Refined Solution Structure of the Glucocorticoid Receptor DNA-Binding Domain[†,Δ]

Herbert Baumann,[‡] Kim Paulsen,[‡] Helena Kovács,[‡,§] Helena Berglund,[‡] Anthony P. H. Wright,[⊥] Jan-Åke Gustafsson,[∥] and Torleif Härd*[,‡]

Center for Structural Biochemistry, Center for BioTechnology, and Department of Medical Nutrition, Karolinska Institutet, NOVUM, S-141 57 Huddinge, Sweden

Received June 4, 1993; Revised Manuscript Received September 20, 1993[⊕]

ABSTRACT: A refined solution structure of the glucocorticoid receptor DNA-binding domain (GR DBD) has been determined using two- and three-dimensional nuclear magnetic resonance (NMR) spectroscopy on an ^{15}N-labeled GR DBD fragment in conjunction with distance geometry and simulated annealing calculations. Thirty structures of the fragment C440-R510 of the rat GR were calculated based on 906 distance constraints obtained from NOE intensities (168 intraresidue and 738 interresidue NOEs) and 43 dihedral constraints. Average atomic root mean square (rms) differences between the 24 best structures and their geometric average are 0.70 Å for backbone atoms and 1.44 Å for all heavy atoms. Several regions that were not well defined in a previous NMR structure determination of a similar protein fragment [Härd, T., Kellenbach, E., Boelens, R., Maler, B. A., Dahlman, K., Freedman, L. P., Carlstedt-Duke, J., Yamamoto, K. R., Gustafsson, J.-Å., & Kaptein, R. (1990b) *Science 249*, 157–160] are now well-defined. The refined structure of the uncomplexed GR DBD is very similar to the crystal structure of GR DBD in a sequence specific DNA complex [Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R., & Sigler, P. B. (1991) *Nature 352*, 497–505], in particular with regard to the presence and relative positions of secondary structure elements. The backbone atom rms difference between the average NMR solution structure and the crystal structure of the DNA-complexed GR DBD is 1.8 Å. The most pronounced differences between the free and DNA-complexed states are found within the fragment C476–C482 in the second zinc-coordinating domain. The different conformations of this fragment, which includes the so-called D-box, can be rationalized in terms of conformational changes induced by protein–protein interactions upon formation of a dimeric DBD–DNA complex.

The glucocorticoid receptor (GR)[1] is a member of the superfamily of ligand-inducible nuclear transcription factors (Evans, 1988). All members of this superfamily contain a highly conserved DNA-binding domain (DBD) consisting of about 70 residues. The DBD mediates specific binding to hormone response elements on DNA, and recombinant protein fragments containing the DBD also bind specifically to these DNA sequences (Freedman et al., 1988; Dahlman et al., 1989). By analogy with the "zinc finger" motifs found in several other transcription factors, the DBD contains two zinc ions that are required for proper folding and DNA binding. However, the zinc-coordinating domains found in steroid/thyroid hormone receptors are structurally distinct (see Kaptein (1991) for a recent review of zinc finger structures).

The structure of the GR DBD (Figure 1) was initially determined in solution using NMR spectroscopy (Härd et al., 1990a,b), and the crystal structure of the same DBD fragment in a dimeric complex with DNA has been determined using X-ray diffraction methods (Luisi et al., 1991). NMR studies of the estrogen receptor (ER) DBD (Schwabe et al., 1990) and retinoic acid receptor-β (RAR) DBD (Katahira et al., 1992; Knegtel et al., 1993) reveal structures that are very similar to that of the GR DBD. The DBD consists of two subdomains that have the general composition [zinc domain–helix–extended region]. The two helices are oriented perpendicular to each other, and hydrophobic side chains of these and of the following extended regions form a protein core. The two zinc domains, which are packed against the protein body, differ structurally and functionally. The first zinc domain provides three nonspecific DNA contacts in the DBD–DNA complex, and the C-terminal residues of this domain constitute the base of the amphiphatic helix that determines sequence-specific binding by making several direct interactions with functional groups in the DNA major groove. The second zinc domain provides the entire dimerization surface for DBD–DBD interactions as well as several contacts with DNA phosphates (Luisi et al., 1991, and references cited therein).

Although the initial NMR structure and the X-ray structure of the GR DBD bound to DNA were very similar with regard to secondary structure elements, overall protein folding, and side-chain packing in the hydrophobic core, there were some remaining ambiguities regarding the conformation and conformational flexibility of the second zinc-coordinating domain.

[†] This work was supported by a research fellowship and a grant from the Swedish Natural Sciences Research Council and a grant from the Magn. Bergvall Foundation (T.H.), a grant from the Swedish Natural Sciences Research Council (A.P.H.W.), and a grant from the Swedish Medical Research Council (No. 13x-2819 to J.-Å.G.).

[Δ] Data has been deposited in the Protein Data Bank under filename 1GDA.

* Author to whom correspondence should be addressed. Telephone: +46 8 608 92 30; Telefax: +46 8 608 92 90.

[‡] Center for Structural Biochemistry.

[§] Present address: Institut für Molekularbiologie und Biophysik, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.

[⊥] Center for BioTechnology.

[∥] Department of Medical Nutrition.

[⊕] Abstract published in *Advance ACS Abstracts*, November 15, 1993.

[1] Abbreviations: GR, glucocorticoid receptor; ER, estrogen receptor; RAR, retinoic acid receptor-β; DBD, DNA-binding domain; GRE, glucocorticoid response element; NOE, nuclear Overhauser enhancement; HMQC, heteronuclear multiple quantum coherence; COSY, correlated spectroscopy; TOCSY, total correlation spectroscopy; DQF, double quantum filtered; rms, root mean square; TPPI, time proportional phase increment; DTT, dithiothreitol.

0006-2960/93/0432-13463$04.00/0

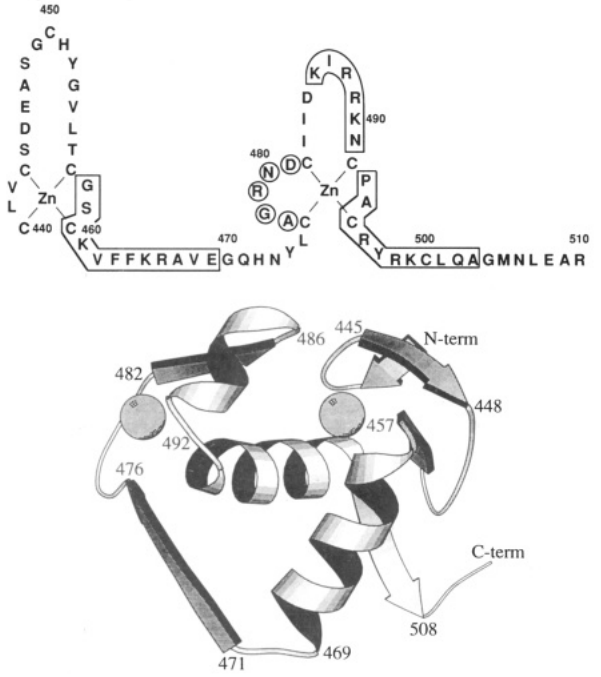

FIGURE 1: (Top) Sequence of residues C440–R510 of the rat GR DBD. (Bottom) The GR DBD solution structure drawn in a ribbon representation (Kraulis, 1991) with the two zinc ions represented by hard spheres. The helical segments of the solution structure are boxed and residues in the "D-box" are circled.

