Lee, J. C., Tweedy, N., & Timasheff, S. N. (1978) *Biochemistry 17*, 2788-2790.
Lee, G., Cowan, N., & Kirschner, M. (1988) *Science 239*, 285-288.
Lewis, S. A., Wang, D., & Cowan, N. J. (1988) *Science 242*, 936-939.
Lowry, O. H., Rosenbrough, N. J., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265-275.
Matsumura, F., & Hayashi, M. (1976) *Biochim. Biophys. Acta 453*, 162-175.
Mukhopadhyay, K., Parrack, P. K., & Bhattacharyya, B. (1990) *Biochemistry 29*, 6845-6850.
Olmsted, J. B., & Borisy, G. G. (1975) *Biochemistry 14*, 2996-3005.
Prasad, A. R. S., Luduena, R. F., & Horowitz, P. M. (1986a) *Biochemistry 25*, 739-742.
Prasad, A. R. S., Luduena, R. F., & Horowitz, P. M. (1986b) *Biochemistry 25*, 3536-3540.
Raybin, D., & Flavin, M. (1977) *Biochemistry 16*, 2189-2194.
Robinson, J., & Engelborghs, Y. (1982) *J. Biol. Chem. 257*, 5367-5371.
Sackett, D. L., Bhattacharyya, B., & Wolff, J. (1985) *J. Biol. Chem. 260*, 43-45.
Schiff, P. B., & Horowitz, S. B. (1981) *Biochemistry 20*, 3247-3252.
Serrano, L., De la Torre, J., Maccioni, R. B., & Avila, J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5989-5993.
Sloboda, R. D., & Rosenbaum, J. L. (1979) *Biochemistry 18*, 48-55.
Sloboda, R. D., & Rosenbaum, J. L. (1982) *Methods Enzymol. 85*, 409-415.
Wilson, L., & Bryan, J. (1974) in *Advances in Cell and Molecular Biology* (DuPraw, E. J., Ed.) Vol. 3, pp 21-72, Academic Press, New York.
Wilson, L., Jordan, M. A., Morse, A., & Margolis, R. L. (1982) *J. Mol. Biol. 159*, 125-149.

# Homo- and Heteronuclear NMR Studies of the Human Retinoic Acid Receptor $\beta$ DNA-Binding Domain: Sequential Assignments and Identification of Secondary Structure Elements†

M. Katahira,‡ R. M. A. Knegtel,‡ R. Boelens,‡ D. Eib,‡ J. G. Schilthuis,§,‖ P. T. van der Saag,§ and R. Kaptein*,‡
*Bijvoet Center for Biomolecular Research, University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands, and Hubrecht Laboratory, Netherlands Institute for Development Biology, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands*
*Received December 16, 1991; Revised Manuscript Received April 21, 1992*

ABSTRACT: An 80 amino acid polypeptide corresponding to the DNA-binding domain (DBD) of the human retinoic acid receptor $\beta$ (hRAR-$\beta$) has been studied by $^1$H homonuclear and $^{15}$N-$^1$H heteronuclear two- and three-dimensional (2D and 3D) NMR spectroscopy. The polypeptide has two putative zinc fingers homologous to those of the receptors for steroid and thyroid hormones and vitamin $D_3$. The backbone $^1$H resonances as well as over 90% of the side-chain $^1$H resonances have been assigned by $^1$H homonuclear 2D techniques except for the three N-terminal residues. The assignments have been confirmed further by means of $^{15}$N-$^1$H heteronuclear 3D techniques, which also yielded the assignments of the $^{15}$N resonances. Additionally, stereospecfic assignments of methyl groups of five valine residues were made. Sequential and medium-range NOE connectivities indicate several elements of secondary structure including two $\alpha$-helices consisting of residues E26-Q37 and Q61-E70, a short antiparallel $\beta$-sheet consisting of residues P7-F9 and S23-C25, four turns consisting of residues P7-V10, I36-N39, D47-C50, and F69-G72, and several regions of extended peptide conformation. Similarly, two helices are found in the glucocorticoid receptor (GR) DBD in solution [Härd et al. (1990) *Science 249*, 157-160] and in crystal [Luisi et al. (1991) *Nature 352*, 497-505], and in the estrogen receptor (ER) DBD in solution [Schwabe et al. (1990) *Nature 348*, 458-461], although the exact positions and sizes of the helices differ somewhat. Furthermore, long-range NOEs suggest the existence of a hydrophobic core formed by the two helices.

Retinoic acid, a vitamin A derivative, has profound effects on vertebrate cellular differentiation, pattern formation, and embryonic development (Brockes, 1989; Thaller & Eichele, 1987; Maden et al., 1988; Durston et al., 1989). Retinoic acid acts through binding to the retinoic acid receptor (RAR),[1] which functions as a ligand-inducible transcription factor. The RAR was cloned and identified as a member of a nuclear receptor superfamily which comprises the receptors for steroid

† This research was supported by the Netherlands Organization for Chemical Research (SON) and the Netherlands Organization of Scientific Research (NWO). M. K. was supported by a grant from Human Frontier Science Program for Long-Term Fellowships.
* To whom correspondence should be addressed.
‡ University of Utrecht.
§ Netherlands Institute for Developmental Biology.
‖ Present address: Ludwig Institute for Cancer Research, Courtauld Building, 91 Riding House St., London W1P 8BT, U.K.

[1] Abbreviations: NMR, nuclear magnetic resonance; DBD, DNA-binding domain; RAR, retinoic acid receptor; hRAR-$\beta$; human RAR-$\beta$; GR, glucocorticoid receptor; TR, thyroid hormone receptor; ER, estrogen receptor; RE, response element; RRE, retinoic acid response element; TRE, thyroid response element; GRE, glucocorticoid response element; ERE, estrogen response element; DTT, dithiothreitol; 2D, two-dimensional; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; HOHAHA, homonuclear Hartmann-Hahn; COSY, 2D *J*-correlated spectroscopy; DQF-COSY, double-quantum-filtered COSY; HMQC, heteronuclear multiple-quantum coherence; TFIIIA, transcription factor IIIA.

