

- Talbot, J. A., & Hodges, R. S. (1982) *Acc. Chem. Res.* 15, 224.
 Wagner, G. (1983) *Q. Rev. Biophys.* 16, 1.
 Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol.* 160, 343.
 Wagner, G., & Wüthrich, K. (1986) *Methods Enzymol.* 131, 307.

- Wagner, G., Pardi, A., & Wüthrich, K. (1983) *J. Am. Chem. Soc.* 105, 5948.
 Woodward, C., Simon, I., & Tüchsen, E. (1982) *Mol. Cell. Biochem.* 48, 135.
 Yesinowski, J. P., Eckert, H., & Rossman, G. R. (1988) *J. Am. Chem. Soc.* 110, 1367.

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

M. L. Remerowski,[‡] E. Kellenbach,[‡] R. Boelens,[‡] G. A. van der Marel,[§] J. H. van Boom,[§] B. A. Maler,^{||} K. R. Yamamoto,^{||} and R. Kaptein^{*,‡}

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

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

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

The glucocorticoid receptor (GR)¹ is a member of the steroid/thyroid hormone receptor superfamily (Danielsen, 1991; Muller & Renkawitz, 1991), which includes receptors for steroid hormones, thyroid hormones, retinoic acid, and vitamin D₃. These receptors act as hormone-regulated transcription factors and display a common functional organization, with a hormone-binding part at the carboxy terminus and an adjacent domain consisting of about 70 residues responsible for DNA binding (Carlstedt-Duke et al., 1987; Rusconi & Yamamoto, 1987).

This DNA binding region has a high degree of sequence homology within the hormone receptor superfamily (Evans, 1988; Beato, 1989). Protein fragments containing the DNA binding domain (DBD) of the glucocorticoid receptor (GR) expressed in *Escherichia coli* exhibit sequence-specific binding to glucocorticoid response elements (GREs) (Freedman et al., 1988a; Tsai et al., 1988; Dahlman et al., 1989). The structured

region corresponding to C440–R510 of the rat glucocorticoid receptor (Figure 1) was shown to contain two “zinc fingers”, i.e., peptide regions in which two Zn(II) ions are tetrahedrally coordinated by eight of nine cysteine residues (Freedman et al., 1988a). These metal ions are required for proper folding and DNA binding activity (Freedman et al., 1988b) but have been shown by NMR to have a conformation which is distinct from other classes of zinc-finger proteins such as TFIIIA (Härd et al., 1990a,b; Kaptein, 1991). Residues C460–E469 and P493–G504 immediately following the fingers are α -helical and are oriented perpendicular to each other with the hydrophilic surfaces exposed to the solvent and the conserved hydrophobic residues forming the protein core (Figure 2). The structure of the estrogen receptor DNA binding domain determined by NMR (Schwabe et al., 1990) is similar.

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

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

[‡] University of Utrecht.

[§] University of Leiden.

^{||} University of California at San Francisco.

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

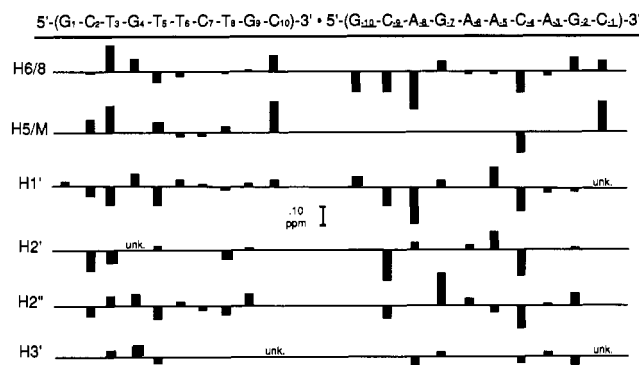


FIGURE 3: Chemical shift changes in a half-site GRE oligonucleotide occurring when GR-DBD is bound to the DNA. The unk. chemical shift changes are due to either missing or overlapping cross peaks in one or both spectra.

recorded with a 40-ms mixing time using the "clean TOCSY" (Griesinger et al., 1988) pulse sequence. Spectra were recorded with 350 or 400 t_1 increments and 2K data points per increment. All data processing was done as described before using the TRITON NMR software package developed at the Department of Chemistry, University of Utrecht. The FID signals were apodized using sine bell, squared sine bell, or Lorentz-Gauss window functions and processed to 1K \times 1K real data. Fourth-order polynomial baseline corrections were applied in both frequency domains.