This domain (residues 476–495 of the rat GR), which protrudes out from the protein core with a zinc ion at the base, is well resolved in the crystal structure of the DBD–DNA complex, with a short loop involving residues A477–D481, a β-strand encompassing C482–D485, and a (distorted) helix consisting of residues K486–N491. The initial NMR solution structure of this region is much less well-defined, and medium-range NOE connectivities corresponding to those secondary structure elements that are observed in the crystal were not observed in previous studies (Härd et al., 1990a,b). It has been proposed that the differences between the NMR and crystal structures in the GR DBD might arise from a stabilization of secondary structure upon formation of the DNA complex (Luisi et al., 1991). According to this view, interactions between DNA and certain residues in DBD help stabilize the dimer interface, thereby causing cooperative enhancement of dimerization and DNA binding. On the other hand, large apparent disorder in loops and terminal regions resulting from too few structural (NOE) constraints is a recurring problem in structure determinations using NMR because it is often impossible to determine if lack of NOE connectivities is due to incomplete spectral analysis or true disorder. The matter is further complicated by the fact that certain structural elements such as loops result in only a few NOE connectivities even though the structure itself is stable and ordered.

To bring some further insight into this problem we recently measured the extent of rapid (picosecond) backbone fluctuations in the GR DBD using ^{15}N relaxation measurements on an ^{15}N-labeled GR DBD sample (Berglund et al., 1992). The results indicated that rapid motions are limited and uniform over the entire protein backbone, as expected for a well-structured protein. In particular, we found no evidence for extensive rapid backbone motions within the second zinc domain. These results strongly suggest that the second zinc domain is not disordered in the uncomplexed state of DBD and that NMR structure determinations of this domain yield poorly defined structures due to lack of structural constraints rather than to inherently high flexibility. A redetermination of the structure of the uncomplexed form of GR DBD in solution at a higher resolution was therefore strongly motivated to resolve the issue of possible conformational changes in this domain upon binding to DNA.

The present paper describes a high-resolution structure determination of the fragment C440–R510 of the rat GR (Figure 1) using an ^{15}N-labeled protein fragment and three-dimensional heteronuclear NMR spectroscopy to facilitate resonance assignments and resolve NOEs that overlap in ordinary NOESY spectra. The structure presented here has an overall backbone rms difference of 0.93 Å for pairwise comparisons of the 24 best structures (0.70 Å for pairwise comparisons with the average structure), whereas the initial NMR structure of the GR DBD (Härd et al., 1990b) had a corresponding average rms difference value of about 2.0 Å. The backbone of the second zinc-coordinating domain, which was poorly resolved in the initial structure, is now well-defined. Comparisons between the uncomplexed (NMR) and complexed (X-ray) forms of the GR DBD reveal a few small, but significant, conformational changes upon DNA binding. The largest conformational change occurs within a region (residues C476–C482) containing the "D-loop", which forms part of the DBD dimerization surface in the dimeric DBD–DNA complex.

MATERIALS AND METHODS

Sample Preparation. Details of the expression, purification, and preparation of ^{15}N-labeled GR DBD (DBD82; residues L439-Q520 of the rat GR, with an N-terminal Met residue) have been described elsewhere (Berglund et al., 1992). For the present study we used an approximately 1 mM solution (90% H_2O/10% ^2H_2O) of protein in 250 mM NaCl, 20 mM phosphate, and 2 mM DTT at pH 6.0. DTT is necessary to prevent oxidation of cysteine residues and, in combination with careful nitrogen purging of the NMR sample, it keeps the protein stable for several months.

NMR Spectroscopy and Data Processing. All spectra were recorded on Varian Unity 500 or 600 NMR spectrometers at magnetic fields of 11.74 or 14.09 T, respectively, using inverse detection probes. ^1H and ^{15}N chemical shifts at 299 K (supplementary material) are referenced to H_2O at 4.75 ppm and to external ^{15}N-benzamide at 105.4 ppm (Levy & Lichter, 1979), respectively. Experiments were carried out at temperatures ranging from 289 to 303K.

All multidimensional NMR experiments were recorded in the phase-sensitive mode, using the hypercomplex method (States et al., 1982) unless otherwise stated. Water supression was achieved by either presaturation during relaxation and evolution delays, presaturation in combination with SCUBA (Brown et al., 1988), or in 3D experiments, by a combination of SCUBA and a jump–return (Plateau & Gueron, 1982) observation pulse.

Two-dimensional homonuclear DQF-COSY (Rance et al., 1983), NOESY (Macura & Ernst, 1980), and clean-TOCSY spectra (Griesinger et al., 1988) were recorded using 512–1024 t_1 increments, 32–96 transients/increment, 2048 complex data points/transient, and WALTZ decoupling of ^{15}N in both dimensions. The mixing time was 50–80 ms in various TOCSY experiments and 60 or 150 ms in NOESY experiments.

Two-dimensional heteronuclear (^1H and ^{15}N) HMQC (Müller, 1979; Bax et al., 1983), HMQC–NOESY, and HMQC–TOCSY (Gronenborn et al., 1988) were recorded using 256 t_1 increments and 2048 complex data points in the acquisition domain.

A *J*-modulated [^{15}N,^1H]-COSY experiment (Billeter et al., 1992) used for determination of $^3J_{HN,H\alpha}$ coupling constants

was recorded with the TPPI method (Marion & Wüthrich, 1983). Eight spectra with different delays (12, 27, 40, 50, 59, 70, 80, and 91 ms) for evolution of the $^3J_{HN,H\alpha}$ coupling were collected in this experiment.

Three-dimensional heteronuclear NOESY–HMQC and TOCSY–HMQC spectra (Marion et al., 1989) were recorded at 299 K using 2 × 1024 complex data points in the acquisition domain and 2 × 160 t_1 increments. A total of 2 × 30 t_2 increments (^{15}N domain), 16 transients/increment, and a mixing time of 150 ms were used in the NOESY–HMQC spectrum, and 2 × 16 t_2 increments, eight transients/increment, and a 60-ms clean-TOCSY spin lock were used in the TOCSY-HMQC spectrum.

Spectra were processed using Varian software (VNMR 4.1) on a Sun Sparc 330 work station or using FELIX from HARE Research. Typically, data processing involved apodization with shifted Gaussian functions in the acquisition domain and shifted sine/cosine bell functions in the other time domains. Processed 2D and 3D spectra normally contained 2048 × 2048 and 512 × 128 × 512 real data points, respectively.

Data Analysis. ^{1}H and ^{15}N resonance assignments were made using standard methods (Wüthrich, 1986; Marion et al., 1989). Stereospecific assignments for several prochiral methylene groups were made according to Hyberts et al. (1987), i.e., by identifying predominant rotameric states using $^3J_{H\alpha,H\beta}$ coupling constants measured in DQF-COSY spectra and intraresidue NOEs measured in NOESY spectra recorded with a short (60 ms) mixing time. Assigned χ^1 rotamers were confirmed by $^3J_{N,H\beta}$ coupling constants measured in a NOESY spectrum recorded without ^{15}N decoupling during t_2, following the method of Montelione et al. (1989). Large (>8.5 Hz) and small (<5.5 Hz) $^3J_{HN,H\alpha}$ coupling constants were identified based on the appearance of the corresponding amide cross peaks as a function of the evolution delay in the J-modulated [^{15}N,^{1}H]-COSY experiment. In this experiment a cross peak changes sign, from positive to negative, when the evolution delay exceeds $(2^3J_{HN,H\alpha})^{-1}$ (Billeter et al., 1992). Thus, negative cross peaks in the spectrum recorded with a 59-ms evolution delay indicate that $^3J_{HN,H\alpha} > 8.5$ Hz and cross peaks that remain positive with an evolution delay of 91 ms indicate that $^3J_{HN,H\alpha} < 5.5$ Hz.

