^{1}H NMR Studies of the Glucocorticoid Receptor DNA-Binding Domain: Sequential Assignments and Identification of Secondary Structure Elements†

T. Härd,‡,§ E. Kellenbach,‡ R. Boelens,‡ R. Kaptein,*,‡ K. Dahlman,∥ J. Carlstedt-Duke,∥ L. P. Freedman,⊥,# B. A. Maler,⊥ E. I. Hyde,○ J.-Å. Gustafsson,∥ and K. R. Yamamoto⊥

Department of Chemistry, University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands, Department of Medical Nutrition and Center for Biotechnology, Karolinska Institute, Huddinge University Hospital, S-141 86 Huddinge, Sweden, Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143-0448, and School of Biochemistry, University of Birmingham, P.O. Box 363, Birmingham B15 2TT, England

Received April 13, 1990; Revised Manuscript Received June 7, 1990

ABSTRACT: Two protein fragments containing the DNA-binding domain (DBD) of the glucocorticoid receptor (GR) have been studied by two-dimensional ^{1}H NMR spectroscopy. The two peptides (93 and 115 residues, respectively) contain a common segment corresponding to residues C440–I519 of the rat GR or residues C421–I500 of the human GR and include two Zn-binding "finger" domains. The structures of this segment are almost identical in the two protein fragments, as judged from chemcial shifts and sequential NOE connectivities. More than 90% of all observable ^{1}H resonances within a 71-residue segment encompassing C440–R510 (rat GR) could be sequentially assigned by standard techniques, and stereospecific assignments could be made for the methyl groups in four valine residues within this segment. Sequential NOE connectivities indicate several elements of secondary structure including two α-helical segments consisting of residues S459–E469 and P493–G504, a type I reverse turn between residues R479 and C482, a type II reverse turn between residues L475 and G478, and several regions of extended peptide conformation. No evidence for α-helical conformation was found within the two putative zinc-finger domains, indicating that the structures of these domains differ from that of TFIIIA-type zinc fingers. The observation of some very slowly exchanging amide protons in the N-terminal (CI) domain of the DBD in combination with slow rotation of the Y452 aromatic ring indicates that this domain has a restricted conformational flexibility compared to the C-terminal (CII) domain. We also observe several long-range NOE connectivities within C440–R510, suggesting that the sequential assigments presented here will provide a basis for a complete structure determination of this segment of the GR.

The DNA-binding activity of steroid/thyroid hormone receptors is localized to distinct domains (Carlstedt-Duke et al., 1987; Rusconi & Yamamoto, 1987) with a high degree of sequence homology within the hormone receptor superfamily (Evans, 1988; Beato, 1989). Protein fragments containing the DNA-binding domain (DBD)[1] of the glucocorticoid receptor (GR) expressed in *Escherichia coli* exhibit sequence-specific binding to GR response elements (GREs) (Freedman et al., 1988a; Tsai et al., 1988). These protein fragments contain two Zn atoms that are required for proper folding and DNA binding activity (Freedman et al., 1988a). The presence of Zn-binding domains is reminiscent of the "zinc-finger" motif found in the *Xenopus* transcription factor IIIA (TFIIIA) (Miller et al., 1985), as well as similar domains found in retroviral nucleic acid binding proteins (Green & Berg, 1989). However, the two Zn-coordinating regions of the steroid receptors (sometimes termed CI and CII, as discussed below) are not homologous to these fingers, indicating that steroid/thyroid receptors possess a unique structural motif for DNA binding (Berg, 1989).

Spectroscopic and mutagenesis studies have revealed some details about the structure and function of various regions of the hormone receptor DBDs. These domains contain nine conserved Cys residues, and spectroscopic (EXAFS) data indicate that eight of these are involved in tetrahedral coordination of Zn^{2+} (Freedman et al., 1988a). An intron occurs in the middle of the DNA sequence coding the DBD, suggesting that it consists of two subdomains (CI and CII) (Ponglikitmongkol et al., 1988). Studies of mutant estrogen receptors, where either CI or CII was replaced by the corresponding segment from the GR, indicate that the CI subdomain specifies the binding site recognition (Green et al., 1988). The CI and CII subdomains are believed to coordinate one Zn atom each as depicted in Figure 1 (Severne et al., 1988),

† This research was supported by the Netherlands Organization for Chemical Research (SON) and the Netherlands Organization of Scientific Research (NWO). The 600-MHz ^{1}H NMR spectra were recorded at the National Dutch HF-NMR Facility in Nijmegen with the assistance of Dr. S. Wijmenga and Mr. J. Joordens. Additional financial support was obtained from the Swedish Medical Research Council (MFR) (Grant 2819), the Swedish National Board for Technical Development (to J.C.-D.), and the U.S. National Science Foundation (to K.R.Y.). T.H. acknowledges a research fellowship from the Swedish National Research Council, a short-term fellowship from the European Molecular Biology Organization, and financial support from NWO. J.C.-D. acknowledges a research fellowship from MFR, and K.D. acknowledges a predoctoral fellowship from MFR. E.I.H. acknowledges the Wellcome Trust for financial support.

‡ University of Utrecht.

§ On leave from the Department of Medical Biophysics, Karolinska Institute, S-104 01 Stockholm, Sweden.

∥ Huddinge University Hospital.

⊥ University of California at San Francisco.

Present address: Cell Biology and Genetics Program, Memorial Sloan-Kettering Center, New York, NY 10023.

○ University of Birmingham.

[1] Abbreviations: GR, glucocorticoid receptor; DBD, DNA-binding domain; GRE, glucocorticoid response element; 2D, two dimensional; DTE, dithioerythritol; NOE, nuclear Overhauser enhancement; COSY, correlated spectroscopy; DQF, double quantum filtered; HOHAHA, homonuclear Hartmann–Hahn; TFIIIA, transcription factor IIIA.

0006-2960/90/0429-9015$02.50/0

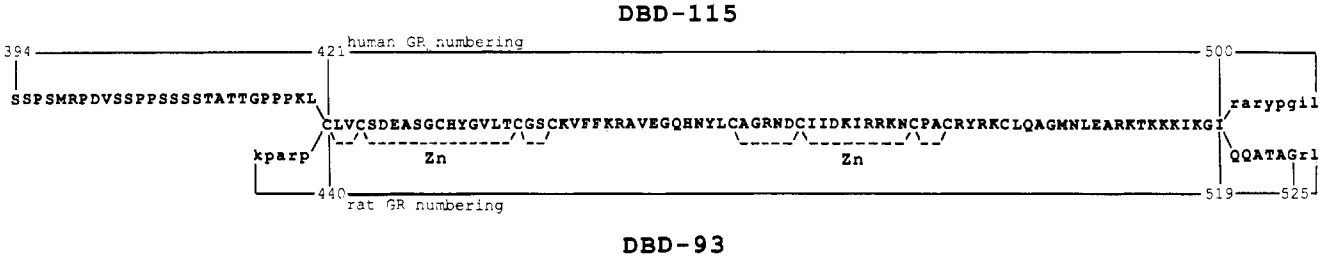

FIGURE 1: Sequences of the two studied protein fragments showing the common 80-residue segment as well as "linker" peptide segments that do not belong to the native GR (lower-case letters). The numbering of DBD-115 and DBD-93 corresponds to the human and rat GR, respectively. The Zn-coordination scheme proposed by Severne et al. (1988) is also indicated.

although an alternative zinc coordination pattern for the CII subdomain has also been noted (Freedman et al., 1988a; Schena et al., 1989). The two residues located between the two cysteines at the C-terminal part of the CI subdomain are essential for discrimination between glucocorticoid and estrogen response elements, consistent with the possibility that this region might directly contact the DNA (Danielsen et al., 1989; Mader et al., 1989; Umesuno & Evans, 1989). It has also been suggested on theoretical grounds that the ten residues immediately following the last cysteine of the CI subdomain might form an amphipatic α-helix, perhaps functioning as a "recognition" helix by analogy with proteins that contain the helix–turn–helix motif (Berg, 1989).

