Flexibility of an Arginine Side Chain at a DNA−Protein Interface

Helena Berglund, Herbert Baumann, Stefan Knapp, Rudolf Ladenstein, and Torleif Härd*

Center for Structural Biochemistry, Karolinska Institutet NOVUM, S-141 57 Huddinge, Sweden

Received September 28, 1995

Nuclear magnetic resonance (NMR) relaxation processes are inherently sensitive to molecular motions,[1] and recent methodological advances have made it possible to employ relaxation measurements in the studies of internal dynamics in proteins and other biomolecules on a wide range of time scales (refs 2−10 and references cited therein). We have studied the picosecond/nanosecond mobility of arginine side chains by measuring the NMR relaxation properties of the arginine $^{15}N^{\epsilon}$ nuclei in isotopically enriched proteins. The objective is to explore the range of mobilities found for arginines in different protein environments and, in particular, to examine changes in the mobility of a surface arginine in a DNA-binding protein as that arginine forms part of the DNA−protein interface upon complexation with DNA.

The protein Sso7d is an abundant 7.1 kDa DNA-binding protein from *Sulfolobus solfataricus* for which we previously reported the structure.[11] The protein contains a two-stranded and a three-stranded β-sheet and a C-terminal α-helix. The topology of the two β-sheets is very similar to that found in SH3 domains.[11] Sso7d contains two arginine residues: Arg24 and Arg42. Arg24 is located close to the protein surface, but the side chain is not completely exposed to the solvent. Arg42 is located on the DNA-binding surface of Sso7d[12] and is more solvent-exposed than Arg24 in the uncomplexed protein.

^{15}N-Enriched Sso7d was expressed and purified as described earlier.[12] The protein was found to form a 1:1 complex with a double-stranded DNA dodecamer with the sequence 5′TATCGCGCGATA3′.[13] Free and bound states of the protein are in slow exchange on the NMR time scale, indicating strong binding of Sso7d to this DNA oligomer. Sample concentrations were 1.1 and 2.4 mM for the free Sso7d and the complex, respectively. For comparison we also measured the relaxation properties of $^{15}N^{\epsilon}$ of Arg496 of the glucocorticoid receptor DNA-binding domain (GR DBD), which is completely buried in the protein core. For these studies we used the same 2 mM ^{15}N-enriched sample of an 82-residue GR DBD fragment as in several of our earlier studies of structure and dynamics of this protein.[8,9,14]

^{15}N T_1 and T_2 relaxation and $\{^1H\}-^{15}N$ cross relaxation (steady state NOE) was measured at a magnetic field of 11.7 T on a Varian Unity 500 spectrometer equipped with pulsed field gradient hardware. The measurements were made at 30 °C following well-established methodology[2,3] using indirect detection pulse sequences (Figure 10 in ref 15) that are designed to avoid effects of cross-correlation between $^{15}N-^1H$ dipolar interactions and ^{15}N chemical shift anisotropy. The CPMG spin-echo delay in the T_2 measurements was 450 μs. The relaxation data was quantified in terms of the Lipari−Szabo formalism.[16−18] The data is summarized in Table 1. We focus the discussion on the amplitude of the internal motions as reflected in the generalized order parameter (S^2). This is because fitted values of S^2 are less dependent on uncertainties such as anisotropy in the overall rotational motions of the proteins/complexes[19] than other parameters in the Lipari−Szabo formalism. The S^2 order parameter can be interpreted as the degree of restriction of the NH bond vector on a time scale faster than the rotational tumbling motions of the molecule, and it ranges from 0 for completely unrestricted mobility to 1 in the absence of internal motions.[16]

We find that the mobility of the Arg24 side chain in the uncomplexed state of Sso7d is more restricted than that of Arg42, as reflected in a higher value of the S^2 order parameter. This difference is presumably due to the fact that Arg42 is essentially completely exposed to the solvent whereas Arg24 is partially buried, but hydrogen-bonding interactions involving the guanidino group of the Arg24 side chain might also contribute. The mobility of Arg24 is still larger than that of backbone amides in ordered secondary structure conformations, for which S^2 normally take values of $\approx 0.8-0.9$.[2] The mobility of Arg496 in the core of the GR DBD, on the other hand, is very restricted, with $S^2 \approx 0.96$. The low flexibility is in this case caused by optimal packing in the GR DBD core, possibly in combination with a hydrogen bond between the guanidino amide and the aromatic ring of Phe462. Thus, we find a large degree of variability in the flexibility of arginine side chains in different protein environments. Our results on arginine side chains are in agreement with the correlation between the S^2 order parameters and extent of solvent exposure noted in a recent study of glutamine and asparagine side chains in lysozyme.[10] On the other hand, arginine side chains in lysozyme all appear to be flexible, with S^2 values within a more narrow range ($0.05 \leq S^2 \leq 0.31$) than the arginines studied here.

