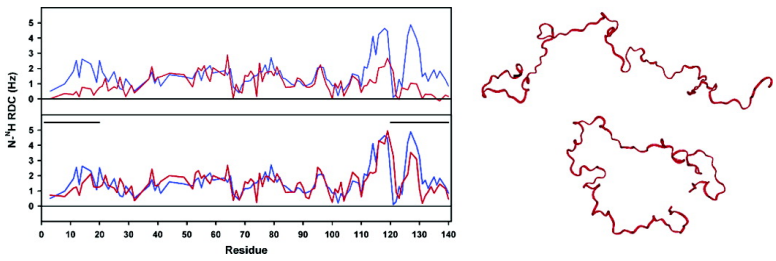
Communication

Defining Long-Range Order and Local Disorder in Native □-Synuclein Using Residual Dipolar Couplings

Pau Bernad, Carlos W. Bertoncini, Christian Griesinger, Markus Zweckstetter, and Martin Blackledge

J. Am. Chem. Soc., **2005**, 127 (51), 17968-17969• DOI: 10.1021/ja055538p • Publication Date (Web): 02 December 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 22 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 12/02/2005

Defining Long-Range Order and Local Disorder in Native α-Synuclein Using Residual Dipolar Couplings

Pau Bernadó, Carlos W. Bertoncini, Christian Griesinger, Markus Zweckstetter, and Martin Blackledge*

Institute de Biologie Structurale Jean-Pierre Ebel, CNRS-CEA-UJF, 41 rue Jules Horowitz, 38027-Grenoble Cedex, France, Department of NMR-Based Structural Biology and Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany

Received August 13, 2005; E-mail: martin.blackledge@ibs.fr

The conformational flexibility inherent to natively unfolded proteins places them beyond the reach of classical structural biology. It is, however, becoming clear that these proteins participate in a vast range of biochemical processes,[1] and that their native plasticity bestows specific functional properties.[2] In contrast to structured proteins, intrinsically unfolded proteins must be described by an ensemble of interconverting conformers. Residual dipolar couplings (RDCs)[3] report on time and ensemble-averaged conformations up to the millisecond time scale[4] and can, therefore, be used to characterize both the structure and dynamics of unfolded proteins.[5] In this study, we present a novel interpretation of RDCs that simultaneously describes long-range structural order and local conformational sampling. This approach is used to describe the structure and dynamics of α-Synuclein (αS), a 140 amino acid protein found in human brain and strongly implicated in the onset of Parkinson's disease (PD).[6]

Recent studies have shown that native αS presents a more compact structure than expected for a completely unfolded chain, and this compactness has been linked to inhibition of fibrillation, due to burial of the NAC domain.[7−9] A network of transient long-range contacts has been proposed from the observation of relaxation enhancement (PRE) in the vicinity of a series of spin labels distributed along the chain.[7,8] RDCs have also been measured in αS, providing important complementary information about the structural response to different experimental conditions[7] and, in particular, showing that long-range interactions are perturbed in familial mutants of αS linked to PD.[10] Here we demonstrate the use of RDCs to study both short- and long-range interactions in αS.

Our approach to the interpretation of RDCs in unfolded proteins is to generate a statistical ensemble of coil conformations and calculate averaged RDCs over the entire population.[11] To achieve this, we have developed *flexible*-meccano, an algorithm that sequentially builds peptide chains using randomly selected ϕ/ψ pairs drawn from a database of amino acid specific conformations present in loop regions of high-resolution X-ray structures.[12] A simple volume-exclusion model is used to avoid steric overlap.[13] The alignment tensor is predicted[14] for each complete conformer on the basis of hydrodynamic shape,[15] 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.

RDCs predicted from αS are compared to experimental data in Figure 1.[7] Close agreement is found for the central region (residues 30−110). This underlines the observation that the range and fine structure of RDCs measured in unfolded proteins can be accurately predicted on the basis of local, amino acid specific conformational propensity.[11] RDCs from urea-unfolded αS are also in broad agreement with these ensemble-averaged values.[16] In native αS,

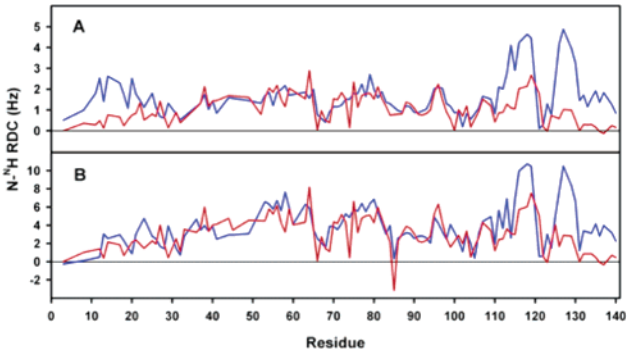


Figure 1. Ensemble-averaged RDCs simulated (red) without long-range contacts and (blue) experimental in (A) Pf1 bacteriophage and (B) lyotropic media (see ref 7). Simulated data are scaled to maximize fit in the region 22−112. For illustration purposes, the RDC sign is inverted compared to conventions found in, for example, ref 19.