We are pursuing NMR studies of two protein fragments containing the GR DBD with the objective of determining the three-dimensional structure in solution. These proteins contain a common segment of 80 residues, including the CI and CII subdomains. In this paper we report on sequential ^1H resonance assignments and secondary structure in a 71-residue segment comprising amino acids C440–R510 of the rat GR. Sequential NOE connectivities indicate several elements of secondary structure within this segment including two helices of 11 and 12 residues, respectively, two reverse turns, and several regions of extended peptide conformation. The structures of the two putative Zn-binding regions are different from that of TFIIIA-type zinc fingers (Pårraga et al., 1988; Lee et al., 1989a,b) and the first zinc-coordinating domain of the *gag* protein p55 of HIV (Summers et al., 1990), which were recently studied by 2D NMR spectroscopy. We note that the CI and CII subdomains of the DBD have a similar "arrangement" of secondary structure elements, including an α-helix immediately following the Zn-coordinating regions.

Materials and Methods

The two protein fragments containing the GR DBD (Figure 1) were KPARP(C440–G525)RL (DBD-93, numbering referring to the rat GR), which was purified and checked for sequence-specific DNA binding activity as described by Freedman et al. (1988), and (S394–I500)RARYPGIL (DBD-115, numbering referring to the human GR), which was purified as described and checked for sequence-specific DNA binding activity by Dahlman et al. (1989). These two proteins contain an 80 amino acid segment with complete sequence identity (C440–I519 of the rat GR or C421–I500 of the human GR) as shown in Figure 1. (All residue numbering below refers to the rat GR.) NMR samples (typically 2–4 mM) in 95/5 $^1H_2O/^2H_2O$ mixtures or 99.9% ^2H_2O were prepared by dialysis and concentration with an Amicon flow cell against a phosphate buffer containing 300 mM NaCl, 0.1 mM NaN_3, and 1–20 mM phosphates at pH 5.3–6.7, with 1 mM DTE to prevent oxidation of cysteines.

^1H NMR spectra were recorded at 294–310 K on Bruker AM500 and AM600 spectrometers equipped with Aspect 3000 computers. Phase-sensitive two-dimensional NOE (NOESY), HOHAHA, and double quantum filtered COSY (DQF-COSY) spectra were recorded according to the time-proportional phase incrementation method (Marion & Wüthrich, 1983). NOESY spectra (Jeener et al., 1979; States et al., 1982) were recorded with a 200-ms mixing time. 2D HOHAHA spectra were recorded with 40–80-ms spin-locking mixing times with either MLEV-17 (Bax & Davis, 1985) or "clean TOCSY" (Griesinger et al., 1988) pulse sequences. A DQF-COSY spectrum was recorded with the pulse sequence and phase cycling devised by Rance et al. (1983). The carrier frequency was placed at the water resonance except in some of the HOHAHA experiments where it was placed at about 6.8 ppm to improve magnetization transfer between amide and $C^\alpha H$ protons. The water resonance was irradiated during the relaxation period, in some cases followed by a "SCUBA" pulse sequence preceding the preparation pulse (Brown et al., 1988) to prevent saturation of $C^\alpha H$ protons. 2D spectra were recorded with 400–700 t_1 increments, 64–256 free induction decays of 2K data points per increment being collected (4K in the DQF-COSY experiment).

All data processing was carried out on a microVAX II cluster using the TRITON NMR software package developed at the Department of Chemistry, University of Utrecht. Sine-bell window functions shifted by $\pi/4$ (NOESY and DQF-COSY) or $\pi/6$ (HOHAHA) were applied in the ω_1 and ω_2 domains, and zero filling in ω_1 yielded spectra with a digital resolution of 5–6 Hz in each dimension for NOESY and HOHAHA spectra and 2.7 Hz in the DQF-COSY spectrum. Fourth-order polynomial baseline corrections were applied in both frequency domains (Boelens et al., 1985).

Results and Discussion

Experimental Conditions. The two protein fragments proved to be stable only within rather limited pH and temperature intervals (pH >5.3 and T <40 °C). Nevertheless, it was possible to obtain good NMR spectra in ^1H_2O (Figure 2), although some NH resonances are not observed due to rapid exchange with water. The protein fragments 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 on a longer time scale (hours), which could confirm the identification of secondary structure elements. A few very slowly exchanging amide NH resonances could still be observed following a change from ^1H_2O to ^2H_2O buffers that could be carried out within about 48 h at 277 K on an Amicon filter apparatus, and these results have been included.

Sequential Assignments. A comparison of the HOHAHA fingerprint regions of DBD-93 and DBD-115 reveals a similar pattern of $C^\alpha H$–NH cross peaks in the two spectra (Figure 3), confirming the notion that the two proteins contain common structural elements. The similar but nonidentical distribution

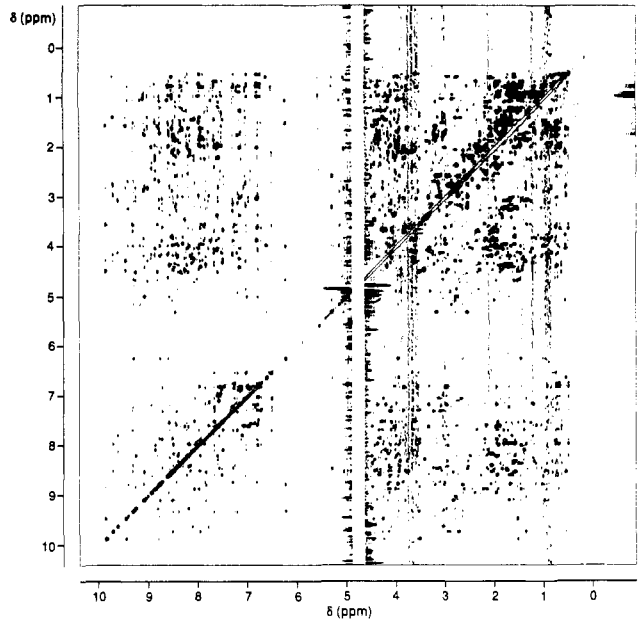

FIGURE 2: 600-MHz NOESY spectrum of DBD-115 in ^1H_2O recorded at 300 K, pH 6.7, with a mixing time of 200 ms.

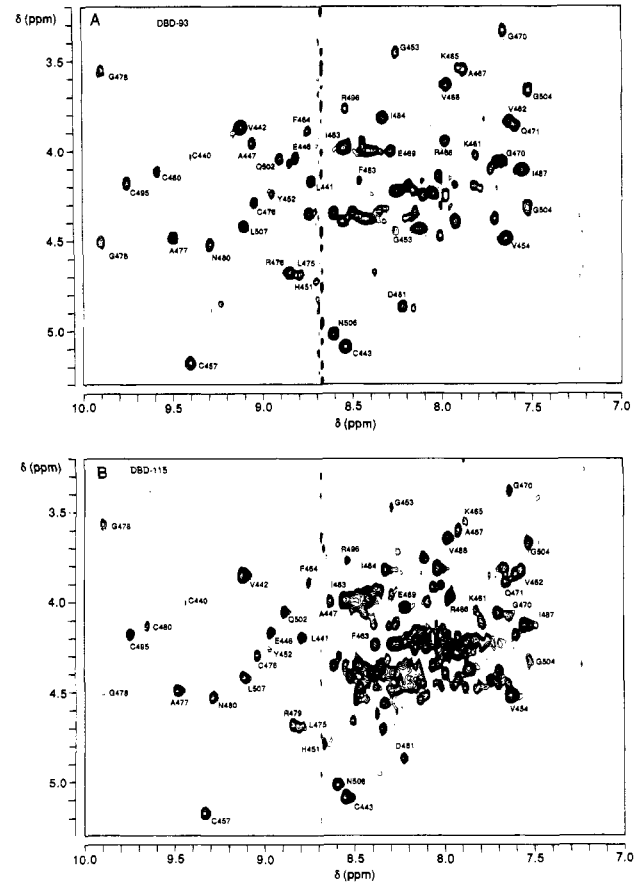

FIGURE 3: Fingerprint regions of 500-MHz HOHAHA spectra of DBD-93 (A) and DBD-115 (B). $C^{\alpha}H$–NH connectivities of some residues belonging to the common 80-amino acid segment are indicated to facilitate comparison. Both spectra were recorded in ^1H_2O at 300 K, pH 6.7, with a 40-ms spin-lock mixing time.