NOEs were quantified as distance constraints based on cross peak volumes measured in a NOESY spectrum recorded with a mixing time of 60 ms. The conversion of volumes into distances was based on calibration of observed intraresidue and sequential NOEs within the two longer helical fragments G458–E469 and P493–A503 against distances expected for a regular α-helix (Wüthrich, 1986) and intraresidue NOEs between the 2,6 and 3,5 protons within tyrosine residues. NOE volumes involving HN protons were corrected for the presence of 10% D_2O in the sample. Distance constraints were divided into four classes: very strong (<2.4 Å), strong (<2.7 Å), medium (<3.3 Å), and weak (<5.0 Å), and lower distance limits were in all cases set to the sum of the van der Waals radii of the involved atoms. NOEs that could only be observed in spectra recorded with a 150 ms mixing time were all interpreted as weak (<5.0 Å) distance constraints. Pseudoatoms with appropriate distance corrections were created for methyl protons, aromatic ring protons and nonstereospecifically assigned methylene protons in leucines. A (reduced) pseudoatom correction of 0.3 Å was used to account for effects due to rapid rotation of methyl groups (Koning et al., 1990).

Measured $^3J_{HN,H\alpha}$ coupling constants were used to derive φ dihedral angle constraints to $-160° < \varphi < -80°$ for residues with $^3J_{HN,H\alpha} > 8.5$ Hz and to $-90° < \varphi < -40°$ for residues with $^3J_{HN,H\alpha} < 5.5$ Hz. Dihedral angle constraints were also added to impose identified χ^1 rotamers, with a tolerance of ±45° about staggered conformations. In a few cases it was not possible to distinguish between two possible rotamers, but the third possibility could definitely be excluded, and the corresponding χ^1 dihedral constraint was then constructed to allow conformations within two rotameric states. The proper tetrahedral zinc coordination was imposed by adding sulfur–sulfur distance constraints between zinc-coordinating cysteines. These distances were allowed to vary between 3.80 and 3.85 Å, calculated assuming a Zn–S distance of 2.35 Å (Pauling, 1967) and tetrahedral zinc coordination. In total, 906 nonredundant NOE distance constraints and 43 dihedral constraints (Table I) could be determined and were used in the calculations together with the 20 zinc-coordinating constraints.

No hydrogen bond constraints were imposed in the structure calculations. Measurements of amide proton exchange rates, which normally are performed to obtain complementary information about hydrogen bonding, cannot easily be carried out on GR DBD fragments, as concluded earlier (Härd et al., 1990a).

Structure Calculations. Structure calculations and analyses were carried out on IBM 6000 and Silicon Graphics 310VGX workstations, respectively, using the InsightII v2.1.2/NMRchitect v1.0 (Biosyn Technologies, San Diego, CA) and CHARMM (Brooks et al., 1983), version 22 (Harvard University, Department of Chemistry, Cambridge, MA), software packages. Three-dimensional structures of the fragment C440–R510 were determined by distance geometry (DG) calculations (Havel et al., 1983), using the DGII program (Havel, 1991) supplied with the Biosym program package followed by restrained energy minimization using the force field in version 22 of CHARMM. In short, sequential tetrangle bound smoothing was followed by embedding in four dimensions with prospective metrization and constant majorization. The simulated annealing optimization protocol involved 40 000 steps of dynamics at 200 K with a time step of 0.2 ps, cooling to 0 K, and conjugate gradient minimization, using standard DGII parameters, constraint violation force constants, and atomic weights (Havel, 1991; *NMRchitect User Guide*, version 1.0 San Diego: Biosym Technologies, 1992). Thirty structures were calculated, and six of these were discarded based on significantly higher error function values. The remaining 24 structures were energy minimized in CHARMM with 200 steps of steepest descent followed by 300 steps of the conjugate gradient Powell method. No potentials for hydrogen bonds or electrostatic interactions were included in the minimizations. The restraints were the same as those used in the distance geometry calculations. The force constant for NOE distance constraints was 32 kcal mol^{-1} Å^{-2}, and the force constants for dihedral angle constraints were 32 and 10 kcal mol^{-1} rad^{-2} for φ and χ^1 dihedrals, respectively. An average structure was calculated by superpositioning the backbone atoms of the calculated structures on the structure with lowest energy in combination with the smallest constraint violations and then calculating the unweighted average atomic positions. The average structure was then minimized as described above and used as a computational reference and for visual comparisons with the X-ray structure.

The structures were analyzed with respect to the precision of atomic positions and dihedral angles and further characterized with respect to backbone dihedral angle conformations and hydrogen bonding. Dihedral angle order parameters, S^{angle}, reflecting the precision of the corresponding dihedral within the ensemble were calculated following Hyberts et al. (1992). A value of S^{angle} approaching unity indicates a very

Table I: Summary of Constraints Used in the Structure Calculations and Structural Statistics[a] of the Calculated Structures of GR DBD

NOEs		906
intrare	168	
sequen ($i=j=1$)	276	
medium range ($2 \le i-j \le 4$)	185	
long ra ($i-j \ge 5$)	277	
dihedral angle		43
φ	28	
χ^1	15	
zinc coordination		20
total no. of constraints		969 (13.6/residue)
total energy (kc mol^{-1})		342 ± 10
bonds (1113)	41 ± 1	
angles (200	194 ± 3	
dihedrals (2 11)	244 ± 5	
impropers (199)	2.4 ± 3.0	
van der Waals	−240 ± 6	
NOE restraints (906)	64 ± 3	
torsion restraints (43)	17 ± 5	
deviations from idealized covalent geometry		
bonds (Å)	0.011 ± 0.0015	
angles (deg)	0.046 ± 0.008	

[a] Energies are calculated using the CHARMM (Brooks et al., 1983) version 22 force field. Force constants for experimental restraints are given in the Materials and Methods section. The numbers of bonds, angles, dihedrals, impropers, distance restraints, and torsion restraints are given in parentheses.

well-defined dihedral angle whereas an isotropic distribution yields $S^{angle} = 0$ (but $S^{angle} = 0$ must not necssarily reflect an isotropic distribution). Criteria for hydrogen bonding were imposed following Brocklehurst & Perham (1993); the distance between the donor hydrogen and acceptor oxygen must be less than 2.5 Å, and the angle defined by hydrogen–oxygen–carbonyl must be larger than 90°. These criteria are fulfilled in at least 18 of the 24 structures for all hydrogen bonds mentioned in the Results and Discussion sections. Structural rms differences quoted in the text indicate comparisons with the average structure unless a pairwise comparison of all structures is explicitly stated. (Table II contains both types of statistics, whereas only pairwise rms differences per residue are reported in Figure 3.) It should be noted that rms difference comparisons containing "all atoms" sometimes might be erroneous and too large due to the specific atom labeling of phenyl and tyrosyl rings and carboxylate groups. This is because the computer program evaluating rms differences do not always consider the inherent symmetry of these groups and therefore might give a large rms difference even in the case of perfect overlap (P. Kraulis, personal communication).

RESULTS

Resonance Assignments and Collection of Experimental Constraints. Proton resonance assignments of the DBD82 protein fragment were carried out following previous NMR studies of the GR DBD (Härd et al., 1990a; Berglund et al., 1992). Backbone ^{15}N resonance assignments of DBD82 at pH 7.6 were also recently reported by us (Berglund et al., 1992). For the present structure refinement, carried out at pH 6.0, we were able to completely assign all backbone ^1H and ^{15}N resonances within C440–R510, except for the backbone amide ^1H and ^{15}N resonances of N473. The Hα and Hβ protons of N473 could still be assigned by identifying sequential NOE connectivities to the NH proton of Y474. In total, about 90% of all proton resonances within C440–R510 were assigned (80% of the remaining unassigned resonances belong to side chains of arginine and lysine residues). ^1H and ^{15}N resonance assignments for C440–R510 are available as supplementary material. Cross peaks corresponding to residues within the C-terminal fragment K511–Q520 could be identified in HMQC and HSQC spectra. However, no interresidue NOE connectivities could be identified within the C-terminal, probably due to a disordered structure of this fragment.

