Periodicity of Amide Proton Exchange Rates in a Coiled-Coil Leucine Zipper Peptide†

Elizabeth M. Goodman‡ and Peter S. Kim*

Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, Massachusetts 02142

Received September 19, 1991; Revised Manuscript Received October 22, 1991

ABSTRACT: The two-stranded coiled-coil motif, which includes leucine zippers, is a simple protein structure that is well suited for studies of helix–helix interactions. The interaction between helices in a coiled coil involves packing of "knobs" into "holes", as predicted by Crick in 1953 and confirmed recently by X-ray crystallography for the GCN4 leucine zipper [O'Shea, E. K., Klemm, J. D., Kim, P. S., & Alber, T. (1991) *Science 254*, 539]. A striking periodicity, extending over six helical turns, is observed in the rates of hydrogen–deuterium exchange for amide protons in a peptide corresponding to the leucine zipper of GCN4. Protons at the hydrophobic interface show the most protection from exchange. The NMR chemical shifts of amide protons in the helices also show a pronounced periodicity which predicts a short H-bond followed by a long H-bond every seven residues. This variation was anticipated in 1953 by Pauling and is sufficient to give rise to a local left-handed superhelical twist characteristic of coiled coils. The amide protons that lie at the base of the "hole" in the "knobs-into-holes" packing show slow amide proton exchange rates and are predicted to have short H-bond lengths. These results suggest that tertiary interactions can lead to highly localized, but substantial, differences in stability and dynamics within a secondary structure element and emphasize the dominant nature of packing interactions in determining protein structure.

The two-stranded parallel coiled coil is a structural motif consisting of a pair of α-helices wrapped around each other with a slight left-handed superhelical twist and characterized by two interspersed heptad repeats of hydrophobic residues: the 4–3 repeat [reviews: Talbot and Hodges (1982) and Cohen and Parry (1990)]. "Leucine zipper" sequences (Landschulz et al., 1988), known to be important for mediating dimerization of many DNA binding proteins [reviews: Johnson and McKnight (1989) and Struhl (1989)], fold as short two-stranded parallel coiled coils (O'Shea et al., 1989). Modeling studies predicted an elegant and simple arrangement for the hydrophobic side chains in coiled coils, referred to by Crick as "knobs-into-holes" packing (Crick, 1953; Pauling & Corey, 1953; McLachlan & Stewart, 1975). This packing arrangement has been confirmed recently with the solution of a high-resolution crystal structure for the leucine zipper of the yeast transcriptional activator GCN4 (O'Shea et al., 1991). Peptides corresponding to this leucine zipper therefore provide an excellent model system for studying the nature of helix–helix interactions (O'Shea et al., 1989).

Amide proton exchange measurements were originally introduced as a way to probe the dynamics of protein structures (Linderstrøm-Lang, 1955). With the advent of two-dimensional nuclear magnetic resonance spectroscopic (2D NMR) techniques, it is possible to measure individual amide proton exchange rates, in some cases for the majority of the residues in a protein [review: Wagner and Wüthrich (1986)]. The interpretation, however, of these exchange rates in terms of structural and dynamic features of the protein is not clear-cut [contrast Woodward et al. (1982), Wagner (1983), and Englander and Kallenbach (1984)]. The mechanism of amide proton exchange depends upon the stability of the protein in the conditions where exchange occurs. In one extreme, where the protein is unstable, the dominant mechanism of exchange involves global unfolding. In strongly native conditions, however, the rate of exchange from a given amide depends on more localized effects.

We have used 2D NMR to measure hydrogen–deuterium (H–D) exchange rates for invididual amide protons in a leucine zipper peptide, called GCN4-p2N. In order to probe local differences in stability, measurements of amide proton exchange were made at low temperature (6 °C). GCN4-p2N corresponds in sequence to the C-terminal 32 residues of GCN4 (Hinnebusch, 1984) with the addition of a Cys-Gly-Gly linker at the N-terminus. The peptide GCN4-p2N differs slightly from GCN4-p1 used in previous studies (O'Shea et al., 1989, 1991; Oas et al., 1990; Rasmussen et al., 1990). GCN4-p2N consists of the C-terminal 32 residues of GCN4 and has a neutral (amidated) C-terminus whereas GCN4-p1 consists of the C-terminal 33 residues of GCN4 and has a free α-carboxylate. The second "N" in GCN4-p2N denotes the addition of the flexible Cys-Gly-Gly linker at the N-terminus. Figure 1 shows the sequence and the coiled-coil helical wheel representation for GCN4-p2N (O'Shea et al., 1989; Hodges et al., 1972; McLachlan & Stewart, 1975). The disulfide-bonded dimer of the peptide was used to avoid complications associated with concentration-dependent exchange rates and also was used because of increased stability; in physiological saline solution, the melting temperature for the disulfide-bonded dimer is 83 °C (E. K. O'Shea, R. Rutkowski, and P. S. Kim, in preparation).

Materials and Methods

Peptide Synthesis. The peptide synthesis and purification procedure used for GCN4-p2N are similar to those described previously for GCN4-p1 (O'Shea et al., 1989). Fast-atom-bombardment mass spectrometry (MIT Mass Spectrometry

† This work was supported by grants from the Rita Allen Foundation, the Pew Memorial Trust, and the NIH (GM44162 and RR05927).

