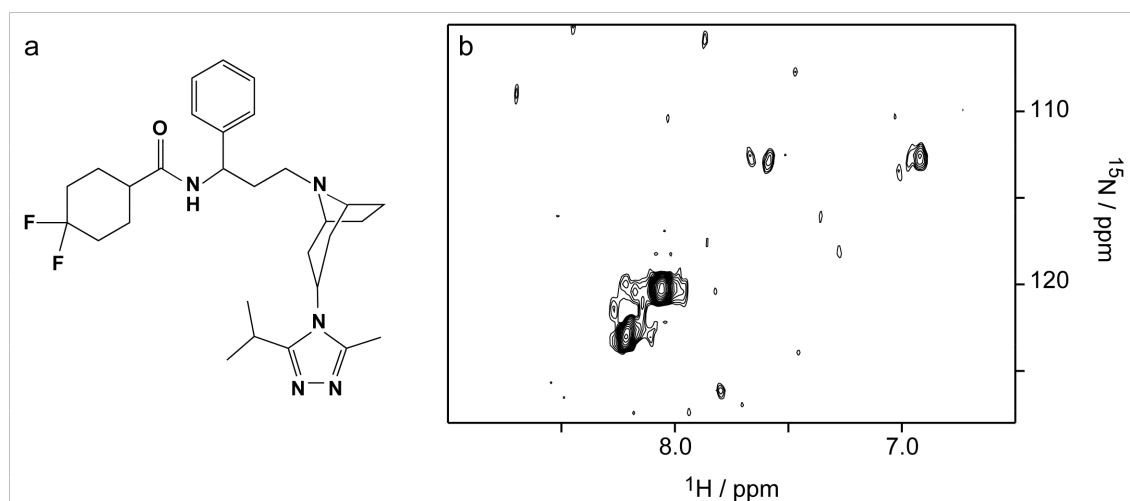

Figure 1.9 **a** Chemical structure of the CCR5 antagonist maraviroc. **b** Two dimensional ^{1}H-^{15}N HSQC spectra of selectively ^{15}N-tyrosine labeled CCR5 in complex with the antagonist maraviroc. The spectrum was recorded at 15 °C for 140 h.

Figure 1.9b shows the resulting ^{1}H-^{15}N HSQC spectrum after 140 h of measurement under the above mentioned conditions. In addition to the two major peaks observed already in shorter experiments, some more resonances appear in the same as well as in the side chain NH$_2$-region (see above). Nevertheless, the emerging peaks represent only a fraction of those expected.

Titration of uniformly ^{15}N-labeled RANTES with detergent-solubilized, unlabeled CCR5

The chemokine RANTES is an endogenous ligand of CCR5 and the most potent naturally occurring anti-HIV-1 agent known to date. The mode of receptor binding is not yet understood but there is evidence that only monomeric RANTES interacts with the receptor (Duma *et al.* 2007). Furthermore it was shown in mutation studies that tyrosine sulfation of CCR5's N-terminus is essential for interaction with chemokines (Farzan

et al. 1999, Bannert *et al.* 2001). Since we could show that the CCR5 produced in insect cells is at least partially sulfated (see chapter 1.2), the unlabeled receptor was used for the NMR titration of ^{15}N-labeled RANTES (RANTES here refers to the less aggregation-prone mutant RANTES-E66S; for more details see chapter 1.4).
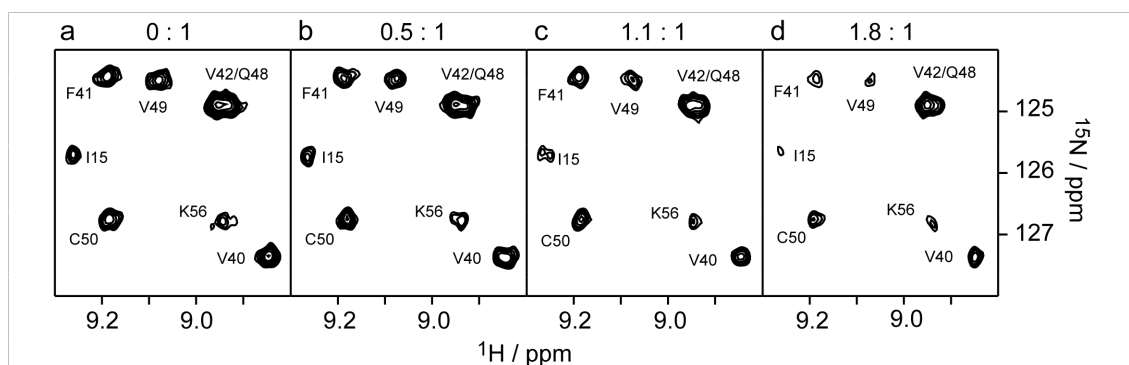

Figure 1.10 Titration of uniformly ^{15}N-labeled RANTES-E66S with non-labeled CCR5. **a** Selected region of the reference ^{1}H-^{15}N HSQC spectrum of 50 μM RANTES-E66S. The spectrum was recorded under exactly the same buffer condition as all further titration points (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 5% D$_2$O, pH 7.5, 2% FosCholine-12, 5 °C). **b − d** With increasing amounts of CCR5 in the sample, the intensities of RANTES-E66S resonances decrease. The CCR5 to RANTES ratios in the shown spectra are 0 to 1.8 from **a** to **d**.

Figure 1.10 shows a sequence of ^{1}H-^{15}N HSQC spectra corresponding to different titration points from 0 : 1 up to 1.8 : 1 (CCR5:RANTES). It should be noted, that all peaks seen in these spectra correspond to dimeric RANTES, since the equilibrium is shifted to this form at the pH used in the experiment (pH 7.5). With increasing amounts of the receptor, signals of the ligand are substantially weakened and no new signals that could be attributed to a bound state are observed. Since the same result is obtained with ^{2}H/^{15}N-RANTES, it is very likely that the particles present in solution are too big to be seen, that means more than hundreds of kDa. This is in contrast to our observations during the characterization of CCR5 by gel filtration-coupled static light scattering as well as one dimensional proton NMR diffusion measurements. It cannot be excluded that RANTES, which is prone to aggregation at elevated pH despite the mutation of glutamic acid 66, is not stable under these conditions. To cope with this problem it would be necessary to find conditions under which both, the ligand and the receptor are equally stable.

Solid state NMR spectroscopy of uniformly ^{13}C/^{15}N-labeled RANTES and its complex with reconstituted CCR5

In addition to solution state NMR spectroscopic investigations of the RANTES – CCR5 interaction, we also pursued a different approach, namely solid state NMR spectroscopy. In this approach, ^{13}C/^{15}N-labeled RANTES and unlabeled, lipid-reconstituted CCR5 was used to investigate the ligand – receptor interaction.

Reconstitution of CCR5 into liposomes   Solid state NMR spectroscopy has not only the advantage of being, in principal, not limited by the size of the system under investigation. With respect to membrane proteins it also offers the possibility to study them in a lipid environment similar to the membranes where they naturally occur. CCR5 has been successfully reconstituted into a lipid mixture (POPC, POPE, and DMPA) on the surface of paramagnetic beads (Mirzabekov *et al.* 2000). The mixture that was found to keep CCR5 functional in this study was used to prepare proteoliposomes for solid state NMR spectroscopy (Figure 1.11a).

Two different reconstitution methods have been used to produce CCR5-containing proteoliposomes. For both attempts, CCR5 was purified in the detergent Cymal-5 since this was found to be easier removed than FosCholine-12. Figure 1.11b shows a Western blot of the reconstitution of CCR5 into lipids after detergent removal with BioBeads (Rigaud *et al.* 1997). Lane 1 shows the protein after purification and before overnight incubation with the lipid mixture (lipid-to-protein ration (LPR) = 4 (*w/w*)). The detergent was removed by addition of 5 mg BioBeads per 5 mg detergent and incubation for 5 h. Removal was completed by further addition of two times 15 mg beads and incubation for 1.5 h. Lanes 2 and 3 show the supernatant after centrifugation of the mixture and washing of the resulting lipid pellet, respectively. To assay the conformation of reconstituted CCR5 the lipid pellet was resuspended and incubated with the conformation-dependent antibody 2D7. Lane 4 shows the supernatant after antibody binding and subsequent centrifugation. The very faint band of 2D7 indicates an excess of the antibody. The lipid pellet was gently washed with buffer. Tight binding of the antibody to CCR5 is obvious from the absence of any antibody band in the washing fraction (lane 5). Lane 6 shows the proteoliposomes solubilized in SDS sample buffer. CCR5 as well as the 2D7 antibody can be seen in the blot. Figure 1.11b also shows the result of a control experiment without CCR5 (lanes 4* – 6*), where the antibody is found only in the supernatant but not in the lipid pellet.

Although the reconstitution of CCR5 using BioBeads works well for high LPRs, it is not suitable for lower amounts of lipids, since the protein aggregates on the beads.

One limitation in solid state NMR spectroscopy is the volume of the sample rotor (e.g.
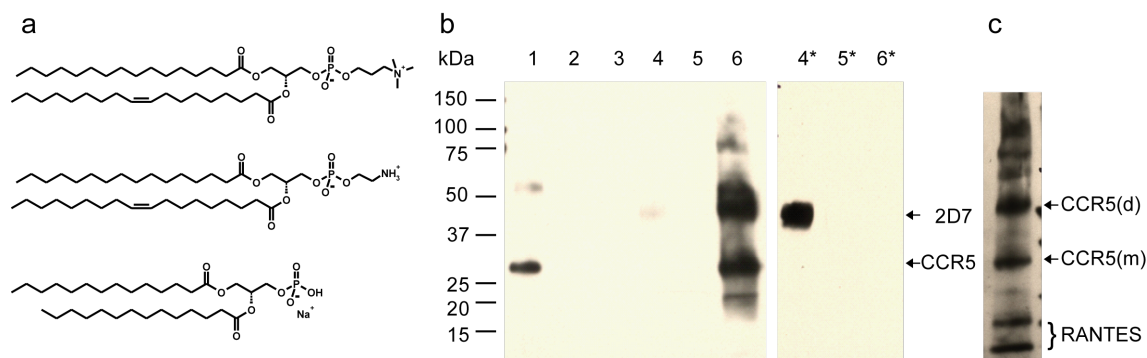
Figure 1.11   **a** Lipids used to reconstitute CCR5. POPC, POPE, and DMPA were mixed in a 6:3:1 ratio. **b** Western blot of the reconstitution and antibody recognition of CCR5 using BioBeads (detection with anti-His antibody). Lanes corresponding to the CCR5-free control are marked (*). Lane 1 shows the purified CCR5. Lanes 2 and 3 correspond to supernatants after reconstitution and washing. Lanes 4 and 5 show the supernatants after addition of the antibody and a washing step. In lane 6 the resolubilized proteoliposomes are shown. **c** Resolubilized proteoliposomes after reconstitution of the CCR5/RANTES complex using methyl-β-cyclodextrin. RANTES monomer and dimer as well as CCR5 monomer, dimer, and some higher order oligomers can be seen.

60 μL active volume for a 4 mm rotor), thus it is desirable to use low LPRs in order to maximize the amount of protein. In the study of a neurotensin-hexapeptide for example, 2.5 mg of the receptor were necessary to study the bound ligand by solid state NMR experiments (Luca *et al.* 2003). To accomplish a low LPR, a second reconstitution method was tested, namely the detergent removal with methyl-β-cyclodextrin (MBCD) (Signorell *et al.* 2007). CCR5 was purified and incubated with lipids overnight. Since the aim of our solid state NMR experiments was to study the CCR5/RANTES interaction the chemokine was added before detergent removal (1:1 RANTES/CCR5 molar ratio). Cymal-5 was complexed by addition of 1.2 equivalents MBCD over the course of 8 h and separated from the proteoliposomes by centrifugation. With this procedure, CCR5 could be reconstituted into lipids with a LPR of 2 while bound to its ligand RANTES as shown in Figure 1.11c. It is important, that the detergent is not removed too fast. If MBCD was added within only 2 h, protein precipitation was observed. It has been reported before that the slower the detergent is removed the better the quality of the reconstitution, in these cases the quality of 2D crystals (e.g. for OmpF or SoPIP2;1 (Signorell *et al.* 2007)). Therefore, it might be possible to obtain reconstitution of CCR5 with even lower LPRs and thus higher protein packing densities, if the MBCD addition time is further extended. This procedure was also used for initial 2D crystallization trials, carried out by Marco Biasini and Hervé Rémigy.
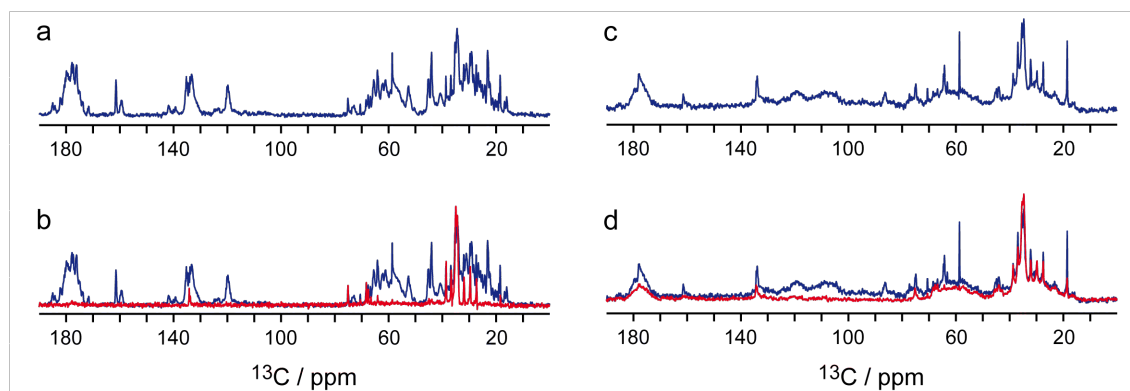
Figure 1.12   One dimensional solid state ^{13}C MAS-NMR spectra of uniformly ^{13}C/^{15}N-labeled RANTES. **a** 1D direct ^{13}C excitation spectrum of RANTES together with a lipid mixture (POPC/POPE/DMPA; 6:3:1). Resonances of both, lipids and protein can be seen. **b** 1D cross-polarization ^{13}C spectrum of RANTES (red). Only lipid resonances can be observed. The direct excitation spectrum is also shown for comparison. **c** 1D direct ^{13}C excitation spectrum of RANTES with proteoliposomes containing CCR5. Again, resonances of lipid and protein can be observed. **d** 1D cross-polarization spectrum of the RANTES/CCR5/lipid mixture (red). In addition to the lipid resonances also protein peaks can be seen. All spectra were recorded at $\approx$ 0 °C (-10 °C VT-gas temperature) and MAS rates of 10 kHz.

One dimensional solid state MAS-NMR spectroscopy   Specific binding of the ligand to its receptor under the conditions used is a prerequisite for any further structural studies. In order to test whether RANTES interacts unspecifically with the lipids used to reconstitute CCR5, ^{13}C/^{15}N labeled protein was added to the lipid mixture and used as reference without CCR5 (2 mg ^{13}C/^{15}N-RANTES and 20 mg lipids). One dimensional solid state ^{13}C MAS-NMR spectra were initially used as a quick test for the interaction. Figure 1.12a shows the spectrum of the reference sample (^{13}C/^{15}N-RANTES and lipids) after direct carbon excitation at $\approx$ 0 °C (VT-gas temperature -10 °C). Between 20 and 40 ppm, signals corresponding to the lipid can be seen in addition to the aliphatic carbons of RANTES (10 – 50 ppm). The protein's α-carbon resonances appear around 60 $\pm$ 10 ppm and its aromatic carbons between 120 and 140 ppm. At 180 ppm, carbonyl carbon resonances occur and can be used as a monitor for the state of the protein (see below). In the direct carbon excitation experiment the resonances of all carbons in the sample are visible.

In contrast, if cross-polarization from ^{1}H to ^{13}C is used (instead of direct excitation) only those carbons yield signals that are immobile. In cross-polarization (CP) experiments, transverse proton magnetization is generated and spin-locked for a certain time (mixing time). Simultaneously, a carbon pulse of a specific strength is applied and if the Hartmann-Hahn conditions are fulfilled (Hartmann & Hahn 1962), polarization is

transfered by dipolar interaction from the protons to carbon. Cross polarization is only effective for molecules that are immobile, since dipolar interactions average to zero in the presence of rotational motion.

Figure 1.12b shows the resulting CP spectrum for the reference sample. Only lipids are visible. This indicates that the lipids are immobile under these conditions whereas RANTES is free to move in the remaining water between lipid layers. Thus, RANTES does not bind unspecifically to the lipid mixture used nor does it form large aggregates under the solid state NMR sample conditions. This is essential for studies of the CCR5/RANTES complex. Figure 1.12c shows the result of an one dimensional direct carbon excitation experiment of RANTES together with CCR5 containing proteoliposomes. The sample contained $\approx$ 1.5 mg ^{13}C/^{15}N-RANTES and 7.5 mg CCR5. The appearance of this spectrum is comparable to the reference, although signal intensities are lower due to a smaller amount of RANTES in the sample. The corresponding crosspolarization experiment yields a spectrum in which not only the lipid resonances but also protein signals are visible, e.g. those of the carbonyl carbons ($\approx$ 180 ppm, Figure 1.12d). This indicates that RANTES is less mobile in this preparation than it is in the reference sample, suggesting that it binds specifically to CCR5.
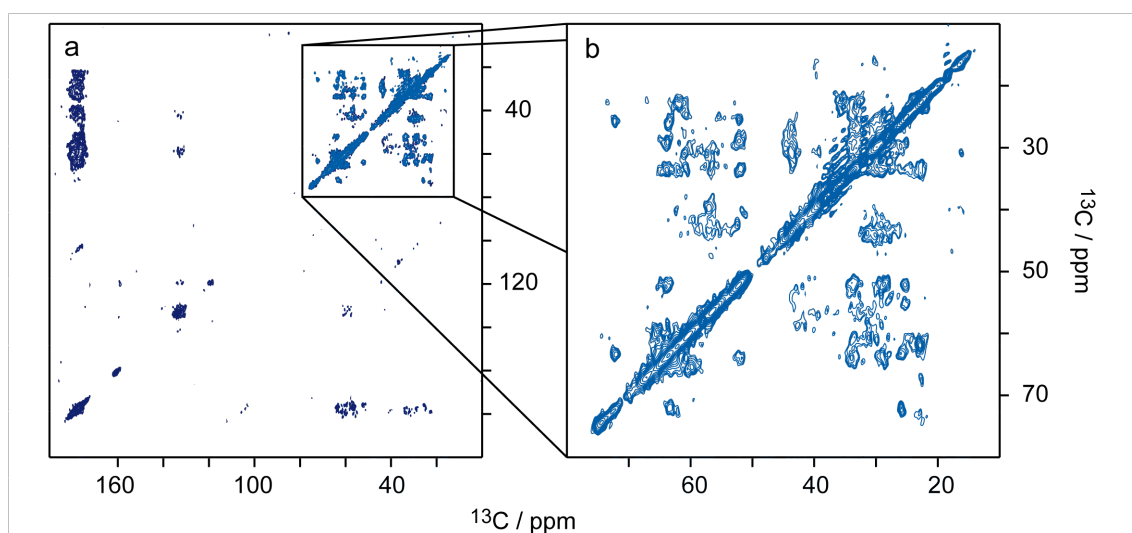
Figure 1.13   Two dimensional ^{13}C/^{13}C proton driven spin diffusion spectrum of RANTES together with a lipid mixture (POPC/POPE/DMPA; 6:3:1). **a** Full-view of the spectrum recorded with 11 kHz MAS at -40 °C (VT-gas temperature). Cross peaks between the carbonyl and C$^\alpha$ as well as the aromatic and aliphatic side chain carbons can be seen. **b** Enlargement of the C$^\alpha$/C^{aliph}-region of the spectrum showing cross peaks between these carbons.

Two dimensional solid state CP-MAS NMR spectroscopy Figure 1.13a shows the two dimensional ^{13}C-^{13}C proton-driven spin diffusion spectrum of the reference sample at $\approx$ -40 °C. The low temperature was chosen in order to obtain cross-polarization spectra even for free RANTES. Correlations between carbonyl carbons and aromatic, aliphatic, and α-carbons can be seen as well as among C^α and side chain carbons. The latter are depicted in the the enlarged spectral region (Figure 1.13b). Comparison to C^α and C^β chemical shifts predicted with ShiftX (Neal *et al.* 2003) from the crystal structure of RANTES (pdb 1eqt) allows several groups of peaks to be assigned to residue types. For example threonine C^α/C^β and C^β/C^γ peaks occur at 65/70 and 25/70 ppm, respectively. The line width obtained in this experiment is approximately 1 ppm. This is slightly broader than was reported for a peptide bound to the KcsA potassium channel (0.7 − 0.8 ppm (Lange *et al.* 2006)). Nevertheless, the resolution is not high enough to assign single peaks directly.

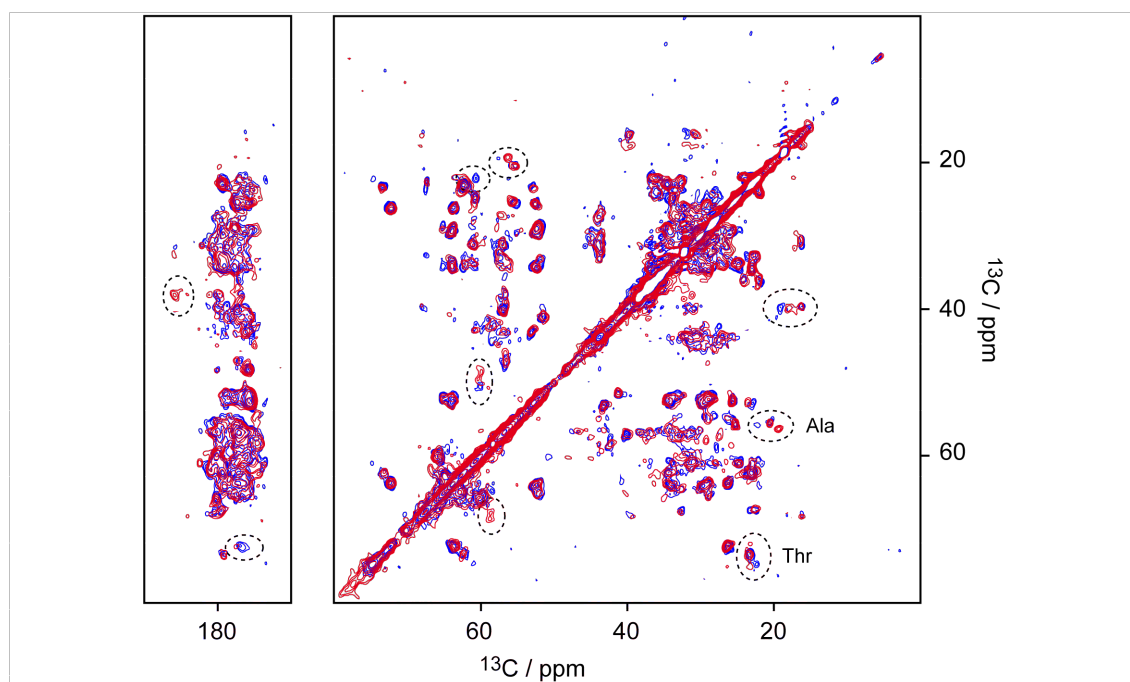**Figure 1.14** Two dimensional $^{13}C/^{13}C$ proton driven spin diffusion spectra of RANTES with lipids (blue) and with CCR5-containing proteoliposomes (red). The carbonyl to C^α/C^{aliph} cross peaks as well as the C^α/C^{aliph}-region is shown. For some resonances small changes in the overlaid spectra can be observed. Both spectra were recorded with 11 kHz magic angle spinning at -40 °C (VT-gas temperature) and a mixing time of 37.5 ms. The total measuring time was 112 h for the reference and 224 h for the CCR5-containing sample.

The sample containing the CCR5 − RANTES complex was measured under the same condition as the reference and data were compared. Figure 1.14 shows an overlay

of the C^{α}/C^{aliph} and the carbonyl region of the ^{13}C-^{13}C spin diffusion spectra of uniformly $^{13}C/^{15}N$-labeled RANTES with either lipids or CCR5-proteoliposomes. The general appearance of the two spectra is very similar, indicating that the overall structure of RANTES is not changed upon binding to the receptor. But for some peaks small changes can be observed, as indicated by dashed circles in Figure 1.14, e.g. threonine and alanine. Comparison to earlier solid state NMR studies of GPCR ligands shows, that also for a neurotensin-peptide in the free and receptor bound form, the side chain resonances did not change significantly upon receptor binding (Luca *et al.* 2003).

1.3.3 Conclusion and Perspective

The membrane protein CCR5 can be amino acid-type selectively labeled using the insect cell – baculovirus system, although for some amino acids scrambling is a problem. This enabled initial studies of the detergent-solubilized receptor using solution NMR spectroscopy. Signals of ^{15}N-isoleucine and ^{15}N-tyrosine CCR5 in heteronuclear $^{1}H/^{15}N$ correlation spectra have been observed for the first time. Although under our experimental conditions only a fraction of the labeled residues can be seen, this is an important starting point for further investigations of CCR5 in solution. The reason for the small number of resonances observed in 2D ^{1}H-^{15}N HSQC spectra could be either due to aggregation or due to conformational flexibility on an intermediate timescale. The first argument would explain why also RANTES, when titrated with detergent-solubilized CCR5, disappears. But it is not in agreement with findings from 1D ^{1}H NMR diffusion experiments and static light scattering data. This contradiction can be the result of higher CCR5 concentrations in the ^{15}N-isoleucine and ^{15}N-tyrosine labeled NMR samples. For further solution state NMR investigations it will be necessary to establish conditions that stabilize the receptor and prevent its aggregation. One possibility to decrease confomational flexibility and trap a single conformation of CCR5 can be the addition of antagonists. We have shown initial results depicting that the small molecule inhibitor maraviroc binds to CCR5. The CCR5 – maraviroc interaction could be further characterized by ^{19}F NMR spectroscopy, since maraviroc contains fluorine, and by saturation transfer difference experiments. With solid state NMR spectroscopy we could show that RANTES binds to CCR5 if reconstituted into liposomes, but does not unspecifically interact with the lipids. Thus, solid state NMR experiments of the RANTES/CCR5 complex are possible, although higher packing densities within the sample rotor are needed to reduce the measurement time. Comparison of first 2D spectra of free and receptor-bound RANTES reveals small differences.

1.3.4 Materials and Methods

Protein Expression

RANTES was prepared as described in chapter 1.4 and the CCR5 expression and purification is described in chapter 1.2. Unlabeled CCR5 was either produced in shaking flasks according to this protocol or in a WAVE bioreactor (GE Healthcare). In the WAVE system $5-10$ l can be cultured at a time. Cells are precultured to a density of 2×10^6 cells/ml in a shaking flask and transfered to the WAVE bag, where they are kept at densities between 0.5 and 2×10^6 cells/ml by stepwise dilution until the desired volume is reached. Then the cells are inoculated with a high-titer baculovirus stock at a multiplicity of infection (MOI) of 1. 24 h post infection 10% fresh medium are added. The culture is warmed to 27 °C and shaken at a rate of 20 (25) rocks per min with an angle of 7° (8°) before (after) induction. Cells are harvested ≈ 44 h post infection and treated as those grown in shaking flasks.

Selectively labeled protein was always expressed in shaking flasks. *Sf9* insect cells were adapted to and grown in SF4 insect cell medium (BioConcept) prior to expression. Cells were grown to a density of 4×10^6 cells/ml, centrifuged at 300g for 5 min and resuspended in SF4 medium deficient of yeast extract and the amino acids that are to be labeled, to reach a density of 2×10^6 cells/ml. This medium was supplemented with the ^{15}N-labeled amino acid and antibiotics (25000 U/l penicillin G, 25 mg/l streptomycin). Cells were inoculated with a MOI of 1 using a high-titer baculovirus stock. Cell culture and harvest as well as membrane preparation and protein purification were carried out as described in chapter 1.2.

Solution NMR sample preparation

Purified CCR5 was concentrated to ≈ 100 μM using a centrifugal filter device (AmiconUltra4 10kDa MWCO, Millipore). During concentration, the buffer was exchanged to 50 mM NaH_2PO_4, 300 mM NaCl, pH 7.5, 5% D_2O. The final concentration of the detergent FosCholine-12 was about 2%. For experiments with maraviroc, the antagonist was added from a 2.5 mM H_2O stock solution. Maraviroc was a gift from the NIH AIDS Research and Reference Reagent Program. The RANTES sample for titration with CCR5 was prepared as a 50 μM solution in the same buffer conditions as the receptor (50 mM NaH_2PO_4, 300 mM NaCl, pH 7.5, 2% FosCholine-12, 5% D_2O).

Solution NMR measurements and data analysis

All solution state NMR experiments were performed on a Bruker DRX 800 spectrometer equipped with a cryogenic triple-resonance, Z-gradient probe. Measurements were performed in Shigemi-NMR sample tubes. The ^{1}H-^{15}N HSQC spectra were recorded with acquisition times of 8.8 ms (^{15}N) and 85 ms (^{1}H). 1D ^{1}H spectra were recorded with the 1-1 sequence using delays of 0.1 ms and 2.1 ms. NMR data were processed with the NMRPipe suite of programs (Delaglio *et al.* 1995) and analyzed using the program PIPP (Garrett *et al.* 1991).

Solid state NMR sample preparation

For solid state NMR spectroscopy, CCR5 was purified as described but the detergent was exchanged from FosCholine-12 to Cymal-5 (0.2% in all buffers) during the two affinity chromatography steps. These and all following steps were carried out at 4 °C. After purification CCR5 was concentrated to 1 mg/ml and mixed with lipids (POPC/POPE/DMPA 6:3:1) solubilized in Cymal-5 to yield a LPR of 2. Lipids were added from a 5 mg/ml stock (50 mM NaH_2PO_4, 300 mM NaCl, pH 7.5, 6% Cymal-5). The mixture was gently shaken overnight before RANTES was added in stoichiometric amounts. After 1 h, Cymal-5 was slowly complexed by methyl-β-cyclodextrin (MBCD) over the course of 8 h. The amount of MBCD to be added was estimated from the amount of Cymal-5 present after the purification and the detergent added together with the lipid. About 1.2 equivalents MBCD were added to ensure complete removal of the detergent. Proteoliposomes and the MBCD/detergent-complex were separated by centrifugation for 30 min at 50000g. The resulting pellet was gently washed with buffer before it was packed into a 4 mm sample rotor with spacers at the bottom and top. The rotor finally contained 7.5 mg CCR5, 1.5 mg $^{13}C/^{15}N$-RANTES, and 15 mg lipids.

Solid state NMR measurements and data analysis

All solid state NMR experiments were performed on a Bruker DMX 600 NMR spectrometer equipped with a MAS probe. For sample cooling the air stream was passed through a bath of liquid nitrogen before entering the probe. 1D ^{13}C NMR spectra were recorded at $\approx$ 0 °C (-10 °C VT-gas temperature) either as direct excitation or as cross-polarization experiments. The MAS rate was 10 kHz. 2D $^{13}C/^{13}C$ proton driven spin diffusion spectra were recorded at $\approx$ -40 °C (VT-gas temperature) and 11 kHz magic angle spinning as data matrices of 192* (indirect) $\times$ 768* (direct) complex points with acquisition times of 13 ms and 18 ms, respectively. The total experimental time was 112 h for the reference sample without CCR5 and 224 h for the CCR5/RANTES complex. NMR data were processed with the NMRPipe suite of programs (Delaglio *et al.* 1995) and plotted using the software SPARKY (Goddard & Kneller n.d.).

1.4 Rational design of a monomeric RANTES mutant and characterization of different RANTES variants

Abstract

The chemotactic protein RANTES (Regulated upon Activation, Normal T cell Expressed and Secreted) belongs to the group of β chemokines and is involved in immune cell activation and migration. To accomplish its function, RANTES binds to CC chemokine receptors on the cell surface. The most prominent of those is CCR5, which is not only involved in chemokine signaling but also functions as the major human immunodeficiency virus 1 (HIV-1) coreceptor. As an endogenous ligand of CCR5, RANTES is the most potent known natural inhibitor of HIV infections. As many other chemokines, RANTES exists in a pH- and concentration-dependent monomer – dimer equilibrium. Previous studies of the interaction between RANTES and CCR5-derived peptides, mimicking the receptors N-terminus, have shown that only monomeric RANTES binds these peptides. Therefore, we engineered a solely monomeric variant of RANTES by disrupting the intermolecular β-sheet that forms in the wild-type protein. In the known mutant RANTES-E66S, which is less prone to aggregation than the wild-type, tyrosine 7 or tyrosine 8 was mutated to lysine to hinder dimerization. The resulting double mutants (RANTES-T7K/E66S and RANTES-T8K/E66S) were characterized by NMR pH titration and serial dilution. The solely monomeric variant RANTES-T7K/E66S shows anti-HIV activity in a p24-based assay similar to wild-type RANTES.

The monomer – dimer equilibrium of the single mutant RANTES-E66S was characterized under the influence of temperature and pressure. The dimerization of this mutant is influenced by enthalpic and entropic effects to a similar extent. At higher pressures, the equilibrium is shifted to the dimer and in addition a new, putatively unfolded monomeric state occurs, that has not been observed before.

1.4.1 Introduction

Chemokines are small bioactive proteins that regulate leukocyte activation and migration (Sodhi *et al.* 2004). Within the chemokine superfamily, these proteins can be classified into four groups according to a conserved cysteine motif in their N-terminus: α (CXC), β (CC), γ (C), and δ (CX$_3$C). Structurally, all chemokines share a similar topology with a C-terminal α-helix, at least three β-pleated sheets and an additional β-strand in the N-terminus. The N-terminus is connected to the rest of the protein by one (γ) or two (α, β, δ) disulphide bonds.

RANTES (CCL5) is a member of the β chemokine family and binds to the CC-type chemokine receptors CCR1, CCR3, and CCR5. Together with MIP-1α and MIP-1β it was identified as a natural suppressor of human immunodeficiency virus (HIV) infections (Cocchi *et al.* 1995). These chemokines antagonize binding of the viral envelope glycoprotein 120 (gp120) to CCR5, a coreceptor playing a crucial role in HIV-1 entry into macrophages. Although the *in vivo* use of full-length, unmodified RANTES as anti-HIV agent is limited due to its inherent proinflammatory activity (Baggiolini *et al.* 1994, Schall & Bacon 1994), functional peptide mapping has shown that RANTES-derived peptides recognize CCR5 and are capable of blocking HIV-1 *in vitro* (Nardese *et al.* 2001). More recently, it was also possible to engineer highly potent variants of RANTES where the G-protein-linked signaling activity could be separated from the HIV inhibiting property (Gaertner *et al.* 2008). One of these candidates (6P4-RANTES) shows a very similar activity than an earlier, N-terminal chemically modified form of RANTES, PSC-RANTES, that was shown to provide protection against vaginal challenge in macaques (Lederman *et al.* 2004).