of chemical shifts in the two proteins proved to be extremely useful in the assignment process, because resonances that overlap in one protein are in many cases resolved in the other. Resonances that overlap in both proteins are in some cases resolved in spectra recorded at different pH and temperature conditions. Sequential resonance assignments were carried

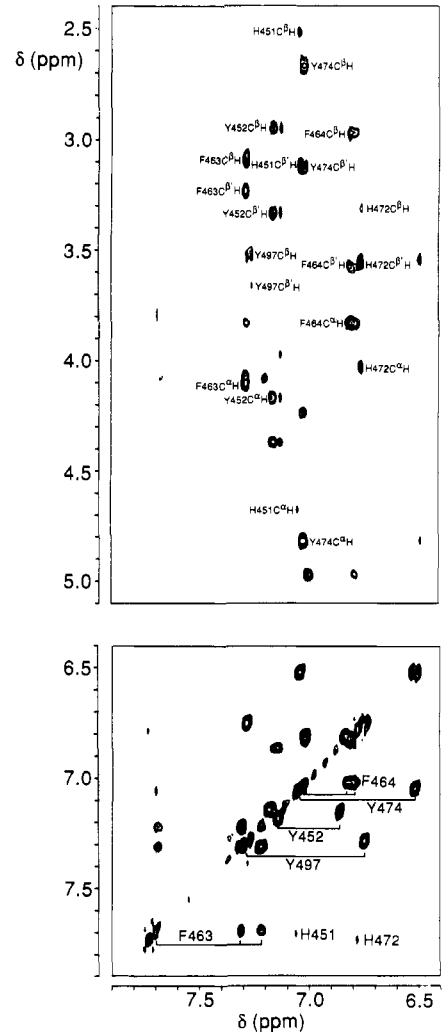

FIGURE 4: Regions of 500-MHz NOESY (top) and HOHAHA (bottom) spectra of DBD-93 in ^2H_2O demonstrating the identification of aromatic residue spin systems (two Phe, three Tyr, and two His) by HOHAHA and ring–$C^{\alpha}H$,$C^{\beta}H$ NOE connectivities. Both spectra were recorded at 300 K, pH 6.7. A mixing time of 200 ms was used for the NOESY spectrum, and a 40-ms spin-lock mixing time was used for the HOHAHA spectrum.

out via a simultaneous identification of spin systems in HOHAHA and DQF-COSY spectra and (strong) sequential connectivities in NOESY spectra, i.e., through a combination of the assignment philosophy initially described by Wüthrich and co-workers (Billeter et al., 1982; Wagner & Wüthrich, 1982; Wüthrich, 1986) and the "main-chain-directed" strategy described by Englander and Wand (1987).

Initially, all four Val, six Leu, and five Ile spin systems, as well as four Gly, seven Ala, and two Thr spin systems, were identified in HOHAHA and DQF-COSY spectra of DBD-93. The aromatic spin systems in DBD-93 (two Phe, three Tyr, and two His) were identified by a comparison of DQF-COSY and HOHAHA spectra, and corresponding $C^{\alpha}H$ and $C^{\beta}H$ resonances were identified from strong NOE connectivities as shown in Figure 4. The four Asp and one of the two Gln spin sytems within the common (C440–I519) sequence were identified from strong NOE connectivities between $C^{\beta}H$ or $C^{\gamma}H$ protons and NH_2 resonances at 6.7–7.8 ppm, respectively. Several of the remaining spin systems could also be classified as being of the AMX type (belonging to remaining Ser, Cys, and Asp residues).

The essential steps of the sequential assignment procedure can be summarized by the following arguments, which, together with Table I and Figures 2–6, should serve as a suf-

Table I: ^{1}H Resonance Assignments of DBD-93 at 300 K, pH 6.7[a]

