AD_____

GRANT NUMBER DAMD17-94-J-4420

TITLE: Crystallization and Structure Determination of the Human Estrogen Receptor by X-Ray Diffraction

PRINCIPAL INVESTIGATOR: Stephen C. Harrison, Ph.D.

CONTRACTING ORGANIZATION: Children's Hospital
Boston, Massachusetts 02115

REPORT DATE: November 1997

TYPE OF REPORT: Final

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	November 1997	Final (15 Sep 94 - 14 Oct 97)

4. TITLE AND SUBTITLE
Crystallization and Structure Determination of the Human Estrogen Receptor by X-Ray Diffraction

5. FUNDING NUMBERS
DAMD17-94-J-4420

6. AUTHOR(S)
Stephen C. Harrison, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Children's Hospital
Boston, Massachusetts 02115

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, MD 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

The project is a postdoctoral training fellowship for Dr. Robert Nolte, to support X-ray crystallographic studies of cancer-related macromolecular complexes. The crystal structure of the six NH2-terminal zinc fingers of Xenopus laevis transcription factor IIIA (TFIIIA) bound with 31 base-pairs of the 5S rRNA gene promoter has been determined at 3.1 Å resolution. Individual zinc fingers are positioned differently in the major groove and across the minor groove of DNA to span the entire length of the duplex. These results show how TFIIIA can recognize several separated DNA sequences using fewer fingers than necessary for continuous winding in the major groove. This structure reveals significant aspects of DNA recognition by Zn-finger proteins, which regulate many loci important for cell growth.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Estrogen, Receptor, Crystal, Structure, Diffraction, Steroid, Hormone, Breast Cancer	33
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

NA Where copyrighted material is quoted, permission has been obtained to use such material.

NA Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

NA Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

NA In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

NA For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

NA In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

NA In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

NA In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

PI - Signature 10/13/97 Date

TES

TABLE OF CONTENTS

Abstract

The project is a postdoctoral training fellowship for Dr. Robert Nolte, to support X-ray crystallographic studies of cancer-related macromolecular complexes. The crystal structure of the six NH2-terminal zinc fingers of Xenopus laevis transcription factor IIIA (TFIIIA) bound with 31 base-pairs of the 5S rRNA gene promoter has been determined at 3.1 Å resolution. Individual zinc fingers are positioned differently in the major groove and across the minor groove of DNA to span the entire length of the duplex. These results show how TFIIIA can recognize several separated DNA sequences using fewer fingers than necessary for continuous winding in the major groove. This structure reveals significant aspects of DNA recognition by Zn-finger proteins, which regulate many loci important for cell growth.

Introduction

The goal of this project is postdoctoral training for Dr. Robert Nolte in the area of structural biology as applied to problems in cancer and to breast cancer in particular. It was originally proposed to crystallize the hormone binding domain (HBD) of the estrogen receptor, in order to determine its structure and the implications of structure for the mode of hormone recognition. Because other groups (Bourquet 1995, Renaud 1995, Wagner 1995, and Brzozowski 1997) have now achieved similar goals, we submitted a revised SOW which shift emphasis to related projects in the structural biology of cancer, in order to ensure rapid progress in Dr. Nolte's training, especially in crystallographic structure determination.

The Problem: Cancer involves anomalies in cellular regulation. Two key regulatory steps are the transduction of signals from the cell surface and the control of specific gene expression. The estrogen receptor, on which the original SOW for this grant focused, participates in both these steps. It receives hormonal signals and in response activates defined genes. The recently discovered BCRA1 gene appears to encode a nuclear protein that may regulate gene expression (Miki , 1994). Progress in structural studies of tyrosine-kinase mediated signaling pathways and protein/DNA complexes that regulate gene expression makes the possibility of structure-based drug discovery and development a real one, but many basic principles remain to be discovered.

Previous work. A number of structural studies of protein/DNA complexes involved in transcriptional regulation have provided the guidelines of how to think about specific recognition (see, for example, Harrison, 1991 and Steitz, 1990). An important notion to emerge both from the structural studies and from studies of transcriptional regulation at many complex promoter/enhancer sites is that large multi-protein complexes bound to DNA are critical for the sort of combination control seen in cells of all higher eukaryotes (and in human cells in particular). Work in our laboratory on the polymerase III transcription factor known as TFIIIA has led to crystallization

of a complex between a part of the factor containing six "zinc fingers" (each finger is a distinct DNA-binding domain) and a 31 BP DNA site. This is a larger complex than any so far studied from this class of transcriptional regulatory proteins.

Purpose of present work. The broad goal is to enhance our understanding of specific protein-protein and protein-DNA interactions in gene regulation and in signal transduction. In particular, we are concluding the structure determination of a complex between a portion of the DNA-binding segment from the polymerase III regulatory factor TFIIIA and its cognate site.

Body of Narrative

Experimental Methods

The NH_2-terminal six zinc fingers of TFIIIA bound in a complex with 31 base-pairs of the 5S rRNA gene has been reconstituted and crystallized as follows. Recombinant TFIIIA (amino acid residues 1–190) was produced from plasmid pRSET B (Invitrogen) in *E. coli* BL21(DE3). Protein was purified on Bio-Rex 70 and heparin Sepharose columns in 7 M urea. Synthetic oligonucleotides were purified by MonoQ chromatography in 7 M urea. Thymines were replaced by 5-iodouracil (5-dIU) at specific positions; T73 and T76 (noncoding strand) and T88' or T93' (coding strand), for heavy atom derivatives. The protein–DNA complex was reconstituted by stepwise dilution (Conlin, 1994) from 0.75–0.25 M NaCl at 25°C. Crystals grew in hanging drops on silanized plastic coverslips from 165 mM NaCl, 35 mM sodium acetate, 3.2 mM dithiotreitol, 9.2% (v/v) glycerol, 1.8 mM NaN_3, 1.8 mM cadavarine-2HCl, 5.5 mM Tris-HCl, pH 8.0 and 22.5% PEG 4000 at 18°C. The complex crystallized in the space group P1 with unit cell parameters a = 64.2 Å, b = 64.7 Å, c =78.0 Å, *alpha*= 90.1°, *beta* = 93.0°, gamma =103.0°. Two complexes are present in the unit cell, which contains 72 percent solvent. Invariably doubled crystals were split under polarized light, and cryoprotectant was introduced in steps over 48 h, to reach the final conditions, which were mother liquor supplemented with 215 mM NaCl, 10 percent sucrose and 15% (v/v) glycerol. Crystals in nylon loops (10-0 suture) were frozen by plunging into liquid nitrogen. Data were collected at –160°C with a MAR image-plate detector at a wavelength of 1.283 Å (the K-edge of zinc) at the NSLS beamline X12B at Brookhaven. Ice rings were deleted from the MAR images and intensities were integrated and merged with the DENZO/SCALEPACK package (Otwinowski, 1997). The structure was determined at low resolution by multiple isomorphous replacement (MIR). A difference Patterson synthesis calculated with the FFT program was searched with a model of the six iodine atoms derived from B-form DNA. The correct constellation of Patterson peaks has noncrystallographic symmetry – a

twofold rotation around and a translation along a direction parallel to the b-axis. 38,000 potential solutions were evaluated with VECREF, and MIR phases were calculated by MLPHARE (Collaborative Computing Project, Number 4, 1994). Zinc anomalous scattering data were derived from four native crystals. |F+ – F–|hkl anomalous differences > 30 percent of |F+ + F–|hkl / 2 were rejected before calculating Fourier maps. Phases were extended from 6 to 4.5 Å by positioning base-pairs into a MIR map averaged with RAVE (Kleywegt, 1993). Homologous zinc fingers taken from known crystal structures (Pavletich, 1991, 1993, Fairall, 1993, Elrod-Erickson, 1996, Houbavity, 1996) were positioned into appropriate electron density and refined using the real-space rigid-body procedure in the O program (Jones, 1993). Combined phases from MIR and the partial model were generated by SIGMAA (CCP4, 1994) as the starting point for several cycles of mask refinement, averaging, solvent flattening, and rigid-body refinement. Phases were extended to 3.1 Å resolution by an iterative procedure involving (a) twofold averaging, solvent flattening and histogram matching with the DM program (Cowtan, 1994); (b) model rebuilding with program O, using a custom zinc finger Lego library, and (c) positional refinement into anisotropically-scaled, B-factor sharpened data with X-PLOR (Brunger, 1996).