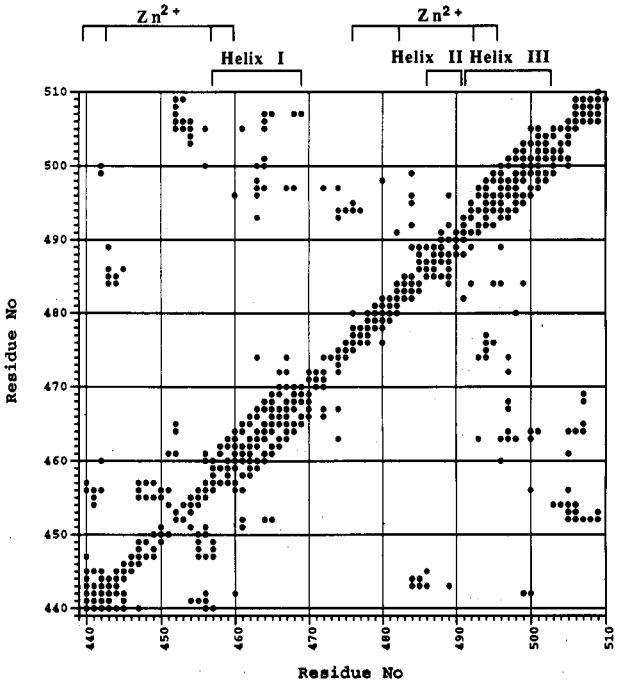

FIGURE 2: Diagonal plot of interresidue NOEs observed within the fragment C440–R510 of the rat GR DBD. Secondary structure elements and zinc coordination are also indicated.

The stereospecific assignments of valine methyl groups were found to be the same as reported previously (Härd et al., 1990a). Resonances of 16 diastereotopic Hβ pairs were stereospecifically assigned, and 15 of the corresponding χ^1 rotamers were unambiguously identified and imposed as χ^1 dihedral angle constraints in the structure calculations.

Over 1000 NOE connectivities could be assigned in NOESY and 3D NOESY–HMQC spectra and were quantified as distance constraints. The DGII program identified 906 of these as nonredundant distance constraints. A summary of the distance and dihedral angle constraints is given in Table I. A distance connectivity plot of interresidue NOE connectivities is shown in Figure 2. An overview of the number of NOE distance constraints per residue is given in Figure 3A. The complete DGII constraint file is available as supplementary material.

Precision of Calculated Structures. The final 24 energy-minimized structures were analyzed with respect to potential energy, atomic rms differences, precision of dihedral angles, and violations of distance and dihedral angle constraints. Minimization of the distance geometry structures removed serious atomic overlaps and yielded structures with VdW energies of −240 ± 6 kcal/mol. Energy statistics are given in Table I. Figure 3B shows the average distance constraint violation per residue, i.e., the sum of violations for each residue in all structures divided by the number of distance constraints for that specific residue and the number of structures. Graphs of this type are useful when identifying "troublesome" regions in which one or several misassignments can be expected to yield significantly higher average violations. The maximum average distance constraint violation per residue and structure is in the order of 0.08 Å, indicating that the quality (or accuracy) of the distance constraint data is good. This conclusion is supported by residual violation statistics of

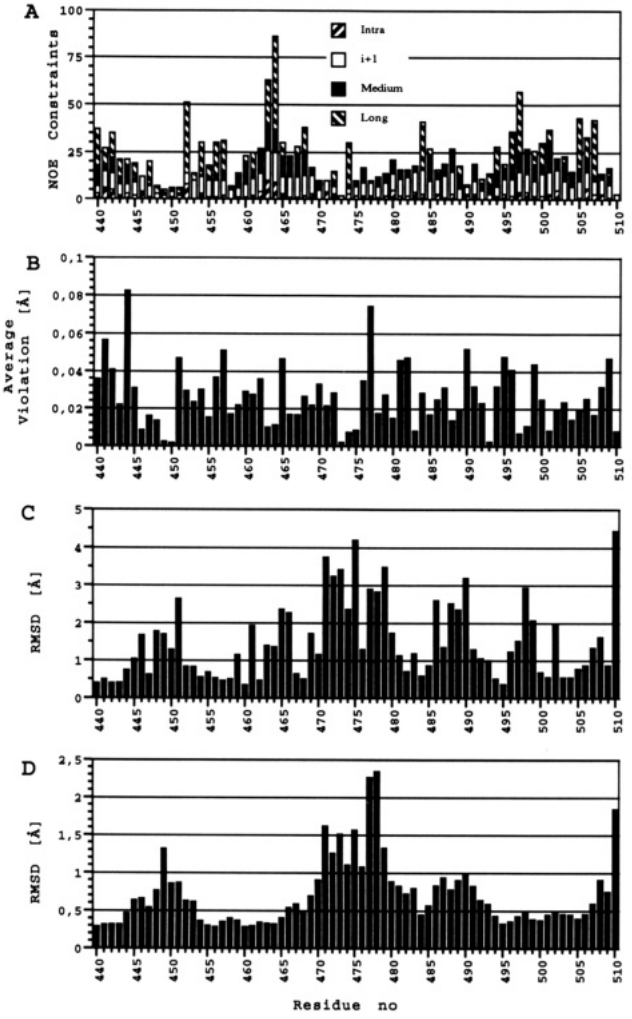

FIGURE 3: NOE distance constraints and violation statistics for the 24 best structures. (A) Histogram of the number of NOE distance constraints per residue used in the structure calculations. Interresidue NOEs have been counted twice, i.e., once for each of the two residues involved in the NOE. Indicated in the figure are intraresidue (filled), sequential (diagonal lines, lower left–upper right), medium-range (nonfilled), and long-range NOE connectivities (diagonal lines, upper left–lower right). (B) The average sum of violations per residue. (C) All atom average pairwise rms differences per residue. (D) Same as C but for the backbone atoms (N, C^{α}, C).

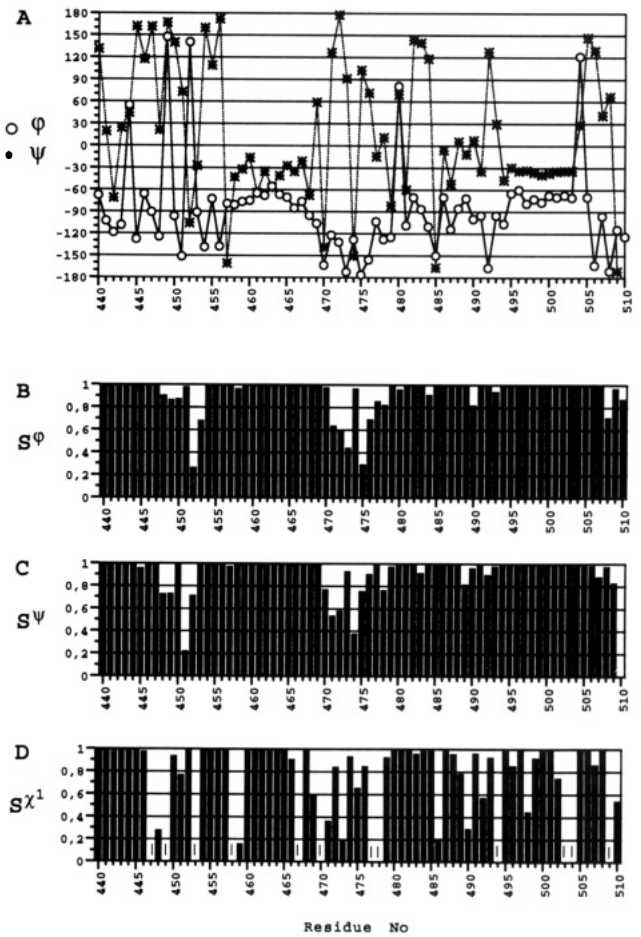

FIGURE 4: (A) Average φ and ψ dihedral angles for all residues in the 24 final structures. Angular order parameter S^{angle} for φ (B), ψ (C), and χ^1 (D) dihedral angles in the 24 final structures.

individual constraints: on average there are less than two distances per structure that are violated by more than 0.4 Å, and the largest distance violation found in any structure is 0.52 Å. No dihedral angle constraint is violated by more than 10° in any of the 24 final structures.

