J|A|C|S
COMMUNICATIONS

Published on Web 02/23/2007

Amino Acid Bulkiness Defines the Local Conformations and Dynamics of Natively Unfolded α-Synuclein and Tau

Min-Kyu Cho,[†] Hai-Young Kim,[†] Pau Bernado,[¶] Claudio O. Fernandez,[§] Martin Blackledge,[¶] and Markus Zweckstetter*[,†,‡]

Department of NMR-Based Structural Biology, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany, DFG Research Center for the Molecular Physiology of the Brain, Institut de Biologie Structurale Jean-Pierre Ebel, CEA-CNRS-UJF UMR 5075, 41 Rue Jules Horowitz, Grenoble 38027, France, and Instituto de Biología Molecular y Celular de Rosario, Universidad Nacional de Rosario, Suipacha 531, S2002LRK Rosario, Argentina

Received October 19, 2006; E-mail: mzwecks@gwdg.de

Natively unfolded proteins play key roles in normal and pathological biochemical processes.[1] This category of proteins remains, however, beyond the reach of classical structural biology because of their inherent conformational heterogeneity. The overall properties of unfolded proteins are similar to the random coil state and are rather insensitive to the details of the amino acid sequence.[2] Spectroscopic measurements, however, suggest the presence of sequence-specific residual secondary and even tertiary structure in unfolded states of proteins.[3−8] NMR residual dipolar couplings (RDCs) are particularly sensitive probes for the structure and dynamics of biomolecules. When confined in weakly aligning media, unfolded proteins display surprisingly variable RDCs as a function of position along the chain, possibly even encoding the native topology.[5,8] This is in clear contrast to the bell-like smooth distribution of RDCs that is expected for a random flight chain[9] and was interpreted as alignment of extended or polyproline II conformations.[5] More recently, it was shown that RDCs in denatured proteins can be predicted from ensembles of unfolded structures that were generated by using a self-avoiding statistical coil model, which was based on residue-specific ϕ/φ propensities from loop regions of a folded protein database.[10,11] Here we show that a much simpler model can also explain many aspects of the profile of RDCs in unfolded proteins: Amino acid bulkiness, the ratio of the side chain volume to its length, predicts clearly observable features reporting on the local conformational behavior of natively unfolded proteins, such as the 140-residue protein α-synuclein (αS).

αS is the major component of abnormal proteinaceous aggregates in the brain of patients with Parkinson's disease. In its monomeric form, αS was classified as natively unfolded.[12] Using RDCs and paramagnetic relaxation enhancement from specifically attached paramagnetic nitroxide radicals, we showed that, despite its high flexibility, native αS adopts an ensemble of conformations that are stabilized by long-range interactions.[13] In the study reported here, αS was weakly aligned in *n*-octylpenta(ethylene glycol)/octanol.[14] RDCs were determined at 15 °C in 20 mM Tris·HCl, pH 7.5, 100 mM NaCl, and with 8 M urea.

Figure 1 compares RDCs observed in αS with the amino acid bulkiness.[15] Bulkiness values were averaged over a five residue window size, and values for prolines were increased by an empirical scaling factor of 1.6. Larger flexibility at the ends of the polypeptide chain was taken into account by combining the bulkiness profile with a bell-shaped curve that is based on the assumption that the influence of neighboring residues decays exponentially as the distance from a given residue (persistence length of the chain = 7).[16] Note that this only affects bulkiness values for residues at the termini. The refined bulkiness profile closely matches the variation of RDCs observed in αS as a function of position along the chain. In regions in which large RDCs were observed, many residues with bulky side chains are present (Figure 1). The regions with large RDCs are separated by residues that showed couplings close to zero. In these linker sequences, mainly amino acids with small side chains such as glycine and alanine are found: ^{29}AAG^{31}, ^{67}GGA^{69}, ^{84}GAGS^{87}, ^{106}GA^{107}. The largest deviations between the RDC pattern and the bulkiness profile were present in the N-terminus and for residues 115−119 and 125−129. Upon addition of urea, these deviations were removed (Figure 1C). Although the interaction between the N- and the C-terminus is expected to be mostly electrostatic, the C-terminus forms hydrophobic interactions with the central part of αS. Thus, the observed changes in RDCs suggest a complex network of long-range interactions, giving rise to a more complex RDC baseline upon which local structure is superimposed.[13,17]

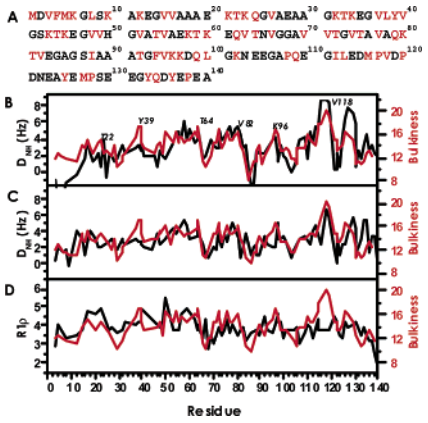

Figure 1. (A) Amino acid sequence of αS. Bulky amino acids (bulkiness > 14) are marked in red. (B, C) Comparison of amino acid bulkiness (five residue window average) (red) with one-bond ^{1}H−^{15}N RDCs (black) observed in αS in (B) buffer and (C) with 8 M urea. (D) Comparison of amino acid bulkiness (red) with ^{15}N $R_{1\rho}$ rates (black). $R_{1\rho}$ values were obtained by collecting six 2D spectra using relaxation delays of 8, 32, 48, 88, and 176 ms, and an on-resonance spinlock pulse of 2.5 kHz. Bulkiness values are only shown for residues with experimental data.