Background

TFIIIA is an essential component of the RNA polymerase III (Pol III) transcription initiation complex for 5S rRNA in *X. laevis* oocytes (Shastry, 1991, 1993, 1996, Hanas, 1992, Pieler, 1993). TFIIIA also participates in the nuclear export (Friedell, 1996) and storage of 5S rRNA, with which it forms a stable cytoplasmic 7S particle (Picard, 1979). The DNA-binding site for a single TFIIIA protein extends over 60 base-pairs of the 5S rRNA gene promoter Engelke, 1980, Pelham, 1980). This site lies within the 5S rRNA coding sequence itself. It is effectively a tripartite promoter (Pieler, 1985) containing separated "box A", "intermediate" (IE) and "box C" sequences (Fig. 1A). Similar regulatory elements exist in tRNA gene promoters. Mapping the

details of this extensive protein-DNA interaction using chemical, biochemical and genetic techniques has continued for almost 20 years (Shastry, 1991). The discovery of nine zinc fingers in TFIIIA (Brown, 1985, Miller, 1985) led to the notion of a transcription factor with repeated modules in its DNA-binding domain (Fig. 1B).

Our present knowledge of how zinc fingers bind specifically to DNA comes from several X-ray structures (Pavletich, 1991, 1993, Fairall, 1993, Elrod-Erickson, 1996, Houbaviy, 1996, Kim, 1996). In all of these protein–DNA complexes, there are contiguous zinc finger interactions with base-pairs in the major groove. In Zif268, for example, three fingers recognize successive, overlapping base-pair quartets in the major groove, covering a total of 11 base-pairs. In the DNA complex of a five-finger segment from Gli, the first finger lies outside the major groove and makes no DNA contacts, but its function in the intact protein is not known. The remaining fingers wrap in the major groove rather like those of Zif268. An extension of this mode of binding is not sufficient to explain the size of the TFIIIA-binding site, however.

Results and Discussion

The crystal structure shows that the six-finger protein stretches along the entire length of the 31 base-pair duplex. The current protein model includes amino acids 10–188 of TFIIIA (Fig. 1B). Residues 1–9, 161 and 189–190 are disordered in the crystal. In the complex, fingers 1–2–3 adopt a completely different configuration than do fingers 4–5–6 (Fig. 2A). Fingers 1–2–3, which are separated by typical linker sequences wrap smoothly around the major groove of DNA rather like those of Zif268. Contacts are made with DNA bases mainly on the noncoding strand of the 5S rRNA gene. In contrast, fingers 4–5–6, which run along one side of the DNA double helix, form an open, extended structure. Of these, only finger 5 makes contacts with bases in the major groove. The two flanking fingers 4 and 6 straddle the neighboring minor grooves and appear to serve primarily as spacer elements in DNA

recognition. In this way the six-finger protein binds in a precise manner to the separated IE and box C sequences (Pieler, 1985).

The DNA is essentially B-form with a mean helical twist of 34.3° and a rise per base-pair of 3.33 Å. Its sequence corresponds to base-pairs 63 to 92 of the intragenic control region (ICR) of the 5S rRNA gene. Terminal 5'-overhanging bases are involved in normal Watson-Crick base-pairs with neighboring duplexes so as to form continuous columns of DNA in the crystal lattice. Analysis of the of the 60 phosphates and 62 bases present in the electron density which make up the double helix was carried out with the program CURVES. The local DNA bend angles are 16.7°, 24.4° and 18.4°.shows that there are three localized bends in the DNA at base-pairs 70, 85 and 90. Fingers 5, 2, and 1 interact with these positions respectively. Furthermore as a result of finger binding there are increases in the depth and width of the major groove in agreement with results from other X-ray structures.

Each TFIIIA finger is folded in the classical way (Berg, 1988, Lee, 1989) around a Zn(II) ion, including finger 6, which lacks some of the conserved amino acid residues. The positions of the six Zn(II) ions were determined independently of the protein structure from an anomalous difference Fourier synthesis. These metal sites are also present in the electron density and indicate the correct path and fold of the polypeptide chain (Fig. 2B).

The consensus pentapeptide linker sequence, Thr-Gly-Glu-Lys-Pro, frequently associated with major-groove binding fingers, appears only twice in this NH_2-terminal segment of TFIIIA. As expected, these linkers, 1–2 and 2–3, do indeed connect fingers that interact with bases in the major groove. The remaining linkers, 3–4, 4–5 and 5–6, have different structures and sequences, which permit the extended configuration for fingers 4, 5 and 6. Linkers 3–4 and 4–5 fold in ways that bring several hydrophobic amino acids into proximity. Five residues of linker 3–4 — Ile^{100}, Ile^{102}, Cys^{103}, Val^{104} and Val^{106} — form a hydrophobic cluster. Likewise four residues of linker 4–

5 — Val^{124}, Phe^{127}, Pro^{134} and Tyr^{135} — come together in another, and Phe^{127} also makes van-der-Waals contacts with Pro^{134}.

For the most part, the alpha helices of fingers 2, 3 and 5 interact with DNA as in previously analyzed structures, with side-chains contacting at least two of four consecutive base-pairs in the major groove (Fig. 3). Base-pair quartets that interact with adjacent fingers overlap by one, and backbone contacts from successive fingers overlap even more extensively. In the previously analysed complexes, most of the major-groove contacts occur between three bases on one strand and amino acid side-chains at alpha helix positions +6,+3, and -1, and the opposite-strand base of the fourth base-pair in the quartet may contact the side-chain of alpha helix position +2 (see shaded bases in Fig. 3E to 3H). In our structure, the "canonical" +6 and +3 contacts are made by fingers 2, 3 and 5, which also have a commonly found His (+7) – phosphate interaction. The +2 and -1 contacts are less standard, and in finger 3, the site is extended by an arginine – guanine interaction from the +10 position of the alpha helix (Figs. 3C and 3G).

Finger 1 is oriented similarly to fingers 2, 3 and 5 with respect to the groove, but it is displaced by over 4 Å toward the COOH-terminus of its recognition helix (Figs. 3A and 3E). As a result, the N1 of Trp^{28} at position +2 in the helix lies opposite the O6 of G89, and Lys^{29} (+3), which would normally contact this guanine, forms a salt bridge with phosphate 88 instead. Lys^{26} (-1) interacts with the opposite-strand guanine "vacated" by the +2 tryptophan. The shifted position of finger 1 requires the small Ala^{32} side-chain at position +6 to avoid steric interference, and it places Tyr^{24} rather than His^{33} within hydrogen-bond distance of phosphate 87.

Together fingers 1–2–3 bind within an 11 base-pair region located between positions 81 and 91 of the ICR, specifying the DNA sequence GGANGGNNGNN (noncoding strand) and NNNNCCNNNNG (coding strand). The structure agrees well with the earlier identification of the Pol III promoter element box C that was derived from site-directed mutagenesis of

the 5S rRNA gene (Pieler, 1985). The local details of finger conformation and DNA contacts, seen in a recent nuclear magnetic resonance (NMR) structure of fingers 1–2–3 bound to 15 base-pairs of DNA (11), closely match our X-ray structure. The relative orientation and position of finger 1 with respect to fingers 2 and 3 are somewhat different, however, so that in the NMR structure finger 1 and the DNA segment to which it binds lie closer to finger 2, than in the X-ray structure.