Atomic rms differences between the final 24 structures were calculated as pairwise rms differences and as the differences between each of the structures and their geometric average (Table II). Pairwise rms differences per residue were also calculated (Figure 3C,D). Figure 4 shows average values of φ and ψ dihedral angles together with the precision of each dihedral angle monitored as a backbone dihedral angle order parameter, S^{angle} (Hyberts et al., 1992), as described in the Materials and Methods section.

Superimposed backbone traces of the 24 calculated DBD structures are shown in Figure 5. The overall protein fold is well-defined, by NMR standards, with overall average rms differences of 0.70 ± 0.15 Å for backbone atoms (N, C^{α}, C) and 1.44 ± 0.16 Å for all heavy atoms. The structure of the N-terminal residues C440–G470 has an average backbone rms differences of 0.39 Å and is better defined than the C-terminal residues Q471–R510, which have a corresponding rms difference of 0.79 Å. Three distinct helices can be observed

Table II: rms Difference Statistics for Calculated GR DBD Structures and Comparison with the Structure of DBD Bound to DNA

		av rms differences between calcd structures[a]		comparison with structure of DBD bound to DNA[b]	
fragment	residues[c]	backbone[d]	all heavy atoms	backbone	all heavy atoms
all residues	(C440–R510)	0.70 ± 0.15(0.93 ± 0.19)	1.44 ± 0.16(1.80 ± 0.19)	1.83 ± 0.11	2.53 ± 0.11
first domain	(C440–G470)	0.39 ± 0.10(0.51 ± 0.12)	0.94 ± 0.11(1.14 ± 0.16)	1.23 ± 0.08	1.87 ± 0.10
second domain	(Q471–R510)	0.79 ± 0.22(1.05 ± 0.27)	1.65 ± 0.21(2.07 ± 0.24)	1.93 ± 0.18	2.72 ± 0.13
first domain	(C440–C460)	0.37 ± 0.130(0.48 ± 0.15)	0.82 ± 0.15(0.99 ± 0.21)	1.36 ± 0.09	1.89 ± 0.14
second zinc domain	(C476–C495)	0.64 ± 0.30(0.86 ± 0.37)	1.31 ± 0.25(1.62 ± 0.30)	2.24 ± 0.24	2.83 ± 0.26
helix I	(G458–E469)	0.18 ± 0.40(0.21 ± 0.05)	0.96 ± 0.11(1.14 ± 0.18)	0.58 ± 0.02	1.54 ± 0.15
helix II	(K486–N491)	0.29 ± 0.11(0.35 ± 0.14)	1.39 ± 0.27(1.64 ± 0.25)	0.44 ± 0.10	1.79 ± 0.20
helix III	(P493–A503)	0.18 ± 0.60(0.24 ± 0.06)	1.03 ± 0.16(1.24 ± 0.20)	0.50 ± 0.06	1.76 ± 0.20
helices I and III		0.24 ± 0.06(0.30 ± 0.06)	1.03 ± 0.10(1.24 ± 0.14)	0.74 ± 0.04	1.79 ± 0.15

[a] Average RMS differences of the 24 structures compared to their geometric average. Numbers in parentheses indicate average pairwise RMS differences. [b] Superposition of calculated structures with the corresponding residues of GR DBD in a sequence-specific DNA complex (Luisi et al., 1991). [c] Residues within superimposed fragments. [d] Backbone N, C^{α}, and C atoms.

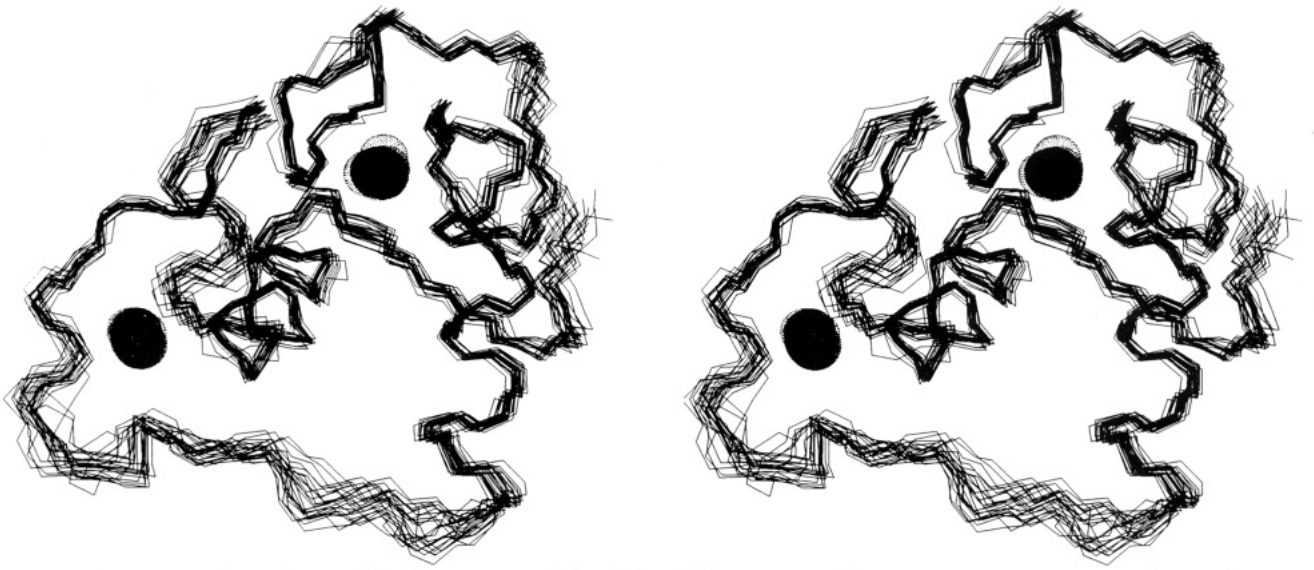

FIGURE 5: Stereoview of superimposed backbone traces of the 24 best DG structures. The structures are superimposed to minimize rms differences of backbone atoms for all residues (C440–R510). The positions of coordinated zinc ions are indicated with dotted VdW surfaces (all 48 zinc ions are shown).

in the structure, and the rms differences within these regions are in the order of 0.2 Å. A close inspection of Figures 3 and 4 reveals that only two parts of the structure are somewhat less ordered: a five-residue fragment of the first zinc-coordinating domain (S448–Y452) and the first 10 residues (G470–R479) of the second zinc-coordinating domain. Higher rms differences per residue in combination with lower S^{angle} values for the backbone dihedrals in these regions are also indicative of less well-defined structures. The backbone dihedral angle precision is very poor ($S^{angle} \sim 0.3$) for the two peptide bonds joining H451 to Y452 and for Y474 to L475, respectively. In the first case low S^{angle} values are due to the presence of two distinct structures, in which the peptide bond joining H451 to Y452 is oriented at approximately 180° relative to each other, thereby yielding low S^{angle} values for the corresponding dihedrals. However, in the case of the peptide bond joining Y474 to L475 the low S^{angle} values reflect an almost isotropic distribution of the angles. A majority of the residues in all other regions have per residue rms differences in the order of 0.5 Å and S^{angle} values for φ and ψ dihedral angles exceeding 0.95, indicating that their conformations are well-defined.