Despite this success, the interaction between CCR5 and RANTES is still not well understood. This is partially due to the fact that wild-type RANTES tends to aggregate above pH 4, if concentrations higher than 1 μM are used (Appay *et al.* 1999). The aggregation can be prevented by mutation of glutamic acid 66 to serine, while retaining the biological activity (Czaplewski *et al.* 1999, Appay *et al.* 1999). Nevertheless, similar to many other chemokines (Clark-Lewis *et al.* 1995) RANTES (wild-type and the E66S-mutant) undergoes a pH- and concentration-dependent shift in the monomer – dimer equilibrium. The dissociation constant of RANTES-E66S is 17.6 μM at pH 3.8 and 37 °C (Duma *et al.* 2007). This means that, under the given conditions, a sample with 52.8 μM RANTES contains equal amounts of monomer and dimer. At higher, more physiological pH values, the concentration at which both forms are equally populated is even lower (8.7 μM at pH 6.0) (Duma *et al.* 2007). This is especially problematic for NMR investigations since the influence of sample concentration on the measurement time is quadratic. As a consequence of dissociation constants in the low-μM range only structures of RANTES

dimers have been elucidated (figure 1.15a) (Skelton *et al.* 1995, Hoover *et al.* 2000). Although the physiological relevance of monomeric, dimeric, and aggregated RANTES is not yet clear, interaction studies of RANTES-E66S with a peptide corresponding to the first 25 amino acids of CCR5 revealed that only monomeric RANTES is able to bind this peptide (Duma *et al.* 2007). These studies were complicated by the fact, that only low RANTES concentrations could be used in order to populate the monomeric species. For the same reason, the study could only be carried out at pH 3.8. Overall, the low dissociation constant of the RANTES dimer hinders interaction studies, especially at physiological pH, which is required for work with the full-length protein CCR5.

In contrast, for the CXC-chemokine SDF-1 it was shown that a constitutively dimeric mutant binds to a N-terminal fragment of its receptor CXCR4 (Veldkamp *et al.* 2008). Here we show the rational design of a monomeric RANTES mutant that is able to bind CCR5 and prevent HIV-1 infection *in vitro*. Two mutants, RANTES-T7K/E66S and RANTES-T8K/E66S, were prepared and characterized by NMR pH titration as well as serial dilution. The T7K variant is monomeric under all tested conditions and was therefore examined for CCR5 binding and *in vitro* HIV-1 inhibition. The T8K mutant does not have the desired properties and was therefore not investigated further.

In addition, we studied the influence of temperature and pressure on the RANTES-E66S monomer − dimer equilibrium. This allows a thermodynamic description of RANTES dimerization in terms of enthalpy and entropy contribution. In the course of these investigations a pressure unfolded state of RANTES was observed. This state is not described here in detail, but since the pressure unfolding of RANTES is reversible (at least under the conditions tested) it can be used as a system to study pressure unfolding by NMR spectroscopy.

1.4.2 Results and Discussion

Rational design of a monomeric RANTES mutant

The RANTES dimerization interface was analyzed for the rational design of a monomeric mutant in which the dimer formation should be disturbed. RANTES dimers form via an antiparallel β-sheet between the N-termini of the two polypeptide chains (figure 1.15b). This β-sheet is made up of the H-bonding amino acids threonine 8, proline 9, and cysteine 10. In addition, the side chain of threonine 7 can interact with the side chain of glutamine 48 in the second chain. On the other side the β-sheet is flanked by cysteine 11.

Proline 9 is structurally important for the protein backbone. This should be preserved, since it is desired to prevent dimer formation while maintaining CCR5 binding. Cysteine
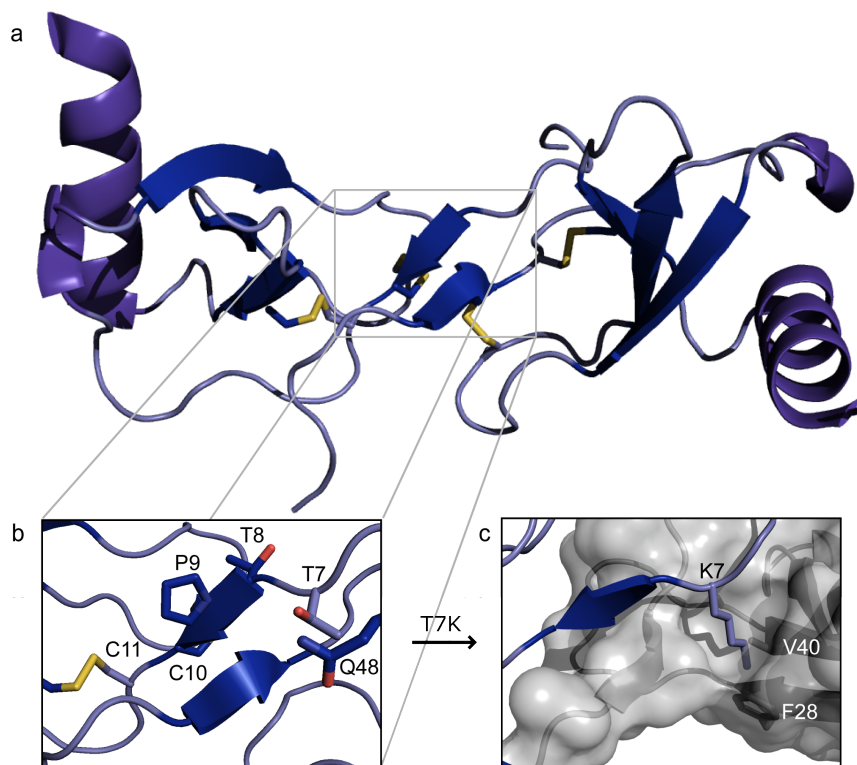
Figure 1.15   Rational design of a monomeric RANTES mutant. **a** Structure of the RANTES dimer (pdb 1eqt) showing the intermolecular β-sheet formed by the residues T8, P9, and C10. Disulfide bonds formed between C10 / C34 and between C11 / C50 within one chain are marked in yellow. **b** Detailed view of the dimer-forming sheet with the amino acid 7 to 11 shown with their side chains. Threonine 7 precedes the β-sheet and points towards the dimerization interface. A side chain hydrogen bond could be formed between T7 of one chain and Q48 of the second chain. **c** Hypothetical view of the T7K mutant of RANTES. Mutation of threonine 7 to lysine introduces a larger and, moreover, charged residue into the interface. This should prevent the formation of the intermolecular β-sheet due to steric hindrance and unfavorable electrostatics. The charged side chain of lysin would face a hydrophobic region with phenylalanine 28 and valine 40 in case the dimer would be formed.

10 forms a disulfide bond with C34 and cysteine 11 with C50, thus limiting the possibilities for mutations in the interaction site to threonines 7 and 8. In two separate mutants these residues were changed from threonine to lysine in order to add a longer side chain which can prevent dimer formation by steric hindrance. In case of the T7K mutation, the additional charge is expected to further repel the two RANTES monomers because it would face hydrophobic residues in the dimer (V40 and F28, figure 1.15c). Since the
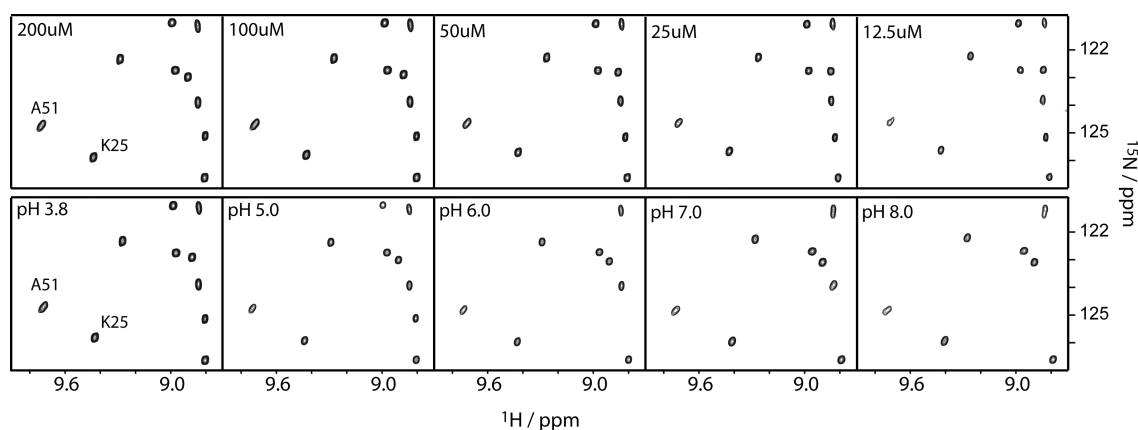
Figure 1.16   Selected region of the ^{1}H-^{15}N HSQC spectrum of RANTES-T7K/E66S. The upper part shows the serial dilution at pH 3.8 while the lower part shows the effect of increasing pH on a 100μM sample. All measurements were performed in 10 mM potassium phosphate buffer with 5% D$_2$O at 37 °C.

side chain of threonine 8 points to the interior of the protein, a mutation might affect the attachment of the N-terminus to the protein core more than it would influence the dimer interface.

NMR spectroscopic characterization of RANTES-T7K/E66S

The monomer – dimer equilibrium of wild-type RANTES (Skelton *et al.* 1995) as well as RANTES-E66S (Duma *et al.* 2007) has been characterized by one- and two-dimensional NMR spectroscopy. For both proteins, the monomeric and dimeric forms occur concurrently as separate species, indicating that they are in slow exchange ($k_{off} = 0.1$ s^{-1}) (Duma *et al.* 2007) with dissociation constants of 17.5 μM (wt RANTES) (Skelton *et al.* 1995) and 17.6 μM (RANTES-E66S) (Duma *et al.* 2007) at pH 3.8 and 37 °C. In order to characterize the monomer – dimer equilibrium of RANTES-T7K/E66S, a series of two-dimensional ^{1}H-^{15}N HSQC spectra was recorded at concentrations between 12.5 and 200 μM (figure 1.16). In contrast to the wild-type and the single mutant (RANTES-E66S), the double mutant RANTES-T7K/E66S shows only one set of resonances that does not change upon dilution. The chemical shift values of these peaks correspond to those of monomeric RANTES-E66S. Since the spectra for the RANTES single and double mutant are very similar (with respect to the monomeric species), the previously obtained assignment of RANTES-E66S (Duma *et al.* 2007) was transfered to RANTES-T7K/E66S.

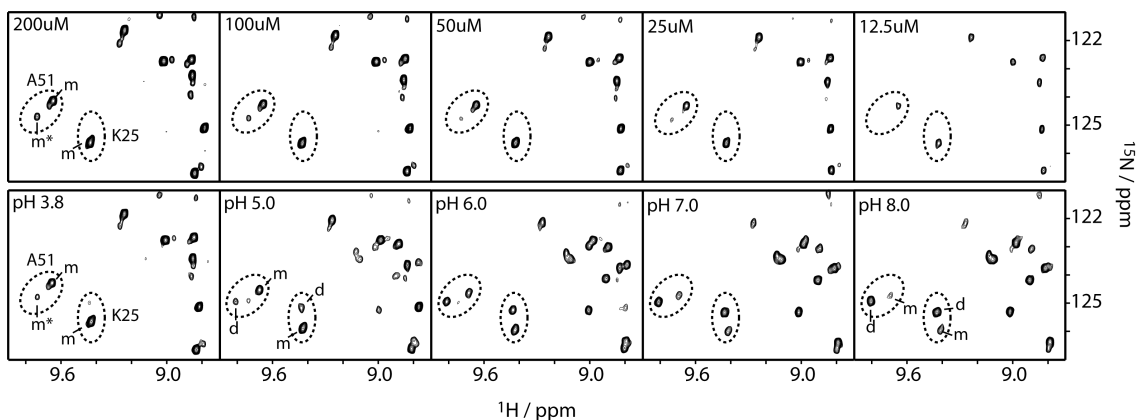The monomer – dimer equilibrium of wild-type RANTES and RANTES-E66S was found

Figure 1.17 Selected region of the ^{1}H-^{15}N HSQC spectrum of RANTES-T8K/E66S. The upper row shows the concentration dependence at pH 3.8 while the lower row shows the effect of pH on a 100μM sample. All measurements were carried out at 37 °C in 10 mM potassium phosphate buffer.

to strongly depend on pH, with the dimer association strengthened at elevated pH values (Skelton *et al.* 1995, Duma *et al.* 2007). Therefore, also the influence of pH on RANTES-T7K/E66S was investigated in a series of two-dimensional ^{1}H-^{15}N HSQC spectra. Data were acquired on a 100 μM sample at pH values between 3.8 and 8.0 (figure 1.16). Independent of the pH, this RANTES double mutant only occurs as monomer. This indicates that mutation of threonine 7 to lysine is sufficient to prevent the formation of an intermolecular β-sheet between the N-termini of two RANTES monomers under all tested conditions. Furthermore, the quality of spectra recorded at high pH is better for RANTES-T7K/E66S compared to the single mutant, due to an even lower tendency to aggregate.

NMR spectroscopic characterization of RANTES-T8K/E66S

Similar to the T7K/E66S-double mutant also RANTES-T8K/E66S was tested for its dimerization behaviour in a serial dilution from 200 μM down to 12.5 μM as well as in a pH titration (pH 3.8 to 8.0; figure 1.17). While RANTES-E66S shows almost only signals from the dimer at 200 μM (pH 3.8, not shown), the double mutant RANTES-T8K/E66S is predominantly monomeric under these conditions. Thus, dilution has no effect as can be seen for residue K25. In contrast, during the pH titration, the equilibrium is shifted towards the dimer with both forms being equally populated at pH 6.0.

For residue A51 the situation is more complicated with at least three species occurring under certain conditions. At the starting point (200 μM, pH 3.8) three peaks can be observed, two of which overlap. Upon dilution, the intensity ratio between the peaks does
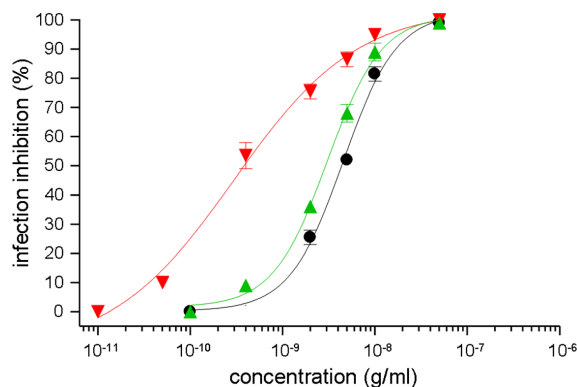
Figure 1.18   Inhibition of acute HIV-1 infection by wild-type RANTES (black), RANTES-E66S (red), and RANTES-T7K/E66S (green). The extracellular HIV-1 p24 antigen release was measured at day 6 post infection relative to an untreated culture as control (Figure courtesy of Luca Vangelista).

not change significantly between 200 and 25 μM. The peak marked m* is probably due to a second monomeric species. In contrast, upon increasing the pH, another separated peak occurs for pH values larger than 5.0 and increases in intensity whereas the other resonances slowly disappear. The newly appearing signal can be assigned to the dimer of RANTES. It is not clear why A51 and some other residues show more than just the monomeric and dimeric form. One possible explanation is that K25 is not very close to the site of mutation whereas A51 is. The mutation might cause some additional local effects leading to a third conformation in spatial vicinity to threonine 8. Since RANTES-T8K/E66S did not show the desired properties of being monomeric over a wide pH and concentration range, it was not investigated further.

CCR5-binding and in vitro activity test of RANTES-T7K/E66S

The release of the HIV-1 capsid protein p24 after infection of PBM cells with HI virus is a measure of infection efficiency. If the cells are treated with HIV-infection inhibitors, their potency can be tested.

Inhibition of acute HIV-1 infection by RANTES-T7K/E66S was tested in a p24-based assay and compared to wild-type RANTES as well as RANTES-E66S. The respective assays were carried out by Massimiliano Secci and Luca Vangelista (San Rafaelle Institute, Milan, Italy). After six days, RANTES-T7K/E66S shows a behavior similar to the wild-type with an average inhibition of extracellular HIV-1 p24 release (ID_{50}) of 0.46 nM (3.6 ng/ml) and 0.58 nM (4.6 ng/ml), respectively. In contrast, RANTES-E66S has a much lower ID_{50} value (0.04 nM (0.3 ng/ml); see Figure 1.18).

The fact that RANTES-E66S is ten times more potent than the wild-type might be explained by its lower tendency to aggregate. Therefore, the concentration of active protein is higher in case of the single mutant. With respect to the double mutant RANTES-T7K-E66S one would expect even lower ID_{50} values if the monomer was the active form. But since the RANTES N-terminus is crucial for CCR5 binding (Oppermann 2004) the single and double mutant cannot be compared easily. Due to the additional charge in its N-terminus, binding of RANTES-T7K-E66S to CCR5 might be weakened. Nevertheless, it seems to be similarly effective as the wild-type protein, making it an interesting alternative for studies of solely monomeric RANTES with both, CCR5-derived peptides as well as the full length receptor.

Influence of temperature and pressure on the monomer – dimer equilibrium of RANTES-E66S

Previously, the monomer – dimer equilibrium of RANTES-E66S was investigated as a function of pH and protein concentration (Duma *et al.* 2007). Here, we extend this study to the influence of temperature and pressure. Since RANTES (wild-type and the E66S mutant) is most stable at pH 3.8, all measurements were carried out at this pH. A protein concentration of 50 μM was chosen because earlier studies have shown that monomer and dimer are equally populated under these conditions (pH 3.8, 37°C) (Duma *et al.* 2007). ^{1}H-^{15}N HSQC spectra were measured at three additional temperatures as well as pressures up to 1500 bar.

Temperature-induced effects   Figure 1.19a shows the temperature-dependent behavior for residues K25 and A51. For lysine 25 two peaks occur at 17 °C (blue). Due to the available assignment of both, RANTES monomer and dimer, the peaks can be unambiguously attributed to the monomeric and dimeric form of the protein. The stronger peak of K25 corresponds to the dimer, the very weak signal arises from monomeric RANTES. With increasing temperature, this peak increases in intensity and, concurrently, the other peak decreases. The same can be observed for the resonances of alanine 51.
In earlier studies, the relative dimer intensity, defined as $I_r = I_d/(I_m+I_d)$, was used to describe the monomer – dimer equilibrium (Duma *et al.* 2007). The temperature dependence of this parameter for K25 and A51 is plotted above the spectrum. At 37 °C and 50 μM concentration, the values fit well to those obtained in earlier studies.

K_d can be calculated for the average of five amino acids (isoleucine 24, lysine 25, glycine 32, valine 49, and alanine 51) according to the equation $K_d = [m]^2/[d]$, where [m] and [d] are the monomer and dimer concentrations. With increasing temperature also the
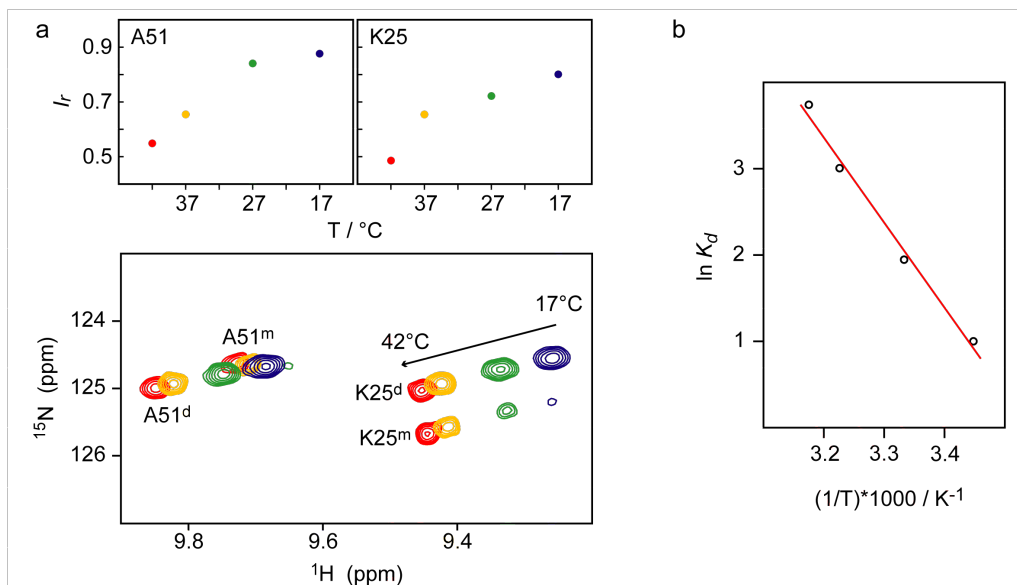
Figure 1.19 **a** Overlay of a selected region of the ^{1}H-^{15}N HSQC spectra of RANTES-E66S at 17 (red), 27 (yellow), 37 (green), and 42 °C (blue) showing the resonances of lysine 25 and alanine 51. With increasing temperature, the amount of monomer increases. On top, for each amino acid the relative dimer intensities are shown as a function of temperature. **b** van't Hoff plot of the temperature-dependence of the RANTES dissociation constant. K_d values obtained as the average of five amino acids are plotted as the natural logarithm against 1/T. The slope and intercept of this graph allow calculation of ΔH and ΔS (see text).

dissociation constant rises (e.g. $K_d(17°C) = 2.7$ μM; $K_d(42°C) = 42.2$ μM), which is characteristic for an endothermic reaction. This is in accordance with isothermal titration measurements (see chapter 1.2), where the RANTES dissociation was shown to consume energy. Using the van't Hoff equation also the dissociation enthalpy and the dissociation entropy can be calculated (figure 1.19b). For this calculation it is assumed that ΔH_d and ΔS_d do not change with temperature. However, the relation between lnK_d and 1/T is not linear for larger temperature ranges. Thus, the calculated values are not as accurate as calorimetrically measured values (Atkins 2000).

The enthalpy for the RANTES dissociation amounts to 82.0 kJ/mol (19.6 kcal/mol) and the dissociation entropy is 0.290 kJ/(mol K) (corresponding to 90.0 kJ/mol (21.5 kcal/mol) at 37 °C). This is in the order of magnitude found for other protein − protein interactions, e.g. (McDonald *et al.* 2008). The dissociation free energy at 37 °C $\Delta G_d(37) = $ -8.01 kJ/mol, which slightly favors dimer dissociation, is influenced by enthalpy and entropy to a similar extent.

The enthalpy ($\Delta H_d(37) = 82.0$ kJ/mol) favors dimerization of RANTES. Using a value
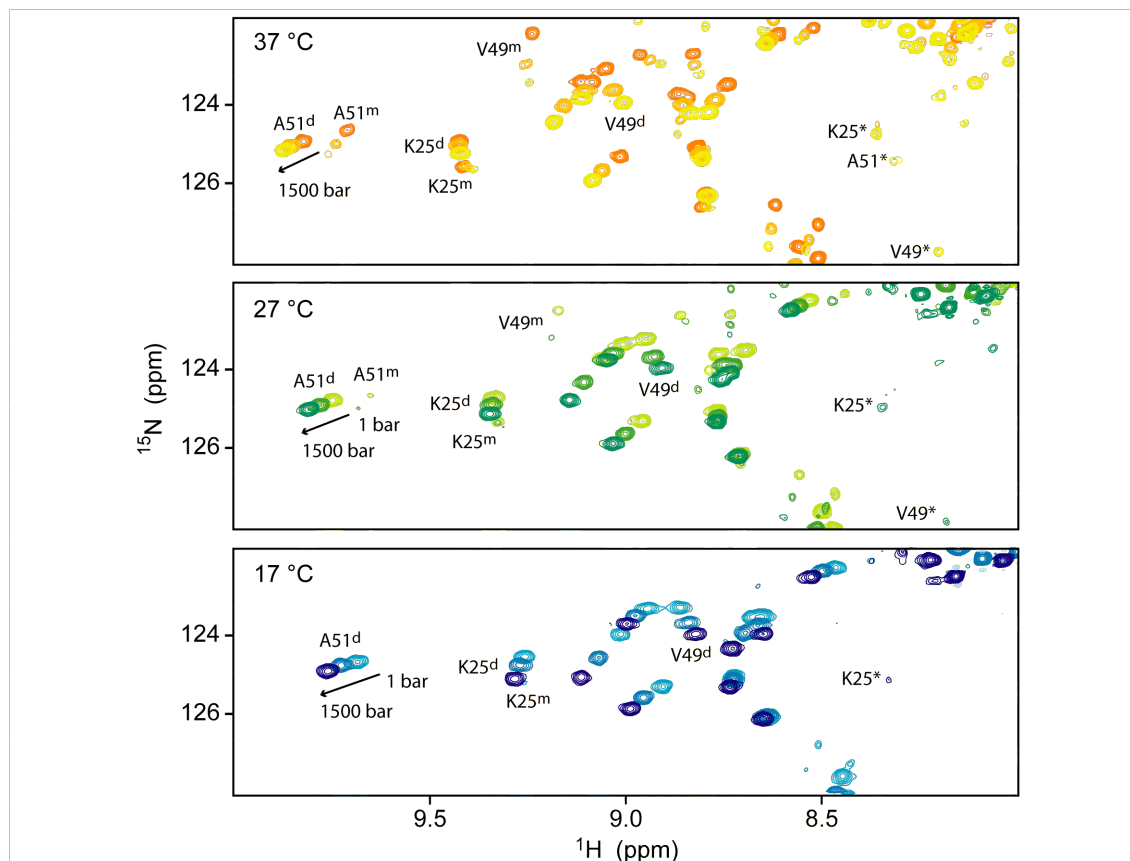
Figure 1.20   Selected region of the ^{1}H-^{15}N HSQC spectrum of RANTES-E66S at 37 °C (top), 27 °C (middle), and 17 °C (bottom). Each plot shows the overlay of spectra recorded at 1, 750, and 1500 bar. For clarity, not all peaks are assigned.

of 10 − 20 kJ/mol per bond (Rose *et al.* 2006), the three backbone hydrogen bonds forming the intermolecular β-sheet in the dimer would account for about 40% − 70% of this enthalpy. The positive dissociation entropy suggests, that the system must gain substantial disorder upon dissociation. Most of this effect can probably be attributed to the presence of two monomers instead of one dimer.

Pressure-induced effects   The effect of pressure on the monomer – dimer equilibrium was investigated in the range from atmospheric pressure up to 1500 bar. Figure 1.20 (top) shows a selected region of the ^{1}H-^{15}N HSQC spectrum of RANTES-E66S at 1, 750, and 1500 bar at 37 °C. In addition to the pressure-induced changes in chemical shifts the resonances corresponding to the monomer are significantly weakened with increasing pressure, while the dimer intensities are only slightly altered. Thus, the monomer is converted into the dimer only to a small extend. Instead, an additional set of peaks appears at 750 bar and increases in intensity at higher pressures. Due to the narrow

distribution range of the amide proton chemical shifts for these resonances, it is very likely that they correspond to an unfolded state of RANTES.

Since the solely monomeric mutant RANTES-T7K/E66S shows a very similar pressure dependent behavior with respect to unfolding, this RANTES variant was used to assign some of the newly appearing peaks from two dimensional ^{15}N exchange spectra. The assignment in this case is easier, because no dimer peaks occur in the spectra and thus the overlap is reduced. The obtained assignment could be easily transfered to the ^{1}H-^{15}N HSQC of RANTES-E66S. Thus, the intensities of monomer, dimer, and the unfolded state can be compared for five residues.

Figure 1.21 shows the relative peak intensities of all three species observed for iso-leucine 24, lysine 25, glycine 32, valine 49, alanine 51, and for the average of those five. At 37 °C, on average the dimer accounts for 64% at 1 bar and slightly increases relative to the other species to 69% at 1500 bar. Concurrently, the monomer population decreases from 36% at 1 bar to 12% at 1500 bar. At the same time the unfolded species rises from 0% at 1 bar to 18% at 1500 bar. These data illustrate that the unfolded state builds up on the expence of monomeric RANTES according to the following equilibria:

$$D \underset{\rightleftharpoons}{K_d} M \underset{\rightleftharpoons}{K_u} U \tag{1.1}$$

where D, M, and U are the dimer, monomer, and unfolded state of RANTES and K_d and K_u are the equilibrium constants for the respective transitions. The dissociation constant under the given conditions at 1 bar is $K_d = 20.2$ μM and decreases to 2.2 μM at 1500 bar. The volume change associated with RANTES dissociation differs between 1 and 1500 bar by $\Delta V_0 = 63$ Å^3. In contrast, the equilibrium constant between monomer and the unfolded state rises from $K_u = 0.4$ at 750 bar to 1.5 at 1500 bar. It can be concluded that the dimer is more compact than the monomer, since it is favored at higher pressures and it seems to be stabilized against pressure denaturation. This is supported by the fact that at lower temperature where there is less monomer present also the unfolded state is less populated (see figure 1.20), e.g. at 17 °C resonances of the unfolded species are barely visible even at 1500 bar. The same is observed for higher concentrated samples in which the dimer is predominant. In contrast, the monomeric mutant RANTES-T7K/E66S changes from the completely monomeric form at 1 bar to a distribution of 40% monomer and 60% unfolded protein at 1500 bar.

For this mutant, a rate constant for the exchange between the monomer and the unfolded state was estimated from exchange spectra (not shown) to be approximately 2.5 Hz.

The pressure-induced unfolding observed for both, RANTES-T7K/E66S and RANTES-
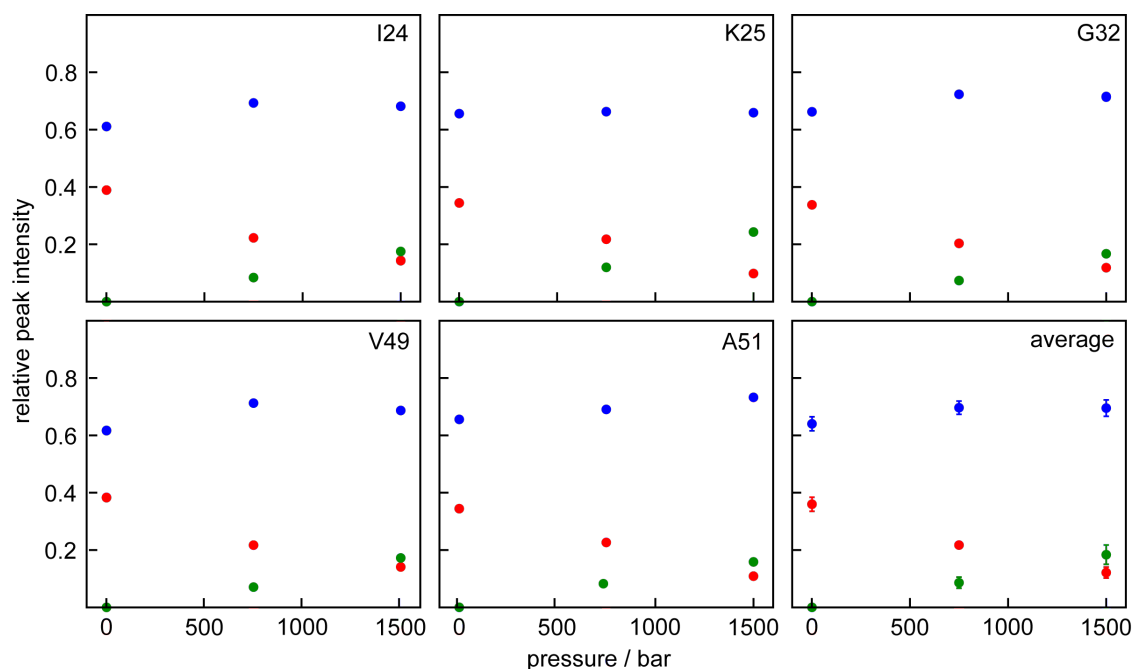
Figure 1.21   Distribution of the monomer (red), dimer (blue), and unfolded state (green) peak intensities at three pressures for five residues of 50 μM RANTES-E66S (pH 3.8, 37 °C) and their average. At atmospheric pressure only the monomer and dimer are populated. With increasing pressure, the dimer peak intensity changes only marginally, whereas the unfolded state builds up on the expense of monomeric RANTES.

E66S, is completely reversible.

1.4.3 Conclusion and Perspective

It was shown that it is possible to obtain monomeric RANTES in a functional form (with respect to *in vitro* HIV inhibition). Obviously the mutation in the ligands N-terminus does not affect the anti-HIV activity very much. RANTES-T7K/E66S is the mutant of choice to study the interaction of a solely monomeric chemokine with both, CCR5-derived peptides as well as the full-length receptor. RANTES-T8K/E66S is in between the single mutant (RANTES-E66S) and the T7K-variant with respect to the monomer – dimer distribution. Since it forms a significant amount of dimer and the mutation of threonine 8 not only affects the dimerization interface but also the protein core, it does not properly resemble the wild-type protein. Thus, it is not well suited for further studies.

In addition, the monomer – dimer equilibrium of RANTES-E66S was investigated as a function of temperature and pressure. Measurements at different temperatures were used to calculate the enthalpy and entropy for the dissociation of RANTES dimer to

monomers.

Increasing pressure shifts the monomer − dimer equilibrium of RANTES-E66S to the dimeric form. This is in line with other studies, where it was shown that high pressures have a similar effect than low temperatures (see chapter 2.4). More interestingly, pressure induces an unfolded state, that is reversibly formed on the expense of monomeric RANTES. In contrast, the RANTES dimer seems to be stable against pressure unfolding. It would be interesting to compare this finding with other ways of unfolding. If the dimer is generally more stable than the monomer this might have implications on the view of chemokine dimers, since their role *in vivo* is very controversial.

Formation of the unfolded state is reversible under all conditions tested and occurs with a rate constant of approximately 2.5 Hz at 37 °C and 1500 bar. Therefore, the monomeric mutant of RANTES, RANTES-T7K/E66S, could be used to further characterize the pressure unfolded state and investigate the folding / unfolding mechanism.

1.4.4 Materials and Methods

Expression and purification

Three RANTES constructs encoding the single (RANTES-E66S) as well as the double mutants (RANTES-T7K/E66S, RANTES-T7K/E66S) were expressed as C-terminal fusions to the protein G domain B1 separated by an 11 amino acid linker (GB1-LVPRGS-DDDDK-RANTES). These mutants differ from wild-type RANTES by one or two point mutations. All three have glutamic acid 66 exchanged to serine and in the double mutants additionally either threonine 7 or threonine 8 is replaced with lysine. Otherwise they consist of the amino acid sequence of mature human RANTES (S1 − S68). All constructs were expressed and purified as described (Duma *et al.* 2007). Briefly, proteins were expressed in *E. coli* BL21(DE21) and cells harvested by centrifugation. Cell rupture was performed in a French press after resuspension in 50 mM sodium citrate, pH 3.9. The resulting lysate was centrifuged for 20 min at 9000g and the pellet dissolved in 7 M urea, 20 mM sodium citrate, pH 3.9. After ultracentrifugation for 30 min at 100,000g the supernatant was applied to an equilibrated SP-sepharose column. A salt gradient (0.05 - 2 M sodium chloride) leads to elution of the protein and RANTES containing fractions were dialyzed against 50 mM Tris-HCl, pH 7.5, 20 mM NaCl for subsequent cleavage with enterokinase (37 °C, 16 h) to remove the fusion protein. After cleavage, the protein was further purified on an equilibrated SP-sepharose column (50 mM Tris-HCl, pH 7.5, 2 mM EDTA). After elution and dialyzation against 5 mM potassium phosphate, pH 3.8, the protein was concentrated by ultrafiltration (VivaSpin 20, MWCO 3 kDa). Masses for all three proteins were confirmed by ESI-MS.