0006-2960/92/0431-6474$03.00/0 

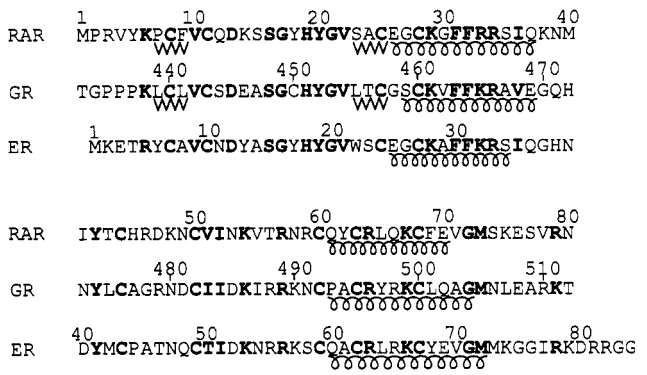


FIGURE 1: Alignment of the RAR-, GR-, and ER-DBDs on the basis of positions of Zn-coordinating cysteines. Conserved residues among three DBDs are indicated in boldface, including conservative exchanges such as K and R and V, I, and T. Spiral lines indicate regions of $\alpha$-helices, and wavy lines indicate regions of $\beta$-sheets. Numbering and secondary structures of the GR- and ER-DBDs in solution are from Härd et al. (1990b) and Schwabe et al. (1990), respectively (see text for details).

and thyroid hormones and vitamin $D_3$ (Petkovich et al., 1987; Giguere et al., 1987). So far, four different RARs, indicated by $\alpha$,$\beta$,$\gamma$, and $\delta$, have been found in human, mouse, and newt (de Thé et al., 1987; Brand et al., 1988; Benbrook et al., 1988; Zelent et al., 1989; Krust et al., 1989; Ragsdale et al., 1989; Giguere et al., 1989). A related receptor, called RXR, has also been found (Mangelsdorf et al., 1990). The DNA-binding domains (DBDs) of these RARs are highly conserved and have two putative zinc fingers homologous to those of other receptors of the superfamily. Although the DNA-binding domains of the receptors of the superfamily are homologous, they bind to response elements (REs) that differ in their nucleotide sequences as well as in spacing and direction (direct or inverted repeats) of the half-sites [for a review, see Martinez and Wahli (1990)]. For example, the retinoic acid (RRE) and thyroid response elements (TRE) are either direct or inverted repeats, while the glucocorticoid (GRE) and estrogen response elements (ERE) are exclusively inverted repeats, implying a difference in the interactions between the two monomeric proteins. Our goal is to understand the structural basis for the binding specificity of this family of zinc-finger proteins. The structures of the GR-DBD (Härd et al., 1990a) and ER-DBD (Schwabe et al., 1990) in solution have been elucidated by NMR spectroscopy. Recently, the crystal structure of the glucocorticoid receptor DNA-binding domain complexed with DNA has also been reported (Luisi et al., 1991). Comparison of RAR, GR, and ER DNA-binding domains aligned on the basis of positions of zinc-coordinating cysteine residues (Figure 1) shows about 45% homology between RAR and GR and 60% between RAR and ER in these domains. Determination of the structure of the retinoic acid receptor DBD and its comparison with those of the glucocorticoid and estrogen receptors should therefore prove useful in understanding interactions of these DNA-binding domains with their response elements.

We have studied a protein fragment of 80 amino acid residues containing the human retinoic acid receptor $\beta$ (hRAR-$\beta$) DNA-binding domain expressed in *Escherichia coli* by means of $^1$H homonuclear two-dimensional (2D) and $^{15}$N–$^1$H heteronuclear three-dimensional (3D) NMR spectroscopy. Here we present the assignments of $^1$H and $^{15}$N resonances and the secondary structure of the RAR-DBD.

## Materials and Methods

A protein fragment corresponding to amino acids 75–153 of the hRAR-$\beta$ with an extra N-terminal methionine residue was expressed in *Escherichia coli* and purified. Sequence-specific DNA-binding activity of the protein fragment was checked by gel retardation as will be reported elsewhere. A uniformly $^{15}$N-labeled protein fragment was obtained by growing cells in minimal medium with $^{15}NH_4Cl$. NMR samples (typically 2–4 mM) in either 95/5% $^1H_2O/^2H_2O$ mixtures or 99.9% $^2H_2O$ were prepared by dialysis and concentration with an Amicon flow cell, the final solution containing 200 mM NaCl, at pH 6.2–6.8, with 1 mM DTT to prevent oxidation of cysteines.

NMR spectra were recorded at 278–305 K on Bruker AM500, AMX500, and AM600 spectrometers. Phase-sensitive 2D and 3D spectra were recorded according to the time-proportional phase-incrementation method (Marion & Wüthrich, 1983). NOESY spectra (Jeener et al., 1979) were recorded with 50-, 100-, 150-, and 200-ms mixing times. 2D HOHAHA spectra (Davis & Bax, 1985) were recorded with 25-, 40-, 50-, and 60-ms spin-locking times with a "clean" pulse sequence (Griesinger et al., 1988). A DQF-COSY spectrum (Rance et al., 1983) was also recorded. For the uniformly $^{15}$N-labeled sample, the 2D $^{15}$N–$^1$H HMQC spectrum (Bax et al., 1983), the 3D $^{15}$N–$^1$H NOESY–HMQC spectrum with a 200-ms NOESY mixing time, and the 3D $^{15}$N–$^1$H HOHAHA–HMQC spectrum (Zuiderweg & Fesik, 1989; Marion et al., 1989) with a 35-ms spin-locking time with a "clean" pulse sequence were recorded. $^{15}$N decoupling was accomplished by means of a GARP sequence (Shaka et al., 1985). The water resonance was irradiated during the relaxation period (1 s), in some cases followed by a "SCUBA" pulse sequence ($t_p$ = 40 ms) preceding the preparation pulse (Brown et al., 1988) to prevent saturation of $C^\alpha H$ resonances. 2D spectra were recorded with 400–800 $t_1$ increments (128 $t_1$ increments in the HMQC experiment); 96–160 free induction decays of 2 K (4K in the DQF-COSY experiment) data points per increment were collected. 3D spectra were recorded with 128 $t_1$ and 96 $t_2$ increments; 16 free induction decays of 1K data points per increment were collected. Typical measuring times of 2D and 3D spectra were 18 and 60 h, respectively.

Data processing was carried out on a VAX Station 2000 and a Silicon Graphics IRIS 4D/35 workstation using the TRITON NMR software package developed at the Bijvoet Center, University of Utrecht. The $t_1$, $t_2$, and $t_3$ data were apodized with a $\pi/4$-shifted sine-bell function. The $t_1$ data of 2D spectra were zero-filled to 1K (2K in the DQF-COSY experiment) points. The $t_1$ and $t_2$ data of 3D spectra were zero-filled to 512 and 128 points, respectively. Fourth-order polynomial base-line corrections were applied in each frequency domain (Boelens et al., 1985).

## Results

*Stability of the Protein.* The protein fragment turned out to be stable only within rather limited pH and temperature ranges ( pH >6 and $T$ <305 K). Nevertheless, spectra obtained under conditions within these limited ranges allowed complete assignments of the backbone proton resonances except for the three N-terminal residues. The protein could not be lyophilized, due to instability, and rapid dissolution in $^2H_2O$ was therefore not possible. This made it difficult to obtain information about NH exchange rates, which could confirm the identification of secondary structure elements.

