Volume 254, number 1,2, 159–164 FEB 07548 August 1989

Complete assignment of the ^1H NMR spectrum of a synthetic zinc finger from *Xfin*

Sequential resonance assignments and secondary structure

Min S. Lee, John Cavanagh and Peter E. Wright

Department of Molecular Biology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, CA 920737, USA

Received 26 June 1989

A 25-residue synthetic peptide corresponding to zinc finger 31 of the *Xenopus* protein *Xfin* adopts a compact, folded conformation in the presence of zinc. Complete ^1H resonance assignments have been made. The peptide contains a helix, beginning as an α-helix and ending as a 3_{10}-helix, that extends from residue 12 to 23. Several positively charged and polar residues located on this helix are likely to be involved in interactions with DNA. Residues 1–10 appear to adopt a hairpin-like structure.

Zinc finger; DNA binding; NMR, 2D; Secondary structure

1. INTRODUCTION

Zinc fingers are important DNA-binding motifs found in a large number of eukaryotic transcriptional regulatory proteins. Transcription factor IIIA (TFIIIA) isolated from *Xenopus laevis* [1] was the first to be identified as a zinc finger protein. These proteins bind to DNA (or RNA) in a sequence-specific manner and are of fundamental importance for studying the developmental regulation of gene expression, as well as basic aspects of transcription mechanisms [2–4]. Only a few of these regulatory proteins have been characterized in higher eukaryotes; these are largely represented by steroid hormone receptors that are cell-specific mediators of the control of transcription by their respective hormones during development. Two principal classes of zinc finger domain that differ in the nature of the zinc ligands have been identified. In the TFIIIA type of finger, zinc is coordinated by 2 Cys and 2 His, whereas in the steroid the zinc is coordinated by 4 Cys [3,4]. It has since been shown that zinc is required for correct folding of both classes of finger domain and is essential for specific DNA recognition [5–9].

EXAFS measurements have confirmed the identity of the zinc ligands for the two principal classes of zinc finger [8,10]. Also, CD and NMR spectroscopy have been used to demonstrate the folding of synthetic single fingers from TFIIIA and ADRI [11,12]. These single fingers have been shown to bind in a non-specific but zinc-dependent manner to DNA. Insight into the three-dimensional structure of the zinc finger motif is essential for our understanding of the molecular basis for sequence-specific recognition of DNA. We have thus undertaken NMR studies of the structure of a single zinc finger from the *Xenopus* protein *Xfin* [13]. A 25-residue peptide with a sequence of finger 31 from *Xfin* (denoted *Xfin-31*) was selected for structural study because it closely corresponds to the consensus sequence derived form 148 zinc finger domains [14]. In particular, it seemed important to choose a finger with only two residues

Correspondence address: P.E. Wright, Department of Molecular Biology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, CA 920737, USA

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

separating the two Cys ligands, since our NMR studies of finger 5 from TFIIIA (with 4 residues between the Cys ligands) showed significant loss of zinc at pH lower than ~6.5. We present here complete sequence-specific resonance assignments and identification of the secondary structure of *Xfin-31*.

2. MATERIALS AND METHODS

2.1. *Peptide preparation*

The *Xfin-31* (Ac-YKCGLCERSFVEKSALSRHQRVHKN-NH_2) peptide was synthesized on an Applied Biosystems synthesizer and purified using a Phenomenex C4 semi-preparative reverse-phase column on a Hitachi HPLC system. The peptide was eluted with a solvent gradient consisting of 0.05% trifluoroacetic acid/acetonitrile/water mixture. The sample used for NMR studies was > 95% pure by analytical HPLC.

The zinc complex of *Xfin-31* was prepared for NMR experiments as a 5.8 mM solution in 90% H_2O/10% D_2O. The free peptide was dissolved in H_2O and the pH value adjusted to pH 7.2 using microliter amounts of 1 M NaOH. To this was added a 1.5-fold molar excess of $ZnCl_2$ in ^2H_2O and the pH was readjusted to pH 5.5. The concentration of the solution of $ZnCl_2$ in ^2H_2O was such that the final solvent composition was 90% H_2O/10% ^2H_2O. All NMR experiments were carried out using this sample. Confirmation that the amino acid sequence is correct was provided by the sequential assignment procedure.