NMR sample preparation

For the pH titrations and serial dilutions ^{15}N-labeled samples containing 50 mM potassium phosphate and 5% D_2O were used. Starting from a concentration of 200 μM, samples were diluted in the same buffer (pH 3.8) to concentrations of 100 μM, 50 μM, 25 μM, and 12.5 μM. pH titrations were carried out on 100 μM samples. The pH was adjusted to 5.0, 6.0, 7.0, and 8.0 with NaOH.

Pressure and temperature-dependent measurements were carried out on 50 μM samples (50 mM potassium phosphate, 5% D_2O, pH 3.8)

NMR spectroscopy and data processing

All NMR experiments were performed on a Bruker DRX 800 spectrometer, equipped with a triple-resonance, Z-gradient cryo-probe. The measuring temperature was 37 °C, unless otherwise stated. Measurements under high pressure were performed in a commercial high-pressure NMR-cell (tube1, see chapter 2.4) connected to a pressure generator. The ^{1}H-^{15}N HSQC spectra were recorded with acquisition times of 50 ms and 85 ms for ^{15}N and ^{1}H, respectively. NMR data were processed with the NMRPipe suite of programs (Delaglio *et al.* 1995) and analyzed using the program PIPP (Garrett *et al.* 1991).

Calculation of thermodynamic parameters

K_d values for different temperatures were calculated from the average monomer and dimer intensities of five well separated peaks according to $K_d = [m]^2/[d]$. The temperature-dependent K_d values in turn were used in the van't Hoff plot and fitted to the equation

$$\ln K_d = \frac{-\Delta H}{RT} + \frac{\Delta S}{R} \qquad (1.2)$$

Equation 1.2 was used to calculate ΔH from the slope and ΔS from the intercept of the fit $\ln K_d = -9.865/T + 34.935$.

The volume change was calculated according to

$$\Delta V_0 = \frac{RT}{\Delta p} * \ln \frac{K_d(1bar)}{K_d(1500bar)} \qquad (1.3)$$

Acute HIV-1 infection assay

The inhibition of acute HIV-1 infection by several RANTES variants was assayed in *in vitro*-activated human peripheral blood mononuclear cells (PBMC) (Nardese *et al.* 2001). The cells were infected with cell-free HIV-1$_{BaL}$ in the presence and absence of RANTES/ RANTES mutants at different concentrations. Extracellular HIV-1 p24 antigen release

was measured at day 4 and day 6 post infection. The inhibitory effect was measured relative to a control (untreated cells).

2 Hydrogen bonds in proteins and nucleic acids

2.1 Introduction

Proteins need sufficient conformational flexibility to accomplish their function yet they also need to be stable. A complete description of proteins requires a multidimensional energy landscape that defines the relative probabilities of the conformational states (thermodynamics) and the energy barriers between them (kinetics) (Henzler-Wildman & Kern 2007). Two parameters necessary for the thermodynamic description are temperature and pressure. Although temperature is the more extensively studied parameter in biological systems, there is no doubt that pressure is equally important as a thermodynamic variable (Smeller 2002). The pressure/temperature – phase diagram for proteins is ellipsoidal as shown in Figure 2.1. Within certain boundaries for temperature and pressure, a given protein can exist in its native state whereas high and low temperature as well as high pressure lead to denaturation.
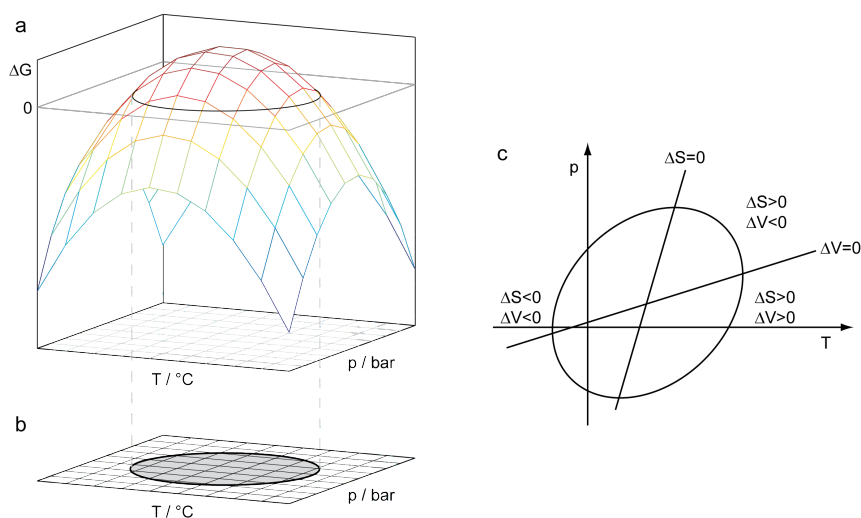

Figure 2.1   Schematic illustration of the elliptic phase diagram of proteins. **a** Three-dimensional representation of the influence of temperature and pressure on the free energy for the denaturation of a protein. Change of temperature can induce either heat or cold denaturation whereas an increase in pressure results in pressure denaturation of the protein. **b** Two-dimensional projection of the temperature and pressure boundaries within which the protein is native. **c** Pressure/temperature – phase diagram of a protein. Figure according to (Smeller 2002).

2.1.1 High-pressure investigations of proteins

The effect of pressure on a physicochemical process at equilibrium is governed by its volume change according to

$$\left(\frac{\delta \ln K}{\delta p}\right)_T = -\frac{\Delta V}{RT} \tag{2.1}$$

where K is the equilibrium constant of the process, δp is the pressure change at constant temperature, ΔV is the change in volume associated with the process, T is the absolute temperature, and R is the gas constant (Atkins 2000). The major factors contributing to volume changes are the electrostriction of charged and polar groups, the minimization of packing defects as well as the solvation of hydrophobic moieties (Frye & Royer 1998). Therefore, if a protein exists as different conformers, at elevated pressure those conformations with lower partial molar volume are favored, making them accessible for investigations.

Since pressure-denatured and pressure-assisted cold-denatured states have been shown to contain substantial residual secondary structure and furthermore, application of pressure allows the stabilization of intermediate conformations (Jonas *et al.* 1998), pressure has been used to study folding intermediates and denatured states of proteins (Kamatari *et al.* 2004). Such states can be characterized in atomic detail in solution using NMR spectroscopy.

2.1.2 High-pressure NMR spectroscopy of proteins

There are two general approaches to obtain NMR spectra under high pressures. One utilizes a special high-pressure probe (Jonas 2002), which allows pressures of up to 9000 bar. The other approach uses a high-pressure sample cell and a conventional probe. With this it is possible to obtain high resolution, but the main limitation until recently was the low sample volume of only ≈ 20 μl. Nevertheless, many proteins have been studied with this method at pressures between 1 and 4000 bar as reviewed by Kamatari *et al.* (Kamatari *et al.* 2004). There are several NMR parameters that can be monitored in dependence of pressure.

A very sensitive probe of conformational changes is the chemical shift since it is highly dependent on the chemical environment and changes therein. ^{1}H and ^{15}N chemical shift changes have often been used to map the effect of pressure on protein structures (Inoue *et al.* 2000, Kamatari *et al.* 2001, Kuwata *et al.* 2002).

Another parameter to characterize changes in the conformational states of proteins is ^{15}N spin relaxation. ^{15}N relaxation rates (R_1 and R_2) have been shown to be very similar at 30 and 2000 bar for BPTI (Sareth *et al.* 2000). Similarly, for ubiquitin the R_1

and $\{^1\text{H}\}$-^{15}N-NOE values do not differ significantly between 30 and 3000 bar, but R_2 showed changes for some residues at 3000 bar (Kitahara *et al.* 2005). These data have been interpreted in the sense that high frequency motions of individual NH bond vectors are not altered by pressure (Li & Akasaka 2006), whereas the variations in R_2 were attributed to μs- to ns-timescale exchange between different conformations (Kitahara *et al.* 2006).

Hydrogen/deuterium exchange experiments were also frequently used to characterize the transition from the native to the pressure denatured state. 1D ^{1}H NMR of RNase in D_2O was, for example, used to monitor residual structure in the pressure denatured protein (5000 bar) (Zhang *et al.* 1995).

2.1.3 Hydrogen bond scalar couplings as monitors of protein conformation

N-H $\rightarrow$ O=C hydrogen bonds defining the secondary structure of proteins are sensitive to conformational changes. Since H-bonding has large effects on NMR observables such as chemical shift and coupling constants of the nuclei involved, these (indirect) parameters have been used to characterize changes in the hydrogen bond network of proteins. Direct observation of hydrogen bonds in proteins and other biomolecules (see chapter 2.2 and 2.3) is possible due to the discovery of hydrogen bond scalar couplings (Dingley & Grzesiek 1998, Pervushin *et al.* 1998, Cordier *et al.* 1999).

The effect of pressure on the $^{3h}J_{\text{NC}'}$ coupling constants was investigated for the immunoglobulin binding domain of streptococcal protein G (GB1) at 30 and 2000 bar (Li *et al.* 2000). Due to the small sample volume (20 μl) and the generally small $^{3h}J_{\text{NC}'}$ coupling constants in proteins (-0.2 − -0.9 Hz) only eleven out of 34 hydrogen bonds could be detected. To be able to observe those coupling constants ($|^{3h}J_{\text{NC}'}| > 0.5$ Hz) measurement times of more than 5.5 days per experiment were required (Li *et al.* 2000). These long experimental times prevented the investigation of more pressure points, but from the differences between 30 and 2000 bar already some conclusions could be drawn. An interesting observation in this study was that pressure induces not only increased but for some residues also decreased $^{3h}J_{\text{NC}'}$ couplings. Decreased coupling constants correspond to a larger H − O distance in the hydrogen bond or a change in the angles away from 180 °. Thus pressure does not compress all parts of protein G but mainly the loop and peripheral regions. Nevertheless, due to the small number of hydrogen bond scalar couplings observed and the lack of more pressure points, a thorough description of the changes occurring in the hydrogen bond network with increasing pressure was not possible.

New set-up for high-pressure NMR measurements

A main problem in high-pressure NMR investigations so far has been the extremely low sample volume in tubes that allowed high pressures to be applied and high resolution to be obtained in the spectra. Therefore, much effort has been made in designing a device that could satisfy both needs.
Now a commercially available high-pressure NMR cell (Daedalus Innovations LLC) made of ceramics can be used for pressures up to 2500 bar in standard NMR probes. The inner diameter of the tube is 3 mm, allowing an active sample volume of 120 μl. This is six times more than was feasible before and thus reduces the measurement time accordingly.

Organization of the chapter

In the following two chapters, the methods are introduced that are used to measure hydrogen bond scalar couplings in nucleic acids (chapter 2.2) and proteins (chapter 2.3). In chapter 2.4 the influence of pressure on the $^{3h}J_{NC'}$ coupling constants is described for the protein ubiquitin. The obtained results are compared to temperature- and pressure-dependent measurements of the $^{2h}J_{NN}$ couplings observed in a small RNA hairpin (chapter 2.5). Finally, the concerted effect of methanol and pressure on ubiquitin is described (chapter 2.6).

2.2 Direct detection of N-H···N hydrogen bonds in biomolecules by NMR spectroscopy

Original Publication

Dingley, A.J., Nisius, L., Cordier, F., and Grzesiek, S.

Direct detection of N-H···N hydrogen bonds in biomolecules by NMR spectroscopy

Nature Protocols 3 (2008) 242-248.

Direct detection of N—H···N hydrogen bonds in biomolecules by NMR spectroscopy

Andrew J Dingley[1,4], Lydia Nisius[2,4], Florence Cordier[3] & Stephan Grzesiek[2]

[1]Department of Chemistry and School of Biological Sciences, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand. [2]Department of Structural Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland. [3]Department of Structural Biology and Chemistry, Institut Pasteur, 28 Rue du Docteur Roux, 75015 Paris, France. [4]These authors contributed equally to this work. Correspondence should be addressed to S.G. (stephan.grzesiek@unibas.ch).

A nuclear magnetic resonance (NMR) experiment is described for the direct detection of N—H···N hydrogen bonds (H-bonds) in ^{15}N isotope-labeled biomolecules. This quantitative HNN-COSY (correlation spectroscopy) experiment detects and quantifies electron-mediated scalar couplings across the H-bond (H-bond scalar couplings), which connect magnetically active ^{15}N nuclei of the H-bond donor and acceptor. Detectable H-bonds comprise the imino H-bonds in canonical Watson–Crick base pairs, many H-bonds in unusual nucleic acid base pairs and H-bonds between protein backbone or side-chain N—H donor and N acceptor moieties. Unlike other NMR observables, which provide only indirect evidence of the presence of H-bonds, the H-bond scalar couplings identify all partners of the H-bond, the donor, the donor proton and the acceptor in a single experiment. The size of the scalar couplings can be related to H-bond geometries and as a time average to H-bond dynamics. The time required to detect the H-bonds is typically less than 1 d at millimolar concentrations for samples of molecular weight $\leq \approx 25$ kDa. A ^{15}N/^{13}C-labeled potato spindle tuber viroid T1 RNA domain is used as an example to illustrate this procedure.

INTRODUCTION

H-bonds are key features of biomacromolecules as these weak, directional, attractive forces not only provide stabilizing rigidity to secondary and tertiary structures, but also play key roles in almost all enzymatic and chemical reactions[1]. Although X-ray and neutron diffraction methods have made invaluable contributions to the present knowledge of biomolecular H-bonds in terms of the exact positions of donor and acceptor groups, only for the few neutron diffraction and ultra-high-resolution X-ray structures of biomacro-molecules have the positions of hydrogen atoms been determined within the H-bond without the use of fixed standard geometries[2–4]. Thus, there is a need for techniques that yield additional quantitative information on the hydrogen atom and the electronic structure within individual H-bonds of biomacromolecules.

Currently, only infrared (IR) and NMR spectroscopy are able to provide such quantitative information on biomolecules. Whereas IR spectroscopy has severe limitations since single H-bonds can only be resolved via individual isotope-labeling schemes[5], NMR spectroscopy can usually resolve many individual H-bonds in uniformly isotope-labeled proteins or nucleic acids in one single experiment. A large number of NMR observables are sensitive reporters of H-bonding. These comprise hydrogen exchange rates, fractionation factors, chemical shifts, and quadrupolar and various scalar coupling constants[6]. The most direct of these parameters are H-bond scalar couplings (HBCs), which connect magnetically active nuclei on both sides of the hydrogen bridge. HBCs can be observed across the H-bonds of regular secondary structure elements in nucleic acids[7,8] and proteins[9,10], as well as in small chemical compounds[11–14]. Due to the requirement that the HBCs connect NMR observable nuclei, their observation in biomolecules has been limited to HBCs between ^{15}N, ^{13}C, ^{1}H or ^{31}P nuclei, but excludes oxygen, for which no suitable isotope exists.

Since the initial observation of HBCs for N—H···N H-bonds in Watson–Crick base pairs of nucleic acids (**Fig. 1**) and backbone N—H···O=C H-bonds in proteins (see ref. 15), there have been numerous reports detailing the measurement of HBCs for other

H-bond moieties in biomacromolecules (for reviews, see refs. 16–19). Besides the experimental observations of HBCs, a number of theoretical quantum chemical studies have improved our understanding of the nature of HBCs (reviewed by Grzesiek et al.[19]). Scalar couplings via H-bonds are caused by the same nucleus → electron → nucleus magnetization transfer that also underlies conventional couplings across covalent bonds. Consequently, the
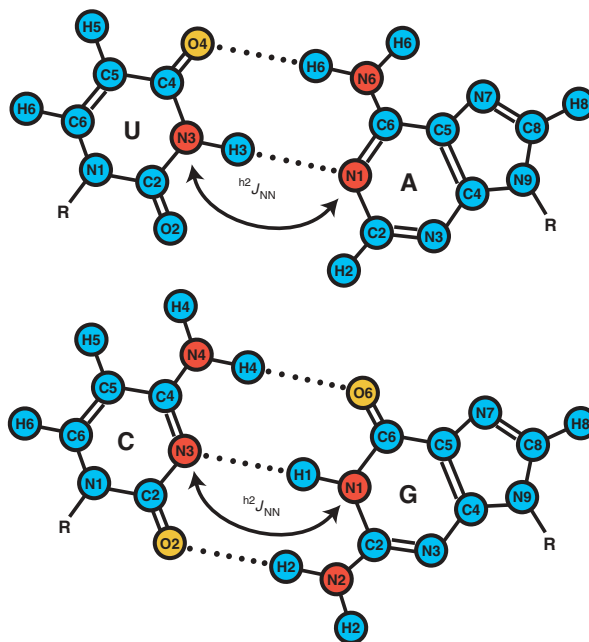

Figure 1 | Watson–Crick RNA base pairs indicating the N—H···N H-bond that can be detected by $^{h2}J_{NN}$ correlations. Similar detectable N—H···N H-bonds are found in many other base pair types.

same experiments for the detection, quantification or magnetization transfer are applicable to HBCs.

HBCs can be used to identify donor and acceptor groups in individual H-bonds from COSY-type experiments. In favorable cases, complete H-bond networks in biomacromolecules can be established, thereby providing valuable data for structure determination. As the HBCs are caused by the overlap of H-bond donor and acceptor electronic wave functions, their magnitude is proportional to the square of the overlap integral[20]. Thus, HBCs provide a very sensitive measure of H-bond geometries, that is, they depend exponentially on donor–acceptor distances[21] as well as on certain H-bond angles[19]. Such applications of HBCs have been reviewed extensively in ref. 19.

The strongest detectable HBCs in nucleic acids are two-bond $^{15}N-^{15}N$ couplings ($^{h2}J_{NN}$), which connect a ^{15}N H-bond donor nucleus to a ^{15}N acceptor nucleus across $N-H\cdots N$ H-bonds (the symbol $^{hn}J_{AB}$ is used for HBCs between nuclei A and B to emphasize that one of the n bonds connecting the two nuclei in the chemical structure is actually an H-bond) (**Fig. 1**). The magnitude of the $^{h2}J_{NN}$ couplings is primarily influenced by H-bond geometry rather than base pair type with sizes ranging between 5 and 11 Hz. The HNN-COSY experiment is the simplest two-dimensional (2D) technique by which $^{h2}J_{NN}$ couplings between ^{15}N imino donor and ^{15}N aromatic acceptor nuclei can be detected and quantified in nucleic acid base pairs. At the same time, the HNN-COSY yields the resonance frequencies of the ^{15}N donor and acceptor nuclei as well as of the H-bonding proton, that is, of all partners within the H-bond. With suitable modifications, the technique is also applicable to other $N-H\cdots N$ H-bonds, for example, to H-bonds involving amino $N-H$ donors[22,23] or to H-bonds between histidine imidazole groups[24].

The principal application of HNN-COSY in nucleic acids is the unequivocal establishment of H-bonding. This applies not only to $N-H\cdots N$ H-bonds in canonical Watson–Crick base pairs[7,8], but also to unusual base pairs that involve U/T-$^{15}N3$ or G-$^{15}N1$ to $^{15}N7$ nuclei, for example, in $G-A$ mismatch, Hoogsteen and reverse Hoogsteen base pairs[19,25]. In combination with residual dipolar couplings, this presents an approach for the fast assignment and structure determination of nucleic acids[26]. Due to the strong dependence on H-bond geometries, the size of the coupling constants provides a sensitive measure for variations in H-bond lengths and angles. Such variations can usually be detected as a result of fraying at the ends of stable structure elements, deviations from straight H-bond geometries or in response to changes in temperature or other physicochemical conditions[19,22,27–29].

As for any solution NMR experiment on biomolecules, the detection of H-bonds is limited by sensitivity and molecular weight. However, as the size of $^{h2}J_{NN}$ is relatively large, these limitations are not very severe, and the HNN-COSY usually yields reasonable results when other ^{15}N-edited experiments are applicable. Thus, detection of $N-H\cdots N$ H-bonds can typically be achieved in less than 1 d for samples of millimolar concentrations and molecular weights $\leq \approx 25$ kDa.

Organization of the protocol

The following protocol gives a step-by-step setup and evaluation procedure for the quantitative HNN-COSY experiment to detect and quantify $^{h2}J_{NN}$ couplings in $N-H\cdots N$ H-bonds between ^{15}N imino donor and ^{15}N aromatic acceptor nuclei. Example results are included from a 250 µl sample of 1.5 mM $^{15}N/^{13}C$-labeled potato spindle tuber viroid (PSTVd) T1 RNA domain (mutant U18C/ A344G), which consists of 69 nucleotides and contains three helical stems[30]. The data were acquired at 298 K on a Bruker DMX 600 MHz spectrometer equipped with a triple resonance TXI room temperature probe using the xwinnmr 3.5 acquisition software. Data processing and analysis were performed with the NMRPipe suite of programs[31]. This software and further information is available at http://spin.niddk.nih.gov/NMRPipe.

Experimental design

Figure 2 depicts the pulse scheme of the quantitative HNN-COSY experiment[7,8] used for the detection of $^{h2}J_{NN}$ couplings mentioned above. The scheme provides both the chemical shifts of the H-bonding nuclei and the magnitude of the $^{h2}J_{NN}$ coupling in a single 2D experiment. During the first INEPT (insensitive nuclei enhanced by polarization transfer) period (time points a to b), magnetization is transferred from the proton within the H-bond (H) via the covalent $^{1}J_{NH}$ coupling to the donor ^{15}N nucleus (N_d). During the $^{15}N-^{15}N$ COSY delay period (time points b to c, total

Figure 2 | Basic HNN-COSY pulse sequence. The color shading marks the boundaries of the various magnetization-transfer steps and the ^{15}N chemical shift evolution period as presented in the Bruker pulse sequence code (**Supplementary Note 1**). Narrow and wide pulses correspond to flip angles of 90° and 180°, respectively. Carrier positions are ^1H_2O (^1H) and 185 ppm (^{15}N). All regular ^1H and ^{15}N pulses are applied at radio frequency field strengths of 31 and 6.3 kHz, respectively. The radio frequency field strengths are specified for a 600 MHz spectrometer. For different magnetic field strength (X MHz) spectrometers, the radio frequency field strengths should be multiplied by $X/600$. Low-power (water flip-back) 90° ^1H pulses

(illustrated as smaller narrow rectangles) are applied at a field strength of 200 Hz. For ^{13}C-labeled samples, a high-power ^{13}C 180° pulse (25 kHz, shaded gray) is applied in the middle of the t_1 evolution period (see text). Delays: $\delta = 2.25$ ms; $\Delta = 15$ ms (delay period dependent on the length of $^{15}N-^{15}N$ COSY transfer employed); $\varepsilon = 0.6$ ms. Unless indicated, the pulses are applied along the x axis. Phase cycling: $\phi_1 = R_1$ with $R_1 = x, y, -x, -y$; $\phi_2 = R_2, -R_2$ with $R_2 = (y, -x, -y, x)$; $\phi_3 = -R_2, -R_2, R_2, R_2$; acq. $= x, -y, -x, y$. Quadrature detection in the t_1 dimension was achieved by simultaneously incrementing ϕ_1 and ϕ_2 in the States-TPPI manner. Gradients are sine-bell shaped, with an absolute amplitude of 30 G cm^{-1} at their center and durations (directions) of $G_{1,2,3,4,5,6,7} = 2.5 (+z), 2.1 (-z), 2 (+y), 2 (-y), 0.2 (+x), 0.4 (+z)$ and 0.404 ms ($-y$).
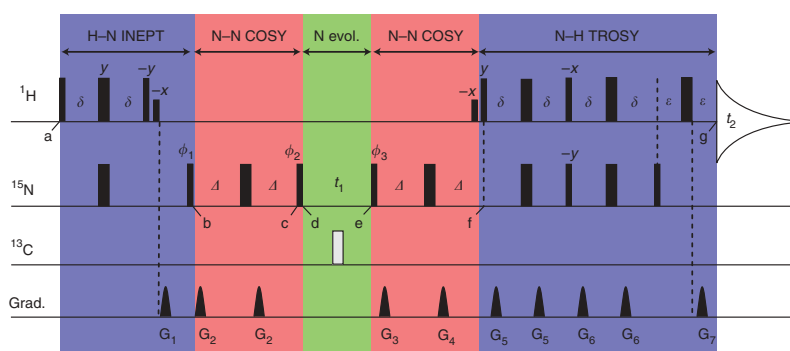
PROTOCOL

Figure 3 | 2D HNN-COSY spectrum recorded on the PSTVd T1 RNA domain, mutant U18C/A344G (see ref. 30). The data matrix consisted of $100*(^{15}N) \times 768*(^1H)$ data points (where $n*$ refers to complex points) with acquisition times of 16.5 (^{15}N) and 64 ms (^1H). Total experimental time was 12.6 h. Positive contours (blue) depict the diagonal resonances for the G and U imino groups. Negative contours (red) correspond to cross-peaks resulting from scalar $^{15}N–^{15}N$ magnetization transfer between the donor imino ^{15}N nucleus and the acceptor ^{15}N nucleus on the opposing base in Watson–Crick base pairs. For many U imino $^{15}N3$ nuclei, additional small cross-peaks are observed due to intramolecular $^2J_{NN}$ transfer to the U $^{15}N1$ nuclei. Resonances are labeled with assignment information. The inset illustrates the definition of the scalar $^{h2}J_{NN}$ correlation via the H-bond.

length 2Δ), part of this magnetization is transferred via the $^{h2}J_{NN}$ coupling to the ^{15}N acceptor nucleus (N_a), whereas another part remains on the donor ^{15}N nucleus. Following this COSY delay, both donor and acceptor ^{15}N chemical shift frequencies are encoded during the t_1 period (time points d and e). Note that, during the t_1 evolution, the small homonuclear $^{15}N–^{15}N$ couplings are active, but are usually not resolved. The magnetization is then transferred back to the donor ^{15}N nucleus during the second $^{15}N–^{15}N$ COSY interval (time points e to f, total length 2Δ). Finally (time points f to g), a transverse relaxation-optimized spectroscopy (TROSY) detection scheme converts the slowly relaxing component of the (^1H)-$^{15}N_d$ doublet into the slowly relaxing component of the ($^{15}N_d$)-^1H doublet, which is detected in the receiver during the t_2 period.

The resulting spectrum (**Fig. 3**) shows 'cross-peaks' at frequency positions (ω_{Na}, $\omega_H+\pi^1J_{NH}$) and 'diagonal peaks' at positions
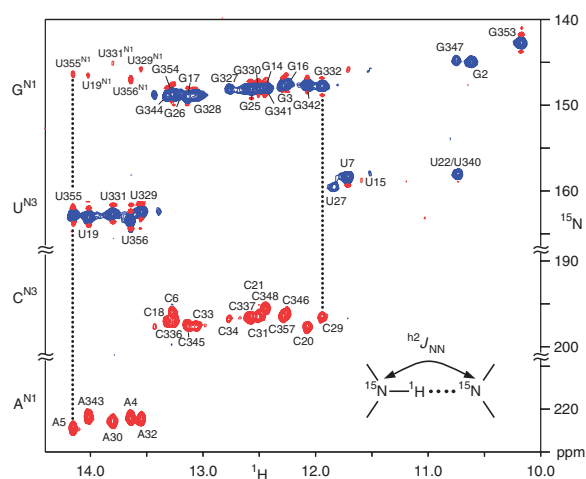
($\omega_{Nd}-\pi^1J_{NH}$, $\omega_H+\pi^1J_{NH}$). The intensities of the cross- and diagonal peaks are proportional to $-\sin^2(2\pi^{h2}J_{NN}\Delta)$ and $\cos^2(2\pi^{h2}J_{NN}\Delta)$, respectively. Hence, the intensity ratio presents a simple measure for the size of the $^{h2}J_{NN}$ coupling. The sign of the coupling cannot be determined from the HNN-COSY. However, in agreement with theory, the sign of the $^{h2}J_{NN}$ coupling has been determined as positive[8,32] by an E.COSY experiment[33].

MATERIALS

REAGENTS
· Nucleic acid samples must be uniformly ^{15}N-enriched. $^{13}C/^{15}N$-enriched samples can also be used. In this case, additional ^{13}C decoupling should be applied. Isotope-labeled nucleic acids for NMR studies are prepared using previously described strategies[34–38].
· Samples should be prepared in a buffered solution containing 5% (vol/vol) deuterium oxide (D_2O).
· For good sensitivity, sample concentrations should be ≥ 0.5 mM.
· Necessary sample volumes should range between 250 and 500 μl depending on the NMR tube type: conventional 5 mm NMR tubes require 400–500 μl for good shimming; Shigemi NMR tubes require only 250 μl or less. Note that, the achieved overall sensitivity will depend on the total sample amount within the active volume of the probe coil.

EQUIPMENT
· A high-resolution NMR instrument (≥ 500 MHz ^1H frequency) with the following characteristics:
  · ^{15}N-enriched nucleic acid samples: two radio frequency channels (^1H, ^{15}N) and double-resonance ($^1H/^{15}N$) probe head optimized for ^1H detection. For $^{15}N/^{13}C$-enriched nucleic acid samples, an additional ^{13}C radio frequency channel and a triple-resonance ($^1H/^{13}C/^{15}N$) probe head should be available.
  · The NMR spectrometer should be equipped with pulsed-field gradients; otherwise, water suppression can be achieved by other techniques[39].
  · As sensitivity is a limiting factor, cryogenic probes are beneficial but not absolutely necessary.

PROCEDURE

Preparation

1| Calibrate ^{15}N (and possibly ^{13}C) power levels on a suitable calibration sample according to **Box 1**.

2| Set the sample temperature of the NMR spectrometer to a suitable value, which is a compromise between sample stability and spectral quality. Place your sample in the magnet.

3| Calibrate ^1H power levels on your sample according to the different pulse lengths given in **Box 1**.

4| If you have no idea from previous experiments, determine the imino proton $^1H^N$ T_2 relaxation time from a suitable spin-echo experiment, for example, using the 1–1 echo sequence[40]. This step is not absolutely necessary, but it takes only a few minutes and is a very good way to characterize the quality of your sample. For the 69-nucleotide PSTVd RNA used in the example, the $^1H^N$ T_2 is about 12 ms at 298 K. Note that, in the absence of chemical exchange and internal motions, $1/T_2$ is proportional to MWT η/T, where MWT, η and T are molecular weight, viscosity of the solvent and absolute temperature of the sample, respectively. Under these conditions also, the ^{15}N T_2 values are proportional to the $^1H^N$ T_2 (slow motion limit). In the example, the T_2 values of the imino donor ^{15}N were about 50 ms.

Parameter setup

5| Load the pulse program for the 2D quantitative J_{NN} HNN-COSY experiment (**Fig. 2** and **Supplementary Note 1** online).

6| Set pulse lengths and power levels as determined in Steps 1 and 3 (see **Box 1**).

BOX 1 | HNN-COSY PARAMETERS

These parameters (see also the pulse program in **Fig. 2**) need to be set up by the operator to optimize acquisition.

1. Radio frequency carrier positions.
   - ^{1}H: exactly on water
   - ^{15}N: 185 ppm
   - ^{13}C: 153 ppm (for ^{13}C-labeled samples only)

2. Pulse lengths and power levels. Power levels should be determined according to the specified pulse length. Pulse lengths are specified for a 600 MHz spectrometer. For different field strength (X MHz) spectrometers, pulse lengths should be multiplied by 600/X.
   - ^{1}H: ≤ 10 μs and 1 ms 90° rectangular pulses
   - ^{15}N: ≤ 40 and 200 μs 90° rectangular pulses
   - ^{13}C: ≤ 15 μs 90° rectangular pulse (for ^{13}C-labeled samples only)

3. Sweep width and acquisition times.
   ^{1}H
   - Sweep width: ≈ 20 ppm (corresponds to a dwell time (dw) ≈ 83 μs at 600 MHz)
   - Number of complex increments (Ninc): choose such that the total acquisition time (aq = dw × Ninc) is about 3–5 times $T_2(^1H^N)$. An acquisition time of 64 ms was used in the example.
   ^{15}N
   - Sweep width: ≈ 100 ppm (corresponds to a dw ≈ 165 μs at 60 MHz)
   - Ninc: choose such that the total acquisition time (aq = dw × Ninc) gives a reasonable resolution, but that it does not exceed the T_2 of donor and acceptor ^{15}N nuclei. Note that due to their large chemical shift anisotropy, the T_2 of the acceptor ^{15}N nuclei is considerably shorter than that of the donor nuclei. An acquisition time of 16.5 ms was used in the example as a compromise.

4. N−N COSY transfer delay.
   Delay $\varDelta$
   - $\varDelta = 4$ ms for the test spectrum 1
   - $\varDelta = 15$ ms for the test spectrum 2 and final spectrum. $\varDelta$ may be set to smaller values for very large molecular weight samples.
   ? TROUBLESHOOTING

7| Adjust the radio frequency carrier positions according to **Box 1**. ^{1}H: set the carrier frequency exactly at the water position (at this carrier position, a simple ^{1}H pulse-acquire experiment should yield a non-oscillating water free induction decay); ^{15}N: set the carrier frequency to 185 ppm. This is the midpoint between the chemical shift of the donor imino nuclei (155 ± 10 ppm) and the chemical shift of the acceptor nitrogen nuclei (210 ± 15 ppm). This ensures maximum excitation of both donor and acceptor ^{15}N nuclei; ^{13}C (for ^{13}C-labeled samples only): set the carrier frequency to 153 ppm. This is the center frequency of the carbon nuclei (i.e., C2 and C6 in purine bases, C2 and C4 in pyrimidine bases), which are directly bonded to the nitrogen donor and acceptor nuclei. A ^{13}C 180° decoupling pulse (shaded gray in **Fig. 2**) is applied at this frequency in the center of the t_1 period to ensure decoupling of these nuclei.
? TROUBLESHOOTING

8| Set ^{1}H and ^{15}N sweep widths and acquisition times according to **Box 1**.