*Sequential Assignments.* Sequential assignments of $^1$H resonances were carried out according to procedures described by Wüthrich (1986) and by Englander and Wand (1987). The HOHAHA fingerprint region of the RAR-DBD is shown in Figure 2. Although there is severe overlap of resonances, careful comparison of spectra recorded at eight different

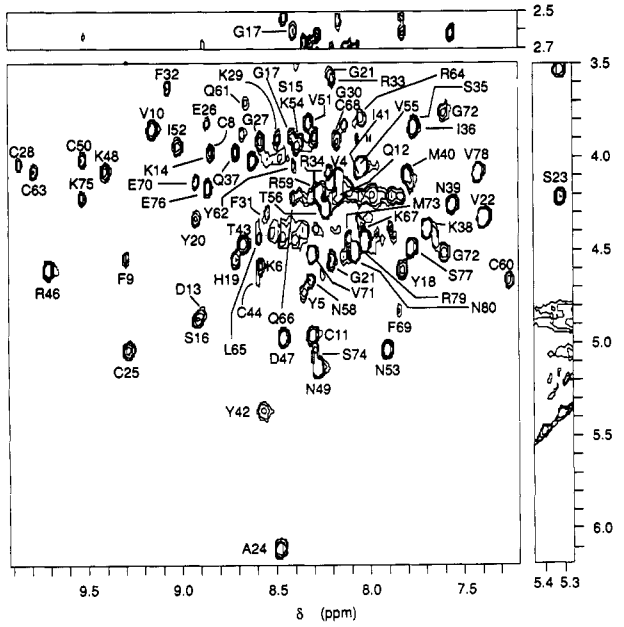


FIGURE 2: Fingerprint region of the 500-MHz HOHAHA spectrum of the RAR-DBD in $^1H_2O$ at 300 K, pH 6.5, with a 40-ms spin-lock mixing time. NH–C$^\alpha$H cross-peaks are labeled.

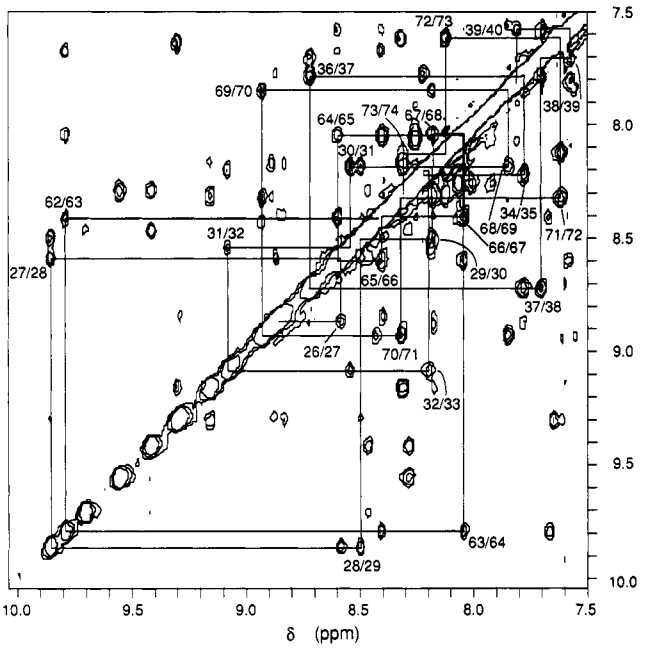


FIGURE 3: Amide region of the 600-MHz NOESY spectrum of the RAR-DBD in $^1H_2O$ at 300 K, pH 6.5, with a 200-ms mixing time, showing NH–NH connectivities in the two segments, E26–M40 and Y62–S74. Two interruptions occur at R33 and R34 and at S35 and I36 due to overlap of the respective amide proton resonances.

temperatures (278–305 K) and at four different pHs (6.2–6.8) proved to be extremely useful in resolving overlapping resonances.

Initially, all three Ile, two Thr, one Ala, and one Leu spin systems, as well as five of seven Val and three of five Gly spin systems, were identified in HOHAHA and DQF-COSY spectra. The spin systems of two His and two of four Phe residues were identified by a comparison of HOHAHA, 2D NOE, and DQF-COSY spectra. The five Tyr and the remaining two Phe residues all gave tyrosine-like cross-peak patterns due to overlap of the aromatic proton resonances of the Phe residues. Partial spin systems corresponding to the five Asn and four Gln residues were identified from strong NOE connectivities between either C$^\beta$H or C$^\gamma$H resonances and $NH_2$ resonances at 6.6–7.6 ppm.

The backbone proton resonances of several segments were assigned from uninterrupted strong to medium $d_{NN}$ or $d_{\alpha N}$ NOEs. For example, the $d_{NN}$ NOEs of the E26–R33 segment and the $d_{\alpha N}$ ones of the G17–C28 segment are shown in Figure 3 and Figure 4, respectively. The assignments of the side-chain protons, obtained from HOHAHA and DQF-COSY spectra, were confirmed by consistent $d_{\beta N}$ (Figure 5) and, occasionally, $d_{\gamma N}$ and $d_{\delta N}$ connectivities. A histidine residue (H45) was connected to R46 by the $d_{NN}$ connectivity found at 290 K, because the amide proton resonance of H45 was not observed above this temperature due to rapid exchange with water. Thus, the backbone proton resonances have been assigned completely except for the three N-terminal residues. Over 90% of the side-chain proton resonances have been assigned as well.

These assignments have been confirmed further through analyzing $^{15}N$–$^1H$ heteronuclear 3D NOESY–HMQC and HOHAHA–HMQC spectra. As an example, Figure 6 shows the $d_{NN}$ connectivities of an E26–R33 segment observed in the NOESY–HMQC spectrum. Several $d_{\alpha N}$, $d_{\beta N}$, and $d_{\gamma N}$ connectivities together with $d_{\alpha N}(i,i+3)$ ones are also indicated. Full analyses of the NOESY–HMQC and HOHAHA–HMQC spectra gave totally consistent results with those obtained from $^1H$ homonuclear 2D spectra. In addition, the 3D spectra yielded the assignments of the side-chain protons of S15, R34, K67, and C68 residues. $^{15}N$ resonance assignments obtained are listed in Table I.

Methyl groups of five Val residues were assigned stereospecifically using the method of Zuiderweg et al. (1985). For the other two Val residues, stereospecific assignments were not possible because of overlap of the methyl group resonances. All resonance assignments are summarized in Table I. The observed short- and medium-range NOEs are presented in Figure 5.

## Discussion

*Elements of Secondary Structure.* Sequential and medium-range NOEs provide a basis for identification of several secondary structure elements. An $\alpha$-helix is characterized by the stretch of the strong $d_{NN}(i,i+1)$ and weak $d_{\alpha N}(i,i+1)$ connectivities in combination with $d_{\alpha N}(i,i+3)$, $d_{\alpha N}(i,i+4)$, and $d_{\alpha\beta}(i,i+3)$ connectivities (Wüthrich, 1986). For the RAR-DBD, two helices are found for residues E26–Q37 (helix I) and Q61–E70 (helix II). The fact that the $d_{NN}(i,i+1)$ connectivities are missing at the R33–R34 and S35–I36 steps of helix I is due to the near-overlap of their amide proton resonances.

