

Internal Dynamics of the Homotrimeric HIV-1 Viral Coat Protein gp41 on Multiple Time Scales**

Nils-Alexander Lakomek, Joshua D. Kaufman, Stephen J. Stahl, John M. Louis, Alexander Grishaev, Paul T. Wingfield, and Ad Bax*

The fusion of viral and cellular membranes is elicited by the human immunodeficiency virus 1 (HIV-1) envelope glycoprotein gp120/gp41. The precursor gp160, encoded by the *Env* gene, is cleaved post-translationally into two chains, gp120 and gp41, which remain noncovalently associated as a homotrimer of heterodimers and form a spike on the viral surface.^[1] Upon binding to the CD4 coreceptor on the surface of the human T cell, the gp120 subunit dissociates from gp41, which remains anchored through its C-terminal transmembrane helix (TM) in the viral membrane. After gp120 dissociation, the N-terminal fusion peptide of gp41 is exposed and can insert into the host-cell membrane.^[2] Because of its high sequence conservation and its accessibility to the humoral immune system,^[3] gp41 is an attractive drug target for antiviral therapy as well as a key protein in vaccine research.^[4,5]

In the current model, membrane fusion is driven by a conformational change in the gp41 ectodomain from an extended prefusion (or so-called prehairpin) intermediate during virus-to-host-cell docking to a trimer of hairpins that forms an antiparallel six-helical bundle (6HB) arrangement; in this way, viral and host-cell membranes are pulled into close juxtaposition to initiate hemifusion.^[6,7] This model is mainly based on X-ray crystal studies of the soluble gp41 ectodomain. These studies revealed that the N- and C-terminal heptad repeat regions (NHR and CHR) pack together as a tight antiparallel 6HB,^[8,9] which represents the

late-fusion or postfusion conformation of the gp41 ectodomain.^[6,10] The fusion mechanism itself is believed to be similar to that of the hemagglutinin system of influenza,^[11,12] which has been analyzed in much more detail, but no structural information on any prefusion or early-fusion intermediate of gp41 is yet available. Whereas the structures of the N-terminal fusion peptide (FP) and the membrane-proximal external region (MPER) have been studied extensively by solution and solid-state NMR spectroscopy as well as by EPR spectroscopy,^[13–18] structural information on full-length gp41 has remained elusive.

Herein, we present a solution NMR spectroscopic study on the structure and dynamics of the homotrimeric gp41 complex encompassing residues 1–194 (512–705 in the numbering of the *Env* precursor, see Figure 1) reconstituted in dodecylphosphatidylcholine (DPC) micelles. Li and Tamm

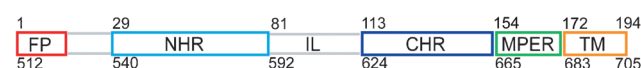


Figure 1. Schematic view of the gp41^{1–194} construct studied, including the fusion peptide (FP), N-terminal heptad repeat (NHR), immunodominant loop region (IL), C-terminal heptad repeat (CHR), membrane-proximal external region (MPER), and transmembrane helix (TM) anchoring gp41 to the viral envelope. The numbering 1–194 refers to the gp41^{1–194} construct; the numbering 512–705 refers to the location in the *Env* precursor.

showed very similar CD spectra, which thus indicated the same helicity, for the gp41 fusion peptide embedded in DPC micelles and in 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine/1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoglycerol (POPC/POPG; 4:1) lipid bilayers. The fusion peptide embedded in liposomes with the same lipid ratio also induced lipid mixing.^[14] Sun et al. observed very similar structures of the MPER peptide in DPC micelles and in 1,2-diheptanoyl-*sn*-glycero-3-phosphocholine/1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DHPC/DMPC) bicelles, and their EPR spectra for the MPER peptide in DPC and in virus-membrane-like liposomes were also similar.^[16] Therefore, DPC micelles appear to be a suitable mimic for a membrane environment.

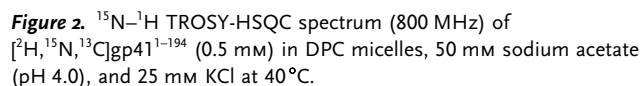
Sedimentation equilibrium centrifugation and size-exclusion chromatography with in-line multiangle light scattering and refractive-index measurements both showed a monomer/trimer equilibrium with a K_D value in the low micromolar range; this equilibrium will be fully shifted towards the homotrimer at the concentrations used for NMR spectroscopy. The total mass of the trimer-detergent complex was 181 kDa (see the Supporting Information).