The minimum deviation between experimental RDCs and the bulkiness pattern was obtained for a five to seven residue window average. Averaging over several residues simulates the effect neighboring residues exert on the local conformation and dynamics of each amino acid in a polypeptide chain. This strongly suggests

[†] Max Planck Institute for Biophysical Chemistry.
[‡] DFG Research Center for the Molecular Physiology of the Brain.
[¶] Institut de Biologie Structurale Jean-Pierre Ebel.
[§] Universidad Nacional de Rosario.

10.1021/ja067482k CCC: $37.00

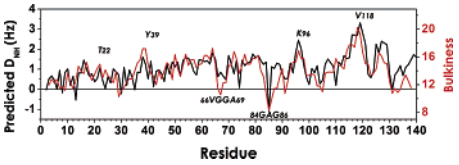

Figure 2. Comparison of amino acid bulkiness (five residue window average; red) with ensemble-averaged RDCs (black). RDCs were predicted from 50 000 coil structures generated by flexible-meccano.

that the Flory isolated-pair hypothesis[18] is not sufficient to explain RDCs in natively unfolded αS. A window size of five to seven is in agreement with other measurements[5,6,16] and calculations[19,20] that estimated the length scale over which spatial correlations decay in denatured proteins to range from six to nine residues.

In the flexible-meccano approach, peptide chains are built using randomly selected ϕ/φ pairs drawn from a database of amino acid specific conformations present in loop regions of high-resolution X-ray structures.[10] The alignment tensor is predicted for each conformer on the basis of the three-dimensional shape using PALES,[21] and associated RDCs are calculated for each NH vector with respect to this tensor. RDCs from each site are then averaged over 50 000 conformers to ensure convergence. Figure 2 compares RDCs predicted from the flexible-meccano ensemble with the pattern of side chain bulkiness. A highly similar variation of values is observed along the polypeptide chain of αS. Slight deviations were observed in the vicinity of residue Y39, for residues 65−70, 85−90, and 125−130, and at the C-terminus of αS, mostly regions in which glycine or proline residues are present.

The similarity between these profiles provides a direct experimental proof for the dominating influence of steric interactions on the composition of the Ramachandran plot. RDCs are reproduced equally well by sampling only residue-specific ϕ/φ distributions or by only considering the bulkiness of amino acid side chains. Removing steric exclusion from the flexible-meccano approach does not affect the results of the simulation very strongly (Supporting Information). In addition, explicit inclusion of nearest neighbor interactions into flexible-meccano was not required. On the other hand, Jha et al. concluded that the identity of neighboring residues needs to be incorporated to improve RDC reproduction.[11] We currently do not know the origin of this disagreement.

How general is the dominance of bulkiness? Previously, RDCs observed in the chemical denatured Snase fragment Δ131Δ were used to argue for the presence of a native-like organization of chain segments in unfolded proteins.[8] Comparison of the experimental RDCs of Δ131Δ with side chain bulkiness averaged over a five residue window, however, suggests that the variation of RDCs along the chain of Δ131Δ can be explained without the need for invoking a native-like topology. Similarly, the RDC pattern observed in chemically denatured eglin C[22] and in a 130-residue fragment of natively unfolded tau closely matches the bulkiness profile (Supporting Information). Thus, a more likely explanation for the variation of RDCs along polypeptide chains is that minimization of steric overlap promotes chain stretching. In more extended parts of the chain, the interaction direction between the dipolar orientation tends to be perpendicular to the external field, resulting in increased RDCs. In addition, more extended backbone conformations are expected to align more effectively, further increasing the magnitude of RDCs in these regions.

Whereas RDCs probe both structure and dynamics, heteronuclear relaxation rates monitor directly backbone motional restrictions. In particular, on-resonance ^{15}N $R_{1\rho}$ transverse relaxation rates report on motions that occur on the pico- to nanosecond and micro- to millisecond time scale. $R_{1\rho}$ values vary along the chain of αS in a similar way as is seen for RDCs, and the bulkiness profile closely matches the $R_{1\rho}$ pattern (Figure 1D). This indicates that the local steric interactions between side chains and the backbone restrict motions on the pico- to nanosecond and micro- to millisecond time scale, in agreement with relaxation time measurements previously reported for acid-unfolded apomyoglobin.[6] The largest deviation between the $R_{1\rho}$ and the bulkiness profile was observed in the vicinity of P117 and P120, suggesting that prolines also restrict slower motions for which $R_{1\rho}$ rates are not sensitive.