‡ Present address: Genetics Institute, One Burtt Road, Andover, MA 01810.

0006-2960/91/0430-11615$02.50/0

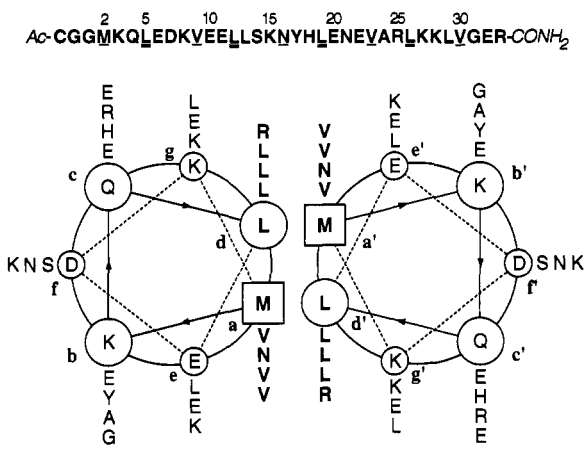

FIGURE 1: (A, top) Amino acid sequence of GCN4-p2N. Residues Met2–Arg33 of GCN4-p2N correspond to the 32 C-terminal residues of GCN4. Residue numbering is the same as used previously (O'Shea et al., 1989, 1991; Oas et al., 1990; Hu et al., 1990). The N-terminal Cys is acetylated and the C-terminal Arg is amidated. The 4–3 hydrophobic repeat is comprised of residues at positions a (underlined) and d (double underlined) of the heptad repeat. Position d is predominantly leucine in most leucine zipper sequences. (B, bottom) The leucine zipper of GCN4-p2N (residues 2–33) represented on a coiled-coil helical wheel. The heptad repeat is described by the letters a–g which comprise two turns of the helix.

Facility) for reduced GCN4-p2N gave a molecular weight (M_r) of 4098.8 (calculated M_r = 4098.9). Reduced GCN4-p2N, purified by reverse-phase HPLC, was allowed to air oxidize overnight in H_2O at room temperature, pH 8.5. The product was lyophilized and used without further purification. Analytical HPLC was used to check the purity of the peptide and to verify complete oxidation.

Amide Proton Exchange. pH* refers here to meter readings in D_2O solutions using a glass pH electrode, without correction for isotope effects. H–D exchange was initiated by dissolving fully protonated GCN4-p2N in exchange buffer (0.2 M NaCl, 50 mM Na_2DPO_4/NaD_2PO_4 in D_2O, pH* 7.0) at 6 °C. The pH* was readjusted quickly to 7.0, and the first time point (referred to as time t = 0) was taken 4–5 min after the start of this procedure. The concentration of GCN4-p2N in the exchange mixture was 6 mM (in monomer) as determined by tyrosine absorption (Edelhoch, 1967). The temperature was maintained at 6 °C with a circulating water bath. H–D exchange was quenched by a 2-fold dilution of aliquots of the exchange mixture into quench buffer (0.2 M deuterated malonic acid in D_2O, pH* 2.0) to give a final pH* of 3.0. Aliquots were removed at the following times: 0, 10, and 40 min, 2, 4, and 12 h, and 1, 2, and 10 days, respectively. The resulting samples (3 mM peptide concentration in monomer) were frozen and stored at −20 °C until required for NMR analysis.

NMR Spectroscopy. Phase-sensitive double quantum filtered (DQF) COSY (Shaka & Freeman, 1983; Rance et al., 1983) spectra were recorded at 20 °C on a Bruker AMX spectrometer operating at 500 MHz. Spectra were acquired consecutively as 256 t_1 × 1024 t_2 complex data points using a sweep width of 5555 Hz in both dimensions. The total aquisition time was approximately 9 h per sample. Spectra were processed identically on a SUN workstation to a final resolution of 1.4 Hz/point in both dimensions, using the program FTNMR (courtesy of Dr. Dennis Hare, Hare Research Inc., Woodinville, WA).

The H^N–H^α fingerprint cross-peak region of GCN4-p2N was assigned by comparison to previously published assignments for the peptide, GCN4-p1, which has an identical sequence between Met2–Arg33 (Oas et al., 1990). In addition, an unambiguous H^N_i–H^N_{i+1} trace from Gly1 to Arg33 was obtained from NOESY spectra of GCN4-p2N at one or more temperatures (20, 30, or 40 °C).

Data Analysis. Cross-peak heights (I) were obtained from the mean of the absolute value of the height for each of the four antiphase wings (or less than four, where overlap was severe) of each H^N–H^α cross-peak. These were normalized relative to the corresponding value for the nonexchangeable H^δ–H^ϵ cross-peak of Tyr17 for each data set. Exchange rates, k_{ex}, were determined by linear regression from the slope of the line of ln (I/I_0) vs time (where I and I_0 are H^N–H^α cross-peak heights at times $t = t$ and $t = 0$, respectively).