[*] Dr. N. A. Lakomek, Dr. J. M. Louis, Dr. A. Grishaev, Dr. A. Bax
Laboratory of Chemical Physics, NIDDK
National Institutes of Health, Building 5, Room 126
9000 Rockville Pike, Bethesda, MD 20892-0520 (USA)
E-mail: bax@nih.gov

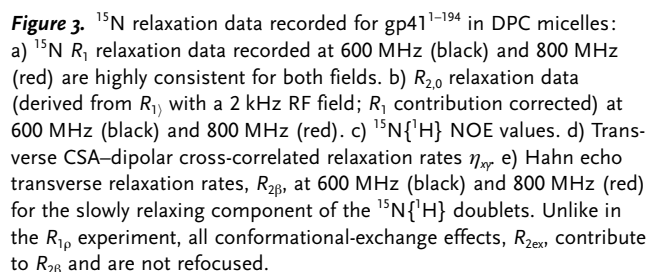
J. D. Kaufman, Dr. S. J. Stahl, Dr. P. T. Wingfield
Protein Expression Laboratory, NIAMS
National Institutes of Health, Building 6B, Room 1B130 (USA)

[**] We thank Yang Shen for expert advice and Jinfa Ying and James Baber for experimental assistance. This research was funded by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (NIH) and the Intramural AIDS-Targeted Antiviral Program of the Office of the Director, NIH. We acknowledge use of the shared scattering beamline resource allocated under the PUP-77 agreement between the National Cancer Institute and the Argonne National Laboratory and thank Dr. Soenke Seifert (Argonne National Laboratory) for his support of the SAXS experiments. Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science under Contract W-31-109-ENG-38. HIV = human immunodeficiency virus.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/ange.201207266>.



We therefore set out to characterize the backbone motions of the gp41^{1–194} trimer by ¹⁵N relaxation. ¹⁵N *R*₁ and *R*_{1ρ} relaxation rates as well as ¹⁵N{¹H} NOE and transverse ¹⁵N chemical-shift anisotropy (CSA)–dipolar cross-correlated relaxation rates, η_{xy} , were measured at both 600 and 800 MHz by TROSY-based methods optimized for perdeuterated proteins^[19] (Figure 3 a–d). Limited sensitivity did not permit the use of 3D methods,^[20] and only data for amides that could be resolved in the 2D ¹⁵N–¹H TROSY-HSQC spectrum are shown in Figure 3. Conformational changes on the micro- to millisecond time scale give rise to exchange contributions to *R*₂. These so-called *R*_{ex} contributions are best measured on the basis of the slowly relaxing ¹⁵N{¹H} doublet component by using a Hahn echo experiment (Figure 3 e; see the Supporting Information for details). The relaxation data showed random-coil-like behavior for the N-terminal residues V2 and G3 and high mobility for regions A15–L26 and L81–Q110, as indicated by smaller *R*_{2,0} (*R*_{1ρ}) and higher *R*₁ rates as well as lower ¹⁵N{¹H} NOE values. The most-rigid visible residues are located at the C terminus in the NHR region (I62–L76). The relaxation data can be clustered into regions of similar dynamics: I4–L12, Q29–R46, I62–L76, T18–



M24, and G86–N105. For a complete overview of the relaxation data in terms of the error-weighted average values and their standard deviations, see Table S2a in the

Supporting Information. Relative to the NHR region, the fully α -helical FP (assigned as residues I4–L12 on the basis of $^{13}\text{C}^\alpha$ secondary chemical shifts) showed lower R_2 ($R_{1\rho}$) and higher R_1 rates, which are indicative of large-amplitude rigid-body motions of the FP relative to the average orientation of the larger ectodomain. The chemical shifts of the most N-terminal 22 residues of gp41^{1–194} agree very closely with those reported previously^[13] for a short FP construct (residues 1–30) in sodium dodecyl sulfate (SDS) micelles (see Figure S9 in the Supporting Information), which indicates that they adopt very similar structures. Increased mobility and reduced α -helical propensity, which are indicative of a transient helical structure (see Figure S9a), were observed for the region stretching from G13 to L26 (compare Figure 3 and Table S2) and thus point to substantial flexibility in this linker region connecting the FP and the NHR.