An extended structure is characterized by the stretch of strong $d_{\alpha N}(i,i+1)$ and weak or nonobservable $d_{NN}(i,i+1)$ connectivities. Extended segments are found for V4–K6, P7–F9, D13–S16, G17–G21, V22–E26, K38–M40, Y42–C44, C50–K54, M73–E76, and V78–N80. Extended backbone conformation is often associated with a $\beta$-sheet, characterized by typical interstrand NOEs, especially interstrand $d_{\alpha\alpha}$ and $d_{\alpha N}$ (Wüthrich, 1986). In the RAR-DBD, this was found only within very short stretches involving residues P7–F9 and S23–C25 (Figure 7). The $d_{\alpha\alpha}$(C8,A24) is shown in Figure 4. Figure 7 suggests the presence of hydrogen bonds between P7 CO and C25 NH, and between F9 NH and S23 CO.

A turn involving four residues is characterized by the observation of a strong $d_{NN}(3,4)$ connectivity in combination with a $d_{\alpha N}(2,4)$ connectivity (Wüthrich, 1986; Chazin & Wright, 1988). Type I and type II turns can be distinguished from each other by the intensity of the $d_{NN}(2,3)$ and $d_{\alpha N}(2,3)$ connectivities (Wüthrich, 1986). In the RAR-DBD, this characteristic pattern of connectivities signifying a turn is

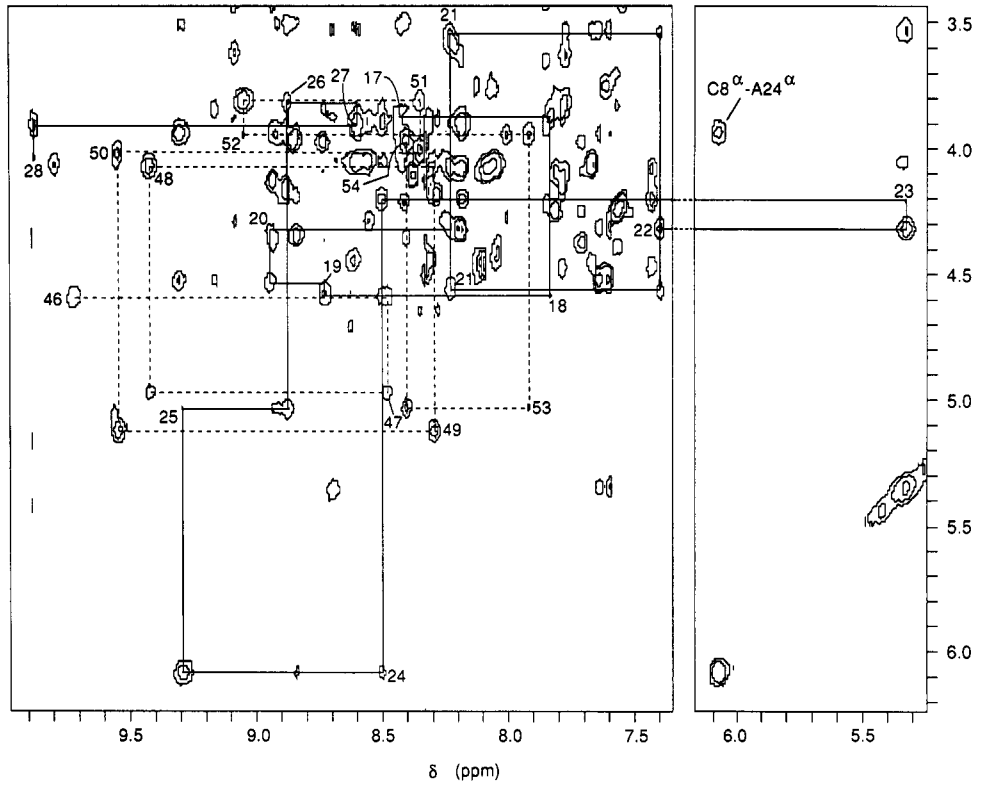


FIGURE 4: Fingerprint region of the 600-MHz NOESY spectrum of the RAR-DBD in $^1H_2O$ at 300 K, pH 6.7, with a 200-ms mixing time, showing the sequential assignments of two segments, G17–C28 (solid lines) and R46–K54 (dotted lines), by intraresidue $C^\alpha H$–NH and sequential $C^\alpha H_i$–$NH_{i+1}$ NOE connectivities. The interstrand $d_{\alpha\alpha}$ connectivity between the C8 and A24 residues is also indicated.