Characterization of the Structure. The secondary structure elements within GR DBD were assigned by analysis of average backbone dihedral angles and the hydrogen bonds identified in the calculated structures. The observed NOE connectivities together with a visual inspection of the structures also aided in this analysis. The central features of the structure are three helical regions: G458–E469 (helix I), K486–N491 (helix II), and P493–A503 (helix III). Helices I and III are oriented perpendicular to each other and form the base of the hydrophobic core. Helix II is part of the loop region forming the second zinc-coordinating domain. Helices I and III are both regular α-helices as judged from consecutive medium range NOE ($d_{i,i+3}$ and $d_{i,i+4}$) connectivities, φ and ψ dihedral angles around –60°, and the typical pattern of hydrogen bonds involving amide protons and carbonyls in residues within the fragments C460–V468 and P493–A503, respectively. Helix II is a somewhat distorted helix with average φ and ψ dihedral angles of -90 ± 25 and -20 ± 25, respectively. Sequential and medium-range NOE connectivities within this region are shown in Figure 6. Several $d_{i,i+3}$ and $d_{i,i+2}$ NOE connectivities make this short helical region more reminiscent of a distorted 3_{10}-helix than of a regular α-helix. However, we do not observe

FIGURE 6: Summary of sequential and medium-range NOE connectivities observed within the fragment K486–N491 (the "distorted" helix II region) in the GR DBD. Relative NOE crosspeak volumes are represented by the thickness of the connecting bars (<2.7 Å thick, <3.3 Å medium, and <5 Å thin).

any distinct hydrogen bonds within K486–N491 (given the criteria mentioned in Materials and Methods), and it is therefore uncertain if this region can be assigned a distinct regular secondary structure. The position of helix II is well defined by long-range NOE connectivities from K486 and R489 to C443 and/or D445.

Six short stretches of β-strands were identified based on sequences of φ and ψ dihedral angles around –100° and +120°, respectively. The first zinc-coordinating domain is built up by three β-strands and three short loops arranged in an alternating fashion. The β-strands formed by residues C440–L441 and V454–C457 form a short antiparallel β-sheet with a hydrogen bond between the amide of L441 and the carbonyl of L455. Helix I is linked to the second zinc-coordinating subdomain by a β-strand involving residues Q471–L475. A loop region between C476 and C482, including the so-called D-box, is followed by a β-strand between residues C482–D485. Helix III is followed by a β-strand involving residues M505–L507.

The φ dihedral angles of S444 and N480 are positive, i.e., located outside the regions of the Ramachandran plot that are normally considered to be sterically allowed for nonglycine residues. The HN,Hα intraresidue NOE cross peak volumes of S444 and N480 are about four times larger than the corresponding volumes measured within the helical domains,

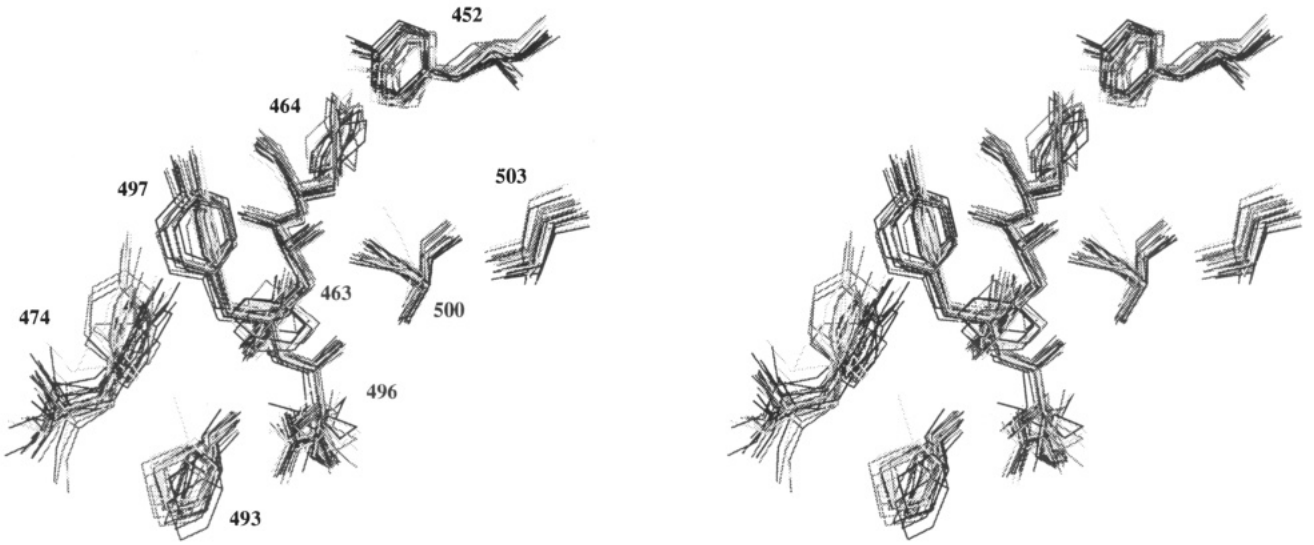

FIGURE 7: Stereoview of residues forming part of the hydrophobic core in the GR DBD (Y452, F463, F464, Y474, P493, R496, Â97, C500, and A503). The structures are superimposed to minimize rms differences of the heavy atoms for all residues (C440–R510).