The limits of the NHR (29–79) and CHR regions (113–155) have previously been deduced from proteolytic digestion studies,^[21] which resulted in proteolysis-resistant soluble gp41 fragments that were used for subsequent crystallization studies.^[9,22] The proteolytically sensitive residues 80–112, which were missing in all X-ray crystal-structure studies, were ascribed to a loop region. Indeed, our ^{15}N relaxation and $\Delta\delta^{13}\text{C}^\alpha$ data indicate that the mobile IL region stretches from L81 to Q110.

To describe the internal dynamics of gp41 and the motion of the FP relative to the NHR region, we analyzed relaxation data, including the transverse cross-correlated relaxation rate, η_{xy} , by the extended model-free approach^[23] as described in detail in the Supporting Information. Fitting of the data was only possible when an internal motion on the nanosecond time scale was invoked.^[24] Internal dynamics are characterized by a generalized order parameter, S^2 , and a correlation time, τ , both for the fast, picosecond (S_f^2 , τ_f) and slower, low-nanosecond time scale (S_s^2 , τ_s). As summarized in Table 1, all domains of gp41 showed a remarkably high degree of internal dynamics on a 2–5 ns time scale, as characterized by average order parameters ranging between $S_s^2 = 0.76$ for the most-rigid NMR-visible part of the NHR domain and $S_s^2 = 0.37$ for the linker region between FP and NHR. The FP itself (I4–L12), with an average order parameter of $S_s^2 = 0.42$, showed high-amplitude motion relative to the NHR region (Q29–R46 and I62–L76). When interpreted in terms of free diffusion of the helical FP in a cone model,^[25] this order parameter

corresponds to a cone semiangle of 42° . The amplitude of the fast motions (on the picosecond time scale), reflected in $S_f^2 = 0.77$, is comparable to that observed for globular proteins and thus confirms that the FP itself remains a well-ordered helix. By contrast, the linker and loop regions T18–M24 and G86–N105 showed strongly elevated dynamics on both picosecond and nanosecond time scales (Table 1).

As noted above, only about 110 spin systems out of 194 gave rise to correlations in the HNCO NMR spectrum, and all such spin systems belong to the FP, the NHR, or the IL region, whereas the CHR, MPER, and TM regions remained unobservable. For the highly hydrophobic TM region, incomplete $^1\text{H}^\text{N}$ back exchange after the expression of the protein in the solvent D_2O cannot be fully excluded, as this region could remain protected from the solvent by the presence of detergent at a low concentration (below the critical micelle concentration) during the purification process. However, the presence of detergent cannot explain the absence of the amide resonances for the CHR and MPER regions. The absence of these signals can therefore be attributed to strongly increased line widths that result in amplitudes below the observable threshold. Increased line widths are due to elevated ^1H and/or ^{15}N R_2 relaxation rates and can result from either slow tumbling in the absence of large internal motions or from conformational-exchange processes on the micro- to millisecond time scale, which add an exchange contribution, R_{ex} , to R_2 . Whereas conformational-exchange effects seem to be largely absent for the FP, linker, and NHR regions, with the exception of residues around A22 and N42, a strong exchange contribution was observed for the IL region around A96 (Figure 3e). All exchange effects disappeared when a $T_{1\rho}$ measurement with a spin-lock radio-frequency (RF) field of 2 kHz was used (Figure 3b); thus, the observed exchange process must have a time constant much longer than 80 μs . The observation that the N-proximal region of the NHR remained visible, although its internal dynamics do not significantly narrow its resonances, together with the existence of a strong conformational-exchange effect in the IL region indicate that the absence of signals for the CHR and MPER regions can be attributed to a conformational-exchange process on an intermediate NMR time scale.

The very different relaxation properties of the MPER (and presumably the TM) from those of the FP, which shows the most intense resonances, excludes a significant interaction between the FP and the TM. Paramagnetic relaxation enhancement (PRE) data recorded for gp41, tagged at the C terminus (position S192C), showed a small but consistent increase in the ^1H R_2 rate by $8 \pm 2 \text{ s}^{-1}$ for the N-terminal 25 residues, but no statistically significant change outside this region (see Figure S11). The small but consistent increase in the ^1H R_2 rate suggests that the FP and its linker to the NHR transiently sample conformations close to the TM region. However, when it is considered how steeply PRE effects scale with distance and how uniform the effects are across the fusion peptide and its linker, contacts between the FP and the TM never get close or specific and must have a low population. In all likelihood, the late-fusion-stage 6HB, for which models predict close proximity of the TM to the FP, is at most very low populated under the conditions of our study.