Intrinsic exchange rates, k_{int}, predicted for the amide proton in the absence of any folded structure, were calculated using data from unfolded peptides (Englander et al., 1979) after correction for sequence-specific nearest-neighbor inductive effects (Molday et al., 1972). A value of 6.6 for the pK_a of His18 (determined by ^1H NMR at 6 °C; data not shown) was used to calculate a weighted average correction at pH* 7.0 using the values of λ and ρ for His^+ and His (Molday et al., 1972).

Conformational-dependent H^N chemical shifts ($\Delta\delta$) were calculated by subtraction of the predicted intrinsic chemical shifts for unstructured tetrapeptides (Bundi & Wüthrich, 1979) from the observed values for each amide (Table I). Positive values of $\Delta\delta$ indicate downfield-shifted resonances relative to intrinsic values. Ring current effects were ignored since the two aromatic residues, Tyr17 and His18, are located at the surface and are not in close proximity to any backbone amide proton (O'Shea et al., 1991). Both aromatic rings are probably freely rotating; Tyr17 shows fast ring flipping on the NMR time scale and His18 has a relatively unperturbed pK_a (data not shown). Predicted H-bond lengths were calculated using $d^3 = [19.2/(\Delta\delta + 2.3)]$, were d = H···O H-bond length in angstroms (Pardi et al., 1983).

Results and Discussion

GCN4-p2N Amide Proton Exchange. Figure 2 shows contour plots of the H^N–H^α region of DQF-COSY spectra of GCN4-p2N after exchange periods of 0 and 12 h, respectively. It is immediately apparent that different amide protons have very different exchange rates. Determination of the exchange rate constants for amide protons with rates that differ by 2 orders of magnitude is illustrated in Figure 3.

Exchange rates could be measured for 24 of the 32 residues in the leucine zipper peptide. These rates, k_{ex}, together with the intrinsic exchange rates predicted for the amide proton in the absence of any structure, k_{int}, are listed in Table I. For the remaining residues (Met2, Lys3, Glu4, Asp7, His18, Gly31–Arg33, and the Cys-Gly-Gly linker), H^N–H^α cross-peaks were not observed, indicating that H–D exchange was too fast to measure. The first residue with a measurably retarded H–D exchange rate is Leu5, implying that it is involved in the first stable helical H-bond, presumably to Gly1 in the disulfide linker sequence. This is supported by the strong H^N_i–H^N_{i+1} (and weak H^N_i–H^N_{i+2}) cross-peaks between Gly1 and Met2 (and Gly1 and Lys3) observed in the NOESY spectrum of GCN4-p2N (data not shown), indicating a helical conformation at Gly1. The C-terminal three residues (Gly31–Arg33) exhibit exchange rates too fast to measure.

The degree of retardation of exchange for the amide protons in GCN4-p2N is up to 10^6-fold (Table I). For comparison, this level of protection from exchange is similar to that of the stable protons in native cytochrome *c* (Jeng et al., 1990), greater than that for apomyoglobin (Hughson et al., 1990)

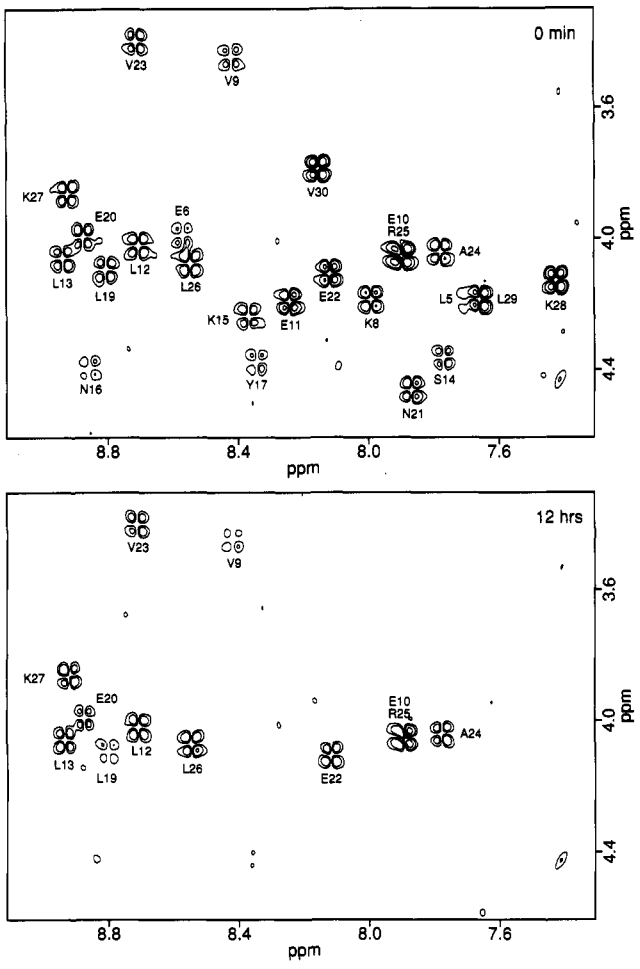

FIGURE 2: Expanded contour plots of the H^N–H^α (fingerprint) region of DQF-COSY spectra for GCN4-p2N before and after H–D exchange for 12 h at 6 °C and pH* 7. The spectra were processed and displayed in an identical manner.

and less than that for the core protons in the β-sheet of BPTI (Wagner & Wüthrich, 1982).