however, the RDC profiles in the N and C-terminal regions deviate significantly from simulation, suggesting the existence of more complex conformational behavior. The striking distribution of RDCs measured in these regions has indeed been attributed to long-range contacts involving the C-terminal, NAC (61−95), and N-terminal domains.[7]

We have investigated the presence of tertiary contacts by creating conformational ensembles that fulfill distance constraints between $^{\beta}$C atoms at specified positions in the primary sequence. The relevance of long-range interactions between different parts of the protein was systematically tested by dividing the 140 amino acid chain into 7 20-residue strands 1−20, 21−40, etc. The *flexible*-meccano procedure was repeated, and conformers were only accepted if a $^{\beta}$C from one of the 20-residue domains was less than 15 Å from a $^{\beta}$C from the other specified domain. Adjacent domains give identical profiles to those in Figure 1, as close contacts between neighboring domains will always be fulfilled, leaving 15 independent ensembles of 50 000 conformers containing contacts between the different regions. The resulting profiles of averaged RDCs were compared to experimental data and fit to experimental profiles by varying a single scaling factor over the whole sequence.

χ^2 values resulting from this procedure are shown in Table 1, and examples of different ensemble averages shown in Figure 2. The distribution of experimental RDCs is closely reproduced throughout the sequence when contacts between the N- and C-terminal regions 1−20 and 121−140 are present. This is true using both alcohol and Pf1 phage alignment.[16] Additional calculations were performed, implicating more specific contacts between these two regions, by further dividing domains 1−20 and 121−140 into 10 and 5 residue segments. The ensemble with contacts between regions (6−10) and (136−140) best reproduces the data

10.1021/ja055538p CCC: $30.25

Table 1. Effect of Long-Range Contacts on Capacity of Conformational Ensembles to Reproduce Experimental RDCs from αS

segment	41–60	61–80	81–100	101–120	121–140
1−20	141[a]	117	120	89	55
21−40		129	119	99	71
41−60			132	140	118
61−80				143	156
81−100					181

[a] Figures denote $\chi^2 = \sum({}^1D_{ij,\mathrm{calc}} - {}^1D_{ij,\mathrm{meas}})^2$ that compares experimental RDCs measured in αS aligned in Pf1 phage with calculated averages over 50 000 conformers having long-range contacts (<15 Å) between $^{\beta}$Cs in the specified ranges. Adjacent domains are equivalent to ensembles with no specified contact shown in Figure 1 (χ^2= 129) and are not shown.

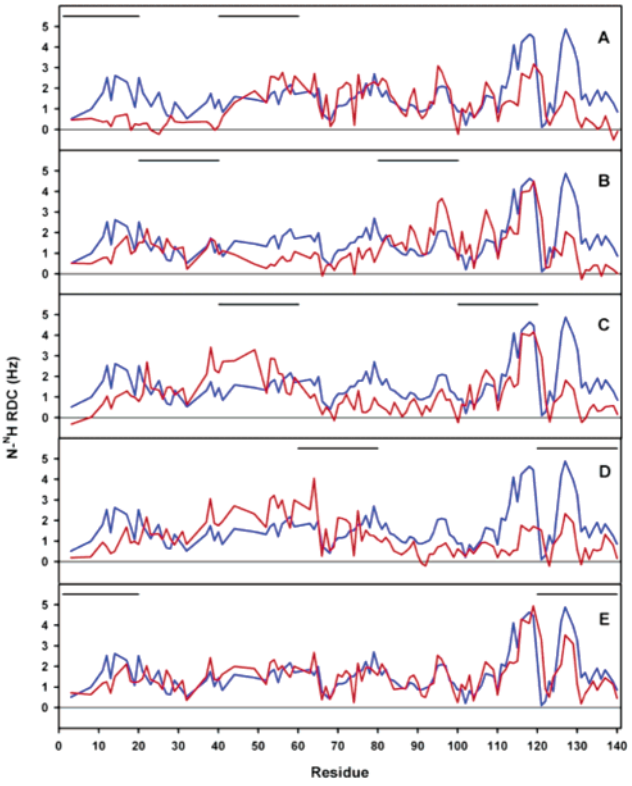


Figure 2. (A−E) Comparison of experimental $^1D_{NH}$ (blue) and simulated from long-range contact ensembles (red). Contact regions are indicated by bars. (A) (1−20, 41−60); (B) (21−40, 81−100); (C) (41−60, 101−120); (D) (61−80, 121−140); (E) (1−20, 121−140).

from both alignment media. The net charge of region 6−10 is +2 (K6, K10) and that of region 136−140 is −2 (E137, E139), indicating that the strongest interaction may be electrostatic.[7] Varying the interaction distance constraint from 15 to 25 Å induced negligible difference in calculated RDCs.