2D test spectra
9| Record a high sensitivity test spectrum with a short N−N transfer time (test spectrum 1). Set the length of the N−N transfer period (2$\varDelta$) close to the smallest possible value, for example, $\varDelta = d27 = 4$ ms, and record a 2D test spectrum with a small number of scans (NS = 4) and an interscan delay (d1) of about 1 s. This should take 10–15 min. Process the data using parameters similar to the ones given below (**Supplementary Note 2** online). The spectrum should look like a normal 2D TROSY-heteronuclear single quantum coherence (HSQC) of the imino resonances, which have chemical shift ranges of 140–165 ppm for ^{15}N and 10–14.5 ppm for ^{1}H nuclei (see **Fig. 3**). Check whether the peaks have good line shapes and phase behavior and whether all the expected imino resonances are observable. The sensitivity of test spectrum 1 should be somewhat less but still similar to a normal 2D TROSY-HSQC. If the sensitivity is not sufficient to observe all relevant resonances, increase the number of scans.

10| Check the signal-to-noise ratio of test spectrum 1. Using appropriate software, determine the ratio of the height (maximum) of typical imino peaks relative to the noise. In the simplest way, take a 1D trace through a typical peak and estimate its height relative to the top level of the noise in a region that does not contain peaks. This top level presents your detection limit for resonances.

   You can determine the signal-to-noise ratio, in a more quantitative way, using, for example, NMRDraw in the following way: (i) press l (for 2D location) and use the left mouse button to select a peak. The header of the window will show you the peak position and its height. Press z (for zoom) and use the middle mouse button to drag the box in an area without peaks. Press the middle mouse button somewhere outside the spectrum. This will open a window in which you can see the standard deviation ("SDev"; second row from bottom) corresponding to the standard deviation of the data points (noise) in the selected area. Typically, this standard deviation is about one-third of the top-level noise.

In the example of the 1.5-mM PSTVd RNA, the ratio of the heights of typical imino resonances relative to the top level of the noise (S/N_{top}) is in the range of 10–70:1 after 17 min total experimental time on a 600 MHz instrument with a room-temperature probe.

11| Record a test spectrum with a longer N−N transfer time (test spectrum 2). Use the same parameters as for test spectrum 1, but set the length of the N−N transfer period ($2\varDelta$) to 30 ms ($d27 = 15$ ms), which is a compromise between good magnetization transfer via $^{h2}J_{NN}$ to the ^{15}N acceptor and sufficient detection of the ^{15}N donor for small- to medium-sized ($\leq \approx 80$ nucleotide) nucleic acids. The purpose of this spectrum is to estimate how long it will take to acquire the final H-bond spectrum. Record this spectrum also with a small number of scans. If the signal-to-noise ratio of test spectrum 1 is similar to our example, NS = 4 may be sufficient.

12| Check the signal-to-noise ratio of test spectrum 2. Process the data of test spectrum 2 in exactly the same manner as in test spectrum 1. Determine the signal-to-noise ratio of the relevant imino resonances. In our example, the S/N_{top} is in the range of 5–20:1 for the imino resonances with NS = 4 after 17 min total experimental time. Many of the cross-peaks to the ^{15}N acceptor nuclei (^{15}N frequencies between 195 and 225 ppm; **Fig. 3**) are also already visible with an S/N_{top} of up to 5:1.

13| Calculate the necessary experimental time (number of scans) required for the quantitative 2D HNN-COSY. Even if the 'cross-peaks' to the H-bond acceptor ^{15}N nuclei are not yet visible in test spectrum 2, the expected cross-peak intensity can be calculated from the intensity of the imino donor ('diagonal') peaks via the relation $I_c/I_d = -\tan^2(2\pi^{h2}J_{NN}\varDelta)$. Thus, if we want to detect this cross-peak with a specified signal-to-noise ratio, the required time for the experiment can be calculated. For example, in our experiment on the PSTVd RNA, certain imino resonances had an S/N_{top} of 5:1 after 17 min (NS = 4). If we assume a $^{h2}J_{NN}$ coupling constant of 5 Hz (which is at the low end of the expected range), the S/N_{top} of the cross-peak should be $5\times \tan^2(2\pi\times 5$ Hz$\times 15$ ms$) \approx 1.3$ after 17 min (neglecting slight differences in the line width of cross- and diagonal peaks). Thus, if we want to observe this cross-peak at an S/N_{top} of 9, the number of scans has to be increased by a factor of $(9/1.3)^2 \approx 48$, as the signal-to-noise ratio increases with the square root of the number of scans. This would lead to a required number of scans of 192, and a total measuring time of 13.6 h for the final experiment. In brief, $NS_{required} = NS_{test} \tan^{-4}(2\pi^{h2}J_{NN}\varDelta)((S/N_{cross,desired})/(S/N_{diagonal,test}))^2$.

Data acquisition
14| Record the final 2D quantitative J_{NN} HNN-COSY experiment using the necessary number of scans determined in Step 13 and all other parameters equal to those in test spectrum 2.

Data processing
15| The data of the experiment are recorded as a phase-sensitive 2D data set, and standard processing is applied. Any suitable software may be used. As an example, we present the processing with the NMRPipe suite. The processing is divided into the following steps (see **Supplementary Note 2**). (a) The raw NMR data are converted into NMRPipe format using the {bruk, var, delta}2pipe programs, depending on the spectrometer data format. (b) A fourth-order polynomial baseline is subtracted in the t_2 time domain for better water suppression. (c) A 60°-shifted squared sine-bell function is applied for filtering in t_2. This maximizes resolution and signal-to-noise ratio under acquisition conditions where $t_{2,max} \approx T_2(^1H^N)$ (see TROUBLESHOOTING). (d) The first data point in t_2 is multiplied by 0.5 to avoid baseline offset effects[41,42]. This is necessary, as the directly detected dimension (t_2) is sampled without any initial delay. (e) Data in the t_2 dimension are zero-filled at least once and Fourier transformed. Proper zero- and first-order phase correction is applied after the Fourier transformation with phases determined empirically from the processed spectrum, for example, by interactive phase manipulation and visualization in the NMRDraw program. (f) After transposition, the sign of the imaginary part of the t_1 domain data is inverted to account for a reversal of the usual States-TPPI phase-sensitive detection in the t_1 dimension, which stems from the simultaneous incrementation of phases ϕ_1 and ϕ_2 in the HNN-COSY experiment (**Fig. 2**). (g) A 60°-shifted sine-bell function is applied for filtering in t_1. This maximizes the resolution and signal-to-noise ratio under acquisition conditions where $t_{1,max} \approx T_2(^{15}N)$ (see TROUBLESHOOTING). (h) The first data point in t_1 is left unchanged (multiplied by 1), as the initial delay of t_1 was set to half a dwell time in the pulse program. This leads to zero- and first-order phase corrections in the ^{15}N dimension of $-90°$ and $180°$, respectively[41]. (i) Data in the t_1 dimension are zero-filled at least once and Fourier transformed. Proper zero- ($-90°$) and first-order ($180°$) phase corrections are applied after the Fourier transformation.
? TROUBLESHOOTING

Data analysis
16| Inspect the 2D spectrum. N−H$\cdots$N H-bonds should be apparent from ^{15}N donor (140–165 ppm) and acceptor (195–225 ppm) cross-peaks, which are aligned according to the common frequency of the imino $^1H^N$ nucleus. This may provide valuable information on the H-bond network of the nucleic acid.

17| For a more quantitative analysis, determine the intensities of the cross- and diagonal peaks by peak integration.
▲ CRITICAL STEP The quantification of the coupling constants from the intensity ratio of cross- and diagonal peaks is made difficult by the different line widths of the ^{15}N donor and acceptor resonances (**Fig. 3**), which are the result of the different relaxation mechanisms for both nuclei. Whereas the width of donor ^{15}N resonances is narrowed by the partial cancellation of the ^1H–$^{15}N_d$ dipole and $^{15}N_d$ chemical shift anisotropy (CSA) relaxation in the TROSY scheme, the very large (300–400 ppm) CSA[43] of the ^{15}N acceptor nuclei in the nucleic bases dominates the relaxation for the acceptor resonances. Due to these differing line widths, it is not sufficient to approximate the intensity ratios of the cross- and diagonal peaks by the amplitude ratios. Instead, proper peak integration

should be used to obtain intensities. As a more precise alternative, the intensities, that is, the amplitudes of the oscillations in the ^{15}N time domain, can be obtained from time domain fitting routines (e.g., nlinLS contained in NMRPipe)[7,31].

18| Calculate the magnitude of the $^{h2}J_{NN}$ couplings using the formula $|^{h2}J_{NN}| = atan[(-I_{Na}/I_{Nd})^{1/2}]/(2\pi\Delta)$, where 2Δ is the COSY transfer time.

● **TIMING**
Following nucleic acid sample preparation and resonance assignments, the entire protocol from Steps 1 to 18 is expected to take no more than 24 h. Longer experimental times may be required for low sample concentrations or higher molecular weight nucleic acids. Time may be saved by using a cryogenic probe rather than a conventional probe.

? TROUBLESHOOTING
Step 7 (finite ^{15}N radio frequency strength)
The large frequency separation between ^{15}N donor and acceptor resonances combined with the finite available strength of ^{15}N radio frequency pulses presents a certain limitation for the simple HNN-COSY experiment depicted in **Figure 2**. For example, in Watson–Crick and Hoogsteen H-bonds, the chemical shift separations of the ^{15}N imino donor and aromatic acceptor resonances are in the range of 50–70 ppm corresponding to 3–4 kHz on a 600 MHz instrument. Typically available ^{15}N radio frequency strengths are slightly larger than 6 kHz (the largest available radio frequency strength is calculated as $1/(4P_{90})$, where P_{90} is the length of the shortest possible 90° pulse). Thus, the simple HNN-COSY is still reasonably efficient when the ^{15}N carrier is placed at the midpoint between donor and acceptor resonances. However, for larger donor–acceptor frequency differences, for example, between amino donor (∼85 ppm) and aromatic acceptor [15]N (190–220 ppm) nuclei or for smaller available radio frequency strengths, band-selective ^{15}N pulses should be used[22,23].

Step 15 (peak overlap)
Peak overlap in crowded regions of the 2D HNN-COSY spectra can impede accurate calculation of the $^{h2}J_{NN}$ couplings. For samples with high signal-to-noise ratios, it may be desirable to increase the maximal t_1 evolution time beyond the T_2 relaxation time of donor and acceptor ^{15}N nuclei at the expense of signal intensity. During processing, peak overlap may also be alleviated by using apodization functions, which provide resolution enhancement (again at the cost of lower signal intensity). For example, in **Supplementary Note 2**, a 60°-shifted sine-bell function is employed in both t_1 and t_2 dimensions. Shifting the sine-bell to 45° or lower (setting "-off" flag to 0.25 or smaller values) will increase resolution.

Larger nucleic acids
The sensitivity for detection of the acceptor nuclei is highest when the total N···N magnetization transfer period (2Δ) does not exceed the transverse relaxation time T_2 of the ^{15}N donor nucleus. For very large nucleic acids (>25 kDa), or at low temperatures, the T_2 may become shorter than the value of 30 ms recommended in Step 11 (**Box 1**). In these cases ($\pi T_2 {}^{h2}J_{NN} \ll 1$), 2Δ should be set to the [15]N T_2 value to achieve optimal transfer.

Unobservable ^{1}H nuclei
For situations where the hydrogen nucleus in the H-bond is not observable due to exchange, or when the experiment is carried out using a nucleic acid sample dissolved in D_2O, additional modifications of the HNN-COSY have been proposed[44–47]. In certain cases, the HNN-COSY can be started and detected on a carbon-bound proton in close vicinity of the acceptor. This is possible for adenosine and guanosine [15]N7 acceptor nuclei as well as for adenosine [15]N1 acceptors, which can be connected to the H8 or H2 proton by means of the covalent $^2J_{H8N7}$ or $^2J_{H2N1}$ couplings of approximately 11–15 Hz, respectively.

ANTICIPATED RESULTS
The result of the quantitative HNN-COSY experiment (total experimental time 13 h) carried out on a 1.5 mM sample of the 69-nucleotide PSTVd T1 RNA domain (mutant U18C/A344G) is shown in **Figure 3**. Diagonal peaks (blue) are observed for almost all imino groups at their respective [1]H[N] (10–14.5 ppm) and [15]N (guanosine N1, 140–150 ppm; uridine N3, 157–165 ppm) frequencies. For the N–H···N H-bonded imino groups, cross-peaks of opposite sign (**Fig. 3**; bottom, red) are also observed, which correspond to [1]H3(U)···[15]N1(A) (∼223 ppm) and [1]H1(G)···[15]N3(C) (∼197 ppm) correlations according to Watson–Crick base pairing (**Fig. 1**). Individual peaks were assigned from additional NOESY (nuclear Overhauser enhancement spectroscopy) data[30]. For most uridines, weak-intensity cross-peaks are also observed to ^{15}N nuclei at resonance frequencies of about 145 ppm (**Fig. 3**, top). These are the results of intrabase $^2J_{NN}$ couplings between uridine N3 and N1 nuclei, which are ≈2 Hz in size[7]. For certain uridine and guanosine imino groups with [1]H[N] frequencies between 10 and 12 ppm (**Fig. 3**, right), for example, U7, no H-bond cross-peaks are observed, as they are not involved in H-bonds of the N–H···N type. Quantification of the $^{h2}J_{NN}$ couplings for the Watson–Crick U–A and G–C base pairs using the time domain fitting program nlinLS, contained within the NMRPipe package[31], yielded values between 6 and 8 Hz in size. This corresponds to the usual size of $^{h2}J_{NN}$ couplings in Watson–Crick base pairs. Smaller couplings are sometimes observed due to fraying at the ends of helical stems or due to nonlinear H-bond geometries[19]. For other base pair types, for example, charged Hoogsteen C+•G, the coupling size may increase up to ∼10 Hz (see ref. 48).

PROTOCOL

1. Jeffrey, G.A. & Saenger, W. *Hydrogen Bonding in Biological Structures* (Springer, New York, 1991).
2. Shu, F., Ramakrishnan, V. & Schoenborn, B.P. Enhanced visibility of hydrogen atoms by neutron crystallography on fully deuterated myoglobin. *Proc. Natl. Acad. Sci. USA* **97**, 3872–3877 (2000).
3. Addlagatta, A., Krzywda, S., Czapinska, H., Otlewski, J. & Jaskolski, M. Ultrahigh-resolution structure of a BPTI mutant. *Acta Crystallogr. D* **57**, 649–663 (2001).
4. Jelsch, C. *et al.* Accurate protein crystallography at ultra-high resolution: valence electron distribution in crambin. *Proc. Natl. Acad. Sci. USA* **97**, 3171–3176 (2000).
5. Gerwert, K. Molecular reaction mechanisms of proteins monitored by time-resolved FTIR-spectroscopy. *Biol Chem* **380**, 931–935 (1999).
6. Becker, E.D. Hydrogen bonding. in *Encyclopedia of Nuclear Magnetic Resonance,* Vol. 4 (eds. Grant, D.M. & Harris, R.K.) 2409–2415 (John Wiley, New York, 1996).
7. Dingley, A.J. & Grzesiek, S. Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^2J_{NN}$ couplings. *J. Am. Chem. Soc.* **120**, 8293–8297 (1998).
8. Pervushin, K. *et al.* NMR scalar couplings across Watson–Crick base pair hydrogen bonds in DNA observed by transverse relaxation-optimized spectroscopy. *Proc. Natl. Acad. Sci. USA* **95**, 14147–14151 (1998).
9. Cordier, F. & Grzesiek, S. Direct observation of hydrogen bonds in proteins by inter-residue $^{3h}J_{NC'}$ scalar couplings. *J. Am. Chem. Soc.* **121**, 1601–1602 (1999).
10. Cornilescu, G., Hu, J.-S. & Bax, A. Identification of the hydrogen bonding network in a protein by scalar couplings. *J. Am. Chem. Soc.* **121**, 2949–2950 (1999).
11. Shenderovich, I.G. *et al.* Nuclear magnetic resonance of hydrogen bonded clusters between F- and (HF)n: experiment and theory. *Ber. Bunsen-Ges. Phys. Chem.* **102**, 422–428 (1998).
12. Golubev, N.S., Shenderovich, I.G., Smirnov, S.N., Denisov, G.S. & Limbach, H.-H. Nuclear scalar spin–spin coupling reveals novel properties of low-barrier hydrogen bonds in a polar environment. *Chem. Eur. J.* **5**, 492–497 (1999).
13. Blake, P.R. *et al.* Quantitative measurement of small through-hydrogen-bond and 'through-space' ^1H-^{113}Cd and ^1H-^{199}Hg J couplings in metal-substituted rubredoxin from *Pyrococcus furiosus*. *J. Biomol. NMR* **2**, 527–533 (1992).
14. Crabtree, R.H., Siegbahn, P.E.M., Eisenstein, O., Rheingold, A.L. & Koetzle, T.F. A new intermolecular interaction: unconventional hydrogen bonds with element-hydride bonds as proton acceptor. *Acc. Chem. Res.* **29**, 348–354 (1996).
15. Cordier, F., Nisius, L., Dingley, A.J. & Grzesiek, S. Direct detection of N—H···O=C hydrogen bonds in biomolecules by NMR spectroscopy. *Nat. Protoc.* **3**, 235–241.
16. Dingley, A.J., Cordier, F. & Grzesiek, S. An introduction to hydrogen bond scalar couplings. *Concepts Magn. Reson.* **13**, 103–127 (2001).
17. Grzesiek, S., Cordier, F. & Dingley, A.J. Scalar couplings across hydrogen bonds. *Methods Enzymol.* **338**, 111–133 (2001).
18. Grzesiek, S., Cordier, F. & Dingley, A.J. Hydrogen bond scalar couplings—a new tool in biomolecular NMR. in *Biological Magnetic Resonance,* Vol. 20 (eds. Krishna, N.R. & Berliner, L.J.) 255–283 (Kluwer Academic/Plenum, New York, 2003).
19. Grzesiek, S., Cordier, F., Jaravine, V.A. & Barfield, M. Insights into biomolecular hydrogen bonds from hydrogen bond scalar couplings. *Prog. Nucl. Magn. Reson. Spectrosc.* **45**, 275–300 (2004).
20. Barfield, M. Structural dependencies of inter-residue scalar coupling $^{3h}J_{NC}$, and donor H-1 chemical shifts in the hydrogen bonding regions of proteins. *J. Am. Chem. Soc.* **124**, 4158–4168 (2002).
21. Cornilescu, G. *et al.* Correlation between $^{3h}J_{NC'}$ and hydrogen bond length in proteins. *J. Am. Chem. Soc.* **121**, 6275–6279 (1999).
22. Majumdar, A., Kettani, A. & Skripkin, E. Observation and measurement of internucleotide $^2J_{NN}$ coupling constants between ^{15}N nuclei with widely separated chemical shifts. *J. Biomol. NMR* **14**, 67–70 (1999).
23. Dingley, A.J., Masse, J.E., Feigon, J. & Grzesiek, S. Characterization of the hydrogen bond network in guanosine quartets by internucleotide $^{3h}J_{NC'}$ and $^{2h}J_{NN}$ scalar couplings. *J. Biomol. NMR* **16**, 279–289 (2000).
24. Hennig, M. & Geierstanger, B.H. Direct detection of a histidine–histidine side-chain hydrogen bond important for folding of apomyoglobin. *J. Am. Chem. Soc.* **121**, 5123–5126 (1999).
25. Majumdar, A. & Patel, D.J. Identifying hydrogen bond alignments in multistranded DNA architectures by NMR. *Acc. Chem. Res.* **35**, 1–11 (2002).
26. Al-Hashimi, H.M., Gorin, A., Majumdar, A., Gosser, Y. & Patel, D.J. Towards structural genomics of RNA: rapid NMR resonance assignment and simultaneous RNA tertiary structure determination using residual dipolar couplings. *J. Mol. Biol.* **318**, 637–649 (2002).
27. Kojima, C., Ono, A. & Kainosho, M. Studies of physicochemical properties of N—H···N hydrogen bonds in DNA, using selective ^{15}N-labeling and direct ^{15}N 1D NMR. *J. Biomol. NMR* **18**, 269–277 (2000).
28. Bytchenkoff, D., Chiarparin, E., Früh, D., Rüdisser, S. & Bodenhausen, G. Temperature dependence of internucleotide nitrogen–nitrogen scalar couplings in RNA. *Magn. Reson. Chem.* **40**, 377–379 (2002).
29. Dingley, A.J., Peterson, R.D., Grzesiek, S. & Feigon, J. Characterization of the cation and temperature dependence of DNA quadruplex hydrogen bond properties using high-resolution NMR. *J. Am. Chem. Soc.* **127**, 14466–72 (2005).
30. Dingley, A.J., Steger, G., Esters, B., Riesner, D. & Grzesiek, S. Structural characterization of the 69-nucleotide potato spindle tuber viroid left-terminal domain by NMR and thermodynamic analysis. *J. Mol. Biol.* **334**, 751–767 (2003).
31. Delaglio, F. *et al.* NMRPipe—a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
32. Barfield, M., Dingley, A.J., Feigon, J. & Grzesiek, S. A DFT study of the inter-residue dependencies of scalar J- coupling and magnetic shielding in the hydrogen-bonding regions of a DNA triplex. *J. Am. Chem. Soc.* **123**, 4014–4022 (2001).
33. Schwalbe, H., Schmidt, P. & Griesinger, C. Coupling constants determined by ECOSY. in *Encyclopedia of Nuclear Magnetic Resonance,* Vol. 3 (eds. Grant, D.M. & Harris, R.K.) 1473–1489 (John Wiley, New York, 1996).
34. Batey, R.T., Inada, M., Kujawinski, E., Puglisi, J.D. & Williamson, J.R. Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy of RNA. *Nucleic Acids Res.* **20**, 4515–4523 (1992).
35. Batey, R.T., Battiste, J.L. & Williamson, J.R. Preparation of isotopically enriched RNAs for heteronuclear NMR. *Methods Enzymol.* **261**, 300–322 (1995).
36. Nikonowicz, E.P. *et al.* Preparation of ^{13}C and ^{15}N labelled RNAs for heteronuclear multi-dimensional NMR studies. *Nucleic Acids Res.* **20**, 4507–4513 (1992).
37. Louis, J.M., Martin, R.G., Clore, G.M. & Gronenborn, A.M. Preparation of uniformly isotope-labeled DNA oligonucleotides for NMR spectroscopy. *J. Biol. Chem.* **273**, 2374–2378 (1998).
38. Rene, B., Masliah, G., Zargarian, L., Mauffret, O. & Fermandjian, S. General method of preparation of uniformly ^{13}C, ^{15}N-labeled DNA fragments for NMR analysis of DNA structures. *J. Biomol. NMR* **36**, 137–146 (2006).
39. Messerle, B.A., Wider, G., Otting, G., Weber, C. & Wuthrich, K. Solvent suppression using a spin lock in 2D and 3D NMR-spectroscopy with H$_2$O solutions. *J. Magn. Reson.* **85**, 608–613 (1989).
40. Sklenar, V. & Bax, A. Spin-echo water suppression for the generation of pure-phase two-dimensional NMR-spectra. *J. Magn. Reson.* **74**, 469–479 (1987).
41. Zhu, G., Torchia, D. & Bax, A. Discrete Fourier transformation of NMR signals. The relationship between sampling delay time and spectral baseline. *J. Magn. Reson. A* **105**, 219–222 (1993).
42. Bax, A., Ikura, M., Kay, L.E. & Zhu, G. Removal of F1 baseline distortion and optimization of folding in multidimensional NMR spectra. *J. Magn. Reson.* **91**, 174–178 (1991).
43. Hu, J., Facelli, J., Alderman, D., Pugmire, R. & Grant, D. ^{15}N chemical shift tensors in nucleic acid bases. *J. Am. Chem. Soc.* **120**, 9863–9869 (1998).
44. Majumdar, A., Kettani, A., Skripkin, E. & Patel, D.J. Pulse sequences for detection of NH$_2$···N hydrogen bonds in sheared G·A mismatches via remote, non-exchangeable protons. *J. Biomol. NMR* **19**, 103–113 (2001).
45. Majumdar, A., Kettani, A., Skripkin, E. & Patel, D.J. Observation of internucleotide N—H···N hydrogen bonds in the absence of directly detectable protons. *J. Biomol. NMR* **15**, 207–211 (1999).
46. Hennig, M. & Williamson, J.R. Detection of N—H···N hydrogen bonding in RNA via scalar couplings in the absence of observable imino proton resonances. *Nucleic Acids Res.* **28**, 1585–1593 (2000).
47. Luy, B. & Marino, J.P. Direct evidence for Watson–Crick base pairs in a dynamic region of RNA structure. *J. Am. Chem. Soc.* **122**, 8095–8096 (2000).
48. Dingley, A.J. *et al.* Internucleotide scalar couplings across hydrogen bonds in Watson–Crick and Hoogsteen base pairs of a DNA triplex. *J. Am. Chem. Soc.* **121**, 6019–6027 (1999).

2.3 Direct detection of N-H···O=C hydrogen bonds in biomolecules by NMR spectroscopy

Original Publication

Cordier, F., Nisius, L., Dingley, A.J., and Grzesiek, S.

Direct detection of N-H···O=C hydrogen bonds in biomolecules by NMR spectroscopy

Nature Protocols 3 (2008) 235-241.

Direct detection of N—H···O=C hydrogen bonds in biomolecules by NMR spectroscopy

Florence Cordier[1,4], Lydia Nisius[2,4], Andrew J Dingley[3] & Stephan Grzesiek[2]

[1]Department of Structural Biology and Chemistry, Institut Pasteur, 28 Rue du Docteur Roux, 75015 Paris, France. [2]Department of Structural Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland. [3]Department of Chemistry and School of Biological Sciences, The University of Auckland, Private Bag 92019, Auckland, New Zealand. [4]These authors contributed equally to this work. Correspondence should be addressed to S.G. (stephan.grzesiek@unibas.ch).

A nuclear magnetic resonance (NMR) experiment is described for the direct detection of N-H···O=C hydrogen bonds (H-bonds) in ^{15}N and ^{13}C isotope-labeled biomolecules. This quantitative 'long-range' HNCO-COSY (correlation spectroscopy) experiment detects and quantifies electron-mediated scalar couplings across the H-bond (H-bond scalar couplings), which connect the magnetically active ^{15}N and ^{13}C nuclei on both sides of the H-bond. Detectable H-bonds comprise the canonical backbone H-bonds in proteins as well as other H-bonds in proteins and nucleic acids with N–H donors and O=C (carbonylic or carboxylic) acceptors. Unlike other NMR observables, which provide only indirect evidence of the presence of H-bonds, the H-bond scalar couplings identify all partners of the H-bond, the donor, the donor proton and the acceptor, in a single experiment. The size of the scalar couplings can be related to H-bond geometries. The time required to detect the N—H···O=C H-bonds in small proteins ($\leq \approx 10$ kDa) is typically on the order of 1 d at millimolar concentrations, whereas H-bond detection for larger proteins ($\leq \approx 30$ kDa) may be possible within several days depending on concentration, isotope composition, magnetic field strength and molecular weight. The proteins ubiquitin (8.6 kDa), dimeric RANTES (2 × 8.5 kDa) and MAP30 (30 kDa) are used as examples to illustrate this procedure.

INTRODUCTION

H-bonds are key features of biomacromolecules[1]. NMR techniques developed in recent years make it possible to observe and quantify H-bonds in the solution state by COSY (correlation spectroscopy) experiments. The techniques are based on H-bond scalar couplings (HBCs), which connect magnetically active nuclei on both sides of the hydrogen bridge via the magnetic polarization of its electron cloud. Observable HBCs comprise H-bonds of regular secondary structure elements in nucleic acids[2,3] and proteins[4,5] as well as in small chemical compounds[6–9]. The HBCs can be used to identify donor and acceptor groups in individual H-bonds from COSY-type experiments. In favorable cases, complete H-bond networks in biomacromolecules can be established, thereby providing valuable data for structure determination. The size of the HBCs is determined by the overlap of H-bond donor and acceptor electronic orbitals[10] and thus provides a sensitive measure of H-bond geometry, that is, HBCs depend exponentially on donor–acceptor distances[11] as well as on certain H-bond angles[12]. Owing to the requirement that HBCs connect NMR-observable nuclei, their observation in biomolecules has been limited to HBCs between ^{15}N, ^{13}C, ^{1}H or ^{31}P nuclei, but excludes oxygen, for which no suitable isotope exists. Experimental procedures for their detection, theoretical descriptions and applications of HBCs have been reviewed extensively[12–15].

In nucleic acids, strong two-bond ^{15}N—^{15}N HBCs ($^{h2}J_{NN}$) can be detected across N—H···N H-bonds between ^{15}N donor and acceptor nuclei in Watson–Crick base pairs and many other noncanonical base pairs. (The symbol $^{hn}J_{AB}$ is used for H-bond scalar couplings between nuclei A and B to emphasize that one of the n bonds connecting the two nuclei in the chemical structure is actually an H-bond.) An accompanying protocol[16] describes their detection by the HNN-COSY experiment and gives a more detailed introduction. As oxygen is not observable, protein backbone N—H···O=C H-bonds cannot be detected via the carbonyl oxygen. However, N—H···O=C H-bonds induce detectable HBCs

(**Fig. 1**) between the amide ^{15}N donor and the next nucleus on the other side of the H-bond, the ^{13}C′ carbonyl acceptor[4,5]. These three-bond ^{15}N—^{13}C′ HBCs ($^{h3}J_{NC'}$) can be detected and quantified by the 'long-range' HNCO experiment, which is a modification of the standard HNCO used for protein assignments. In a two-dimensional version, the 'long-range' H(N)CO yields the resonance frequencies of the amide proton ^{1}H^{N} and the carbonyl ^{13}C′ nucleus, whereas in a three-dimensional HNCO version, the frequency of the amide ^{15}N donor nucleus is also detected such that all partners of the H-bond are visible.
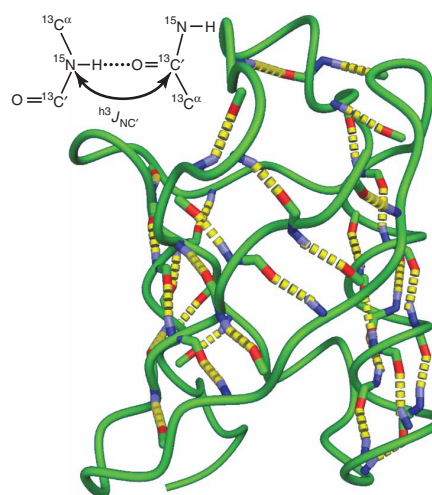

Figure 1 | Backbone H-bond network of ubiquitin detected by $^{h3}J_{NC'}$ correlations[4]. Nitrogen donors, amide protons and oxygen acceptors are shown in dark blue, light blue and red, respectively. The insert illustrates the definition of the scalar $^{h3}J_{NC'}$ coupling via the N—H···O=C H-bond.

The principal application of the 'long-range' HNCO is the unequivocal establishment of H-bonding for structural analysis. This primarily applies to protein backbone N−H···O=C H-bonds. However, H-bonds involving nucleic acid base[17] or protein side-chain N−H or O=C groups[11,18,19] can also be detected by this experiment. Owing to the strong dependence on H-bond geometries, the size of the coupling constants provides a sensitive measure for variations in H-bond lengths and angles[4,10,11]. Such variations can be detected as a result of ligand binding[20] and protein folding[21], in response to changes of pressure[22], temperature[23] and amino-acid variations[24,25] or other changes of physicochemical conditions. Applications of this type have been reviewed in ref. 12.

As for any solution NMR experiment on biomolecules, the detection of H-bonds is limited by sensitivity and molecular weight. $^{h3}J_{NC'}$ couplings have typical sizes in the range of −0.1 to −0.9 Hz and are considerably smaller than $^{h2}J_{NN}$ couplings (5−11 Hz) in nucleic acids. Thus, detection of N−H···O=C H-bonds is significantly more difficult than that of N−H···N H-bonds, especially for larger proteins. Nevertheless, small (10–15 kDa) proteins in $^{13}C/^{15}N$-labeled form are easily accessible, provided they can be measured at concentrations of $\geq \sim 0.5$ mM. Additional ^2H labeling and the use of TROSY techniques allows detection of $^{h3}J_{NC'}$ HBCs even for medium-sized proteins of up to about 30 kDa (see ref. 26).

Organization of the protocol

The following protocol gives a step-by-step setup and evaluation procedure for the quantitative 'long-range' HNCO experiment to detect and quantify $^{h3}J_{NC'}$ couplings in N−H···O=C H-bonds between ^{15}N amide donor and $^{13}C'$ carbonyl acceptor nuclei (**Fig. 1**). Example results are included from samples of ubiquitin (8.6 kDa), dimeric RANTES (2 × 8.5 kDa) and MAP30 (30 kDa), recorded on Bruker 800 and 750 MHz spectrometers using the xwinnmr 3.5 acquisition software. Data processing and analysis were performed with the NMRPipe suite of programs[27]. This software and further information are available at http://spin.niddk.nih.gov/NMRPipe.

Experimental design

The 'long-range' HNCO experiment[4,5] (**Fig. 2**) used to detect the $^{h3}J_{NC'}$ couplings is a variant of the standard HNCO experiment for protein backbone assignments with the N−C′ INEPT transfer time set to larger values. In brief, magnetization is transferred from the amide proton (H) to the H-bond donor ^{15}N nucleus during the first INEPT period (time points a to b) via the covalent $^1J_{NH}$ coupling. During the subsequent N−C′ INEPT period (time points b to c), both $^{h3}J_{NC'}$ and the large covalent $^1J_{NC'}$ ($\approx$ −15 Hz) couplings are active for a time $2T − 2\varepsilon$, where the length of the delay ε is determined by the position of the $^{13}C'$ 180° pulse (**Fig. 2**). Optimal transfer from the ^{15}N nucleus via the H-bond to the $^{13}C'$ nucleus of the acceptor is achieved when the one-bond coupling $^1J_{NC'}$ is approximately refocused, that is, $2T − 2\varepsilon = n/^1J_{NC'}$, and when the total transfer time is set to a value that is close to the T_2 relaxation time of the ^{15}N nucleus. In practice, a total transfer time of about 133 ms ($n = 2$) works well for small proteins, which is realized by setting the delays $T = 66.5$ ms and $\varepsilon = 4$ µs ≈ 0. During the following t_1-evolution period (time points d to e), the frequencies of the $^{13}C'$ are encoded, and magnetization is then transferred back to the ^{15}N nucleus via the C′−N reverse-INEPT step (time points f to g, total active transfer time $2T − 2\varepsilon$). A final TROSY (transverse relaxation optimized spectroscopy) transfer (time points h to i) converts the slowly relaxing component of the {^1H}–^{15}N doublet into the slowly relaxing component of the {^{15}N}–^1H doublet, which is detected in the receiver during the t_3 period. The resulting spectrum of the 2D H(N)CO shows 'cross-peaks' at frequency positions ($\omega_{C'j}$, $\omega_{Hi} + \pi^1J_{NH}$) (**Fig. 3**). The intensity of these cross-peaks (I_{cross}) is proportional to $\cos^2(\pi^1J_{NC'}2T)\sin^2(\pi^{h3}J_{NC'}2T)$.