It is of particular interest to obtain knowledge about the flexibility of side chains at a DNA−protein interface. This is because a decrease in flexibility upon the formation of a complex can be expected to contribute unfavorably to the free energy of binding.[20] Reduced side chain mobility at the interface might also result in altered vibrational modes,[21,22] and this effect has been suggested to contribute to the large negative change in heat capacity which is often observed upon binding of proteins

(1) Abragam, A. *Principles of Nuclear Magnetism*; Clarendon Press: Oxford, 1961.
(2) Palmer, A. G. *Curr. Opin. Biotechnology* **1993**, *4*, 385−391.
(3) Kay, L. E.; Torchia, D. A.; Bax, A. *Biochemistry* **1989**, *28*, 8972−8979.
(4) Peng, J. W.; Wagner, G. *Methods Enzymol.* **1994**, *239*, 563−596.
(5) Szyperski, T.; Luginbühl, P.; Güntert, P.; Wüthrich, K. *J. Biol. NMR* **1993**, *3*, 151−164.
(6) Clore, G. M.; Driscoll, P. C.; Wingfield, P. T.; Gronenborn, A. M. *Biochemistry* **1990**, *29*, 7387−7401.
(7) Mandel, A. M.; Akke, M.; Palmer, A. G. *J. Mol. Biol.* **1995**, *246*, 144−163.
(8) Berglund, H.; Kovács, H.; Dahlman-Wright, K.; Gustafsson, J.-Å.; Härd, T. *Biochemistry* **1992**, *31*, 12001−12011.
(9) Eriksson, M. A. L.; Berglund, H.; Härd, T.; Nilsson, L. *Proteins* **1993**, *17*, 375−390.
(10) Buck, M.; Boyd, J.; Redfield, C.; MacKenzie, D. A.; Jeenes, D. J.; Archer, D. B.; Dobson, C. M. *Biochemistry* **1995**, *34*, 4041−4055.
(11) Baumann, H.; Knapp, S.; Lundbäck, T.; Ladenstein, R.; Härd, T. *Nature Struct. Biol.* **1994**, *1*, 808−819.
(12) Baumann, H.; Knapp, S.; Karshikoff, A.; Ladenstein, R.; Härd, T. *J. Mol. Biol.* **1995**, *247*, 840−846.
(13) Baumann, H.; Knapp, S.; Ladenstein, R.; Härd, T. Unpublished results.
(14) Baumann, H.; Paulsen, K.; Kovács, H.; Berglund, H.; Wright, A. P. H.; Gustafsson, J.-Å.; Härd, T. *Biochemistry* **1993**, *32*, 13463−13471.
(15) Farrow, N. A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Forman-Kay, J. D.; Kay, L. E. *Biochemistry* **1994**, *33*, 5984−6003.
(16) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.
(17) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559−4570.
(18) Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1990**, *112*, 4989−4991.
(19) Schurr, J. M.; Babcock, H. P.; Fujimoto, B. S. *J. Magn. Reson.* **1994**, *B105*, 211−224.
(20) Akke, M.; Brüschweiler, R.; Palmer, A. G. *J. Am. Chem. Soc.* **1993**, *115*, 9832−9833.
(21) Sturtevant, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 2236−2240.
(22) Tidor, B; Karplus, M. *J. Mol. Biol.* **1994**, *238*, 405−414.

0002-7863/95/1517-12883$09.00/0

Table 1. Relaxation Data and Order Parameters for Studied Arginine Side Chains[a]

protein	residue	$^{15}N^{\epsilon}$ T_1 (s)	$^{15}N^{\epsilon}$ T_2 (s)	{^{1}H}−$^{15}N^{\epsilon}$ NOE	S^2
Sso7d[b]	Arg24[e]	0.570 ± 0.007	0.247 ± 0.004	0.258 ± 0.027	0.51 ± 0.01
	Arg42[e]	1.096 ± 0.290	0.235 ± 0.005	−2.365 ± 0.060	0.08 ± 0.01
Sso7d-DNA[c]	Arg24[e]	0.789 ± 0.009	0.124 ± 0.001	0.108 ± 0.010	0.58 ± 0.01
	Arg42[f]	0.711 ± 0.022	0.125 ± 0.003	0.270 ± 0.023	0.60 ± 0.02
GR DBD[d]	Arg496[g]	0.409 ± 0.013	0.109 ± 0.007	0.686 ± 0.142	0.96 ± 0.03