2.2. *NMR measurements*

The ^1H NMR spectra were recorded at 278, 288 and 298 K using a Bruker AM500 spectrometer equipped with an Aspect 3000 computer and digital phase-shifting hardware. All data processing was carried out on CONVEX C1 and C240 computers using software provided by Dr Dennis Hare.

Two-dimensional NMR spectra were recorded in the phase-sensitive mode with quadrature detection in the ω_1 dimension using time-proportional phase incrementation [15,16]. All spectra were measured with the carrier placed on the ^1H_2O resonance.

Double quantum filtered COSY (2QF-COSY) spectra were acquired using the standard pulse sequence and phase cycling [17]. Spectra were acquired with 48 or 64 scans per t_1 value and 512 t_1 values. A double quantum spectrum was recorded with a 30 ms excitation period and without water saturation during mixing [18]. 400 t_1 points were acquired with 16 scans per t_1 value. Phase-sensitive NOESY spectra were acquired using standard methods [19] with 100, 200 and 300 ms mixing periods at a sample temperature of 278 K. In each case 512 t_1 were recorded with either 64 or 128 scans per t_1 poin. A NOESY spectrum with 300 ms mixing period was also acquired at 298 K. TOCSY spectra were acquired by the method of Rance [20] but using the DIPSI-2 sequence for isotropic mixing [21]. The DIPSI sequence effectively eliminates chemical shift evolution and leaves only the scalar coupling interaction active thus allowing optimal magnetization transfer. Such a TOCSY spectrum was obtained with a mixing timing of 100 ms at a spin-locking field strength of 6.25 kHz.

Spectra were Fourier transformed using a Lorentzian to Gaussian weighting function in the ω_2 dimension and a shifted sine bell weighting function for the ω_1 dimension. Baseline correction [22] and t_1 ridge suppression [23] were routinely used for cosine modulated experiments. For most experiments, the final spectrum contained 2048 real points in both dimensions.

3. RESULTS AND DISCUSSION

Substantial changes occur in the one-dimensional ^1H NMR spectrum of the *Xfin-31* peptide upon binding of zinc (fig.1). Most notably, several methyl proton resonances are shifted both upfield and downfield from the overlapped group of methyl resonances between 0.8 and 0.9 ppm in the zinc-free peptide. Aromatic proton resonances are also perturbed (not shown) and the dispersion of the amide proton resonances increases considerably when zinc is bound (fig.1). This characteristic shift of proton resonances was observed for both *Xfin-31* and finger 5 of TFIIIA (fig.1). The spectral changes observed indicate formation of a well-defined folded structure in the

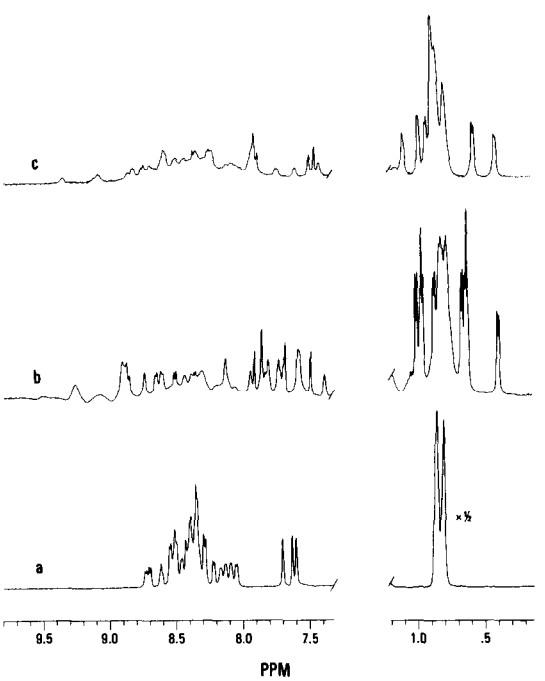

Fig.1. Amide and methyl regions of ^1H NMR spectra of Zn finger peptides in H_2O. (a) Free *Xfin-31* peptide in the absence of zinc at 278 K and pH 5.4 (b) *Xfin-31* in the presence of zinc at 278 K and pH 5.5. (c) Finger 5 of TFIIIA in the presence of zinc at 278 K and pH 5.5.

