Order Number 9405525

The role of flexibility in the specific interactions of even zinc fingers to DNA

Feng, Waldo C., Ph.D.

City University of New York, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

THE ROLE OF FLEXIBILITY IN THE SPECIFIC INTERACTIONS OF "EVEN" ZINC FINGERS TO DNA

by

Waldo C. Feng

A dissertation submitted to the Graduate Faculty in Biomedical Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1993

This manuscript has been read and accepted for the Graduate Faculty in Biomedical Sciences in satisfaction of the dissertation requirements for the degree of Doctor of Philosophy.

May 28, 1993
Date

Chair of Examining Committee

May 28, 1993
Date

Executive Officer

James Bieker
Carter Bancroft
Roman Osman
David Setzer

The City University of New York

ABSTRACT

THE ROLE OF FLEXIBILITY IN THE SPECIFIC INTERACTIONS OF "EVEN" ZINC FINGERS TO DNA

BY

WALDO C. FENG

Advisor: Dr. Harel Weinstein

The molecular basis of specificity and affinity in protein-DNA complexes was explored with the study of interactions between the canonical (CC/HH) zinc finger motifs of Transcription Factor IIIA (*X.laevis*) and the 5S gene. The nine zinc fingers of TFIIIA can be classified as "odd" or "even" based upon the number of intervening residues (3 or 4) between the conserved histidine ligands in the zinc cluster. The systematic arrangement of the "odd" and "even" zinc fingers in TFIIIA and other multi-zinc finger proteins suggests an underlying functional role in protein-DNA binding.

To assess the molecular basis of such a functional role, the structural and dynamic properties of the zinc fingers were explored with molecular dynamics simulations which revealed major differences in the properties of the two classes of zinc figers. The time dependent motions of the zinc cluster were found to be correlated with the structure of the DNA-binding helix in the "even" zinc finger only,

producing a degree of structural flexibility not available in the "odd" zinc finger. The binding of zinc finger proteins containing both "odd" and "even" fingers is proposed to be dependent on the observed difference in the flexibility of the two types.

This hypothesis was probed in gel retardation assays with native and mutant TFIIIA zinc fingers obtained from an *in vitro* transcription/translation system. A deletion mutant missing one residue between the conserved histidines of the third zinc finger of TFIIIA (i.e., making an "even" to "odd" ZF) was constructed. In gel shift experiments, TFIIIA containing teh mutant finger 3 was found to have decreased DNA binding affinity, in agreement with the inferences developed from theoretical studies that the "even" zinc fingers of TFIIIA are required for positioning groups of "odd" for sequence specific recognition.

A detailed molecular model, consistent with known structural data (e.g., footprinting) of the TFIIIA-DNA complex was constructed. The model incorporates the structural implications of the observed flexibility of the "even" zinc fingers and proposes a novel mode of interaction for zinc fingers 4, 7, and 9 to DNA.

ACKNOWLEDGEMENTS

This thesis is dedicated in loving memory to Howard Feng (or "third uncle") who first encouraged my curiosity in the "nature of things." I would like to thank my family for their support and understanding during the difficult times as well as the joyous ones - my accomplishments are to be shared and celebrated by all of us. I would also like to acknowledge Ms. Gloria Jan for her unswaying confidence and belief in me.

This thesis dissertation would not have been possible without the aid and guidance of several people. Most of all I would like to acknowledge my mentor Dr. Harel Weinstein who has taught me the scientific method. More importantly, he has inspired me with his dignity and his humanity. The molecular biology experiments would not have been possible without the generosity of Dr. Carter Bancroft and his laboratory. I would like to especially acknowledge the help and advice of Brian Kloss in the cloning and gel shift studies. The computational experiments were facilitated with much greater ease by the system manager Janne Raviantti. I would also like to acknowledge the help and counsel of Ms. Mildred Tolson for these past four years.

Finally, I would like to acknowledge the MSTP (Medical Scientist Training Program) for its financial support.

TABLE OF CONTENTS

APPROVAL PAGE ii

ABSTRACT iii

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF TABLES ix

LIST OF FIGURES x

1. INTRODUCTION 1

1.1. Sequence Specific DNA Recognition 1

1.2. Classical (CC/HH) Zinc Fingers as a DNA Binding Motif 7

1.2.1. Historical Background of Zinc Fingers 7

1.2.2. Structure of Zinc Fingers 9

1.3. The TFIIIA Paradox 12

1.4. Proposed models of TFIIIA/5S complex 14

2. CHAPTER 2 22

2.1. Introduction 22

2.2. Methods 33

2.2.1. Homology Modeling 33

2.2.2. Zinc Ion Parameters 38

2.2.3. Minimization Procedure 42

2.2.4. Molecular Dynamics 42

2.3. Results 45

2.3.1 Molecular Dynamics Trajectories Analyses 45

2.3.2. Simulation of the Xfin31 Zinc Finger 52

2.3.2.1. Xfin31 molecular dynamics structural comparison to 2D NMR structure....52
2.3.3. Zinc Metal Cluster....57
2.3.4. Comparison of the TFIIIA(iii) "Even" and TFIIIA(iv) "Odd" Simulations....65
2.3.4.1. Dynamic properties of "even" and "odd" zinc fingers....65
2.3.4.2. Superpositioning of 5 MD (10 psec) Averaged Structures....76
2.3.5. Correlation Analysis of the Zinc Ion Cluster to Hydrogen Bonding....81
2.4. Discussion and Conclusions....87
3. CHAPTER 3....92
3.1. Introduction....92
3.2. Methods....93
3.2.1. Constrained Molecular Dynamics....93
3.2.2. Deletion Mutation - TFIIIA(iiimf)....94
3.3. Results....98
3.3.1. MD analysis of TFIIIA(iiic) and TFIIIA(ivc)....98
3.3.1.1. The effect of constrained cluster....108
3.3.2. TFIIIA(iiimf) Simulation....118
3.4. Discussion and Conclusions....122
4. CHAPTER 4....124
4.1. Introduction....124
4.2. Materials and Methods....125
4.2.1. Subclone of the TFIIIA cDNA into the pGEM1 vector....125

4.2.2. PCR (Polymerase Chain Reaction) Deletion Mutant TFIIIA(iiimf) 131
4.2.3. *In vitro* transcription/translation 135
4.2.4. Gel retardation 137
4.3. Results 138
4.3.1. *In vitro* transcription/translation 138
4.3.2. Gel Retardation 144
4.4. Discussion and Conclusions 151
5. CHAPTER 5 154
5.1. Model of the TFIIIA-5S Complex 154
5.1.1. Conformation of the 5S DNA 154
5.1.2. Conformation of the zinc finger 154
5.1.3. General Features of the Model 155
5.1.4. Construction of the Model 163
5.1.5. Alignment for the rest of the fingers 172
6. CONCLUSION 180
REFERENCES 183
BIBLIOGRAPHY 193

LIST OF TABLES

Table Page

1. Molecular systems used in Molecular Dynamics simulations with Periodic Boundary Conditions (PBC) 37

2. Zn+2 covalent force field 39

3. Modified charges for the histidine ligands (HIZ) 40

4. Modified charges for the cysteine ligands (CYZ) 41

5. Geometry parameters of the metal cluster 60

6. RMS fluctuation values for the metal ligands only of "even" TFIIIA(iii) and "odd" TFIIIA(iv) zinc fingers. 75

7. Helix mainchain hydrogen bonding profile of TFIIIA(iii) and TFIIIA(iv) 80

8. Cross correlation table of hydrogen bonds to metal ligand angles of TFIIIA(iv). 84

9. Cross correlation table of hydrogen bonds to metal ligand angles of TFIIIA(iii) 86

10. Description of MD simulations discussed 97

11. Geometry of the metal ligands during the constrained simulations 109

12. Comparison of the RMS metal ligand fluctuations of the TFIIIA(iii) and TFIIIA(iiic) simulations 115

13. The DNA sequence of PCR Primers 134

14. Different lengths of mRNA can be transcribed 136

15. Summary of experimental data of the TFIIIA-5S Complex 162

LIST OF ILLUSTRATIONS

Figure Page

1. Hydrogen bonding scheme 2

2. Primary sequence of TFIIIA. 23

3. Compilation of zinc finger proteins with "even" and "odd" zinc fingers 25

4. Consequences of an a helical structure and metal binding for "odd" HX3H spacing finger 31

5. Homology modeling of TFIIIA zinc fingers 35

6. AMBER force field equation 44

7. Potential energy trajectory of TFIIIA(iii) "even," TFIIIA "odd," and xfin31 zinc fingers 47

8. RMS structural deviation trajectory of TFIIIA(iii) "even," TFIIIA "odd," and xfin31 zinc fingers 50

9. 2DNMR and MD structure of xinf31 are superimposed 53

10. Comparison of the proton-proton distances 56

11. Metal cluster of the xfin31 50 psec time averaged structure. 58

12. Comparison of the metal cluster distances of 50 psec time averaged structures from TFIIIA(iii), TFIIIA(iv), and xfin31 62

13. Comparison of the metal cluster angles of 50 psec time averaged structures from TFIIIA(iii), TFIIIA(iv), and xfin31 64

14. 50 psec time averaged structure of TFIIIA(iii) "even" zinc finger 67

15. 50 psec time averaged structure of TFIIIA(iv) "odd" zinc finger 68

16. RMS fluctuation for each amino acid of TFIIIA(iii) and TFIIIA(iv) 50 psec time averaged structure 70

17. Residue fluctuation differences between TFIIIA(iii) and TFIIIA(iv 73

18. Five superimposed structures (10 psec time averaged) of "even" TFIIIA(iii) 77

19. Five superimposed structures (10 psec time averaged) of "odd" TFIIIA(iv) 78

20. Amino acid sequence comparison of the wildtype and mutant zinc finger 95

21. RMS structural deviation trajectory of TFIIIA(ivc) 99

22. RMS structural deviation/potential energy trajectory of TFIIIA(iiic) 100

23. Temperature and total energy trajectory of TFIIIA(iiic) 102

24. Kinetic and potential energy trajectory of TFIIIA(iiic) 103

25. Structural comparison of the 50 psec time averaged structures of TFIIIA(iv) and TFIIIA(ivc) 105

26. RMS difference between the 50 psec structures of TFIIIA(iiic) and TFIIIA(iii) 107

27. Comparison of RMS fluctuations of TFIIIA(iv) and the constrained finger TFIIIA(ivc) 111

28. RMS fluctuation of TFIIIA(iii) and the constrained finger TFIIIA(iiic) 113

29. Five superimposed 10 psec average structures of TFIIIA(iiic) 117

30. RMS structural deviation trajectory of TFIIIA(iiimf) 119

31. RMS fluctuation for each amino acid of TFIIIA(iv) and TFIIIA(iii) (from Figure 16.) and TFIIIA(iiimf)....121
32. Subcloning strategy....127
33. Determination of insert orientation in pGEM1 vector....130
34. Single deletion mutation strategy with PCR....132
35. RNA transcripts from *in vitro* transcription....139
36. 15% SDS denaturing gel of the *in vitro* translated products....141
37. *In vitro* translated product of TFIIIA(iiimf)....143
38. Gel retardation of *in vitro* translated products....146
39. Binding of the seven zinc fingers construct....148
40. Binding of the three zinc fingers construct....150
41. Zif268 co-crystal....156
42. Proposed DNA contacting residues of TFIIIA....158
43. Three N terminus zinc fingers and their DNA contacts....160
44. Model of the three N terminal TFIIIA zinc fingers with DNA....164
45. Mainchain hydrogen bonds between the linker and the HX3H spacing....166
46. The graphic depiction of the orientation of the zinc finger recognition helices (ZF1-4) around the central axis of the DNA....168
47. Φ torsional angles from TFIIIA(iv) and TFIIIA(iii)....170
48. Ψ torsional angles from TFIIIA(iv) and TFIIIA(iii)....171
49. Proposed DNA contacting residues of ZF5....173
50. Model of the N terminal 6 zinc fingers with DNA....175
51. Proposed DNA contacting residues of ZF7-8....177
52. TFIIIA fingers 7, 8, and 9 with DNA....178

53. Model of the TFIIIA-5S complex 179

1. INTRODUCTION

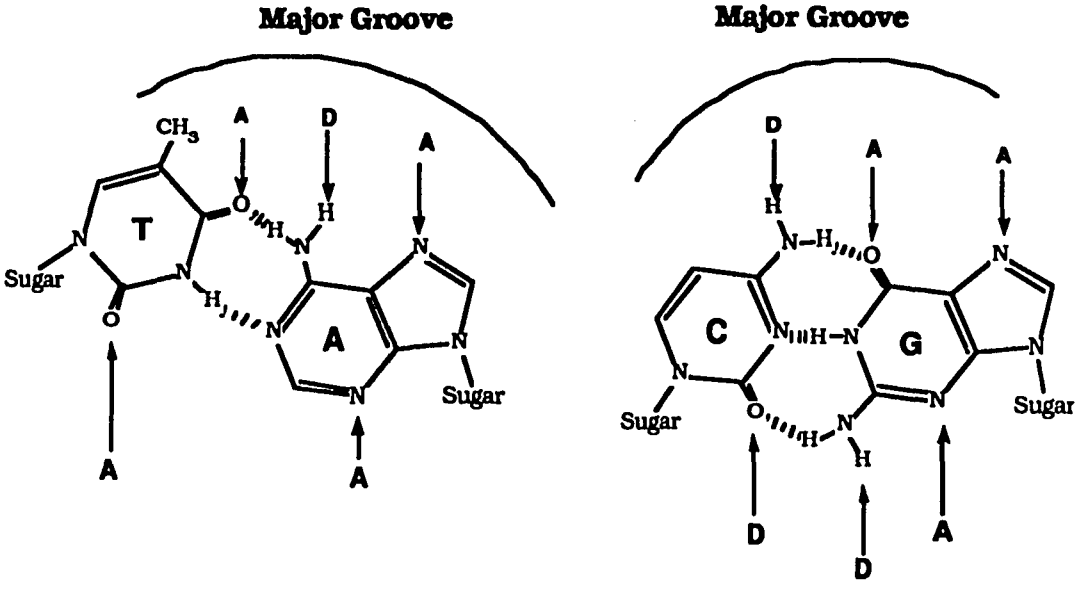
Sequence Specific DNA Recognition

The control of genetic information through the activation and repression of genes by DNA-binding proteins is a fundamental process of life. However, the essential properties of selectivity and activity that underlie the mechanism of protein-DNA interaction are not well understood. For example, it is not clear what are the nature and structural basis of the molecular forces that determine DNA sequence recognition, and how they contribute to specific binding of proteins to DNA. Sequence and structural analysis of myriad of DNA-binding proteins have identified certain structural motifs[1, 2, 3, 4](e.g., zinc fingers, helix-turn-helix, and leucine zippers, binuclear clusters, β-barrel, etc.) that are conserved among a variety DNA-binding proteins. What remains to be established is the relationship between the amino acid sequence and the three dimensional structure of these protein motifs, as well as the manner in which the structural elements determine the functionality of these units and their repeats. Given their ubiquity, the physico-chemical forces governing these particular structures and their roles as DNA-binding elements constitute an important question in current structural biology.

Seeman et al[5] had proposed that specific protein-DNA recognition involved direct hydrogen bonds from the side chains of proteins to the base pairs of DNA forming a hydrogen bonding code, (see figure 1).

Figure 1

Hydrogen bonding scheme as a basis for specific recognition. Potential sites for donor hydrogen bonds (D) or acceptor hydrogen bonds (A) are identified. The major groove of the DNA is above the base pair and the minor groove is depicted below. Watson-Crick hydrogen is shown as dashed lines between the base pairs.



The relative juxtapositioning of potential acceptor (**A**) and donor atoms (**D**) forms a pattern which allows for the discrimination of G•C versus A•T base pairing. For example, the corresponding side-chains of the DNA-binding protein would have a complementary hydrogen bonding pattern D•A•D in order to discriminate the T-A base pair from the G-C base pair. Thus, a series of these contacts can specifically read a cognate DNA sequence. Hydrogen bonds themselves are distance and vectorially dependent, making them ideal candidates as sources of specific interaction. The distance between the donors and acceptors should be approximately 2.5 - 3.2 Å and the angle formed by the three atoms (D-H•••A) should be colinear for energetically favorable hydrogen bonding interaction[6].

As the knowledge base for protein-DNA complexes increasingly expands from crystallographic, biochemical, and molecular biological data, the notion that "direct" readout of DNA sequences by hydrogen bonds is the sole source of specificity appears increasingly inadequate to account for the findings. For example, the structural analysis of the *trp* repressor/operator complex[7] at atomic resolution (2.4 Å) indicated that 12 of the 14 "direct" hydrogen bonding contacts from the protein to the DNA involved phosphate groups rather than base pairs. For one of the contacts to the base pair (the sidechain of R69 to G•C), a mutation of the operator sequence to C•G did not affect binding affinity. Furthermore, a large number of hydrogen bonds to the DNA were made through water-mediated contacts, or "indirect" readout. The authors proposed that the *trp* repressor is able to recognize the sequence of the operator through the sequence-dependent structural

variation of the DNA rather than through direct base pair interactions hypothesized earlier by Seeman.

Wharton and Ptashne identified the role of van der Waals contacts as a component of sequence specific recognition in 434 repressor protein[8, 9]. The repressor (helix-turn-helix protein) binds to its receptor by inserting the recognition helix into the major groove of the operator. They identified a mutant repressor protein, "repressorala28," in which a residue in the recognition helix (glutamine 28) was mutated to an alanine. RepressorAla28 binds to a mutant operator in which a cytosine was substituted by thymine. Modeling of the complex suggested that alanine 28 forms a favorable van der Waals contact to the thymine methyl of the mutant operator. This hypothesis was supported by further binding studies in which the substituted thymine was replaced by bromouracil.

A series of experiments of two different protein-DNA complexes demonstrated that DNA flexibility also could be a source of specific interaction. Koudelka et al[10, 11] have shown that the "twistability" of the 434 operator site is an important contributor to 434 λ repressor binding. 434 λ repressor binds to its operator site as a dimer, with each monomer contacting a half-site. From the analysis of the crystal structure, the recognition helix of each monomer does not contact the inner two base pairs of each half-site; however, mutations of the non-contacted bases affect DNA affinity. In order to show that the affinity is dependent in part on the inner four base pairs which affect the ease with which the DNA can be deformed, the authors introduced single strand cuts at the center of the operator sequence to enhance the

"twistability" of the DNA resulting in increased affinity. Gartenberg and Crothers[12] have shown through mutations of the non-consensus binding site of catabolite activator protein (CAP) and subsequent electrophoretic mobility studies that the non-consensus sequence can change affinity by affecting the overall bending of the DNA.

These experiments have shown that the molecular forces which determine specific protein-DNA recognition involve more than just hydrogen bonds in a "direct" base-pair readout mechanism. Furthermore, if DNA flexibility is an important component of specificity, then it is reasonable to suggest that specificity involves protein flexibility as well. A study of the role of flexibility in DNA complexation would then require a dynamic interpretation of the protein-DNA interaction. A static representation assumes that the components for specificity (e.g., hydrogen bonds, steric, van der Waals, etc.) are pre-positioned before binding and specificity is achieved when complementary elements of the protein and DNA interact[13]. In contrast, a dynamic interpretation assumes that the sequence-specific complementary components of the DNA and protein involved in recognition "find" each during complexation. Thus for large protein-DNA complexes where the contacts are numerous, the entropic component of the free energy of binding may be reduced. Flexibility would allow a portion of the DNA-binding domain initially to form energetically favorable interactions. Subsequently, other regions of the DNA-binding domain would be brought closer to their DNA site and would bind to their base pairs with a substantially reduced entropic term. A dynamic interpretation would require an

understanding of the way in which the flexibility of the protein is introduced in the primary structure, and how this dynamic property is required for specificity. I intend to show that protein flexibility is an essential component of the specific interaction of transcription factor IIIA (TFIIIA) of *Xenopus laevis* with the 5S gene.

TFIIIA is one of the three eukaryotic transcription factors required to initiate transcription of the 5S gene by RNA polymerase III[14]. The DNA-binding domain constitutes approximately 80% of the protein. Within the DNA-binding region of the protein are nine independent folding domains of the zinc finger motif (for a complete description see below). Putative zinc fingers have been identified in over 300 genes representing a dominant class of DNA-binding motif in eukaryotic transcription factors[15, 16, 17]. Since its first discovery over ten years ago, a wealth of information has been gathered about the TFIIIA-5S complex and zinc fingers with DNA in general. However, the fundamental questions of the manner in which the nine zinc fingers are able to recognize the long stretch of the internal control region (ICR) of the 5S gene and bind with DNA are not answered completely. This thesis focuses on the role of protein flexibility in sequence-specific recognition of DNA and provides new evidence for the specialized role of certain zinc fingers in TFIIIA for binding with the 5S gene.

Classical (CC/HH) Zinc Fingers as a DNA Binding Motif

Historical Background of Zinc Fingers

In 1983, Jay Hanas et al[18] reported that transcription factor IIIa in a complex formed with 5S RNA (7S complex) contained multiple zinc ions which were tightly bound. The isolated protein was shown also to have specific binding to the 5S gene. Furthermore, treatment of the protein with EDTA, a dechelating agent, resulted in lost of DNA-binding activity which could later be recovered by the addition of zinc metal. The functional role of *Xenopus laevis* TFIIIA, correct initiation of transcription by RNA polymerase III at the internal control region, was described earlier by Donald Brown's groups [19, 20, 21].

An initial analysis of the cDNA clone of *X. laevis* TFIIIA[22] indicated no homologies with any known DNA-binding protein at that time. However, an inspection of the primary sequence by Miller et al[23] hinted at the idea of small independent folding domains. Furthermore, atomic absorption spectrophotometry indicated that TFIIIA contains 7-11 tightly bound zinc ions. Prolonged proteolytic digestion of the protein with trypsin yielded 3 kDa fragments, each consisting of approximately 30 residues and containing two invariant pairs of cysteines and histidines, the most common ligands for zinc metal ion. Based on these observations, the authors named the structural units formed by the tandem repeated sequences "zinc fingers."

These early experiments suggested the requirement of zinc for the functional activity of TFIIIA (i.e., sequence-specific recognition), but the role zinc ions played in DNA-binding remained unclear. Frankel et al[24] proposed that zinc is required for the correct folding of the zinc finger. They were able to clone a single zinc finger of TFIIIA, finger 2,

and were able to demonstrate significant changes in the circular dichroism spectrum upon the addition of zinc. Absorption spectral of metal substitution of the finger peptide with cobalt suggested that the conserved cysteines may chelate the metal. However, zinc was readily able to displace cobalt in competition studies. In addition to the CD data, they showed that zinc protects the finger peptide from tryptic digestion.