	residue	NH	$C^{\alpha}H$	$C^{\beta}H$	others
1	K		4.67	1.78, 1.86	
2	P		4.43	1.80, 2.28	2.03 ($C^{\gamma}H$); 3.69, 3.89 ($C^{\delta}H$)
3	A	8.70	4.33	1.43	
4	R	8.67	4.70	1.67, 1.80	
5	P		4.46	1.76, 2.12	3.51, 3.61 ($C^{\delta}H$)
6	C440	9.34	4.02	2.73, 3.21	
7	L441	8.67	4.13	1.68, 1.71	2.05 ($C^{\gamma}H$); 1.04, 1.16 ($C^{\delta}H$)
8	V442	9.06	3.83	2.85	0.89 ($C^{\gamma^1}H$); 1.16 ($C^{\gamma^2}H$)
9	C443	8.48	5.05	1.96, 2.80	
10	S444	7.88	4.35	4.15, 4.40	
11	D445	8.91	4.73	2.63, 2.92	
12	E446	8.77	3.99	1.90, 2.30	2.17 ($C^{\gamma}H$)
13	A447	9.00	3.92	1.18	
14	S448	7.64[b]	3.80	3.57, 3.66	
15	G449				
16	C450				
17	H451	8.56[b]	4.68	2.52, 3.09	
18	Y452	8.91	4.19	2.98, 3.36	7.09, 7.12 (2,6);[c] 6.77 (3,5)
19	G453	8.21	3.39, 4.40		
20	V454	7.59	4.45	1.88	0.63 ($C^{\gamma^2}H$); 0.94 ($C^{\gamma^1}H$)
21	L455	8.34	4.62	1.38, 1.45	1.39 ($C^{\gamma}H$); 0.67, 0.73 ($C^{\delta}H$)
22	T456	8.62	6.25	4.64	1.47 ($C^{\gamma}H$)
23	C457	9.36	5.14	2.79, 3.50	
24	G458	9.12	3.93, 4.23		
25	S459	8.09	4.28	4.12 ($C^{\beta'}H$)	
26	C460	9.53	4.07	3.05	
27	K461	7.76	3.99	1.66, 1.90	
28	V462	7.58	3.80	2.16	1.06 ($C^{\gamma^1}H$); 1.16 ($C^{\gamma^2}H$)
29	F463	8.42	4.12	3.12, 3.26	7.26 (2,6); 7.18 (3,5); 7.65 (4)
30	F464	8.68	3.82	3.02, 3.62	6.74 (2,6); 6.99 (3,5); 6.78 (4)
31	K465	7.85	3.51	1.97, 2.22	1.27, 1.37 ($C^{\gamma}H$); 1.76 ($C^{\delta}H$); 3.05 ($C^{\epsilon}H$)
32	R466	7.91	3.91	1.75, 1.81	1.58 ($C^{\gamma}H$); 3.93 ($C^{\delta}H$)
33	A467	7.85	3.52	0.77	
34	V468	7.93	3.60	1.60	0.50 ($C^{\gamma^2}H$); 0.68 ($C^{\gamma^1}H$)
35	E469	8.20	3.97	1.88, 1.94	2.12, 2.46 ($C^{\gamma}H$)
36	G470	7.60	3.32, 4.03		
37	Q471	7.59	3.83	1.96, 1.99	2.17, 2.20 ($C^{\gamma}H$); 6.77, 7.53 ($N^{\epsilon}H_2$)
38	H472	7.71	4.05	3.32, 3.56	
39	N473	8.23[b]	4.67	2.56, 2.69	6.75, 7.44 ($N^{\delta}H_2$)
40	Y474	8.08	4.83	2.70, 3.15	6.99 (2,6); 6.48 (3,5)
41	L475	8.76	4.65	1.49, 1.56	1.55 ($C^{\gamma}H$); 0.67, 0.73 ($C^{\delta}H$)
42	C476	8.99	4.24	2.79, 2.81	
43	A477	9.43	4.44	1.53	
44	G478	9.85	3.52, 4.47		
45	R479	8.80	4.63	1.63, 1.97	1.84 ($C^{\gamma}H$); 3.19 ($C^{\delta}H$)
46	N480	9.24	4.48	2.97, 3.06	6.95, 7.57 ($N^{\delta}H_2$)
47	D481	8.18	4.83	2.31, 2.75	
48	C482	8.39	3.91	2.51, 2.97	
49	I483	8.28	3.78	1.62	0.82, 1.20 ($C^{\gamma}H$); 0.75 ($C^{\gamma}H_3$); 0.85 ($C^{\delta}H$)
50	I484	8.50	3.94	2.08	0.82, 1.57 ($C^{\gamma}H$); 0.56 ($C^{\gamma}H_3$); 0.58 ($C^{\delta}H$)
51	D485	6.62	5.28	2.55, 3.15	
52	K486	8.40	3.93	1.80, 1.83	1.35 ($C^{\gamma}H$); 1.65 ($C^{\delta}H$); 3.00 ($C^{\epsilon}H$)
53	I487	7.51	4.07	1.94	1.28, 1.56 ($C^{\gamma}H$); 0.93 ($C^{\gamma}H_3$); 0.85 ($C^{\delta}H$)
54	R488	8.55	4.30	1.72, 1.79	3.31 ($C^{\delta}H$)[d]
55	R489	8.24	4.19		
56	K490	7.65	4.02	1.71, 1.79	1.40, 1.43 ($C^{\gamma}H$); 3.05 ($C^{\epsilon}H$)
57	N491	7.96	4.43	2.93, 2.95	7.07, 7.76 ($N^{\delta}H_2$)
58	C492	6.78	4.86	2.70, 2.90	
59	P493		4.08	1.29, 1.55	1.86 ($C^{\gamma}H$); 3.31, 3.99 ($C^{\delta}H$)
60	A494	7.94	4.27	1.35	
61	C495	9.70	4.13	2.73, 2.94	
62	R496	8.49	3.73	1.55, 2.03	1.76, 1.80 ($C^{\gamma}H$); 3.17 ($C^{\delta}H$)
63	Y497	9.30	4.67[b]	3.50, 3.64	7.23 (2,6); 6.70 (3,5)
64	R498	8.39	3.95	1.85, 2.07	1.65, 1.67 ($C^{\gamma}H$); 3.34, 3.37 ($C^{\delta}H$)
65	K499	8.14	4.15	1.96, 2.04	
66	C500	8.47	3.94	2.97, 3.06	
67	L501	7.97	4.10	1.46, 1.89	1.48 ($C^{\gamma}H$); 0.70, 0.95 ($C^{\delta}H$)
68	Q502	8.85	4.00	2.15, 2.24	2.45, 2.63 ($C^{\gamma}H$); 6.91, 7.52 ($N^{\epsilon}H_2$)
69	A503	8.06	4.18	1.42	
70	G504	7.48	3.62, 4.27		
71	M505	7.65	4.33	2.02, 2.11	2.40 ($C^{\gamma}H$); 1.95 ($C^{\epsilon}H_3$)
72	N506	8.55	4.98	2.74, 2.80	7.05, 7.56 ($N^{\delta}H_2$)
73	L507	9.05	4.37	1.62, 1.73	1.66 ($C^{\gamma}H$); 0.75, 0.86 ($C^{\delta}H$)
74	E508	8.44	4.52	1.97, 2.10	2.28 ($C^{\gamma}H$)
75	A509	8.11	4.19	1.43	
76	R510	8.39[b]	4.29[b]	1.85, 1.92	1.70 ($C^{\gamma}H$); 3.25 ($C^{\delta}H$)
77	K511				
78	T512	8.14	4.36	4.22	1.23 ($C^{\gamma}H$)
79	K513				
80	K514				
81	K515				
82	I516	8.20	4.18	1.88	1.20, 1.48 ($C^{\gamma}H$); 0.91 ($C^{\gamma}H_3$); 0.89 ($C^{\delta}H$)
83	K517	8.42[d]	4.42		
84	G518				
85	I519	8.02	4.20	1.89	1.22, 1.48 ($C^{\gamma}H$); 0.92 ($C^{\gamma}H_3$); 0.87 ($C^{\delta}H$)
86	Q520	8.52	4.37	2.01, 2.14	2.39 ($C^{\gamma}H$)
87	Q521	8.43		2.00, 2.15	2.40 ($C^{\gamma}H$)
88	A522				
89	T523				
90	A524				
91	G525				
92	R	8.10	4.40	1.66, 1.91	1.81 ($C^{\gamma}H$); 3.24 ($C^{\delta}H$)
93	L	7.96	4.22	1.59, 1.62	1.58 ($C^{\gamma}H$); 0.88, 0.92 ($C^{\delta}H$)

[a] Uncertainties are ±0.02 ppm, except for overlapping $C^{\beta}H$, $C^{\gamma}H$, and $C^{\delta}H$ resonances for which uncertainties are ±0.04 ppm. [b] Chemical shifts refer to 294 K, pH 6.7. [c] 2,6-Resonance for Y452 occurs as a doublet (see text for discussion). [d] Chemical shifts refer to 310 K, pH 5.7.

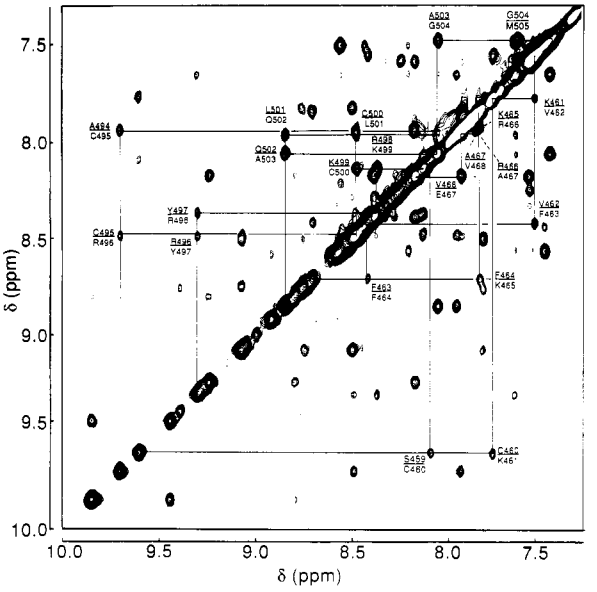

FIGURE 5: Amide region of 600-MHz NOESY spectrum of DBD-115 in ^1H_2O where NH–NH connectivities of the two segments S459–E469 and A494–M505 are indicated (below and above the diagonal, respectively). The spectrum was recorded at 300 K, pH 6.7 with a mixing time of 200 ms.

ficient guide for a "retracing" of DBD spectra. The guidelines refer to assignment of DBD-93 at 300 K, pH 6.7, unless otherwise stated. The segments S459–G470 and A494–M505 were assigned first from uninterupted sequences of strong d_{NN} NOE connectivities (Figure 5 shows these connectivities for DBD-115), in combination with (weak) $d_{\alpha N}$ and (stronger) $d_{\beta N}$ NOEs. In the first of these segments, sequential $d_{\alpha N}$ connectivities could not be observed for S459-C460 and C460-K461. However, the position of the K461 amide is inferred by d_{NN}(C460,K461) and $d_{\beta N}$(C460,K461) connectivities. The position of the S459 amide resonance is unambiguous from the d_{NN}(S459,C460), but the assignment of the S459 C$^\alpha$H is less firm. The assignments of K465–V468 are also somewhat tricky since the amides of these four residues are located within the same narrow spectral region. Sequential assignments could still be made because the V468 and R466 amides are resolved in DBD-93 spectra recorded at lower temperature (294 K) and the K465 amide is shifted to lower field (by about 0.1 ppm) in DBD-115 compared to DBD-93. The S459–G470 segment was initially identified as X-AMX-X-V-F-F-X-X-A-V-X-G, and subsequent tracing of scalar connectivities in HOHAHA and DQF-COSY spectra yielded additional assignments of resonances in the (previously) unidentified spin systems. Figure 6 shows the tracing of the $d_{\alpha N}$ connectivities of F463–G470 of this segment.