presence of zinc. *Xfin-31* binds zinc to an extent greater than 90% at pH 5.5 and is fully bound at pH 7. In contrast, zinc tends to dissociate from the TFIIIA peptide at pH < 7, making sequential assignments difficult. The decreased stability of this TFIIIA zinc finger is probably due to the increased size of the Cys-XXXX-Cys chelate ring. Two-dimensional NMR spectra indicate that the zinc complex of the *Xfin-31* peptide adopts a single conformation in aqueous solution at pH > 5.5 at temperatures ranging from 5 to 25°C. The spectral changes observed on zinc binding are very similar to those described by Klevit and co-workers [12] for a synthetic zinc finger from ADR1. In particuar, the chemical shifts of the His $C^{\epsilon}H$ and $C^{\delta}H$ resonances are very similar in the ADR1 and *Xfin* zinc fingers and differ substantially from the values observed for the free peptides. The two His and two Cys residues have been strongly implicated in zinc binding [10–12]. Our present NMR data are fully in accord with Cys_2/His_2 coordination (see below).

Complete sequence-specific resonance assignments for the backbone and side chain protons of *Xfin-31* were obtained using standard methods. Complete spin system assignments were made first using 2QF COSY, 2Q, TOCSY and NOESY spectra [24,25] followed by sequence-specific assignment using the standard sequential assignment procedure [26]. The sequential NOE connectivities observed are summarized in fig.2 and the assignments are summarized in table 1. Representative regions of the NOESY spectra showing $C^{\alpha}H_i$–NH_{i+1} and NH_i–NH_{i+1} sequential NOE connectivities are shown in fig.3. The well-resolved NH-$C^{\alpha}H$ fingerprint region of the NOESY spectrum allowed sequential assignments to be made in a straightforward manner. Both the NH and $C^{\alpha}H$ proton resonances are well dispersed with chemical shifts ranging from 7.0 to 9.3 ppm (NH) and from 3.15 to 5.16 ppm ($C^{\alpha}H$). All of the NH-$C^{\alpha}H$ cross-peaks could be assigned in the 2QF COSY spectrum except for Cys 6, the $C^{\alpha}H$ resonance of which is under the water resonance. However, the Cys 6 NH-$C^{\alpha}H$ connectivity could be identified in a 2Q spectrum ($\tau = 30$ ms) recorded without water saturation during mixing.

The observed pattern of sequential and medium range NOE connectivities (fig.2) provides information about the secondary structure of the *Xfin-31*

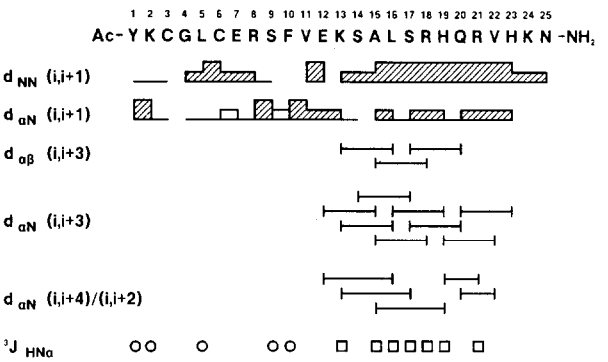

Fig.2. Summary of sequential and medium-range NOE connectivities for the *Xfin-31* zinc finger. Sequential $(i,i+1)$ NOEs are represented by lines or hatched blocks. The height of the block is a qualitative meaure of NOE intensity in the $\tau_m = 300$ ms NOESY spectrum at 278 K. The blank blocks represent sequential connectivities observed in a 300 ms NOESY spectrum at 278 K. Medium-range $(i,i+2)$, $(i,i+3)$, and $(i,i+4)$ NOEs are represented by bars connecting the appropriate residues. These NOEs were observed in a 200 ms NOESY spectrum at 278 K. Small (< 6.0 Hz) and large (> 8.0 Hz) values of $^3J_{HN\alpha}$ are indicated by squares and circles, respectively.