Periodicity of Amide Proton Exchange. There is a periodic variation of the degree of protection from exchange with residue position in GCN4-p2N that correlates with position in the coiled-coil heptad repeat (a–g), as shown in Figure 4A. Apart from residues at the N- and C-terminus of GCN4-p2N, amide protons at positions a, b, d, and e in the coiled-coil heptad show the greatest protection and those at positions c and f show the least protection from exchange. In some cases, amide protons that are adjacent in primary sequence have protection factors that differ by approximately 100-fold (Table I).

Of the three complete heptad repeats for which exchange rates were measured (Val9–Leu29), one might expect the first and third repeats, which are closer to the helix termini, to be less stable than the central repeat and therefore to exhibit faster exchange rates. Surprisingly, however, the central repeat (Asn16–Glu22) shows faster exchange rates on average (Figure 4A). This heptad is unusual since it contains a polar residue (Asn16) at the predominantly hydrophobic a position. Although this Asn is conserved in other leucine zipper proteins, genetic studies indicate that Asn16 is not required for dimerization of the GCN4 leucine zipper (Hu et al., 1990). In the crystal structure of GCN4-p1, the Asn16 residues in the homodimer have different side-chain conformations (O'Shea et al., 1991). NMR studies, however, suggest that the entire GCN4 leucine zipper, including Asn16, is symmetrical in solution (Oas et al., 1990). Rapid interconversion of the Asn16

Table I: Amide Proton Exchange Rates and Chemical Shifts for GCN4-p2N[a]

residue	k_{ex} ($\times 10^{-3}$ min^{-1})	k_{int} (min^{-1})	log (k_{ex}/k_{int})	H^N chemical shift (ppm)
Ac-Cys(−1)	>300[b]	144		8.63
Gly0	>300[b]	256		8.95
Gly1	>300[b]	256		8.63
Met2	>300[b]	91.0		7.89
Lys3	>300[b]	57.4		8.35
Gln4	>300[b]	129		8.14
Leu5	90[c]	64.4	−2.9	7.68
Glu6	6.7	45.6	−3.8	8.55
Asp7	>300[b]	57.4	>−2.3[d]	8.37
Lys8	5.3	64.4	−4.1	7.97
Val9	0.59	91.0	−5.2	8.41
Glu10	0.35[e]	45.6	−5.1	7.88
Glu11	2.8	45.6	−4.2	8.23
Leu12	0.20	45.6	−5.4	8.68
Leu13	0.14	45.6	−5.5	8.92
Ser14	24	144	−3.8	7.77
Lys15	2.8	115	−4.6	8.35
Asn16	2.3	229	−5.0	8.85
Tyr17	1.5	102	−4.9	8.32
His18	>300[b]	197	>−2.8[d]	8.05
Leu19	0.75	145	−5.3	8.77
Glu20	0.32	45.6	−5.2	8.85
Asn21	8.5	115	−4.1	7.84
Glu22	0.59	129	−5.3	8.11
Val23	0.052	45.6	−6.0	8.69
Ala24	0.21	45.6	−5.3	7.76
Arg25	0.35[e]	57.4	−5.2	7.88
Leu26	0.17	91.0	−5.7	8.51
Lys27	0.23	57.4	−5.4	8.91
Lys28	13	115	−4.0	7.41
Leu29	9.0	91.0	−4.0	7.63
Val30	23	45.6	−3.3	8.11
Gly31	>300[b]	129		7.88
Glu32	>300[b]	91.0		7.98
Arg33	>300[b]	57.4		8.12

[a] k_{ex} and k_{int} are the observed and predicted intrinsic exchange rates, respectively, at 6 °C and pH* 7. Amide proton chemical shifts were measured in H_2O at 20 °C and pH 3. [b] Exchange too fast to be measured under these conditions; estimated $k_{ex} > 0.3$ min^{-1}. [c] Partial cross-peak overlap with Leu29; estimated range $0.04 < k_{ex} < 0.14$ min^{-1}. [d] Estimated lower limit for log (k_{ex}/k_{int}). [e] Cross-peaks for Glu10 and Arg25 are unresolved; k_{ex} for these residues are the same to within a factor of 3.

side-chain conformations in solution could reconcile these observations. Our results support the suggestion that the region near Asn16 is slightly destablized and involved in greater motional flexibility. In addition, Val23 was found to be the least tolerant to substitution in the hydrophobic core in genetic experiments (Hu et al., 1990), and this residue exhibits the slowest amide exchange rate, presumably reflecting a higher degree of local stability.