From a structural standpoint there is nothing different about fingers 4 and 6. They do not wrap around the double helix but instead traverse the minor groove. As spacers they increase the range of the TFIIIA protein, making possible a more economical use of fingers in binding to the separate promoter elements, IE and box C. In spanning the minor groove a few contacts are made with the DNA backbone. Gln^{121} and conserved Tyr^{135} of finger 4 both contact phosphate 75', and Lys^{175} of finger 6 contacts phosphate 68.

Finger 5 binds to bases in the major groove at the IE element, positions 70 to 73, seven base-pairs upstream of box C. A standard major-groove finger interaction is supplemented by contacts between Leu^{148} (-1) and the base of T74', Ser^{150} (+2) and phosphate 74', and Lys^{153} and phosphate 73'. The DNA sequence specified by finger 5 is GGNNN (noncoding strand) and NNNAT (coding strand), the consensus IE sequence (Pieler, 1985).

TFIIIA specificity is highly conserved. This is apparent at positions in the recognition helix that are involved in binding to DNA (Table 2) in eight aligned TFIIIA sequences (Ginsberg, 1984, Taylor, 1986, Gaskins, 1990, 1992, Archambault, 1992, Arakawa, 1995). Similarly the nucleotide bases that make contacts with fingers 1–2–3 and 5 are invariant with some exceptions in *I. puntatus* and *S. cerevisae* 5S rRNA genes (Wegnez, 1972, Korn, 1978, Gaskins, 1982, Forget, 1969, Olson, 1977, Maxwell, 1986). Finger 1 has almost no amino acid sequence variation. Substitutions are often conservative. Moreover, to the extent that Lys replaced by Arg may still specify guanine and

that Asn replaced with Gln may specify adenine, the substitutions do not affect DNA recognition (Pavletich, 1991, 1993, Fairall, 1993, Elrod-Erickson, 1996, Houbavity, 1996). In *S. cerevisae* TFIIIA the sequence identity is limited to the recognition helix of finger 2. Nevertheless the same methylation pattern of guanine residues in the ICR that interfere with binding of fingers 1–2–3 was also found (Rowland, 1996).

The characteristics of individual linkers are the same for various TFIIIA sequences (see Table 2). This conserved pattern points to a common structural organization for these proteins. In addition invariant helix residues and corresponding DNA bases suggests that the topology of fingers 1 to 6 in other TFIIIA–DNA complexes will be identical to our structure.

Fingers 7–8–9, not present in our structure, bind to the Pol III element box A. It has been proposed that these fingers wrap, like fingers 1–2–3, around the major groove of base-pairs 48 to 62 (Hayes, 1992, Clemens, 1992) based on results from DNA methylation protection and binding interference and on site-directed mutagenesis experiments (Fairall, 1986, Sakonju, 1982, Clemens, 1992, Pieler, 1985, McConkey, 1987, Veldhoen, 1994, Rawlings, 1996, Smith, 1991, Choo, 1993, Zang, 1995). In the model shown in Fig. 4, fingers 7–8–9 have been placed so that Arg^{271} at helix position +6 in finger 9 can recognize G51. Finger 6 can then connect to finger 7 with only a small displacement from its position in our crystal structure. We note that linkers 7–8 and 8–9 have amino acid sequences resembling the Thr-Gly-Glu-Lys-Pro consensus characteristic of sets of fingers that wrap around the major groove.

The X-ray structure of the TFIIIA–DNA complex shows how zinc fingers have been deployed to bind to separated promoter elements. Local folding of the protein orients fingers with respect to each other for a "custom fit" to the extended site. In this sort of design, some fingers will contact base-pairs and some will not. It is likely that other multifingered proteins will use a similar strategy to recognize regulatory elements in DNA. Bridging fingers may also serve additional functions in the multiprotein assemblies that activate transcription. Fingers 4–5–6 in this structure form a continuous,

platform-like surface, which could dock against other components of a Pol III transcription complex.

The structure shows that major groove insertion is not obligatory for zinc fingers. Three of the four fingers that do insert in the major groove align essentially in the manner previously seen in other complexes, but the fourth (finger 1) is displaced by about one base-pair and has idiosyncratic interactions. Recent efforts to design zinc finger proteins with desired DNA specificity have concentrated on the recognition helix (Greisman, 1997). Our structure further justifies the focus on a roughly standard orientation for this helix, but it also demonstrates the critical role of the linker sequences in determining the overall protein conformation. Mutagenesis and selection of linkers is likely to be particularly important in engineering the best fit to more complex DNA and RNA sites.

Conclusions:

The crystallographic study of a complex between (TFIIIA fingers 1-6) and cognate DNA has yielded significant new information about modes of specific DNA recognition by this family of transcription factors. Other work in our laboratory on gene regulation at complex promoter/enhancers will is now poised to capitalize on the results. Dr. Nolte having completed this training has launched an independent career participating in a drug development effort targeted at the nuclear receptor family of proteins.

NOTE: The text of this report is adapted from a draft manuscript, to be submitted for publication to *The National Academy of Sciences.*

Thus, the structural work reported here has contributed significant new information to an understanding of transcrriptional control elements, and it has also achieved its goal of training a young structural biologist to participate in structure-based drug development. Dr. Nolte's current work on nuclear receptors is likely to bear directly on breast cancer and related disorders.

References

Archambault, J., et al., *J. Biol. Chem.* **267**, 3282 (1992)

Arakawa, H., et al., *Cytogenet. Cell. Genet.* **70**, 235 (1995)

Arnold, S.F., Obourn, J.D., Jaffe, H., and Notides. AC., Molecular Endocrinology 9, 24-33 (1995)

Berg, J., *Proc. Natl. Acad. Sci. USA* **85**, 99 (1988)

Bourguet, W., Ruff, M., Chambon, P., Gronmeyer, H., and Moras, D.. Nature 375, 377-382 (1995)

Brinkmann, U. Mattes, R.E., and Buckel, P., Gene 85, 109-114 (1989)

Brown, R.S., Sander, C., and Argos, P., *FEBS Lett.* **186**, 271 (1985)

Brünger, A.T., *X-PLOR version 3.8, A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1996)

Brzozowski, A.M., Pike, A.C.W., Dauter, Z., Hubbard, R.E., Bonn, T., Engstrom, O., Ohman, L., Greene, G.L., Gustafsson, J-A., and Carlquist. M., Nature, 389, 753-758 (1997)

Carson, M., and Bugg, C.E., *J. Mol. Graphics* **4**, 121 (1986)

Choo, Y., and Klug, A., *Nucl. Acids Res.* **21**, 3341 (1993)

Clemens, K.R., Liao, X., Wolf, V., Wright, P.E., and Gottesfeld, J.M., *Proc. Natl. Acad. Sci. USA* **87**, 10822 (1992)

Collaborative Computational Project, Number 4, *Acta Crystallogr.* **D50**, 760 (1994)

Conlin, R.M., and Brown, R.S., *Methods. Mol. Biol.* **30**, 357 (1994)

Cowtan, K., *Joint CCP4 ESF-EACBM Newsletter Protein Crystallogr.* **31**, 34 (1994)

Eck, M.J., Atwell, S.K., Shoelson, S.E. & Harrison, S.C., Nature, 368, (1994)

Eck, M.J. Shoelson, S.E., Harrison, S.C., Nature, 362, (1993)

Elrod-Erickson, M., Rould, A.M., Nekludova, L., and Pabo, C.O. , *Structure* **4**, 1171 (1996)

Engelke, D.R. , Ng, S.-Y., Shastry, B.S., and Roeder, R.G., *Cell* **19**, 717 (1980)

Fairall, L., Rhodes, D., and Klug, A., *J. Mol. Biol.* **192**, 577 (1986);

Fairall, L., Schwabe, J.W.R., Chapman, L., Finch, J.T., and Rhodes,D., *Nature* **366**, 483 (1993)

Forget, B.G., and Weissman, S.M., *J. Biol. Chem.* **244**, 3148 (1969)

Foster, M.P. et al., *Nature Struc. Biol.* **4**, 605 (1997)

Friedell, R.A. et al., *Proc. Natl. Acad. Sci. USA* **93**, 2936 (1996)

Gaskins, C.J., and Hanas, J.S. , *Nucl. Acids Res.* **18**, 2117 (1990)

Gaskins, C.J., Smith, J.F., Ogilvie, M.K., and Hanas, J.S., *Gene* **120**, 197 (1992)

Ginsberg, A.M., King, B.O., and Roeder, R.G., *Cell* **39**, 479 (1984)

Greisman, H.A., and Pabo, C.O., *Science* **275**, 657 (1997).