zinc finger. Most striking is the continuous series of d_{NN} NOE connectivities from Lys 13 to Asn 25 at the C-terminus, together with many medium-range $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ NOEs throughout the sequence between Glu 12 and His 23. This pattern of NOE connectivities is diagnostic of helix [27]. The presence of helix is confirmed by the small coupling constants ($^3J_{HN\alpha} < 6$ Hz) observed for Lys 13, Ala 15, Leu 16, Ser 17, Arg 18, His 19 and Arg 21. In addition, serveral $d_{\alpha N}(i,i+4)$ NOEs are observed in the region Glu 12–His 19, indicating the presence of α-helix [27]. There are no $d_{\alpha N}(i,i+4)$ NOEs to residues following His 19: instead $d_{\alpha N}(i,i+2)$ NOEs occur between His 19/Arg 21 and Gln 20/Val 22, suggesting that the helix changes to 3_{10} at the C-terminal end.

Sequential $d_{\alpha N}$ NOE connectivities of medium and strong intensity are observed from Tyr 1 to Lys 2, Cys 6 to Glu 7 and in the region from Arg 8 to Lys 13. Sequential d_{NN} NOE connectivities in this region are mostly very weak or absent. The large $^3J_{HN\alpha}$ coupling constants at Tyr 1, Lys 2, Ser 9 and Phe 10 suggest the presence of extended chain conformations at these residues, and the observation of a Tyr 1–Phe 10 NH-NH NOE suggests that these regions pack against each other. Also, a Cys 3 to Ser 9 NH-$C^{\alpha}H$ NOE was observed in a 300 ms

Table 1

^1H *NMR* Chemical Shifts of *Xfin-31* Zinc Finger at p*H* 5.5 and 278 K

	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	$C^{\delta}H$	Others
1 Tyr	8.90	4.48	3.03,2.79		7.08	6.89($C^{\epsilon}H$)
2 Lys	8.70	4.69	1.81,1.66	1.52,1.32	1.72,1.54	2.95,2.95($C^{\epsilon}H$) 7.56 ($N^{\zeta}H$)
3 Cys	8.95	4.34	3.45,2.81			
4 Gly	8.95	4.23,3.88				
5 Leu	9.30	4.45	0.88,0.73	1.42	0.74,0.70	
6 Cys	7.76	4.92	3.37,3.26			
7 Glu	8.65	4.31	2.24,2.06	2.33,2.18		
8 Arg	8.47	4.01	1.36,1.26	1.82,1.53	3.13,2.94	7.19($N^{\epsilon}H$)
9 Ser	7.91	5.18	3.45,3.45			
10 Phe	8.65	4.70	3.41,2.69		7.27	6.84($C^{\epsilon}H$) 6.13 ($C^{\zeta}H$)
11 Val	8.92	4.06	2.24	1.09,1.04		
12 Glu	7.88	4.83	2.04,2.04	2.36,2.22		
13 Lys	8.78	3.17	1.45,1.13	0.95,0.95	1.48,1.44	2.82,2.82($C^{\epsilon}H_2$) 7.60 ($N^{\zeta}H$)
14 Ser	8.94	4.09	3.89,3.89			
15 Ala	7.02	4.06	1.68			
16 Leu	7.09	3.22	2.02,1.34	1.53	1.05,0.95	
17 Ser	8.17	4.25	3.90,3.90			
18 Arg	7.85	3.99	1.84,1.80	1.53,1.53	3.27,3.16	7.32($N^{\epsilon}H$)
19 His	7.62	4.31	3.13,2.83		7.03	7.90($C^{\epsilon}H$)
20 Gln	7.99	3.63	2.23,2.23	2.87,2.70		7.63,7.43($N^{\epsilon}H$)
21 Arg	7.23	4.03	1.86,1.86	1.72,1.72	3.21,3.21	7.28 ($N^{\epsilon}H$)
22 Val	7.78	3.92	1.96	0.71,0.47		
23 His	7.21	4.74	3.25,3.06		6.46	7.96($C^{\epsilon}H$)
24 Lys	7.86	4.28	1.89,1.79	1.42,1.42	1.67,1.67	2.96,2.96($C^{\epsilon}H_2$) 7.60($N^{\zeta}H$)
25 Asn	8.55	4.68	2.85,2.77		7.53,7.28($N^{\delta}H$)	