To determine the size of the $^{h3}J_{NC'}$ coupling by quantitative J-correlation, a reference experiment is recorded using the same total lengths $2T$ for the N−C′ INEPT and C′−N reverse-INEPT periods, but with the two $^{13}C'$ 180° pulses shifted by $\varepsilon = 1/(4^1J_{NC'}) \sim 16.5$ ms relative to the ^{15}N 180° pulses. This reduces the effective time for ^{15}N−$^{13}C'$ de- and refocusing to $(2T − 2\varepsilon) = (n/^1J_{NC'} − 1/(2^1J_{NC'}))$ while keeping the relaxation losses identical to the long-range experiment. In this experiment, the nitrogen-to-carbonyl magnetization transfer occurs primarily via the large $^1J_{NC'}$ couplings. The resulting reference spectrum shows resonances at frequency positions ($\omega_{C'i-1}$, $\omega_{Hi} + \pi^1J_{NH}$) with intensities proportional to $\sin^2(\pi^1J_{NC'}2(T − \varepsilon))\cos^2(\pi^{h3}J_{NC'}2(T − \varepsilon))$. Thus, the intensity ratio of cross- and

Figure 2 | 2D long-range H(N)CO TROSY pulse sequence. Color-coding highlights the various magnetization transfer and chemical shift evolution periods as presented in the Bruker pulse sequence code (**Supplementary Box 2**). Pulse durations or field strengths are given for an 800 MHz spectrometer. For different magnetic field strength (X MHz) spectrometers, the radio frequency field strengths should be multiplied by X/800. Narrow and wide pulses correspond to flip angles of 90° and 180°, respectively. Carrier positions are ^1H_2O (^1H), 116.5 ppm (^{15}N), 177 ppm ($^{13}C'$) and 56 ppm ($^{13}C^\alpha$). All regular ^1H and ^{15}N pulses are applied at radio-frequency field strengths of 25 and 6.25 kHz, respectively. Small narrow ^1H pulses are weak water flip-back 90° pulses with a duration of 1 ms. The shaped (90°) pulse has the profile of the center lobe of a sinx/x function with a duration of ∼2 ms. All the $^{13}C'$ pulses have the shape of the center lobe of a sinx/x function and durations of 110 µs (220 µs) for the narrow 90° (wide 180°) pulses. GARP $^{13}C'$ decoupling sequence is used at a field strength of 1.25 kHz. $^{13}C^\alpha$ 180° pulses are rectangular pulses applied at a field strength of 14.0 kHz. Delays: $\delta = 2.7$ ms; T, ε (see text). Unless otherwise indicated, the phases of all pulses are along the x axis. Phase cycling: $\phi1 = 4(x)$, $4(−x)$; $\phi2 = x$, $−x$; $\phi3 = y$, y, $−x$, $−x$; receiver: x, $−x$, y, $−y$, $−x$, x, $−y$, y. Quadrature detection in the t_1 dimension is achieved by incrementing $\phi2$ in the States-TPPI manner. For the 3D version of the experiment, sensitivity-enhanced quadrature detection in the t_2 dimension is achieved by applying $\phi3$ either as y, y,$−x$,$−x$ or as y, y, x, x for p- and n-type detection. Gradients (sine-bell shaped, absolute amplitude at the center of 30 G cm$^{−1}$) with durations: $G_{1,2,3,4,5,6,7,8} = 10$, 3, 2.7, 2.5, 3.75, 0.888, 0.7, 0.4 ms.
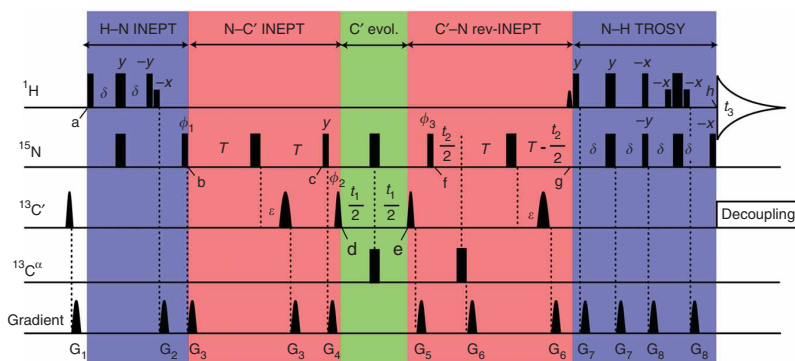
Figure 3 | Results of the long-range HNCO. (**a**) Selected region of the 2D long-range H(N)CO TROSY spectrum recorded on a 2 mM sample of $^{15}N/^{13}C$ RANTES, pH 3.8, 310 K, 800 MHz spectrometer with cryogenic probe. The data matrix consisted of $80*(^{13}C', t_1) \times 1,024*(^{1}H, t_3)$ data points (where $n*$ refers to complex points) with acquisition times of 40 ($^{13}C'$) and 85 ms (^{1}H). The total experimental time was 46.5 h. Cross-peaks marked as $Res_i \cdots Res_j$ are due to $^{h3}J_{NiC'j}$ HBCs between the ^{15}N nucleus of residue i and $^{13}C'$ nucleus of residue j (see inset for definition). Intramolecular and intermolecular $^{h3}J_{NiC'j}$ correlations are labeled in red and orange, respectively. Residue names marked in blue denote incompletely suppressed sequential correlations between the ^{15}N nucleus of residue i and $^{13}C'$

nucleus of residue i − 1. Residues marked in green correspond to intraresidue two-bond $^{15}N_i$–$^{13}C'_i$ correlations. (**b**) Strip plot of the 3D long-range HNCO TROSY spectrum recorded previously[26] on a 0.7 mM sample of $^{15}N/^{13}C/^{2}H$ MAP30, pH 5.5, 313 K, 750 MHz spectrometer equipped with a room-temperature probe. The total experimental time was 91 h. The data matrix consisted of $32*(^{13}C', t_1) \times 99*(^{15}N, t_2) \times 768*(^{1}H, t_3)$ data points with acquisition times of 65 (^{15}N), 14 ($^{13}C'$) and 79 ms (^{1}H). The strips are labeled by the residue of the N—H moiety. H-bond (red), sequential (blue) and intraresidue (green) correlations to carbonyls are marked with assignment information.
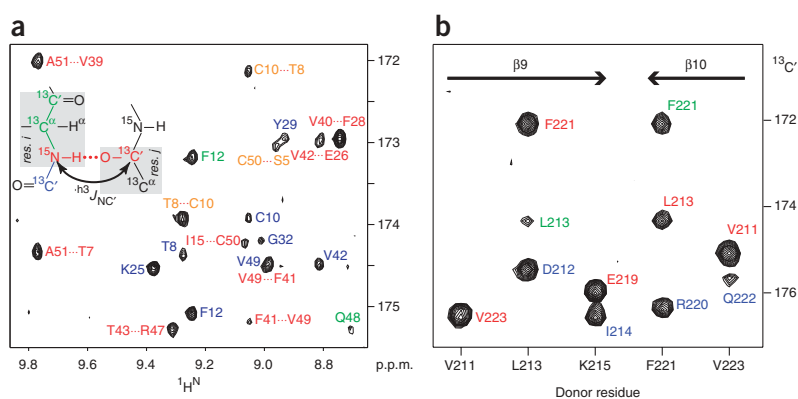
reference peaks is given by $I_{cross}/I_{ref} = \cos^2(\pi^1 J_{NC'}2T)\sin^2(\pi^{h3}J_{NC'}2T)/[\sin^2(\pi^1 J_{NC'}(2T − 2\varepsilon))\cos^2(\pi^{h3}J_{NC'}(2T − 2\varepsilon))]$ and the size of the coupling can be determined by a numerical inversion of this implicit equation with the known values of $^1J_{NC'}$. As an alternative, the coupling can be estimated to good approximation as $|^{h3}J_{NC'}| \approx (I_{cross}/I_{ref})^{1/2}/(2\pi T)$, since $I_{cross}/I_{ref} \approx \sin^2(\pi^{h3}J_{NC'}2T) \approx (\pi^{h3}J_{NC'}2T)^2$ for $^1J_{NC'}$ close to 15 Hz and $|\pi^{h3}J_{NC'}2T| \ll 1$.

A 3D version of the long-range HNCO experiment can easily be implemented at no additional cost in signal intensity by inserting a ^{15}N constant time evolution (t_2) period within the C'−N reverse-INEPT period[26]. For the 2D H(N)CO version of the experiment described in the following protocol, the incrementation in t_2 is simply omitted.

For the following protocol, example results are included from the following 270 μl samples:

- 1 mM $^{15}N/^{13}C$-labeled or $^{15}N/^{13}C/^{2}H$-labeled ubiquitin (MWT 8.6 kDa) pH 6.5, recorded at 298 K on a Bruker DMX 800 MHz spectrometer equipped with a triple resonance TCI cryogenic probe.
- 2 mM (monomer concentration) $^{15}N/^{13}C$-labeled dimeric RANTES (2×8.5 kDa), pH 3.8, recorded at 310 K on a Bruker DMX 800 MHz spectrometer equipped with a triple resonance TCI cryogenic probe.
- 0.7 mM $^{15}N/^{13}C/^{2}H$-labeled MAP30 (30 kDa) recorded at 313 K on a Bruker DMX 750 MHz spectrometer equipped with a triple-resonance TXI room-temperature probe.

MATERIALS
REAGENTS
- Protein samples must be uniformly $^{13}C/^{15}N$-enriched. Isotope-labeled proteins are prepared using previously described strategies[28–31]
- For highest sensitivity, proteins should also be labeled with deuterium (^{2}H) for carbon-bonded hydrogen atoms. As this is usually achieved by growing bacteria in D_2O (see refs. 28,32,33), amide hydrogen atoms must be back-exchanged to ^{1}H nuclei during or after protein purification
- Samples should be prepared in a buffered solution containing 5% (vol/vol) D_2O
- Sample concentrations should be ≥ 0.5 mM
- Necessary sample volumes range between 250 and 500 μl depending on the NMR tube type (see accompanying protocol[16] on the detection of N–H⋯ N H-bonds)

EQUIPMENT
- A high-resolution NMR instrument (≥ 500 MHz ^{1}H frequency) with the following characteristics:
- Three radio frequency channels ($^{1}H/^{15}N/^{13}C$) and triple-resonance ($^{1}H/^{13}C/^{15}N$) probe-head optimized for ^{1}H detection
- For highest sensitivity, the magnetic field strength of the NMR spectrometer should be 18.8 Tesla (800 MHz ^{1}H frequency) or higher to make optimal use of the TROSY effect
- The NMR spectrometer should be equipped with pulsed-field gradients; otherwise, water suppression can be achieved by other techniques[34]
- As sensitivity is a limiting factor, cryogenic probes are beneficial, but are not absolutely necessary

PROCEDURE
Preparation
1| Calibrate ^{15}N and ^{13}C power levels on a suitable calibration sample according to **Box 1**.

2| Set the sample temperature of the NMR spectrometer to a suitable value, which is a compromise between sample stability and spectral quality. Insert the sample into the magnet.
? TROUBLESHOOTING

3| Calibrate the ^{1}H power levels on the sample according to the different pulse lengths given in **Box 1**.
▲ **CRITICAL STEP** Both the low-power and the shaped pulses are necessary for the water flip-back scheme to maximize water suppression and to improve signal to noise of the solvent labile amide proton signal.

BOX 1 | LONG-RANGE HNCO PARAMETERS

These parameters (see also the pulse program in **Fig. 2**) need to be set up by the operator to optimize acquisition.

1. Radio frequency carrier positions.
   - ^{1}H: exactly on water
   - ^{15}N: 116.5 ppm
   - ^{13}C: 177 ppm

2. Pulse lengths and power levels. Power levels should be determined according to the specified pulse length. Pulse lengths are specified for an 800 MHz spectrometer with cryogenic probe. For different field strength (X MHz) spectrometers, pulse lengths should be multiplied by 800/X.
   - ^{1}H: $\leq$12 μs and 1 ms 90° rectangular, $\leq$2 ms 90° sinc pulse (see **Fig. 2**)
   - ^{15}N: $\leq$40 μs 90° rectangular pulse
   - ^{13}C: 17.2 μs 90° rectangular and 110 μs 90° sinc pulses

3. Sweep widths and acquisition times.

^{1}H
   - Sweep width: $\approx$15 ppm (corresponds to a dwell time dw $\approx$ 83 μs at 800 MHz)
   - Number of complex increments (Ninc): choose such that the total acquisition time (aq = dw × Ninc) is about 3–5 times $T_2(^1H^N)$.

An acquisition time of 85 ms was used in the example.

^{13}C'
   - Sweep width: $\approx$20 ppm (corresponds to dw $\approx$ 250 μs at 201 MHz)
   - Ninc: for best sensitivity, choose such that the total acquisition time (aq = dw × Ninc) gives reasonable resolution, but does not exceed the T_2 of the ^{13}C' nuclei. An acquisition time of 40 ms was used in the examples.

4. N$-$C' transfer period.

Delays T and ε
   - T = 16.5 ms, ε = 4 μs for the test experiment
   - T = 66 ms, ε = 16.5 ms for the reference experiment
   - T = 66 ms, ε = 4 μs for the cross H-bond experiment

4| If you have no idea from previous experiments, determine the amide proton $^1H^N$ T_2 relaxation time from a suitable spin-echo experiment, for example, using the 1-1 echo sequence[35]. This step is not absolutely necessary, but it only takes a few minutes and is a very good way to characterize the quality of your sample. For the 8.6 kDa protein ubiquitin used in the example, the $^1H^N$ T_2 is about 35 ms at 298 K for the protonated and 75 ms for the deuterated sample. For the RANTES dimer (MWT = 2 × 8.5 kDa), the $^1H^N$ T_2 is about 28 ms at 310 K.

Parameter setup
5| Load the pulse program for the two-dimensional long-range H(N)CO TROSY experiment (**Fig. 2** and **Supplementary Box 1** online).

6| Set pulse lengths and power levels as determined in Steps 1 and 3 (see **Box 1**).

7| Adjust the carrier frequencies according to **Box 1**:
   - ^{1}H: set the carrier frequency exactly at the water position. (At this carrier position, a simple ^{1}H pulse-acquire experiment should yield a non-oscillating water FID (free induction decay).)
   - ^{15}N: set the carrier frequency to 116.5 ppm
   - ^{13}C: set the carrier frequency to 177 ppm

8| Set ^{1}H and ^{13}C sweep widths and acquisition times according to **Box 1**.

9| Optimize the recycling delay between scans, d1, to maximize the signal to noise. The amide proton longitudinal relaxation time (H^N T_1) increases with the size of the protein and depends on the level of deuteration. It is, therefore, important to have a good estimate for H^N T_1 to evaluate the optimum recycling delay d1 between scans. Generally, a good compromise for d1 is about 1 s for small- to medium-sized protonated proteins and 2–3 s for larger or deuterated proteins.

$^1J_{NC'}$ test spectrum
10| *Record a high-sensitivity $^1J_{NC'}$ test spectrum*. This spectrum is a normal HNCO and is a quick check on whether your parameter setup is correct. Set the total length of the N$-$C' transfer period to a small value, which is optimal for $^1J_{NC'}$ detection, for example, T = d16 = 16.5 ms, ε = d18 = 4 μs. Record the test spectrum with a small number of scans (NS = 4). This should take about 20–40 min. Process the data using parameters similar to the ones given in **Supplementary Box 2** online.

Details are described in the data-processing section (Step 15) of the HNN-COSY procedure of the accompanying protocol[16]. This should result in a normal 2D H(N)CO spectrum. Check whether the peaks have good line shapes and phase behavior and whether all the expected resonances are observable.

11| *Check the signal-to-noise ratio of the test spectrum.* Using appropriate software, determine the ratio of the height (maximum) of typical peaks relative to the noise. See Step 10 of the HNN-COSY procedure. In our examples, the S/N_{top} of typical resonances was about 43 (77):1 for protonated (deuterated) ubiquitin and 10:1 for protonated RANTES with NS = 4 after 20 min total experimental time on an 800 MHz machine equipped with a cryoprobe. If the sensitivity is not high enough, increase the number of scans.
? TROUBLESHOOTING

$^1J_{NC'}$ reference spectrum
12| *Record a $^1J_{NC'}$ reference spectrum.* Set the total length of the N–C' transfer period to the same value as will be used for the cross-experiment (T = d16 = 66 ms) and the delay between the ^{15}N and ^{13}C 180° pulses tuned for the observation of the $^1J_{NC'}$ 'diagonal' peaks (ε = d18 = 16.5 ms). Set all other parameters as for the test spectrum and acquire the data.

13| *Check the signal-to-noise ratio of the reference spectrum.* Process the data of the reference spectrum exactly as the test spectrum. Determine the signal-to-noise ratio of the relevant resonances. In our examples, the S/N_{top} of typical resonances was about 13 (27):1 for protonated (deuterated) ubiquitin and 4:1 for RANTES with NS = 4 after 20 min total experimental time. Note that such signal-to-noise ratios for the reference experiment are usually sufficient, since the error in the value of $^{h3}J_{NC'}$ coupling is dominated by the (small) signal-to-noise ratio of the cross-experiment. However, if the sensitivity of the reference experiment is not high enough, increase the number of scans.
? TROUBLESHOOTING

14| *Estimate the experimental time (number of scans) needed for the H-bond experiment.* Determine the minimal observable $^{h3}J_{NC'}$ coupling if a cross-experiment (long-range experiment) would be recorded with the number of scans of the reference experiment (NS_{ref}): $|J_{min}(NS_{ref})| = (S_{ref}/N_{top,ref})^{-1/2}/(2\pi T)$, where $N_{top,ref}$ corresponds to the top level of the noise and S_{ref} to the signal height of the reference spectrum (Step 13). If the final cross-experiment will be carried out with NS_{cross} number of scans, the respective minimal observable $^{h3}J_{NC'}$ coupling will be $|J_{min}(NS_{cross})| = (NS_{ref}/NS_{cross})^{1/4} \times |J_{min}(NS_{ref})| = (NS_{ref}/NS_{cross})^{1/4} \times (S_{ref}/N_{top,ref})^{-1/2}/(2\pi T)$. Typical $|^{h3}J_{NC'}|$ values are in the range of −0.2 to −0.9 Hz, with $|^{h3}J_{NC'}|$ values in β-sheets larger than in α-helices. Check whether the minimal calculated coupling for your anticipated experiment is sufficient for the detection of this $|^{h3}J_{NC'}|$ range. Otherwise increase NS_{cross}. Note that the available experimental time rapidly becomes limiting. For the protonated (deuterated) ubiquitin example, a total experimental time of 12.3 h (NS_{cross} = 136) yields $|J_{min}(NS_{cross})|$ = 0.28 (0.07) Hz, which covers most of the usual $^{h3}J_{NC'}$ values. For the example of RANTES, a total experimental time of 46.5 h (NS_{cross} = 512) yields $|J_{min}(NS_{cross})|$ = 0.36 Hz, which still allows detection of most β-sheet H-bonds.
? TROUBLESHOOTING

Long-range $^{h3}J_{NC'}$ experiment
15| *Record the $^{h3}J_{NC'}$ spectrum.* Set the length of the N—C' transfer period ($2T$) to the same value used for the reference experiment (T = d16 = 66 ms) and the delay between the ^{15}N and ^{13}C 180° pulses close to zero (ε = d18 = 4 μs). Adjust the number of scans according to your calculation in Step 14.

Data processing
16| Process the data in the same way as for the reference experiment.

Data analysis
17| Inspect the long-range HNCO spectrum. N—H⋯O=C H-bonds should be apparent from cross-peaks at the positions of the respective ($\omega_{C'j}$, ω_{Hi} + $\pi^1 J_{NH}$) frequencies. This may provide valuable information on the H-bond network of the protein.

18| Determine the size of the $^{h3}J_{NC'}$ couplings from the ratio of cross-peak ($\omega_{C'j}$, ω_{Hi} + $\pi^1 J_{NH}$) and reference peak ($\omega_{C'i-1}$, ω_{Hi} + $\pi^1 J_{NH}$) intensities. The simple approximation $|^{h3}J_{NC'}| \approx (I_{cross} \times NS_{ref}/(I_{ref} \times NS_{cross}))^{1/2}/(2\pi T)$ yields good estimates for values of $|^1J_{NC'}|$ close to 15 Hz and $|\pi^{h3}J_{NC'}2T| \ll 1$. However, for larger deviations of $|^1J_{NC'}|$ from 15 Hz (the typical range of $|^1J_{NC'}|$ is 13—17 Hz) or when $|\pi^{h3}J_{NC'}2T| \ll 1$ is not valid, the $^{h3}J_{NC'}$ values should be derived from a numerical solution of the equation $I_{cross} \times NS_{ref}/(I_{ref} \times NS_{cross}) = \cos^2(\pi^1 J_{NC'}2T)\sin^2(\pi^{h3}J_{NC'}2T)/[\sin^2(\pi^1 J_{NC'}(2T - 2\varepsilon))\cos^2(\pi^{h3}J_{NC'}(2T - 2\varepsilon))]$. This then requires the independent determination of the $^1J_{NC'}$ coupling constants. The determined values of $^{h3}J_{NC'}$ should be in the range of −0.1 to −0.9 Hz for most protein backbone H-bonds.

19| If no cross-peaks are observed for a certain H-bond, an upper limit of the coupling size can still be estimated from the top level of the noise in the cross-experiment ($N_{top,cross}$) and the intensity of the reference peak (I_{ref}), as $|^{h3}J_{NC'}|$ $< (N_{top,cross} \times NS_{ref}/(I_{ref} \times NS_{cross}))^{1/2}/(2\pi T)$.

● TIMING

Following protein sample preparation and resonance assignments, the test of the experiment, data acquisition of the reference and long-range experiments, processing and calculation of the $^{h3}J_{NC'}$ coupling is expected to take no more than 16 h for a small and deuterated protein like ubiquitin at 1 mM concentration. Longer acquisition times will be required for larger and non-deuterated proteins as well as for lower sample concentrations and if experiments are recorded at lower field strengths and on a conventional rather than a cryogenic probe.

? TROUBLESHOOTING
Insufficient sensitivity (Steps 2, 11, 13 and 14)
The low sensitivity of the long-range HNCO experiment stems from the necessary N→C′ transfer times of several hundred milliseconds. This low sensitivity presents the largest problem in the application of the experiment. If the test spectra reveal that the signal-to-noise is not high enough, several possibilities may exist to enhance the signal-to-noise.
- Increasing the concentration: if not ruled out by aggregation or lack of available protein material, this may be a simple solution.
- Deuteration: as evident from the examples of ubiquitin, deuteration significantly increases the sensitivity even for a small protein.
- Increasing the temperature: the temperature should be as high as possible without compromising protein stability. As the relaxation rates are proportional to η/T, a temperature increase by, for example, 10 K will have significant effects. However, the exchange of amide protons with water also increases at higher temperatures[23], especially for samples at pH $> \sim 6.5$, which may compromise sensitivity for less-protected amides. In addition, the size of the couplings decreases with increasing the temperature (($\Delta^{h3}J_{NC'}/\Delta T)/$ $^{h3}J_{NC'} \approx 2 \cdot 10^{-3}$ K^{-1}), because of the thermal expansion of proteins. Thus, the most suitable temperature has to be chosen with care[23].
- Optimizing the transfer delay T: the sensitivity of cross-peak detection is highest, when the transfer delay is chosen such that $2T$ is close to the ^{15}N relaxation time T_2 under the condition $2T = n/^1J_{NC'}$ (with $\varepsilon \approx 0$). For small- to medium-sized proteins, $2T = 133$ ms ($n = 2$) presents a reasonable compromise. However, for larger proteins, $2T = 66$ ms ($n = 1$) may be more sensitive. Vice versa, for smaller proteins or very long transverse relaxation times, for example, in unfolded proteins, $2T$ should be set to larger values.
- Increasing the experimental time: increasing the NS very rapidly becomes impractical, as the detection limit of the $^{h3}J_{NC'}$ couplings decrease only with the fourth root of NS.

Peak overlap
For small- to medium-sized proteins, peak overlap in the 2D H(N)CO is usually not a severe problem due to the large dispersion of ^{1}H^N and ^{13}C′ chemical shifts. Recording the 2D H(N)CO experiment at different temperatures may help to separate cross-peaks, as there is usually a significant variation in the temperature behavior of the different amide groups. Slight overlap may also be overcome by increasing the acquisition time of the ^{13}C′ dimension or by processing with apodization functions that provide higher resolution (although at the expense of sensitivity).

As an alternative, a 3D version of the long-range HNCO-TROSY can be recorded at no additional cost in sensitivity due to the sensitivity enhancement scheme implemented in the sequence in **Figure 2**. For this, the flag NITROGEN has to be defined in the pulse program in addition to CARBONYL (see **Supplementary Box 1**). For larger proteins, the 3D version is recommended, as overlap in the 2D version is unavoidable.

ANTICIPATED RESULTS
Figure 3a shows the result of the 2D long-range H(N)CO experiment carried out on a 2 mM sample of the RANTES dimer (total MWT 17 kDa) with a total measuring time of 46.5 h. A total of 13 intramolecular backbone H-bond $^{h3}J_{NC'}$ correlations (red) could be detected. In addition, four intermolecular H-bond correlations (orange) are visible, which connect the two RANTES monomers across an intermolecular antiparallel β-sheet. Besides these H-bonds, the spectrum shows a number of intraresidue two-bond ^{15}N$_i$-^{13}C′$_i$ (green) and sequential one-bond ^{15}N$_i$-^{13}C′$_{i-1}$ (blue) connectivities.

Figure 3b shows the 3D long-range HNCO TROSY spectrum of MAP30 (MWT 30 kDa) acquired previously[26] with a total experimental time of 91.2 h. Strip plots are given for five residues, which are part of the antiparallel β9/β10 β-sheet. A number of H-bond correlations (red) are visible in addition to the intraresidue (green) and sequential (blue) ^{15}N–^{13}C′ connectivities.

A quantitative evaluation was carried out previously for protonated[4] and deuterated[23] ubiquitin. For both deuterated and protonated samples of ubiquitin, typically more than 30 backbone $^{h3}J_{NC'}$ correlations are detected in 12 h experiments using about 1 mM samples. This corresponds to $>80\%$ of all backbone H-bonds expected from the ubiquitin crystal structure with a proton to carbonyl oxygen distance smaller than 2.2 Å. Quantification using a reference experiment yields coupling constants

in the range −0.2 to −0.9 Hz. The size of the coupling constants depends strongly on the H-bond geometry, which has been parameterized to good approximation by several simple geometric formulae[10], the simplest one being[36]

$$^{h3}J_{NC} = -357 \text{ Hz } \exp(-3.2 r_{HO}/\text{Å}) \cos^2 \theta$$

where r_{HO} presents the $H \cdots O$ distance and θ the $H \cdots O = C$ angle. As β-sheet H-bonds are typically shorter than α-helical H-bonds, differences can be detected in the values of $^{h3}J_{NC'}$ coupling constants: for the two proteins ubiquitin and protein G, the average $^{h3}J_{NC'}$ for H-bonds in β-sheet regions is −0.53 ± 0.15 Hz, whereas α-helical H-bonds have a weaker average $^{h3}J_{NC'}$ of −0.36 ± 0.18 Hz (ref. 12).

1. Jeffrey, G.A. & Saenger, W. *Hydrogen Bonding in Biological Structures* (Springer, New York, 1991).
2. Dingley, A.J. & Grzesiek, S. Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^2J_{NN}$ couplings. *J. Am. Chem. Soc.* **120**, 8293–8297 (1998).
3. Pervushin, K. *et al.* NMR scalar couplings across Watson–Crick base pair hydrogen bonds in DNA observed by transverse relaxation-optimized spectroscopy. *Proc. Natl. Acad. Sci. USA* **95**, 14147–14151 (1998).
4. Cordier, F. & Grzesiek, S. Direct observation of hydrogen bonds in proteins by interresidue $^{3h}J_{NC'}$ scalar couplings. *J. Am. Chem. Soc.* **121**, 1601–1602 (1999).
5. Cornilescu, G., Hu, J.-S. & Bax, A. Identification of the hydrogen bonding network in a protein by scalar couplings. *J. Am. Chem. Soc.* **121**, 2949–2950 (1999).
6. Shenderovich, I.G. *et al.* Nuclear magnetic resonance of hydrogen bonded clusters between F^- and $(HF)_n$: experiment and theory. *Ber. Bunsenges. Phys. Chem.* **102**, 422–428 (1998).
7. Golubev, N.S., Shenderovich, I.G., Smirnov, S.N., Denisov, G.S. & Limbach, H.-H. Nuclear scalar spin-spin coupling reveals novel properties of low-barrier hydrogen bonds in a polar environment. *Chem. Eur. J.* **5**, 492–497 (1999).
8. Blake, P.R. *et al.* Quantitative measurement of small through-hydrogen-bond and 'through-space' $^1H-^{113}Cd$ and $^1H-^{199}Hg$ J couplings in metal-substituted rubredoxin from *Pyrococcus furiosus*. *J. Biomol. NMR* **2**, 527–533 (1992).
9. Crabtree, R.H., Siegbahn, P.E.M., Eisenstein, O., Rheingold, A.L. & Koetzle, T.F. A new intermolecular interaction: unconventional hydrogen bonds with element-hydride bonds as proton acceptor. *Acc. Chem. Res.* **29**, 348–354 (1996).
10. Barfield, M. Structural dependencies of interresidue scalar coupling (h3)J(NC), and donor H-1 chemical shifts in the hydrogen bonding regions of proteins. *J. Am. Chem. Soc.* **124**, 4158–4168 (2002).
11. Cornilescu, G. *et al.* Correlation between $^{3h}J_{NC'}$ and hydrogen bond length in proteins. *J. Am. Chem. Soc.* **121**, 6275–6279 (1999).
12. Grzesiek, S., Cordier, F., Jaravine, V.A. & Barfield, M. Insights into biomolecular hydrogen bonds from hydrogen bond scalar couplings. *Prog. Nucl. Magn. Reson. Spectrosc.* **45**, 275–300 (2004).
13. Dingley, A.J., Cordier, F. & Grzesiek, S. An introduction to hydrogen bond scalar couplings. *Concepts Magn. Resonance* **13**, 103–127 (2001).
14. Grzesiek, S., Cordier, F. & Dingley, A.J. Scalar couplings across hydrogen bonds. *Methods Enzymol.* **338**, 111–133 (2001).
15. Grzesiek, S., Cordier, F. & Dingley, A.J. Hydrogen bond scalar couplings—a new tool in biomolecular NMR. in *Biological Magnetic Resonance*, Vol. 20 (eds. Krishna, N.R. & Berliner, L.J.) 255–283 (Kluwer Academic/Plenum, 2003).
16. Dingley, A.J., Nisius, L., Cordier, F. & Grzesiek, S. Direct detection of N-H···N hydrogen bonds in biomolecules by NMR spectroscopy. *Nat. Protoc.* **3**, 242–248.
17. Dingley, A.J., Masse, J.E., Feigon, J. & Grzesiek, S. Characterization of the hydrogen bond network in guanosine quartets by internucleotide $^{h3}J_{NC'}$ and $^{h2}J_{NN}$ scalar couplings. *J. Biomol. NMR* **16**, 279–289 (2000).
18. Liu, A., Hu, W., Majumdar, A., Rosen, M.K. & Patel, D.J. Detection of very weak side chain-main chain hydrogen bonding interactions in medium-size $^{13}C/^{15}N$-labeled proteins by sensitivity-enhanced NMR spectroscopy. *J. Biomol. NMR* **17**, 79–82 (2000).
19. Liu, A., Hu, W., Majumdar, A., Rosen, M.K. & Patel, D.J. NMR detection of side chain-side chain hydrogen bonding interactions in $^{13}C/^{15}N$-labeled proteins. *J. Biomol. NMR* **17**, 305–310 (2000).
20. Cordier, F., Wang, C., Grzesiek, S. & Nicholson, L.K. Ligand-induced strain in hydrogen bonds of the c-Src SH3 domain detected by NMR. *J. Mol. Biol.* **304**, 497–505 (2000).
21. Jaravine, V.A., Alexandrescu, A.T. & Grzesiek, S. Observation of the closing of individual hydrogen bonds during TFE-induced helix formation in a peptide. *Protein Sci.* **10**, 943–950 (2001).
22. Li, H., Yamada, H., Akasaka, K. & Gronenborn, A.M. Pressure alters electronic orbital overlap in hydrogen bonds. *J. Biomol. NMR* **18**, 207–216 (2000).
23. Cordier, F. & Grzesiek, S. Temperature-dependence properties of protein hydrogen bonds as studied by high-resolution NMR. *J. Mol. Biol.* **317**, 739–752 (2002).
24. Bougault, C.M., Eidsness, M.K. & Prestegard, J.H. Hydrogen bonds in rubredoxins from mesophilic and hyperthermophilic organisms. *Biochemistry* **42**, 4357–4372 (2003).
25. Assadi-Porter, F.M., Abildgaard, F., Blad, H. & Markley, J.L. Correlation of the sweetness of variants of the protein brazzein with patterns of hydrogen bonds detected by NMR spectroscopy. *J. Biol. Chem.* **278**, 31331–31339 (2003).
26. Wang, Y.-X. *et al.* Measurement of $^{h3}J_{NC'}$ connectivities across hydrogen bonds in a 30 kDa protein. *J. Biomol. NMR* **14**, 181–184 (1999).
27. Delaglio, F. *et al.* nmrPipe—a multidimensional spectral processing system based on unix pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
28. Marley, J., Lu, M. & Bracken, C. A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* **20**, 71–75 (2001).
29. Reilly, D. & Fairbrother, W.J. A novel isotope labeling protocol for bacterially expressed proteins. *J. Biomol. NMR* **4**, 459–462 (1994).
30. Cai, M. *et al.* An efficient and cost-effective isotope labeling protocol for proteins expressed in *Escherichia coli*. *J. Biomol. NMR* **11**, 97–102 (1998).
31. Goto, N.K. & Kay, L.E. New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Curr. Opin. Struct. Biol.* **10**, 585–592 (2000).
32. Grzesiek, S., Anglister, J., Ren, H. & Bax, A. C-13 line narrowing by H-2 decoupling in $^2H/^{13}C/^{15}N$-enriched proteins—application to triple-resonance 4d J-connectivity of sequential amides. *J. Am. Chem. Soc.* **115**, 4369–4370 (1993).
33. Venters, R.A. *et al.* High-level $^2H/^{13}C/^{15}N$ labeling of proteins for NMR studies. *J. Biomol. NMR* **5**, 339–344 (1995).
34. Messerle, B.A., Wider, G., Otting, G., Weber, C. & Wuthrich, K. Solvent suppression using a spin lock in 2d and 3d NMR-spectroscopy with H_2O solutions. *J. Mag. Reson.* **85**, 608–613 (1989).
35. Sklenar, V. & Bax, A. Spin-echo water suppression for the generation of pure-phase two-dimensional NMR-spectra. *J. Mag. Reson.* **74**, 469–479 (1987).
36. Sass, H.J., Schmid, F.F. & Grzesiek, S. Correlation of protein structure and dynamics to scalar couplings across hydrogen bonds. *J. Am. Chem. Soc.* **129**, 5898–5903 (2007).