Our results demonstrate that, although various types of intramolecular interactions, such as electrostatic and solvent interactions, play important roles, simple considerations of the bulkiness of amino acids predict a major component of diverse parameters dependent on the local conformation and dynamics of αS and other natively unfolded proteins. Deviations from this random coil behavior, as evidenced by RDCs in the N- and C-terminal domain of αS, can provide insight into residual secondary structure and long-range transient interactions in weakly structured proteins. The local steric restrictions in the unfolded state can also bias the conformational search toward native-like elements and thereby reduce the Levinthal paradox.

Acknowledgment. We thank C. Griesinger and T.M. Jovin for discussions, S. Becker for help with the production of αS, and M.D. Mukrasch for the RDCs observed in tau. This work was supported by a DFG-Graduiertenkolleg scholarship to M.-K.C., through ANR NT05-4_42781 to M.B., by the DFG through ZW71/1-5, ZW71/2-1, and ZW71/3-1 to M.Z., by the EU through UPMAN, and by the Max Planck society.

Supporting Information Available: Comparison of RDCs of αS, Δ131Δ, eglin C, and K18-tau with residue bulkiness. Influence of steric exclusion on RDCs predicted by flexible-meccano. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Baldwin, R. L. *Adv. Protein Chem.* **2002**, *62*, 361−367.
(2) Tanford, C.; Kawahara, K.; Lapanje, S. *J. Biol. Chem.* **1966**, *241*, 1921.
(3) Fiebig, K. M.; Schwalbe, H.; Buck, M.; Smith, L. J.; Dobson, C. M. *J. Phys. Chem.* **1996**, *100*, 2661−2666.
(4) Klein-Seetharaman, J.; Oikawa, M.; Grimshaw, S. B.; Wirmer, J.; Duchardt, E.; Ueda, T.; Imoto, T.; Smith, L. J.; Dobson, C. M.; Schwalbe, H. *Science* **2002**, *295*, 1719−1722.
(5) Mohana-Borges, R.; Goto, N. K.; Kroon, G. J.; Dyson, H. J.; Wright, P. E. *J. Mol. Biol.* **2004**, *340*, 1131−1142.
(6) Schwarzinger, S.; Wright, P. E.; Dyson, H. J. *Biochemistry* **2002**, *41*, 12681−12686.
(7) Shi, Z.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9190−9195.
(8) Shortle, D.; Ackerman, M. S. *Science* **2001**, *293*, 487−489.
(9) Louhivuori, M.; Paakkonen, K.; Fredriksson, K.; Permi, P.; Lounila, J.; Annila, A. *J. Am. Chem. Soc.* **2003**, *125*, 15647−15650.
(10) Bernado, P.; Blanchard, L.; Timmins, P.; Marion, D.; Ruigrok, R. W.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17002−17007.
(11) Jha, A. K.; Colubri, A.; Freed, K. F.; Sosnick, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13099−13104.
(12) Weinreb, P. H.; Zhen, W.; Poon, A. W.; Conway, K. A.; Lansbury, P. T., Jr. *Biochemistry* **1996**, *35*, 13709−13715.
(13) Bertoncini, C. W.; Jung, Y. S.; Fernandez, C. O.; Hoyer, W.; Griesinger, C.; Jovin, T. M.; Zweckstetter, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 1430−1435.
(14) Ruckert, M.; Otting, G. *J. Am. Chem. Soc.* **2000**, *122*, 7793−7797.
(15) Zimmerma, Jm.; Eliezer, N.; Simha, R. *J. Theor. Biol.* **1968**, *21*, 170−201.
(16) Schwalbe, H.; Fiebig, K. M.; Buck, M.; Jones, J. A.; Grimshaw, S. B.; Spencer, A.; Glaser, S. J.; Smith, L. J.; Dobson, C. M. *Biochemistry* **1997**, *36*, 8977−8991.
(17) Bernado, P.; Bertoncini, C. W.; Griesinger, C.; Zweckstetter, M.; Blackledge, M. *J. Am. Chem. Soc.* **2005**, *127*, 17968−17969.
(18) Flory, P. J. *Statistical Mechanics of Chain Molecules*; Hanser Publishers: Munich, Germany, 1969.
(19) Ohkubo, Y. Z.; Brooks, C. L., III. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13916−13921.
(20) Tran, H. T.; Pappu, R. V. *Biophys. J.* **2006**, *91*, 1868−1886.
(21) Zweckstetter, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 3791−3792.
(22) Ohnishi, S.; Lee, A. L.; Edgell, M. H.; Shortle, D. *Biochemistry* **2004**, *43*, 4064−4070.

JA067482K