Comparison with Predicted H-Bond Lengths. A striking periodicity in the chemical shifts of the amide protons of GCN4-p1 has been reported previously (Kuntz et al., 1991). This periodicity is also observed for GCN4-p2N (Figure 4B). Proton chemical shifts have been shown to correlate with H-bond length for amide protons of bovine pancreatic trypsin inhibitor (BPTI) and homologous proteins in aqueous solution (Pardi et al., 1983; Wagner et al., 1983) and of lysozyme (Redfield & Dobson, 1990), with shorter H-bonds giving more downfield shifted NMR signals. Predicted H-bond lengths for GCN4-p2N were estimated using a previously determined correlation (Pardi et al., 1983) of conformational-dependent chemical shifts, $\Delta\delta$, and the resulting values are shown in Figure 4B. We emphasize that the correlation between $\Delta\delta$ and H-bond length used here was determined empirically for BPTI (Pardi et al., 1983). A slightly different correlation has been found in lysozyme (Redfield & Dobson, 1990). In the

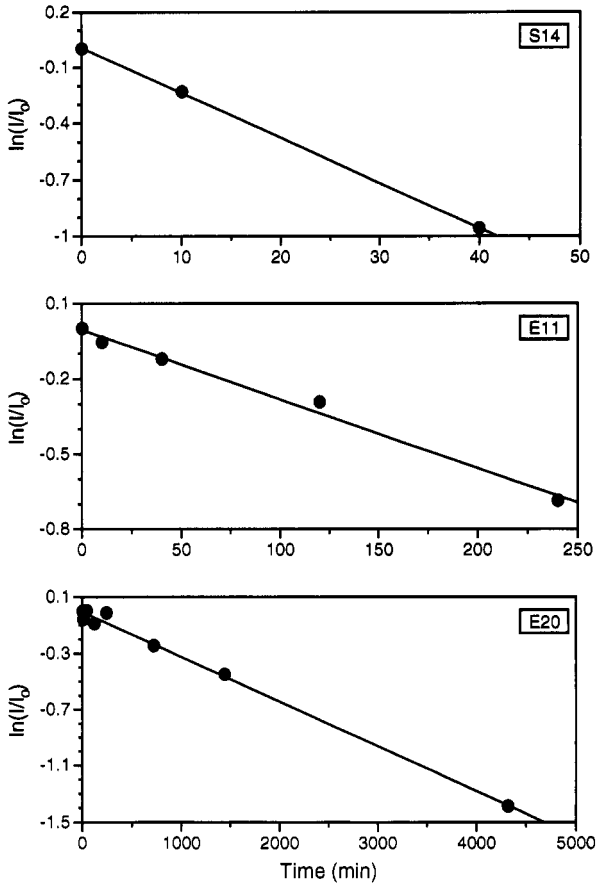

FIGURE 3: Plots of ln (I/I_0) vs time (where I and I_0 are normalized mean H^N–H^α cross-peak heights at times $t = t$ and $t = 0$, respectively) for Ser14, Glu11, and Glu20. The exchange rates were obtained from the slope of the lines using a least-squares fit to the data.

solid state, a significantly different correlation was found for the hydroxyl protons of minerals (Yesinowski et al., 1988).

The periodicity of $\Delta\delta$ and the predicted H-bond length correlates with position in the coiled-coil heptad repeat (Figure 4B). In particular, amides at position e consistently have downfield-shifted H^N resonances and are predicted to have short H-bonds, whereas amides at position f have upfield-shifted H^N resonances and are predicted to have relatively long H-bonds. This predicted increase in H-bond length between positions e and f has been confirmed recently by X-ray crystallography (O'Shea et al., 1991). Periodicity in the primary sequence along a helix is thought to give rise to a variation in H-bond length, thereby producing a curving of the helix axis so that the short H-bonds lie on the concave side (Pauling & Corey, 1953). Such curving has also been observed for some amphipathic α-helices in globular proteins [e.g., Blundell et al. (1983) and Chakrabarti et al. (1986)].

Periodicity is also observed in the amide exchange data (Figure 4A). A consistently sharp change in protection from exchange between adjacent residues is observed, again involving positions e (most retarded) and f (least retarded) of the heptad repeat. A comparison of the periodicity of amide proton exchange and of $\Delta\delta$ for positions in the heptad repeat is given in Figure 4C.

Mechanism of Amide Proton Exchange. The major factors thought to be responsible for retardation of amide proton exchange rates in native proteins are H-bonding and solvent exclusion [reviews: Englander and Kallenbach (1984), Wagner (1983), and Woodward et al. (1982)]. All of the amide protons studied here (residues 5–30 of GCN4-p2N) are involved in intrahelical H-bonds within a single secondary