Parraga et al[25] reported changes in NMR spectra upon the addition of zinc to a synthetic zinc-finger peptide ADR1b, a yeast transcription activator. CD experiments indicated that helical structure was induced in the peptide by the addition of zinc. Spectroscopic evidence with cobalt substitution suggested tetrahedral geometry of the zinc binding site. Chemical shifts of the imidizole protons at different pH's implicate them in metal chelation. An early EXAFS (Experimental Absorption X ray Fine Spectroscopy) study suggested that the zinc ligands were the conserved cysteines and histidines[26] and Berg et al[27] were able to demonstrate also with EXAFS that the metal ion induced folding - the transition of the ligands to a tetrahedral cluster in the native peptide - is specific for zinc ion. Redemann et al[28] demonstrated that a mutational disruption (cysteine to serine) of the invariant cysteine in the Kruppel zinc finger results in a loss of its biological activity. Presumably, the DNA-binding function of the zinc finger was lost because of the disruption of the constituents of the metal cluster. Parraga et al [29] were able to further define the role of the conserved cysteine and histidine ligands in zinc chelation and protein folding. They observed no changes in the ^{1}HNMR spectrum

upon the addition of zinc to thiol-alkylated cysteines suggesting that the finger peptide had not folded. Furthermore, the alkylation of the cysteine abolished the α helical formation usually observed after the addition of zinc to the apoprotein. The role of histidines in zinc-mediated folding was studied with ^{1}HNMR because of the well defined protons of the imidazole. These protons do not shift if the pH is lowered below 5.5 indicating that only the unprotonated imidazole group serve as zinc metal ligands. Moreover, decreasing pH also affected other methyl proton reasonances suggestive of an unfolded protein.

Structure of Zinc Fingers

Before structural data were available from 2D NMR, two different methods of computational modeling were concurrently initiated to elucidate the three dimensional structure of the zinc finger motif. Through the use of secondary structure predictions and sequence comparisons with metallo-proteins of deduced structure, structure-knowledge-based modeling by Jeremy Berg[30] proposed a zinc finger model composed of a single anti-parallel ß-sheet followed by an α helix. A separate study using interactive modeling techniques in conjunction with refined molecular dynamics simulation[31] also converged to a structure which consists of an anti-parallel β sheet and a β turn (type III) turn followed by an α helix.

Lee et al[32] were the first group to report a detailed solution structure of a single zinc finger with two dimensional NMR technique.

The structure, xfin31 of *Xenopus laevis*, was remarkably similar to the earlier predicted models (ββα motif). Of the ten final structures which represented the possible solutions to the 2D NMR distance constraints, five structures contained an α helix while the other five suggest a 3_{10} helix at the C terminus. Since xfin31, subsequent 2D NMR structures of single and double zinc fingers have revealed the structural consistency of the individual "classical" zinc finger motif : mKr2[33], ADR1[34], SWI5[35], ZFY[36], MBP-1[37] and a co-crystal complex Zif268[1].

The determinants of the structural architecture and their spacing are highly conserved in zinc fingers to form a consensus sequence $(Y/FXCX_{2,4}CX_3F/YX_5LX_2HX_{3,4}H)$,[30] where X represents any amino acid. Michael et al[38] successfully constructed a "minimalist" zinc finger where the variable amino acids (X_{nn}) were substituted with alanines to demonstrate that the metal binding and the conserved hydrophobic core residues within the motif are sufficient to determine the zinc finger folding. Competition metal binding with cobalt indicated a disassociation constant similar to other zinc finger peptides. Mutations of the hydrophobic residues to alanines or deletion of residues between the cysteines to histidines resulted in malformed aggregates of the finger protein. Mortishire-Smith et al[39] constructed a xfin31 zinc finger peptide in which the conserved aromatic residue in the second β strand was mutated to leucine. Although the mutant retained the overall folding of a canonical zinc finger structure, they observe faster proton exchange rates in the α helix suggesting that the mutant was less structurally stable than the wildtype finger.

Recent detailed 2D NMR data of zinc fingers from ADR1, a yeast transcriptional activator and ZFY, a human male-associated protein factor suggest that some structural variability can be found in the classical TFIIIA-like zinc fingers. 2D NMR spectrum analyses of two consecutive zinc fingers from ADR1, ADR1a and ADR1b, suggested dynamic structural differences that may be attributed to the number of residues between the invariant histidines[40]. Although ADR1a and ADR1b are similar zinc fingers in most respects (conserved residues of the zinc finger motif), ADR1a has four residues between the two histidine ligands while ADR1b has three. In contrast to ADR1b, 2D NMR data of ADR1a suggested that it may be found in two distinct forms, 70% major and 30% minor. The differences in resonance peaks in the two forms of ADR1a can be localized to three regions of the finger peptide including the helical domain. Concurrently, Kochoyan et al[36], [41] also reported systematic differences in 2D NMR data between consecutive zinc fingers in ZFY which also differ in their spacing between the metal binding histidines. They described the zinc fingers with respect to the number of residues between the histidines, "odd" for a three spacing segment (HX_3H) and "even" fingers for a four spacing segment (XH_4X) and proposed that the "even" fingers form "jumping linkers" between the "odd" zinc fingers.

The TFIIIA Paradox

The abundance of the classical (CC/HH) zinc fingers and their consistency in supersecondary structure among a variety of organisms

highlight their importance as a structural motif for specific DNA interaction. The modular nature of zinc fingers suggests they may utilize analogous structural elements for sequence-specific recognition. Detailed structural comparison of helix-turn-helix DNA-binding proteins from numerous sources (e.g., λ Cro. λ 434, *E. coli* CAP, etc.) revealed common structural features shared in the protein-DNA complex[1]. Previous molecular biology experiments of zinc finger alterations/swapping[42, 43, 44] and the detailed analyses of the high resolution co-crystal of Zif268[1] held the promise that the rules of zinc finger/DNA specificity can be understood and applied to novel zinc finger complexes[45, 46]. For example, each finger acting independently would "read" three bases by three positionally defined residues of the recognition helix with a predilection for guanines by the arginines. The spatial orientation of the zinc fingers relative to the DNA would also be conserved so that the position of each finger is related to the next by simple transformations and translations. Thus it would be possible to predict the cognate DNA sequence from an inspection of the zinc finger residues and to design novel zinc finger proteins to recognize any sequence of choice. Yet, recent attempts of zinc finger engineering with Sp1 from Jeremy Berg's group[47] suggest that the application of these simplistic recognition guidelines do not provide an adequate explanation for the sequence-specificity. They found that sequence-specificity requires amino acid changes in the helix at proposed non-DNA contact positions.

Ironically, these simple rules of zinc finger recognition cannot be wholly applied to the TFIIIA zinc fingers in which they were first

identified. For example, if each zinc finger of TFIIIA were in contact with 3 base pairs of DNA as suggested by the co-crystal, then all nine zinc fingers should contact a region of approximately 27 base pairs. However, the footprinting (with DNAseI and hydroxyl radical), methylation protection, and missing nucleotide studies have consistently shown that the protection region of the 5S gene extends over 45 base pairs. Furthermore, as described below, detailed hydroxyl radical protection studies have consistently demonstrated that the protection along the ICR is not continuous. This would suggest that all nine zinc fingers of TFIIIA do not make equivalent contacts in the major groove of the DNA.

Thus the intriguing paradox which confronts the TFIIIA-5S complex is this - how can a structurally defined and repetitive protein domain motif such as the zinc finger motif allow for variability in the interaction with DNA? In other words, how can the zinc fingers be alike in overall folding but be different enough to allow for one zinc finger to bind in the major groove of the DNA and while another to bind in the minor groove in an unconventional manner? What is the structural basis for the differences between the zinc fingers of TFIIIA that could account for the variability in binding and how is this structural difference related to molecular properties of sequence-specific recognition of DNA? The brief survey outlined below describes the source of these questions and demonstrates that they have not yet received a satisfactory answer.

Proposed models of TFIIIA/5S complex

Early qualitative studies with DNAse-I footprinting indicated that TFIIIA protects the 5S RNA gene at positions +45 to +97[48] Rhodes and Klug[49] attempted to relate careful analysis of the protection profile from DNAse I along the 5S gene to the structure of the DNA. The pattern of cleavage sites repeated approximately every 5 1/2 base pairs correlated with the periodicity of the appearance of guanines. They proposed that this underlying DNA structural repeat characterized an A-type DNA conformation with each finger in contact with 1/2 turn (5 1/2 base pairs) of the DNA in an alternating manner.

Fairall et al[50] probed the accessibility of the TFIIIA-5S complex with dimethylsulphate and micrococcal nuclease. Dimethylsulphate methylates the N-7 atoms of guanines in the major groove if its access is not hindered by protein. The guanines between positions 41 and 92 showed reduced methylation as a consequence of TFIIIA binding with a gradient of protection, highest at the 3' end and lowest at the 5' end. The strongest protections from methylation were located approximately every 10 base pairs. The authors proposed two models of the TFIIIA: 5S complex. One, the "wrap around" model in which the zinc fingers would be in continuous contact with the DNA in the major groove implied that each zinc finger would make equivalent contacts with the DNA. This model predicts a continuous protection pattern from DNA cleavage[23]. Another model is the "alternating" model that is consistent with Fairall's and Rhodes' experiments in which the DNA interdigitates between two successive pairs of zinc fingers with every other finger crossing the minor groove. This model proposes that the protein lies on one face of the DNA double helix, thus making contacts

to both the major and the minor groove. In contrast to the "wrap around" model, a detailed protection experiment would reveal local regions of exposed bases to DNA cleavage within the overall protected area.

The experiments of Vrana et al[51] represented a significant advancement in the understanding of how TFIIIA binds to the 5S gene. They were able to characterize the protein-DNA interaction through deletion mutations of TFIIIA in conjunction with protection studies using hydroxyl radical oxidation. Instead of a continuous contact along the internal control region (ICR), they concluded that TFIIIA contacts the 5S gene in two principal regions (+46 to +58, and +75 to +92) in the major groove, and a smaller region in the minor groove. Their model proposed that the N terminus of the protein binds with the most 3' end of the internal control region while the C-terminal residues contact the base pairs at the 5' end of the ICR (internal control region). By selectively truncating the N terminus and C-terminal fingers and comparing the footprinting maps to the wild-type profiles, they concluded that fingers at the amino terminus contribute more to the overall binding energy than the fingers at the carboxyl end. Furthermore, deletion of the first finger (finger closest to the N terminus) did not significantly alter the footprint map suggesting that the zinc fingers may not bind with equivalent affinity.

In contrast to Vrana, Hanas et al[52] made internal zinc finger deletions, the entire fourth finger and partial seventh finger of TFIIIA. Surprisingly, the finger 4 deletion mutant only protected positions +96 to +78 from DNAse I cleavage. Instead of simply shifting the

binding sites of fingers 5 through 9 up (toward the 3' direction), the protection pattern suggested that only the N terminal fingers bound to their cognate site on the ICR. Furthermore, the partial seventh zinc finger deletion showed protection consistent with a model in which only fingers 1-6 are bound. These experimental results suggested that the nine zinc fingers of TFIIIA may bind in groups of fingers and that the binding of N-terminal fingers positions the subsequent fingers in an interdependent manner.

Christensen et al[53] demonstrated that just the first three N terminal fingers of TFIIIA can fully account for the footprint pattern at positions +77 to +96 in the ICR. Liao et al[54] recently demonstrated that the first three fingers of TFIIIA are sufficient for sequence-specificity and high affinity. Moreover, quantitative gel retardation assays suggested that the majority of the free energy of binding resides in the first three zinc fingers (5.6 $\pm$0.9 nM disassociation constant) when compared to the entire protein (2.2 $\pm$0.4 nM disassociation constant). In addition, Liao et al. showed that constructs of fingers 1-2 or fingers 2-4 were not capable of sequence-specific contacts.

Churchill et al[55] conducted hydroxyl radical footprinting experiments to probe the TFIIIA-5S complex. Hydroxyl radicals cleave at every backbone phosphate and provide a more detailed protection pattern. They also observed a periodicity of exposed regions every 10-11 base pairs with the protected region of the non-coding strand offset by 3 base pairs from the coding region. This protection pattern was suggestive of a model in which the fingers bound to only one face

of the right-handed DNA. Minor groove interactions by the linkers are implied in this "alternating" model of binding.

In a review article, Jeremy Berg[56] proposed another model of the TFIIIA-5S complex which represented a departure from either the "alternating" of "wrapping around" model. Supported by the published footprinting data and closer inspection of the primary sequence of TFIIIA, Berg suggested that the fingers of TFIIIA are grouped into two domains which contact two corresponding protected regions of the ICR with finger 6 crossing the minor groove. He hypothesized that because the 5S gene is closer to the classical B conformation, it must undergo significant bending upon binding if these two protected regions are brought closer together in their interaction with TFIIIA. This hypothesis from the modeling has a basis in an earlier experimental study by Schroth et al[57] who demonstrated with electrophoretic methods that the 5S DNA bends upon complexation .

Early characterizations of the TFIIIA-5S complex suggested that a single zinc finger made contact to every 5.5 base pairs. However, Lee et al proposed that the zinc finger of XFIN31 binds to 4 base pairs[58]. In an elegant experiment to determine which residues of the zinc finger helix contacts the DNA, Nardelli et al[42] conducted gel retardation studies of two proteins, Sp1, a SV40 transcription factor and KROX-20, a serum-inducible transcription activator. Each DNA-binding protein contains 3 zinc fingers, but binds to only 9 base pairs. By switching proposed sequence-specific amino acids between the Sp1 and Krox-20 zinc fingers, they were able to demonstrate specific binding of hybrid zinc fingers (Krox-20/Sp1) to hybrid DNA probes

suggesting that the modularity between the zinc fingers to its corresponding DNA element is confined to 3 bases.

Recently, a high resolution (2.1 Å) crystal structure of three zinc fingers in a complex with its cognate DNA was obtained[1]. The results of the interpretation of the crystal data are far-reaching. The zinc finger-DNA contacts in the crystal complex showed a 3 base-pair periodicity to one zinc finger as suggested by Nardelli. With simple rotations and transformations, the three zinc fingers follow the major groove of the DNA arranged in a semicircular "C" with respect to the DNA helical axis. The authors also described which residues of the finger consistently contacted the DNA in B-like conformation with local distortions and bending. In contrast to the prevailing notion that there is no "code" between the amino acid residues for base pairs[59], the arginines in the helical region of these zinc fingers show a predilection for guanines. Furthermore, their direct interaction is stabilized by aspartic acids in the helix. Finally, the juxtaposition of the imidazole ring of the first invariant histidine (zinc-binding residue) relative to the phosphate backbone of the DNA suggests that it may participate in DNA-binding.

In contrast to the co-crystal data and zinc finger swapping experiments which support the "wrapping around" model, Fairall et al[60] re-examined earlier protection data with DNAse I by considering the molecular details of interaction between the asymmetric binding of DNAse I to the minor groove of DNA. TFIIIA showed overall protection from DNAse I at base pairs 45 to 97 on the coding strand and 45 to 95 on the noncoding strand. While the coding strand was

almost entirely protected, two regions of the noncoding strand remained unprotected from DNAse I, bp's 60-63 and 73-76. In addition, the region of bp's 60-63 was not a site for cleavage by DNAse I in the unprotected DNA suggesting that protein binding induced some conformational change in the DNA. The author then subtracted the DNAse I binding region from the overall protected areas and concluded that TFIIIA binds to the 5S gene in a tripartite manner, with three major areas of protection, bp's 49-58, 65-71, and 78-88.

Hayes et al[61] conducted missing nucleoside protection studies to determine the delineation of individual TFIIIA zinc fingers and their base-pair contacts on the 5S gene. After introducing random gaps in the radiolabelled DNA by hydroxyl radical cleavage, the DNA mixture was incubated with excess TFIIIA protein and electrophoresed in a non-denaturing gel. The gaps that occurred in the bound and free DNA were determined by denaturing gel electrophoresis. By comparing nucleosides of the bound and unbound DNAs, the delineation of finger contact can be elucidated. Because the missing nucleoside experiment allows for the evaluation each base pair's importance to binding, a detailed picture of the TFIIIA-5S complex emerged where the fingers at the two ends of the complex bind in the same manner, while the middle fingers bind in a different fashion. Based on the experimental results, Hayes and Tullius proposed a new "intermediate" model of TFIIIA-5S complex. This model combines aspects of the "alternating" and "wrapping" model described earlier with finger 1-3 and fingers 7-9 bind with the C and A box respectively in a continuous "wrapping" around manner reminiscent of the Zif268 complex. However, they

proposed that fingers 4-6 bind in a novel manner, parallel to the DNA helical axis.

In another approach to delineate the base pair boundaries of individual finger contacts, Clemens et al[62] employed DNAse I footprinting, methylation interference, and primer extension in conjunction with differential TFIIIA footprinting constructs (e.g., ZF1-3, ZF1-4, ZF1-5, ZF1-6, ZF1-8, ZF1-9) to map the individual TFIIIA zinc finger protection regions of the 5S gene. Their results supported the "intermediate" model where fingers 1-3 bind to the C box, fingers 7-9 bind to the A box, and fingers 4-6 span these two promoter regions. In order to achieve a more detailed protection profile, Hayes and Clemens[63] analyzed the differential zinc finger constructs this time with hydroxyl-radical cleavage.

The model of the TFIIIA/5S complex that has emerged from these experiments is unlike the previous "wrapping around or alternating" model. Instead, these experiments suggest that the binding of TFIIIA to the DNA occurs in three groups corresponding to the A, C, and IE box[64] of the 5S gene. The nine zinc fingers of TFIIIA are organized in a tripartite fashion with fingers 1-3 and 7-9 binding in modular manner, analogous to the Zif268 complex. However, these authors contend that zinc fingers 4-6 bind in a parallel orientation to the central axis of the double helix. This model implies that the mode of binding which was thought to be conserved for all classic (CC/HH) zinc fingers does not apply for certain zinc fingers of TFIIIA. Instead of binding in the major groove, the new model requires that certain zinc fingers cross and may even bind to the minor groove. Furthermore, in order to account

for the total protein-protected area of the DNA and the differential mode of binding for each group of zinc fingers, the number of base pairs contacted by each zinc finger varies.

CHAPTER 2

Introduction

Zinc fingers can form independent folding domains upon the addition of zinc metal allowing for the folding of isolated individual fingers in solution. The structural components that determine the characteristic anti-parallel ß sheet at the N-terminus and the α helix at the C-terminus of a correctly folded zinc finger include the conserved cysteines and histidines for metal binding and the aromatic and hydrophobic residues for the formation of a small hydrophobic core. The spacing between these conserved elements of the primary structure may also influence the folding of zinc fingers and hence their molecular properties. Four example, the nine zinc finger of TFIIIA can be defined in terms of the number or residues between the conserved histidine ligands. Fingers 3, 6, and 8 have a four residue spacing (or "**even**" zinc finger) whereas fingers 1, 2, 4, 5, 7, and 9 have the typical three residue spacing (or "**odd**" zinc finger).

Figure 2

The primary sequence of TFIIIA. The amino acid sequence is aligned according to the nine zinc fingers domains. Conserved zinc ligands are boldfaced. The hydrophobic core residues are italicized. Each finger (I-IX) is designated "even" or "odd" with respect to the number of intervening residues between the histidines

M G E K A L P V V Y K
R *Y* I **C** S F A D **C** G A A *Y* N K N W K *L* Q A **H** L C K • **H** T G E K (I) "ODD"
P *F* P **C** K E E G **C** E K G *F* T S L H H *L* T R **H** S L T • **H** T G E K (II) "ODD"
N *F* T **C** D S D G **C** D L R *F* T T K A N *M* K K **H** F N R F **H** N I K I C (III) "EVEN"
V *Y* V **C** H F E N **C** G K A *F* K K H N Q *L* K V **H** Q F S • **H** T Q Q L (IV) "ODD"
P *Y* E **C** P H E G **C** D K R *F* S L P S R *L* K R **H** E K V • **H** A • • • (V) "ODD"
G *Y* P **C** K K D DS**C** S • • *F* V G K T WT*L* Y K **H** V A E C **H** Q D • • (VI) "EVEN"
L A V **C** • D V • **C** N R K *F* R H K D Y *L* R D **H** Q K T • **H** E K E R T (VII) "ODD"
V *Y* L **C** P R D G **C** D R S *Y* T T A F N *L* R S **H** I Q S F **H** E E Q R (VIII) "EVEN"
P *F* V **C** E H A G **C** G K C *F* A M K K S *L* E R **H** S V V • **H** D P E K (IX) "ODD"
R K L K E K C P R P K R S L A S R L T G Y I P P K S K E K
N A S V S G T E K T D S L V K N K P S G T E T N G S L V L
D K L T I Q.

Several other zinc finger proteins have arrangements of "odd" and "even" fingers.