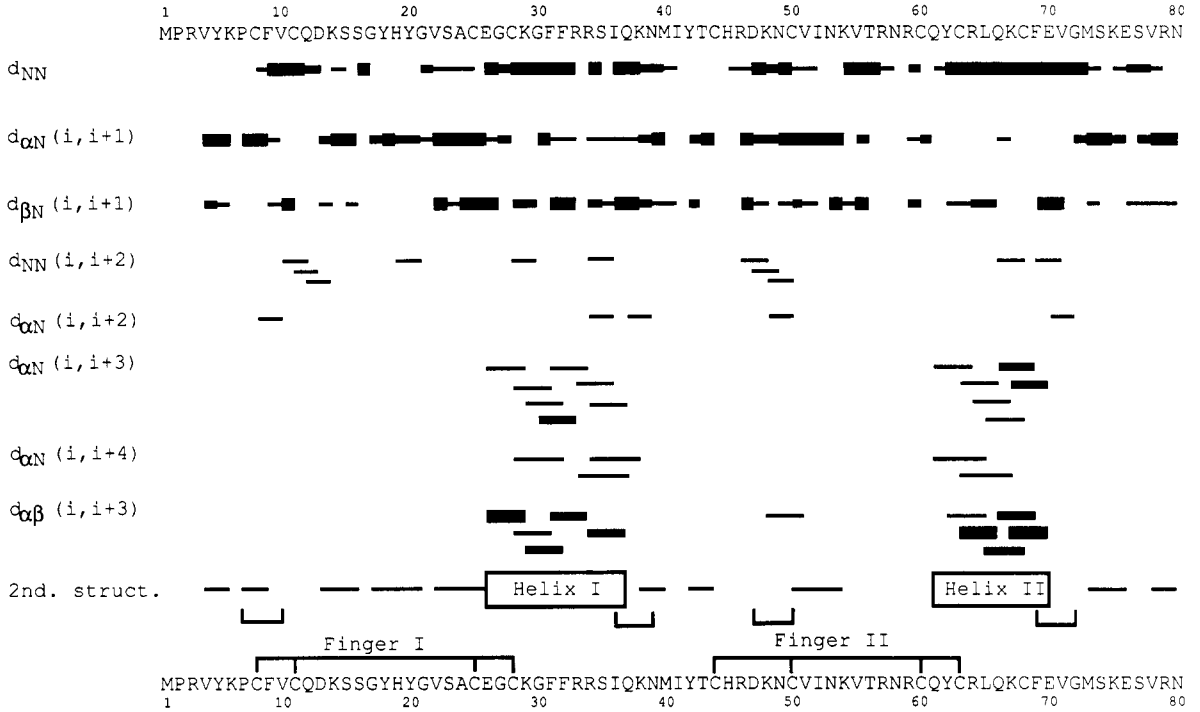


FIGURE 5: Summary of short- and medium-range NOEs and elements of secondary structure of the RAR-DBD. The intensities of NOEs are categorized as strong, medium, and weak, and represented by the thickness of the lines. The secondary structure elements indicated, α-helices (boxes), turns (brackets), and regions of extended peptide conformation (lines), were concluded on the basis of the sequential and medium-range NOE connectivities, as discussed in the text. The two putative zinc fingers are also indicated with lines connecting Zn-coordinating cysteine residues.

observed for P7–V10, I36–N39, D47–C50, and F69–G72 segments. Strong $d_{\alpha N}(2,3)$ and weak $d_{NN}(2,3)$ connectivities were found for P7–V10, indicating that this segment forms a type II turn. Several amide proton resonances were observed in a one-dimensional spectrum recorded after solvent exchange from $^1H_2O$ to $^2H_2O$ at 277 K and pH ~6.5 by dialysis with an Amicon flow cell, which took about 1 day (data not shown). One of these slowly exchanging resonances could be assigned unambiguously to V10 NH because of its unique resonance position. The slow exchange rate of V10 NH is consistent with the presence of a turn for the P7–V10 segment where the hydrogen bond between P7 CO and V10 NH is expected. The

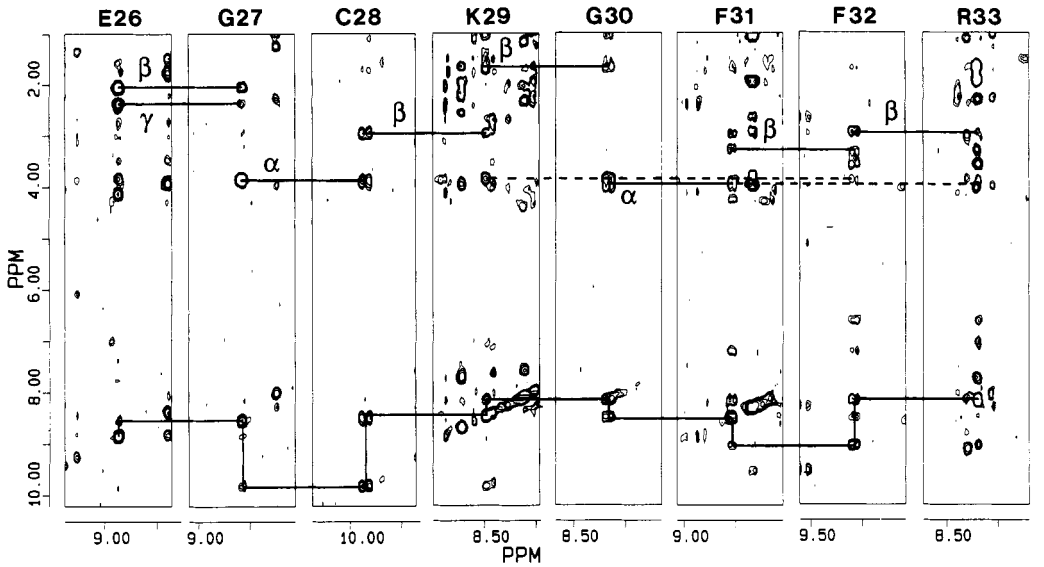


FIGURE 6: $\omega_1$–$\omega_3$ slices of the 500-MHz 3D NOESY–HMQC spectrum of the uniformly $^{15}N$-labeled RAR-DBD in $^1H_2O$ at 298 K, pH 6.7. Slices are taken at the $\omega_2$ axis position corresponding to the backbone amide nitrogen frequency of residues E26–R33: 119.7, 110.5, 130.2, 120.6, 105.3, 123.2, 126.9, and 117.4 ppm, respectively. For clarity, each slice contains only 1 ppm of the $\omega_3$ axis. The $d_{NN}$ connectivities are shown in the lower portion. Several $d_{\alpha N}$, $d_{\beta N}$, and $d_{\gamma N}$ connectivities are denoted by $\alpha$, $\beta$, and $\gamma$, respectively. $d_{\alpha N}(29,32)$ and $d_{\alpha N}(30,33)$ are indicated by dashed lines.

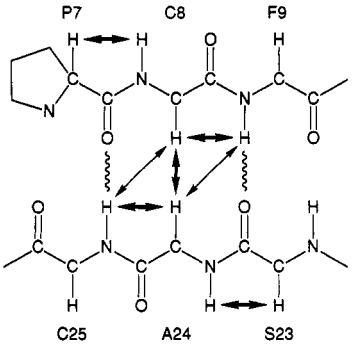


FIGURE 7: Strong sequential and long-range backbone NOEs indicating the presence of a short antiparallel $\beta$-sheet involving residues P7–F9 and S23–C25. Strong and medium NOEs are represented by thick and thin arrows, respectively. Wavy lines indicate possible interstrand hydrogen bonds (see text for details).

P7–F9 segment forms a short antiparallel $\beta$-sheet with the S23–C25 segment, as described above, indicating that the $\beta$-sheet is immediately followed by a turn. In the I36–N39 and F69–G72 segments, conversely, there are strong $d_{NN}(2,3)$ and weak or no $d_{\alpha N}(2,3)$ connectivities, suggesting that these segments form type I turns. It is noted that both helices are followed by type I turns.

*Extremely Shifted Resonances.* Several $^1H$ resonances are found at extremely shifted positions with respect to the ranges expected for proteins (Wüthrich, 1986; Gross & Kalbitzer, 1988). They are G17 $C^\alpha H$ (2.61 ppm), S23 NH (5.32 ppm), and A24 $C^\alpha H$ (6.09 ppm) (Figure 2). Due to the presence of S23 in a $\beta$-sheet (Figure 7), it seems likely that the high-field shift of the NH resonance could be explained by the ring current effect caused by the ring of the F9 residue. In fact, several long-range NOEs between S23 NH and aromatic proton resonances of the F9 residue are observed (data not shown).

In the case of the GR-DBD, a low-field shift of the $C^\alpha H$ resonance (6.25 ppm) is observed for the T456 residue (Härd et al., 1990b), the position of which corresponds to the A24 residue of the RAR-DBD (Figure 1). A similar short antiparallel $\beta$-sheet is also found in the GR-DBD. Therefore, the low-field shift may be related to generally observed tendency that $C^\alpha H$ in a $\beta$-sheet resonates at lower-field (Pardi et al., 1983; Szilagi & Jardetsky, 1989) in combination with the fact in the GR-DBD (and presumably in the RAR-DBD as well) this proton is within 5 Å from the metal ion of the first finger.