RESULTS AND DISCUSSION

1D NMR spectra taken in the course of titration of the DNA with increasing protein showed a gradual appearance of new resonance lines as those of the free DNA decreased in intensity, consistent with a slow exchange between free and bound DNA on the NMR chemical shift time scale. The complex precipitates slowly over the course of the data acquisition.

Resonance lines of most protons in the complex were shifted, but it was possible to assign the resonances of all the base protons, H6 pyrimidine and H8 purine, and most of the deoxyribose sugar protons, H1', H2'/H2'', and H3', in the complex using a sequential assignment strategy (Scheek et al.,

Table I: NOEs Observed between GR-DBD 93 and Half-Site GRE

GR-DBD 93	DNA	GR-DBD 93	DNA
Tyr-452 H2,6	A ₈ H2'	Val-462 C _α H	T ₅ CH ₃
Tyr-452 H2,6	A ₈ H3'	Val-462 C _β H	T ₅ CH ₃
Tyr-452 C _β H	A ₈ H3'	Val-462 C _{γ1} H ₃	T ₅ CH ₃
Tyr-452 C _β H	A ₈ H1'	Val-462 C _{γ2} H ₃	T ₅ CH ₃
Tyr-452 C _β H	A ₈ H8	Val-462 C _{γ1} H ₃	G ₄ H8
		Val-462 C _{γ2} H ₃	G ₄ H8

1983; Hare et al., 1983). The pattern of NOE cross peaks and their relative intensities indicated that the oligonucleotide was in a B-DNA conformation.

The chemical shift differences between free and complexed oligonucleotide are shown in Figure 3. From the model in Figure 2, nucleotides T₃, G₄, T₅, A₈, G₇, and A₆ are closest to the protein and thus would be expected to have the greatest chemical shift perturbation. Figure 3 shows that this is generally true. But, chemical shift differences of similar magnitude were also found for the cytosine (C₄) and the preceding adenine (A₅) in the "A-strand" oligonucleotide. The A₅-C₄ are base paired to nucleotides G₄-T₅ which also displayed chemical shift perturbations and NOEs to protein residues. Therefore, we interpret the chemical shift changes of A₅-C₄ as indirect effects from the amino acid residues close to G₄-T₅.

The assignment of the protein resonances in the complex is more complicated than those of the DNA. Some pairs could be assigned by comparison of 2D spectra of the complex with that of the assigned free GR-DBD by Härd et al. (1990b). The assignment of the tyrosine residues in the complex was done using photo-CIDNP measurements (Kellenbach et al., 1991), but chemical shift differences between the free and bound form of the protein did not allow unambiguous assignments of all residues using only D₂O spectra. Thus far we have identified two amino acid residues, Tyr-452 and Val-462, with unambiguous NOEs to the DNA (cf. Figures 4 and 5). These residues could be identified using the patterns of cross peaks, both intra- and interresidue, in both TOCSY and NOESY spectra.

The summary of the observed protein-DNA NOEs (cf. Table I) contains contacts involving amino acid residues, Tyr-452 and Val-462, and three nucleotides, G₄, T₅, and A₈.