2.4 Pressure-dependence of protein hydrogen bonds studied by NMR spectroscopy

Abstract

The pressure-dependence of $^{3h}J_{NC'}$ scalar couplings through hydrogen bonds was measured by nuclear magnetic resonance (NMR) spectroscopy in the range from atmospheric pressure up to 2500 bar in ubiquitin as a model protein. The quantitative 'long-range'-HNCO experiment was used to detect changes in electronic orbital overlap between the hydrogen and acceptor nuclei in ubiquitin's H-bonds. On average, these couplings are strengthened with increasing pressure. For some hydrogen bonds in the β-sheets the $^{3h}J_{NC'}$ couplings show a non-linear pressure-dependence with initially increasing coupling constants that decrease above pressures of $1200 - 1500$ bar. Together with H-bonds that are weakened already at low pressures, these residues mark the probably most pressure-labile regions of ubiquitin. Only for the first β-hairpin and the α-helix of ubiquitin, where the secondary structure is preserved at high pressures, a correlation between the pressure-dependent changes of the $^{3h}J_{NC'}$ coupling constants and amide proton chemical shift changes is observed. This correlates with earlier findings that the N-terminal half of ubiquitin is more stable and also preserved in its A-state under changed solvent conditions. Comparison of the pressure-induced changes in ubiquitin's hydrogen bond network with temperature-induced effects shows that for most residues an increase in pressure corresponds to a decrease in temperature with respect to the effect on $^{3h}J_{NC'}$ coupling constants. Hydrogen bonds that do not follow the global trends mark points, where local unfolding is likely to be initiated.

2.4.1 Introduction

The three-dimensional structure of proteins is stabilized by covalent (disulfide bonds) as well as non-covalent interactions (hydrogen bonds, ionic interactions, metal complexation, and hydrophobic interactions). Among the latter, hydrogen bonds are crucial for the formation of secondary and tertiary structural elements.

Scalar couplings across hydrogen bonds allow the direct observation, not only of the donor and acceptor atoms involved in a given hydrogen bond, but also its strength (Cordier *et al.* 1999, Cornilescu *et al.* 1999b). Since the $^{3h}J_{NC'}$ scalar couplings connecting the ^{15}N of the hydrogen bond donor with the carbonyl ^{13}C of the acceptor amino acid depend on the electron orbital overlap between hydrogen and the acceptor atom (Barfield 2002), they are an extremely sensitive probe of hydrogen bond length (Cornilescu *et al.* 1999a) and angles (Grzesiek *et al.* 2004). Thus, these coupling constants are a good parameter to observe changes in proteins subjected to different conditions, such as ligand binding (Cordier *et al.* 2000), exposure to organic solvents (Jaravine *et al.* 2001), or variation in temperature (Cordier & Grzesiek 2002) and pressure (Li *et al.* 2000). A number of studies have shown that although the $^{3h}J_{NC'}$ couplings in N-H $\rightarrow$ O=C hydrogen bonds are small in size, they can be measured reproducibly. Particularly, the influence of temperature on the hydrogen bond network in ubiquitin was studied by NMR spectroscopy using $^{3h}J_{NC'}$ scalar couplings as well as other hydrogen bond parameters (Cordier & Grzesiek 2002). This study demonstrated that hydrogen bond scalar couplings can be measured precisely enough to deduce thermally induced changes in H-bond geometry from variations in the coupling constants. Over the temperature range from 5 °C to 65 °C the average length of hydrogen bonds expands by 0.03 Å. The most thermolabile region in ubiquitin is the N-terminal end of β-strand β5 whereas the end of β-sheet β1/β2 is even stabilized at higher temperatures.

The influence of pressure on a protein's hydrogen bond network was investigated in the immunoglobulin binding domain of streptococcal protein G (Li *et al.* 2000). In this study it was shown that $^{3h}J_{NC'}$ couplings in the core of the protein decrease under pressure whereas those in loop and peripheral regions increase. Since the experimental set-up in this study allowed only very small sample volumes, just eleven out of 34 hydrogen bond scalar couplings could have been detected with sufficient accuracy, namely those stronger than 0.5 Hz. Due to these problems with measurements under high pressure, usually not the $^{3h}J_{NC'}$ couplings but other H-bond related, better observable parameters, such as amide proton chemical shift, have been measured.

Here, we present the pressure dependence of $^{3h}J_{NC'}$ coupling constants in ubiquitin in the range from 1 bar up to 2500 bar. Due to a new high-pressure NMR cell that allows much larger sample volumes than previous set-ups (120 μl vs 20 μl), we could

follow the changes in $^{3h}J_{NC'}$ couplings for almost all hydrogen bonds. Interestingly, no correlation between hydrogen bond scalar couplings and amide proton chemical shifts has been observed for some part of ubiquitin, namely the β-sheets 3/4, 3/5, and 1/5. Furthermore, the pressure-induced variation of the amide proton, nitrogen and carbonyl carbon chemical shifts has been monitored at different temperatures.
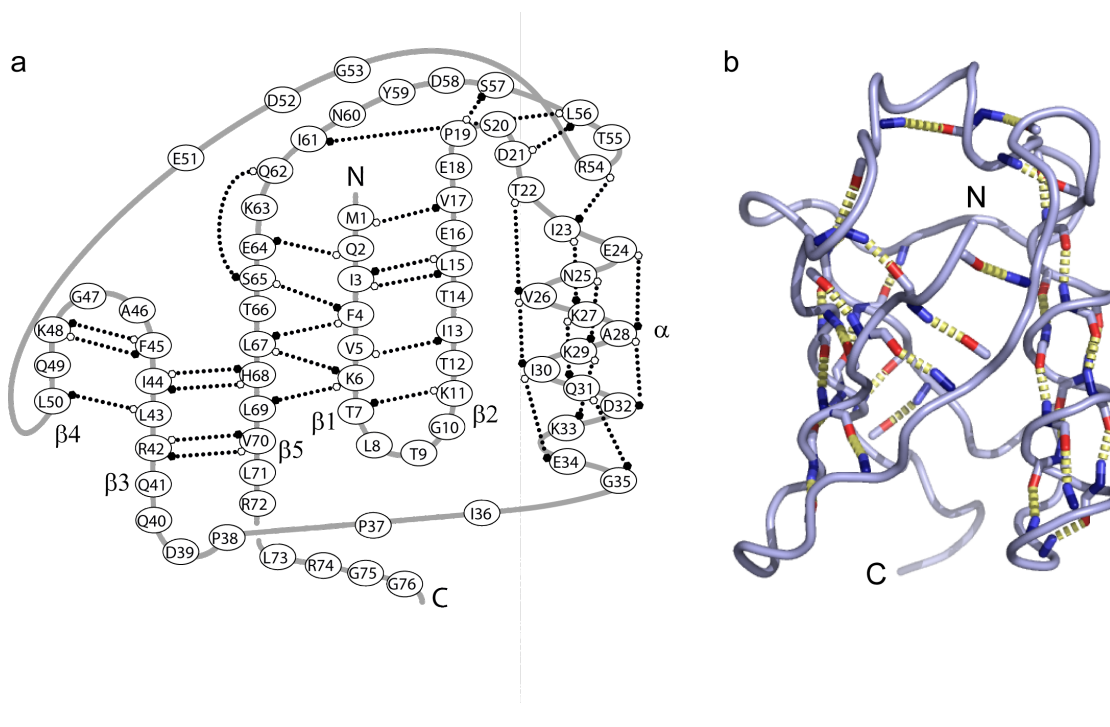

Figure 2.2   **a** Topology of Ubiquitin with the backbone hydrogen bonds (dotted lines) between the amide proton (filled circles) and the carbonyl oxygen (open circles). Figure taken from (Cordier & Grzesiek 2002). **b** High-resolution X-ray structure of ubiquitin (pdb 1ubq). The hydrogen bond network is shown with the amide donor group in blue and the carbonyl acceptor in grey and red. H-bonds between NH and CO, that could be detected by NMR spectroscopy with the 'long-range'-HNCO experiment, are represented as dashed yellow lines.

2.4.2 Results and Discussion

Pressure-induced changes in ^{3h}J$_{NC'}$ hydrogen bond scalar couplings

Figure 2.3 shows a plot of the $^{3h}J_{NC'}$ scalar couplings of ubiquitin measured at 1, 300, 600, 900, 1200, 1500, 2000, and 2500 bar. The data represent the average of three independent measurements. The hydrogen bond scalar coupling constants could be measured very reproducibly, with maximum errors of 0.012 Hz.

On average, the coupling through H-bonds gets stronger up to pressures of 1200 – 1500 bar. Afterwards, two main trends can be observed: some couplings become even

stronger whereas for others the direction is reversed. There are few hydrogen bonds that do not fit into these two groups, but show a decrease in strength already at low pressures.

As indicated by the change in scalar couplings through hydrogen bonds, pressure influences ubiquitin non-uniformly. The largest effects can be observed for hydrogen bonds in the parallel $\beta1/\beta5$ and in the antiparallel $\beta3/\beta5$ β-sheets (R42, L67, and L69) as well as for residue A28 in the α-helix.
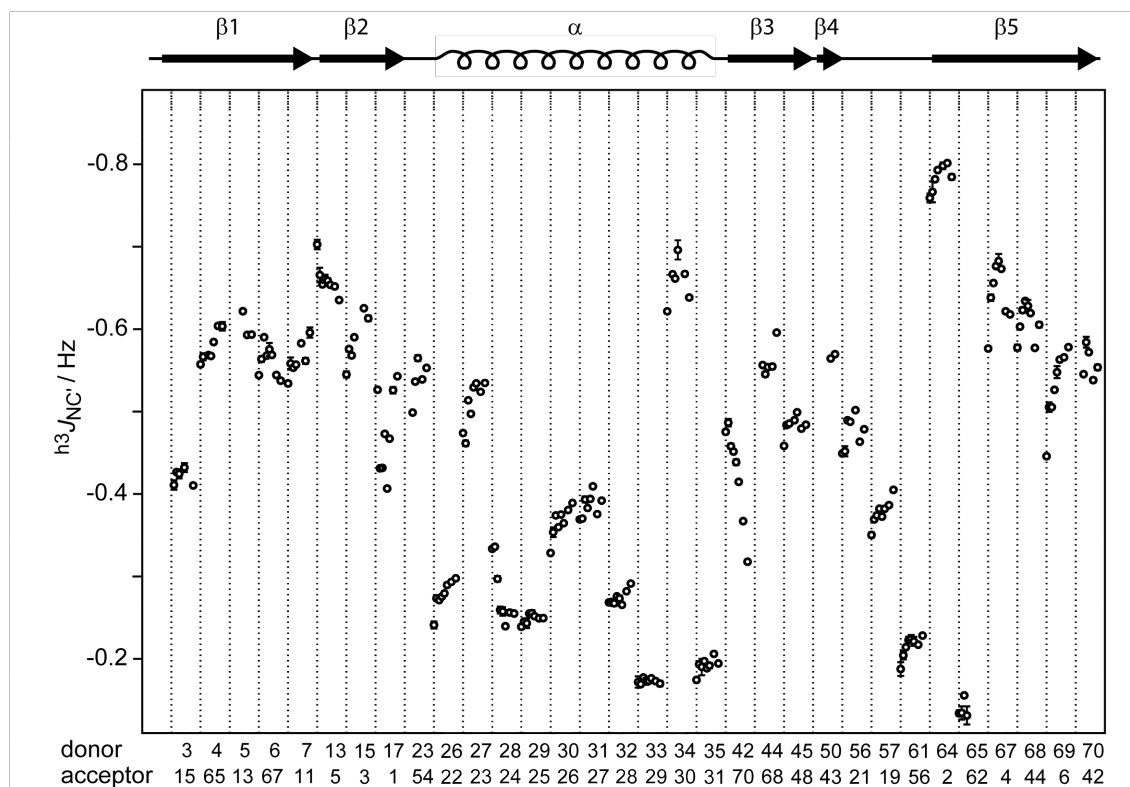

Figure 2.3   Pressure-induced changes in $^{3h}J_{NC'}$ coupling constants in ubiquitin as a function of the amino acid sequence. For each hydrogen bond donor-acceptor pair, the coupling constants at eight pressures (1, 300, 600, 900, 1200, 1500, 2000, and 2500 bar) are shown as the mean of three measurements. Depicted error bars correspond to the resulting standard deviation. Donor and acceptor amino acids are given at the bottom while on top the secondary structural elements are represented. Note that the coupling constants are negative and thus a numerical decrease corresponds to a stronger coupling. For this reason, the y-axis scale is inverted.

Pressure-dependence of H-bond properties in different secondary structural elements
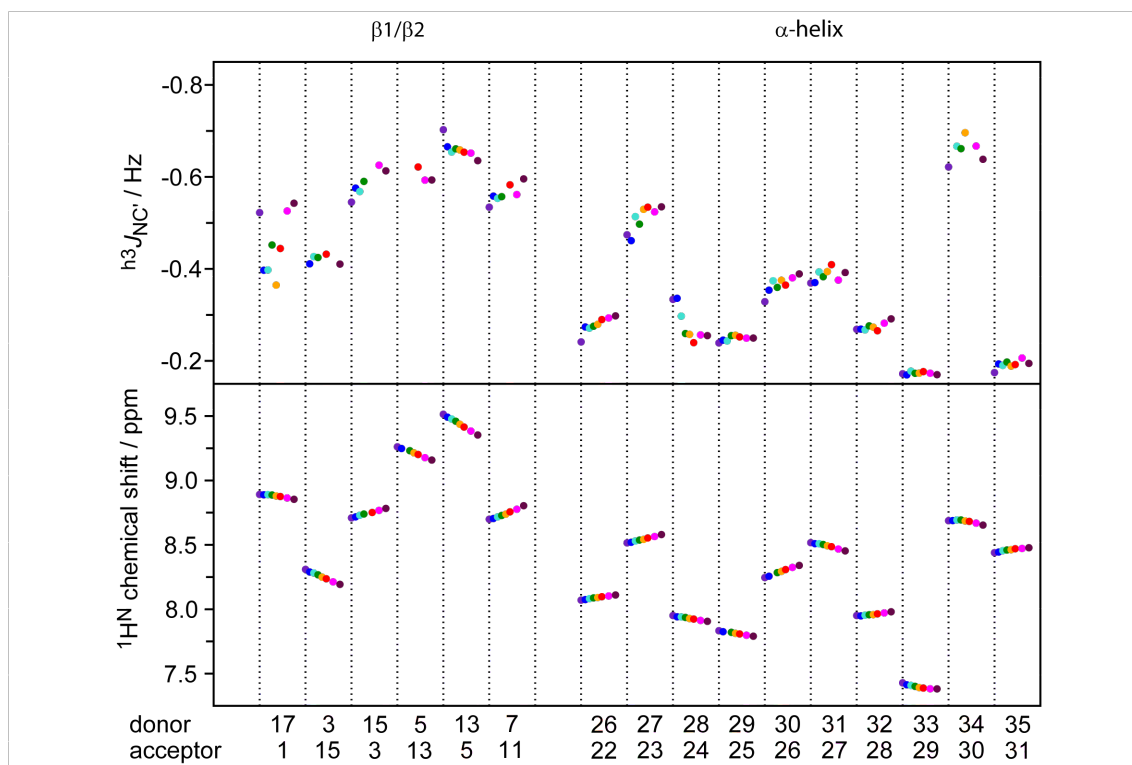
Figure 2.4 $^{3h}J_{NC'}$ coupling constant (top) and amide proton chemical shift (bottom) for the hydrogen bonds within the $\beta1/\beta2$ β-sheet and the α-helix of ubiquitin. Values are shown for eighth different pressures between 1 and 2500 bar: 1 (purple), 300 (blue), 600 (turquoise), 900 (green), 1200 (orange), 1500 (red), 2000 (magenta), and 2500 bar (brown). Spacing between data points represents the corresponding difference in pressure.

The $\beta1/\beta2$ β-sheet and the α-helix The geometry of the $\beta1/\beta2$-sheet and the α-helix in ubiquitin has been described before (Cordier & Grzesiek 2002). Briefly, the β-sheet 1/2 is influenced by the irregular geometry of the whole $\beta1/\beta2/\beta5$-region, leading to a very short I13$\rightarrow$V5 H-bond ($r_{OH} = 1.71$ Å, according to the X-ray structure 1ubq).

Under pressure, the antiparallel $\beta1/\beta2$-sheet shows non-uniform behavior with three couplings becoming clearly stronger and one becoming weaker. Starting from the N-terminus, the first hydrogen bond occurs between V17 and M1, for which the strongest increase in coupling is observed (-0.40 Hz at 300 bar to -0.54 Hz at 2500 bar). Points that do not fit very well could be due to a second conformation. Note that the $^{3h}J_{NC'}$ coupling constants are negative and thus a lower value corresponds to a stronger coupling. The hydrogen bonds between residues L15/I3 and T7/K11 strengthen to a similar

extend between 1 and 2500 bar. In contrast, the coupling across the H-bond between I13 and V5 decreases, but is nevertheless the strongest coupling within β-sheet 1/2 over the whole range of pressures. Another interesting observation is that the hydrogen bond scalar couplings for the $\beta 1/\beta 2$-sheet vary much less at higher pressure, indicating a harmonization of the $^{3h}J_{NC'}$ coupling constants. This reflects a more uniform geometry of the β-sheet at higher pressure.

The α-helix is uniform in its backbone (ϕ and ψ angles) between residues I23 and D32 and slightly more varied at the C-terminal end (K33 to G35). The residues involved in the hydrogen bond network of the α-helix are V26 to G35 with $H^N(i) \rightarrow O(i-4)$ H-bonds. The helix is bent by about 9° (Cordier & Grzesiek 2002).
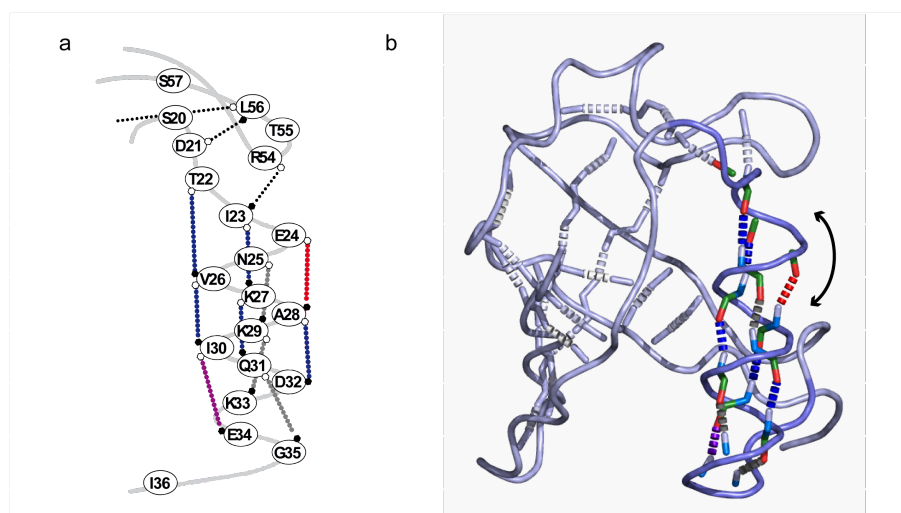
Figure 2.5   Pressure-induced changes of $^{3h}J_{NC'}$ within ubiquitin's α-helix. Couplings that become stronger are depicted in blue, H-bonds with weakened couplings are shown in red. The E34→I30 hydrogen bond shows a curved behavior. H-bonds colored grey do not show any changes.

Compared to the β-sheets, pressure-induced changes in the α-helix are smaller, with the largest overall difference observed for the A28→E24 hydrogen bond (see Figure 2.4). Figure 2.5 illustrates the major changes in the $^{3h}J_{NC'}$ coupling constants. The coupling between A28 and E24 on the helix outside is weakened, indicating an increase in H-bond length. On the opposite side, the couplings between V26/T22, K27/I23, I30/V26, and Q31/K27 become stronger. Thus the initial bend in the helix becomes even more pronounced under pressure. Assuming constant angles, the changes in coupling constants would correspond to an H-bond expansion for A28/E24 by 0.067 Å. At the same time, the hydrogen bond between V26 and T22 shortens by a similar amount (0.053 Å).

The hydrogen bond scalar coupling for E34→I30 is the largest in the helix and shows a unique behavior: after an initial strengthening the coupling constant weakens at 2500 bar to nearly the same value as for 1 bar.

The fact that the couplings for the H-bonds K33→K29 and G35→Q31 are very weak and flank the hydrogen bond between E34 and I30 with an extremely strong coupling is another indication for a bended helix. In contrast to the N-terminal end, this C-terminal part of the helix does not seem to be very much influenced by the pressure as indicated by nearly constant $^{3h}J_{NC'}$ couplings (D32→A28, K33→K29, G35→Q31). The different behavior of E34→I30 might be rather explained in terms of a slight movement of the helix towards β-strand 2. This should result in a strengthening of the salt bridge between E34 and K11, which could induce a change in angles for the E34/I30 hydrogen bond.
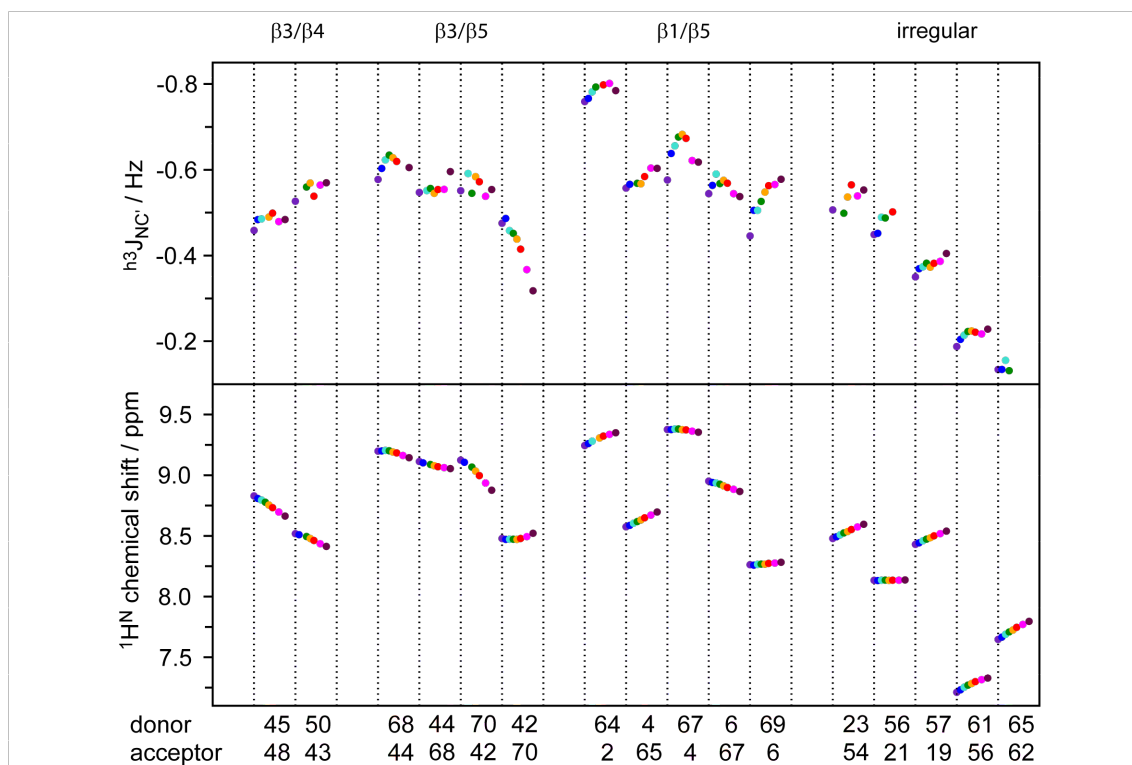
Figure 2.6 $^{3h}J_{NC'}$ coupling constant (top) and amide proton chemical shift (bottom) for the hydrogen bonds within the $\beta3/\beta4$, $\beta3/\beta5$, and $\beta1/\beta5$ β-sheets and those not involved in regular secondary structural elements in ubiquitin (see figure 2.4 for a description of the color coding).

The $\beta3/\beta4$ and $\beta3/\beta5$ β-sheets   The backbone ϕ and ψ angles for the hydrogen bonded residues in these two β-sheets are uniform with only the C-terminal end of β-strand 4 (L50) being slightly distorted (Cordier & Grzesiek 2002).

In the β-sheet 3/4 only two hydrogen bonds are observable via $^{3h}J_{NC'}$ couplings (Figure 2.6). The two observable cross-H-bond coupling constants suggest that the $\beta3/\beta4$ β-sheet is only marginally influenced by pressures up to 2500 bar. Nevertheless, the corresponding amide proton chemical shifts decrease in a nearly linear fashion with pressure by ≈ 0.08 ppm.

As evident from the crystal structure, the $\beta3/\beta5$-sheet shows very uniform geometries for the hydrogen bonds H68→I44, I44→H68, and V70→R42 (Cordier & Grzesiek 2002). The $^{3h}J_{NC'}$ coupling constants across these hydrogen bonds are -0.58 Hz, -0.55 Hz, and -0.55 Hz at 1 bar, respectively (Figure 2.6). Although these values are in a very narrow range, their pressure-dependent behavior is different. The first hydrogen bond scalar coupling shows a curved behavior, while the coupling across the second hydrogen bond is nearly constant over the whole pressure range. In contrast, the coupling across the H-bond V70→R42, although very similar for 1 and 2500 bar, fluctuates rather strongly in between. This could be due to a second species that becomes visible for residue V70 under pressure. Nevertheless, compared to the β-sheets 1/2 and 1/5, the influence of pressure on the coupling constants for these three hydrogen bonds is small ($\Delta^{3h}J_{NC'} = 0.10$ Hz vs 0.34 Hz for $\beta1/\beta2$ and 0.35 Hz for $\beta1/\beta5$).

On the contrary, the $^{3h}J_{NC'}$ coupling between R42→V70 behaves in a very different way: the coupling drastically weakens with higher pressures from -0.49 Hz down to -0.32 Hz at 2500 bar. In fact, this coupling is the weakest detected within the β-sheet structures in ubiquitin. The extreme change in $^{3h}J_{NC'}$ by 0.17 Hz indicates an opening of the $\beta3/\beta5$-sheet from its C-terminus. Surprisingly, the proton chemical shift for R42 is constant up to 1200 bar before rising. The proton chemical shifts of the second hydrogen bond donor between these residues, V70, decreases over the whole pressure range and exhibits the largest overall change ($\Delta\delta = 0.25$ ppm).

The $\beta1/\beta5$ β-sheet   The parallel $\beta1/\beta5$ β-sheet shows a wide distribution of $^{3h}J_{NC'}$ coupling constants between -0.45 Hz (L69→K6 at 1 bar) and -0.80 Hz (E64→Q2 at 2000 bar; Figure 2.6). There is a trend of decreasing hydrogen bond scalar couplings from the N- to the C-terminus in this β-sheet (at 1 bar) which goes in hand with increasing O–H distances in the crystal structure (1ubq). The β-sheet starts with an irregular hydrogen bond in strand $\beta5$, connecting the amide donor of S65 with the carbonyl acceptor of Q62 instead of the regular acceptor Q2 (see also Figure 2.7b). Glutamine 2, therefore, is able to form a hydrogen bond to the amide group of E64 (Cordier & Grzesiek 2002). Coupling across this H-bond is the strongest observed in ubiquitin at all pressures.

All hydrogen bond scalar couplings in this β-sheet become stronger with pressures of up to $900 - 1200$ bar. Two out of five couplings increase further at higher pressures, whereas the others decrease if the pressure is increased further. The strongest effect of all hydrogen bond scalar couplings in the $\beta 1/\beta 5$ β-sheet occurs for the coupling across the terminating H-bond at the sheets C-terminus, L69$\rightarrow$K6, which changes over the whole pressure range from -0.45 Hz at 1 bar to -0.58 Hz at 2500 bar. This might be induced by an opening of the neighboring β-sheet (R42$\rightarrow$V70), which releases the $\beta 5$-strand so that it can move closer to $\beta 1$. Assuming a sole distance dependency, the hydrogen bond between R42 and V70 would increase in length by 0.101 Å, while that between L69 and K6 would decrease by 0.065 Å.

It can be speculated that the initial strengthening in $^{3h}J_{NC'}$ coupling constants in the $\beta 1/\beta 5$ as well as in parts of the $\beta 1/\beta 2$ β-sheet is not only due to a shortening of the hydrogen bonds but also to a release of the initial twist in this mixed parallel, antiparallel β-sheet.

Residues not involved in regular secondary structural elements There are five hydrogen bonds in ubiquitin that are not within a regular secondary structural element. These are found in the loops flanking the 3_{10}-helix and connect I23$\rightarrow$R54, L56$\rightarrow$D21, S57$\rightarrow$P19, I61$\rightarrow$L56, and S65$\rightarrow$Q62. The $^{3h}J_{NC'}$ couplings across all these hydrogen bonds increase with pressure (Figure 2.6). In more detail, the strongest coupling is observed for I23$\rightarrow$R54, concatenating the loop between $\beta 4$ and the 3_{10}-helix to the α-helix (see also Figure 2.7b). The weakest coupling on the other hand is found for the unusual hydrogen bond between S65 and Q62 at the N-terminal end of the $\beta 5$-strand. With respect to the amide proton chemical shift, the five hydrogen bonds outside regular secondary structural elements behave similarly showing a downfield shift. Only H^N of L56 is not changed over the whole pressure range.

Implications of pressure-induced changes of $^{3h}J_{NC'}$ on ubiquitin's structure

On average, the hydrogen bond scalar couplings in ubiquitin slightly increase over the pressure range from 1 to 2500 bar (see Figure 2.7a).

The changes in hydrogen bond scalar couplings indicate, that the first two β-strands and the α-helix in ubiquitin behave rather uniformly under pressure with most couplings becoming stronger. Only the couplings across the hydrogen bond A28$\rightarrow$E24 in the helix and I13$\rightarrow$V5 in the first β-sheet decrease under pressure. In contrast, the hydrogen bonds between β-strands 1 and 5 show a curved behavior with similar couplings for 1 and 2500 bar and a maximum around 1200 bar. The strength of H-bonds between the β-strands 3 and 5 decreases, particularly for R42 / V70. These changes might mark the

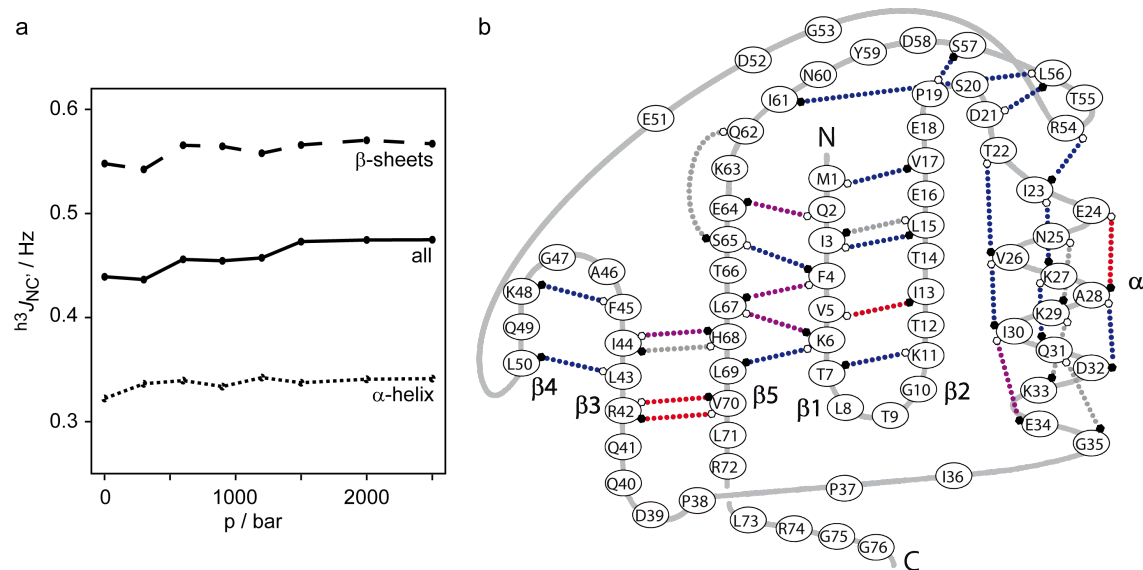starting point for pressure unfolding.

Figure 2.7   **a** Average pressure-induced change of the $^{3h}J_{NC'}$ coupling constants in ubiquitin for all hydrogen bonds (solid line), for those in β-sheet structures (dashed line), and for H-bonds in the α-helix (dottet line). **b** Direction of changes for individual hydrogen bond scalar couplings marked on the protein's topology. Couplings that become stronger are depicted in blue, H-bonds with weakened couplings are shown in red, purple color marks a curved change in $^{3h}J_{NC'}$. H-bonds colored grey do not show any changes.

Correlation between the $^{3h}J_{NC'}$ coupling constant and the $^1H^N$ chemical shift

It has been shown before that the hydrogen bond scalar coupling constant correlates with the amide proton chemical shift (Cordier *et al.* 1999). For ubiquitin, a correlation between these parameters was shown for the temperature-dependence (Cordier & Grzesiek 2002) as well as for the influence of the solvent (Cordier & Grzesiek 2004).

Figure 2.8 shows the change of the $^{3h}J_{NC'}$ coupling constant under pressure (2500 bar) *versus* the $^1H^N$ chemical shift difference between 1 and 2500 bar. The values are normalized for the pressure change in both dimensions. Regarding the whole protein, there is no correlation between the two parameters. This may indicate that the $^{3h}J_{NC'}$ coupling constant and the $^1H^N$ chemical shift have a different geometry dependency that only becomes apparent under pressure. Furthermore, the amide proton chemical shift can be influenced not only by the hydrogen bond length and angles, but also by electric field effects.