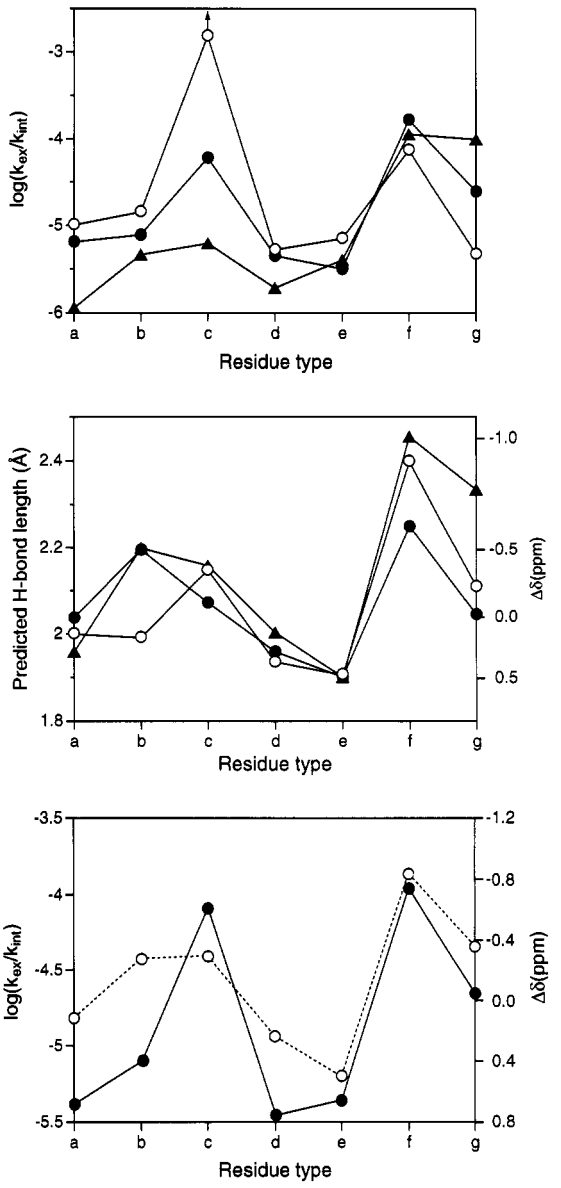

FIGURE 4: (A, top) Relative amide exchange rates for GCN4-p2N plotted as log (k_{ex}/k_{int}) vs residue type in the coiled-coil heptad repeat. k_{ex} and k_{int} are the observed and intrinsic exchange rates, respectively. Symbols: (●) first heptad repeat (Val9–Lys15); (○) second heptad (Asn16–Glu22); (▲) third heptad (Val23–Leu29). (B, middle) Conformation-dependent chemical shifts, $\Delta\delta$, and predicted H-bond length as a function of residue type at the NH position. Symbols as in (A). (C, bottom) Mean values taken over three heptad repeats (Val9–Leu29) vs residue type for (●) log (k_{ex}/k_{int}) and (○) $\Delta\delta$.

structure element (Oas et al., 1990; O'Shea et al., 1991). Nonetheless, the degrees of protection from exchange for these amide protons are highly variable, ranging from approximately 10^3- to 10^6-fold (Table I). Thus, the dominant mechanism for exchange under these conditions does not involve complete unfolding of the helices but rather highly local, dynamic events that most likely occur in the otherwise folded coiled coil.

The occurrence of the most slowly exchanging amide protons at positions a, b, d, and e can be understood using the knobs-into-holes packing scheme (Figure 5) for a two-stranded coiled coil (Crick, 1953). The knobs formed by the side chains of one helix pack into the holes of the other helix. As illustrated in Figure 5, the knob of residue 3a′ packs into the hole formed by the N–H of 3a, while the knob of residue 2d′ packs into the hole formed by the N–H of 2e. Thus the amide protons at positions a and e lie at the base of the holes into which the knobs, formed by hydrophobic side chains of the

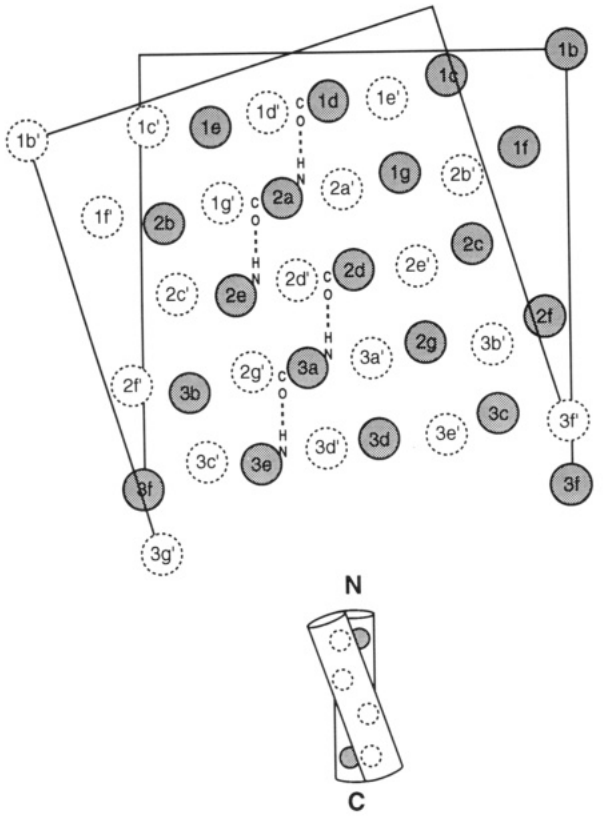

FIGURE 5: A helical net representation of the surfaces of two α-helices (oriented parallel with a crossover angle of +18°) illustrating the meshing of side chains in a knobs-into-holes packing scheme (Crick, 1953). The hydrophobic "stripe" created by residues at a and d runs down the side of the helices and forms the hydrophobic interface characteristic of a coiled coil. The side-chain knobs of the front helix (shown as dotted open circles) intercalate between the knobs of the back helix (shown as filled gray circles), thereby packing into the holes formed by the N–H and C=O groups which are shown for the back helix (i.e., filled circles). Note that the amide protons at positions a and e (shown for the back helix) lie at the base of holes into which the side-chain knobs of residues a′ and d′ of the front helix (i.e., open circles) are packed.

other helix at positions a′ and d′, are packed. The crystal structure of GCN4-p1 (O'Shea et al., 1991) shows that the hydrophobic interface extends to residues e and g, and thus the amide protons at positions d and b are also expected to be substantially buried in the dimer interface. Amide protons at positions g, f, and c are progressively further from the dimer interface, and protons at these positions show correspondingly faster exchange rates.