Figure 3

Compilation of zinc finger proteins with "even" and "odd" zinc fingers. The name of the protein and its reference are italicized. The "even" zinc fingers are denoted by boldface.

cf2 (Shea et al, Genes Dev.(1990) 4:1128-1140)
H K C P D C P K T F K T P G T L A M H R K I H(odd)
Y T C S Y C G K S F T Q S N T L K Q H T R I H(odd)
F H C G Y C E K S F S V K D Y L T K H R T T H(odd)
Y T C P Y C D K R F T Q R S A L T V H T T K L H(**even**)

cid (Orenic et al, Genes Dev.(1990) 4:1053-1067)
T N C H W R R C R I E F I T Q D E L V K H I N N D H(**even**)
F V C R W E D C TRGEKPF K A Q Y M L V V H M R R H(odd)
H K C T F E G C F K A Y S R L E N L K T H L R S H(odd)
Y T C E Y P G C S K A F S N A S D R A K H Q N R T H(**even**)
Y I C K A P G C T K R Y T D P S S L R K H V K T V H(**even**)

evi1 (Morshita et al, Cell (1988) 54:831-840)
H R C E D C D Q L F E S K A E L A D H Q K F P C(**even**)
Q D C K E C D R V F P D L Q S L E K H M L S H(odd)
Y K C D Q C P K A F N W K S N L I R H Q M S H(odd)
Y E C E N C A K V F T D P S N L Q R H I R S Q H(**even**)
H A C P E C G K T F A T S S G L K Q H K H I H(odd)
F I C E V C H K S Y T Q F S N L C R H K R M H(odd)
I K C K D C G Q M F S T T S S L N K H R R F C(odd)
Y T C R Y C G K I F P R S A N L T R H L R T H(odd)
Y R C K Y C D R S F S I S S N L Q R H V R N I H(**even**)
F K C H L C D R C F G Q Q T N L D R H L K K H(odd)

gli (Kinzler et al, Nat.(1988) 332:371-374)
T D C R W D G C S Q E F D S Q E Q L V H H I N S E H(**even**)
F V C H W G G C SRELRPF K A Q Y M L V V H M R R H(odd)
H K C T F E G C R K S Y S R L E N L K T H L R S H(odd)
Y M C E H E G C S K A F S N A S D R A K H Q N R T H(odd)
Y V C K L P G C T K R Y T D P S S L R K H V K T V H(**even**)

kr5 (Chowdhury et al, Nuc. Acids Res.(1988) 16:9995-10011)
F Q C N E C K K T F T Q S S S L T V H Q R I H(odd)
Y K C N Q C G K A F S D G S S F A R H Q R Y H(odd)
Y E C P E C G K A F I Q N T S L V R H W R Y Y H(**even**)
F D C I D C G K A F S D H I G L N Q H R R I H(odd)
Y T C E V C H K S F R Y G S S L T V H Q R I H(odd)
Y E C E I C R K A F S H H A S L T Q H Q R V H(odd)
F K C K E C G K A F R Q N I H L A S H W R I H(odd)
F E C G E C G K S F S I S S Q L A T H Q R I H(odd)
F E C K V C R K A F R Q N I H L A S H W R I H(odd)
F E C G E C G K S F S I S S Q L A T H Q R I H(odd)

zg17 (Nietfeld et al, J. Mol. Biol.(1989) 208:639-659)
I S C S E C G K C F I K S S E L T V H Q M T H(odd)
Y S C S E C G K C F A S L S H L R V H Q K I H(odd)
F S C S E C G K C F L N R G S L V R H H R T H(odd)
F F C S E C G K R F A A S S D L R V H R R T H(odd)
F S C S E C E K R F L N P W S L V R H Y R T H(odd)
F S C S E C G K C F A R S S D L T V H R R R S H(**even**)
F S C S E C G K C F T S S S E L T V H L R T H(odd)

zg26 (Nietfeld et al, J. Mol. Biol.(1989) 208:639-659)
F D C T E C G K S F T D L K T L Q H H Y K I H(odd)
F I C A E C G K G F N Q K T T L L N H S K I H(odd)
F P C T E C G K C F T E R K T L L N H N K I H(odd)
F I C A E C G I G F T Q K T T L L N H S K I H(odd)
F P C T E C G K C F T E R K T L L N H N K I H(odd)
F I C A E C G I G F T Q K T T L L N H S K I H(odd)
F P C T E C G K C F A E K K T L Q N H N K I H(odd)
F T C T D C G K S F T Q R T S L Q N H V K I H(odd)
F T C T E C G K S F S E K K T L R E H N K I H(odd)
F T C T Y C G K S F S Q R I S L Q N H F K I H(odd)
F S C T E C G K C F T I K S T L Q S H L K R T H(**even**)
F T C T E C G K S F T K K K I L L K H N K I H(odd)

zg62 (Nietfeld et al, J. Mol. Biol.(1989) 208:639-659)
F I C T E C G K C F S E K R T L K H H I R T H(odd)
F I C T D C G K C F S F E I C L N R H Y K T H(odd)
F I C T E C G K S F S D K S R L R V H H R S H(odd)
F T C T D C G K C F S V K S I L N H H R Q A I H(**even**)
F I C T E C G K G F A S K H Y L H G H K R T H(odd)
F V C T E C G K G F A S N Y Y L H V H K R T H(odd)

zo72 (Knochel et al, PNAS(1989) 86:6097-6100)
F P C S E C G K C F I N Q S T L A R H Y R T H(odd)
Y P C S E C G K C F A S S T Y L R D H R R I H(odd)
S S C S E C G K Y F L N C W S L A R H H R T H(odd)
Y S C S E C G K S F A I S S D L A G H R R R T H(**even**)

Adr1 (Hartshorne et al, Nat.(1986) 320:283-287)
F V C E V C T R A F A R Q E H L K R H Y R S H(odd)
Y P C G L C N R C F T R R D L L I R H A Q K I H(**even**)

p43 (Joho et al, Cell(1990) 61:293-300)
L R C P A A G C K A F Y R K E G K L Q D H M A G H(odd)
W K C G I K D C D K V F A R K R Q I L K H V K R H(odd)
L S C P T A G C K M T F S T K K S L S R H K L Y K H(**even**)
L K C F V P G C K R S F R K K R A L R R H L S V H(odd)
S V C D V P G C S W K S S S V A K L V A H Q K R H(odd)
Y R C S Y E G C Q T V S P T W T A L Q T H V K K H(odd)
L Q C A A C K K P F K K A S A L R R H K A T H(odd)
L P C P R Q D C D K T F S S V F N L T H H V A R K L H
H R C P H S G C T R S F A M R E S L L R H L V V H(odd)

zo71 (Knochel et al, PNAS(1989) 86:6097-6100)

Y S C N E C H E Y L I H K R D F G K H Q M T H(odd)
F S C S K C G K C F A F L S D L T V H R R I H(odd)
F S C S E C G K G F T R P N A L I I H H R T H(odd)
F S C S E C G K C F S K Q S S L V H H Q R T H(odd)
F C C S E C D K C F A S S S E L N I H Q R T H(odd)
F S C S E C G K C F T N H S H F A H H Q M I H(odd)
F C C S K C G K C F A S S S D L T F H R R T H(odd)
F S C S E C G K C F S N H S H L A R H Q M I H(odd)
F C C S E C G K C F S S S S G L T A H Q Q R T H(**even**)
F S C S A C G K C F S N R S H L I R H Q M I H(odd)
F S C F E C R K C F S N P S N L A R H Q M T H(odd)
F S C S E C G K C F A S S S D L T F H H R T H(odd)
F S C S E C G K C Y S K K S S L V H H Q R T H(odd)
F S C S K C D K C F A S S S E L N I H Q R T H(odd)
F S C S E C G K C F T N R S Q L S R H Q M I H(odd)
I S C P E C E E C F V S S S Q L T A H Q Q Q A H(**even**)
F S C L E C G K C F S N R S N F A R H Q M I H(odd)
F S C S E C R K G F S N Q S S L A R H Q M T H(odd)
F S C S E C G K R F S N Q S H L A R H Q M I H(odd)
F S C S E C A K G F S N Q S G L A R H Q M T H(odd)
F A C S E C G K C F A S S S K L T A H Q R T H(odd)

s03679 (Chowdhury et al, Nuc. Acids. Res.(1988) 16:9995-10011)

P Y E C P E C G KAF I Q N T S L V R H W R Y Y H(**even**)
F D C I D C G K A F S D H I G L N Q H R R I H(odd)
Y T C E V C H K S F R Y G S S L T V H Q R I H(odd)
Y E C E I C R K A F S H H A S L T Q H Q R V H(odd)
F K C K E C G K A F R Q N I H L A S H W R I H(odd)
F E C G E C G K S F S I S S Q L A T H Q R I H(odd)
F E C K V C R K A F R Q N I H L A S H W R I H(odd)
F E C G E C G K S F S I S S Q L A T H Q R I I H(**even**)
Y E C K E C R K T F I Q I G H L N Q H K R V H(odd)

S.cere_tfiiia (Woychick et al, PNAS(1992) 89:3999-4003)

Y F C D Y D G C D K A F T R P S I L T E H Q L S V H(**even**)
F Q C D K C A K S F V K K S H L E R H L Y T H(odd)
F Q C S Y C G K G V T T R Q Q L K R H E V T H(odd)
F I C P E E G C N L R F Y K H P Q L R A H I L S V H(**even**)
L T C P H C N K S F Q R P Y R L R N H I S K H H(**even**/odd)
Y Q C T F A G C C K E F R I W S Q L Q S H I K N D H(odd)
L K C P I C S K P C V G E N G L Q M H M I I H(odd)
W K C H I C P DMS F S R K H D L L T H Y G S I H(**even**)
Y R C F Y N C S R T F K T K E K Y E K H I D K H(odd)

zfy (Palmer et al, PNAS(1989) 87:1681-1685)

Y P C M I C G K K F K S R G F L K R H M K N H(odd)
Y H C T D C D Y T T N K K I S L H N H L E S H(odd)
I E C D E C G K H F S H A G A L F T H K M V H(odd)
H K C K F C E Y E T A E Q G L L N R H L L A V H(**even**)
H I C V E C G K G F R Y P S E L R K H M R I H(odd)
Y Q C Q Y C E Y R S A D S S N L K T H I K T K H(**even**)
F K C D I C L L T F S D T K E V Q Q H T L V H(odd)
H Q C L H C D H K S S N S S D L K R H V I S V H(**even**)
H K C E M C E K G F H R P S E L K K H V A V H(odd)
H Q C R H C D F K I A D P F V L S R H I L S V H(**even**)
F R C K R C R K G F R Q Q N E L K K H M K T H(odd)
Y Q C E Y C E Y S T T D A S G F K R H V I S I H(**even**)
H R C E Y C K K G F R R P S E K N Q H I M R H H(**even**/odd)

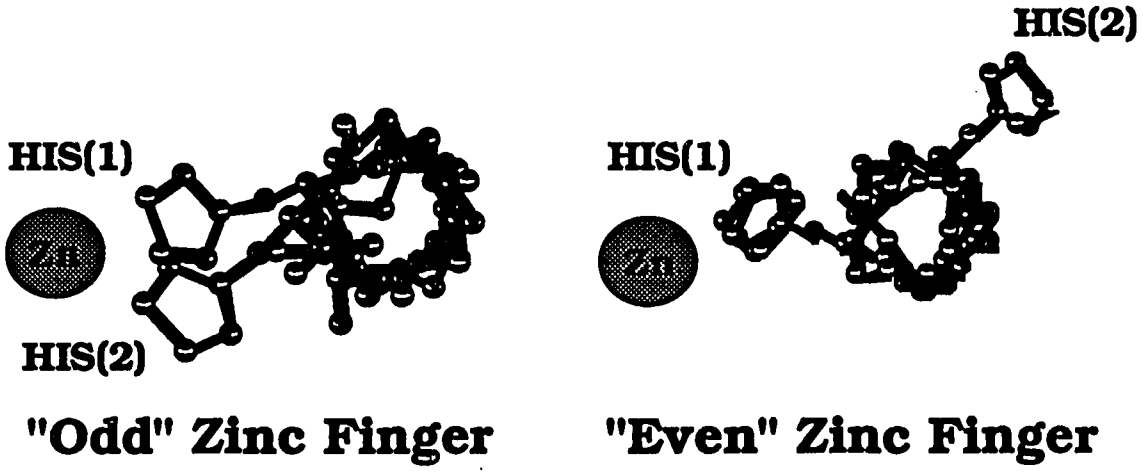
zo22 (Koester et al, EMBO J.(1988) 7:1735-1741)

H S C S Q C G K C F S S S S D L L A H R R Q S H(**even**)
F S C S E C G K C F S F R S R L I D H Q R T H(odd)
F C C F Q C G K S F S V R S R F L D H R R T H(odd)
F S C L E C G K C F L F R S R L L E H Q R T H(odd)
F S C L K C G K C F S V R S R L K D H Q R T H(odd)
F S C L E C G K S F S F R P C L I D H Q R T H(odd)
F S C F Q C G K C F S F Q S R L I N H Q R T H(odd)
F S C S E C G K S F S N Q S C L R V H Q R T H(odd)
Y S C S E C G K S F V T S S Q L A V H R R R T H(**even**)
F S C S E C G K C F S N Q S C L R V H Q R T H(odd)
F S C S E C G K S F V T S S K L A S H Q R Q T H(**even**)
V S C S E C G K C F T R K R S L K V H F K I H(odd)

A consequence of metal induced folding for zinc finger peptides is the formation of an α helix (see above, 1.2). The zinc-finger helix has two important functions : its two invariant histidines form metal ligands to zinc and three of its residues participate in specific DNA recognition[1]. α helices form stable secondary structures with the backbone carbonyl oxygen of each residue making a near linear hydrogen bond to each backbone amide hydrogen. Consequently the physical forces that maintain α helices create a highly repetitive structure. The right handed α helix has 3.6 residues per turn and a translation of 1.5 Å per residue (pitch of 5.41 Å/turn)[65]. Therefore, residues on the same side or "face" of an α helix are separated by three amino acids. In this manner, the conserved three residue spacing between the invariant histidines positions the imidazole sidechains of each histidine on the same face of the α helix.

Figure 4

Consequences of an α helical structure and metal binding for "odd" HX3H spacing finger, right, and for "even" HX4H zinc finger, left. Only the sidechain of the two histidine metal ligands, HIS(1) and HIS(2), are shown. The helical axis is perpendicular to the plane of the paper.



This orientation is appropriate for the positioning of the HIS sidechain as chelating ligands of the zinc, so that the zinc finger α helix forms a complementary arrangement with the major stabilizing component of the protein, the metal cluster.

In contrast, the four residue spacing of "even" fingers appears incompatible with simultaneous zinc chelation and would disrupt this complementarity without a modification of the α helix. Because the strong backbone hydrogen bonding network of helices involves residues separated by three spacing, local perturbations within the helix may affect other residues along the chain.

In order to evaluate the relationship between the structural constraints imposed by the difference of spacing between the "even" and "odd" zinc fingers and their properties related to DNA binding, I have carried out molecular dynamics simulations of the structural properties of single "odd" (xfin31 and the fourth finger of TFIIIA) and "even" (third finger of TFIIIA) zinc fingers in an aqueous environment. Molecular dynamics simulations allow us to probe at the molecular level the effect of the difference in spacing between "even" and "odd" zinc fingers on their dynamic properties such as flexibility. Understanding the mechanisms which result in conformational flexibility or time-dependent structural heterogeneity of "even" fingers compared to "odd" fingers may provide a basis for understanding their organization in multi-zinc finger proteins such as TFIIIA and may suggest the functional importance of such an organization for specific recognition of long stretches of DNA.

To make possible the use of molecular dynamics algorithms for the simulation of zinc-finger systems, I incorporated new Zn^{+2} force fields from *ab initio* quantum mechanical methods into the AMBER3revA[66] molecular dynamics program. Since there are presently no structural data on TFIIIA zinc fingers, the third finger TFIIIA(iii), an "even" zinc finger, and fourth finger, TFIIIA(iv), an "odd" zinc finger, were constructed by homology modeling from the known structure of the xfin31 zinc finger and refined with energy minimization and extensive simulation.

Methods

2.2.1. Homology Modeling

The initial starting conformation of the zinc finger peptide xfin31[32] was obtained from the 2D NMR structural coordinates available in the Brookhaven protein database (PDB). Because detailed structural data for TFIIIA(iii) and TFIIIA(iv) are not presently available, the initial structure of each zinc finger from TFIIIA was modeled with the Quanta (release 3.2) program (Polygen Corporation, Waltham, MA). The mainchain and α carbon atoms of xfin31 were used as a structural template on which to construct the TFIIIA zinc fingers. The conservation pattern of certain amino acids among all the CC/HH zinc fingers and their spacing in a very large number of zinc finger proteins as evidenced in Figure 3 make this class of zinc fingers ideal candidates for homology modeling[67]. Replacement of xfin31 amino

acids with TFIIIA(iii) and TFIIIA(iv) residues were done whenever possible including equivalent Cβ and Cγ atoms, (see Figure 5).

Figure 5

Homology modeling of TFIIIA zinc fingers TFIIIA(iii) and TFIIIA(iv). TFIIIA are shown in homology alignment with xfin31 zinc finger. The conserved residues are boldfaced.

TFIIIA(iii)	N	**F**	T	**C**	D	S	D	G	**C**	D	L	R	**F**	T	T	K	A	N	**M**	K	K	**H**	F	N	R	F	H	N	I
TFIIIA(iv)	V	**Y**	V	**C**	H	F	E	N	**C**	G	K	A	**F**	K	K	H	N	Q	**L**	K	V	**H**	Q	F	S	•	H	T	Q
xfin31	•	**Y**	K	**C**	•	G	L	•	**C**	E	R	S	**F**	V	E	K	S	A	**L**	S	R	**H**	Q	R	V	•	H	K	N

When replacements with equivalent sidechain atoms were not possible, the initial orientation of the sidechains were derived from a standard residue topology file available in the Quanta program. The CYS loop (residues between the two conserved cysteine ligands) of xfin31 was extended in order to accommodate the two extra residues of TFIIIA fingers while conserving the β-turn. The Ψ torsion angle of the last histidine ligand was modified in TFIIIA(iii) "even" in order to "swing" the imidazole ring around to face the zinc ion and to position the Nε atom near its coordination site. All finger structures were expressed in an all-atom representation.

Structural refinement of the models was performed with energy minimization. After an initial energy minimization of the zinc finger structure alone, (see below), each finger was solvated in a water bath (TIP3 description) derived from a Monte Carlo simulation provided in the AMBER program package. Any water molecule with a hydrogen within 2.0 Å or oxygen within 2.8 Å from a solute atom was deleted from the system. Any water at a distance greater than 6.0 Å from any solute atom along each of the three cartesian axes was also deleted. The procedure was repeated for three single zinc finger systems: the original xfin31, TFIIIA(iii) representing the solvated third finger of TFIIIA which is of the "even" type, and TFIIIA(iv) representing the solvated fourth zinc finger in TFIIIA and is of the "odd" type. All resulting systems had the geometric configuration of a box with the dimensions as described in Table 1.

Table 1. Molecular systems used in Molecular Dynamics simulations with Periodic Boundary Conditions (PBC).

	Protein		Water		PBC Dimensions(Å^3)		
	Atoms	*Residues*	*Atoms*	*Residues*	*X*	*Y*	*Z*
xfin31	427	28	2745	915	37.81	32.99	29.32
TFIIIA(iii)	480	32	3213	1071	41.14	35.74	29.63
TFIIIA(iv)	473	31	2913	971	42.86	31.90	28.88

Zinc Ion Parameters

In order to fully represent the dynamics of the system, the interaction of the components of the zinc cluster (zinc ion to imidazole group of histidine and thiol group of cysteine) was included in the simulations. These calculations were performed with a set of parameters derived by Dr. David Garmer (Department of Physiology and Biophysics, Mount Sinai School of Medicine) from fits to results from *ab-initio* quantum chemical calculations of Zn-clusters. Early attempts (data not shown) with non-bonded parameters for the zinc ion were found to be inappropriate by failing to reproduce the tetrahedral geometry and ligand distance of the zinc coordination complex determined from EXAFS[27, 26] for TFIIIA zinc fingers. Therefore final parameters were constructed to mimic covalent zinc ion to ligand interactions and were represented by a harmonic force field. The novel zinc-ion force field was implemented into the AMBER3revA force field with the PREP module. In addition, the zinc-coordinated histidines (HIZ) and zinc coordinated cysteines (HIZ) were also modified to represent the quantum chemical results for the zinc metal cluster.

Table 2. Zn^{+2} covalent force field. The parameter file "parm89a" which contains the force fields for AMBERrev3a was modified with the following zinc ion parameters: charge , 0.77; molecular weight, 65.38; the van der Waals parameters, (Ro= 0.44 E_O=0.27).

Parameter	Force Constant (kcal/mol)	Value (distance, angle)
Zn-N	73.5	2.093
Zn-S	110.0	2.400
S-Zn-S	15.0	169.5
S-Zn-N	25.3	130.1
N-Zn-N	15.7	104.8
Zn-S-C	23.7	104.4
Zn-N-CR	21.4	127.1
Zn-N-CW	21.4	128.1

Table 3. Modified charges for the histidine ligands (HIZ)

ATOM	ORIGINAL CHARGE	NEW CHARGE
N	-0.4630	-0.4630
HN	0.2520	0.2520
CA	0.0350	0.0350
HA	0.0480	0.0480
CB	-0.0980	-0.0980
HB1	0.0380	0.0570
HB2	0.0380	0.0570
CG	-0.0320	0.0050
ND1	-0.1460	-0.1100
HND	0.2280	0.2600
CE1	0.2410	0.1890
HE	0.0360	0.0860
NE2	-0.5020	-0.3220
CD2	0.1950	0.1010
HD	0.0180	0.0730
C	0.6160	0.6160
O	-0.5040	-0.5040

Table 4. Modified charges for the cysteine ligands (CYZ)

ATOM	ORIGINAL CHARGE	NEW CHARGE
N	-0.4630	-0.4630
HN	0.2520	0.2520
CA	0.0350	0.0350
HA	0.0480	0.0480
CB	-0.0600	0.0090
HB1	0.0380	0.0385
HB2	0.0380	0.0385
SG	0.8270	-0.7360
C	0.6160	0.6160
O	-0.5040	-0.5040

Minimization Procedure

All the molecular mechanics and molecular dynamics simulations were carried out with the AMBER3revA program.The structure of the protein in the absence of any water molecule was first energy-minimized with a distance-dependent dielectric constant (ε=r) in the electrostatic term of the potential energy function. The cutoff distance for non-bonded calculations was set to 15 Å, and the non-bonded list was updated every 100 timesteps. A steepest descent (first derivative) minimization algorithm was applied and convergence was established to 0.01 kcal/mol. The energy minimized zinc finger was then solvated as described previously. The atoms of the zinc finger were then positionally constrained while energy minimization was performed for the solvent component of the system alone with the steepest descent algorithm and periodic boundary conditions. The dielectric constant was set to 1.0 and the non-bonded cutoff distance was set to 12 Å. Minimization was completed for 4000 steps and halted prior to convergence. Then the entire system (zinc finger with waters) was minimized for 4000 steps with no applied constraints on any atom.

Molecular Dynamics

The method of molecular dynamics (MD) rests on a computer simulation that solves the classic equations of motion to determine the positions of each atom as a function of time. The forces acting upon each individual atom are derived from a predefined interatomic

potential function which includes: bond distance, bond angle, torsional angle, improper torsional angle, a 6-12 Lennard-Jones parameter for van der Waals interaction, and electrostatic interaction terms. While spectroscopy provides dynamic information for a limited number of atoms and crystallography provides information of the static structure of all atoms, the advantage of MD simulations is the possibility for investigating the motion of all atoms[68, 69].