However, if different structural elements of ubiquitin are regarded separately, the picture changes. In regions where the hydrogen bonds are strengthened and hence the secondary

structure is stabilized, a good correlation can be found. In figure 2.8, hydrogen bonds in the first β-sheet and the α-helix of ubiquitin are depicted in green. If only these values are fitted, the empirical correlation $\Delta^{3h}J_{NC'} = 0.0096 + 0.3735\Delta\delta^1H^N$ can be found. The values for V17 and A28 do not fit well. In case of V17 the reason for this is probably that this hydrogen bond is the first in ubiquitin and marks the beginning of the $\beta1/2$-sheet. Furthermore, increased exchange has been shown for this residue. The hydrogen bond A28→E24 changes dramatically due to helix bending, which could influence $\Delta^{3h}J_{NC'}$ and Δ^1H^N differently. In contrast, for regions where the hydrogen bonds are weakened, a random change of δ^1H^N is observed. Fitting of the remaining data points (orange) shows no correlation between the $^{3h}J_{NC'}$ coupling constant and the $^1H^N$ chemical shift for the C-terminal part of ubiquitin.
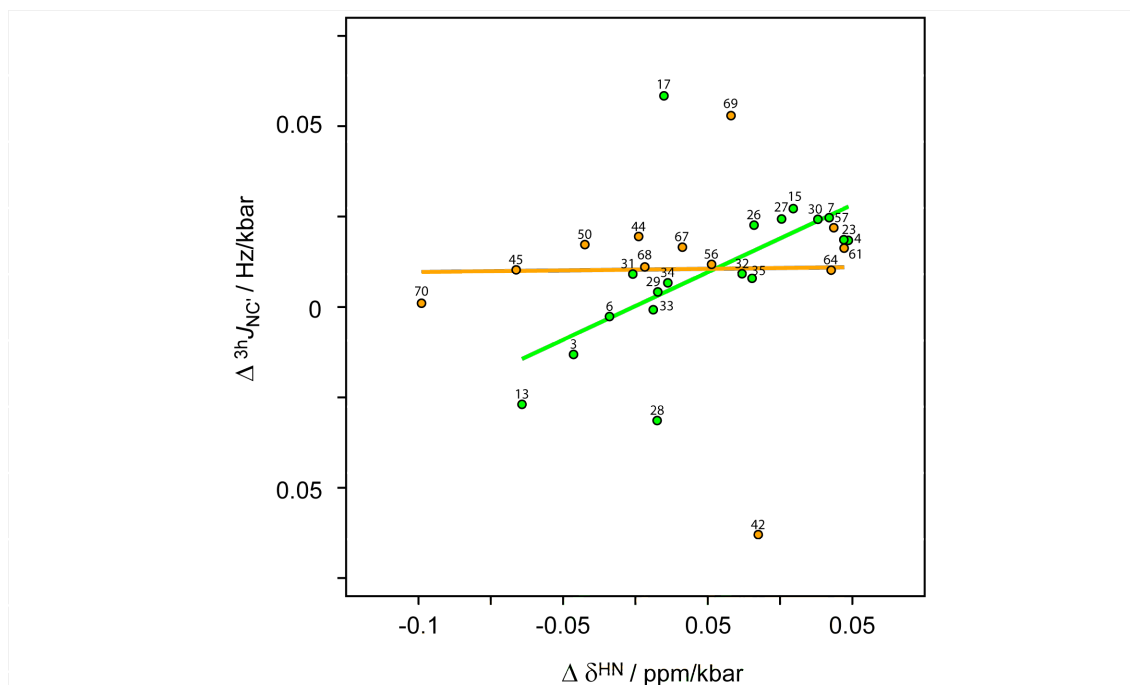
Figure 2.8   Pressure-induced changes in the $^{3h}J_{NC'}$ coupling constant *versus* pressure-dependent amide proton chemical shift change. Hydrogen bonded residues in the first β-sheet as well as the α-helix are colored green and grouped together, amino acids involved in the remaining H-bonds are shown in orange. Straight lines represent the corresponding linear fits. The correlation is depicted for a pressure difference of 2500 bar (all values are normalized for this pressure change). The outlayers (V17, R42, and L69) mark beginnings or ends of β-sheets.

Comparison of pressure and temperature effects on $^{3h}J_{NC'}$

For several proteins, phase diagrams showing the effect of temperature and pressure have been obtained (see figure 2.1) (Smeller 2002). According to these phase diagrams,

an expansion of the protein is expected with increasing temperature, which leads to a destabilization and finally to heat denaturation. In contrast, increasing pressure, initially leads to stabilization of the protein, due to a volume reduction. Thus the effect of high temperature and moderately high pressure oppose each other. Above a certain limit high pressure also destabilizes and finally unfolds proteins.

The temperature-dependence of the $^{3h}J_{NC'}$ coupling has been measured before (Cordier & Grzesiek 2002). In the following, this temperature-dependent change of $^{3h}J_{NC'}$ is compared with its pressure-dependent variation.

Figure 2.9 illustrates the effect of pressure and temperature on $^{3h}J_{NC'}$ for the observable H-bonds in ubiquitin. For example, leucine 15 and isoleucine 30 show a stronger coupling constant at higher pressures, corresponding to a stronger hydrogen bond (smaller volume). With increasing temperature, these two hydrogen bonds become weaker, as indicated by a weakened coupling constant. Thus, the behavior of these amino acids under the influence of pressure and temperature agrees with that expected from the whole-protein phase diagram.

In contrast, there are hydrogen bonds which show other behavior than predicted by the phase diagram that would represent the whole protein. The coupling constant for the H-bond of isoleucine 13 becomes weaker with both, high pressure and high temperature. A similar behavior is observed for valine 70, corresponding to an increase in volume in both cases.

But also an increase in temperature can lead to a strengthening of H-bonds as observable for isoleucine 3.

Thus, although globally the $^{3h}J_{NC'}$ coupling constants become stronger under the pressures applied (max. 2500 bar) and weaker with increasing temperature, there are local effects, that do not fit into this picture. Therefore, the combined investigation of pressure and temperature effects on the hydrogen bond network of a protein might give valuable information about local unfolding. Furthermore, such studies might be used to compare different ways of denaturation. Whereas heat seems to have a more uniform effect on ubiquitin, high pressure has more pronounced local effects.

Differences in the pressure-dependence of $^1H^N$ chemical shifts at high and low temperature

In order to get a more complete picture of ubiquitin's behavior under both, different temperatures and pressures, HSQC and HNCO spectra, giving the $^1H^N$, ^{15}N, and $^{13}C'$ chemical shifts, were also measured at 5 °C as a function of pressure. For technical reasons this was possible only up to pressures of 1500 bar and for rather short times (1–2 hours). Therefore, no $^{3h}J_{NC'}$ coupling constants have been obtained but only the
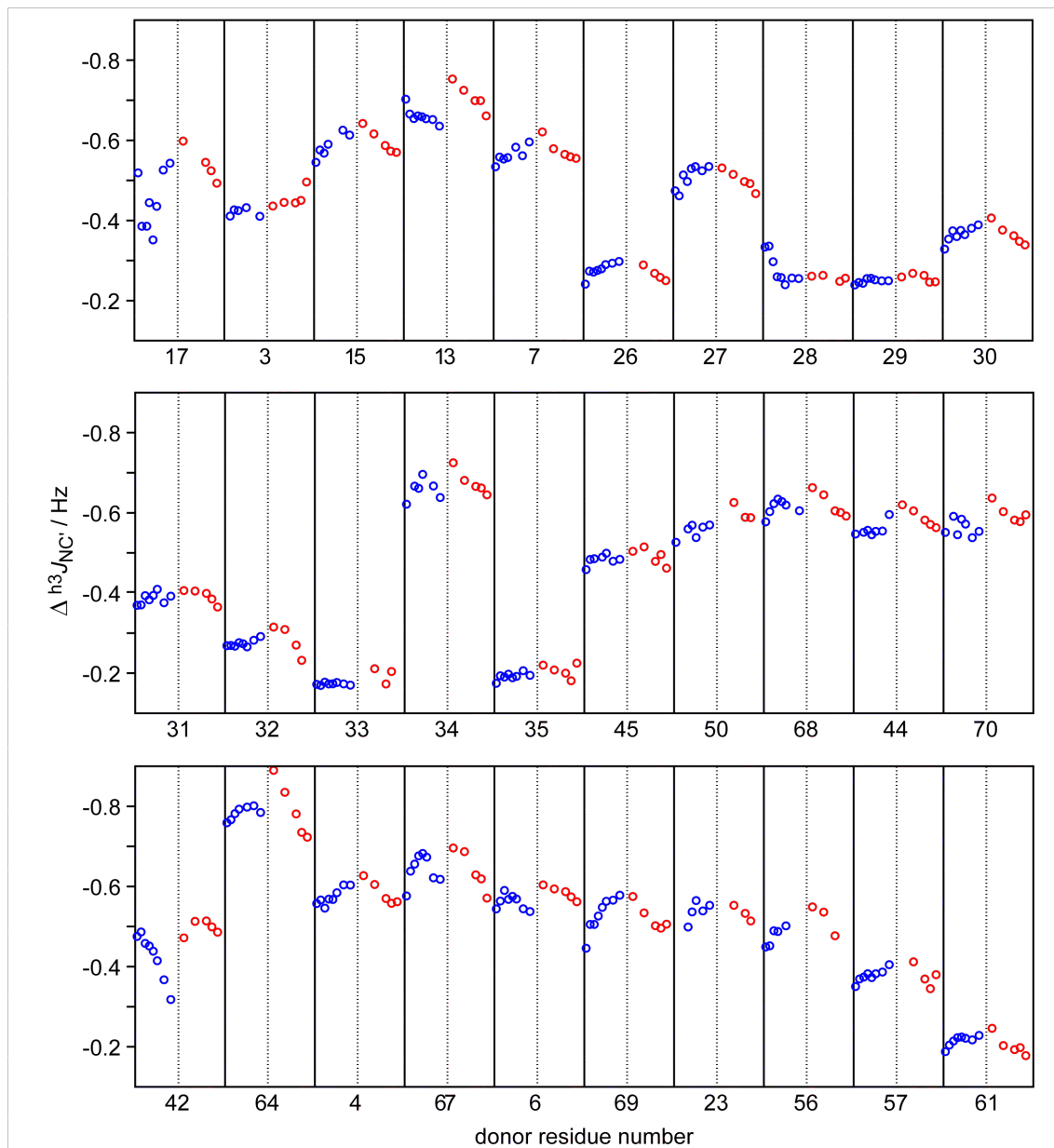
Figure 2.9   Change of the $^{3h}J_{NC'}$ coupling constant under the influence of pressure (blue circles from left to right: 1, 300, 600, 900, 1200, 1500, 2000, and 2500 bar at 35 °C) and with increasing temperature (red circles from left to right: 5, 25, 45, 55, and 65 °C; data taken from (Cordier & Grzesiek 2002)). Data are shown for the observable hydrogen bond scalar couplings in ubiquitin and labeled with the donor residue number. For most H-bonds, the temperature effect opposes the pressure-dependent behavior, only some residues show a different pattern (see text).

changes of the chemical shift of the amide proton and nitrogen as well as the carbonyl carbon have been measured.
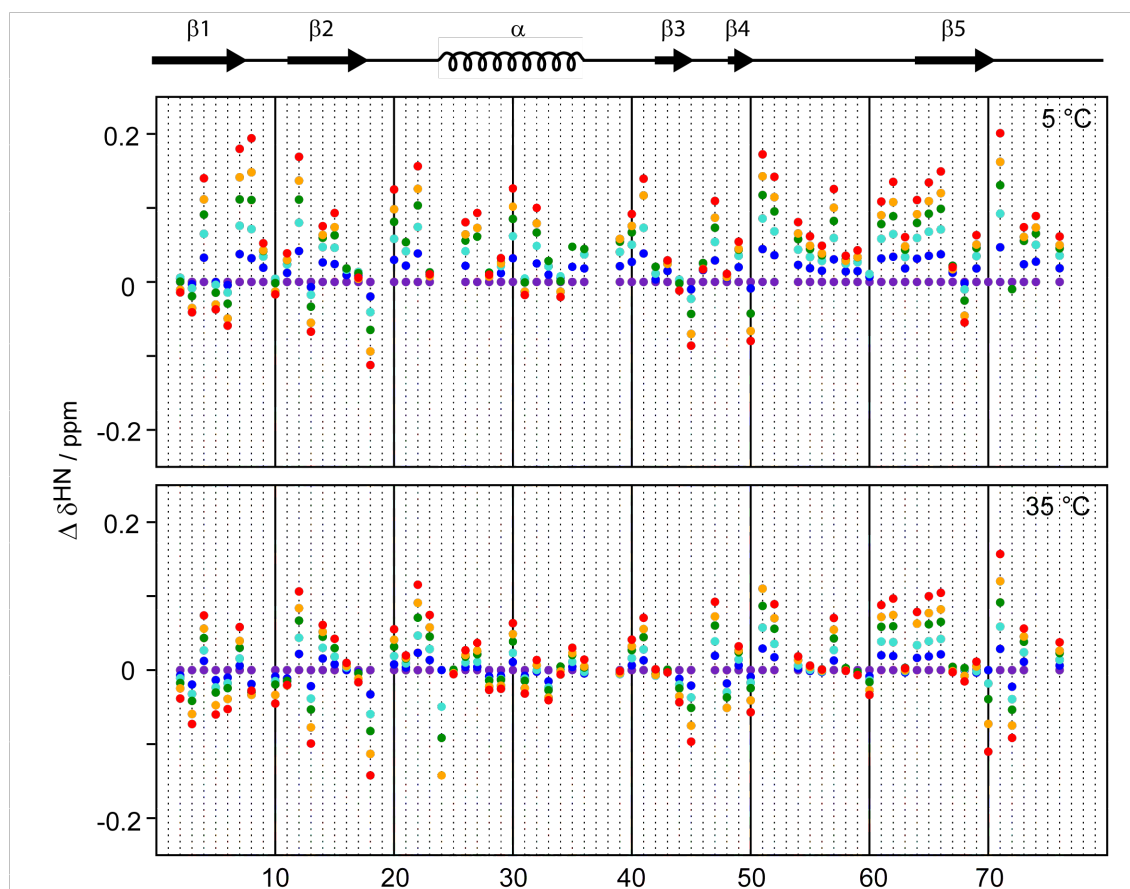
Figure 2.10   Pressure-induced changes in the amide proton chemical shift as a function of the amino acid sequence at 5 °C (top) *versus* 35 °C (bottom). Data are shown for six different pressures between 1 and 1500 bar (color coding: 1 (purple), 300 (blue), 600 (turquoise), 900 (green), 1200 (orange), 1500 (red)).

Figure 2.10 compares the pressure-induced amide proton chemical shift changes at 5 °C and 35 °C as a function of the amino acid sequence. On average, the $^1\text{H}^\text{N}$ chemical shift clearly shows larger pressure-induced changes at 5 °C compared to 35 °C. For most residues the resonances are shifted further downfield with pressure at the lower temperature, indicating a more compact structure. There are only few amino acids whose proton chemical shift is less influenced by pressure at low temperatures. Most of these residues are found in the $\beta 1/\beta 2$-sheet (Q2, I3, V5, T10, and I13), thus in the region where an effect on the $^{3h}J_{\text{NC}'}$ coupling constant can be deduced from the amide proton chemical shift.

Those amino acids with an upfield shift, are less influenced by pressure at the lower temperature. In contrast, residues within the first half of the protein with a pressure-induced

downfield shift show a stronger effect at 5 °C. Thus, for the coupling constants in the first β-sheet and the α-helix of ubiquitin one can expect a smaller effect for H-bonds that are weakened with pressure and a larger effect for those where the coupling becomes stronger. Thus, for the N-terminal half of ubiquitin, residues that are stabilized by pressure are even more stabilized at the lower temperature and H-bonds that are weakened are less weakened at the lower temperature.

2.4.3 Conclusion

The pressure-induced changes in ubiquitin's hydrogen bond network have been described and compared to temperature-dependent differences. Both, the $^{3h}J_{NC'}$ coupling constants and the amide proton chemical shifts have been monitored over the range from 1 to 2500 bar at 35 °C.

For some hydrogen bonds a curved pressure-dependence has been observed. Assuming that at higher pressures the volume of the protein is lower, the initial increase in strength of hydrogen bond scalar couplings could be the result of protein compression, for example due to minimization of internal cavities and release of water from voids. Since the compression is limited, above a certain pressure water penetrates into the protein core and hydration leads to destabilization of the protein. In this context, intra-protein hydrogen bonds can be replaced by intermolecular H-bond between the protein and water. Although the pressure applied to ubiquitin seems to be not high enough to denature the protein, a decrease in coupling constants indicates a destabilization. This is in line with the general finding that pressure below $1000 - 1500$ bar stabilized proteins against thermal unfolding (Cioni & Strambini 1999), but higher pressure can lead to destabilization. In contrast to all earlier findings, a correlation between the pressure-induced changes in the $^{3h}J_{NC'}$ coupling constant and the amide proton chemical shift has been found only for the first β-sheet and the α-helix of ubiquitin. In the second part of the protein, no correlation is observed. This suggests that in the study of pressure-dependent changes in protein hydrogen bonds by observing chemical shifts, additional factors, such as angular-dependencies or electric field effects, have to be taken into account. Clearly, the sole measurement of amide proton chemical shifts is not sufficient to describe pressure effects on hydrogen bonds. Interestingly, a good correlation between the two parameters ($^{3h}J_{NC'}$ and amide proton chemical shift) has been found only in those region, where the secondary structure is mainly unchanged or even stabilized by pressure. In contrast, in regions that are destabilized the correlation is not valid.

Most hydrogen bonds in the N-terminal part of ubiquitin (β-sheet 1/2 and α-helix) increase in strength as is evident from the H-bond scalar couplings. The average streng-

thening in $^{3h}J_{NC'}$ for this part of the protein amounts to 0.04 Hz at 2500 bar considering only coupling constants that get stronger and 0.02 Hz taking all hydrogen bond scalar couplings into account. This corresponds to a shortening of the hydrogen bonds by 0.025 Å (0.012 Å), if angular dependencies are neglected.

In contrast, ubiquitin's C-terminal half shows only initially stronger cross-hydrogen bond coupling constants. Above 1200–1500 bar about 50% of the H-bonds in this part of the protein are weakened. Together with the attenuation of the $^{3h}J_{NC'}$ couplings between residues R42 and V70 this marks the most pressure-labile region in ubiquitin, namely the β-sheets 3/5 and 1/5.

Comparison to the temperature-dependent changes in the $^{3h}J_{NC'}$ coupling constants reveals, that most hydrogen bonds within ubiquitin become stronger under pressure up to 2500 bar and weaker with increasing temperature between 5 °C and 65 °C. Nevertheless, there are H-bonds for which the coupling constant becomes weaker with increasing pressure, thus marking points where the pressure and cold denaturation of the protein might start.

2.4.4 Materials and Methods

High-pressure sample cell and pressure generator

All experiments were carried out using commercial high-pressure NMR cells (Daedalus Innovations LLC) with an inner diameter of 3 mm and an active volume of 120 μl. Tube 1 was rated to 1500 bar and used in combination with a static pressure cell made of bronze. Tube 2 was rated to 2500 bar and used in an aluminium alloy static pressure cell. Both cells were connected to a high pressure generator (High Pressure Equipment Company) via a pressure line. This line as well as the pressure generator were filled with extra-low viscosity paraffin wax (Sigma-Aldrich product number 95369, CAS 8002-74-2).

NMR measurement and data analysis

All spectra were recorded on a 7.5 mM solution of uniformly $^{2}H/^{13}C/^{15}N$-labeled ubiquitin in 10 mM phosphate buffer, pH 6.5, supplemented with 10% D_2O. A sample volume of 250 μl was placed in the high-pressure NMR tube and overlaid with paraffin wax. NMR spectra were measured on a Bruker DRX 800 spectrometer, equipped with a triple-resonance, Z-gradient cryo-probe. NMR data were processed with the NMRPipe suite of programs (Delaglio *et al.* 1995) and analyzed using the program PIPP (Garrett *et al.* 1991).

The $^{3h}J_{NC'}$ *trans* H-bond scalar couplings were measured with a long-range water flip-

back 2D-HNCO experiment as described previously (Cordier *et al.* 1999, 2008). Long-range as well as reference spectra were recorded at eight pressures (1, 300, 600, 900, 1200, 1500, 2000, and 2500 bar) with a measurement time of 23.25 hours and 1.5 hours, respectively. Data matrices consisted of 80*(^{13}C', t_1) × 1024*(^{1}H^N, t_3) data points (where n* refers to the complex points) with acquisition times of 40 ms (^{13}C') and 85 ms (^{1}H^N). $^{3h}J_{NC'}$ couplings were calculated from the intensity ratio of cross- and reference peaks as described (Cordier *et al.* 1999, 2008). Errors were estimated on the basis of three distinct experiments at each pressure. Data reported here represent the mean of all measurements with their standard deviation.

2.5 Pressure- and temperature-dependence of hydrogen bonds in nucleic acids

2.5.1 Introduction

RNA hairpins are very important secondary structural elements, which are involved in RNA folding and intermolecular interactions. Tetraloops are the most abundant hairpin motif. The 14mer tetraloop GGCACUUCGGUGCC was used as a model RNA in NMR spectroscopy before (Fürtig *et al.* 2004).

Here, we describe the temperature- and pressure-dependent changes of the hydrogen bond scalar coupling constants, $^{2h}J_{NN}$, in this RNA as measured by NMR spectroscopy. The HNN-COSY experiment has been used to quantify the two-bond ^{15}N-^{15}N scalar coupling between the imino donor and the acceptor nuclei in A/U and G/C base pairs. This experiment is described in more detail in chapter 2.2. The hydrogen bond scalar coupling constants normally range between 5 and 11 Hz. Thus they are almost one order of magnitude larger than the $^{3h}J_{NC'}$ coupling constants observed in proteins.

The pressure-induced changes in hydrogen bond scalar coupling constants in the model RNA are compared with data obtained for the protein ubiquitin.
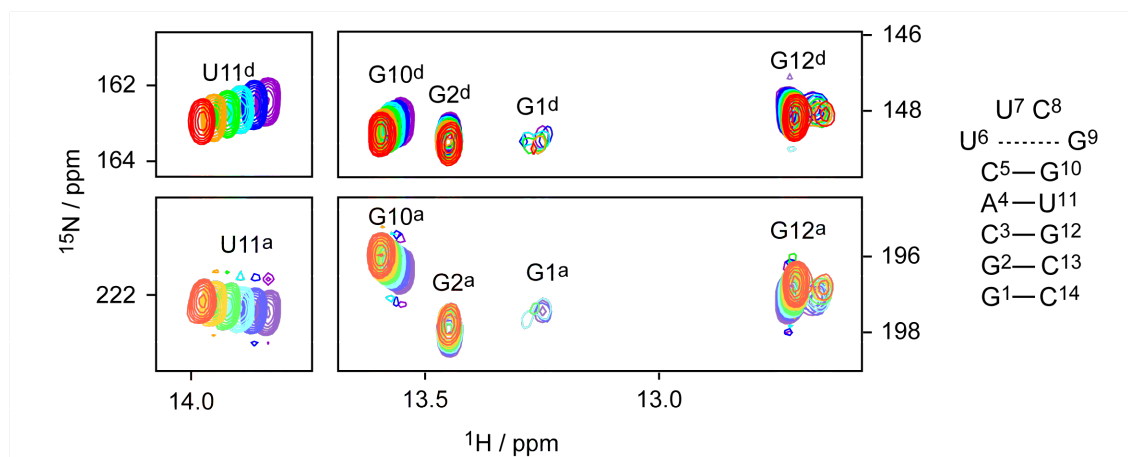
Figure 2.11   HNN-COSY spectra of the RNA hairpin shown on the right. Donor (top) and acceptor (bottom) resonances are shown for the five hydrogen bonded base pairs at six different pressures between 1 and 2500 bar at 25 °C: 1 (purple), 500 (blue), 1000 (turquoise), 1500 (green), 2000 (orange), and 2500 bar (red).

2.5.2 Results and Discussion

Influence of pressure on hydrogen bonds in RNA

Figure 2.11 shows the HNN-COSY spectra of the GGCACUUCGGUGCC RNA hairpin at 25 °C and different pressures ranging from 1 to 2500 bar. Obviously, the imino proton chemical shift changes are small for G1 and G2, slightly bigger for G10 and G12, and largest for U11. Figure 2.12 shows the $^{2h}J_{NN}$ coupling constant as a function of pressure at 25 and 35 °C. Experiments were performed at eleven pressures between 1 and 2500 bar ($\Delta p = 250$ bar) with three measurements per data point at 25 °C. At 35 °C not all experiments have been repeated three times (spectra at 1000 and 2000 bar have been recorded only once, no data are available for 250 bar).

The hydrogen bond between G1 and C14 cannot be detected easily, due to low signal intensities. This is the reason for the large error in the $^{2h}J_{NN}$ coupling constant for this H-bond. The low signal intensity is due to conformational exchange between a state with an open and another state with a closed hydrogen bond. For all other hydrogen bond scalar couplings measured, the error is much smaller (maximum error at 25 °C: 0.18 Hz; at 35 °C: 0.26 Hz). The coupling between G2 and C13 is weakened at elevated pressure, which indicates an opening of the end of the RNA stem. The couplings G12→C3 and U11→A4 increase with pressure, with the latter being overall larger. The last hydrogen bond before the loop, G10→C5, does not change significantly with pressures between 1 and 2500 bar. Interestingly, although the errors for the single pressure points are small, there are substantial fluctuations between different pressure points in the order of 0.2 to 0.3 Hz.

At higher temperature, the overall trends are the same, but the observed fluctuations are larger. Only the G2→C13 coupling constant shows a different pressure-dependent behavior with initially decreasing, than increasing and finally again decreasing values at the higher temperature. One reason for the observed fluctuations can be the large frequency separation between the ^{15}N donor and acceptor nuclei (3 − 4 kHz with a 600 MHz spectrometer). This can result in non-optimal excitation of the donor and acceptor nuclei.

Comparison of hydrogen bond scalar couplings in nucleic acids and proteins

The pressure-dependence of $^{3h}J_{NC'}$ coupling constants in ubiquitin as a model protein has been measured at 35 °C (see Chapter 2.4). For this protein, the hydrogen bond length decreases on average by 0.02 Å between 1 and 2500 bar.

In the RNA, the first two base pairs are influenced by fraying of the hairpin end. The average change in H-bond length for the three remaining base pairs is very similar to the
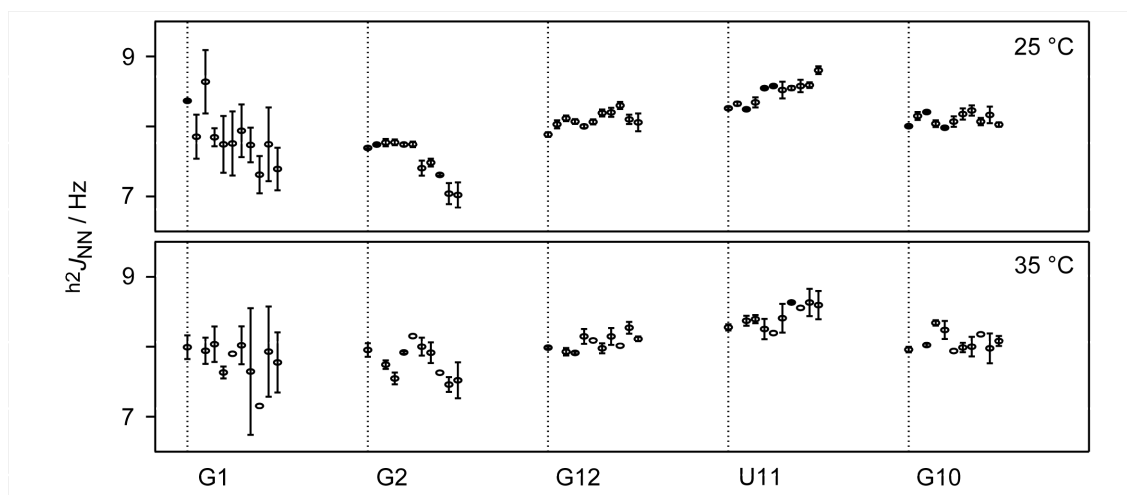
Figure 2.12 $^{2h}J_{NN}$ coupling constant as a function of pressure for the five hydrogen bonded base pairs in the 14mer RNA. Spectra were recorded at eleven equally spaced pressures between 1 and 2500 bar ($\Delta p = 250$ bar) at 25 and 35 °C. The data represent the mean and standard deviation of three independent experiments (at 35 °C only one set of data was recorded at 1000 and 2000 bar, no data were recorded at 250 bar).

difference observed in ubiquitin.

2.5.3 Materials and Methods

Uniformly $^{13}C/^{15}N$ labeled RNA (GGCACUUCGGUGCC) was purchased from Silantes GmbH (Munich, Germany). All measurements were carried out on a ≈ 0.5 mM RNA sample in 20 mM phosphate buffer, pH 6.4, 0.4 mM EDTA and 10% D_2O. The sample was heated to 70 °C and cooled rapidly on ice prior to the measurement to ensure hairpin formation of monomeric RNA. Spectra were acquired at different temperatures between 15 and 50 °C and at pressures between 1 and 2500 bar. All experiments were done in a 2500 bar – high-pressure NMR sample cell (tube2, see chapter 2.4) on a Bruker DRX 600 spectrometer equipped with a broadband probe with Z-gradient. Each spectrum consisted of 100* (t_1, ^{15}N) × 768* (t_2, $^1H^N$) complex data points with acquisition times of 17.3 ms (^{15}N) and 64 ms ($^1H^N$). The $^{15}N - {^{15}N}$ transfer period was set to 40 ms. NMR data were processed with the NMRPipe suite of programs (Delaglio *et al.* 1995) and analyzed using the program PIPP (Garrett *et al.* 1991).

2.6 The concerted effect of methanol and pressure on ubiquitin

2.6.1 Introduction

Many proteins adopt partially folded states under certain non-native conditions, such as extreme pH, presence of denaturants, or in organic solvents. Investigation of these states can provide information about the initial stages of protein folding. Ubiquitin exists in such a state in a mixture of water and methanol (40%/60%) at pH 2 and room temperature (Wilkinson & Mayer 1986). This so-called A-state of ubiquitin was investigated by CD (Wilkinson & Mayer 1986) and NMR spectroscopy (Brutscher *et al.* 1997). The NMR study revealed that the first β-sheet present in the native protein (M1 – V17) is preserved in the A-state and also the core of the α-helix is still present under the changed solvent condition. In contrats, the β-sheets 3, 4, and 5 change completely to a helix. This structural model of the methanol-induced A-state was confirmed by the measurement of hydrogen bond scalar couplings (Cordier & Grzesiek 2004). Although the A-state of ubiquitin has no stable hydrophobic core, it was shown that there is a well defined hydrogen bond network and thus secondary structural elements.

Application of pressure can lead to protein unfolding. Pressure-induced unfolding is induced by solvent, penetrating into the proteins core (Harano & Kinoshita 2006). The free energy of the protein-solvent system for the native state is lower at atmospheric pressure, but at sufficiently high pressure, a (partially) denatured state can be favored (Harano *et al.* 2008).

Here, the combined influence of methanol and pressure on ubiquitin is described. Most probably, the A-state is induced by disruption of the hydrophobic core in the presence of methanol. This hypothesis was tested by the influence of pressure on a sample containing not enough methanol to induce the A-state on its own.

2.6.2 Results and Discussion

Figure 2.13 shows the overlaid HSQC-spectra of ubiquitin in buffer containing 45% methanol at pH 4.6, measured at 1, 750, and 1500 bar. At 1 bar, only peaks corresponding to the native state of ubiquitin can be seen. If pressure is applied, new resonances appear and can be assigned to the A-state. At 1500 bar the ratio between the native and the partially folded state is $\approx 0.6 : 0.4$.

Normally, the A-state is induced by increasing the methanol content and for the given buffer conditions, the two states are equally populated at 56% methanol (Bobby 2008). Thus, pressure has a similar effect as larger amounts of methanol, indicating that the effective concentration of methanol is increased at relevant sites in ubiquitin under pressure.
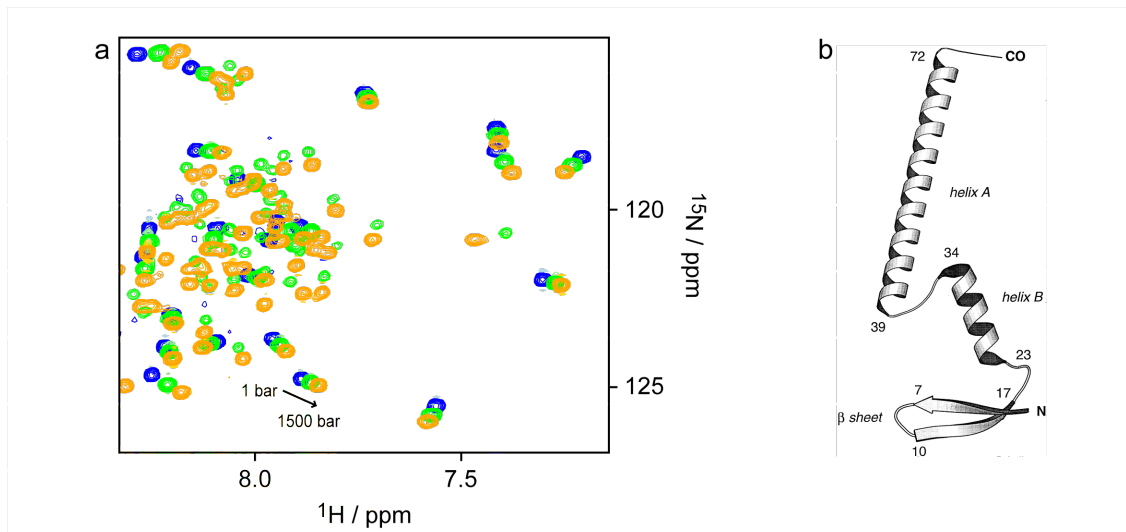
Figure 2.13   **a** Selected region of the ^{1}H-^{15}N HSQC spectra of ubiquitin in a solution containing 45% methanol at 1 bar (blue), 750 bar (green), and 1500 bar (yellow). Conditions: 250 μM ^{15}N-Ubiquitin, 20 mM acetate, pH 4.6, 45% methanol, 10% D$_2$O; 35 °C. With increasing pressure new peaks appear, which correspond to the A-state of ubiquitin. **b** Model of ubiquitin's A-state showing the preserved $\beta1/\beta2$-sheet and the α-helix (helix B). Helix A is unique in the A-state, the respective amino acids form a β-strand-rich region in native ubiquitin. Figure taken from (Brutscher *et al.* 1997).

While 45% methanol are not sufficient to induce the A-state in ubiquitin under the given conditions, pressure assists in partial unfolding. The A-state seems to be induced under pressure due to an increased packing density of the solvent due to exposure of hydrophobic groups in the A-state. This leads to an overall lower volume of the system. Using equation 1.3, a change in the volume difference between 750 and 1500 bar can be calculated to be $\Delta V_0 = 56$ Å^3. Higher pressures should lead to a complete conversion of the native to the A-state.