Hayes, J.J., and Tullius, T.D., *J. Mol. Biol.* **227**, 407 (1992);

Jones, T.A., and Kjeldgaard, K., "O" Computer Graphics Program (Uppsala Univ., Sweden, 1993)

Hanas, J.S. Gaskins, C.J., Smith, J.F., and Ogilvie, M.K., *Prog. Nucl. Acid Res. Mol. Biol.* **43**, 205 (1992)

Harrison, S. C. "A structural taxonomy of DNA-binding domains" Nature, 353, (1991)

Houbaviy, H.B., Usheva, A., Shenk, T., and Burley, S.K., *Proc. Natl. Acad. Sci. USA* **93**, 13577 (1996)

Kim, C.A., and Berg, J.M., *Nature Struct. Biol.* **3**, 940 (1996)

Kleywegt, G.J., and Jones, T.A., *Proceedings of the CCP4 Study Weekend,* (SERC Daresbury, Daresbury, UK, 1994), p. 59

Korn, L.J., and Brown, D.D., *Cell* **15**, 1145 (1978)

Lavery, R., and Sklenar, H.J., *Biomol. Struct. Dyn.* **6**, 655 (1989)

Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A., and Wright, P.E., *Science* **245**, 635 (1989)

Mao X., and Darby, M.K., *Mol. Cell. Biol.* **13**, 7496 (1993)

Maxwell, E.S., and Martin, T.E. , *Nucl. Acids Res.* **14**, 5741 (1986)

McConkey, G.A., and Bogenhagen, D.F., *Mol. Cell. Biol.* **7**, 486 (1987)

Miki, Y. , Swensen, J., Shattuck-Eidens, D., Futreal,P.A., Harshman,K., Tavtigian,S. et al., Science 266, 66-71 (1994)

Miller, J., McLachlan, A.D., and Klug, A., *EMBO J.* **4**, 1609 (1985)

Olson, M.V., et al., *Nature* **267**, 641 (1977)

Otwinowski, Z., and Minor, W., *Meth. Enzymol.* **276**, 307 (1997)

Pavletich, N.P. , and Pabo, C.O. , *Science* **252**, 809 (1991)

Pavletich, N.P., and Pabo, C.O. , *Science* **261**,1701 (1993)

Pelham, H. R. B., and Brown, D. D., *Proc. Natl. Acad. Sci. USA* **77**, 4170 (1980)

Picard, B. and Wegnez, M., *Proc. Natl. Acad. Sci. USA* **76**, 241 (1979)

Pieler, T., Oei, S.-L., Hamm, J., Engelke, U., and Erdmann, V.A.,*EMBO J.* **4**, 3751 (1985)

Pieler, T. and Theunissen, O., *Trends Biochem. Sci.* **18**, 226 (1993)

Rawlings, S.L., Matt, G.D., and Huber, P.W., *J. Biol. Chem.* **271**, 868 (1996)

Renaud, J.P., Rochel, N., Ruff, M., Vivat, V., Chambone, P., Gronemeyer, H., and Moras, D. , Nature 278, 681-689 (1995)

Rowland, O., and Segall, J., *J. Biol. Chem.* **271**, 12103 (1996)

Sakonju, S., and Brown, D.D., *Cell* **31**, 395 (1982)

Shastry, B.S. *Prog. Biophys. Mol. Biol.* **56**, 135 (1991)

Shastry, B.S. *Experientia* **49**, 831 (1993)

Shastry, B.S., *J. Cell Sci.* **109**, 535 (1996)

Smith, J.F., Hawkins, J., Leonard, R.E. and Hanas, J.J., *Nucl. Acids Res.* **19**, 6871 (1991)

Steitz, T. A., Quarterly Reveiws of Biophysics, 23, 205-280 (1990)

Taylor, W., Jackson, I.J., Siegel, N., Kumar, A., and Brown, D.D., *Nucl. Acids. Res.* **14**, 6185 (1986);

Veldhoen, N., You, Q.M., Setzer, D.R., and Romaniuk, P.J., *Biochemistry* **33**, 7568 (1994)

Wagner, R.L., Apriletti, J.W., McGrath, M.E., West, B.E., Baxter, J.D., and Fletterick, R.J., Nature 378, 690-696 (1995)

Wegnez, M., Monier, R., and Denis, H., *FEBS Lett.* **25**, 13 (1972)

Zang, W.Q., Veldhoen, N., and Romaniuk, P.J., *Biochemistry* **34**, 15545 (1995)

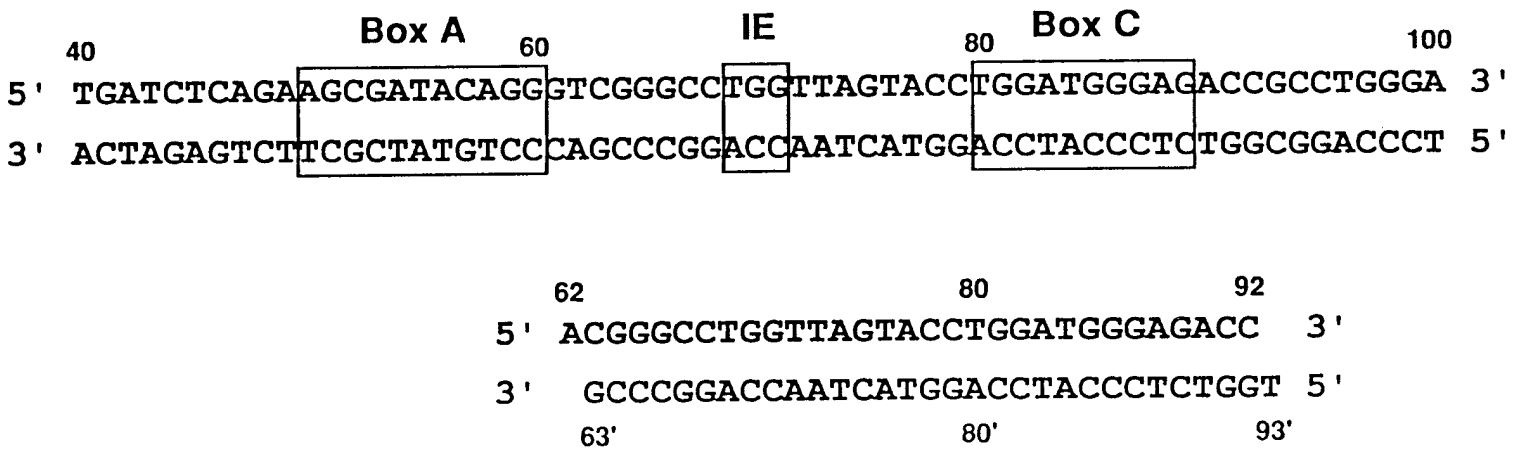
Bibliography

Abstract and Poster Presentation : Nolte , R.T., Conlin, R.M., Brown, R.S. and Harrison, S. C., " CRYSTAL STRUCTURE OF A SIX FINGER TFIIIA - DNA COMPLEX", DOD Era of Hope Meeting, Washington DC, Oct 31-Nov 4, 1997.