The force fields for AMBERrev3a were derived with an emphasis on an accurate representation of the electrostatic properties of the molecules. This was accomplished by fitting partial charges of the atoms to the electrostatic potential determined from *ab initio* quantum mechanical calculations. These models were then tested by comparing the hydrogen-bonded structures and energies as well as conformational energies and vibrational energies with experimental data derived from spectroscopy as well as accurate *ab initio* calculations[70, 71]. These force fields were then implemented into the AMBERrev3a software. The form of the AMBER force field potential was as follows:

Figure 6

AMBER force field equation. The bond and angle force constants are represented by K_r and K_θ respectively. The equilibrium bond length and bond angle are describe by R_{eq} and θ_{eq}. The dihedral angles, ϕ, are represented by a truncated Fourier series potential (V is the barrier height, η is the periodicities of the torsion and γ is the dihedral phase shift angle). The terms of the non-bonded term are the distance between the two atoms (ij), rij, the van der Waals radius R^*_{ij} and van der Waals well-depth ε^*_{ij}. Hydrogen bonds are represented by a 10-12 term in which the C_{ij} and D_{ij} parameters are coefficient specific for that atom pair. The electrostatic term is proportional to the product of the point charges and inversely proportional to the dielectric constant.

$$
\begin{aligned}
V_{\text{total}} &= \sum_{\text{bonds}} K_r\left(r - r_{eq}\right)^2 + \sum_{\text{angles}} K_\theta\left(\theta - \theta_{eq}\right)^2 \\
&+ \sum_{\text{dihedral}} \sum_{\eta} \frac{V_\eta}{2}\left[1 + \cos(\eta\phi - \gamma)\right] \\
&+ \frac{1}{VDW_{\text{scale}}} \sum_{j=1}^{\text{atoms}} \sum_{i>j}^{\text{atoms}} \varepsilon^*_{ij}\left[\left(\frac{R^*_{ij}}{r_{ij}}\right)^{12} - \left(\frac{R^*_{ij}}{r_{ij}}\right)^{6}\right] \\
&+ \sum_{j=1}^{\text{Hbonds}} \sum_{i>j}^{\text{Hbonds}} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right) \qquad (1) \\
&+ \frac{1}{EEL_{\text{scale}}} \sum_{j=1}^{\text{atoms}} \sum_{i>j}^{\text{atoms}} \frac{q_i q_j}{\varepsilon r_{ij}},
\end{aligned}
$$

where $\varepsilon^*_{ij} = \sqrt{\varepsilon^*_i \varepsilon^*_j}$ (2) and $R^*_{ij} = R^*_i + R^*_i$. (3).

Molecular dynamics simulation of each solvated individual zinc finger was done with velocity scaling (constant temperature), constant volume, and constant number of atoms. The system was weakly coupled (τ=0.1) to an adiabatic heat bath at a temperature of 300° K. The temperature fluctuation allowed in the constant temperature runs was kept at 5° K. The temperature relaxation constant was set to 0.1 to avoid any abrupt changes in the system should the temperature exceed the allowed window temperature. Periodic Boundary Conditions (PBC) were applied to the system with isotropic position scaling to maintain constant volume. The non-bonded cutoff was set to 12.0 Å, and the non bonded list was updated every 100 steps (1 femtosecond per step). All bonds involving hydrogens were constrained using the (SHAKE) option in order to reduce computational time. Coordinates were saved every 0.1 picosecond. The dielectric constant was kept at 1.0. Two harmonic bonds between the backbone atoms of residue 3 to 14 and 5 to 12 (characteristic of all classic zinc fingers of this family[72] were used to maintain the integrity of the anti-parallel ß-strand.

2.3. Results

2.3.1 Molecular Dynamics Trajectories Analyses

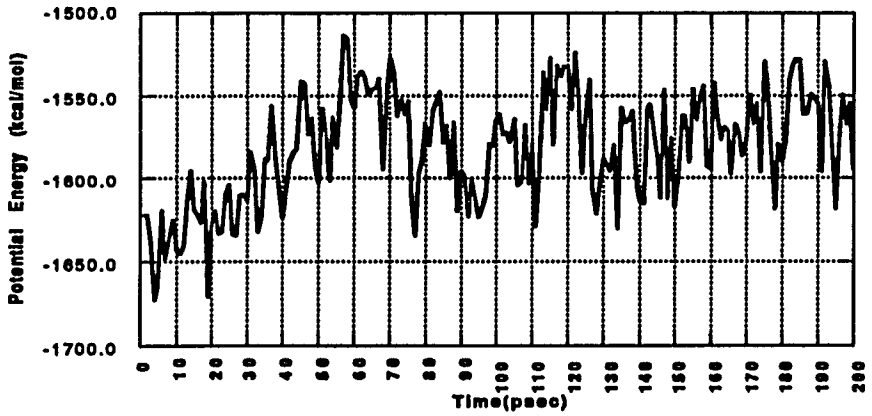
Molecular dynamics trajectories for the three individual zinc finger simulations are shown in Figures 7 and 8. The potential energy trajectories plotted as a function of time are shown in Figures 7a, b, c -

TFIIIA(iii) "even," TFIIIA(iv) "odd," and xfin31, respectively. The RMS deviation of each zinc finger at a specified timepoint is shown relative to the initial MD structure (fully minimized solvated structure before the onset of molecular dynamics). All three zinc fingers had an initial large increase in RMS by 10 psec followed by a slower upward drift of RMS. The convergence of the RMS trajectories of all three zinc fingers coincided with the same time period in which their potential energy converged. In general, the analyses of the MD trajectories suggest that the extended simulation of all three solvated zinc fingers with the newly developed zinc ion force field resulted in dynamically equilibrated zinc fingers. This conclusion is further supported by the comparison of the results from simulation to available data for xfin31 from 2D NMR measurements.

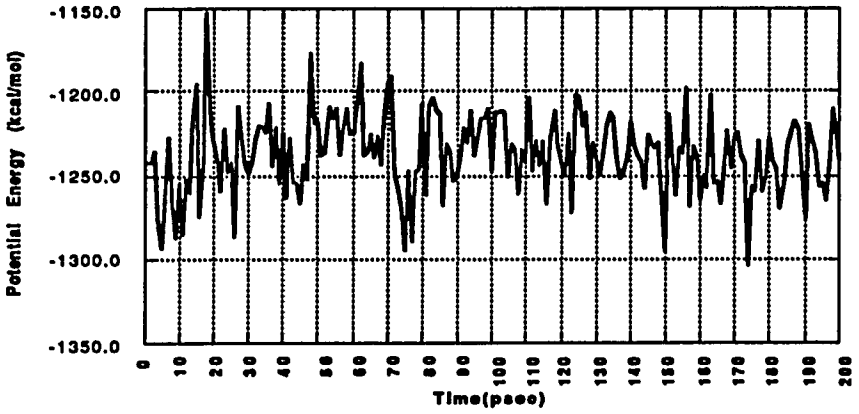
Figure 7

The potential energies trajectories of the fully solvated zinc finger simulations are shown as a function of time: a) TFIIIA(iii) "even" zinc finger of TFIIIA, b) TFIIIA(iv) "odd" zinc finger, and c) xfin31 "odd" zinc finger.

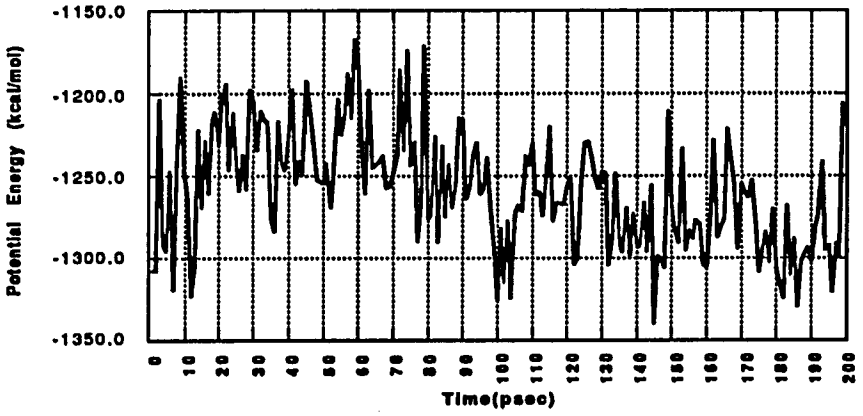
a)


-1500.0
-1550.0
-1600.0
-1650.0
-1700.0
Potential Energy (kcal/mol)
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200
Time(psec)

b)


-1150.0
-1200.0
-1250.0
-1300.0
-1350.0
Potential Energy (kcal/mol)
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200
Time(psec)

c)

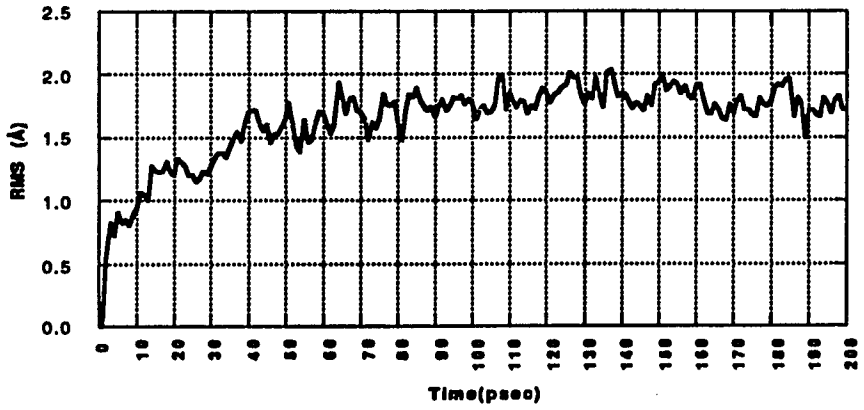

-1150.0
-1200.0
-1250.0
-1300.0
-1350.0
Potential Energy (kcal/mol)
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200
Time(psec)

After an initial heating and equilibration period of approximately 100 psec, the potential energy of each finger converged to a value with an energetic fluctuation of approximately 25 kcal/mol. The total simulation time for each zinc finger was 200 psec, with at least 50 psec of production time for data collection. Dynamic structural analyses were performed for a period of 50 psec during the production phase of the simulation. The time evolution of the root mean square (RMS) deviation of the structure is shown in Figures 8a, b, c - TFIIIA(iii) "even," TFIIIA(iv) "odd," and xfin31, respectively.

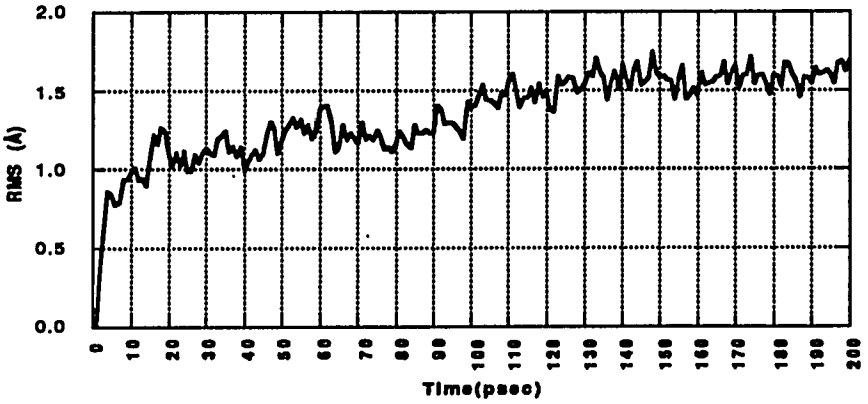
Figure 8

The RMS (root mean square) deviation trajectories of the fully solvated zinc finger simulations are shown as a function of time: a) TFIIIA(iii) "even" zinc finger of TFIIIA, b) TFIIIA(iv) "odd" zinc finger, and c) xfin31 "odd" zinc finger.

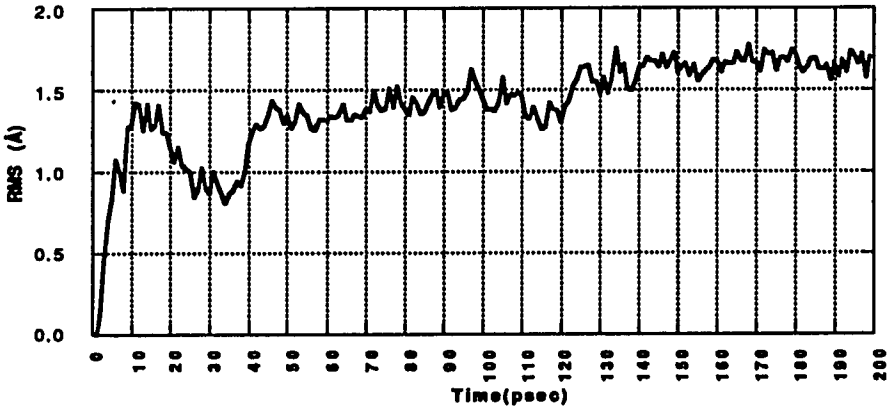
a)


RMS (Å)
2.5
2.0
1.5
1.0
0.5
0.0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200
Time(psec)

b)


RMS (Å)
2.0
1.5
1.0
0.5
0.0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200
Time(psec)

c)


RMS (Å)
2.0
1.5
1.0
0.5
0.0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200
Time(psec)

Simulation of the Xfin31 Zinc Finger

Xfin31 molecular dynamics structural comparison to 2D NMR structure

In order to evaluate the novel zinc ion parameters and the MD simulation conditions, the structural results of the xfin31 is compared to results from 2D NMR structure[32]. A 50 psec time averaged structure (taken from the converged time period, 150-200 psec) of xfin31 is nearly identical to the 2D NMR structure reported by Lee et al.

Figure 9

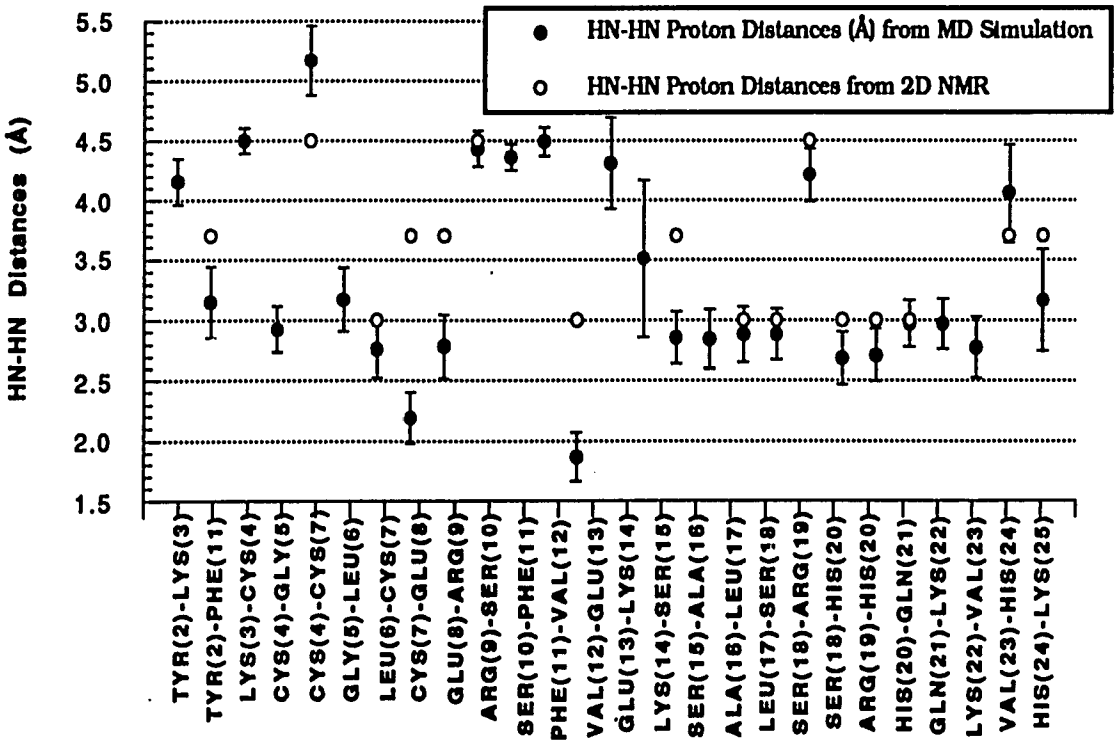
2D NMR and MD structure of xfin31 are superimposed. Cα trace of 2D NMR (yellow) and 50 psec MD structure (green) are superimposed. The sidechains of the two cysteines and two histidines of each zinc finger are shown. The zinc ion is shown in purple

The RMS deviation of the backbone heavy atoms (N, Cα, C, and O) between initial structure and structures along the trajectory averaged 1.42 Å (with a standard deviation of 0.30 Å); and the averaged RMS deviation for all atoms was 3.14 Å (with standard deviation of 0.64 Å). Because xfin31 is a small protein with many charged/polar residues (13/25 total residues) and has a relatively small buried surface area, most of its sidechains are solvent exposed. Thus its sidechains had numerous interactions with the solvent water molecules and moved freely during the simulation. However, the hydrogen bonding between the mainchain atoms remained relatively consistent. The simulation maintained the identifying secondary structural features of classical (CC/HH) zinc fingers- the anti-parallel ß sheet and the α helical folding. Major RMS deviations of the amino acids from the 2D NMR structure were near the C terminus beyond the histidine ligands and the loop turn region ("fingertip") connecting the second ß-strand to the α helix. Because the amino acids in these two regions do not make any significant contacts with the rest of the finger, they are free to explore a larger conformational space during an extended simulation. The results of the simulation of the 2D NMR structure of xfin31 is similar to a recent molecular dynamics simulation of solvated xfin31 by Palmer et al[73] who concluded that only solvated MD simulations for xfin31 are able to reproduce the NMR relaxations. They reported a mainchain heavy atoms RMS deviation of 1.22 Å and all atom RMS deviation of 2.87 Å. However, their simulation time was only 100 psec.

A comparison of the MD simulated and experimentally obtained HN-HN proton distances is shown in Figure 10.

Figure 10

Comparison of the proton-proton distance derived from the MD simulation to the 2D NMR Data. The error bars for the simulation values represent the fluctuations around the mean proton-proton distances for the 50 psec time period in which the averaged xfin31 structure was taken.



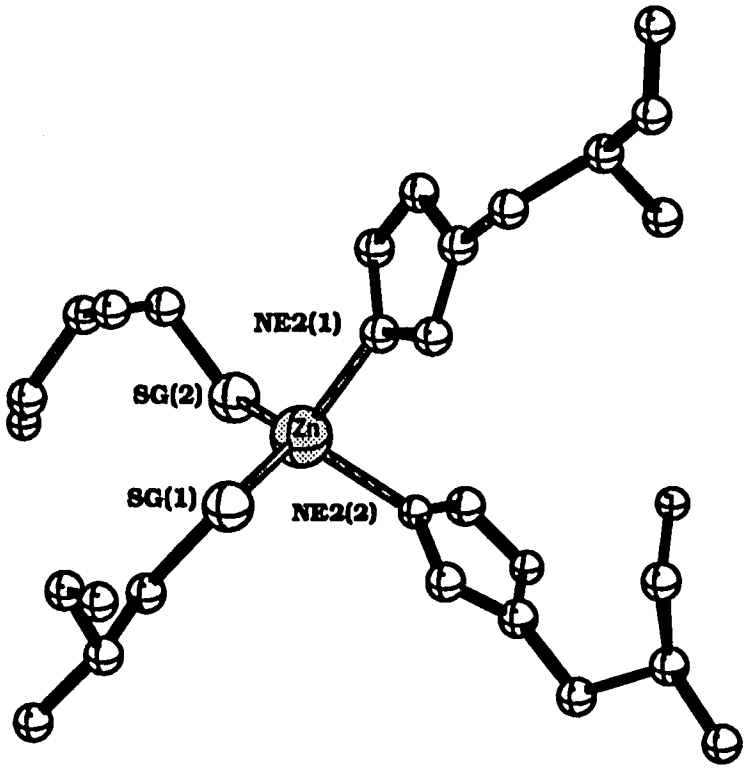
The values derived from the 2D NMR data are similar to the MD values, especially in the helical region. Again, the greatest differences are found with those residues in the connecting loop region described earlier and the CYS-CYS loop.

2.3.3. Zinc Metal Cluster

The zinc cluster for the time averaged xfin31 structure is shown in Figure 11.

Figure 11

Metal cluster of the xfin31 50 psec time averaged structure. The sulfur atom of residues 4 and 7 are depicted as Sγ(1) and Sγ(2), respectively. The nitrogen atom of residues 20 and 24 are depicted as Nε(1) and Nε(2), respectively.



The mean and RMS values which describe the geometry of the metal cluster are listed in Table 5.

Table 5. Geometry parameters of the metal cluster. The averaged values and their means were derived from the xfin31 50 psec time averaged structure.

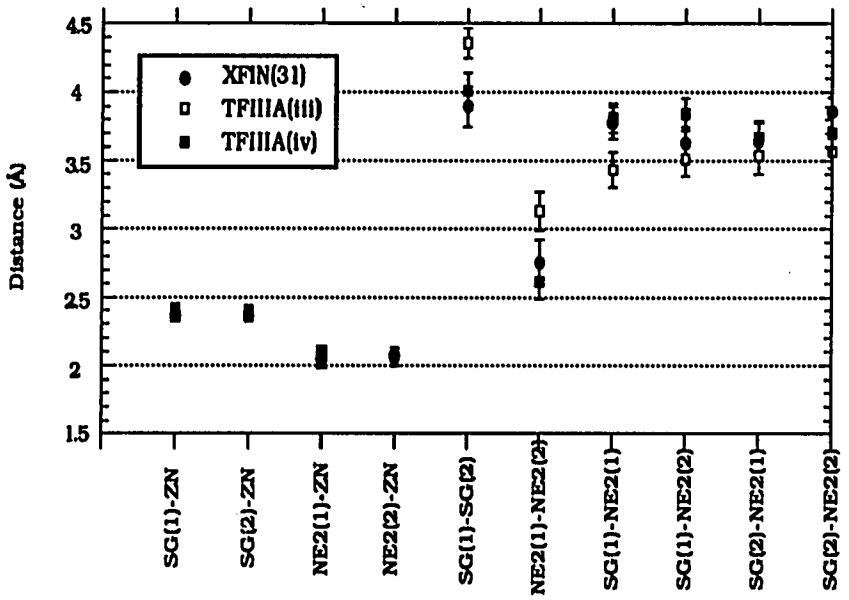
Parameter (Distance/Angle)]	Value (Ångstrom/degrees)	Root Mean Square (Ångstrom)
Sγ(1)-Zn	2.37 Å	0.05 Å
Sγ(2)-Zn	2.37 Å	0.05 Å
Nε2(1)-Zn	2.05 Å	0.07 Å
Nε2(2)-Zn	2.07 Å	0.06 Å
Sγ(1)-Sγ(2)	3.90 Å	0.15 Å
Nε2(1)-Nε2(2)	2.75 Å	0.17 Å
Sγ(1)-Nε2(1)	3.78 Å	0.12 Å
Sγ(1)-Nε2(2)	3.63 Å	0.12 Å
Sγ(2)-Nε2(1)	3.64 Å	0.14 Å
Sγ(2)-Nε2(2)	3.86 Å	0.11 Å
Sγ(1)-Zn-SG(2)	111.0 °	6.0 °
Sγ(1)-Zn-Nε2(1)	117.8 °	5.4 °
Sγ(2)-Zn-Nε2(1)	110.9 °	5.7 °
Sγ(1)-Zn-Nε2(2)	109.7 °	4.8 °
Sγ(2)-Zn-Nε2(2)	120.8 °	5.0 °
Nε2(1)-Zn-Nε2(2)	84.0 °	5.0 °

The ligand-metal distances are consistent with the reported values determined from EXAFS experiments[27]. The overall geometry of the metal cluster is tetrahedral with some angles deviating from standard tetrahedral value of 109.5° which was predicted from the quantum chemical calculations (data not shown). Because the zinc ion force field approximates the molecular interactions of the ligands to the metal and the ligands between themselves, these values represent a complete description of the dynamics of the zinc finger.