Chemical shifts are referenced to internal dioxane and are generally accurate to ± 0.01 ppm (± 0.03 ppm for geminal protons separated by < 0.1 ppm)

NOESY spectrum recorded at 298 K. Strong and medium sequential d_{NN} NOEs in the region between Gly 4 and Arg 8 and a Cys 3–Cys 6 NH-NH NOE indicate a turn in the vicinity of the cysteine ligands. Thus, residues 1–10 appear to adopt a hairpin-like structure.

Numerous long-range NOEs are observed between residues distant in the sequence (fig.4), indicating that the peptide folds into a compact structure in the presence of zinc. NOEs are observed between Tyr 1 and protons of Phe 10, Val 11, Glu 12 and Lys 13, providing evidence that the N-terminus packs against the core of the molecule. NOEs are also observed between the conserved hydrophobic residues Phe 10 and Leu 16, which, from the patterns of NOEs observed to other amino acids (fig.4), appear to be in the interior of the zinc finger. Note that NOEs are observed between sidechain protons of Cys 3 and His 19, Cys 6 and His 23, consistent with the involvement of these residues in binding of zinc.

The observed patterns of long-range and sequential NOE connectivities suggest that the secondary structure and three-dimensional structure of the *Xfin* zinc finger is rather like the model proposed by Berg [28] on the basis of the conformations of homologous metal-binding loops in metalloproteins. The peptide clearly contains a helix, beginning as an α-helix and ending as a 3_{10}-helix, that extends from Glu 12 to His 23 and is substantially longer than that in Berg's model [28] or suggested by previous NMR studies [12]. Several positively charged and polar residues are located on this helix and are likely to be involved in interactions with DNA [14,28]. Residues 1–10 appear to adopt a hairpin-like structure, also in agreement with Berg's model. Given that the zinc is coordinated to the Cys 3, Cys 6, His 19 and His 23 residues, it would be situated at one end of the molecule. Since the peptide is unfolded in the absence of zinc, the metal clearly plays an important role in stabilization of the folded structure. A schematic diagram