*Observation of the Hydroxyl Proton Resonance of T56.* Usually, the OH resonance of a Thr residue cannot be detected due to rapid exchange with water (Wüthrich, 1986). In the case of the RAR-DBD, however, the OH resonance of T56 is observed. The assignment is based on the scalar connectivity between T56 $C^\gamma H_3$ and $O^\gamma H$. The detection of the OH resonance suggests the presence of a hydrogen bond. Although identification of the acceptor of the hydrogen bond is impossible at this moment, the OH resonance gives NOEs to the proton resonances of I52 and N53 residues (data not shown), which must serve to determine the structure of the loop region between the C60 and C60 residues of the second finger. This is important because involvement of the loop region in the protein–DNA interaction is observed in the crystallographic complex of the GR-DBD with DNA (Luisi et al., 1991).

*Comparison of the Secondary Structure of the RAR-DBD with Those of the GR- and ER-DBDs and with TFIIIA-Type Zn-Finger Peptides.* The RAR-DBD has two $\alpha$-helices as do the GR- and ER-DBDs in solution (Härd et al., 1990a; Schwabe et al., 1990). The exact position and size of the helices differ slightly from those of the GR- and ER-DBDs (cf. Figure 1). Helix I of the RAR-DBD, as well as that of the ER-DBD, begins one residue earlier than helix I of the GR-DBD. Helix I of the RAR-DBD ends at the same position as that of the GR-DBD, although helix I of the ER-DBD ends two residues earlier. Helix II of all three DNA-binding domains begins at the same position, but ends at different positions, the RAR-DBD being shorter than the GR- and ER-DBDs by two and three residues, respectively. GR-DBD complexed with DNA in the crystal has another distorted $\alpha$-helix in the second finger in addition to the two $\alpha$-helices (Luisi et al., 1991). None of the three DNA-free DNA-binding domains of RAR, GR, and ER have a distorted $\alpha$-helix in this region. The short antiparallel $\beta$-sheet found in the RAR-DBD involving residues P7–F9 and S23–C25 corresponds excactly to that in the GR-DBD in solution involving residues L439–L441 and T456–C457, while a $\beta$-sheet is not

Table I: $^{15}N$ and $^{1}H$ Resonance Assignments of the Retinoic Acid Receptor DNA-Binding Domain at 300 K, pH 6.7[a]