A comparison of the zinc cluster geometry (distances and angles) from all three simulations is shown in Figures12 and 13 .

Figure 12

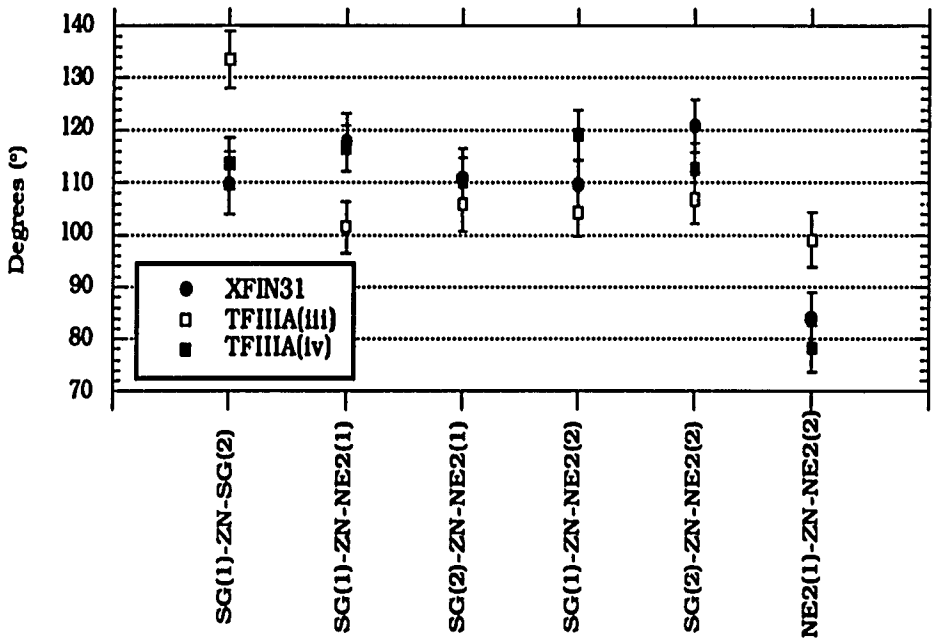
Comparison of the metal cluster distances of 50 psec time averaged structures from TFIIIA(iii), TFIIIA(iv), and xfin31. The atoms are SG(1), sulfur of the first cysteine ligand; SG(2), sulfur of the second cysteine ligand; NE2(1), nitrogen of the first histidine ligand; NE2(2), nitrogen of the second histidine ligand; ZN, zinc.



The values for xfin31 and TFIIIA(iv), both "odd" zinc fingers, are statistically the same while some values for TFIIIA(iii), the "even" one, differ. Although zinc ion to ligand distances are similar for all three zinc clusters, three ligand to ligand distances for the "odd" zinc fingers are significantly different from the ones in the "even" structure: Sγ(1) to Sγ(2), Nε2(1) to Nε2(2), and Sγ(1) to Nε2(1). A comparison of the bond angles form by the metal and two atoms of the sidechain which form the metal ligands is shown in Figure 13.

Figure 13

Comparison of the metal cluster angles of 50 psec time averaged structures from TFIIIA(iii), TFIIIA(iv), and xfin31. The atoms are SG(1), sulfur of the first cysteine ligand; SG(2), sulfur of the second cysteine ligand; NE2(1), nitrogen of the first histidine ligand; NE2(2), nitrogen of the second histidine ligand; ZN, zinc.



The distortion of the cluster geometry observed in the distance analysis is localized to three bond angles: Sγ(1)-Zn-Sγ(2), Sγ(1)-Zn-Nε2(1), and Nε2(1)-Zn-Nε2(2). The greatest difference in bond angle is found in Sγ(1)-Zn-Sγ(2) and Sγ(1)-Zn-Nε2(1). The differences between the "odd" and "even" ligand to ligand distances and the bond angles suggest that the overall geometry of the "even" zinc finger metal cluster is slightly distorted while preserving the same ligand to metal distances. These results are consonant with the findings of Xu et al [40] who reported that ADR1a, the "even" zinc finger of ADR1, a number of cross peaks localized to metal cluster that was observed in the major form, but not in the minor form of the peptide. They attributed the major differences between the two forms of ADR1a to the distortions of the metal cluster.

2.3.4. Comparison of the TFIIIA(iii) "Even" and TFIIIA(iv) "Odd" Simulations

2.3.4.1. Dynamic properties of "even" and "odd" zinc fingers

Time averaged structures were determined for a 50 psec time period during the production phase of TFIIIA(iii) "even" (130 to 180 psec) and TFIIIA(iv) "odd" (150 to 200 psec). The average rms deviations relative to the initial energy minimized structure of the mainchain heavy atoms (N, Cα, C, and O) were 1.64 Å (S.D. 0.31) and 1.37 Å (S.D. 0.28) for TFIIIA(iii) and TFIIIA(iv), respectively. The backbone rms difference for equivalent amino acids between the two zinc fingers of TFIIIA was 1.65 Å. Both zinc fingers retained their

secondary structures throughout the simulation, (see Figures 14 and 15).

Figure 14

50 psec time averaged structure of TFIIIA(iii) "even" zinc finger.

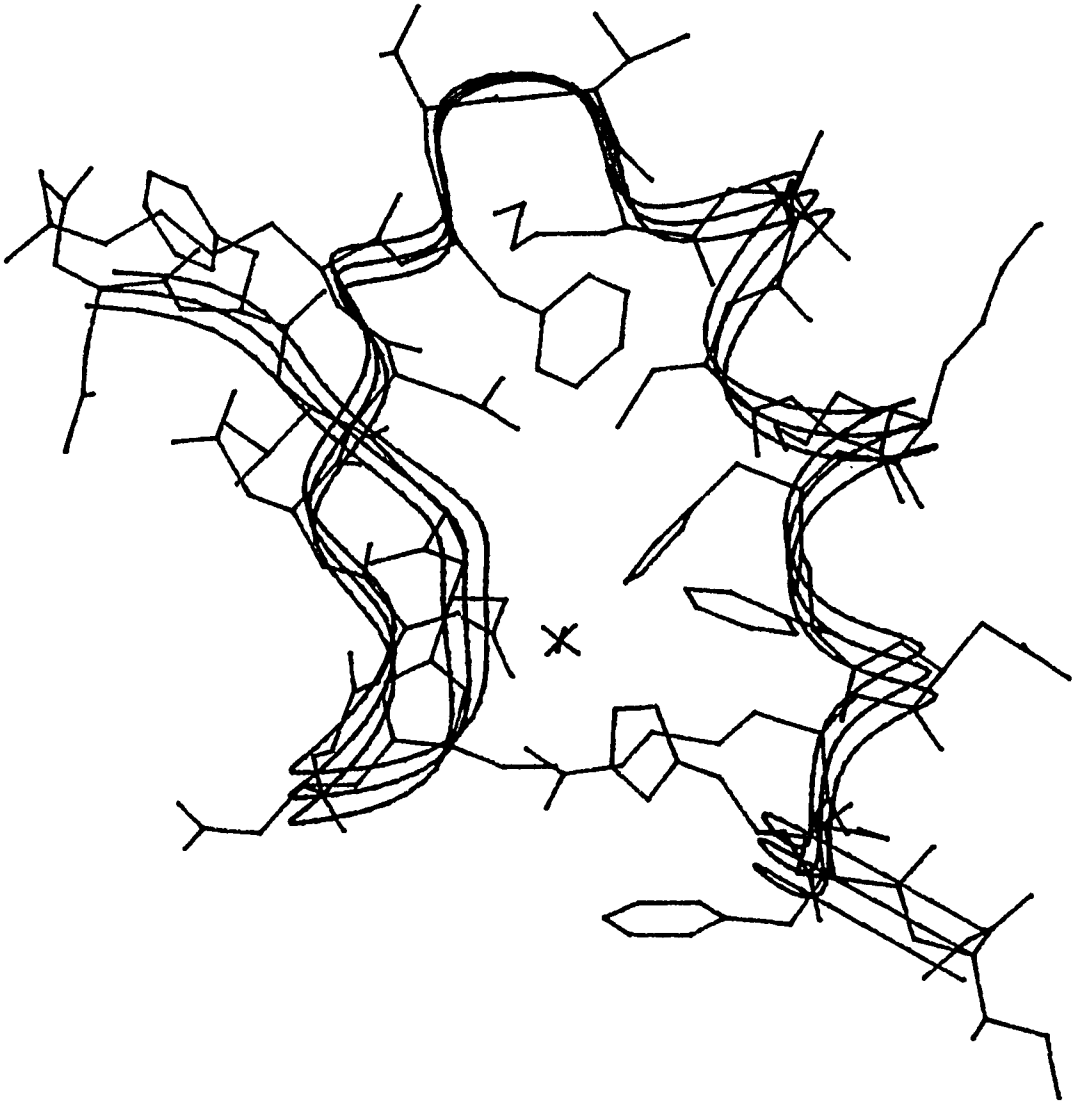
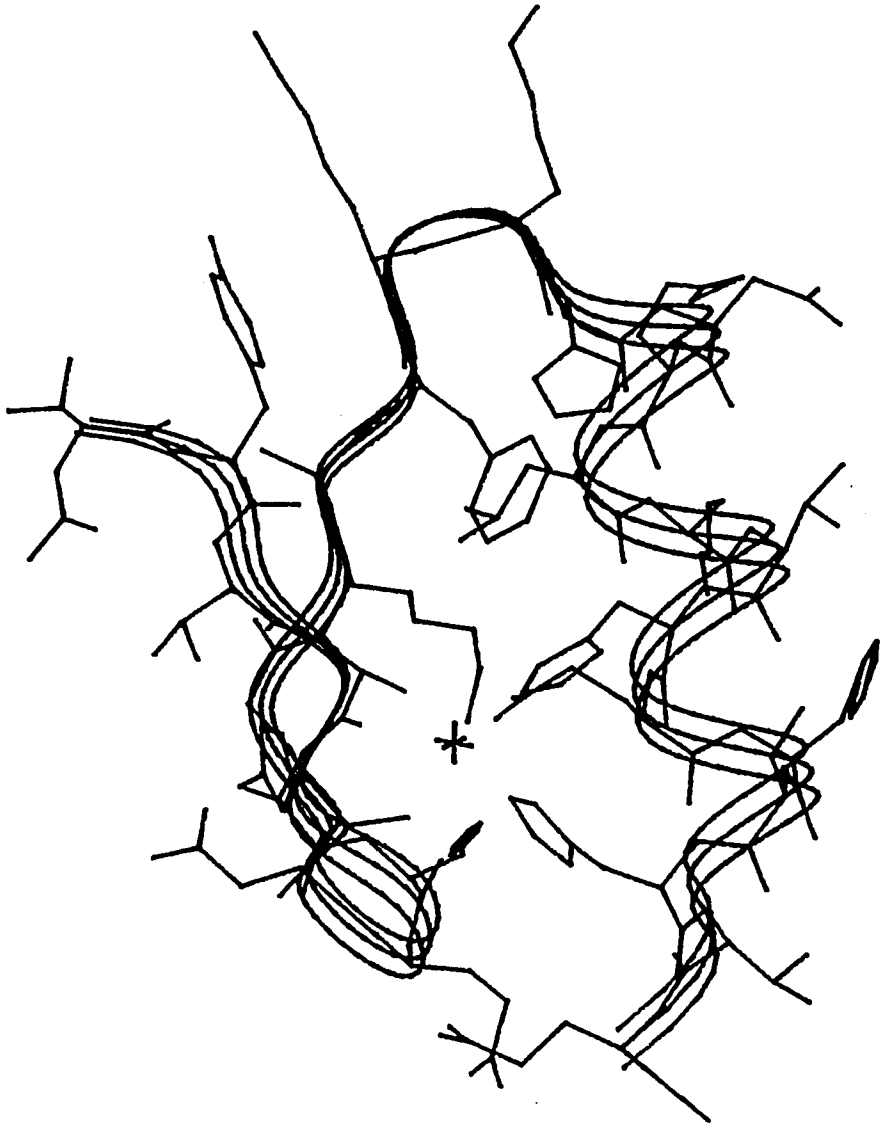


Figure 15

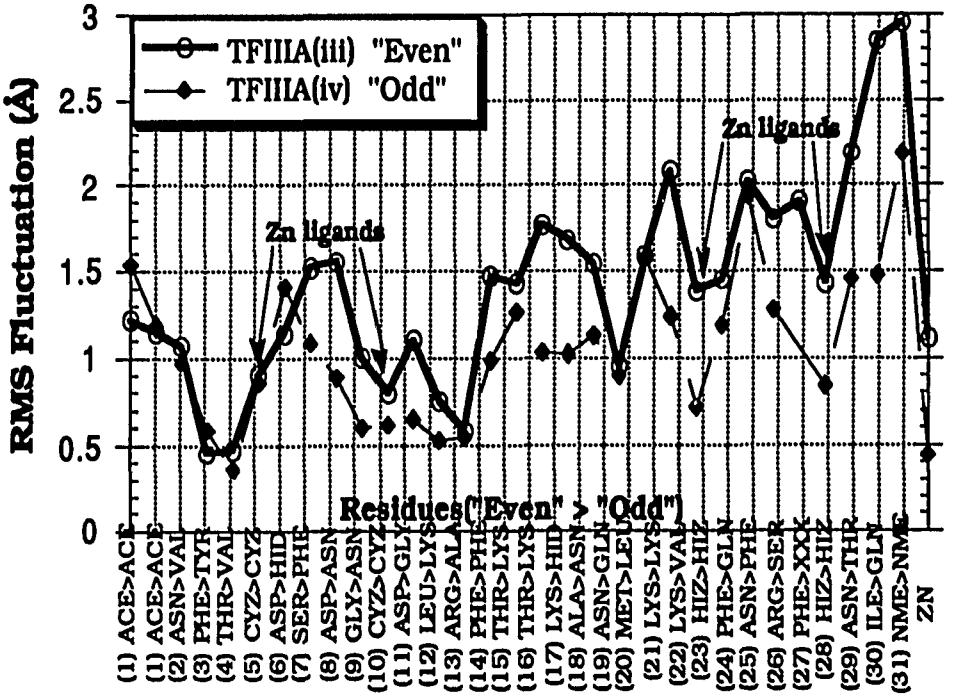
50 psec time averaged structure of TFIIIA(iv) "odd" zinc finger.



In spite of the similarities in their time averaged structure, the "even" TFIIIA zinc finger differs in its dynamic property from the "odd" zinc finger. A plot of the fluctuation of each residue in the sequence is shown in Figure 16.

Figure 16

RMS fluctuation for each amino acid of TFIIIA(iii) and TFIIIA(iv) 50 psec time averaged structure. Zn ligands identify the 4 residues which form the zinc cluster.



The mean position of each amino during the 50 psec time period (the time period in which the average structure was derived) was first determined, and the relative displacement of each residue to the mean was then calculated and expressed in terms of a RMS fluctuation.

$$\text{The RMS fluctuation} = \left\langle d_{ij}^2 \right\rangle^{1/2} \quad (4)$$

$$\text{where } d_{ij} = \left[\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 + \left(z_i - z_j\right)^2 \right]^{1/2} \quad (5).$$

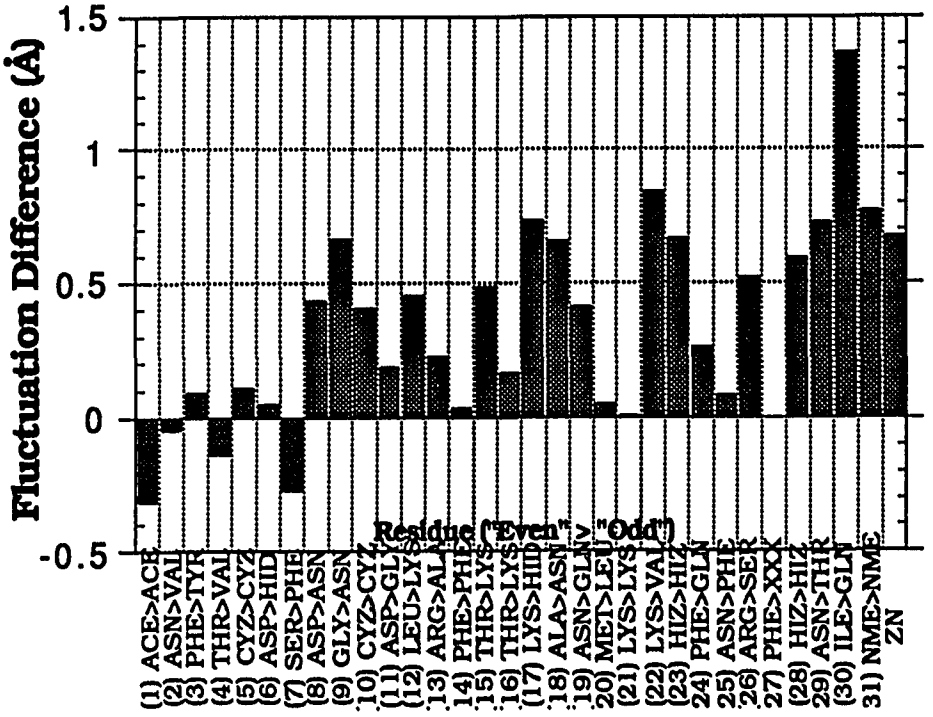
The average MD structure is represented by the subscript *j*, and the cartesian coordinates (x, y, and z) of each structure (represented by subscript *i*) during the time averaged period was compared to the average. A total of 500 structures (10 structures per psec) were used in the determination of fluctuation. Thus the RMS values represent the extent of excursion of each amino acid from the mean for a defined length of time. Consequently, the greater the fluctuation observed in the dynamics trajectory, the greater the flexibility attributable to that region of the protein structure.

The "even" zinc finger TFIIIA(iii) had a greater overall fluctuation during the MD simulation than the "odd" TFIIIA(iv). The mean fluctuation for all the residues of TFIIIA(iii) was 1.41 Å (S. D. 0.59) while the mean fluctuation for TFIIIA(iv) was 1.05 Å (S. D. 0.43). A student pairwise T-test revealed that the difference in the mean was significant to $\rho < 0.05$ level. The difference in RMS fluctuation

between the "even" and "odd" zinc finger at each residue position is illustrated in Figure 17.

Figure 17

Residue fluctuation differences between TFIIIA(iii) and TFIIIA(iv). The RMS fluctuations from Figure 16 of the "odd" TFIIIA(iv) zinc finger was subtracted from the "even" TFIIIA(iii) zinc finger. Thus values above the line indicate the residues with greater fluctuation observed in the "even" finger.



The pattern of fluctuations in Figure 17 shows that TFIIIA(iii) exhibited greater fluctuation than TFIIIA(iv) for the second strand of the β-sheet (residues 10-14) and in the α helix (residues 17-28). More important, however, is the difference in the C terminus of the "even" zinc finger which shows much greater fluctuation than in the "odd" zinc finger. Residue 30 of TFIIIA(iii) demonstrates the largest difference (1.37 Å).

It is noteworthy that in the "odd" zinc finger TFIIIA(iv), the invariant residues which form the zinc metal cluster, CYZ(5, 10) and HYZ(23, 28) have typically smaller fluctuations compared to the residues within the particular secondary structure (ß sheet or α helix) in which they are found. Compared to TFIIIA(iv), the residues which constitute the zinc ligands in the "even" TFIIIA(iii) demonstrated a much greater fluctuation (Table 6).

Table 6. RMS fluctuation values for the metal ligands only of "even" TFIIIA(iii) and "odd" TFIIIA(iv) zinc fingers.

Residue (position)	TFIIIA(iii) Å	TFIIIA(iv) Å	Difference Å
CYZ (5)	0.47	0.36	0.11
CYZ (10)	1.00	0.60	0.41
HIZ (23)	1.39	0.72	0.67
HIZ (28)	1.439	0.84	0.59
Zn	1.12	0.44	0.68

The greatest difference is found in HYZ (23, 28), the histidine residues which form part of the helix and which also participate in zinc chelation. The zinc ion itself has a greater fluctuation in an "even" finger than "odd." Thus, all the components of the zinc cluster which is a major stabilizing force in zinc fingers are more flexible in "even" zinc fingers than "odd" zinc finger

Superpositioning of 5 MD (10 psec) Averaged Structures

Five 10 psec MD averaged structures were obtained during the 50 psec time period described earlier for TFIIIA(iii) and TFIIIA(iv). These five structures were then positionally superimposed, as shown Figures 18 and 19 where only the heavy atoms are depicted and the hydrogen bonds are represented as dashed lines.

Figure 18

Five superimposed structures (10 psec time averaged) of "even" TFIIIA(iii). The hydrogen bonds are represented by (dashed) yellow lines. The Cα trace of each finger is drawn in red.