Table 1: Results of the extended model-free analysis of ^{15}N relaxation data.

| Cluster | τ_c [ns] ^[a] | S_f^2 ^[b] | τ_s [ns] ^[c] | S_s^2 ^[d] |
|----------|------------------------------|------------------------|------------------------------|------------------------|
| I4–L12 | 44 \pm 3 | 0.77 \pm 0.02 | 5.4 \pm 0.7 | 0.42 \pm 0.04 |
| Q29–R46 | 44 \pm 3 | 0.85 \pm 0.02 | 4.0 \pm 0.9 | 0.64 \pm 0.05 |
| I62–L76 | 44 \pm 3 | 0.90 \pm 0.04 | 3.1 \pm 0.9 | 0.76 \pm 0.03 |
| T18–M24 | 44 \pm 3 | 0.62 \pm 0.02 | 2.3 \pm 0.2 | 0.37 \pm 0.02 |
| G86–N105 | 44 \pm 3 | 0.67 \pm 0.03 | 2.5 \pm 0.7 | 0.48 \pm 0.04 |

[a] Rotational correlation time. [b] Order parameter for the fast time scale (order parameters can range between 0 and 1, whereby 0 corresponds to full mobility and 1 to complete rigidity). The correlation time of the motion on the fast time scale was fixed to 50 ps during the fitting.

[c] Correlation time of the internal motion on the slow time scale ($\ll \tau_c$).

[d] Order parameter for the slow time scale.

Small-angle X-ray scattering (SAXS) data show a pairwise distribution function with a maximum length vector in the 150–170 Å range (see Figure S5). The late-fusion-stage 6HB arrangement positions the FP and TM domains at the same end of this bundle, and the sole population of this state appears incompatible with the dimensions extracted from the scattering data. The extended prefusion three-helical bundle arrangement could reach a length exceeding 200 Å when fully extended and is also incompatible with the SAXS data. However, as the approximately 30 residue IL region is dynamically highly disordered, the relative orientation of the NHR and CHR domains will fluctuate. The SAXS data are compatible with an ensemble in which the CHR helices sample a bundle of orientations with an average interhelical angle of about 50° relative to the NHR (see the Supporting Information).

Thus, the ^{15}N -relaxation, PRE, and SAXS data are compatible with a prehairpin intermediate that samples a range of relative orientations of the CHR to the NHR, possibly in exchange with a low population of the late-fusion 6HB. When transitioning from the prehairpin intermediate (Figure 4A) to the late-fusion 6HB (Figure 4C), the CHR region is believed to either jackknife and pack against the NHR^[26] or zipper up along the inner NHR coiled coil.^[27] This transition would require breaking up of the CHR three-helical bundle and at least transient disorder of the CHR and MPER regions during the transition period (Figure 4B). Such a process is anticipated to have a high energy barrier and would dramatically impact the chemical shifts of the IL, CHR, MPER, and TM regions. The large exchange broadening observed for the CHR and MPER regions in our measurements is compatible with such a scenario. Interestingly, the high flexibility of the linker region connecting the FP and the NHR in the prefusion intermediate suggests that in the “harpoon model”,^[9,22] the main purpose of the FP during the

docking stage is to act as an anchor that is attached to or injected into the host-cell membrane, whereas stress on the membrane curvature,^[28] at least at this stage, is not transferred from the NHR to the FP and, if at all, must be generated by the FP itself. The flexibility of the linker region adjacent to the FP then enables the repositioning of the ectodomain during the transition to the 6HB by driving membrane juxtaposition and subsequent hemifusion stalk formation.