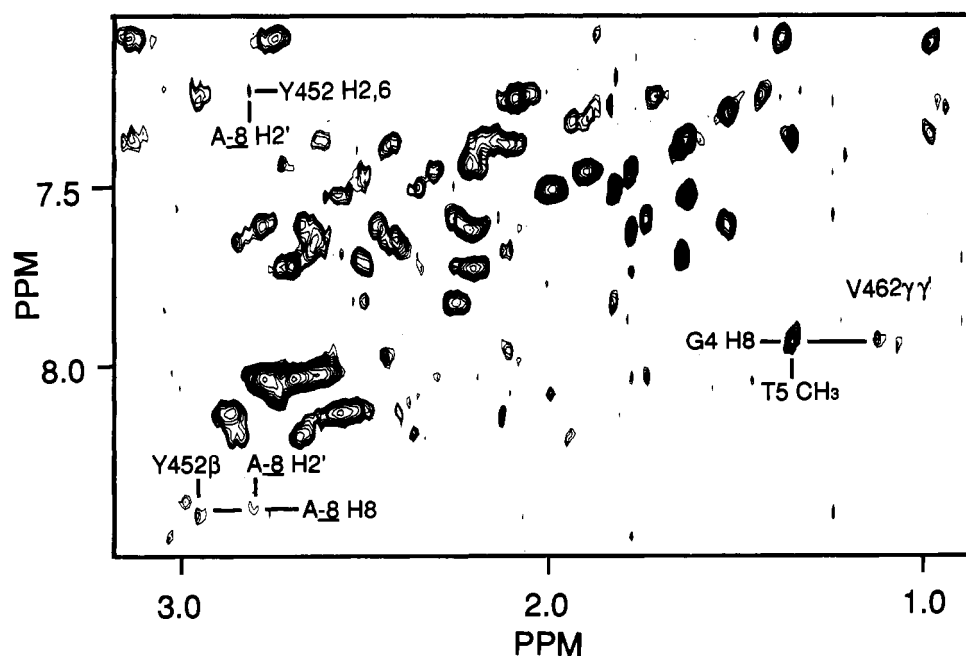


FIGURE 4: Region of a NOESY spectrum showing protein-DNA NOEs between G₄ and Val-462 and between A₈ and Tyr-452. The NOESY spectrum was recorded with a 120-ms mixing time.

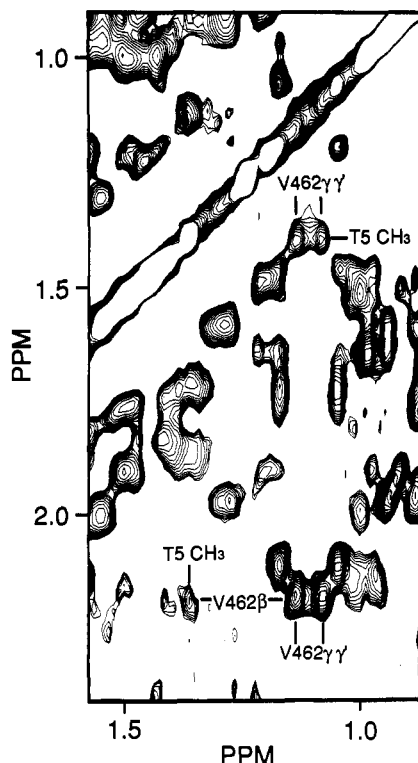


FIGURE 5: Region of a NOESY spectrum showing protein-DNA NOEs between T₅ and Val-462. The NOESY spectrum was recorded with a 120-ms mixing time.

Most of the NOEs are shown in Figures 4 and 5. The existence of several NOEs between atoms of each amino acid-nucleotide pair is self-consistent. From these observations it is possible to draw several conclusions: (i) The DNA protons contacted by the protein, and the relative intensities of those NOEs, position the recognition helix (C460-E469) in the major groove of the DNA helix. (ii) The protein-DNA contacts occur at opposite ends of the half-site, orienting the GR-DBD relative to the DNA as proposed by the model (Figure 2). This orientation places the D-loop (A477-D481) at the end of the half-site, which makes it available for protein-protein interaction proposed for dimerized GR-DBD bound to an intact consensus GRE. (iii) A specific interaction between Val-462 and the 5-methyl group of the T₅ nucleotide in the protein-DNA binding is identified. (iv) Tyr-452 is implicated in the protein-DNA interaction, a result that is in apparent disagreement to the photo-CIDNP results reported by Kellenbach et al. (1991).