Assignments of A494–M505 were more straightforward, because the only "missing" sequential connectivity is $d_{\alpha N}$(R496,Y497). However, the Y497 amide resonance is not overlapped in the low-field spectral region and was therefore unambiguously assigned on the basis of d_{NN} and $d_{\beta N}$ connectivities. The A494–M505 segment was identified as A-AMX-X-Y-X-X-AMX-L-Q-A-G, and additional assignments were made on the basis of scalar connectivities. The position of the P493 C$^\alpha$H resonance is inferred from NOE connectivities to the amides of A494 and R496, and this assignment was confirmed at a later stage, as described below.

There is no d_{NN}(M505,N506) connectivity, but the location of the N506 amide is evident from a strong $d_{\alpha N}$(M505,N506). The following residues (N506–A509) were then traced by use of strong $d_{\alpha N}$ and weak d_{NN} connectivities, and the N506, L507, and A509 spin systems are easily recognized in DQF-

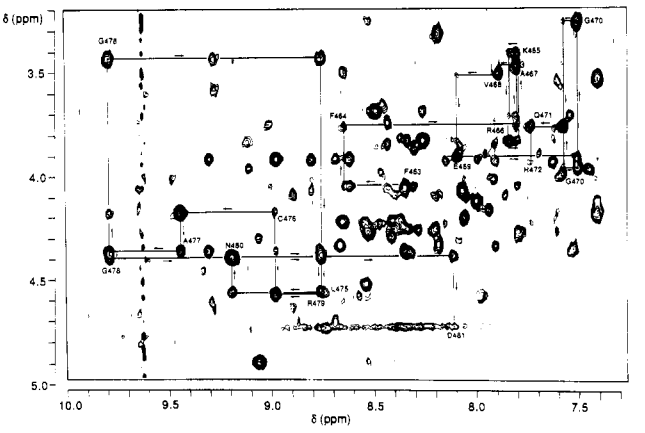

FIGURE 6: Fingerprint region of 500-MHz NOESY spectrum of DBD-93 in ^1H_2O showing the sequential assigment of two peptide fragments by intraresidue C$^\alpha$H–NH and sequential C$^\alpha$H$_i$–NH$_{i+1}$ NOE connectivities. The fragments are F463–H472 and L475–D481. The spectrum was recorded at 294 K, pH 6.7, with a 200-ms mixing time.

COSY and HOHAHA spectra. The sequential step A509-R510 is difficult owing to severe spectral overlap but could be carried out by comparison of NOESY spectra recorded at various temperature and pH conditions.

The segment L475–I484 was assigned following strong $d_{\alpha N}$ connectivities located in the low-field part of the amide region as shown for L475–D481 in Figure 6. The L475–I484 segment was initially recognized as L-AMX-A-G-X-N-AMX-AMX-I-I. The D481 C$^\alpha$H resonance is located close to the water resonance, but the intraresidue D481 C$^\alpha$H–NH NOE and sequential $d_{\alpha N}$(D481,C482) NOEs are observable in DBD-115 spectra as well as in DBD-93 spectra recorded at temperatures higher than 305 K. The D485 amide has a rather high-field chemical shift (about 6.6 ppm) and was easily located with the strong $d_{\alpha N}$(I484,D485) connectivity (D485 could be identified as an AMX spin system in HOHAHA and DQF-COSY spectra). The position of the K486 amide resonance is evident from the (isolated) $d_{\alpha N}$(D485,K486) connectivity. The intraresidue K486 C$^\alpha$H–NH NOE is more difficult to identify due to spectral overlap, but the K486 C$^\alpha$H resonance could be located with the (weak) $d_{\alpha N}$(K486,I487) connectivity in combination with d_{NN}(K486,I487) and $d_{\beta N}$-(K486,I487) connectivities. The I487–R488 sequential step is straightforwardly based on a weak $d_{\alpha N}$(I487,R488) connectivity and strong d_{NN}(I487,R488) and $d_{\beta N}$(I487,R488) connectivities.

R489 is a "weak" link in the chain of sequential assignments. The position of the R489 amide resonance is given by a strong d_{NN}(R488,R489), but the intraresidue R489 C$^\alpha$H–NH and sequential $d_{\alpha N}$(R489,K490) NOEs are very difficult to observe. The chemical shift of the R489 C$^\alpha$H resonance was deduced by locating the scalar C$^\alpha$H–NH connectivity in HOHAHA spectra that corresponded to the known chemical shift of the R489 amide resonance.

The K490–C492 segment was assigned as an X-N-AMX tripeptide from strong d_{NN} and (weaker) $d_{\alpha N}$ and $d_{\beta N}$ connectivities, and the strong $d_{\alpha\delta}$(C492,P493) and $d_{\alpha\delta'}$(C492,P493) NOEs were also easily located. The complete P493 spin system could then be traced in HOHAHA and DQF-COSY starting from the P493 C$^\delta$H–C$^{\delta'}$H connectivity, and the P493 C$^{\beta\beta'}$H–C$^\alpha$H connectivities could be used to confirm the previously assigned P493 C$^\alpha$H resonance.

The L441–C443 segment serves as a nice starting point for sequential assignments of residues at the N-terminal part of the DBD (the first zinc finger). These residues are easily recognized as an L-V-AMX tripeptide with strong d_{NN} and

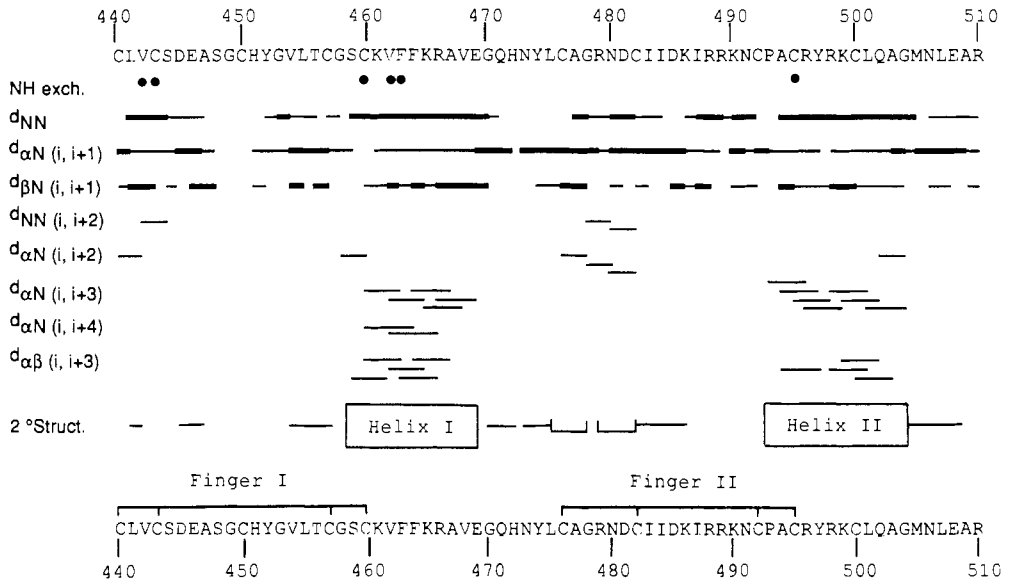

FIGURE 7: Summary of short- and medium-range NOEs and elements of secondary structure of the C440–I519 segment of the GR DBD (the segment that is common to the two proteins studied here). Strong and weak NOEs are distinguished by thick and thin lines, respectively. All but a few (three or four) sequential NOEs shown here could be unambiguously identified in both the DBD-93 and the DBD-115 protein fragments. The secondary structure elements indicated—α-helices (boxes), type I and II turns (—), and regions of extended peptide conformation (—)—were concluded on basis of the sequential and medium-range NOE connectivities, as discussed in the text. The two putative zinc-finger domains are also indicated with lines connecting Zn-coordinating cysteine residues [according to Severne et al. (1988)].