2.6.3 Materials and Methods

Experiments with ubiquitin in the presence of methanol were performed using a 250 μM sample of uniformly ^{15}N-labeled ubiquitin in 20 mM sodium acetate buffer, pH 4.6, with 45 % methanol and 10 % D$_2$O. All measurements were done in the 1500 bar – high-pressure NMR sample cell (tube 1, see chapter 2.4) at 35 °C. ^{1}H-^{15}N HSQC spectra were recorded with acquisition times of 50 ms and 85 ms for ^{15}N and ^{1}H, respectively. NMR data were processed with the NMRPipe suite of programs (Delaglio *et al.* 1995) and analyzed using the program PIPP (Garrett *et al.* 1991).
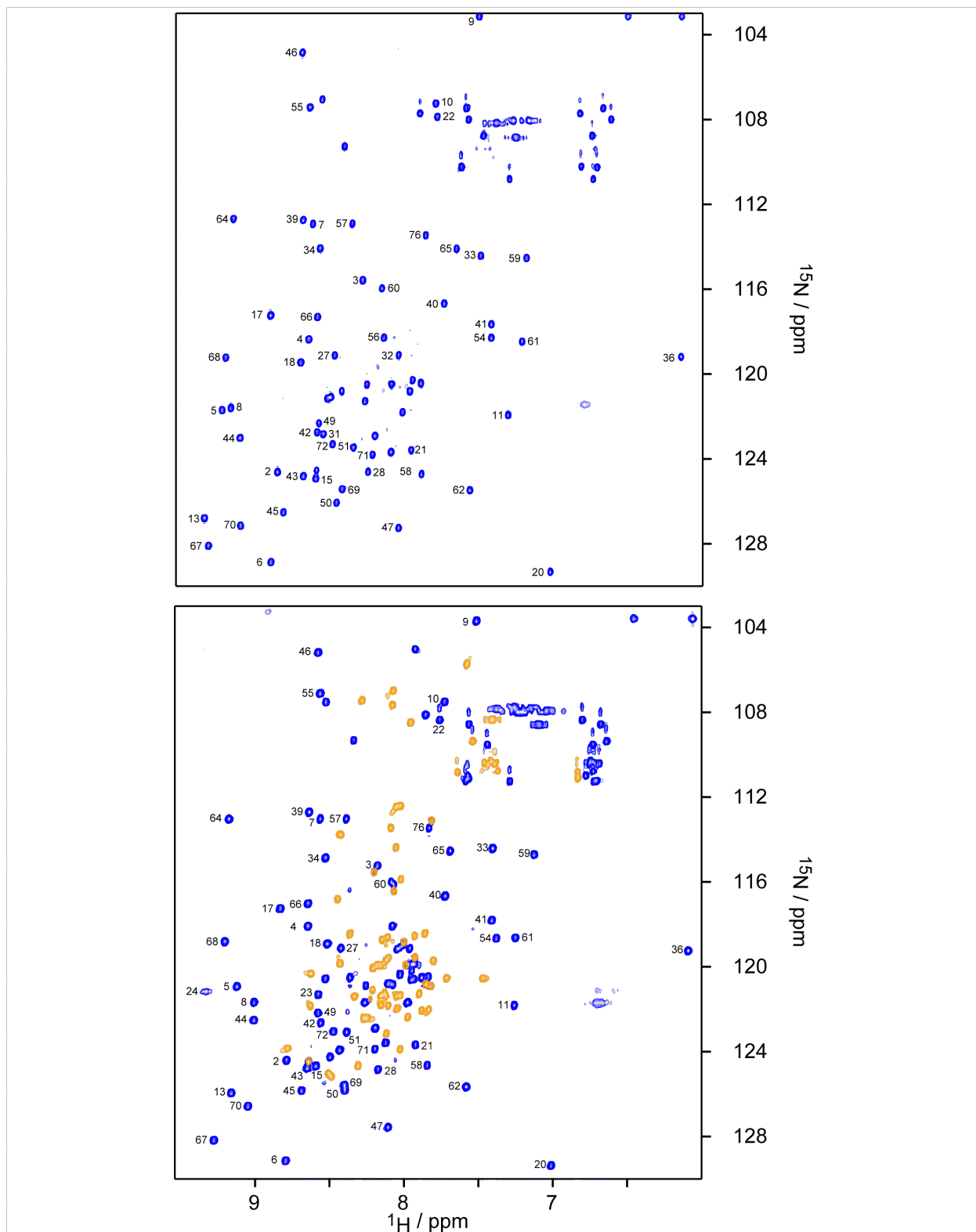
Figure 2.14   **a** ^{1}H-^{15}N HSQC spectrum of ubiquitin in a solution containing 45% methanol at 1 bar. Only the native state is visible. **b** After application of pressure (1500 bar), also the A-state appears. For clarity, resonances corresponding to the A-state are colored yellow. Resonances for the A-state have not been assigned. Conditions were as described for figure 2.13.

Bibliography

Adam, P., Gtlich, M., Oschkinat, H., Bacher, A., & Eisenreich, W. 2005. Studies of the intermediary metabolism in cultured cells of the insect Spodoptera frugiperda using ^{13}C- or ^{15}N-labelled tracers. *BMC Biochem*, **6**, 24.

Andrew, E.R., Bradbury, A., & Eades, R.G. 1958. Nuclear magnetic resonance spectra from a crystal rotated at speed. *Nature*, **182**, 1659.

Appay, V., Brown, A., Cribbes, S., Randle, E., & Czaplewski, L. G. 1999. Aggregation of RANTES is responsible for its inflammatory properties. Characterization of nonaggregating, noninflammatory RANTES mutants. *J Biol Chem*, **274**(39), 27505–27512.

Atkins, P. 2000. *Physical Chemistry*. Oxford University Press.

Baba, M., Nishimura, O., Kanzaki, N., Okamoto, M., Sawada, H., Iizawa, Y., Shiraishi, M., Aramaki, Y., Okonogi, K., Ogawa, Y., Meguro, K., & Fujino, M. 1999. A small-molecule, nonpeptide CCR5 antagonist with highly potent and selective anti-HIV-1 activity. *Proc Natl Acad Sci U S A*, **96**(10), 5698–5703.

Baggiolini, M., Dewald, B., & Moser, B. 1994. Interleukin-8 and related chemotactic cytokines–CXC and CC chemokines. *Adv Immunol*, **55**, 97–179.

Bannert, N., Craig, S., Farzan, M., Sogah, D., Santo, N. V., Choe, H., & Sodroski, J. 2001. Sialylated O-Glycans and Sulfated Tyrosines in the NH$_2$-Terminal Domain of CC Chemokine Receptor 5 Contribute to High Affinity Binding of Chemokines. *J. Exp. Med.*, **194**(11), 1661–1674.

Barfield, M. 2002. Structural Dependencies of Interresidue Scalar Coupling $^{h3}J_{NC'}$ and Donor ^{1}H Chemical Shifts in the Hydrogen Bonding Regions of Proteins. *Journal of the American Chemical Society*, **124**(15), 4158–4168.

Blanpain, C., Lee, B., Vakili, J., Doranz, B. J., Govaerts, C., Migeotte, I., Sharron, M., Dupriez, V., Vassart, G., Doms, R. W., & Parmentier, M. 1999. Extracellular cysteines of CCR5 are required for chemokine binding, but dispensable for HIV-1 coreceptor activity. *J Biol Chem*, **274**(27), 18902–18908.

Blanpain, C., Wittamer, V., Vanderwinden, J. M., Boom, A., Renneboog, B., Lee, B., Poul, E. Le, Asmar, L. El, Govaerts, C., Vassart, G., Doms, R. W., & Parmentier, M. 2001. Palmitoylation of CCR5 is critical for receptor trafficking and efficient activation of intracellular signaling pathways. *J Biol Chem*, **276**(26), 23795–23804.

Bobby, R. 2008. *Characterization of two salt bridges in Ubiquitin and Structural characterization of HES1*. M.Phil. thesis, University of Basel.

Brutscher, B., Bruschweiler, R., & Ernst, R. R. 1997. Backbone Dynamics and Structural Characterization of the Partially Folded A State of Ubiquitin by 1H, 13C, and 15N Nuclear Magnetic Resonance Spectroscopy;. *Biochemistry*, **36**(42), 13043–13053.

Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G. F., Thian, F., Kobilka, T., Choi, H., Kuhn, P., Weis, W. I., Kobilka, B. K., & Stevens, R. C. 2007. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science*, **318**(5854), 1258–1265.

Cioni, P., & Strambini, G. B. 1999. Pressure/temperature effects on protein flexibilty from acrylamide quenching of protein phosphorescence. *Journal of Molecular Biology*, **291**(4), 955–964.

Clark-Lewis, I., Kim, K. S., Rajarathnam, K., Gong, J. H., Dewald, B., Moser, B., Baggiolini, M., & Sykes, B. D. 1995. Structure-activity relationships of chemokines. *J Leukoc Biol*, **57**(5), 703–711.

Cocchi, F., DeVico, A. L., Garzino-Demo, A., Arya, S. K., Gallo, R. C., & Lusso, P. 1995. Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells. *Science*, **270**(5243), 1811–1815.

Cordier, F., & Grzesiek, S. 2002. Temperature-dependence of protein hydrogen bond properties as studied by high-resolution NMR. *J Mol Biol*, **317**(5), 739–752.

Cordier, F., & Grzesiek, S. 2004. Quantitative comparison of the hydrogen bond network of A-state and native ubiquitin by hydrogen bond scalar couplings. *Biochemistry*, **43**(35), 11295–11301.

Cordier, F., Rogowski, M., Grzesiek, S., & Bax, A. 1999. Observation of through-hydrogen-bond 2hJHC' in a perdeuterated protein. *J Magn Reson*, **140**(2), 510–512.

Cordier, F., Wang, C., Grzesiek, S., & Nicholson, L. K. 2000. Ligand-induced strain in hydrogen bonds of the c-Src SH3 domain detected by NMR. *J Mol Biol*, **304**(4), 497–505.

Cordier, F., Nisius, L., .Dingley, A. J, & Grzesiek, S. 2008. Direct detection of N-H···O=C hydrogen bonds in biomolecules by NMR spectroscopy. *Nat Protoc*, **3**(2), 235–241.

Cornilescu, G., Ramirez, B. E., Frank, M. K., Clore, G. M., Gronenborn, A. M., & Bax, A. 1999a. Correlation between $^{3h}J_{NC}$ and Hydrogen Bond Length in Proteins. *Journal of the American Chemical Society*, **121**(26), 6275–6279.

Cornilescu, G., Hu, J., & Bax, A. 1999b. Identification of the Hydrogen Bonding Network in a Protein by Scalar Couplings. *Journal of the American Chemical Society*, **121**(12), 2949–2950.

Czaplewski, L. G., McKeating, J., Craven, C. J., Higgins, L. D., Appay, V., Brown, A., Dudgeon, T., Howard, L. A., Meyers, T., Owen, J., Palan, S. R., Tan, P., Wilson, G., Woods, N. R., Heyworth, C. M., Lord, B. I., Brotherton, D., Christison, R., Craig, S.,

Cribbes, S., Edwards, R. M., Evans, S. J., Gilbert, R., Morgan, P., Randle, E., Schofield, N., Varley, P. G., Fisher, J., Waltho, J. P., & Hunter, M. G. 1999. Identification of amino acid residues critical for aggregation of human CC chemokines macrophage inflammatory protein MIP-1alpha, MIP-1beta, and RANTES. Characterization of active disaggregated chemokine variants. *J Biol Chem*, **274**(23), 16077–16084.

De Vries, L., Zheng, B., Fischer, T., Elenko, E., & Farquhar, M. G. 2000. The Regulator of G Protein Signaling Family. *Annu. Rev. Pharmacol. Toxicol.*, **40**(1), 235–.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., & Bax, A. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*, **6**(3), 277–293.

Dingley, A. J., & Grzesiek, S. 1998. Direct Observation of Hydrogen Bonds in Nucleic Acid Base Pairs by Internucleotide $^2J_{NN}$ Couplings. *Journal of the American Chemical Society*, **120**(33), 8293–8297.

Dingley, A. J., Nisius, L., Cordier, F., & Grzesiek, S. 2008. Direct detection of N-H···N hydrogen bonds in biomolecules by NMR spectroscopy. *Nat Protoc*, **3**(2), 242–248.

Dorr, P., Westby, M., Dobbs, S., Griffin, P., Irvine, B., Macartney, M., Mori, J., Rickett, G., Smith-Burchnell, C., Napier, C., Webster, R., Armour, D., Price, D., Stammen, B., Wood, A., & Perros, M. 2005. Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity. *Antimicrob Agents Chemother*, **49**(11), 4721–4732.

Drews, M., Doverskog, M., Ohman, L., Chapman, B. E., Jacobsson, U., Kuchel, P. W., & Hggstrm, L. 2000. Pathways of glutamine metabolism in Spodoptera frugiperda (Sf9) insect cells: evidence for the presence of the nitrogen assimilation system, and a metabolic switch by ^{1}H/ ^{15}N NMR. *Journal of Biotechnology*, **78**(1), 23–37.

Duma, L., Häussinger, D., Rogowski, M., Lusso, P., & Grzesiek, S. 2007. Recognition of RANTES by extracellular parts of the CCR5 receptor. *J Mol Biol*, **365**(4), 1063–1075.

Etzkorn, M., Martell, S., Andronesi, O.C., Seidel, K., Engelhard, M., & Baldus, M. 2007. Secondary Structure, Dynamics, and Topology of a Seven-Helix Receptor in Native Membranes, Studied by Solid-State NMR Spectroscopy. *Angewandte Chemie International Edition*, **46**(3), 459–462.

Farzan, M., Mirzabekov, T., Kolchinsky, P., Wyatt, R., Cayabyab, M., Gerard, N. P., Gerard, C., Sodroski, J., & Choe, H. 1999. Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. *Cell*, **96**(5), 667–676.

Fernández, C., Hilty, C., Wider, G., Gntert, P., & Wthrich, K. 2004. NMR structure of the integral membrane protein OmpX. *J Mol Biol*, **336**(5), 1211–1221.

Frye, K. J., & Royer, C. A. 1998. Probing the contribution of internal cavities to the volume change of protein unfolding under pressure. *Protein Sci*, **7**(10), 2217–2222.

Fürtig, B., Richter, C., Bermel, W., & Schwalbe, H. 2004. New NMR experiments for RNA nucleobase resonance assignment and chemical shift analysis of an RNA UUCG tetraloop. *J Biomol NMR*, **28**(1), 69–79.

Gaertner, H., Cerini, F., Escola, J.M., Kuenzi, G., Melotti, A., Offord, R., Rossitto-Borlat, I., Nedellec, R., Salkowitz, J., Gorochov, G., Mosier, D., & Hartley, O. 2008. Highly potent, fully recombinant anti-HIV chemokines: reengineering a low-cost microbicide. *Proc Natl Acad Sci U S A*, **105**(46), 17706–17711.

Gardner, K. H., & Kay, L. E. 1998. The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct*, **27**, 357–406.

Garrett, D. S., Powers, R., Gronenborn, A. M., & Clore, G. M. 1991. A common sense approach to peak picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. *Journal of Magnetic Resonance (1969)*, **95**(1), 214–220.

Gautier, A., Kirkpatrick, J. P., & D.Nietlispach. 2008. Solution-State NMR Spectroscopy of a Seven-Helix Transmembrane Protein Receptor: Backbone Assignment, Secondary Structure, and Dynamics. *Angewandte Chemie International Edition*, **47**(38), 7297–7300.

Goddard, T.D., & Kneller, D.G. SPARKY 3.

Gonen, T., Cheng, Y., Sliz, P., Hiroaki, Y., Fujiyoshi, Y., Harrison, S. C., & Walz, T. 2005. Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature*, **438**(7068), 633–638.

Goodman, O. B., Krupnick, J. G., Santini, F., Gurevich, V. V., Penn, R. B., Gagnon, A. W., Keen, J. H., & Benovic, J. L. 1996. Beta-arrestin acts as a clathrin adaptor in endocytosis of the beta2-adrenergic receptor. *Nature*, **383**(6599), 447–450.

Grzesiek, S., Cordier, F., Jaravine, V., & Barfield, M. 2004. Insights into biomolecular hydrogen bonds from hydrogen bond scalar couplings. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **45**(3-4), 275–300.

Harano, Y., & Kinoshita, M. 2006. On the physics of pressure denaturation of proteins. *Journal of Physics: Condensed Matter*, **18**, 107–113.

Harano, Y., Yoshidome, T., & Kinoshita, M. 2008. Molecular mechanism of pressure denaturation of proteins. *The Journal of Chemical Physics*, **129**, 145103.

Hartmann, & Hahn. 1962. *Physical Reviews*, **128**, 2042.

Henderson, R., & Unwin, P.N. 1975. Three-dimensional model of purple membrane obtined by electron microscopy. *Nature*, **257**, 28–32.

Henzler-Wildman, K., & Kern, D. 2007. Dynamic personalities of proteins. *Nature*, **450**(7172), 964–972.

Herzfeld, J., & Lansing, J. C. 2002. Magnetic resonance studies of the bacteriorhodopsin pump cycle. *Annu Rev Biophys Biomol Struct*, **31**, 73–95.

Hilf, R. J. C., & Dutzler, R. 2008. X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature*, **452**(7185), 375–379.

Hoover, D.M., Shaw, J., Gryczynski, Z., Proudfoot, A.E.I., & Wells, T. 2000. The crystal structure of MET-RANTES: Comparison with Native RANTES and AOP-RANTES. *Protein Peptide Letters*, **7**, 73–82.

Huang, C., Lam, S., Acharya, P., Tang, M., Xiang, S., Hussan, S. Shahzad-Ul, Stanfield, R. L., Robinson, J., Sodroski, J., Wilson, I. A., Wyatt, R., Bewley, C. A., & Kwong, P. D. 2007. Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. *Science*, **317**(5846), 1930–1934.

Inoue, K., Yamada, H., Akasaka, K., Herrmann, C., Kremer, W., Maurer, T., Dker, R., & Kalbitzer, H. R. 2000. Pressure-induced local unfolding of the Ras binding domain of RalGDS. *Nat Struct Biol*, **7**(7), 547–550.

Jaakola, V. P., Griffith, M. T., Hanson, M. A., Cherezov, V., Chien, E. Y. T., Lane, J. R., IJzerman, A. P., & Stevens, R. C. 2008. The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist. *Science*, **322**(5905), 1211–1217.

Jaravine, V. A., Alexandrescu, A. T., & Grzesiek, S. 2001. Observation of the closing of individual hydrogen bonds during TFE-induced helix formation in a peptide. *Protein Science*, **10**(5), 943–950.

Jonas, J. 2002. High-resolution nuclear magnetic resonance studies of proteins. *Biochim Biophys Acta*, **1595**(1-2), 145–159.

Jonas, J., Ballard, L., & Nash, D. 1998. High-resolution, high-pressure NMR studies of proteins. *Biophys J*, **75**(1), 445–452.

Kamatari, Y. O., Yamada, H., Akasaka, K., Jones, J. A., Dobson, C. M., & Smith, L. J. 2001. Response of native and denatured hen lysozyme to high pressure studied by ^{15}N/^{1}H NMR spectroscopy. *Eur J Biochem*, **268**(6), 1782–1793.

Kamatari, Y. O., Kitahara, R., Yamada, H., Yokoyama, S., & Akasaka, K. 2004. High-pressure NMR spectroscopy for characterizing folding intermediates and denatured states of proteins. *Methods*, **34**(1), 133–143.

Kitahara, R., Yokoyama, S., & Akasaka, K. 2005. NMR snapshots of a fluctuating protein structure: ubiquitin at 30 bar-3 kbar. *J Mol Biol*, **347**(2), 277–285.

Kitahara, R., Okuno, A., Kato, M., Taniguchi, Y., Yokoyama, S., & Akasaka, K. 2006. Cold denaturation of ubiquitin at high pressure. *Magnetic Resonance in Chemistry*, **44**(S1), S108–S113.

Klein-Seetharaman, J., Getmanova, E. V., Loewen, M. C., Reeves, P. J., & Khorana, H. G. 1999. NMR spectroscopy in studies of light-induced structural changes in mammalian rhodopsin: applicability of solution ^{19}F NMR. *Proc Natl Acad Sci U S A*, **96**(24), 13744–13749.

Klein-Seetharaman, J., Yanamala, N. V. K., Javeed, F., Reeves, P. J., Getmanova, E. V., Loewen, M. C., Schwalbe, H., & Khorana, H. G. 2004. Differential dynamics in the G protein-coupled receptor rhodopsin revealed by solution NMR. *Proc Natl Acad Sci U S A*, **101**(10), 3409–3413.

Kobilka, B. K. 2007. G protein coupled receptor structure and activation. *Biochim Biophys Acta*, **1768**(4), 794–807.

Kondru, R., Zhang, J., Ji, C., Mirzadegan, T., Rotstein, D., Sankuratri, S., & Dioszegi, M. 2008. Molecular interactions of CCR5 with major classes of small-molecule anti-HIV CCR5 antagonists. *Mol Pharmacol*, **73**(3), 789–800.

Kuwata, K., Li, H., Yamada, H., Legname, G., Prusiner, S. B., Akasaka, K., & James, T. L. 2002. Locally disordered conformer of the hamster prion protein: a crucial intermediate to PrPSc? *Biochemistry*, **41**(41), 12277–12283.

Lacapere, J. J., Pebay-Peyroula, E., Neumann, J. M., & Etchebest, C. 2007. Determining membrane protein structures: still a challenge! *Trends Biochem Sci*, **32**(6), 259–270.

Landau, E.M., & Rosenbusch, J.P. 1996. Lipidic cubic phases: A novel concept for the crystallization of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **93**(25), 14532–14535.

Lange, A., Giller, K., Hornig, S., Martin-Eauclaire, M. F., O.Pongs, Becker, S., & Baldus, M. 2006. Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. *Nature*, **440**(7086), 959–962.

Lederman, M. M., Veazey, R.S., Offord, R., Mosier, D. E., Dufour, J., Mefford, M., Piatak, M., Lifson, J. D., Salkowitz, J. R., Rodriguez, B., Blauvelt, A., & Hartley, O. 2004. Prevention of vaginal SHIV transmission in rhesus macaques through inhibition of CCR5. *Science*, **306**(5695), 485–487.

Lederman, M. M., Penn-Nicholson, A., Cho, M., & Mosier, D. 2006. Biology of CCR5 and its role in HIV infection and treatment. *JAMA*, **296**(7), 815–826.

Li, H., & Akasaka, K. 2006. Conformational fluctuations of proteins revealed by variable pressure NMR. *Biochim Biophys Acta*, **1764**(3), 331–345.

Li, H., Yamada, H., Akasaka, K., & Gronenborn, A. M. 2000. Pressure alters electronic orbital overlap in hydrogen bonds. *J Biomol NMR*, **18**(3), 207–216.

Liu, R., Paxton, W. A., Choe, S., Ceradini, D., Martin, S. R., Horuk, R., MacDonald, M. E., Stuhlmann, H., Koup, R. A., & Landau, N. R. 1996. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell*, **86**(3), 367–377.

Lodowski, D. T., & Palczewski, K. 2009. Chemokine receptors and other G protein-coupled receptors. *Curr Opin HIV AIDS*, **4**(2), 88–95.

Lopez, J. J., Shukla, A. K., Reinhart, C., Schwalbe, H., H.Michel, & Glaubitz, C. 2008. The structure of the neuropeptide bradykinin bound to the human G-protein coupled receptor bradykinin B2 as determined by solid-state NMR spectroscopy. *Angew Chem Int Ed Engl*, **47**(9), 1668–1671.

Luca, S., White, J. F., Sohal, A. K., Filippov, D. V., van Boom, J. H., Grisshammer, R., & Baldus, M. 2003. The conformation of neurotensin bound to its G protein-coupled receptor. *Proc Natl Acad Sci U S A*, **100**(19), 10706–10711.

Luca, S., Heise, H., Lange, A., & Baldus, M. 2005. Investigation of ligand-receptor systems by high-resolution solid-state NMR: recent progress and perspectives. *Arch Pharm (Weinheim)*, **338**(5-6), 217–228.

Lyukmanova, E. N., Shenkarev, Z. O., Paramonov, A. S., Sobol, A. G., Ovchinnikova, T. V., Chupin, V. V., Kirpichnikov, M. P., Blommers, M. J. J., & Arseniev, A. S. 2008. Lipid-protein nanoscale bilayers: a versatile medium for NMR investigations of membrane proteins and membrane-active peptides. *J Am Chem Soc*, **130**(7), 2140–2141.

Magnani, F., Shibata, Y., Serrano-Vega, M. J., & Tate, C.r .G. 2008. Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor. *Proc Natl Acad Sci U S A*, **105**(31), 10744–10749.

McDonald, C. B., Seldeen, K. L., Deegan, B. J., Lewis, M. S., & Farooq, A. 2008. Grb2 adaptor undergoes conformational change upon dimerization. *Archives of Biochemistry and Biophysics*, **475**, 25–35.

McDonald, P. H., Chow, C. W., Miller, W. E., Laporte, S. A., Field, M. E., Lin, F. T., Davis, R. J., & Lefkowitz, R. J. 2000. Beta-arrestin 2: a receptor-regulated MAPK scaffold for the activation of JNK3. *Science*, **290**(5496), 1574–1577.

Mirzabekov, T., Kontos, H., Farzan, M., Marasco, W., & Sodroski, J. 2000. Paramagnetic proteoliposomes containing a pure, native, and oriented seven-transmembrane segment protein, CCR5. *Nat Biotechnol*, **18**(6), 649–654.

Mustafi, D., & Palczewski, K. 2009. Topology of Class A G Protein-Coupled Receptors: Insights Gained from Crystal Structures of Rhodopsins, Adrenergic and Adenosine Receptors. *Mol Pharmacol*, **75**(1), 1–12.

Nardese, V., Longhi, R., Polo, S., Sironi, F., Arcelloni, C., Paroni, R., DeSantis, C., Sarmientos, P., Rizzi, M., Bolognesi, M., Pavone, V., & Lusso, P. 2001. Structural determinants of CCR5 recognition and HIV-1 blockade in RANTES. *Nat Struct Biol*, **8**(7), 611–615.

Neal, S., Nip, A. M., Zhang, H., & Wishart, D.S. 2003. Rapid and accurate calculation of protein ^{1}H, ^{13}C and ^{15}N chemical shifts. *Journal of Biomolecular NMR*, **26**(3), 215–240.

Nisius, L., Rogowski, M., Vangelista, L., & Grzesiek, S. 2008. Large-scale expression and purification of the major HIV-1 coreceptor CCR5 and characterization of its interaction with RANTES. *Protein Expr Purif*, **61**(2), 155–162.

Oberlin, E., Amara, A., Bachelerie, F., Bessia, C., Virelizier, J. L., Arenzana-Seisdedos, F., Schwartz, O., Heard, J. M., Clark-Lewis, I., Legler, D. F., Loetscher, M., Baggiolini, M., & Moser, B. 1996. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature*, **382**(6594), 833–835.

Opella, S. J., & Marassi, F. M. 2004. Structure determination of membrane proteins by NMR spectroscopy. *Chem Rev*, **104**(8), 3587–3606.

Opella, S. J., Marassi, F. M., Gesell, J. J., Valente, A. P., Kim, Y., Oblatt-Montal, M., & Montal, M. 1999. Structures of the M2 channel-lining segments from nicotinic acetylcholine and NMDA receptors by NMR spectroscopy. *Nat Struct Biol*, **6**(4), 374–379.

Oppermann, M. 2004. Chemokine receptor CCR5: insights into structure, function, and regulation. *Cell Signal*, **16**(11), 1201–1210.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Trong, I. Le, Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., & Miyano, M. 2000. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*, **289**(5480), 739–745.

Park, S., Mrse, A. A., Nevzorov, A. A., Mesleh, M. F, Oblatt-Montal, M., Montal, M., & Opella, S. J. 2003. Three-dimensional structure of the channel-forming transmembrane domain of virus protein u (Vpu) from HIV-1. *J Mol Biol*, **333**(2), 409–424.

Park, S. H., Prytulla, S., Angelis, A. A. De, Brown, J. M., Kiefer, H., & Opella, S. J. 2006. High-resolution NMR spectroscopy of a GPCR in aligned bicelles. *J Am Chem Soc*, **128**(23), 7402–7403.

Pervushin, K., Riek, R., Wider, G., & Wthrich, K. 1997. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A*, **94**(23), 12366–12371.

Pervushin, K., Ono, A., Fernndez, C., Szyperski, T., Kainosho, M., & Wthrich, K. 1998. NMR scalar couplings across Watson-Crick base pair hydrogen bonds in DNA observed by transverse relaxation-optimized spectroscopy. *Proc Natl Acad Sci U S A*, **95**(24), 14147–14151.

Pierce, K. L., Premont, R. T., & Lefkowitz, R. J. 2002. Seven-transmembrane receptors. *Nat Rev Mol Cell Biol*, **3**(9), 639–650.

Rasmussen, S. G. F., Choi, H., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M., Ratnala, V. R. P., Sanishvili, R., Fischetti, R. F., Schertler, G. F. X., Weis, W. I., & Kobilka, B. K. 2007. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature*, **450**(7168), 383–387.

Rastogi, V. K., & Girvin, M. E. 1999. Structural changes linked to proton translocation by subunit c of the ATP synthase. *Nature*, **402**(6759), 263–268.

Riek, R., Wider, G., Pervushin, K., & Wüthrich, Kurt. 1999. Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. *Proceedings of the National Academy of Sciences of the United States of America*, **96**(9), 4918–4923.

Rigaud, J. L., Mosser, G., Lacapere, J. J., Olofsson, A., Levy, D., & Ranck, J. L. 1997. Bio-Beads: an efficient strategy for two-dimensional crystallization of membrane proteins. *J Struct Biol*, **118**(3), 226–235.

Robinson, C. V., Sali, A., & Baumeister, W. 2007. The molecular sociology of the cell. *Nature*, **450**(7172), 973–982.

Rose, G. D., Fleming, P. J., Banavar, J. R., & Maritan, A. 2006. A backbone-based theory of protein folding. *Proceedings of the National Academy of Sciences*, **103**(45), 16623–16633.

Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H., Yao, X., Weis, W. I., Stevens, R. C., & Kobilka, B. K. 2007. GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science*, **318**(5854), 1266–1273.

Sanders, C. R., & Sönnichsen, F. 2006. Solution NMR of membrane proteins: practice and challenges. *Magn Reson Chem*, **44 Spec No**(Jul), S24–S40.

Sareth, S., Li, H., Yamada, H., Woodward, C. K., & Akasaka, K. 2000. Rapid internal dynamics of BPTI is insensitive to pressure. ^{15}N spin relaxation at 2 kbar. *FEBS Lett*, **470**(1), 11–14.

Schall, T. J., & Bacon, K. B. 1994. Chemokines, leukocyte trafficking, and inflammation. *Curr Opin Immunol*, **6**(6), 865–873.

Schlyer, S., & Horuk, R. 2006. I want a new drug: G-protein-coupled receptors in drug development. *Drug Discov Today*, **11**(11-12), 481–493.

S.Hiller, Garces, R. G., Malia, T. J., Orekhov, V. Y., Colombini, M., & Wagner, G. 2008. Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science*, **321**(5893), 1206–1210.

Signorell, G. A., Kaufmann, T. C., Kukulski, W., Engel, A., & Rémigy, H. W. 2007. Controlled 2D crystallization of membrane proteins using methyl-beta-cyclodextrin. *J Struct Biol*, **157**(2), 321–328.

Skelton, N. J., Aspiras, F., Ogez, J., & Schall, T. J. 1995. Proton NMR assignments and solution conformation of RANTES, a chemokine of the C-C type. *Biochemistry*, **34**(16), 5329–5342.

Smeller, L. 2002. Pressure-temperature phase diagrams of biomolecules. *Biochim Biophys Acta*, **1595**(1-2), 11–29.

Sodhi, A., Montaner, S., & Gutkind, J. S. 2004. Viral hijacking of G-protein-coupled-receptor signalling networks. *Nat Rev Mol Cell Biol*, **5**(12), 998–1012.

Strauss, A., Bitsch, F., Cutting, B., Fendrich, G., Graff, P., Liebetanz, J., Zurini, M., & Jahnke, W.g. 2003. Amino-acid-type selective isotope labeling of proteins expressed in Baculovirus-infected insect cells useful for NMR studies. *J Biomol NMR*, **26**(4), 367–372.

Takeda, S., Kadowaki, S., Haga, T., Takaesu, H., & Mitaku, S. 2002. Identification of G protein-coupled receptor genes from the human genome sequence. *FEBS Lett*, **520**(1-3), 97–101.

Veldkamp, C. T., Seibert, C., Peterson, F. C., De la Cruz, N. B., Haugner, J. C., III, Basnet, H., Sakmar, T. P., & Volkman, B. F. 2008. Structural Basis of CXCR4 Sulfotyrosine Recognition by the Chemokine SDF-1/CXCL12. *Sci. Signal.*, **1**(37), ra4–.

von Heijne, G. 1996. Principles of membrane protein assembly and structure. *Progress in Biophysics and Molecular Biology*, **66**(2), 113–139.

von Heijne, G. 2007. The membrane protein universe: what's out there and why bother? *J Intern Med*, **261**(6), 543–557.

Wang, J., Kim, S., Kovacs, F., & Cross, T. A. 2001. Structure of the transmembrane region of the M2 protein H(+) channel. *Protein Sci*, **10**(11), 2241–2250.

Warne, T., Serrano-Vega, M. J., Baker, J. G., Moukhametzianov, R., Edwards, P. C., Henderson, R., Leslie, A. G. W., Tate, C. G., & Schertler, G. F. X. 2008. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature*, **454**(7203), 486–491.

Werner, K., Lehner, I., Dhiman, H., Richter, C., Glaubitz, C., Schwalbe, H., Klein-Seetharaman, J., & Khorana, H. 2007. Combined solid state and solution NMR studies of a,e- ^{15}N labeled bovine rhodopsin. *Journal of Biomolecular NMR*, **37**(4), 303–312.

Werner, K., Richter, C., Klein-Seetharaman, J., & Schwalbe, H. 2008. Isotope labeling of mammalian GPCRs in HEK293 cells and characterization of the C-terminus of bovine rhodopsin by high resolution liquid NMR spectroscopy. *J Biomol NMR*, **40**(1), 49–53.

Wilkinson, K. D., & Mayer, A. N. 1986. Alcohol-induced conformational changes of ubiquitin. *Arch Biochem Biophys*, **250**(2), 390–399.

Wise, A., Gearing, K., & Rees, S. 2002. Target validation of G-protein coupled receptors. *Drug Discov Today*, **7**(4), 235–246.

Wood, A., & Armour, D. 2005. The discovery of the CCR5 receptor antagonist, UK-427,857, a new agent for the treatment of HIV infection and AIDS. *Prog Med Chem*, **43**, 239–271.

Wüthrich, K. 1986. *NMR of Proteins and Nucleic acids*. John Wiley & Sons.

Yagi, H., Tsujimoto, T., Yamazaki, T., Yoshida, M., & Akutsu, H. 2004. Conformational change of H+-ATPase beta monomer revealed on segmental isotope labeling NMR spectroscopy. *J Am Chem Soc*, **126**(50), 16632–16638.

Zhang, J., Peng, X., Jonas, A., & Jonas, J. 1995. NMR study of the cold, heat, and pressure unfolding of ribonuclease A. *Biochemistry*, **34**(27), 8631–8641.

Acknowledgments