Nolte , R.T., Conlin, R.M., Brown, R.S. and Harrison, S. C., "Differing Roles for Zinc Fingers in DNA Recognition: Structure of a Six-Finger TFIIIA Complex." To be submitted to *Proceedings of the National Academy of Sciences*. November, 1997

Fig. 1. Sequences of the DNA and protein used for crystallization. (**A**) Pol III elements of the *X. laevis* oocyte 5S rRNA ICR are shown boxed. The 31 base-pair duplex is numbered according to the 5S rRNA gene. (**B**) The six-finger protein corresponds to amino acid residues 1–190 of *X. laevis* TFIIIA. Zinc fingers are aligned to show their secondary structure. Beta sheet is indicated by zigzag lines and the alpha helix as an open box. The "TA" region is required for transcription activation (Mao, 1993) and "NE" for nuclear export (Friedell, 1996).

A


Box A
IE
Box C
40
60
80
100
5' TGATCTCAGAAGCGATACAGGGTCGGGCCTGGTTAGTACCTGGATGGGAGACCGCCTGGGA 3'
3' ACTAGAGTCTTCGCTATGTCCCAGCCCGGACCAATCATGGACCTACCCTCTGGCGGACCCT 5'
62
80
92
5' ACGGGCCTGGTTAGTACCTGGATGGGAGACC 3'
3' GCCCGGACCAATCATGGACCTACCCTCTGGT 5'
63'
80'
93'

B

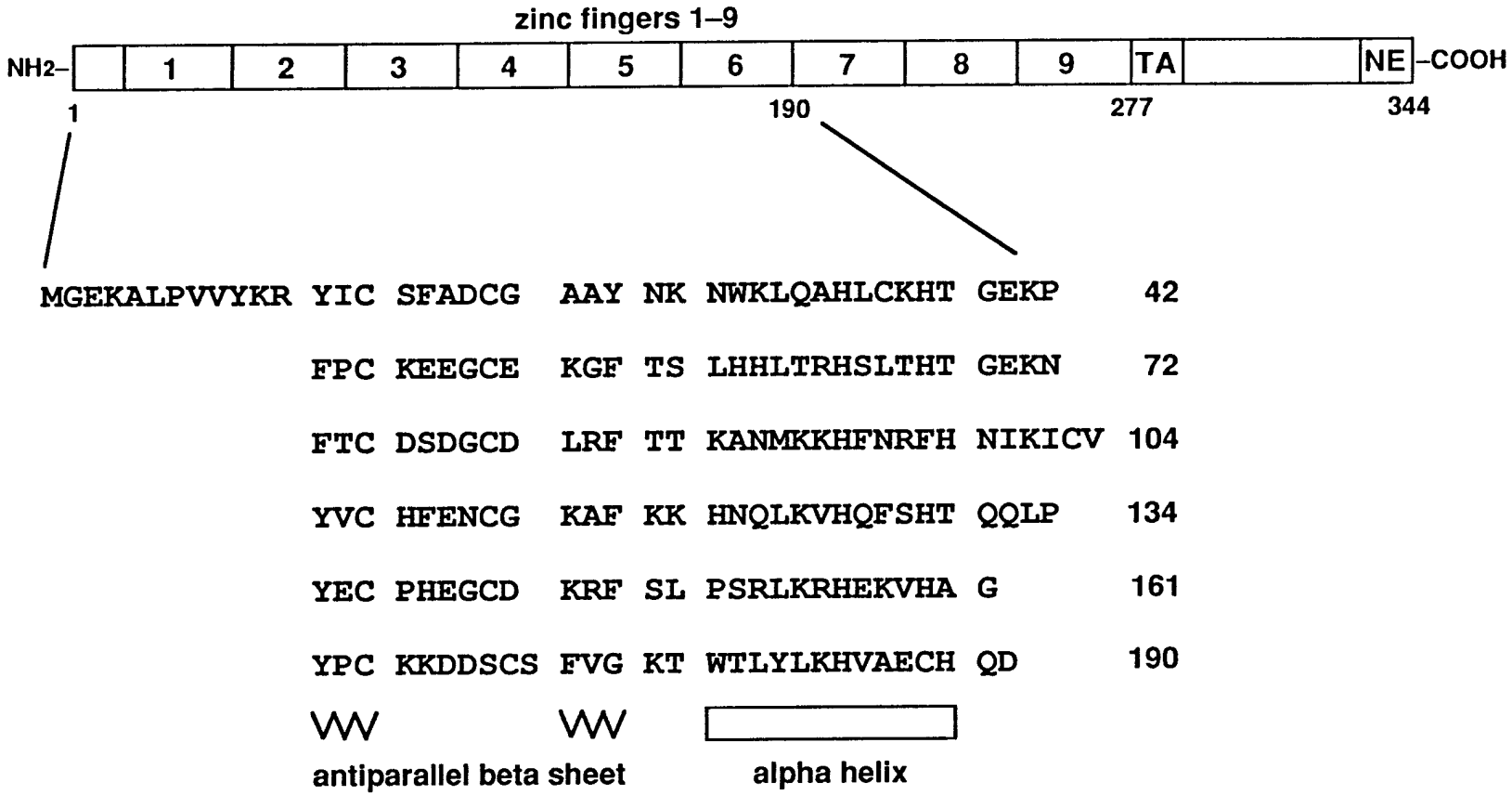
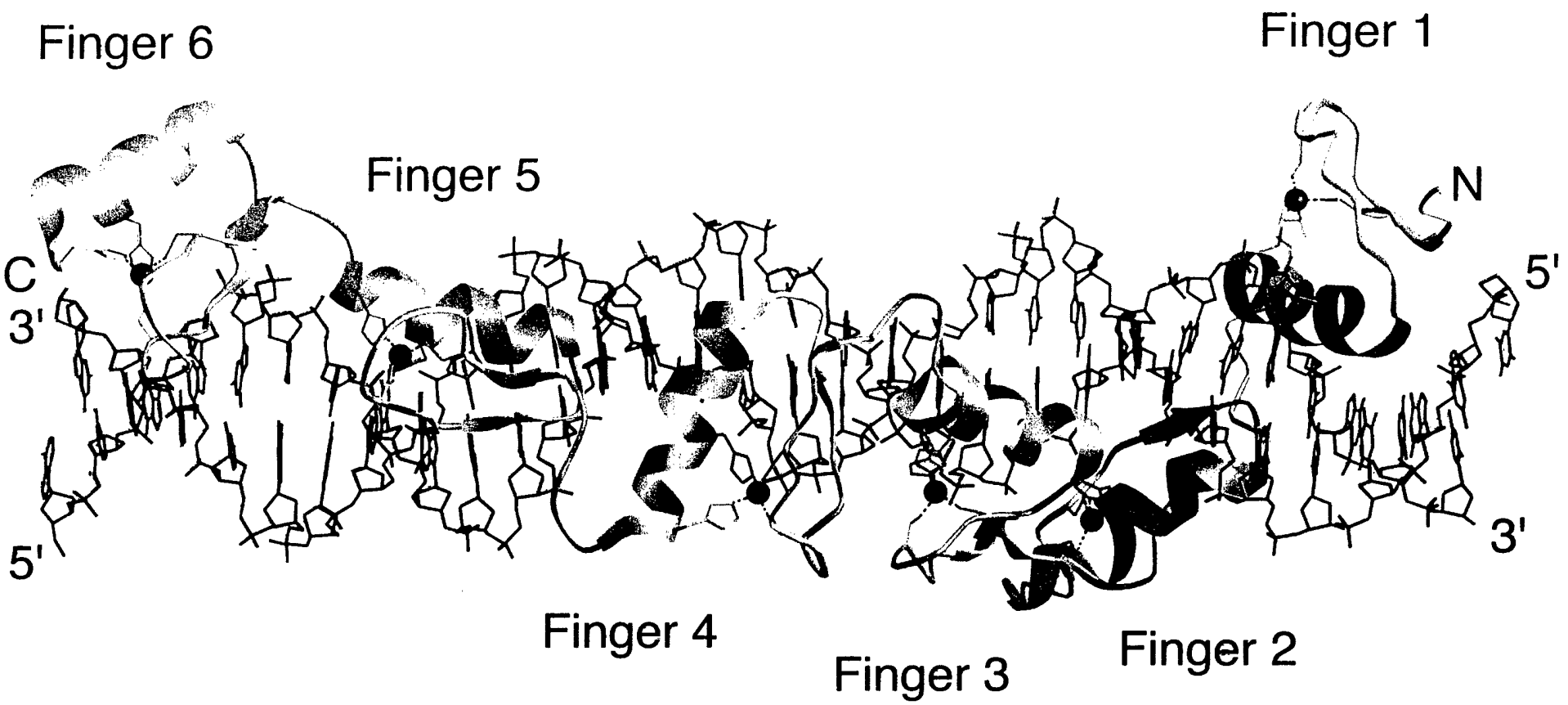