Bending of the helices, so that only one or a few H-bonds are broken at any given moment, is a simple mechanism that would explain the faster exchange rates for residues that are on the outside of the dimer. For the exchange conditions studied here, our results are consistent with models that involve penetration of solvent to the site of exchange (Woodward et al., 1982), although our results also emphasize the role of hydrogen bonding in retarding amide proton exchange. The correlation between the rates of amide proton exchange and the predicted H-bond lengths is intriguing but could be secondary to the effects of solvent accessibility on H-bond lengths and proton exchange rates.

Conclusion

Our results suggest that tertiary packing interactions can give rise to large differences in local stability within a secondary structural unit (and even between residues adjacent in primary sequence). This indicates that tertiary packing interactions can have a dominant role in determining local stability. Such regions of local stability involving tertiary interactions are also likely to be key determinants of protein folding.

Acknowledgments

We thank Dr. Lawrence P. McIntosh for advice and discussions in all aspects of this work and for helpful comments on an earlier version of the manuscript. We also thank Erin K. O'Shea, Pehr Harbury, and Dr. Terrence G. Oas for discussions and Rheba Rutkowski for peptide synthesis.

References

Blundell, T., Barlow, D., Borkakoti, N., & Thornton, J. (1983) *Nature 306*, 281.
Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285.
Chakrabarti, P., Bernard, M., & Rees, D. C. (1986) *Biopolymers 25*, 1087.
Cohen, C., & Parry, D. A. D. (1990) *Proteins: Struct., Funct., Genet. 7*, 1.
Crick, F. H. C. (1953) *Acta Crystallogr. 6*, 689.
Edelhoch, H. (1967) *Biochemistry 6*, 1948.
Englander, J. J., Calhoun, D. B., & Englander, S. W. (1979) *Anal. Biochem. 92*, 517.
Englander, S. W., & Kallenbach, N. R. (1984) *Q. Rev. Biophys. 16*, 521.
Hinnebusch, A. G. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6442.
Hodges, R. S., Sodek, J., Smillie, L. B., & Jurasek, L. (1972) *Cold Spring Harbor Symp. Quant. Biol. 37*, 299.
Hu, J. C., O'Shea, E. K., Kim, P. S., & Sauer, R. T. (1990) *Science 250*, 1400.
Hughson, F. M., Wright, P. E., & Baldwin, R. L. (1990) *Science 249*, 1544.
Jeng, M.-F., Englander, S. W., Elöve, G. A., Wand, A. J., & Roder, H. (1990) *Biochemistry 29*, 10433.
Johnson, P. F., & McKnight, S. L. (1989) *Annu. Rev. Biochem. 58*, 799.
Kuntz, I. D., Kosen, P. A., & Craig, E. C. (1991) *J. Am. Chem. Soc. 113*, 1406.
Landschulz, W. H., Johnson, P. F., & McKnight, S. L. (1988) *Science 240*, 1759.
Linderstrøm-Lang, K. U. (1955) *Chem. Soc., Spec. Publ. 2*, 1.
McLachlan, A. D., & Stewart, M. (1975) *J. Mol. Biol. 98*, 293.
Molday, R. S., Englander, S. W., & Kallen, R. G. (1972) *Biochemistry 11*, 150.
Oas, T. G., McIntosh, L. P., O'Shea, E. K., Dahlquist, F. W., & Kim, P. S. (1990) *Biochemistry 29*, 289.
O'Shea, E. K., Rutkowski, R., & Kim, P. S. (1989) *Science 243*, 538.
O'Shea, E. K., Klemm, J. D., Kim, P. S., & Alber, T. (1991) *Science 254*, 539.
Pardi, A., Wagner, G., & Wüthrich, K. (1983) *Eur. J. Biochem. 137*, 445.
Pauling, L., & Corey, R. B. (1953) *Nature 171*, 59.
Rance, M., Sorenensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479.
Rasmussen, R., Benvegnu, D., O'Shea, E. K., Kim, P. S., & Alber, T. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 561.
Redfield, C., & Dobson, C. M. (1990) *Biochemistry 29*, 7201.
Shaka, A. J., & Freeman, R. (1983) *J. Magn. Reson. 51*, 169.
Struhl, K. (1989) *Trends Biochem. Sci. 14*, 137.

Talbot, J. A., & Hodges, R. S. (1982) *Acc. Chem. Res. 15*, 224.
Wagner, G. (1983) *Q. Rev. Biophys. 16*, 1.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 160*, 343.
Wagner, G., & Wüthrich, K. (1986) *Methods Enzymol. 131*, 307.
Wagner, G., Pardi, A., & Wüthrich, K. (1983) *J. Am. Chem. Soc. 105*, 5948.
Woodward, C., Simon, I., & Tüchsen, E. (1982) *Mol. Cell. Biochem. 48*, 135.
Yesinowski, J. P., Eckert, H., & Rossman, G. R. (1988) *J. Am. Chem. Soc. 110*, 1367.