The 5-methyl of T₅ which is important for DNA binding of GR (Truss et al., 1990) is occupied by an adenine in the consensus ERE, which suggests that protein-DNA interaction at that position would be with an amino acid residue of the binding domain which differentiates the specificity of the two binding domains. Three residues, Gly-458, Ser-459, and Val-462 (shown boxed in Figure 1), have been implicated to account for the differentiation. Our results identify an interaction between the T₅ methyl group and Val-462, one of those amino acids. No protein-DNA interaction sites have been found for Gly-458 or Ser-459 to date, but we have been unable to establish the chemical shifts of any of those protons due to the severe overlap of intraprotein and intra-DNA cross peaks in that region of the spectrum. It is interesting to note, however, that in the X-ray structure of the specific GR-DBD DNA complex only a possible role for Val-462 could be established. No clear interactions were found for Gly-458 and Ser-459.

Protein-DNA interactions also occur between the aromatic and C_β-protons of Tyr-452 and the base and ribose moieties of A₈. Previously, we performed photo-CIDNP experiments on the GR DBD-GRE complex with the result that Tyr-474 polarization quenched after addition of DNA to the protein, while Tyr-452 polarization was unaffected. From these CIDNP experiments it was concluded that Tyr-474 was involved in protein-DNA interaction, while Tyr-452 was not. An explanation for the CIDNP results compatible with the present study may be that Tyr-452 interacts with the DNA, but its hydroxyl group is still accessible to the photoexcited dye. The quenched polarization of Tyr-474 may be a consequence of interaction of the hydroxyl with the DNA in a more shielded environment, while the aromatic ring protons are too far from the DNA for observable NOE effects.

These NMR results are compatible with the X-ray structure (Luisi et al., 1991) in so far as they can be compared. The X-ray structure is based on a complex between a symmetrized GRE containing four spacer base pairs, which binds two molecules of GR-DBD. One of the molecules is supposed to bind "specifically" and the other binds "nonspecifically". Our results as well as the X-ray structure of the specifically bound molecule indicate a hydrophobic interaction between the T₅ methyl group and Val-462. The X-ray structure shows hydrogen bonds between the phosphate backbone and both Tyr-452 and Tyr-474. For Tyr-452 this means either that the hydroxyl hydrogen bond seen in the X-ray structure does not exclude the hydroxyl from reaction with the photoexcited dye as seen in the CIDNP experiment or that the hydrogen bond in solution is present only part of the time. The general features, i.e., orientation and placement of the recognition helix in the major groove, appear the same in both the X-ray and NMR results.

In conclusion, a complex between GR-DBD and a half-site GRE has been made which is suitable for NMR analysis. NOESY spectra of the complex show several protein-DNA interaction sites which are sufficient to confirm the principal features of the model suggested by Härd et al. (1990) and the structure determined by X-ray crystallography (Luisi et al., 1991). It is interesting to note that, in spite of the differences between the helix-turn-helix DNA binding proteins and the hormone receptor DNA binding domains, both have a recognition helix which lies in the major groove of DNA with a second helix lying on top and oriented perpendicularly to it. Other residues on the surface of the recognition helix are prime candidates for further protein-DNA interactions, which could be identified using 2D NMR and 3D heteronuclear NMR spectra in water.

ACKNOWLEDGMENTS

We thank Dr. Delia Picone for her help in the preparation of the complex sample.

REFERENCES

- Bax, A., & Davis, D. G. (1985) *J. Magn. Reson.* 65, 355-360.
- Beato, M. (1989) *Cell* 56, 335-344.
- Berg, J. M. (1989) *Cell* 57, 1065-1068.
- Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson.* 53, 521-528.
- Carlstedt-Duke, J., Strömstedt, P.-E., Wrangé, Ö., Bergman, T., Gustafsson, J.-Å., & Jörnvall, H. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 4437-4440.
- 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.