| residue | $^{15}N$ | NH | $C^{\alpha}H$ | $C^{\beta}H$ | others |
|---|---|---|---|---|---|
| M1 | | | | | |
| P2 | | | | | |
| R3 | | | | | |
| V4 | 122.8 | 8.16 | 4.10 | 1.99 | 0.91 ($C^{\gamma}H_3$) |
| Y5 | 125.1 | 8.37 | 4.70 | 2.68, 3.06 | 7.04 ($C^{\delta}H$); 6.72 ($C^{\epsilon}H$) |
| K6 | 125.2 | 8.55 | 4.57 | 1.82, 2.10 | 1.69, 1.93 ($C^{\gamma}H$) |
| P7 | | | 4.36 | 2.14, 2.24 | 2.07 ($C^{\gamma}H$); 3.65, 3.93 ($C^{\delta}H$) |
| C8 | 121.6 | 8.83 | 3.93 | 2.66, 3.17 | |
| F9 | 130.6 | 9.30 | 4.53 | 3.15, 3.29 | 7.64 ($C^{\delta}H$); 7.59 ($C^{\epsilon}H$, $C^{\zeta}H$) |
| V10 | 121.3 | 9.15 | 3.85 | 3.02 | 1.11 ($C^{\gamma 1}H_3$); 0.89 ($C^{\gamma 2}H_3$) |
| C11 | 117.4 | 8.30 | 4.95 | 2.00, 2.75 | |
| Q12 | 114.7 | 8.17 | 4.20 | 2.50, 2.57 | 2.38 ($C^{\gamma}H$); 6.57, 7.32 ($N^{\epsilon}H_2$); 113.5 ($N^{\epsilon}$) |
| D13 | 121.7 | 8.88 | 4.79 | 2.68, 3.09 | |
| K14 | 121.8 | 8.85 | 3.97 | 1.78 | 1.53 ($C^{\gamma}H$) |
| S15 | 119.3 | 8.37 | 3.92 | 3.51, 3.84 | |
| S16 | 120.1 | 8.93 | 4.87 | 3.97, 4.28 | |
| G17 | 114.7 | 8.43 | 2.61, 3.88 | | |
| Y18 | 119.5 | 7.80 | 4.59 | 2.52, 2.62 | 6.80 ($C^{\delta}H$); 6.82 ($C^{\epsilon}H$) |
| H19 | 124.2 | 8.69 | 4.54 | 2.29, 2.45 | 6.80 ($C^{\delta 2}H$); 7.80 ($C^{\epsilon 1}H$) |
| Y20 | 119.7 | 8.90 | 4.32 | 2.98, 3.38 | 7.02 ($C^{\delta}H$); 7.09 ($C^{\epsilon}H$) |
| G21 | 102.6 | 8.21 | 3.54, 4.55 | | |
| V22 | 113.5 | 7.40 | 4.31 | 1.61 | 0.76 ($C^{\gamma 1}H_3$); 0.66 ($C^{\gamma 2}H_3$) |
| S23 | 115.4 | 5.32 | 4.20 | 3.13, 3.53 | |
| A24 | 129.9 | 8.49 | 6.09 | 1.34 | |
| C25 | 119.0 | 9.29 | 5.03 | 2.82, 3.51 | |
| E26 | 119.7 | 8.86 | 3.82 | 2.06 | 2.37, 2.44 ($C^{\gamma}H$) |
| G27 | 110.5 | 8.58 | 3.90 | | |
| C28 | 130.2 | 9.85 | 4.03 | 2.93, 3.00 | |
| K29 | 120.6 | 8.49 | 3.89 | 1.54, 1.69 | 1.76 ($C^{\gamma}H$) |
| G30 | 105.3 | 8.17 | 3.91 | | |
| F31 | 123.2 | 8.54 | 4.29 | 3.04, 3.28 | 7.25 ($C^{\delta}H$); 7.03 ($C^{\epsilon}H$); 7.57 ($C^{\zeta}H$) |
| F32 | 126.9 | 9.08 | 3.62 | 2.93, 3.42 | 6.65 ($C^{\delta}H$); 6.49 ($C^{\epsilon}H$, $C^{\zeta}H$) |
| R33 | 117.4 | 8.19 | 3.58 | 2.33 | 1.96 ($C^{\gamma}H$) |
| R34 | 116.8 | 8.22 | 4.07 | 1.75, 1.89 | 1.65 ($C^{\gamma}H$); 3.29 ($C^{\delta}H$) |
| S35 | 114.5 | 7.76 | 3.78 | 3.53, 3.65 | |
| I36 | 120.5 | 7.77 | 3.84 | 1.68 | 0.88 ($C^{\gamma}H$); 0.74 ($C^{\gamma}H_3$); 0.40 ($C^{\delta}H_3$) |
| Q37 | 120.5 | 8.71 | 3.97 | 1.97, 2.15 | 2.30, 2.58 ($C^{\gamma}H$); 6.68, 7.41 ($N^{\epsilon}H_2$); 111.2 ($N^{\epsilon}$) |
| K38 | 113.4 | 7.69 | 4.37 | 1.68, 2.01 | 1.44, 1.56 ($C^{\gamma}H$) |
| N39 | 118.3 | 7.56 | 4.25 | 2.62, 2.99 | 6.77, 7.49 ($N^{\delta}H_2$); 113.1 ($N^{\delta}$) |
| M40 | 114.9 | 7.80 | 4.09 | 1.22, 1.38 | 2.32, 2.56 ($C^{\gamma}H$); 2.12 ($C^{\epsilon}H_3$) |
| I41 | 122.4 | 8.05 | 4.08 | 1.78 | 1.13 ($C^{\gamma}H$); 0.91 ($C^{\gamma}H_3$) |
| Y42 | 127.2 | 8.57 | 5.36 | 2.77, 3.53 | 7.16 ($C^{\delta}H$); 6.71 ($C^{\epsilon}H$) |
| T43 | 114.8 | 8.66 | 4.46 | 3.87 | 0.98 ($C^{\gamma}H_3$) |
| C44 | 129.2 | 8.58 | 4.64 | 2.42, 2.69 | |
| H45 | | 9.63[b] | 4.61 | 3.12, 3.46 | 7.27 ($C^{\delta 2}H$); 8.25 ($C^{\epsilon 1}H$) |
| R46 | 128.5 | 9.70 | 4.59 | 1.55, 1.64 | 1.96 ($C^{\gamma}H$); 3.35 ($C^{\delta}H$) |
| D47 | 124.5 | 8.46 | 4.97 | 2.53, 2.82 | |
| K48 | 120.0 | 9.41 | 4.07 | 2.19, 2.24 | 0.97, 1.31 ($C^{\gamma}H$); 1.63 ($C^{\delta}H$) |
| N49 | 118.7 | 8.28 | 5.11 | 2.65, 2.75 | 6.84, 7.42 ($N^{\delta}H_2$); 112.1 ($N^{\delta}$) |
| C50 | 126.2 | 9.55 | 4.01 | 2.65, 2.94 | |
| V51 | 123.4 | 8.32 | 3.82 | 1.97 | 1.06 ($C^{\gamma 1}H_3$); 0.89 ($C^{\gamma 2}H_3$) |
| I52 | 131.6 | 9.04 | 3.93 | 2.10 | 0.96, 1.62 ($C^{\gamma}H$); 0.58 ($C^{\gamma}H_3$); 0.50 ($C^{\delta}H_3$) |
| N53 | 125.3 | 7.91 | 5.03 | 2.78, 3.31 | 6.97, 7.57 ($N^{\delta}H_2$); 113.8 ($N^{\delta}$) |
| K54 | 118.7 | 8.39 | 3.94 | 1.79, 1.91 | 1.67 ($C^{\gamma}H$) |
| V55 | 116.5 | 8.05 | 4.04 | 2.29 | 1.02 ($C^{\gamma 1}H_3$); 0.94 ($C^{\gamma 2}H_3$) |
| T56 | 111.2 | 8.25 | 4.27 | 4.65 | 1.25 ($C^{\gamma}H_3$); 5.52 ($O^{\gamma}H$) |
| R57 | 120.0 | 8.00 | 4.27 | | |
| N58 | 115.6 | 8.32 | 4.66 | 2.66, 2.86 | 7.12, 7.46 ($N^{\delta}H_2$); 113.4 ($N^{\delta}$) |
| R59 | 119.1 | 8.27 | 4.18 | 2.00 | 1.75 ($C^{\gamma}H$) |
| C60 | 119.3 | 7.27 | 4.66 | 2.78, 2.93 | |
| Q61 | 124.2 | 8.65 | 3.72 | 0.52, 1.52 | 2.21, 2.37 ($C^{\gamma}H$); 6.86, 7.49 ($N^{\epsilon}H_2$); 113.1 ($N^{\epsilon}$) |
| Y62 | 120.2 | 8.40 | 4.07 | 2.69, 2.89 | 7.22 ($C^{\delta}H$); 6.95 ($C^{\epsilon}H$) |
| C63 | 124.3 | 9.79 | 4.06 | 2.69, 3.08 | |
| R64 | 123.0 | 8.04 | 3.77 | 1.51 | |
| L65 | 121.5 | 8.59 | 4.42 | 1.82, 2.42 | 1.67 ($C^{\gamma}H$); 0.83, 1.14 ($C^{\delta}H_3$) |
| Q66 | 116.6 | 8.40 | 4.21 | 2.01, 2.24 | 2.14, 2.35 ($C^{\gamma}H$); 7.10, 7.40 ($N^{\epsilon}H_2$); 113.2 ($N^{\epsilon}$) |
| K67 | 121.3 | 8.03 | 4.38 | 2.26 | |
| C68 | 118.0 | 8.17 | 3.91 | 3.08 | |
| F69 | 116.4 | 7.84 | 4.79 | 3.23, 3.42 | 7.54 ($C^{\delta}H$); 7.42 ($C^{\epsilon}H$); 7.29 ($C^{\zeta}H$) |
| E70 | 124.3 | 8.93 | 4.13 | 2.40 | 2.32, 2.50 ($C^{\gamma}H$) |
| V71 | 109.0 | 8.31 | 4.52 | 2.76 | 1.17 ($C^{\gamma}H_3$) |
| G72 | 106.0 | 7.60 | 3.76, 4.52 | | |
| M73 | 120.8 | 8.11 | 4.44 | 0.58, 0.76 | 2.02, 2.33 ($C^{\gamma}H$); 1.74 ($C^{\epsilon}H_3$) |
| S74 | 116.0 | 8.30 | 5.04 | 3.89, 4.09 | |
| K75 | 128.8 | 9.54 | 4.21 | 1.94, 2.04 | 1.59 ($C^{\gamma}H$) |
| E76 | 119.7 | 8.86 | 4.14 | 2.06, 2.12 | 2.37, 2.45 ($C^{\gamma}H$) |
| S77 | 114.0 | 7.76 | 4.47 | 4.14 | |
| V78 | 121.6 | 7.42 | 4.09 | 2.12 | 0.58 ($C^{\gamma 1}H_3$); 0.89 ($C^{\gamma 2}H_3$) |
| R79 | 125.4 | 8.02 | 4.46 | 1.85, 2.00 | 1.71 ($C^{\gamma}H$) |
| N80 | 125.1 | 8.08 | 4.50 | 2.74, 2.85 | 6.87, 7.55 ($N^{\delta}H_2$); 113.0 ($N^{\delta}$) |

[a] Chemical shfits in ppm, with $NH_4Cl$ and water resonances being set at 22.3 and 4.79 ppm, respectively, for $^{15}N$ and $^{1}H$. [b] Chemical shifts refer to 288 K, pH 6.2.

reported for the ER-DBD. The other regions of extended peptide conformation found in the RAR-DBD generally correspond well to those found in the GR-DBD in solution. The two turns for P7–V10 and D47–C50 segments in the RAR-DBD correspond exactly to those for L439–V442 and R479–C482 segments in the GR-DBD in solution as well.