$d_{\beta N}$, and weaker $d_{\alpha N}$ connectivities. Starting from L441, the peptide chain was traced backward via strong $d_{\alpha N}$(C440,L441) and $d_{\alpha N}$(P5,C440) connectivities (the residues K1–P5 which precede C440 in DBD-93 are not part of the native GR). Given the chemical shift of P5 $C^{\alpha}H$, the complete P5 spin system could be traced in HOHAHA spectra to locate the P5 $C^{\delta\delta'}H$ resonances, which show strong sequential NOE connectivities with the R4 $C^{\alpha}H$. The R4 $C^{\alpha}H$–NH connectivity is not observable in any of our spectra, possibly owing to exchange broadening of the R4 amide resonance, but the three N-terminal residues (K1–A3) were assigned as an X-P-A tripeptide, where the identification of the P2 spin system was analogous to the assignment of P5. The P2-A3 dipeptide is distinguished from P493-A494, because the corresponding proton resonances are not present in DBD-115 spectra.

The sequential step C443-S444 was made with the strong d_{NN}(C443,S444) NOE, and the following step (S444-D445) is based on (weak) d_{NN}, $d_{\alpha N}$, and $d_{\beta N}$ connectivities. The D445 $C^{\alpha}H$ resonance is overlapping with the water resonance at room temperature but is observed at 294K (AMX spin system). The sequential $d_{\alpha N}$(D445,E446) NOE was also identified in low-temperature spectra, and the sequential assignments of E446 [AM(PT)X spin system] and A447 were straightforward from sequential d_{NN}, $d_{\alpha N}$, and $d_{\beta N}$ connectivities.

Residues S448, G449, and C450 represent a troublesome part of the sequential assignments in DBD. The assignment of S448 is not completely unambiguous due to the absence of an obvious d_{NN}(A447,S448) in combination with the presence of additional (long-range) NOE connectivities involving the $C^{\alpha}H$ and $C^{\beta}H$ resonances of A447 and amide protons of other residues. No sequential assignments could be made for G449 or C450.

The segment Y452–G458 contains residues that can be uniquely identified in HOHAHA and DQF-COSY spectra and was therefore relatively easy to assign on the basis of strong or intermediate $d_{\alpha N}$ and weak d_{NN} connectivities, although the T456 $C^{\alpha}H$ also shows a strong interstrand NOE with the L441 NH as discussed below. (The C457 amide overlaps with the C440 amide in DBD-93 spectra, but the two resonances are separated in DBD-115.) A sequential $d_{\alpha N}$(H451,Y452) connectivity could be found in NOESY spectra recorded at 294 K, and the remaining His spin system could therefore simultaneously be assigned as H472.

At this point, sequential assignments are lacking for G449-C450 (as mentioned above), Q471, N473-Y474, and the 17 C-terminal residues starting with K511. The Q471 NH resonance overlaps with the G470 NH in DBD-93, but the two resonances [and the two sequential $d_{\alpha N}$(G470,Q471) connectivities] could be distinguished by careful examination of NOESY spectra of DBD-93 and DBD-115 recorded at various temperature and pH conditions, and a $d_{\alpha N}$(Q471,H472) NOE could subsequently also be identified. N473 and Y474 were assigned as the remaining Asn and Tyr spin systems, respectively, and these assignments were confirmed by the observation of a $d_{\alpha N}$(N473,Y474) connectivity in DBD-115 and a $d_{\alpha N}$(Y474,L475) connectivity in a NOESY spectrum of DBD-93 recorded at 294 K.

Sequential assignments of resonances in the C-terminal are difficult owing to extensive spectral overlap of lysine resonances and also owing to rapid exchange of amide protons at the high-pH conditions required for protein stability. However, some isolated residues and dipeptides could still be assigned: T512 was assigned as the remaining Thr spin system present in both DBD-93 and DBD-115. The two remaining Ile spin systems in DBD-93, identified in HOHAHA and DQF-COSY spectra, could be assigned on basis of a $d_{\alpha N}$(I519,Q520) connectivity, and Q521 was subsequently assigned as the remaining AM(PT)X spin system. The assignment of I516 also made it possible to locate K515 and K517 on the basis of weak $d_{\alpha N}$ connectivities, and a similarly weak $d_{\alpha N}$ connectivity was found for the C-terminal R92-L93 dipeptide (which is not part of the native GR).

Assignments of ^1H resonances in DBD-93 are listed in Table I, and observed short- and medium-range NOE connectivities within the C440–I519 peptide segment are shown in Figure 7. The present assignments account for about 80% of all observable protons in DBD-93 and over 90% of the resonances in the C440–R510 segment. It should be noted that spin systems and connectivities found in DBD-93 spectra were, with

a few exceptions, also found in DBD-115 spectra. One of the DBD-115 NOESY spectra recorded at 600 MHz with a sample with high protein concentration also contained a few (three or four) additional (weak) sequential NOE connectivities, and these are included in Figure 7. The only structural difference detected within the common sequence of the two proteins occurs at the N-terminus, where a strong $d_{\alpha N}(i,i+2)$ connectivity was found between C440 and V442 in DBD-115, but not in DBD-93 (to be discussed below).

Elements of Secondary Structure. Sequential and medium-range NOEs shown in Figure 7 provide a basis for identification of several secondary structure elements. Stretches of strong d_{NN} and weak $d_{\alpha N}$ connectivities (Figures 3 and 4) in combination with $d_{\alpha N}(i,i+3)$, $d_{\alpha N}(i,i+4)$, $d_{\alpha\beta}(i,i+3)$, and strong intraresidue $d_{\alpha N}$ connectivities are found for residues S459–E469 and P493–G504, indicating that these regions are α-helical (Wüthrich, 1986). (Below, these segments are referred to as helix I and II, respectively.) However, the exact starting points of the two helices are not completely unambiguous. For instance, S459 shows a strong d_{NN} connectivity with C460, suggesting that this residue could be the starting point for helix I. On the other hand, we could not find medium-range $d_{\alpha N}(i,i+3,4)$ NOEs between S459 and V462 or F463, and it is therefore not possible to conclude whether the S459 carbonyl forms a hydrogen bond with the F463 amide, although a $d_{\alpha\beta}$(S459,V462) NOE connectivity supports the notion that S459 indeed is the first residue of helix I. It is also not completely clear if helix II starts with P493 or A494. These two residues cannot have a d_{NN} connectivity and NOEs between P493 C$^\delta$H and A494 NH could not be found, but we do find a $d_{\alpha N}$(P493,R496) connectivity, suggesting that P493 is the first residue of helix II with the P493 CO forming a hydrogen bond with the Y497 NH.

Isolated reverse turns involving four residues (1–4) can be distinguished from α-helical, extended, or irregular conformations by observation of a strong $d_{NN}(3,4)$ in combination with a $d_{\alpha N}(2,4)$ connectivity, although several consecutive type I turns cannot in general be distinguished from a 3_{10} helix (Wüthrich, 1986; Chazin & Wright, 1988). Type I and II turns can also be distinguished from each other by the intensity of the $d_{NN}(2,3)$ and $d_{\alpha N}(2,3)$ connectivities (Wüthrich, 1986). In our case the characteristic pattern of connectivities signifying reverse turns is observed for L439–V442 (DBD-115 only), L475–G478, and R479–C482 (Figure 7). Strong $d_{\alpha N}(2,3)$ connectivities, but no $d_{NN}(2,3)$ connectivities, were found for L475–G478 in both proteins and for the L439–V442 segment in DBD-115, indicating that these segments form type II turns. In the R479–C482 segment, on the other hand, there is a strong $d_{NN}(2,3)$ connectivity, suggesting that these residues form a type I turn. It should be noted that the complete N-terminal part of the DBD-115 protein is part of the GR and that the turn involving residues L439–V442 therefore is very likely to be a structural element of the "native" GR, whereas the presence of a proline in place of L439 in DBD-93 might force C440–V442 into a different conformation. We also note that the three turns identified here all involve regions of the protein that might be expected to be constrained by closely spaced Zn-coordinatin cysteine resiudes (Figure 7).