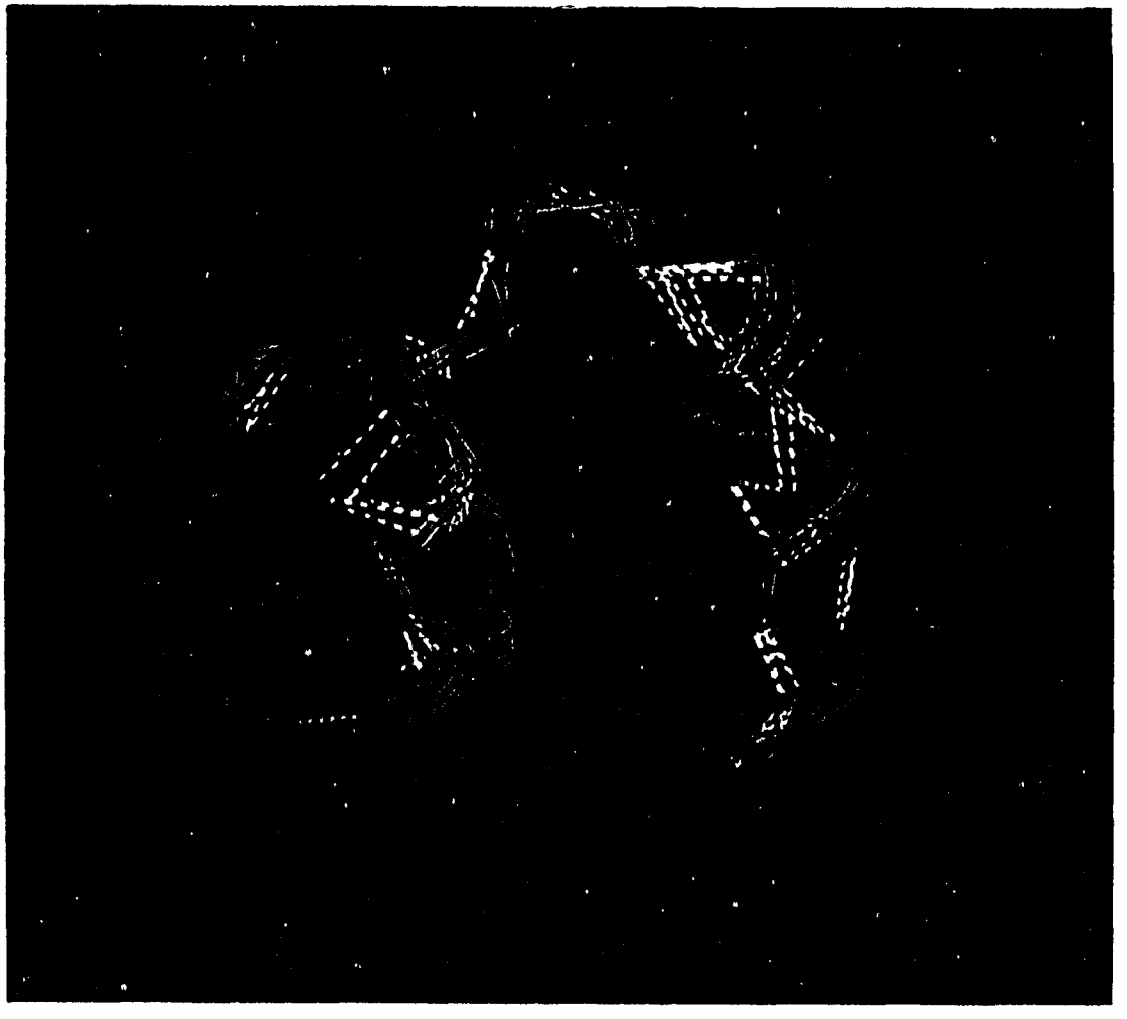
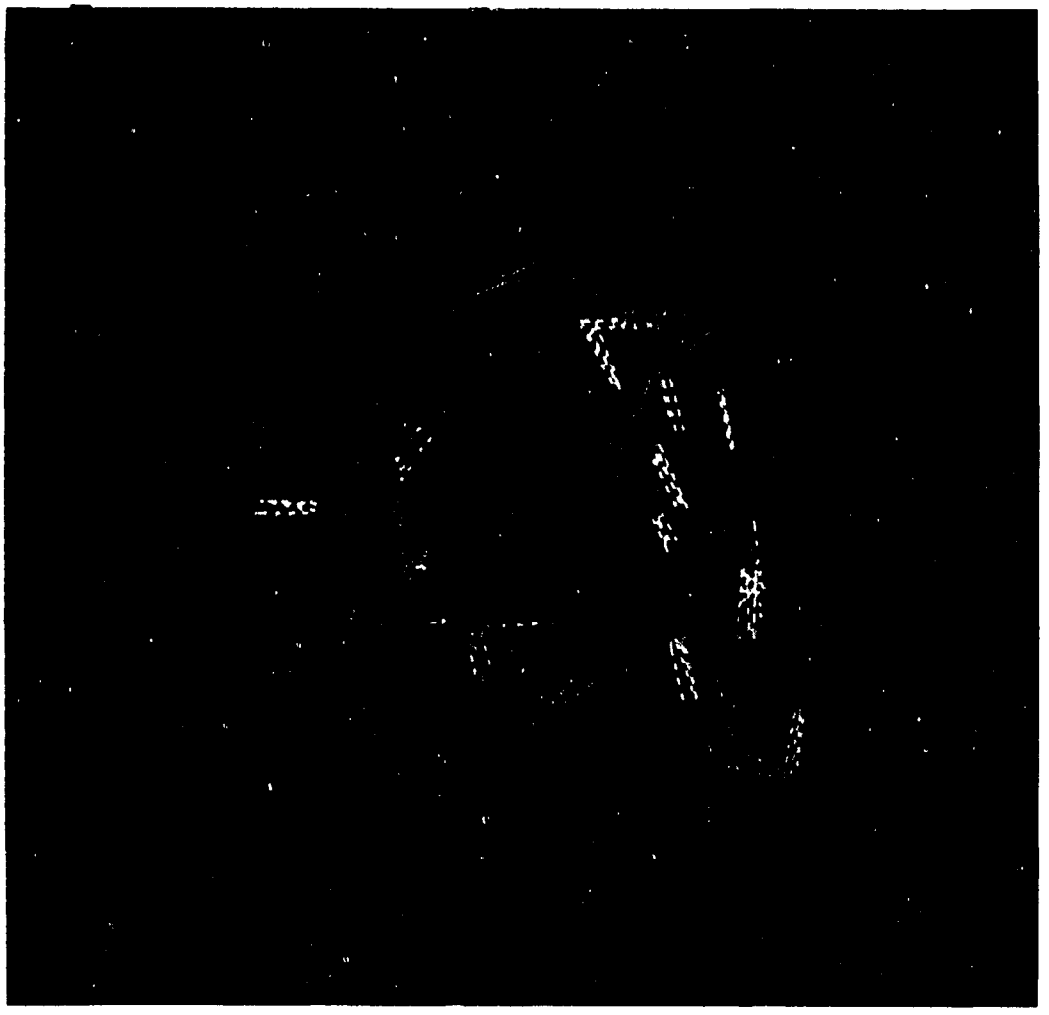


Figure 19

Five superimposed structures (10 psec time averaged) of "odd" TFIIIA(iv). The hydrogen bonds are represented by (dashed) yellow lines. The Cα trace of each finger is drawn in red.



The superposition of the zinc fingers offers an illustration of the structural implications of the dynamic motion of the protein. Clearly, the five TFIIIA(iv) structures are more conformationally homogeneous during the examined time period and thus appear to have less dynamic motion than TFIIIA(iii). The difference between TFIIIA(iv) and TFIIIA(iii) is more pronounced in the C-terminal region where TFIIIA(iii) covers a much greater variety of conformations than TFIIIA(iv). More importantly, the predominantly α helical mainchain hydrogen bonds pattern (i to i+4) identifiable in the superposition of TFIIIA(iv) structures appear to be consistently maintained throughout the time interval for all five structures. In contrast, the hydrogen bonding pattern of TFIIIA(iii) shows much greater variability; it departs from the (i to i+4) pattern of an α helix, and at times displays an (i to i+3) pattern of a 3_{10}. The greater variability of the "even" TFIIIA(iii) zinc finger during the 50 psec time interval is best illustrated in Table 7.

Table 7. Helical mainchain hydrogen bonding profile of TFIIIA(iii) and TFIIIA(iv). The mainchain hydrogen bonds from the five10 psec time averaged structures shown in Figures 18 and 19 were queried. If a hydrogen bond were found between the backbone residues listed in parentheses, a • was scored. Thus "• • • • •" would indicate that particular hydrogen bond was maintained throughout the 50 psec time interval.

TFIIIA(iii) "Even"			TFIIIA(iv) "Odd"		
ASN(19)	THR(16)	• • • •	GLN(19)	LYS(16)	•
MET(20)	THR(16)	• • •	LEU(20)	LYS(16)	• • • • •
LYS(21)	LYS(17)	• •	LYS(21)	HIS(17)	• • • • •
LYS(21)	ALA(18)	•	VAL(22)	ASN(18)	• • • • •
LYS(22)	ASN(19)	• • • • •	HIZ(23)	LEU(20)	• • • • •
HIZ(23)	ASN(19)	• • •	HIZ(23)	GLN(19)	• • • • •
HIS(23)	MET(20)	•	GLN(24)	LEU(20)	• • • • •
PHE(24)	MET(20)	• • •	PHE(25)	LYS(21)	• • • • •
PHE(24)	LYS(21)	•	SER(26)	VAL(22)	• • • •
ASN(25)	LYS(22)	• • • • •	SER(26)	HIZ(23)	• •
ARG(26)	HIS(23)	• • • • •	HIZ(27)	HIS(23)	• • • • •
PHE(27)	PHE(24)	• • • •	THR(28)	GLN(24)	• • • • •
HIZ(28)	PHE(24)	• • • •	GLN(29)	HIS(27)	• • • • •

Correlation Analysis of the Cluster to Hydrogen Bonding

In order to elucidate more precisely the association between the molecular details of metal cluster dynamics to the flexibility of the "even" zinc finger, time cross-correlation analysis of the MD simulation was performed. Since the MD trajectory files generated during the simulation contained the cartesian coordinates or "frames" of all the atoms during the simulation (10 frames were recorded for every picosecond) the distance between two atoms can be determined at any point during the simulation. Thus the correlated displacement or coupling of any two sets of atomic distances during a specified length of time can be calculated directly from the trajectory files[74 75].

Displacement cross-correlation *c(i,j)* is given by equation:

$$c(i,j) = \left\langle (d_i - \overline{d_i})(d_j - \overline{d_j}) \right\rangle \quad (6)$$

where d_i and d_j represent any two atomic distances and the angle brackets denote the ensemble average. The mean distance ($\overline{d_i}$) for one set of distances (i) is given by:

$$\overline{d_i} = \frac{\sum_{1}^{N} d_i(N)}{N} \quad (7)$$

where N represent the total number of frames, and $d_i(N)$ is the distance of i at the N^{th} time frame. A normalized distance cross-correlation is described by:

$$C(i,j) = \frac{\left\langle (d_i - \overline{d_i})(d_j - \overline{d_j}) \right\rangle}{\left\langle (d_i - \overline{d_i})^2 \right\rangle^{1/2} \left\langle (d_j - \overline{d_j})^2 \right\rangle^{1/2}} \quad (8)$$

Therefore, for complete correlated motion, $C(i,j)$= 1 and complete anti-correlated motion, $C(i,j)$= -1. For non-correlated displacement, $C(i,j)$= 0. $C(i,j)$ values ranging between 0 to 1 (-1) would imply some degree of correlated motion. Normalized correlated motion can also be calculated for angles a_i and a_j,

$$C(i,j) = \frac{\left\langle (a_i - \overline{a_i})(a_j - \overline{a_j}) \right\rangle}{\left\langle (a_i - \overline{a_i})^2 \right\rangle^{1/2} \left\langle (a_j - \overline{a_j})^2 \right\rangle^{1/2}} \quad (9)$$

or between a set of displacements and a set of angles :

$$C(i,j) = \frac{\left\langle (d_i - \overline{d_i})(a_i - \overline{a_i}) \right\rangle}{\left\langle (d_i - \overline{d_i})^2 \right\rangle^{1/2} \left\langle (a_i - \overline{a_i})^2 \right\rangle^{1/2}} \quad (10)$$

The working hypothesis proposes that the dynamic motion of the metal cluster in a zinc finger is related to the helical mainchain

hydrogen bond formation or deformation during the simulation. Preliminary displacement cross correlation analyses of ligand-ligand distances and helical hydrogen bonding distances indicated that some motions in "even" TFIIIA(iii) are correlated. In order to determine which metal cluster angle(s) is responsible for the changes in hydrogen bonding pattern, metal cluster angle (X-Zn-X) to displacement (mainchain hydrogen bonds) correlation analyses were performed. The results for TFIIIA(iv) are shown in Table 8.

Table 8. Cross correlation table of hydrogen bonds to metal ligand angles of TFIIIA(iv). The atomic coordinates from the 50 psec of trajectory (from which the time averaged structure was derived) was taken from the TFIIIA(iv) simulation. A total of 500 frames (10 frames per picosecond) were used for the correlation calculations. The angle and hydrogen bonds are referenced below the table. Hydrogen bonds were represented by the distance between the mainchain amide hydrogen (left) of the amino acid and the carbonyl oxygen (right). The hydrogen bonds were identified earlier in Table 7.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(I)	.087	.106	.224	.152	-.101	-.033	.149	-.037	.027	.076	.048	.101	.030
(II)	-.033	-.027	-.056	-.043	.127	.004	-.122	.015	.139	.072	.043	.021	.082
(III)	-.047	.136	.140	.038	-.120	-.012	.058	-.021	-.078	-.009	.042	-.072	-.118
(IV)	.002	-.185	[-.316]	.236	.118	.003	-.019	.105	-.015	-.139	-.115	-.240	.059
(V)	-.120	-.103	-.185	-.031	.122	.098	-.220	-.015	-001	-.060	-.142	.045	.081
(VI)	.035	-.043	.018	.005	-.064	-.052	.081	-.018	-.086	.028	.131	.092	-.053

Metal Cluster Angles

I- Sγ(1)-Zn-Sγ(2)
II- (II) Sγ(1)-Zn-Nε2(1)
III- Sγ(1)-Zn-Nε2(2)
IV- Sγ(2)-Zn-Nε(1)
V- Sγ(2)-Zn-Nε2(2)
VI- Nε2(1)-Zn-Nε2(2)

Mainchain Helical Hydrogen Bonding

1- Gln(19)-Lys(16)
2- Leu(20)-Lys(16)
3- Lys(21)-His(17)
4- Val(22)-Asn(18)
5- Hiz(23)-Leu(20)
6- His(23)-Gln(19)
7- Gln(24)-Leu(20)
8- Phe(25)-Lys(21)
9- Ser(26)-Val(22)
10- Ser(26)-Hiz(23)
11- Hiz(27)-Hiz(23)
12- Thr(28)-Gln(24)
13- Gln(29)-Hiz(27)

There was only one strong correlation (>0.300) between the metal ligand angles to the hydrogen bonds in the TFIIIA(iv) simulation, Sγ(2)-Zn-Nε(1) and Lys(21)-His(17). This reference value of 0.300 for strong correlation was derived from a molecular dynamics study of the collective motions of bovine pancreatic trypsin inhibitor [74]. The authors reported strong correlation values of 0.35 to 0.55 for the correlated motions of residues, 47 to 56 in the helix, defined as residues that formed mainchain hydrogen bonds between positions *i* and *i+3* or *i+4* .

An identical correlation analysis was performed for the TFIIIA(iii) simulation, see Table 9.

Table 9. Cross correlation table of hydrogen bonds to metal ligand angles of TFIIIA(iii). The atomic coordinates from the 50 psec of trajectory (from which the time averaged structure was derived) was taken from the TFIIIA(iii) simulation. A total of 500 frames (10 frames per picosecond) were used for the correlation analyses. The angle and hydrogen bonds are referenced below the table. Hydrogen bonds were represented by the distance between the mainchain amide hydrogen (left) of the amino acid and the carbonyl oxygen (right). The hydrogen bonds were identified earlier in Table 7.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(I)	-.022	.031	-.001	.070	-.129	-.128	-.067	-.098	.009	-.140	-.044	-.111	.020
(II)	-.083	.070	.073	-.084	.114	-.054	.006	.366	.360	.409	.349	.096	.032
(III)	.122	.026	.016	-.001	.041	.015	-.005	.123	.101	.150	.026	.087	.140
(IV)	.090	-.093	-.065	.009	-.002	.164	.067	-.285	-.387	-.276	-.295	-.001	-.001
(V)	.029	.067	-.023	.031	.082	.199	.041	-.063	-.078	-.083	-.089	.007	-.112
(VI)	.143	.012	-.050	-.014	-.018	-.100	-.036	.055	.104	-.100	-.027	.059	.137

Metal Cluster Angles

I- Sγ(1)-Zn-Sγ(2)
II- Sγ(1)-Zn-Nε2(1)
III- Sγ(1)-Zn-Nε2(2)
IV- Sγ(2)-Zn-Nε2(1)
V- Sγ(2)-Zn-Nε2(2)
VI- Nε2(1)-Zn-Nε2(2)

Mainchain Helical Hydrogen Bonding

1- Asn(19)-Thr(16)
2- Met(20)-Thr(16)
3- Lys(21)-Lys(17)
4- Lys(21)-Ala(18)
5- Lys(22)-Asn(19)
6- Hiz(23)-Asn(19)
7- Hiz(23)-Met(20)
8- Phe(22)-Met(20)
9- Phe(24)-Lys(21)
10- Asn(25)-Lys(22)
11- Arg(26)-Hiz(23)
12- Phe(27)-Phe(24)
13- Hiz(28)-Phe(24)

Using the same criterion of a correlation value of 0.300 for strong correlation, two metal bond angles Sγ(1)-Zn-Nε2(1) and Sγ(2)-Zn-Nε(1) and 6 hydrogen bonds indicated strong correlations in the "even" TFIIIA(iii) zinc finger. Both metal ligand angles involve the first histidine residue. The hydrogen bonds with positive correlation were found along the helix before the first histidine ligand, Phe(22)-Met(20), and among the spacing residues.

Discussion and Conclusions

The main conclusion from the MD simulations of TFIIIA(iii) and TFIIIA(iv) that included the chemico-physical interactions of the components of the metal cluster is that the differentiating property between the four residue spacing of an "even" TFIIIA(iii) zinc finger and the three residue spacing of an "odd" TFIIIA(iv) finger is in the dynamic behavior rather than in the average structure. While the overall β-sheet and α helical folding is preserved for both types of fingers, TFIIIA(iii) is more flexible than TFIIIA(iv). The explicit use of an appropriate potential describing the interactions in the zinc cluster revealed the greater fluctuations of "even" TFIIIA(iii), and identified their origin in the more flexible and distorted metal cluster. Since this cluster is the major stabilizing component in the zinc finger protein, its flexibility has major implications for the properties of the entire zinc finger domain.

An HX3H zinc finger incorporates the structural properties of α helices in the structural arrangement leading to high-affinity metal

binding. Three-residue spacing between the two conserved histidine ligands allows the imidazoles to occupy the same face (metal binding) of the helix. In contrast, a four residue spacing between the conserved histidines found in "even" zinc fingers disrupts this seemingly complementary arrangement between the structural requirement for high-affinity ligand binding and for helix stability. The binding of zinc forces a modification of the torsional angles in the spacing region which includes the histidine ligands. This results in the disruption of the orderly mainchain hydrogen bonding pattern found in the helix. A distorted metal geometry results in "even" fingers from the compromise of chelation of the metal with the requirements for of optimal helical hydrogen-bonding pattern.

The local distortion between the histidines is propagated to the entire helix and to the second strand of the β-sheet. The large fluctuations of the "even" zinc finger alter the mainchain hydrogen-bonding contacts of the histidine ligands including their intervening residues with amino acids 4 positions away. Thus the entire helical hydrogen bonding pattern is affected by the local fluctuations of the cluster in a dynamic manner. In contrast to the "even" finger, TFIIIA(iv) had a consistent pattern of mainchain hydrogen bonds during an identical length of trajectory. Since the mainchain hydrogen bonds in helices are the major components of structural stability, the variability of the hydrogen bonds in "even" fingers accounts for the conformational heterogeneity which is expressed in the dynamic property of flexibility.

Cross-correlation analyses between the zinc cluster angles and the mainchain helical hydrogen bonds were performed to learn how the dynamic angular movements within the cluster are propagated to the hydrogen bonding network. Since the hydrogen bonds in an α helix stabilize secondary the structure, the disruption of the mainchain hydrogen bonds of the helix in "even" zinc fingers resulted in a more flexible finger.

An HX4H zinc finger requires flexibility in order to maintain zinc ion binding and helical conformation. Metal chelation is a prerequisite for the folding of the zinc finger into its canonical β-sheet and α helix, and proper folding of the zinc finger is a prerequisite for specific DNA recognition through the helix. Thus without an alteration of the overall folding of zinc fingers, a dynamic property of flexibility is introduced in this particular domain solely as a result of the increased spacing between the invariant histidines.

The implications of these findings from the MD simulations for the properties of the entire TFIIIA protein result from the systematic arrangement of its "odd" and "even" fingers. The third, sixth and eighth zinc fingers are of the "even" type while the first, second, fourth, fifth, seventh, and ninth fingers are of the "odd" type. Previously, reported hydroxyl-radical protection studies of the TFIIIA-5S complex suggested that the nine zinc fingers may bind to the A and C boxes of the ICR as groups of zinc fingers[51]. Furthermore Churchill et al[55] have hypothesized that the binding of one or one group of zinc fingers may preposition the next group of zinc fingers near their cognate DNA-binding site. Perhaps the "even" zinc fingers delineate

groups of zinc fingers and the flexibility of the "even" fingers allows the subsequent group of zinc fingers to bind to their respective base pairs.

Recently Hayes et al have accomplished extensive footprinting analyses in conjunction with missing nucleotide experiments[63] of TFIIIA constructs with variable numbers of zinc fingers[61] in order to probe specific zinc finger-to-DNA contact sites in the TFIIIA/5S complex. Briefly, their model proposes that TFIIIA binds to the 5S gene in a tripartite manner. The nine zinc fingers of TFIIIA can be organized in three groups of zinc fingers with each groups consisting of three zinc fingers. Furthermore, their model suggests that the first (ZF1-3) and third group (ZF7-9) of zinc fingers bind to the DNA in a Zif268-like manner (i.e., the helices of the fingers follow and bind to the base pairs in the major groove of the DNA). However, their findings support the notion that the second group of zinc fingers (ZF 4-6) binds to the minor groove of the DNA and somehow orients itself parallel to the major axis of the B-DNA.

Since to date the canonical (CC/HH) zinc fingers have similar folding regardless of the variable amino acids[38], what distinguishing features of the second group of TFIIIA zinc fingers allow them to bind to the DNA in this unconventional manner? I propose that the greater flexibility (especially at the C terminus) of the "even" third TFIIIA zinc finger "allows" for the transition of the major groove-binding zinc fingers of the first group to the minor groove binding zinc fingers of the second groove. Because all zinc fingers of TFIIIA bind to one face of the DNA, the minor groove in which ZF4-6 binds is adjacent to the

major groove in which ZF1-3 binds. The flexibility of the third zinc finger may "throw" the subsequent zinc finger, finger 4, out of the major groove to allow for minor groove interaction. In turn, the flexibility of the sixth zinc finger ("even") may "throw" the next group of zinc fingers back into the major groove.

CHAPTER 3

Introduction

The analysis of the results from "even" TFIIIA(iii) and "odd" TFIIIA(iv) MD simulations suggested that a four residue spacing between the invariant histidine ligands confers a significant degree of detectable flexibility for only "even" zinc fingers. The HX4H spacing of TFIIIA(iii) appears to disrupt the complementary arrangement between the helical structure to metal binding, a more flexible metal cluster, and a helix with frequently disrupted hydrogen bonds patterns.

The results of the simulation showed that the zinc metal ion and three of the four residues which constitute the zinc ligands in TFIIIA(iii) had greater RMS fluctuations than in TFIIIA(iv). A direct relationship between the flexibility of the metal cluster and the overall flexibility of the "even" zinc finger can be probed by artificially removing the dynamic motion of the metal to its four ligands. Such a "constrained" molecular dynamics simulation would demonstrate that an accurate dynamic representation of the entire molecular system is required for observing the fluctuations of the "even" zinc finger and would support the hypothesis that the overall greater flexibility observed in TFIIIA(iii) is attributable to the dynamics of the metal cluster. A constrained MD simulation should have minimal effect for the "odd" TFIIIA(iv) finger since its ligands and zinc metal ion already

have low fluctuations. However, if the flexibility of the metal cluster were directly associated with the flexibility of the entire "even" zinc finger, a constrained simulation for the "even" TFIIIA(iii) finger would reduce the flexibility of whole protein compared to the unconstrained results.