zinc fingers 1–9
NH2–
1
2
3
4
5
6
7
8
9
TA
NE
–COOH
1
190
277
344
MGEKALPVVYKR YIC SFADCG AAY NK NWKLQAHLCKHT GEKP 42
FPC KEEGCE KGF TS LHHLTRHSLTHT GEKN 72
FTC DSDGCD LRF TT KANMKKHFNRFH NIKICV 104
YVC HFENCG KAF KK HNQLKVHQFSHT QQLP 134
YEC PHEGCD KRF SL PSRLKRHEKVHA G 161
YPC KKDDSCS FVG KT WTLYLKHVAECH QD 190
antiparallel beta sheet
alpha helix

Fig. 2. Structure of the six-finger TFIIIA–DNA complex. (**A**) A RIBBONS (Carson, 1986) representation, in which alpha helices and beta sheets of TFIIIA are colored yellow; Zn(II) ions are red spheres; and the DNA double helix is light blue. (**B**) Crystallographic assignment of zinc fingers and DNA bases to locations within the complex. A zinc anomalous difference Fourier map (calculated at 4 Å resolution, white contour levels > 4*s*) and 5-IdU difference Fourier maps (calculated at 5 Å resolution, red contour levels > 5*s*) are shown superimposed on a molecular model generated with the program O (Jones, 1993). Carbon atoms are colored yellow; oxygen, red; nitrogen, blue; sulfur, green; and phosphate groups, magenta. The direction of view in (B) is oriented approximately perpendicular to the direction in (A).