Experimental Section

NMR spectroscopic measurements were carried out on a uniformly $^2\text{H}/^{15}\text{N}/^{13}\text{C}$ -enriched sample of homotrimeric gp41^{1–194} (10 mg mL^{−1}, 0.5 mM) in 50 mM sodium acetate buffer (pH 4.0) with 25 mM KCl and 330 mM DPC (ca. 115 mg mL^{−1}) at 313 K. Comparison with a spectrum recorded at pH 7.1 in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and 200 mM DPC showed that the HSQC spectrum retains a very similar appearance over this pH range (see the Supporting Information), and all subsequent NMR spectroscopic measurements were conducted at pH 4, the pH value at which the protein was most stable and the spectra were not compromised by the exchange of amide hydrogen atoms with the solvent. Oligomerization of the construct was examined by sedimentation equilibrium ultracentrifugation and size-exclusion chromatography with in-line multi-angle light scattering. NMR spectroscopic backbone assignment was performed as described in the Supporting Information. Backbone dynamics were studied by ^{15}N NMR R_1 and $R_{1\rho}$ relaxation and $\{^1\text{H}\}^{15}\text{N}$ NOE spectroscopic experiments with a TROSY detection scheme^[19] as well as transverse ^{15}N CSA-dipolar cross-correlated relaxation (η_{xy}) measurements. The η_{xy} rates were measured as the difference in the relaxation rates of the NH doublet components exhibiting slow and fast relaxation in a TROSY-based experiment.^[29] The presence of conformational-exchange effects was studied by measuring the T_2 relaxation time of the slowly relaxing ^{15}N - $\{^1\text{H}\}$ doublet component in a simple but informative extension of the regular TROSY-HSQC experiment. ^1H R_2 rates were measured to study PRE effects for an *S*-(2,2,5,5-tetramethyl-2,5-dihydro-1*H*-pyrrol-3-yl)methyl methanesulfonothioate (MTSL) spin-labeled S192C mutant of gp41^{1–194} and the reference sample by using a TROSY-based detection scheme. Spectra were recorded on 600 MHz Bruker Avance II, 800 MHz Bruker Avance III, and 900 MHz Bruker Avance II spectrometers, all equipped with cryogenic-probe-head technology. Details of the experimental setup of all experiments can be found in the Supporting Information. Spectra were processed with the NMRPipe/NMRDraw software package.^[30] Secondary chemical shifts were corrected for ^2H isotope effects by using the random-coil values of Maltsev et al.^[31] ^{15}N relaxation and η_{xy} rates were fitted by using the extended model-free approach,^[23] implemented with a home-written MATLAB script.

Received: September 8, 2012

Revised: February 1, 2013

Published online: February 28, 2013

Keywords: glycoproteins · membrane proteins · NMR spectroscopy · protein dynamics · viruses

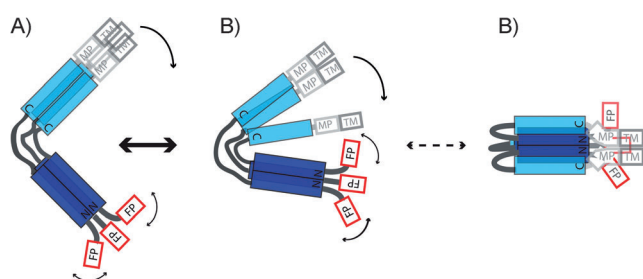


Figure 4. Homotrimeric gp41 shows a high degree of intrinsic mobility on the nanosecond and microsecond time scale. NMR and SAXS data are compatible with the sampling of a range of relative orientations of the CHR to the NHR by the prehairpin intermediate (A), possibly in exchange with a low population of the late-fusion 6HB (C). A conformational change from the prehairpin intermediate to the late-fusion 6HB requires the breaking up of the CHR three-helical bundle with concomitant disorder of the CHR and MPER region during the transition period (B). The large exchange broadening observed for the CHR and MPER region is indicative of such a scenario. No NMR spectroscopic information on the relative arrangement of the CHR, MPER, and TM domains is available, and the arrangement shown is purely schematic. Intense resonances observed for the FP exclude a strong interaction between the FP and the TM, although weak PRE interactions point to their transient proximity.

- [1] P. Zhu, J. Liu, J. Bess, E. Chertova, J. D. Lifson, H. Grise, G. A. Ofek, K. A. Taylor, K. H. Roux, *Nature* **2006**, *441*, 847.
- [2] Q. Sattentau, J. Moore, *J. Exp. Med.* **1991**, *174*, 407.
- [3] D. S. Dimitrov, R. Blumenthal, *J. Virol.* **1994**, *68*, 1956.
- [4] J. Münch, L. Ständker, K. Ademann, A. Schulz, M. Schindler, R. Chinnadurai, S. Pöhlmann, C. Chaipan, T. Biet, T. Peters, B. Meyer, D. Wilhelm, H. Lu, W. Jing, S. Jiang, W.-G. Forssmann, F. Kirchhoff, *Cell* **2007**, *129*, 263.