In order to test the hypothesis that the flexibility observed in the TFIIIA(iii) simulation is a property intrinsic to the HX4H spacing and not to the primary structure, a MD simulation of a single residue deletion (between the invariant histidines) of TFIIIA(iii) was performed. These computational experiments are described below. Mutating the "even" TFIIIA(iii) zinc finger to an "odd" zinc finger should reduce the flexibility of the metal cluster and therefore the zinc finger itself.

Methods

Constrained Molecular Dynamics

MD simulations of the third "even" zinc finger TFIIIA(iii) and the fourth "odd" zinc finger TFIIIA(iv) were repeated as described in section 2.2.4, but with substantially increased force constant (1000 kcal/mol/Å^2) for Zn^{+2} to Nε2 and Zn^{+2} to Sγ bonds. With these parameters, the five atoms which constitute the core metal cluster becomes effectively "frozen." The initial configurations of the system for constrained simulation of "even" TFIIIA(iiic) and "odd" TFIIIA(ivc) for MD simulation were the fully minimized solvated structures of

TFIIIA(iii) and TFIIIA(iv). The methods of molecular construction and molecular modeling are described in sections 2.2.1 and 2.2.3. The MD simulation conditions are given in section 2.2.4.

Deletion Mutation - TFIIIA(iiimf)

Figure 20 compares the sequence of a single deletion mutation of the third zinc finger of TFIIIA to the wildtype sequence. The deletion between the invariant histidines converts the "even" zinc finger into an "odd" one. Because the conserved residues and their spacing is kept for the mutant zinc finger TFIIIA(iiimf), the folding of the mutant finger should be similar to an "odd" finger. Based on crystallographic data of Zif268, a three zinc finger protein complexed with DNA[1], residues at this position of the zinc finger do not come in direct contact with the DNA. The nearest proposed direct DNA contact position, -1 residue from the first histidine ligand, would position the phenylalanine in the wildtype "even" zinc finger faced away from the major groove of the DNA.

Figure 20

Amino acid sequence comparison of the wildtype and mutant zinc finger. The phenylalanine (+1 residue from the first conserved histidine ligand, **bold**) of the third "even" zinc finger of TFIIIA (TFIIIA(iii)) was deleted to form TFIIIA(iiimf). The residue spacing which defines an "even" zinc finger is underlined.

```
TFIIIA(iii)   N F T C D S D G C D L R F T T K A N M K K H F N R F H N I
TFIIIA(iiimf) N F T C D S D G C D L R F T T K A N M K K H • N R F H N I
```

The mutant was constructed, solvated and minimized as describe in sections 2.2.1 and 2.2.3. The molecular dynamics simulation conditions are described in sections 2.2.4. A summary of the MD simulations that will be compared and their descriptions is shown in Table 10.

Table 10.

Simulation	Description
TFIIIA(iii)	Third "even" zinc finger of TFIIIA
TFIIIA(iv)	Fourth "odd" zinc finger of TFIIIA
TFIIIA(iiic)	Third "even" zinc finger of TFIIIA with constrained zinc metal cluster
TFIIIA(ivc)	Fourth "odd" zinc finger of TFIIIA with constrained zinc metal cluster
TFIIIA(iiimf)	Single residue deletion between the invariant histidine ligands of the third zinc finger of TFIIIA

Results

MD analysis of TFIIIA(iiic) and TFIIIA(ivc)

Both constrained MD simulations, TFIIIA(ivc) and TFIIIA(iii), were performed for a total of 150 psec. Each structure converged after 100 psec of simulation, and a 50 psec time averaged structure (100 to 150 sec) was derived for each finger. Figure 21 and Figure 22 describe the RMS deviations of the mainchain atoms (N, C, O, Cα) along the trajectory of the constrained "odd" zinc finger and "even" zinc finger respectively. The RMS deviation is represented by the equation $\sqrt{\frac{\sum(x_i - x_j)^2}{n}}$ (11), where x_i is the cartesian coordinates of the initial structure (fully energy minimized solvated structure before the onset of molecular dynamics), x_j is the atomic coordinates of the atoms at the j^{th} picosecond, and n represents the total number of atoms.

Figure 21

RMS deviation of TFIIIA(ivc).

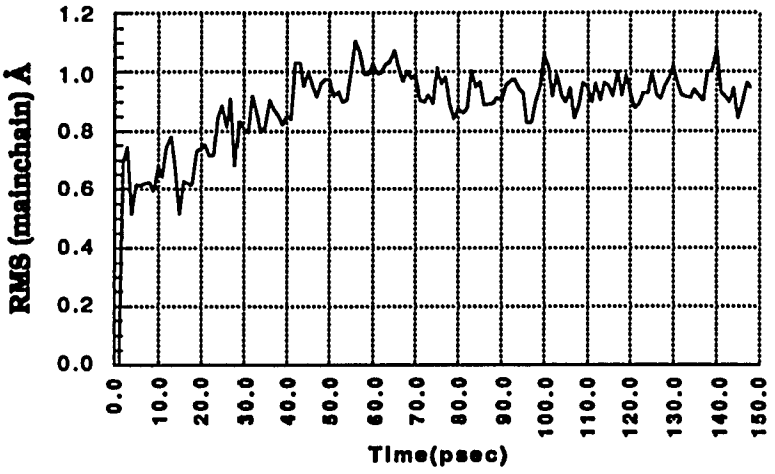
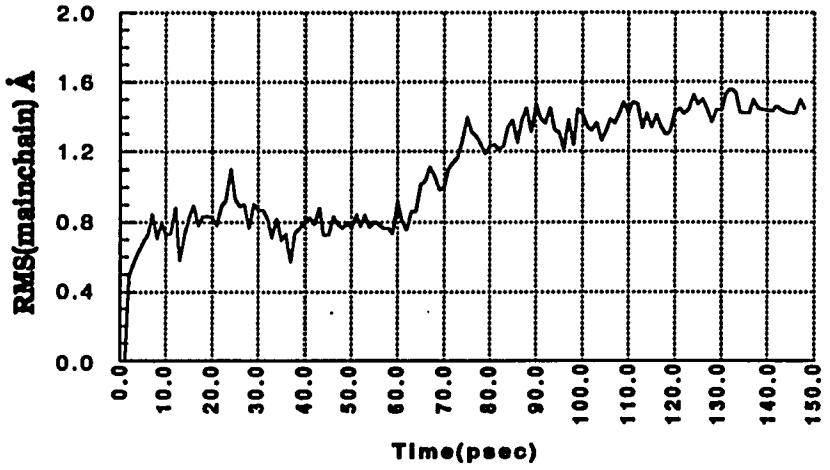
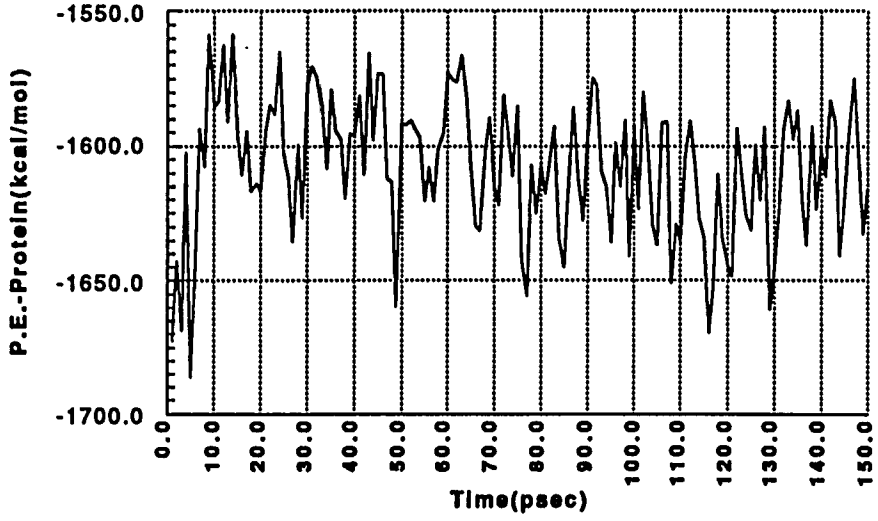


Figure 22

a) RMS deviation of TFIIIA(iiic)



b) Potential energy trajectory of TFIIIA(iiic)

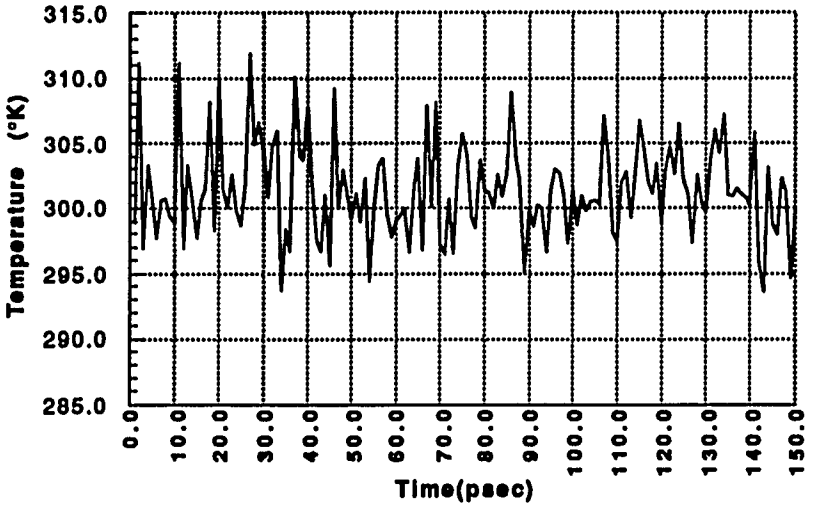


The average RMS deviations for the mainchain atoms were 0.89 Å (S. D. 0.14) for the "odd" TFIIIA(ivc) zinc finger and 1.10 Å (S.D. 0.32) for the "even" zinc finger. The potential energy of the constrained "even" zinc finger is shown in Figure 23b. Constraining the metal cluster resulted in less structural deviation and faster structural convergence than the "unconstrained" simulation.

The temperature and the total energy trajectory of the entire TFIIIA(iiic) molecular system (protein and solvent) are shown in Figure 23a and Figure 23b, respectively. The components of the total energy, the kinetic and potential energy, are shown in Figure 24a and Figure 24b, respectively. In summary, constraining the zinc ion and its ligand atoms did not adversely affect the molecular dynamics simulation.

Figure 23

a) Temperature trajectory of the TFIIIA(iiic) simulation



b) Total energy trajectory of TFIIIA(iiic) simulation

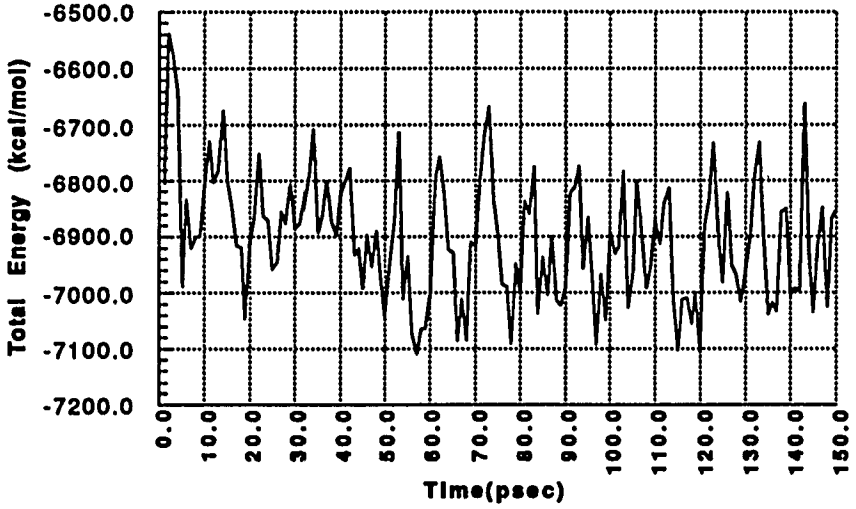
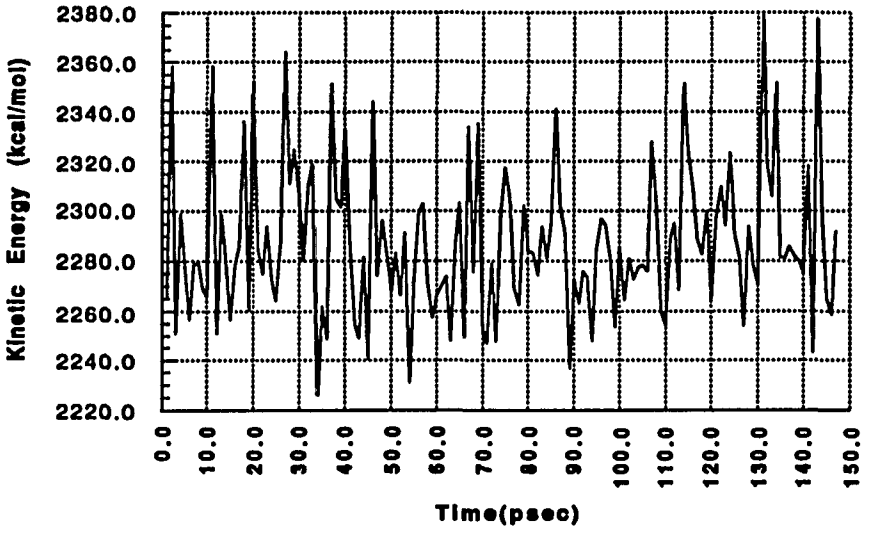
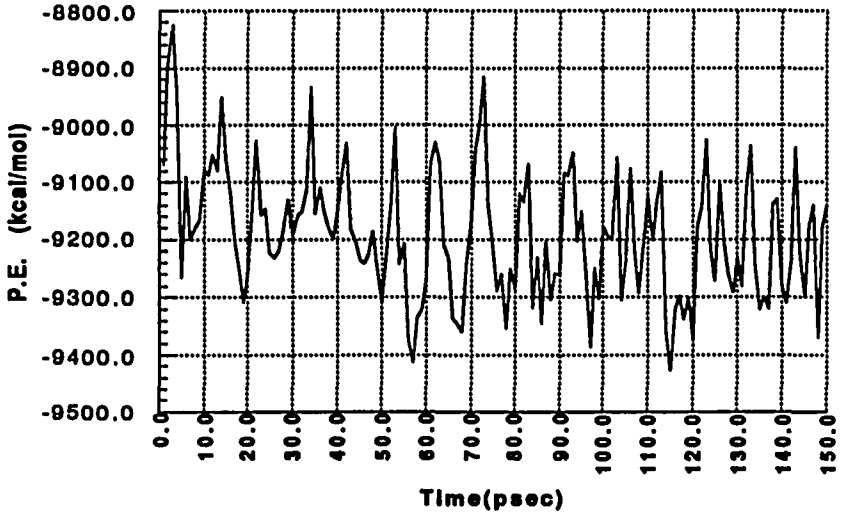


Figure 24

a) Kinetic energy trajectory of the TFIIIA(iiic) simulation



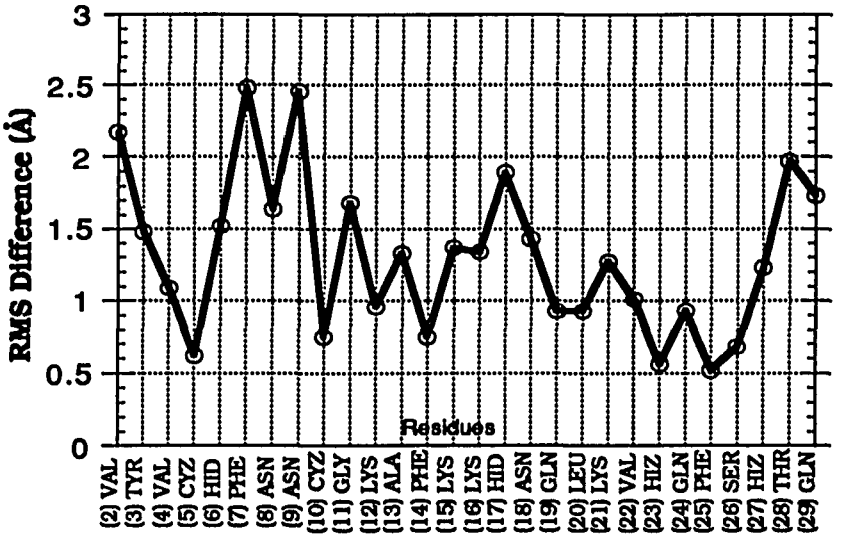
b) Potential energy of the TFIIIA(iiic) simulation



The 50 psec time averaged structure of the constrained "odd" finger TFIIIA(ivc) was compared to the unconstrained finger TFIIIA(iv). A plot of the RMS difference is shown in Figure 25.

Figure 25

Structural comparison of the 50 psec time averaged structures of TFIIIA(iv) and TFIIIA(ivc). The Molecular Similarity module of the Quanta program initially aligns all the Cα atoms of the two structures using a least-squares minimization algorithm. This results in the smallest possible RMS difference between the structures. Then the scalar displacement of each pair of Cα atoms (e.g., (2)VAL Cα of TFIIIA(iv) and Cα of TFIIIA(ivc)) is plotted.

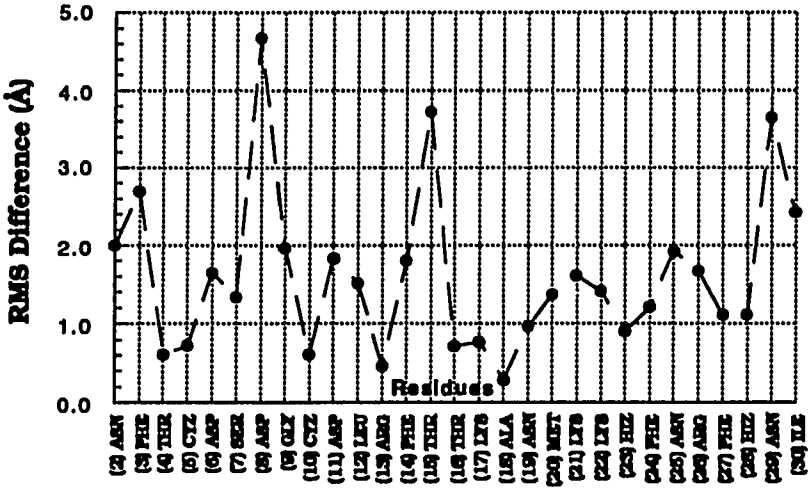


The total rms deviation of the Cα atoms between 50 psec time averaged TFIIIA(iv) (sec. 2.3.1) and TFIIIA(ivc) was 1.42 Å. RMS deviations may be used to identify large structural differences between two proteins. However, the RMS deviation may be misleading because regional flexibility may bias the molecular alignment and thus may not be a reliable means in which to assess structural differences. For example, the largest RMS differences are found in the cysteine loop, the "fingertip," the N and C terminus. These regions are conformationally heterogeneous because of the lack of consistent hydrogen bonds in the protein backbone and/or any intermolecular interactions with other residues in the zinc finger for structural stabilization. The sidechains of the cysteine loop are completely solvated and move freely during the simulation. Thus the difference in the time averaged conformation in this region bias alignment of the entire protein.

The RMS deviation of the Cα atoms between the 50 psec time averaged TFIIIA(iii) "even" zinc finger (sec. 2.3.1) and the constrained "even" zinc finger was 1.97 Å and 2.11 Å for the constrained finger at 20 to 70 psec and 100 to 150 psec, respectively. An analyses of the RMS deviation trajectory of the "even" zinc finger suggested that the structure may have equilibrated by 10 psec. However, the RMS deviation increased at 65 psec and re-equilibrated after 100 psec. Thus a structural comparison of TFIIIA(iiic) at these two 50 psec intervals is shown in Figure 26.

Figure 26

RMS difference between the two 50 psec structures of TFIIIA(iiic). The Molecular Similarity module of the Quanta program initially aligns all the Cα atoms of the two structures of TFIIIA(iiic) (20 to 70 psec and 100 to 150 psec) using a least-squares minimization algorithm. This results in the smallest possible RMS difference between the structures. Then the scalar displacement of each pair of Cα atoms (e.g., (2)ASN Cα at 10-60 psec and Cα at 100 to 150 psec) is plotted.



The RMS difference plot revealed that origin of the structural differences between the two equilibrated structures are found in the flexible regions of the finger. The greatest differences are located in the cysteine loop and the "fingertip" region (sec. 2.3.2.1) observed in the xfin31 structure. Thus these two structures are qualitatively similar, and the time averaged structure derived from 100 to 150 psec is representative structure of TFIIIA(iii).

The effect of constrained cluster

Table 11 describes the metal-ligand distances and geometry of the 5 atoms which constitute the zinc metal cluster during the simulation.

Table 11. Geometry of the metal ligands during the constrained simulations.

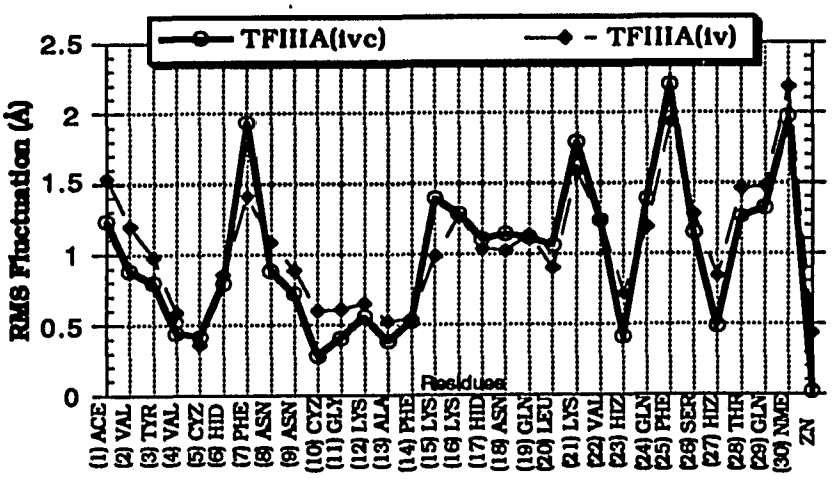
Parameter (Distance/Angle)	**TFIIIA(ivc)** (Ångstrom/degrees)	**TFIIIA(iiic)** (Ångstrom/degrees)
Sγ(1)-Zn	2.37 Å	2.36 Å
Sγ(2)-Zn	2.37 Å	2.37 Å
Nε2(1)-Zn	2.06 Å	2.05 Å
Nε2(2)-Zn	2.05 Å	2.05 Å
Sγ(1)-Zn-SG(2)	125.0° ± 0.7°	119.0° ± 0.8°
Sγ(1)-Zn-Nε2(1)	106.7° ± 0.8°	106.3° ± 0.9°
Sγ(2)-Zn-Nε2(1)	105.5° ± 0.7°	104.4° ± 0.8°
Sγ(1)-Zn-Nε2(2)	109.9° ± 0.9°	109.9° ± 0.9°
Sγ(2)-Zn-Nε2(2)	112.9° ± 0.8°	116.2° ± 0.9°
Nε2(1)-Zn-Nε2(2)	90.0° ± 0.8°	92.9° ± 0.7°

Because of the large harmonic force constant added to the distance dependent parameter, the metal to ligand distances did not vary during the simulation. In addition, the cluster bond angles which included the zinc metal ion had reduced angular movement. Although only the metal to ligand distances were effectively "frozen" in the "constrained" TFIIIA(iiic) and TFIIIA(ivc) simulations, the cluster bond angles also had much less RMS fluctuations (<1°) than the "unconstrained" simulations of TFIIIA(iii) and TFIIIA(iv) (5-7°). Because the dynamic bond stretching motion of the metal to its ligand is coupled to the bond angle bending motion (sec. 2.2.2), constraining the covalent bonds also effectively constrained the bond angles.