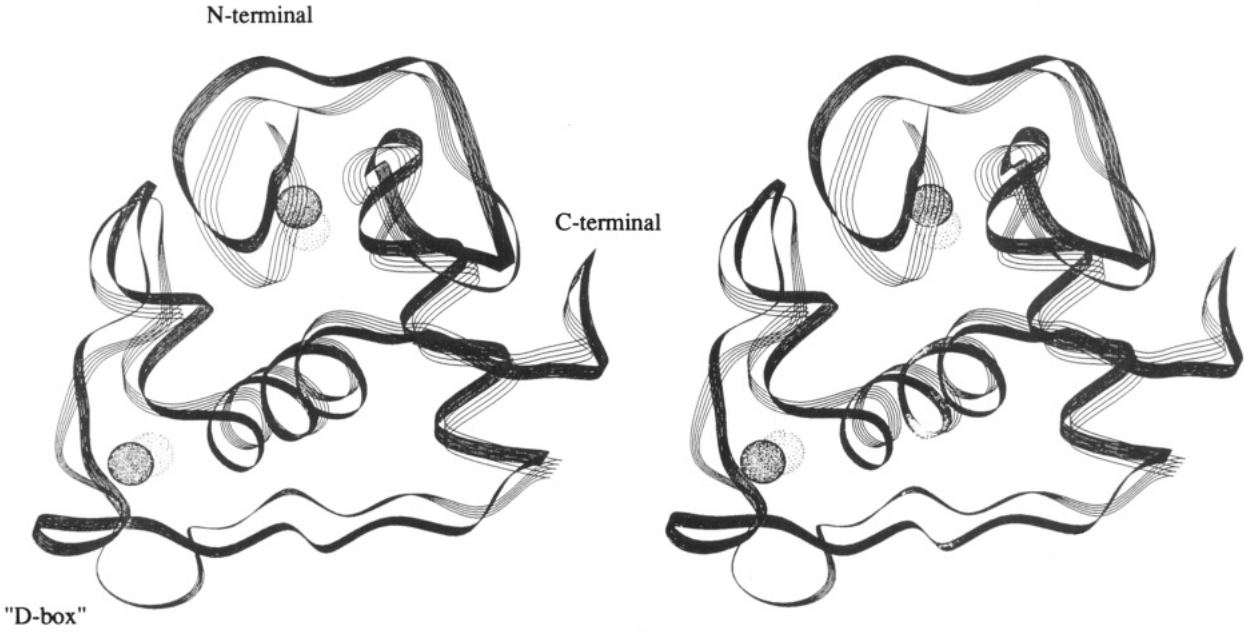

FIGURE 8: Stereoview of a ribbon representation of the average NMR structure of the uncomplexed DBD (DBD_{NMR}) superimposed on the crystal structure of DBD in a sequence specific DNA complex ($DBD_{X\text{-}ray}$) (Luisi et al., 1991). The DBD_{NMR} structure is represented by a lighter outline and $DBD_{X\text{-}ray}$ by a darker outline.

indicating that the corresponding distances in these cases are about 2.4 Å as expected for residues with positive φ dihedral angles (Ludvigsen & Poulsen, 1992).

The structures of the residues forming part of the hydrophobic core are shown in Figure 7. The four aromatic residues Y452, F463, F464, and Y497 all have well-defined side-chain conformations, and the large number of NOE constraints (50–85 each) defines, to a large extent, the tertiary fold of the GR DBD. The fifth aromatic residue (Y474) also seems to be part of the hydrophobic core, but the side chain conformation of this residue is less well-defined. The hydrophobic core also contains well-defined side chains of P493, R496, C500, and M505. It is somewhat uncommon to find an arginine side chain with a protein interior, but the R496 side chain conformation is indeed very well-defined, and a similar conformation is present in the crystal structure of GR DBD described by Luisi et al. (1991).

Comparison with the Crystal Structure of DBD Bound to DNA. The 2.9-Å crystal structure of the GR DBD–DNA complex reported by Luisi et al. (1991) contains a dimeric DBD entity in which only one of the monomers binds in a sequence specific manner of a GRE half site. The structures of the two DBD monomers in this complex are very similar, with a backbone rms difference of 0.6 Å. Below we compare our refined structure of the uncomplexed GR DBD to the crystal structure of the specifically bound DBD monomer. This structure is in the following referred to as "$DBD_{X\text{-}ray}$" whereas our refined solution structure is termed "DBD_{NMR}".

Average atomic rms differences between the 24 DBD_{NMR} structures and $DBD_{X\text{-}ray}$ (Luisi et al., 1991) are summarized in Table II. The average DBD_{NMR} structure was also superimposed on $DBD_{X\text{-}ray}$ by minimizing the rms differences for backbone atoms of the three helices (G458-E469, K486-N491, and P493-A503), and the result is shown in Figure 8. The overall backbone rms difference was then 1.83 Å with the largest differences found in the second zinc-coordinating domain (2.24 Å). The conformation and relative orientation of the three helices are nearly identical in the two structures with a backbone rms difference of 0.5 Å. In fact, there are only two regions of the protein where there seem to be

significant differences between the backbone conformation of DBD_{NMR} compared to $DBD_{X\text{-}ray}$, as described in the following text.

The N-terminal has a somewhat different conformation in DBD_{NMR} compared to $DBD_{X\text{-}ray}$ (Figure 8). In DBD_{NMR}, the fragment C440–A447 forms an almost planar loop, whereas the same loop in $DBD_{X\text{-}ray}$ is nonplanar with a right-handed conformation. As a consequence, the backbone atoms of C440 are located on different "sides" of the backbone atoms of A477 in the two structures. The NMR structure is very well-defined in this region, and the observed conformation is defined by several NOE constraints which would be violated in the conformation found in $DBD_{X\text{-}ray}$.

A close examination of the dimer interface, including the D-box region (residues A477–D481), reveals that the average planes through the loops formed by the residues C476–C482 are oriented at an angle of about 90° relative to each other in the DBD_{NMR} and $DBD_{X\text{-}ray}$ structures (Figure 8). The difference originates in the backbone φ and ψ dihedral angles of residues L475 and C476 which all differ by approximately 60° in the two structures. The DBD_{NMR} structure is less well defined in this part of the molecule, but there is still no single structure in which the loop has an orientation similar to that in $DBD_{X\text{-}ray}$. Several NOE sequential and medium-range distance constraints would also be violated with the loop conformation found in the $DBD_{X\text{-}ray}$ structure.

In addition to these two more obvious differences between the DBD_{NMR} and $DBD_{X\text{-}ray}$ structures there are also some more minor differences between backbone conformations within the two zinc-coordinating domains, as can be seen in Figure 8. However, these differences are rather subtle with only minor average differences in, for instance, backbone dihedrals.

DISCUSSION

Relation of the Present Study to Previous NMR Studies of the GR DBD. The present study represents a redetermination of the solution structure of the GR DBD. The precision of the present structure of the GR fragment C440–R510 is significantly higher than in the previously published NMR structure of the same fragment (Härd et al., 1990b). The overall improvement in precision is evident when comparing pairwise rms differences between individual structures: in the present calculations we obtain 0.93 and 1.80 Å for backbone atoms and all heavy atoms, respectively, whereas the corresponding numbers were 2.0 and 3.1 Å in the set of previously determined structures calculated using a similar refinement protocol. The higher precision of the present structure is due to a larger number of experimental constraints: 906 NOE distance constraints and 43 dihedral angle constraints were used in the present calculations, compared to the 470 NOE distance constraints used in the previous structure determination. A larger number of constraints could in this case be collected because we studied a smaller protein fragment, which was ^{15}N-labeled in order to extend the number of observable NOE connectivities in the crowded amide proton region of the spectrum using 3-dimensional heteronuclear spectroscopy. The high and almost uniform precision of the present structures is in accordance with our previous observation that the picosecond mobility is limited and uniform along the GR DBD backbone (Berglund et al., 1992), which is what is normally observed in protein regions for which well-defined NMR structures can be determined (see, for instance, Redfield et al., 1992, and references cited therein).

The overall peptide folding of the present GR DBD structure is very similar to the previously determined NMR structure when the conformations of the two longer α-helices, the hydrophobic core, and parts of the first zinc-coordinating domain are compared. The largest improvement in the present refinement is found within the second zinc-coordinating domain, i.e., residues C476–C495. This fragment was not very well-defined in the previous NMR structure (pairwise backbone rms difference of 2.7 ± 0.6 Å), but it is now much better resolved (backbone rms difference of 0.86 ± 0.37 Å). A significant improvement in precision is also observed for residues S448–Y452 which were disordered in the previous NMR structure. The structure of this region is now more well defined, due to an extended set of NOE distance constraints in this region.

The extended region connecting the two zinc-coordinating domains, i.e., residues H471–L475, is still not very well-defined in the present structure, possibly because of the missing assignment for the amide proton of N473 in combination with low NOE intensities for the other amide proton signals in the region which makes it difficult to observe medium- and long-range NOE's.