Regions of strong $d_{\alpha N}$ and weak or nonobservable d_{NN} connectivities, indicative of extended peptide conformations, are found for D445–A447, V454–C457, G470–H472, N473–L475, C482–K486, and M505–A509. Extended peptide segments are often associated in β-sheets, characterized by typical interstrand NOEs, especially interstrand $d_{\alpha\alpha}$ and $d_{\alpha N}$ connectivities (Wüthrich, 1986). In our case this typical

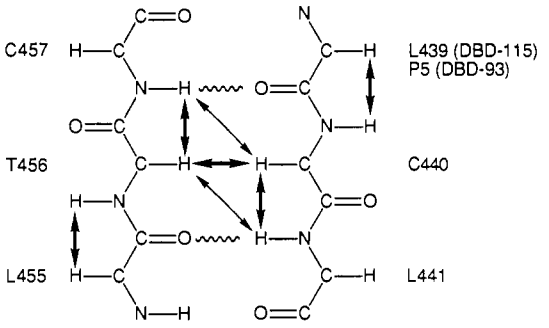

FIGURE 8: Strong sequential and long-range backbone NOEs indicating the presence of a short stretch of antiparallel β-sheet involving residues C440-L441 and T456-C457. Strong and weak NOEs are represented by thick and thin arrows, respectively. Dashed lines indicate possible interstrand hydrogen bonds (see text for details).

pattern can only be found within a very short stretch involving residues C440–L441 and T456–C457, as shown in Figure 8. The observed connectivities imply the presence of hydrogen bonds between C457 NH and L439 CO in DBD-115 (or C457 NH and P5 CO in DBD-93) and between L455 CO and L441 NH. In DBD-115 (but not in DBD-93) we observe a $d_{\alpha N}(i,i+2)$ connectivity between C440 and V442, as described above, indicating that the β-strand is immediately followed by a reverse turn.

Given the putative Zn-binding domains and the secondary structure elements identified here, it is interesting to note that the pattern "finger–helix–extended region" is repeated twice within the DBD (Figure 7). Naturally, at this point one cannot draw any conclusions about the significance of this observation, but it is still possible that it is indicative of an approximate tandem repeat in the overall protein folding.

Slowly Exchanging Amide Protons. At least eight amide NH protons in DBD-93 have exchange rates in the order of days^{-1} even at pH values above 6.0. Six of these could be unambiguously assigned to V442, C443, C460, V462, F463, and C495 by intraresidue and sequential NOE connectivities observed in a spectrum recorded at 300 K shortly after a change from ^1H_2O to ^2H_2O buffers at 277 K, and the chemical shifts of the remaining two indicate that they belong to either C440 or C457 and either L441 or F464. The slow exchange rate of F463 is consistent with the location within an α-helix (Figure 7), but the cause of the low solvent accessibility of the others cannot be deduced at this stage of structure determination, although the slow exchange of the L441 NH would be consistent with the interstrand NOEs shown in Figure 8. It can be noted that seven (of eight) slowly exchanging amide protons belong to the N-terminal (or CI) subdomain of the DBD, suggesting that finger I and helix I exhibit lower conformational flexibility than other parts of the protein.

Aromatic Ring of Y452 Is Slowly Rotating. The J coupling connectivities of the resonances in the Y452 aromatic ring are reminiscent of an ABCD spin system, with the 2,6-resonances at 7.09 and 7.12 ppm ($\Delta\nu \approx 15$ Hz in 500-MHz spectra recorded at 300 K, pH 6.7) and the 3,5-resonances centered around 6.77 ppm ($\Delta\nu \leq 10$ Hz). The 2,6-resonances also show nonequal NOE connectivities with other protons (Figure 6). The effect might be due to slow rotational motions of the aromatic ring or, less likely, to the presence of two slowly exchanging conformational states of the DBD. The former possibility is strongly supported by the absence of any other "doubled" NOE connectivities in the DBD spectra. Slow rotational exchange of aromatic rings has also been observed in other proteins, e.g., in BPTI (Wüthrich & Wagner, 1978), where the aromatic resonances of Y35 display ABCD spin system characteristics at lower temperatures (<26 °C) but

collapse into an AA'BB' spin system as the temperature is increased (>61 °C). We do not observe a similar effect when the temperature is increased to 37 °C, although the splitting between the two 2,6-resonances is somewhat smaller ($\Delta\nu \approx$ 10 Hz) at this temperature, but we do observe a large splitting in 600-MHz spectra (approximately 25 Hz at 300 K), which supports the notion of slow rotational exchange. It should be noted that slow rotational motions of the Y452 ring are consistent with a restricted conformational flexibility within the CI domain, as suggested above.

Stereospecific Assignments of Valine Methyl Groups. Spin–spin coupling constants $^3J_{\alpha\beta}$ for V442, V454, V462, and V468, estimated from a DQF-COSY spectrum processed with a high digital resolution (2.7 Hz/point), were all found to be fairly large (≥10 Hz), indicating that the C^α- and C^β-protons are oriented in a trans conformation (g^+ conformation for χ_1). In this conformation, one of the methyl groups (γ^2) is expected to be in close spatial proximity to the amide proton of the same residue. Thus, if one of the intraresidue $C^\gamma H_3$–NH NOE connectivities is significantly stronger than the other, then the corresponding $C^\gamma H_3$ resonance position can be stereospecifically assigned to the γ^2 methyl group (Zuiderweg et al., 1985). Different amide–methyl NOE intensities (at a mixing time of 80 ms) could be observed for the four valines mentioned above, and stereospecific assignments of methyl resonances could therefore be made (Table I).

Structures of the DBD Zinc-Finger Domains Differ from Those of the TFIIIA-Type Fingers. The secondary structure elements identified in the GR DBD can be compared with the results of recent NMR studies of zinc fingers from other DNA-binding proteins. Lee et al. (1989a,b) determined the solution structure of a peptide corresponding to zinc finger 31 of the *Xenopus* protein *Xfin*, and a similar peptide (from the yeast transcription activator ADR1) was studied by Pärraga et al. (1988). Both these peptides belong to the class of TFIIIA-type zinc fingers (Miller et al., 1985) with characteristic sequence homologies that distinguish them strongly from the Zn-binding domains of the steroid receptors. The TFIIIA fingers also differ from steroid receptor zinc fingers in that the Zn atom is coordinated by two Cys and two His, whereas it is coordinated by four Cys in the steroid receptors. The finger in the 25-residue peptide studied by Lee et al. consists of an N-terminal hairpin-like structure followed by a helix comprising 12 residues at the C-terminal end. We find no evidence for helical structures within the two DBD fingers per se, even though such regions would be relatively easy to detect in NOESY spectra. Neither does the pattern of sequential NOE connectivities found in the N-terminal parts of the two fingers in DBD resemble those of the *Xfin* peptide. Our data therefore confirm the notion that the TFIIIA-type and steroid receptor type zinc fingers contain different structural elements (Freedman et al., 1988b; Berg, 1989).