The agreement of experimental and simulated RDCs along the whole sequence for conformational sub-ensembles containing a contact between the N- and C-terminal domains is striking. The local structure found in the C-terminus corresponds to the presence of three prolines and numerous bulky side chains that combine to induce increased order that is accentuated in the regions implicated in the long-range contact. The results indicate that the proposed contact is present in native αS in solution, although they do not exclude the presence of other contacts within the same conformer. Indeed, a model with simultaneous 15 Å contacts between the C-terminal domain and regions 1−20 and the NAC domain fits the data almost as closely (χ^2 = 59, data not shown). These results are therefore complementary to PRE-based detection of long-range contacts in αS.[7,8] The existence of long-range contacts between the N- and C-termini is supported by the observation of enhanced fibrillation rates upon addition of polyamines, on acidification, with increasing ionic strength, and in site-specific mutants[10] as well as deletion of the C-terminus of αS.[17] Under these conditions, an electrostatic interaction between the terminal domains may be disrupted, resulting in increased solvent accessibility of the hydrophobic NAC domain, provoking onset of fibrillation.[18]

The structural model proposed here indicates that two features are required to describe αS in solution: local conformational fluctuations based on random sampling of residue-specific ϕ/ψ distributions, and long-range contacts between domains that are distant in primary sequence. Inclusion of both aspects accurately reproduces nonaveraged couplings measured in αS. The model is validated from RDCs predicted from the shape of each conformer approximated as an ellipsoid.[15] An atomic resolution approach to alignment prediction[14] may deliver more precise detail, and we are currently actively investigating this possibility. Although RDCs have previously been shown to report on local conformational preferences in unstructured proteins,[11,19] this study demonstrates their additional sensitivity to long-range order in highly flexible systems. The demonstration that RDCs can be accurately reproduced using simple models of local and long-range structure has important implications for our understanding of the conformational behavior of unfolded proteins in solution.

Acknowledgment. Support from the Commisariat à l'Energie Atomique, the CNRS and by the MPG, the DFG (CMPB and GRK 782) and the FCI. P.B. received an EMBO fellowship, and M.Z. a DFG Emmy Noether grant (ZW 71/1-5).

Supporting Information Available: RDC profiles representing Table 1. No-contact simulation compared to RDCs from αS aligned in alcohol and in 4 and 8 M urea. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Uversky, V. N. *Protein Sci.* **2002**, *11*, 739−756.
(2) Dyson, H. J.; Wright, P. E. *Nat. Rev.* **2005**, *9*, 197−208.
(3) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.
(4) Prestegard, J. H.; al-Hashimi, H. M.; Tolman, J. R. *Q. Rev. Biophys.* **2000**, *33*, 371−424.
(5) (a) Shortle, D.; Ackerman, M. S. *Science* **2001**, *293*, 487−489. (b) Louhivouri, M.; Pääkkönen, K.; Fredriksson, K.; Permi, P.; Lounila, J.; Annila, A. *J. Am. Chem. Soc.* **2003**, *125*, 15647−15650.
(6) Spillantini, M. G.; Schmidt, M. L.; Lee, V. M.; Trojanowsky, J. Q.; Jakes, R.; Goedert, M. *Nature* **1997**, *388*, 839−840.
(7) 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.
(8) Dedmon, M. M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. *J. Am. Chem. Soc.* **2005**, *127*, 476−477.
(9) Uversky, V. N.; Li, J.; Souillac, P.; Millett, I. A.; Doniach, S.; Jakes, R.; Goedert, M.; Fink, A. L. *J. Biol. Chem.* **2002**, *277*, 11970−11978.
(10) Bertoncini, C. W.; Fernandez, C. O.; Griesinger, C.; Jovin, T. M.; Zweckstetter, M. *J. Biol. Chem.* **2005**, *280*, 30649−30652.
(11) (a) Bernado, P.; Blanchard, L.; Timmins, P.; Marion, D.; Ruigrok, R.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17002−17007. (b) Jha, A.; Colubri, A.; Freed, K.; Sosnick, T. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13099−13104.
(12) Lovell, S. C.; Davis, I. W.; Arendall, W. B., III; de Bakker, P. I. W.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. *Proteins* **2003**, *50*, 437−450.
(13) Levitt, M. *J. Mol. Biol.* **1976**, *104*, 59−107.
(14) Zweckstetter, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 3791−3792.
(15) Almond, A.; Axelsen, J. B. *J. Am. Chem. Soc.* **2002**, *124*, 9986−9987.
(16) See Supporting Information.
(17) (a) Fernández, C. O.; Hoyer, W.; Zweckstetter, M.; Jares-Erijman, E. A.; Subramaniam, V.; Griesinger, C.; Jovin, T. M. *EMBO J.* **2004**, *23*, 2039−2046. (b) Hoyer, W.; Cherny, D.; Subramaniam, V.; Jovin, T. M. *Biochemistry* **2004**, *43*, 16233−16242.
(18) El-Agnaf, O. M.; Irvine, G. B. *Biochem. Soc. Trans.* **2001**, *30*, 559−565.
(19) (a) Louhivuori, M.; Fredriksson, K.; Pääkkönen, K.; Permi, P.; Annila, A. *J. Biomol. NMR* **2004** *29*, 517−54. (b) Mohana-Borges, R.; Goto, N. K.; Kroon, G. J. A.; Dyson, H. J.; Wright, P. E. *J. Mol. Biol.* **2004**, *34*, 1131−1142.

JA055538P