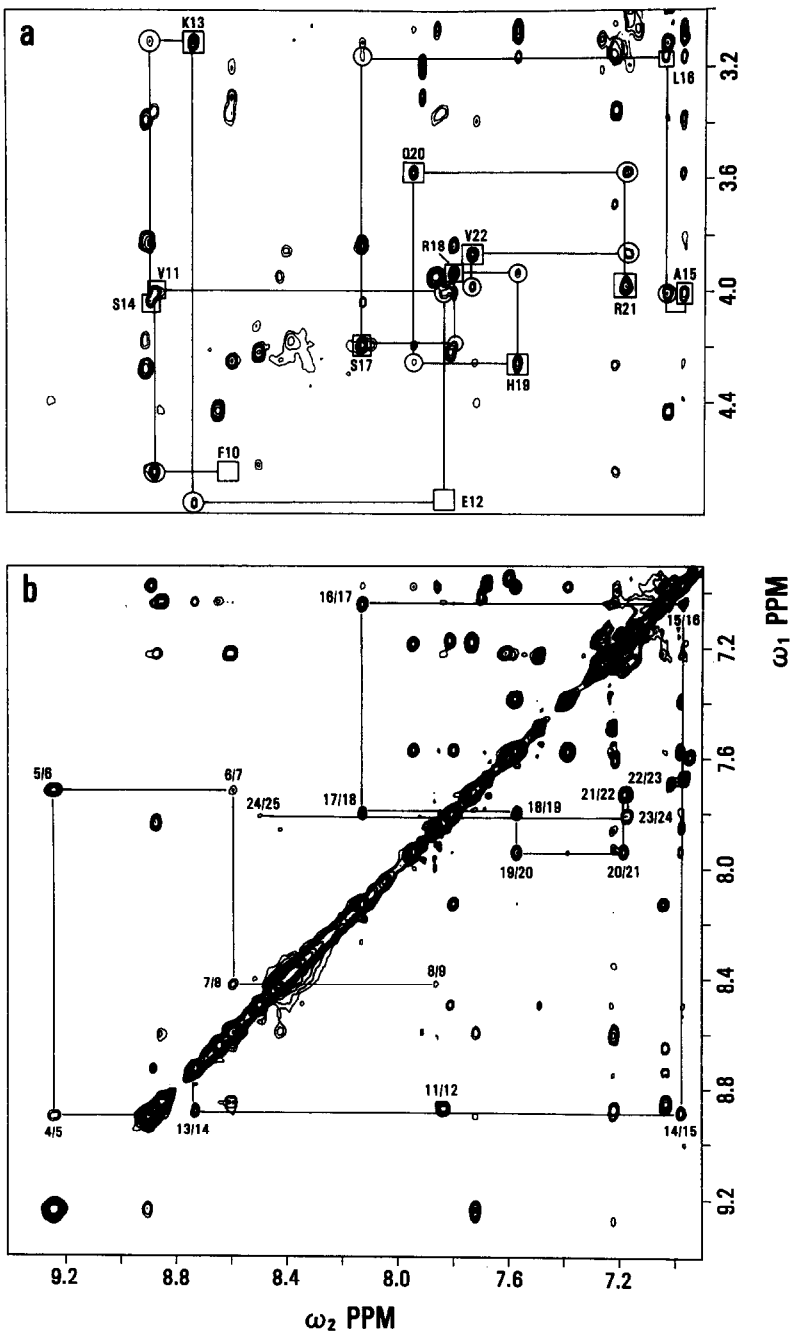

Fig.3. Two regions of a NOESY spectrum ($\tau_m = 300$ ms, 278 K) of the *Xfin-31* zinc finger plotted at the same contour levels. (a) Region showing some of the sequential connectivities. Cross-peaks corresponding to $d_{\alpha N}(i,i+1)$ NOE connectivities are encircled. Boxes indicate the position of COSY cross-peaks. (b) Region containing d_{NN} sequential connectivities.

of a zinc finger fold consistent with the present NMR studies is shown in fig.5. Determination of the three-dimensional solution structure of the *Xfin-31* zinc finger using distance geometry and restrained molecular dynamics methods is presently in progress in this laboratory and will be describ-

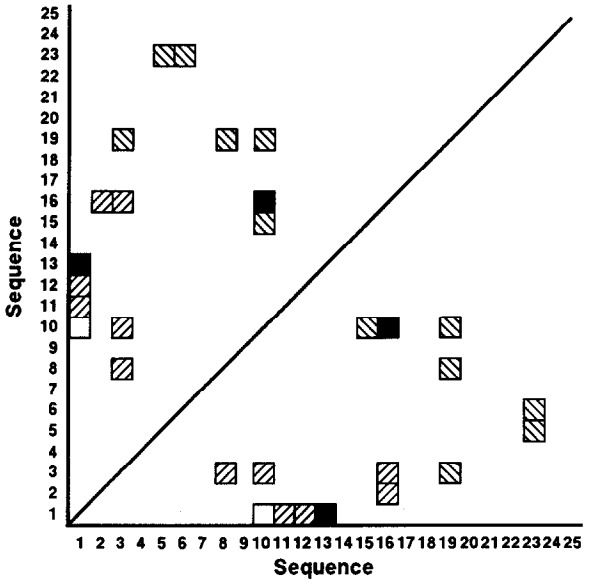

Fig.4. Diagonal plot of $(i,i+5)$ and longer range NOEs observed for *Xfin-31* in a 200 ms NOESY spectrum at 278 K. The squares represent pairs of residues linked by one or more NOEs. Backbone-backbone (□), backbone-side chain (▨), side chain-side chain (▧), and both backbone-side chain and side chain-side chain (■) NOEs are indicated.