- [5] D. C. Douek, P. D. Kwong, G. J. Nabel, *Cell* **2006**, 124, 677.
- [6] S. C. Harrison, *Nat. Struct. Mol. Biol.* **2008**, 15, 690.
- [7] S. C. Harrison, *Virus Structure and Assembly*, Vol. 64, Academic Press, New York, **2005**, p. 231.
- [8] D. C. Chan, P. S. Kim, *Cell* **1998**, 93, 681.
- [9] W. Weissenhorn, A. Dessen, S. C. Harrison, J. J. Skehel, D. C. Wiley, *Nature* **1997**, 387, 426.
- [10] K. H. Roux, K. A. Taylor, *Curr. Opin. Struct. Biol.* **2007**, 17, 244.
- [11] I. A. Wilson, J. J. Skehel, D. C. Wiley, *Nature* **1981**, 289, 366.
- [12] J. J. Skehel, D. C. Wiley, *Annu. Rev. Biochem.* **2000**, 69, 531.
- [13] C. P. Jaroniec, J. D. Kaufman, S. J. Stahl, M. Viard, R. Blumenthal, P. T. Wingfield, A. Bax, *Biochemistry* **2005**, 44, 16167.
- [14] Y. L. Li, L. K. Tamm, *Biophys. J.* **2007**, 93, 876.
- [15] W. Qiang, Y. Sun, D. P. Weliky, *Proc. Natl. Acad. Sci. USA* **2009**, 106, 15314.
- [16] Z. Y. J. Sun, K. J. Oh, M. Y. Kim, J. Yu, V. Brusic, L. K. Song, Z. S. Qiao, J. H. Wang, G. Wagner, E. L. Reinherz, *Immunity* **2008**, 28, 52.
- [17] M. Kim, Z. Y. J. Sun, K. D. Rand, X. M. Shi, L. K. Song, Y. X. Cheng, A. F. Fahmy, S. Majumdar, G. Ofek, Y. P. Yang, P. D. Kwong, J. H. Wang, J. R. Engen, G. Wagner, E. L. Reinherz, *Nat. Struct. Mol. Biol.* **2011**, 18, 1235.
- [18] A. L. Lai, A. E. Moorthy, Y. L. Li, L. K. Tamm, *J. Mol. Biol.* **2012**, 418, 3.
- [19] N. A. Lakomek, J. F. Ying, A. Bax, *J. Biomol. NMR* **2012**, 53, 209.
- [20] M. Caffrey, J. Kaufman, S. J. Stahl, P. T. Wingfield, A. M. Gronenborn, G. M. Clore, *J. Magn. Reson.* **1998**, 135, 368.
- [21] M. Lu, S. C. Blacklow, P. S. Kim, *Nat. Struct. Biol.* **1995**, 2, 1075.
- [22] D. C. Chan, D. Fass, J. M. Berger, P. S. Kim, *Cell* **1997**, 89, 263.
- [23] G. M. Clore, A. Szabo, A. Bax, L. E. Kay, P. C. Driscoll, A. M. Gronenborn, *J. Am. Chem. Soc.* **1990**, 112, 4989.
- [24] D. M. Korzhnev, V. Y. Orekhov, A. S. Arseniev, *J. Magn. Reson.* **1997**, 127, 184.
- [25] G. Lipari, A. Szabo, *J. Am. Chem. Soc.* **1982**, 104, 4546.
- [26] R. Blumenthal, S. Durell, M. Viard, *J. Biol. Chem.* **2012**, 287, 40841.
- [27] G. Frey, S. Rits-Volloch, X. Q. Zhang, R. T. Schooley, B. Chen, S. C. Harrison, *Proc. Natl. Acad. Sci. USA* **2006**, 103, 13938.
- [28] Y. Kozlovsky, M. M. Kozlov, *Biophys. J.* **2002**, 82, 882.
- [29] D. Lee, C. Hilty, G. Wider, K. Wüthrich, *J. Magn. Reson.* **2006**, 178, 72.
- [30] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, 6, 277.
- [31] A. S. Maltsev, J. F. Ying, A. Bax, *J. Biomol. NMR* **2012**, 54, 181.