Observation of many long-range NOEs between the two helices, particularly concerning F31, F32, I36, Y62, L65, and F69 residues (data not shown), suggests that the two helices pack against each other, forming a hydrophobic core, as found in the GR- and ER-DBDs.

The secondary structure of the RAR-DBD is clearly different from those of TFIIIA-type Zn-finger peptides like peptides from *Xfin* (Lee et al., 1989), ADR1 (Parraga et al., 1988), *mKr2* (Carr et al., 1990), SWI5 (Neuhaus et al., 1990), and Zif268 (Pavletich & Pabo, 1991). Whereas in TFIIIA-type fingers and $\alpha$-helix and a $\beta$-sheet are integral parts of the finger structure, in the nuclear receptors $\alpha$-helices occur mainly in the peptide region following the Zn finger.

In general, the secondary structure of the RAR-DBD is similar to those of the GR- and ER-DBDs. Some differences have been discussed above. Determination of the three-dimensional structure of the RAR-DBD is now in progress, and will undoubtedly increase our understanding of sequence-specific DNA binding, by allowing further structural comparisons among members of this superfamily.

## Acknowledgments

The 600-MHz $^{1}H$ NMR spectra were recorded at the National Dutch HF-NMR facility in Nijmegen with the assistance of Dr. S. Wijminga and Mr. J. Joordens.

## References

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Am. Chem. Soc. 105*, 7188–7190.

Benbrook, D., Lernhardt, E., & Pfahl, M. (1988) *Nature 333*, 669–672.
Boelens, R., Scheek, R. M., Dijkstra, K., & Kaptein, R. (1985) *J. Magn. Reson. 62*, 378–386.
Brand, N. J., Petkovich, M., Krust, A., Chambon, P., de The, H., Marchio, A., Tiollais, P., & Dejean, A. (1988) *Nature 332*, 850–853.
Brockes, J. P. (1989) *Neuron 2*, 1285–1294.
Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 77*, 166–169.
Carr, M. D., Pastore, A., Gausepohl, H., Frank, R., & Roesch, P. (1990) *Eur. J. Biochem. 188*, 455–461.
Chazin, W. J., & Wright, P. E. (1988) *J. Mol. Biol. 202*, 623–636.
Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2821–2822.
de Thé, H., Marchio, A., Tiollais, P., & Dejean, A. (1987) *Nature 330*, 667–670.
Durston, A. J., Timmermans, J. P. M., Hage, W. J., Hendriks, H. F., de Vries, N. J., Heideveld, M., & Nieuwkoop, P. D. (1989) *Nature 340*, 140–144.
Englander, S. W., & Wand, A. J. (1987) *Biochemistry 26*, 5953–5958.
Giguere, V., Ong, E. S., Sequi, P., & Evans, R. M. (1987) *Nature 330*, 624–629.
Giguere, V., Ong, E. S., Evans, R. M., & Tabin, C. J. (1989) *Nature 337*, 566–569.
Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870–7872.
Gross, K. H., & Kalbitzer, H. R. (1988) *J. Magn. Reson. 76*, 87–99.
Härd, T., Kellenbach, E., Boelens, R., Maler, B. A., Dahlman, K., Freedman, L. P., Carlstedt-Duke, J., Yamamoto, K. R., Gustafsson, J. A., & Kaptein, R. (1990a) *Science 249*, 157–160.
Härd, T., Kellenbach, E., Boelens, Kaptein, R., Dahlman, K., Carlstedt-Duke, J., Freedman, L. P., Maler, B. A., Hyde, E. I., Gustafsson, J. A., & Yamamoto, K. R. (1990b) *Biochemistry 29*, 9015–9023.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Krust, A., Kastner, P., Petkovich, M., Zelent, A., & Chambon, P. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5310–5314.
Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., & Wright, P. E. (1989) *Science 245*, 635–637.
Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R., & Sigler, P. B. (1991) *Nature 352*, 497–505.
Maden, M., Ong, D. E., Summerbell, D., & Chytil, F. (1988) *Nature 335*, 733–735.
Mangelsdorf, D. J., Ong, E. S., Dyck, J. A., & Evans, R. M. (1990) *Nature 345*, 224–229.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A., & Clore, G. M. (1989) *Biochemistry 28*, 6150–6156.
Martinez, E., & Wahli, W. (1991) in *Nuclear Hormone Receptors* (Parker, M. G., Ed.) pp 125–153, Academic Press, London.
Neuhaus, D., Nakaseko, Y., Nagai, K., & Klug, A. (1990) *FEBS Lett. 262*, 179–184.
Pardi, A., Wagner, G., & Wüthrich, K. (1983) *Eur. J. Biochem. 137*, 445–454.
Parraga, G., Horvath, S. J., Eisen, A., Taylor, W. E., Hood, L., Young, E. T., & Klevit, R. E. (1988) *Science 241*, 1489–1492.
Pavletich, N. P., & Pabo, C. O. (1991) *Science 252*, 809–817.
Petkovich, M., Brand, N. J., Krust, A., & Chambon, P. (1987) *Nature 330*, 444–450.
Ragsdale, C. W., Petkovich, M., Gates, P. B., Chambon, P., & Brockes, J. P. (1989) *Nature 341*, 654–657.
Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.
Schwabe, J. W. R., Neuhaus, D., & Rhodes, D. (1990) *Nature 348*, 458–461.
Shaka, A. J., Barker, P. B., & Freeman, R. J. (1985) *J. Magn. Reson. 64*, 547–552.
Szilagyi, L., & Jardetzky, O. (1989) *J. Magn. Reson 83*, 441–449.
Thaller, C., & Eichele, G. (1987) *Nature 327*, 625–628.
Wüthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Zelent, A., Krust, A., Petkovich, M., Kastner, P., & Chambon, P. (1989) *Nature 339*, 714–717.
Zuiderweg, E. R. P., & Fesik, S. W. (1989) *Biochemistry 28*, 2387–2391.
Zuiderweg, E. R. P., Boelens, R., & Kaptein, R. (1985) *Biopolymers 24*, 601–611.