The present data might also be compared with a recent NMR structure determination of a peptide segment containing the first zinc-finger-like domain from the *gag* protein p55 of HIV, which is very similar to the iron-coordinating domains of rubredoxin (Summers et al., 1990). There are no obvious homologies between this segment and the hormone receptor zinc-binding domains, and the *gag* protein domain is, in particular, very different with regard to the "spacing" between the second, third, and fourth zinc-coordinating residues, suggesting that it folds in a different manner. Analogous sequential NOE connectivities are observed within the C-X-X-C-X segments of the N-terminal parts of the *gag* fragment and finger I in DBD, respectively, indicating certain structural similarities within this subregion. However, some very strong long-range ($C^\alpha H$–$C^\alpha H$) NOE connectivities in the *gag* fragment cannot be found between corresponding residues in DBD, indicating that the two protein fragments are generaly distinct in structure.

Prospects for a Complete Structure Determination. We have, in addition to the sequential NOE connectivities shown in Figure 7, also identified several (>130) long-range NOE connectivities between residues within the C440–R510 segment. This analysis is not yet completed and will therefore not be described in detail, but the results show that long-range NOEs occur in groups within and between distinct protein domains. These domains are finger I, helix I, the extended regions within G470–L475, finger II, helix II, and the extended region within M505–A509. Several long-range NOEs can be found between helixes I and II, between finger I and M505–A509, and between helix I and M505–A509. Long-range NOEs are also found between a "cluster" of aromatic and other hydrophobic residues (including F463, F464, A467, V468, Y474, A494, Y497, L501, and L507), suggesting that these might be part of a hydrophobic protein core. These observations indicate that the present assignments might provide a sufficient basis for determination of the tertiary structure within the C440–R510 segment of the glucocorticoid receptor DNA-binding domain (Härd et al., 1990).

References

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Beato, M. (1985) *Cell 56*, 335–344.

Berg, J. M. (1989) *Cell 57*, 1065–1068.

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.

Boelens, R., Scheek, R. M., Dijkstra, K., & Kaptein, R. (1985) *J. Magn. Reson. 62*, 378–386.

Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 77*, 166–169.

Carlstedt-Duke, J., Strömstedt, P.-E., Wrange, Ö., Bergman, T., Gustafsson, J.-Å., & Jörnvall, H. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4437–4440.

Chazin, W. J., & Wright, P. E. (1988) *J. Mol. Biol. 202*, 623–636.

Dahlman, K., Strömstedt, P.-E., Rae, C., Jörnvall, H., Flock, J.-I., Carlstedt-Duke, J., & Gustafsson, J.-Å. (1989) *J. Biol. Chem. 264*, 804–809.

Danielsen, M., Hinck, L., & Ringold, G. (1989) *Cell 57*, 1131–1138.

Englander, S. W., & Wand, A. J. (1987) *Biochemistry 26*, 5953–5958.

Evans, R. M. (1988) *Science 240*, 889–895.

Freedman, L. P., Luisi, B. F., Korszun, Z. R., Basavappa, R., Sigler, P. B., & Yamamoto, K. R. (1988a) *Nature 334*, 543–546.

Freedman, L. P., Yamamoto, K. R., Luisi, B. F., & Sigler, P. B. (1988b) *Cell 54*, 444.

Green, L. M., & Berg, J. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4047–4051.

Green, S., Kumar, V., Theulaz, I., Wahli, W., & Chambon, P. (1988) *EMBO J. 7*, 3037–3044.

Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870–7872.

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. (1990) *Science 249*, 157–160.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Lee, M. S., Gippert, P., Kizhake, V. S., Case, D. A., & Wright, P. E. (1989a) *Science 245*, 635–637.
Lee, M. S., Cavanagh, J., & Wright, P. E. (1989b) *FEBS Lett. 254*, 159–164.
Mader, S., Kumar, V., de Verneuil, H., & Chambon, P. (1989) *Nature 338*, 271–274.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Miller, J., McLachlan, A. D., & Klug, A. (1985) *EMBO J. 4*, 1609–1614.
Pârraga, G., Horvath, S. J., Eisen, A., Taylor, W. E., Hood, L., Young, E. T., & Klevit, R. E. (1988) *Science 241*, 1489–1492.
Ponglikitmongkol, M., Green, S., & Chambon, P. (1988) *EMBO J. 7*, 3385–3388.
Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.
Rusconi, S., & Yamamoto, K. R. (1987) *EMBO J. 6*, 1309–1315.
Schena, M., Freedman, L. P., & Yamamoto, K. (1989) *Genes Dev. 3*, 1590–1601.
Severne, Y., Wieland, S., Schaffner, W., & Rusconi, S. (1988) *EMBO J. 7*, 2503–2508.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–297.
Summers, M. F., South, T. L., Kim, B., & Hare, D. R. (1990) *Biochemistry* 29, 329–340.
Tsai, S. Y., Carlstedt-Duke, J., Weigel, N. L., Dahlman, K., Gustafson, J.-Å., Tsai, M.-J., & O'Malley, B. W. (1988) *Cell 55*, 361–369.
Umesuno, K., & Evans, R. M. (1989) *Cell 57*, 1139–1146.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347–366.
Wüthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., & Wagner, G. (1978) *Trends Biochem. Sci. 3*, 227–230.
Zuiderweg, E. R. P., Boelens, R., & Kaptein, R. (1985) *Biopolymers 24*, 601–611.

Specific Binding of *lac* Repressor to Linear versus Circular Polyoperator Molecules†

Henri M. Sasmor‡ and Joan L. Betz*

Department of Biochemistry, Biophysics and Genetics, University of Colorado School of Medicine, Denver, Colorado 80262

Received November 9, 1989; Revised Manuscript Received May 31, 1990

ABSTRACT: Gel shift assays were used to examine the binding of the lactose (*lac*) repressor to polyoperator DNA molecules. Specific binding was differentiated from nonspecific DNA association by (i) equilibrating repressor–operator complexes below the nonspecific association constant and (ii) demonstrating the effects of the inducer isopropyl β-D-thiogalactoside (IPTG) on the formation of repressor–operator complexes. With the linear polyoperator molecules, all eight operator sites could be simultaneously bound by distinct repressors. However, with circular molecules, the eight operator sites were saturable by repressor only in the nicked circular state and not in the covalently closed circular form. Under the experimental conditions used, there was no evidence of bifunctional repressor binding or loop formation. The results suggest that the conformational perturbation of DNA that occurs upon specific repressor binding was retained in topologically closed molecules and could modify other operator sites so as to make them unavailable for specific binding.

Because many regulatory proteins have been reported to induce structural distortions in their target DNA sequences, such as bending, untwisting, or looping (Kotlarz et al., 1986; Dunn et al., 1984; Hatfull et al., 1987; Koudelka et al., 1988), one determinant of protein–DNA interactions may be local conformation. The *lac* repressor–operator system has long been a paradigm for investigation of sequence-specific interactions that regulate initiation of transcription. *lac* repressor is known to bind tighter to supercoiled DNA than to relaxed circular DNA (Wang et al., 1974; Sadler et al., 1977; Whitson et al., 1987). Additionally, specific binding is characterized by a conformational change of the operator DNA, a result of either helix unwinding, DNA bending, or a combination of the two (Wang et al., 1974; Kim & Kim, 1983). Assuming one repressor bound per operator site, an unwinding of approximately 55° per specific repressor–operator interaction can be calculated from the equilibrium binding of repressor to plasmids containing 15 operator sites (Kim & Kim, 1983). However, kinetic analysis of similar polyoperator plasmids suggests that only a single repressor can be accommodated on adjacent multiple binding sites until the number of operators comprising that target exceeds four (Sadler et al., 1980). This study utilized nicked circular plasmid DNA in which no apparent steric or structural features blocked the additional operator sites.

A variety of complex geometries have been inferred for the binding of *lac* repressor to its operator, depending on the concentration of protein and the size, length, and configuration of the operator fragment. A single *lac* repressor can bifunctionally bind two appropriately separated operators in a linear DNA fragment (Kramer et al., 1987; Whitson et al., 1987). The resultant looped protein–DNA complexes show retarded

† Supported by the National Institutes of Health (GM30709).
* Author to whom correspondence should be addressed at the Department of Biology, Regis College, Denver, CO 80221.
‡ Present address: Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309.

0006-2960/90/0429-9023$02.50/0