A



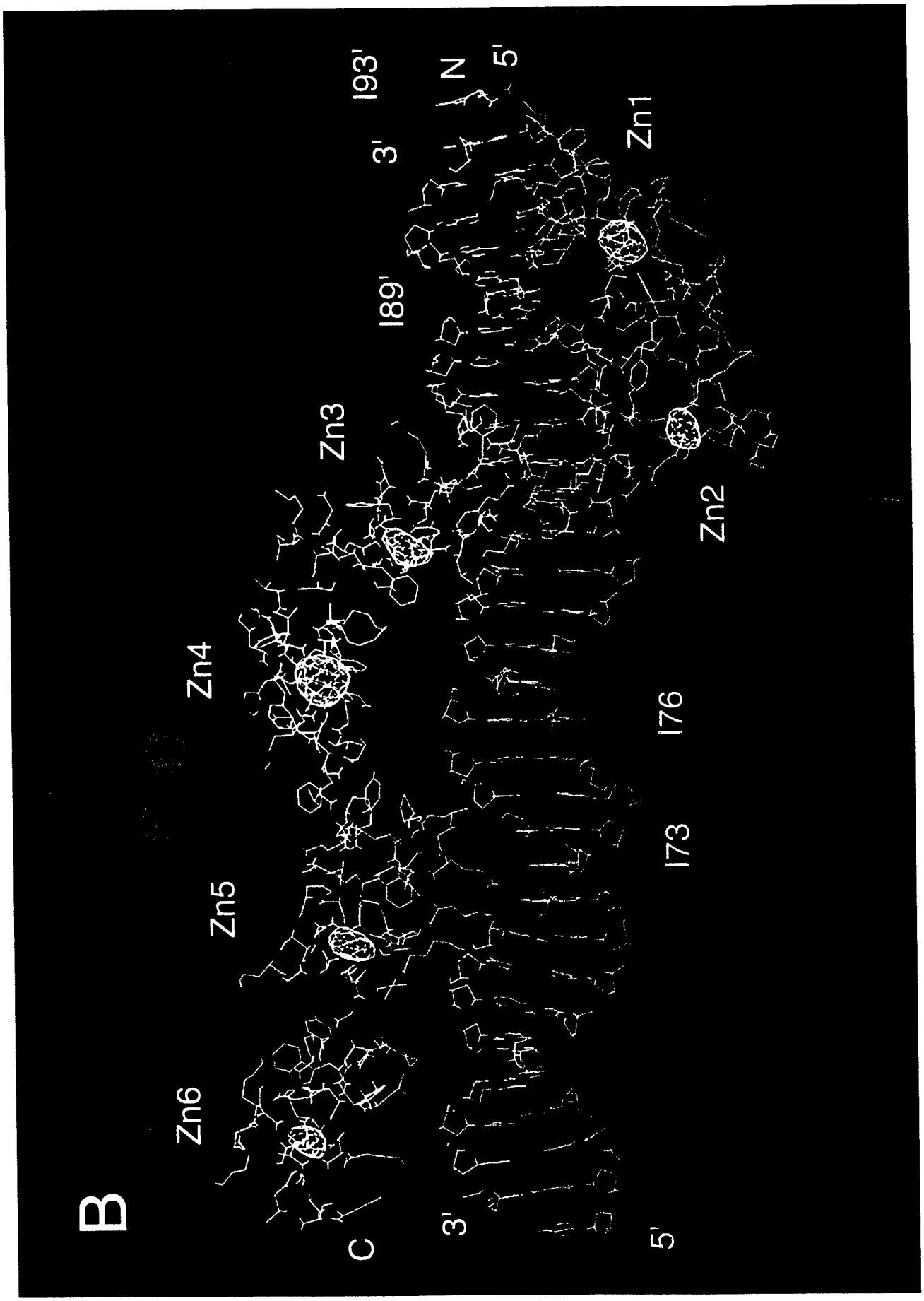
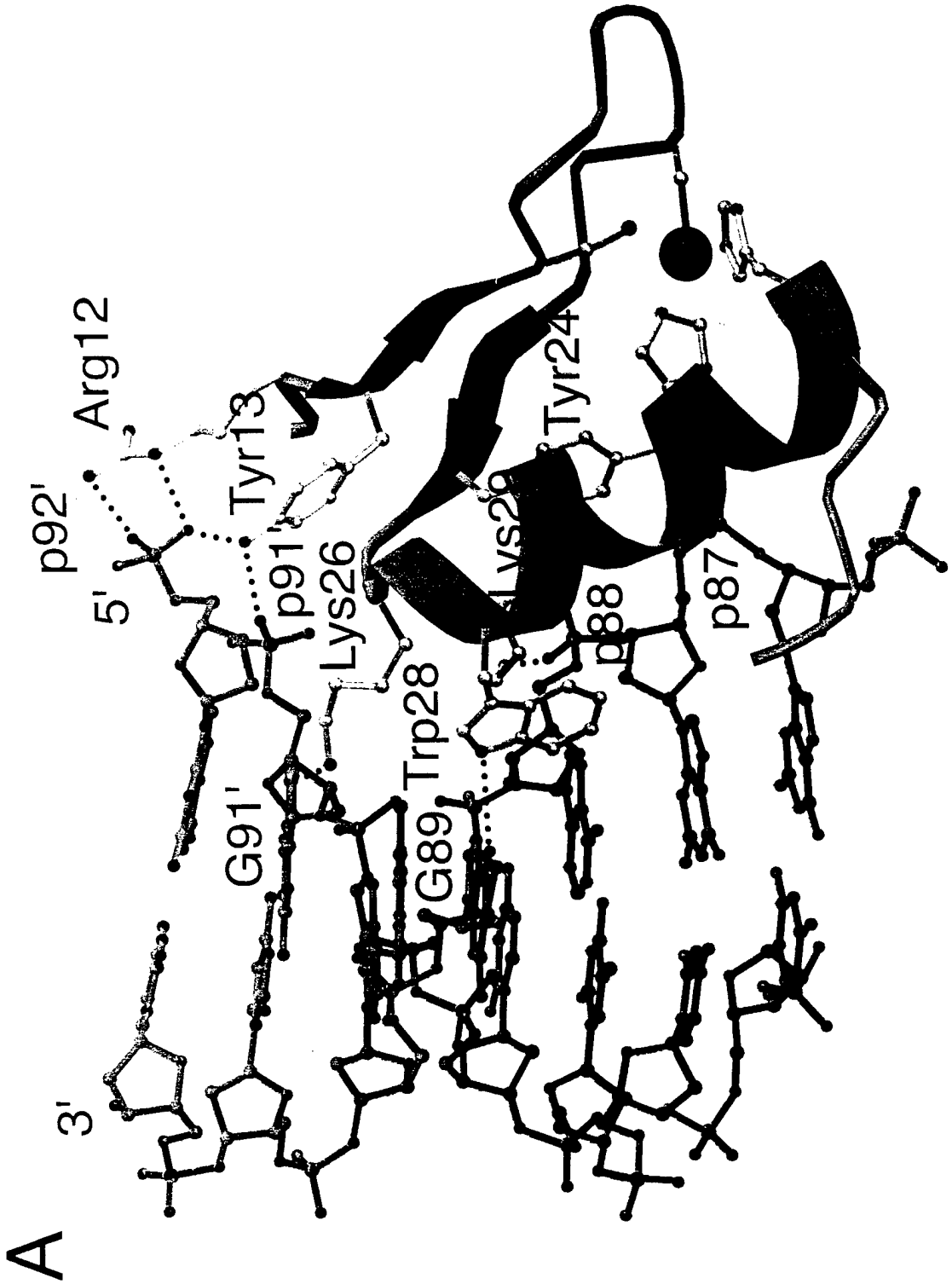
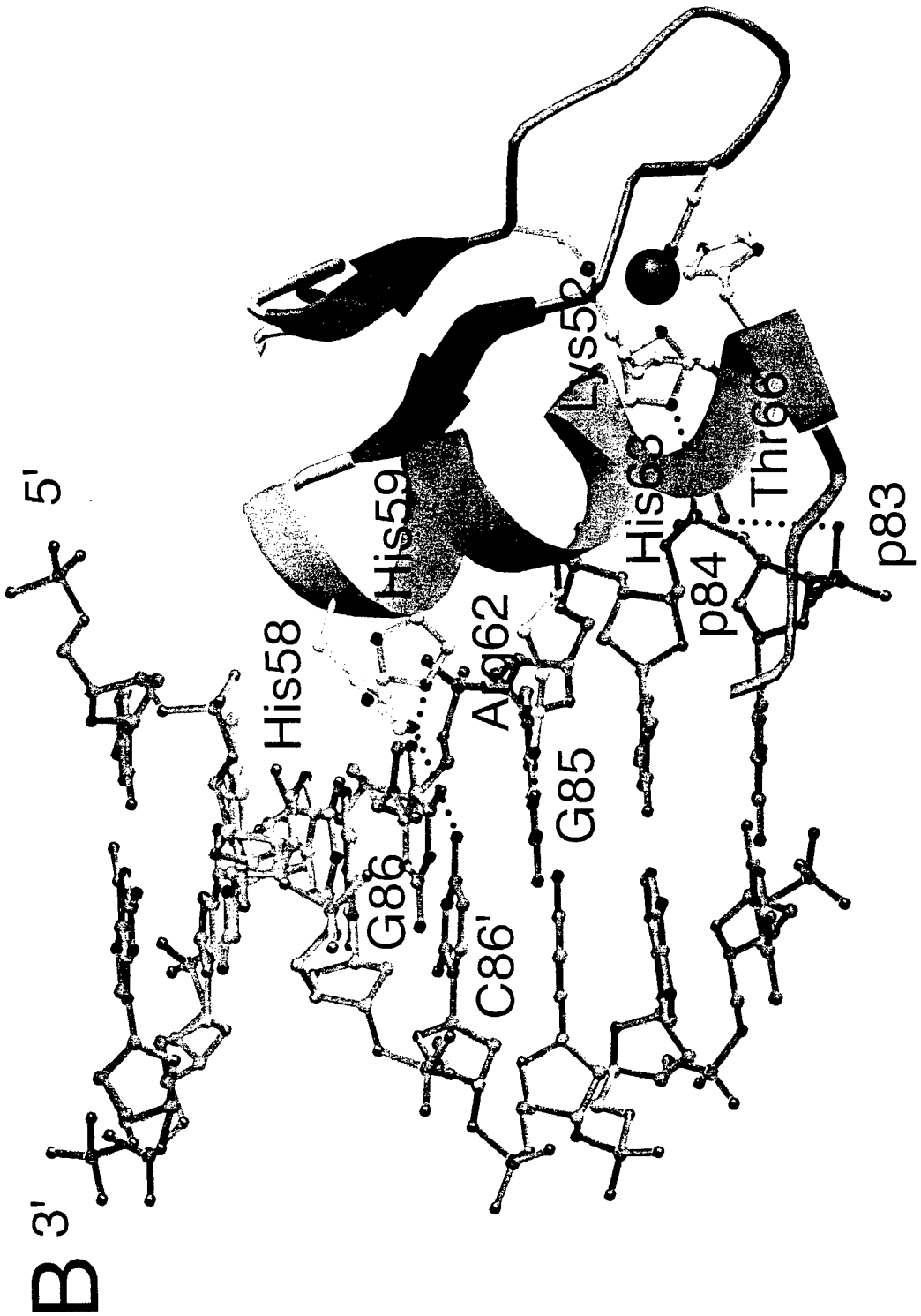
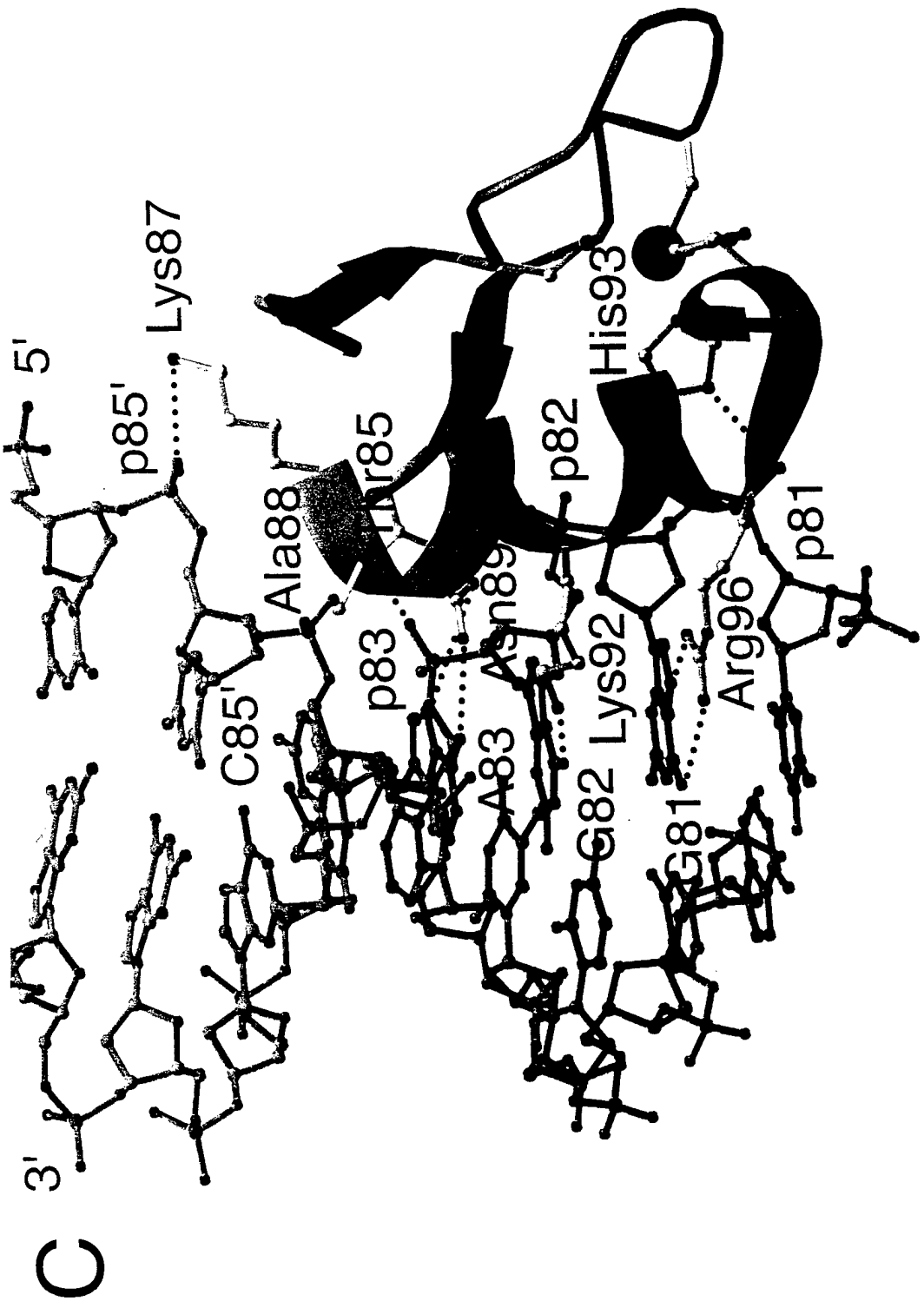