Structural Changes in DBD upon Binding to DNA. The structural definition of the second zinc-coordinating domain was previously (Härd et al., 1990b) unsatisfactory with regard to conclusions about structural changes that might occur as a consequence of dimerization upon DNA binding, as discussed in the Introduction. The present refinement makes it possible to analyze these changes in detail. The short distorted helical region involving residues K486–N491, which could not be observed in the previous NMR structure (Härd et al., 1990a,b), is now clearly identified as a result of an extended set of sequential and long-range NOE connectivities. Local conformations of the residues within this helix as well as its overall orientation within the molecule are very similar to the $DBD_{X\text{-}ray}$ molecule (Luisi et al., 1991). Thus, the distorted helix is not formed as a consequence of dimerization to two DBD molecules in the DNA complex. The "D-box" region including residues A477–D481, on the other hand, seems to undergo a conformational change upon DNA binding. This fragment, which also forms part of the dimerization surface in the dimeric DBD–DNA complex, is not as well-defined as other parts of GR DBD in the present refinement. However, the peptide backbone of A477–D481 in DBD_{NMR} does not overlap with the corresponding $DBD_{X\text{-}ray}$ residues in any of our 24 calculated structures when superimposing the full proteins or their second zinc-coordinating domains only. A conformational change within the D-box upon dimerization is not unreasonable because it forms a surface loop which seems loosely anchored to other parts of the protein in the uncomplexed state of the DBD.

The other regions that are involved in intermolecular interactions in the dimeric DBD–DNA complex is the helix encompassing G458–E469, which interacts with the DNA major groove, and residues C450–Y452 of the first zinc-coordinating domain, which form several contacts with the DNA phosphate backbone (Luisi et al, 1991). We observe no structural differences between the backbone of the recognition helix in the free and complexed states. The backbone of the first zinc-coordinating domain differs only slightly (1.4 Å) between the free and complexed states of the DBD, but the difference is still larger than that observed within the DBD core (0.30 Å). Most of this difference can be attributed to the different conformation of the N-terminal, and the (more minor) differences in the conformation in other parts of the first zinc-coordinating domain might not be significant, although it would not be surprising if the observed (small) differences in the fragment G449–G453 were a

consequence of the DNA contacts mentioned above.

The origin of the different conformation of the N-terminal cysteine residue (C440) in the DBD_{NMR} and the $DBD_{X\text{-}ray}$ structures is not obvious, but it is possible that it is a consequence of different N-terminal sequences of the two recombinant proteins. The protein fragment used for crystallization contains six "non-GR" residues at the N-terminal with a proline residue preceding C440 whereas our DBD82 fragment contains a leucine residue at position 439, as in the native GR.

None of the two hydrogen bonding networks described by Luisi et al. (1991), involving residues H451–S448–G458 and R489–(D445 and S459), respectively, in the sequence-specific DBD–DNA complex can be detected in the solution structure of DBD. The networks seem to stabilize the DNA phosphate contacts made by H451 and R489. The involved side chains are in both cases located on the protein surface and are not very well-defined in our DBD structures, as is often the case with surface side chains in protein structures determined with NMR (Wagner et al., 1992). If strong hydrogen bonds were present, we would expect to observe several NOE connectivities which would yield more well-defined side-chain conformations. Thus, it is plausible that the two hydrogen bond networks are stable only when DBD is bound to DNA.

CONCLUSIONS

We have determined a high-resolution solution structure of the GR DBD using NMR spectroscopy. The structure contains all elements of secondary structure found in the crystal structure of the GR DBD in complex with a specific DNA binding site. Positions of atoms in the core of the free protein are very similar to those found in the DNA complex, but small differences can be observed in the two zinc-coordinating domains. The largest differences are found for a short surface loop (the D-box region) within the second zinc-coordinating domain and for the very N-terminal residues of the protein. The different conformations of the D-box loop can be rationalized in terms of a conformational change induced by dimerization interactions upon DNA binding.

ACKNOWLEDGMENT

Helmi Siltala and Dr. Karin Dahlman-Wright are acknowledged for assistance with protein purification. We thank Prof. Paul Sigler for providing coordinates for the crystal structure of the GR DBD–DNA complex. Dr. Per Kraulis is acknowledged for fruitful discussions.

SUPPLEMENTARY MATERIAL AVAILABLE

Assigned ^1H and ^{15}N chemical shifts for the residues C440–R510 in GR DBD (the DBD82 fragment) at pH 6.0, 299 K and DGII constraint list (13 pages). Ordering information is given on any current masthead page.

REFERENCES

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

Berglund, H., Dahlman-Wright, K., Kovacs, H., Gustafsson, J.-Å., & Härd, T. (1992) *Biochemistry 31*, 12001–12011.

Billeter, M., Neri, D., Otting, G., Qian, Y. Q., & Wüthrich, K. (1992) *J. Biomol. NMR 2*, 257–274.

Brocklehurst, S. M., & Perham, R. N. (1993) *Protein Science 2*, 626–639.

Brooks, B. R., Bruccoleri, R. E., Olafsen, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.

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

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

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

Gronenborn, A. M., Bax, A., Wingfield, P. T., & Clore, G. M. (1989) *FEBS Lett. 243*, 93–98.

Härd, T., Kellenbach, E., Boelens, R., Kaptein, R., Dahlman, K., Carlstedt-Duke, J., Freedman, L. P., Maler, B. A., Hyde, E. I., Gustafsson, J.-Å., & Yamamoto, K. R. (1990a) *Biochemistry 29*, 9015–9023.

Härd, T., Kellenbach, E., Boelens, R., Maler, B. A., Dahlman, K., Freedman, L. P., Carlstedt-Duke, J., Yamamoto, K. R., Gustafsson, J.-Å., & Kaptein, R. (1990b) *Science 249*, 157–160.

Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Mathem. Biology 45*, 665–720.

Havel, T. F. (1991) *Progr. Biophys. Molec. Biol. 56*, 43–78.

Hyberts, S. G., Märki, W., & Wagner, G. (1987) *Eur. J. Biochem. 164*, 625–635.

Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736–751.

Kaptein, R. (1991) *Curr. Opin. Struct. Biol. 1*, 63–70.

Katahira, M., Knegtel, R. M. A., Boelens, R., Eib, D., Schilthius, J. G., van der Saag, P. T., & Kaptein, R. (1992) *Biochemistry 31*, 6474–6480.

Knegtel, R. M. A., Katahira, M., Schilthius, J. G., Bonvin, A. M. J. J., Boelens, R., Eib, D., van der Saag, P. T., & Kaptein, R. (1993) *J. Biomol. NMR 3*, 1-17.

Koning, T. M. G., Boelens, R., & Kaptein, R. (1990) *J. Magn. Reson. 90*, 111–123.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.

Levy, G. C., & Lichter, R. L. (1979) in *Nitrogen-15 Nuclear Magnetic Resonance Spectroscopy*, p 59, John Wiley, New York.

Ludvigsen, S., & Poulsen, F. M. (1992) *J. Biol. NMR 2*, 227–233.

Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R., & Sigler, P. B. (1991) *Nature 352*, 497–505.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.

Marion, D., & Würthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989) *Biochemistry 28*, 6150–6156.

Montelione, G. T., Winkler, M. E., Rauenbuehler, P., & Wagner, G. (1989) *J. Magn. Reson. 82*, 198–204.

Müller, L. (1979) *J. Am. Chem. Soc. 101*, 4481–4484.

Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741–751.

Pauling, L. (1967) in *The Chemical Bond*, Cornell Univ. Press, Ithaca, NY.

Plateau, P., & Guéron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.

Rance, M., Sorenson, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 69*, 979–987.

Redfield, C., Boyd, J., Smith, L. J., Smith, R. A. G., & Dobson, C. M. (1992) *Biochemistry 31*, 10431–10437.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.

Schwabe, J. W. R., Neuhaus, D., & Rhodes, D. (1990) *Nature 348*, 458–461.

Wagner, G., Hyberts, S. G., & Havel, T. F. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 167–198.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.