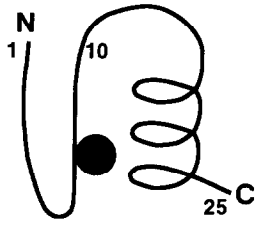

Fig.5. Schematic diagram of zinc finger structure consistent with the present NMR studies.

ed in detail elsewhere. The availability of such a structure should provide new insights into the mode of interaction of the Cys_2, His_2 class of zinc finger with DNA.

Acknowledgements: We thank Ms L. Harvey for preparation of the manuscript and the National Institutes of Health for financial support (Grants GM36643 and GM38794). M.L. is the recipient of an American Cancer Society Postdoctoral Fellowship (Grant PF-3265).

REFERENCES

[1] Miller, J., McLachlan, A.D. and Klug, A. (1985) *EMBO J.* 4, 1609–1614.
[2] Klug, A. and Rodes, E. (1987) Trents Biochem. Sci., 464–469.
[3] Evans, R.M. and Hollenberg, S.M. (1988) *Cell* 52, 1–3.
[4] Payre, F. and Vincent, A. (1988) *FEBS Lett.* 234, 245–250.
[5] Nagai, K., Nakaseko, Y., Nasmyth, K. and Rhodes, D. (1988) *Nature* 332, 284–286.
[6] Blumberg, H., Eisen, A., Sledziewski, A., Bader, D. and Young, E.T. (1987) *Nature* 328, 443–445.
[7] Redemann, N., Gaul, U. and Jäckle, H. (1988) *Nature* 332, 90–92.
[8] Freedman, L.P., Luisi, B.F., Korszun, Z.R., Basavappa, R., Sigler, P.B. and Yamamoto, K.R. (1988) *Nature* 334, 543–546.
[9] Severne, Y., Wieland, S., Schaffner, W. and Rusconi, S. (1988) *EMBO J.* 7, 2503–2508.
[10] Diakun, G.P., Fairall, L. and Klug, A. (1986) *Nature* 324, 698–699.
[11] Frankel, A.D., Berg, J.M. and Pabo, C.O. (1987) *Proc. Natl. Acad. Sci. USA* 84, 4841–4845.
[12] Párraga, G., Horvath, S.J., Eisen, A., Taylor, W.E., Hood, L., Young, E.T. and Klevit, R.E. (1988) *Science* 241, 1489–1492.
[13] Altaba, A.R., Perry-O'Keefe, H. and Melton, D.A. (1987) *EMBO J.* 6, 3065–3070.
[14] Gibson, T.B., Postma, J.P.M., Brown, R.S. and Argos, P. (1988) *Protein Engineering* 2, 209–218.
[15] Bodenhausen, G., Vold, R.I. and Vold, R.R. (1980) *J. Magn. Reson.* 37, 93–106.
[16] Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* 113, 967.
[17] Rance, M., Sorensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R., Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* 117, 479–485.
[18] Rance, M. and Wright, P.E. (1986) *J. Magn. Reson.* 66, 372–378.
[19] Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.* 71, 4546–4553.
[20] Rance, M. (1987) *J. Magn. Reson.* 74, 557–564.
[21] Shaka, A.J. Lee, C.J. and Pines, A. (1988) *J. Magn. Reson.* 77, 274.
[22] Zuiderweg, E.R.P., Boelens, R. and Kaptein, R. (1985) *Biopolymers*, 24, 601–611.
[23] Otting, G., Widmer, H., Wagner, G. and Wüthrich, K. (1986) *J. Magn. Reson*, 66, 187–193.
[24] Chazin, W.J. and Wright, P.E. (1987) *Biopolymers* 26, 973–977.
[25] Chazin, W.J. Rance, M. and Wright, P.E. (1988) *J. Mol. Biol.* 202, 603–622.
[26] Billeter, M., Braun, W. and Wüthrich, K. (1982) *J. Mol. Biol.* 155, 321.
[27] Wüthrich, K., Billeter, M., Braun, W. (1984) *J. Mol. Biol.* 180, 715–740.
[28] Berg, J.M. (1988) *Proc. Natl. Acad. Sci. USA* 85, 99–102.