B
Zn6
Zn4
Zn5
Zn3
Zn1
Zn2
I93'
I89'
I76
I73
N
5'
3'
C
3'
5'

Fig. 3. DNA major-groove contacts with each of the zinc fingers 1, 2, 3 and 5. The zinc fingers are placed in similar orientations (**A** to **D**). The protein is shown as a ribbon with alpha helix, blue, and beta sheet, green. The DNA is light blue.The amino acid side-chains that contact nucleotide bases are yellow, and hydrogen-bond contacts are shown as dotted lines. Oxygen atoms are red, and nitrogen, magenta. There are many biochemical results (Fairall, 1986, Sakonju, 1982, Clemens, 1992, Pieler, 1985, McConkey, 1987, Veldhoen, 1994, Rawlings, 1996, Smith, 1991, Choo, 1993, Zang, 1995) that support these interactions. (**E** to **H**). The major groove of DNA is represented schematically in cylindrical projection. The noncoding strand is numbered as in the 5S rRNA gene. Nucleotide bases of the "canonical" quartet for contacts by zinc fingers in previously analyzed structures, are shown shaded, as are two phosphates that frequently receive hydrogen bonds. Contacts between amino acids and DNA are drawn as arrows.


A
3'
5'
p92'
Arg12
Tyr13
p91'
Lys26
G91'
G89
Trp28
Lys29
Tyr24
p88
p87


B
3'
5'
His58
His59
Lys52
His63
Thr66
p84
p83
A62
G85
G86
C86'


C
3'
5'
p85'
Lys87
C85'
Ala88
Thr85
p83
Asn89
A83
p82
His93
G82
Lys92
G81
Arg96
p81

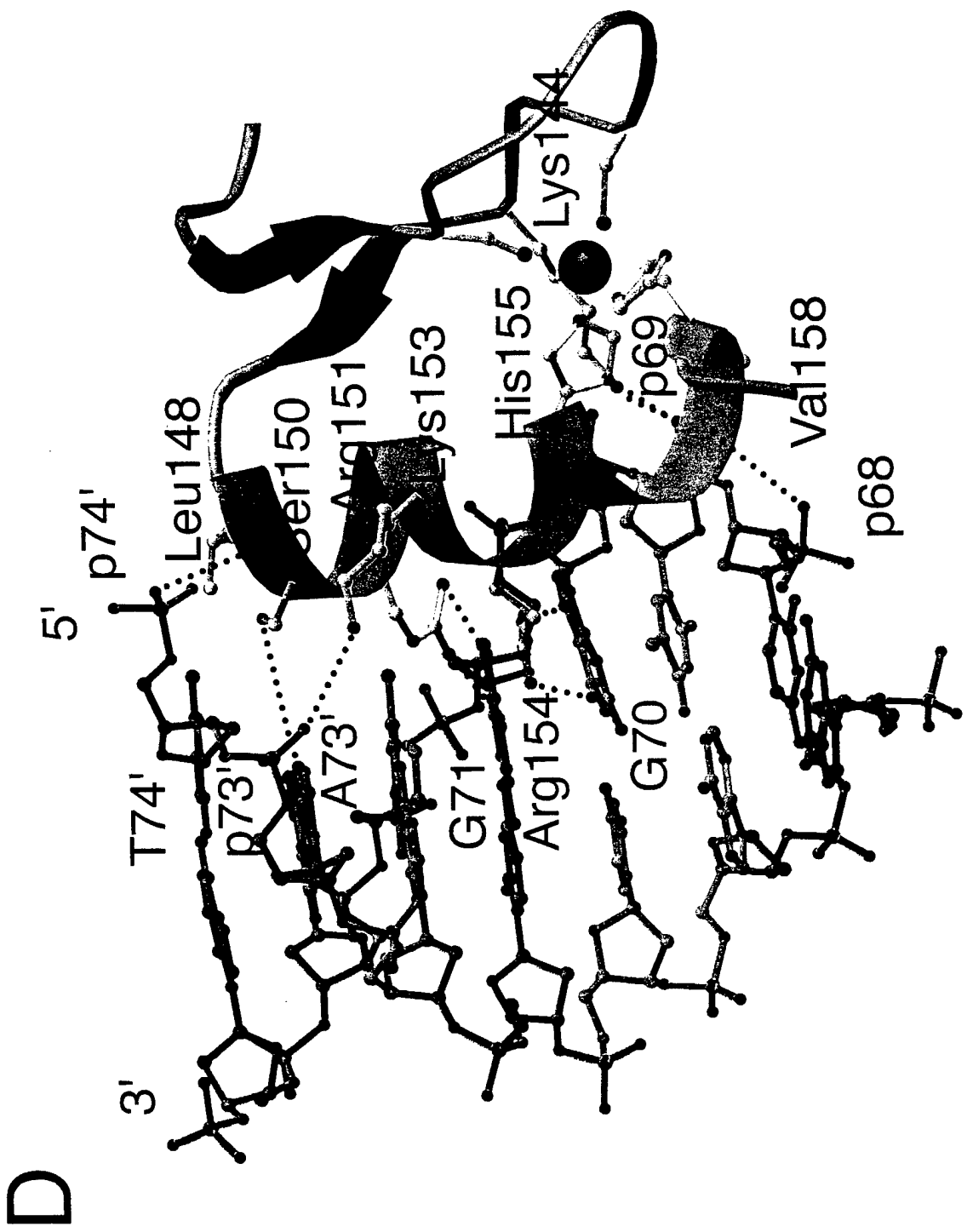

D
5'
p74'
Leu148
Ser150
Arg151
Lys153
His155
Lys154
p69
Val158
p68
3'
T74'
p73'
A73'
G71
Arg154
G70

Fig. 4. A plausible model for the nine-finger TFIIIA–DNA complex. The DNA double helix (purple) and the TFIIIA zinc fingers (green) are shown as ribbons. Zn(II) ions are red spheres. The positioning of the COOH-terminal fingers 7–8–9 in the major groove of DNA is derived from biochemical results (Hayes, 1992, Clemens, 1992).