- Dahlman-Wright, K., Siltala-Roos, H., Carlstedt-Duke, J., & Gustafsson, J.-Å. (1990) *J. Biol. Chem.* 265, 14030–14035.
- Dahlman-Wright, K., Wright, A., Gustafsson, J.-Å., & Carlstedt-Duke, J. (1991) *J. Biol. Chem.* 266, 3107–3122.
- Danielsen, M. (1991) in *Nuclear Hormone Receptors* (Parker, M. G., Ed.) pp 39–68, Academic Press, London.
- Danielsen, M., Hinck, L., & Ringold, G. M. (1989) *Cell* 57, 1131–1138.
- Drobny, G. P., Pines, A., Sinton, S., Weitekamp, D. P., & Wemmer, D. E. (1979) *Symp. Faraday Soc.* 13, 49–55.
- 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.
- Glass, C. K., Holloway, J. M., Devary, O. V., & Rosenfeld, M. G. (1988) *Cell* 54, 313–323.
- Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc.* 110, 7870–7872.
- Ham, J., Thomson, A., Needham, M., Webb, P., & Parker, M. (1988) *Nucleic Acids Res.* 16, 5263–5276.
- Hård, T., Kellenbach, E., Boelens, R., Maler, B. A., Dahlman, K., Freedman, L. P., Carlstedt-Duke, J., Yamamoto, K. R., Gustafsson, J.-Å., & Kaptein, R. (1990a) *Science* 249, 157–160.
- Hård, 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. (1990b) *Biochemistry* 29, 9015–9023.
- Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G. P., & Reid, B. B. (1983) *J. Mol. Biol.* 171, 319–336.
- Ivarie, R. (1987) *Nucleic Acids Res.* 15, 9975–9983.
- Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys.* 71, 4546–4553.
- Kaptein, R. (1991) *Curr. Opin. Struct. Biol.* 1, 63–70.
- Kellenbach, E., Hård, T., Boelens, R., Dahlman, K., Carlstedt-Duke, J., Gustafsson, J.-Å., van der Marel, G. A., van Boom, J. H., Maler, B., Yamamoto, K. R., & Kaptein, R. (1991) *J. Biomol. NMR* 1, 105–110.
- Klein-Hitpass, L., Schorpp, M., Wagner, U., & Ryffel, G. U. (1986) *Cell* 46, 1053–1061.
- Klock, G., Strähle, U., & Schütz, G. (1987) *Nature* 329, 734–736.
- Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R., & Sigler, P. B. (1991) *Nature* 352, 497–505.
- Mader, S., Kumar, V., de Verneuil, H., & Chambon, P. (1989) *Nature* 338, 271–274.
- Martinez, E., Givel, F., & Wahli, W. (1987) *EMBO J.* 6, 3719–3727.
- Muller, M., & Renkawitz, R. (1991) *Biochim. Biophys. Acta* 1088, 171–182.
- Rusconi, S., & Yamamoto, K. R. (1987) *EMBO J.* 6, 1309–1315.
- Scheek, R. M., Russo, N., Boelens, R., & Kaptein, R. (1983) *J. Am. Chem. Soc.* 105, 2914–2916.
- Schwabe, J. W. R., Neuhaus, D., & Rhodes, D. (1990) *Nature* 348, 458–461.
- Seiler-Tuyns, A., Walker, P., Martinez, E., Mérellat, A.-M., Givel, F., & Whali, W. (1986) *Nucleic Acids Res.* 14, 8755–8770.
- States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson.* 48, 286–297.
- Steitz, T. A. (1990) *Q. Rev. Biophys.* 23, 205–280.
- Strähle, U., Klock, G., & Schütz, G. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 7871–7875.
- Truss, M., Chalepakidis, G., & Beato, M. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 7180–7184.
- Tsai, S. Y., Carlstedt-Duke, J., Weigel, N. L., Dahlman, K., Gustafsson, J.-Å., Tsai, M. J., & O'Malley, B. W. (1988) *Cell* 55, 361–369.
- Umesono, K., & Evans, R. M. (1989) *Cell* 57, 1139–1146.
- van der Marel, G. A., van Boeckel, C. A. A., Wille, G., & van Boom, J. H. (1981) *Tetrahedron Lett.* 22, 3887–3891.
- Zilliacus, J., Dahlman-Wright, K., Wright, A., Gustafsson, J.-Å., & Carlstedt-Duke, J. (1991) *J. Biol. Chem.* 266, 3101–3106.