^{1}H NMR Studies of DNA Recognition by the Glucocorticoid Receptor: Complex of the DNA Binding Domain with a Half-Site Response Element†

M. L. Remerowski,‡ E. Kellenbach,‡ R. Boelens,‡ G. A. van der Marel,§ J. H. van Boom,§ B. A. Maler,‖ K. R. Yamamoto,‖ and R. Kaptein*,‡

Department of Chemistry, University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands, Gorlaeus Laboratory, University of Leiden, P.O. Box 9502, 2300 RA Leiden, The Netherlands, and Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143-0448

Received September 17, 1991; Revised Manuscript Received October 17, 1991

ABSTRACT: The complex of the rat glucocorticoid receptor (GR) DNA binding domain (DBD) and half-site sequence of the consensus glucocorticoid response element (GRE) has been studied by two-dimensional ^{1}H NMR spectroscopy. The DNA fragment is a 10 base-pair oligonucleotide, 5′d(GCTGTTCTGC)3′·5′d-(GCAGAACAGC)3′, containing the stronger binding GRE half-site hexamer, with GC base pairs at each end. The 93-residue GR-DBD contains an 86-residue segment corresponding to residues 440–525 of the rat GR. Eleven NOE cross peaks between the protein and DNA have been identified, and changes in the chemical shift of the DNA protons upon complex formation have been analyzed. Using these protein–DNA contact points, it can be concluded that (i) the "recognition helix" formed by residues C460–E469 lies in the major groove of the DNA; (ii) the GR-DBD is oriented on the GRE half-site such that residues A477–D481, forming the so-called D-loop, are available for protein–protein interaction in the GR-DBD dimer on the intact consensus GRE; and (iii) the 5-methyl of the second thymine in the half-site and valine 462 interact, confirming indirect evidence [Truss et al. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 7180–7184; Mader et al. (1989) *Nature 338*, 271–274] that both play an important role in GR-DBD DNA binding. These findings are consistent with the model proposed by Härd et al. [(1990) *Science 249*, 157–160] and the X-ray crystallographic complex structure determined by Luisi et al. [(1991) *Nature 352*, 497–505].

The glucocorticoid receptor (GR)[1] is a member of the steroid/thyroid hormone receptor superfamily (Danielsen, 1991; Muller & Renkawitz, 1991), which includes receptors for steroid hormones, thyroid hormones, retinoic acid, and vitamin D_3. These receptors act as hormone-regulated transcription factors and display a common functional organization, with a hormone-binding part at the carboxy terminus and an adjacent domain consisting of about 70 residues responsible for DNA binding (Carlstedt-Duke et al., 1987; Rusconi & Yamamoto, 1987).

This DNA binding region has a high degree of sequence homology within the hormone receptor superfamily (Evans, 1988; Beato, 1989). Protein fragments containing the DNA binding domain (DBD) of the glucocorticoid receptor (GR) expressed in *Escherichia coli* exhibit sequence-specific binding to glucocorticoid response elements (GREs) (Freedman et al., 1988a; Tsai et al., 1988; Dahlman et al., 1989). The structured region corresponding to C440–R510 of the rat glucocorticoid receptor (Figure 1) was shown to contain two "zinc fingers", i.e., peptide regions in which two Zn(II) ions are tetrahedrally coordinated by eight of nine cysteine residues (Freedman et al., 1988a). These metal ions are required for proper folding and DNA binding activity (Freedman et al., 1988b) but have been shown by NMR to have a conformation which is distinct from other classes of zinc-finger proteins such as TFIIIA (Härd et al., 1990a,b; Kaptein, 1991). Residues C460–E469 and P493–G504 immediately following the fingers are α-helical and are oriented perpendicular to each other with the hydrophilic surfaces exposed to the solvent and the conserved hydrophobic residues forming the protein core (Figure 2). The structure of the estrogen receptor DNA binding domain determined by NMR (Schwabe et al., 1990) is similar.

Glucocorticoid response elements (Strähle et al., 1987; Ham et al., 1988) commonly consist of a partially palindromic sequence composed of two half-site hexamers separated by three

† This research was supported by The Netherlands Organization for Chemical Research (SON) and The Netherlands Organization of Scientific Research (NWO). K.R.Y. acknowledges the support of the National Science Foundation, and M.L.R. was supported by an NSF–NATO postdoctoral fellowship award (RCD-9050102).

‡ Univesity of Utrecht.
§ University of Leiden.
‖ University of California at San Francisco.

[1] Abbreviations: CIDNP, chemically induced dynamic nuclear polarization; DBD, DNA binding domain; DTT, dithiothreitol; ERE, estrogen response element; FID, free induction decay; GR, glucocorticoid receptor; GRE, glucocorticoid response element; NMR, nuclear magnetic resonance spectroscopy; NOE, nuclear Overahuser effect; NOESY, 2D NOE spectroscopy; TRE, thyroid response element; TOCSY, 2D total correlation spectroscopy; TFIIIA, transcription factor IIIA.

0006-2960/91/0430-11620$02.50/0