[a] Model free[16−18] parameters were fitted using the program ModelFree 3.1 (kindly provided by Dr. A. G. Palmer, University of Columbia) essentially following the procedures for model selection described by Farrow *et al.*[15] That is, a model was accepted if all three relaxation parameters were fitted within 95% confidence intervals and if the errors were less than fitted values of the parameters in the model. An additional requirement was that the sum-squared error of fitted relaxation data should fall within the $\alpha = 0.05$ critical value of the distribution of sum-squared errors from fits of simulated relaxation data. Correlation times, τ_c, for overall tumbling motions were determined on the basis of relaxation data for all backbone amides. The fitted parameters are indicated for each case, but fitted values for parameters other than S_2 and τ_c, such as correlation times for internal motions and contributions due to chemical exchange, are not shown as descussed in the text. (The reader is referred to refs 7 and 18 for a comprehensive description of the parametrization of various models describing internal motions.) The $N^{\epsilon}H$ bond length was set to 1.02 Å. The chemical shift anisotropy of arginine $^{15}N^{\epsilon}$ was set to $\Delta\sigma = -160$ ppm as for amide nitrogens in peptides.[26] A variation of $\Delta\sigma$ within the interval −120 to −200 ppm results in variations in fitted S^2 values of <15%, with approximately equal relative effects on all best-fit values of S^2. [b] Overall rotational correlation time $\tau_c = 3.1$ ns. [c] $\tau_c = 9.4$ ns. [d] $\tau_c = 5.8$ ns. [e] Fitted parameters: S^2, τ_e, and R_{ex}. [f] Fitted parameters: S^2_s, S^2_f, and τ_{es}; $S^2 = S^2_s S^2_f$. [g] Fitted parameter: S^2.

to DNA.[23] We find that the S^2 of Arg42 in Sso7d increases from 0.08 to 0.60 upon binding to DNA. Thus, we conclude that there is a reduced mobility of Arg42 in the complex compared to the uncomplexed state. Still, the mobility of Arg42 in the Sso7d−DNA complex is not as low as that of amides in an ordered protein backbone or that of an arginine residue in a protein core.

It is important to note that there is an entropic penalty associated with the restriction of conformational flexibility. Akke et al. recently showed how the free energy loss due to restricted mobility may be estimated on the basis of measured S^2 values, provided that $S^2 > 0.5$.[20] Thus, a complete immobilization of Arg42 at the protein−DNA interface, i.e., an additional increase of S^2 from ≈0.6 to ≈0.9, would correspond to an expected entropy cost of 0.5−1.0 kcal mol^{-1} if one assumes, for instance, a diffusion-in-a-cone model for the motions in both the free and complexed states. This number is comparable to the free energy of a hydrogen bond between biomolecules in solution (ref 24 and work cited therein) or the expected entropy of a bound water molecule,[25] and corresponds to a reduction/increase in the binding constant by a factor of 2−5. The observed restriction of Arg42 motions (S^2 increase from 0.08 to 0.60 upon complexation) is also associated with an entropy loss, but the theoretical treatment by Akke et al. does not in this case allow for a reliable estimate of the magnitude.[20] In general, the DNA-binding surface of a protein contains several amino acid side chains, and the combined entropy loss due to the restriction of their mobilities upon binding to DNA might become large compared to the binding free energy. The extent of immobilization of individual side chains on the surface of a DNA-binding protein might therefore be a decisive determinant for the DNA-binding affinity and/or specificity. Our results should therefore motivate further studies on other DNA−protein complexes in order to investigate the generality of the observations.

We also measured the relaxation properties of all backbone amide nitrogens in Sso7d to compare the backbone dynamics of the protein in the free and DNA-bound states. The flexibility varies somewhat along the Sso7d backbone. Still, measured S^2 values fall in the range $0.75 \leq S^2 \leq 0.90$ for most of the protein, as expected considering the compact Sso7d fold. Regions which are notably more flexible, with $S^2 \leq 0.75$, include a glycine-rich surface loop and the five C-terminal residues. These two regions do not form part of the DNA-interacting surface.[12] The backbone flexibility does not change dramatically upon complexation with DNA, although there might be a slight average increase in measured S^2 values of the complex compared to the free state. Furthermore, the relative variation in S^2 along the backbone found in the uncomplexed protein is retained in the complex. Thus, as far as amplitudes of picosecond motions are concerned, we conclude that complexation with DNA in this case affects the side chain mobility at the DNA-interacting surface to a larger extent than the backbone mobility.

In summary, we present experimental evidence for a wide variation in the flexibility of arginine side chains in different protein environments. We also find that the flexibility of an arginine side chain at the DNA−Sso7d interface is restrained compared to the uncomplexed state, but that it still is more flexible than an arginine side chain which is packed in a protein core. A comparison between side chain and backbone mobilities in the free and DNA-complexed states of Sso7d indicates that the backbone mobility is affected less upon DNA binding than the side chain mobility at the DNA-interacting surface.

Acknowledgment. This work was supported by grants from the Swedish Natural Sciences Research Council and the Magn. Bergvall Foundation and by the European Community (project "Biotechnology of Extremophiles") through a grant from NUTEK.

JA953307K

(23) Ladbury, J. E.; Wright, J. G.; Sturtevant, J. M.; Sigler, P. B. *J. Mol. Biol.* **1994**, *238*, 669−681.
(24) Smith, S. A.; Rajur, S. B.; McLaughlin, L. W. *Nature Struct. Biol.* **1994**, *1*, 18−22.
(25) Dunitz, J. D. *Science* **1994**, *264*, 670.
(26) Hiyama, Y.; Niu, C.-H.; Silverton, J. V.; Bavoso, A.; Torchia, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 2378−2383.