The consequences of the metal cluster constraints on the "odd" zinc finger are shown in Figure 27.

Figure 27

RMS fluctuation of TFIIIA(iv) and the constrained finger TFIIIA(ivc)

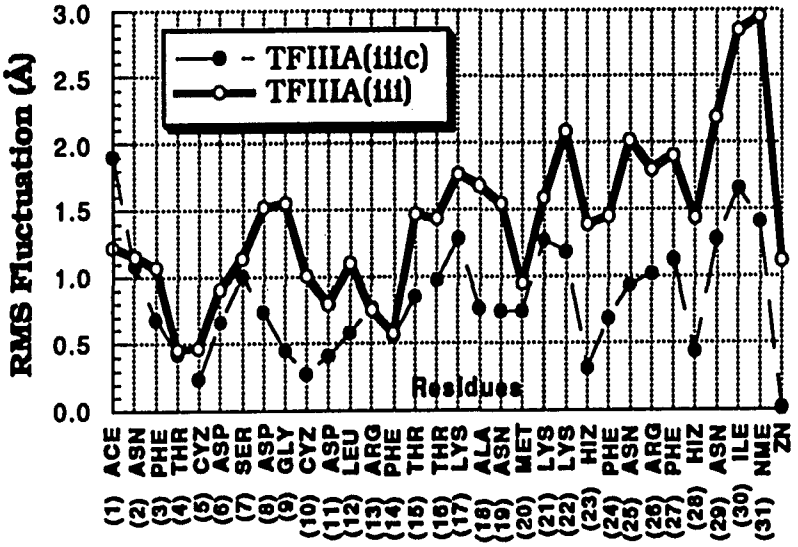


As predicted, the low RMS fluctuation of zinc in the TFIIIA(iv) simulation observed earlier was reduced to near zero. The constrained simulation did not significantly decrease the low fluctuations of the conserved cysteines and histidine ligands in the TFIIIA(iv) MD simulation. However, one residue (F7) in the cysteine loop of TFIIIA(ivc) and one residue (K15) in fingertip region of TFIIIA(ivc) had greater RMS fluctuations than the unconstrained simulation. These regions were previously identified as flexible regions in the xfin31 simulation (sec. 2.3.2.1). The overall TFIIIA(ivc) zinc finger mean fluctuation was 1.02 Å (S.D. 0.56) and not significantly different from the unconstrained simulation 1.05 Å (S.D. 0.43).

In contrast, the constrained zinc cluster had a more pronounced effect on the "even" TFIIIA zinc finger. Figure 28 compares RMS fluctuation of the amino acids of the constrained "even" zinc finger TFIIIA(iiic) simulation to the TFIIIA(iii) simulation described earlier.

Figure 28

RMS fluctuation of TFIIIA(iii) and the constrained finger TFIIIA(iiic)



Obvious differences in fluctuations are observed for the CYZ(5, 10) and HYZ(23, 28) residues.

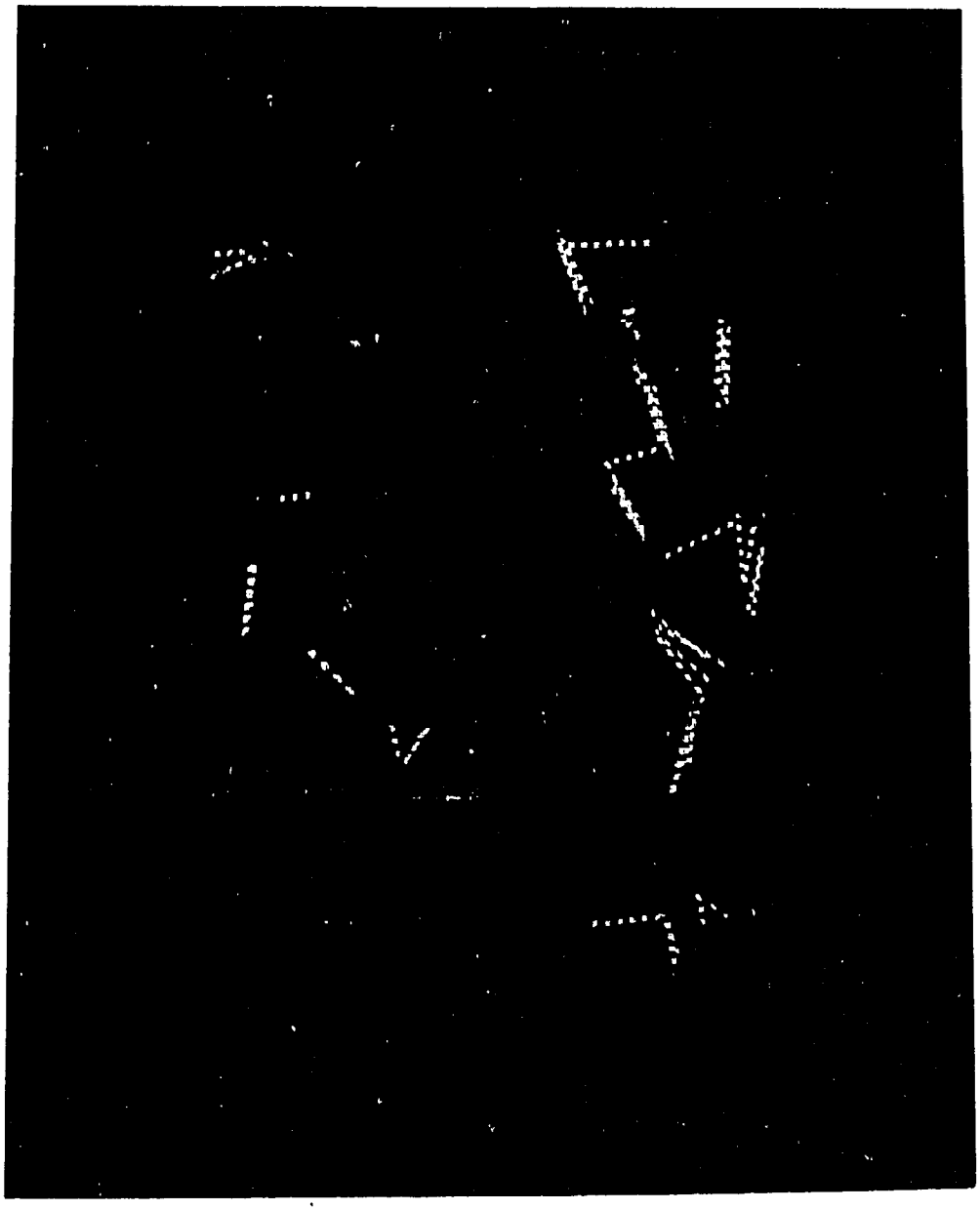
Table 10. Comparison of the RMS metal ligand fluctuations of the TFIIIA(iii) and TFIIIA(iiic) simulations

Residue (position)	TFIIIA(iii) Å	TFIIIA(iiic) Å	Difference Å
CYZ (5)	0.47	0.23	0.23
CYZ (10)	1.01	0.28	0.73
HIZ (23)	1.39	0.32	1.07
HIZ (28)	1.43	0.44	1.00
Zn	1.11	0.02	1.10

Amino acids fluctuations were also substantially reduced near the C terminus. The mean fluctuation for the TFIIIA(iiic) zinc finger was 0.83 Å (S. D. 0.43) while the unconstrained simulation's mean was 1.41 Å (S.D. 0.59). A student t-test confirmed that the difference between the two means was significant. Thus the flexibility observed in TFIIIA(iii) is removed once the zinc cluster environment becomes rigidified with fixed atomic positions and geometry. In Figure 29, five 10 psec time averaged structures of TFIIIA(iiic) taken during the 50 averaged time period are superimposed. The heterogeneity once observed in the TFIIIA(iii) simulation are no longer present when the zinc cluster is "frozen."

Figure 29

Five 10 psec average structures of TFIIIA(iiic) are superimposed. The hydrogen bonds are the dashed white lines. The helical portion of the finger is red.

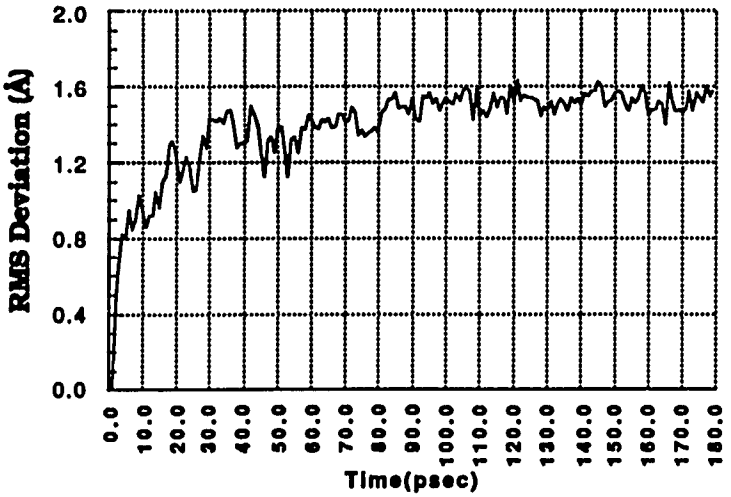


TFIIIA(iiimf) Simulation

In order to probe whether the flexibility of the "even" zinc finger is a physical property of the HX4H spacing only and not of its primary structure, a single residue deletion mutation of TFIIIA(iii), TFIIIA(iiimf) was simulated under the same conditions as the other zinc finger simulations. The RMS deviation of the zinc finger mainchain atoms (N, Cα, C, O) during the simulation compared to the initial structure is shown in Figure 30.

Figure 30

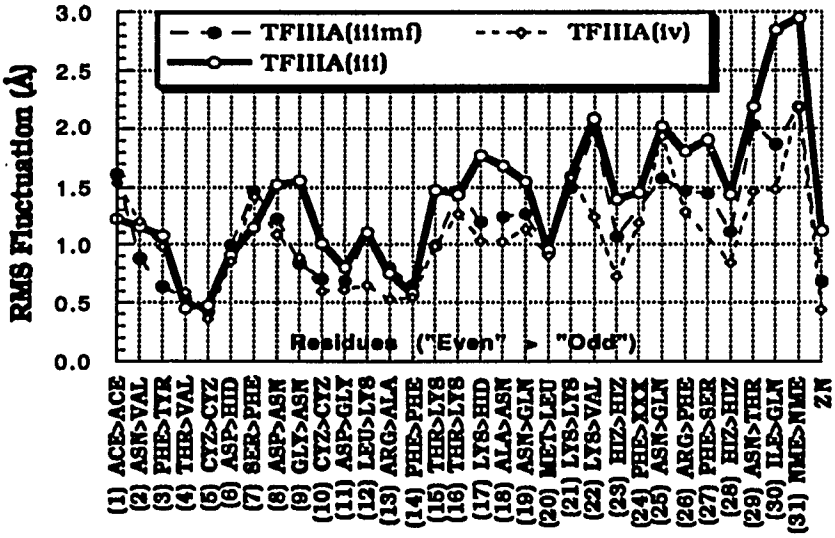
RMS deviation of TFIIIA(iiimf)



The total MD simulation time was 180 psec. The average RMS deviation was 1.47 Å (S.D. 0.23). A time averaged 50 psec structure was derived from 120 to 170 psec. The RMS fluctuation of each amino acid is plotted in Figure 31.

Figure 31

RMS fluctuation for each amino acid of TFIIIA(iv) and TFIIIA(iii) (from Figure 16.) and TFIIIA(iiimf). The mutated residue, phenylalanine 24, is represented by XXX.



Deleting one spacing between the conserved histidines reduced the flexibility observed in the wildtype "even" zinc finger. The mean fluctuation for the TFIIIA(iiimf) zinc finger was 1.18 Å (S.D. 0.46) while the TFIIIA(iii) mean was 1.41 Å (S.D. 0.59). An unpaired t-test indicated that the differences between means were statistically significant to 0.08 level for the entire finger. The mean RMS fluctuation for the helical residues was 1.44 Å (S.D. 0.42) for TFIIIA(iiimf) and 1.77 Å (S.D. 0.53) for TFIIIA(iii) and the mean difference was significant to 0.05 level.

Discussion

Previous limited molecular dynamics of zinc fingers used to refine 2D NMR data[32, 33, 36, 37], and an MD simulation of a solvated zinc finger[73] assumed a static zinc metal cluster in a tetrahedral geometry with fixed experimentally determined zinc ion to Nε2 and Sγ distances. For zinc fingers of the HX3H type, such an assumption may not have affected the interpretation of the structure. The MD simulations of "odd" zinc fingers (e.g., TFIIIA(iv), xfin31, and TFIIIA(iiimf)) have consistently demonstrated the reproducibility of the classic zinc finger folding and relatively low atomic fluctuations. Constraining the metal cluster in TFIIIA(ivc) was shown to have minimal effect on the results.

However, a simulation of a HX4H zinc finger clearly confirmed the necessity for a complete and accurate representation of all the components in the MD simulation including the zinc cluster.

TFIIIA(iiic) simulation have shown that increased fluctuations of the "even" zinc finger are not observed if the constituents of the cluster are simply fixed in their tetrahedral geometry. Thus a description is required that allows for time dependent changes in the structure of the cluster under the influence of the conflicting strains induced by the interaction with the HIS residues and the maintenance of a helical motif. Such a description leads to a better understanding at the molecular level of the effects of four residue spacing to the metal cluster and to the entire finger protein.

Recently, two HX4H "even" zinc fingers solution structures were solved from 2D NMR data. Without any assumptions about the metal cluster, Xu et al[40] used only a distance geometry algorithm to construct the coordinated of ADR1a, an "even" zinc finger. Their results suggested that this particular zinc finger of ADR1 existed as at least two similar conformations in solution. In contrast, Kochoyan et al[36]did not observe this property in their 2D NMR "even" zinc finger ZFY. However they used a distance geometry/simulated annealing algorithm which assumed a fixed tetrahedral metal cluster for their zinc finger. The TFIIIA(iiic) simulation caried out for this work have shown that a constrained or fixed metal cluster may reduce the flexibility of the "even' zinc finger. In addition, the ZFY "even" zinc finger has a slightly different positioning of the conserved hydrophobic core residues which may affect the dynamic properties of the "even" zinc finger.

CHAPTER 4

Introduction

The structural roles of the conserved amino acids of the classical (CC/HH) zinc finger motif have been elucidated over the years. For example, the two cysteines and two histidines chelate zinc [27, 76, 26] and the hydrophobic residues form a small hydrophobic core [33, 34, 39] at the "fingertip" region (transition from the β strand to the α helix). In order to probe whether the conserved residues were the only determinants of folding, recently Michael et al[38] constructed a "minimalist" zinc finger with all the variable amino acids (X_{nn}) mutated to alanines. They were able to show that this "minimalist" finger folded into the classical zinc finger structure and retained the proper tetrahedral coordination with zinc metal.

In contrast, less attention have been focused upon amino acid spacing between the conserved residues. Because these residues are found within definitive secondary structures (ββα), pertubation in their spacing may have structural consequences not readily deduced from an initial inspection of the primary sequence. For example, Xu et al [40] attributed the flexibility of ADR1a observed from the 2D NMR data to the four residue spacing between the conserved histidine ligands in the α helix. Because an α helix is a highly ordered structure with every 3.6 residues on same face of the helix, a four residue spacing (compared to a three residue spacing) would not "easily" allow both histidines to appear on the same "face." The consequences of a HX4H

spacing to the third TFIIIA zinc finger are described in chapters 2 and 3. Briefly, the four spacing imparts a new dynamic property to the zinc finger-flexibility.

The importance of the spacing between the histidine ligands is hinted in the systematic arrangement of three "odd" HX3H or four "even" HX4H residue spacing zinc fingers in TFIIIA. Fingers 1, 2, 4, 5, 7, 9 are of the "odd" type while fingers 3, 6 and 8 are of the "even" type. In order to relate the structural consequences of an "even" zinc finger observed in the molecular dynamics simulations to its functional importance, DNA-binding, the third zinc finger of TFIIIA was mutated from an "even" finger to an "odd" one, TFIIIA(iiimf), and its binding compared to the wildtype TFIIIA. Previous computational simulations suggested that a mutant "odd" third TFIIIA zinc finger would be less flexible than the wildtype "even" one.

In order to investigate whether the spacing between the histidine ligands is important for sequence-specific recognition, the third "even" zinc finger of TFIIIA was mutated to an "odd" zinc finger with the extension overlap method [77]. Because zinc-finger motifs can be translated and folded in an *in vitro* reaction[78], the DNA affinity of the mutated and wildtype translated TFIIIA protein was assessed with gel retardation studies with the 69mer DNA representing the internal control region (ICR) of the 5S gene.

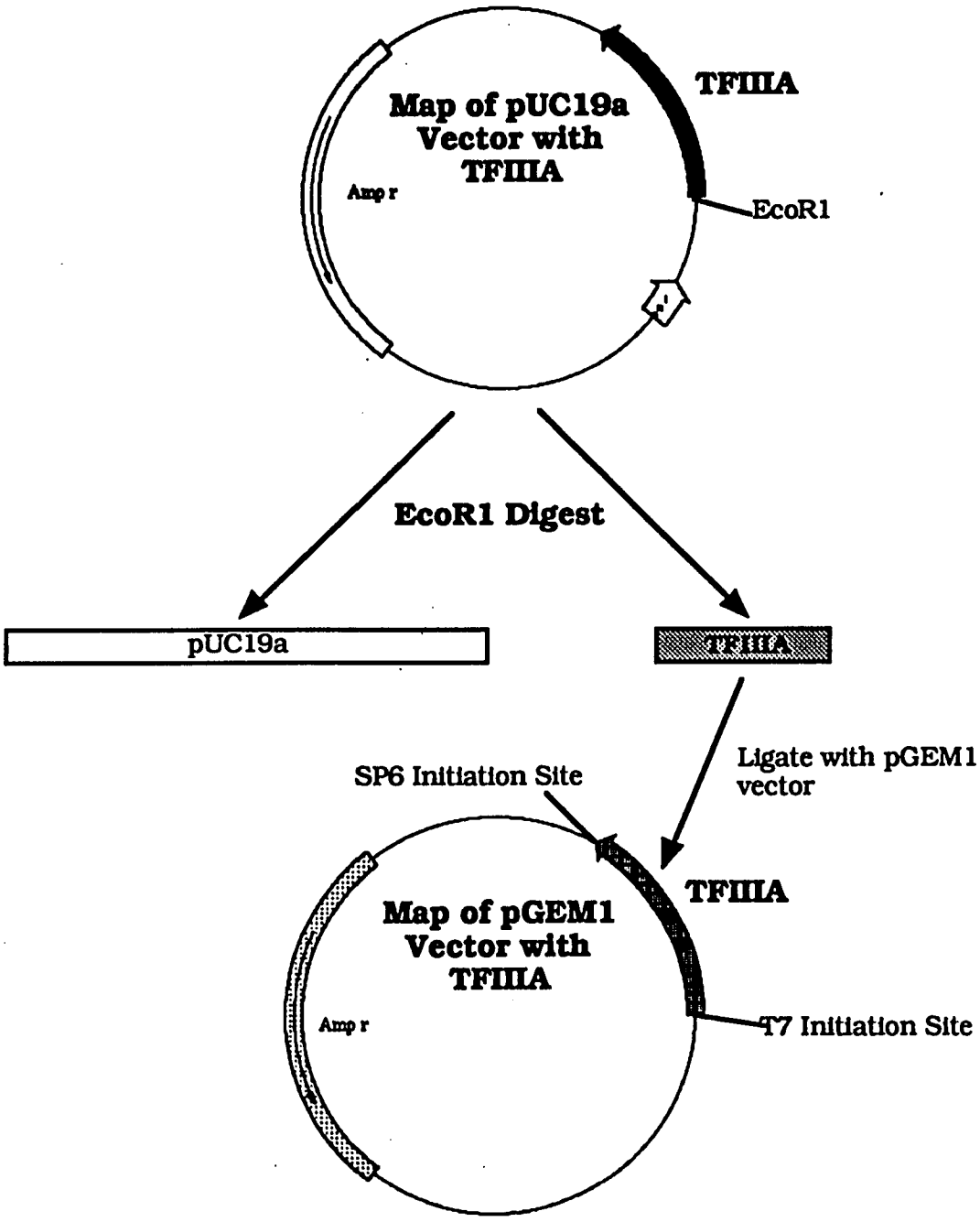
Materials and Methods

Subclone of the TFIIIA cDNA from Bieker into the pGEM1 vector

The cDNA of transcription factor IIIA[22] was generously donated by Dr. James Bieker (Mount Sinai School of Medicine, Brookdale Center for Molecular Biology). The original TFIIIA construct inserted in a pUC19a vector was subcloned into a pGEM1 vector for *in vitro* transcription. The subcloning strategy is depicted in Figure 32.

Figure 32

Subcloning strategy



The original plasmid was transformed into MC1000 *E. Coli* competent cells and plated onto LB (Luria-Bertani Medium)[79] with ampicillin (20 μg/ml) agar plates for overnight incubation at 37° C. Six colonies were selected and inoculated in 10 mls of TB (Terrific Broth) and ampicillin (20 μg/ml). The colonies were shaken overnight in a incubation chamber at 37° C. The following day, plasmid minipreps[80] were performed on each sample to obtain plasmids from each colony. The following protocol was changed from the original procedure described by Birnbohm: 1) after the addition of ARS III solution and 30' minute incubation, the samples were pelleted, 2) the supernatant was phenol:chloroform extracted 2X, 3) the DNA was ether-extracted and precipitated from solution with 95% ice cold ethanol. The isolated plasmids were electrophoresed in 1% agarose gel to insure the TFIIIA cDNA was retained along with the pUC19a vector.

One plasmid sample preparation (20 μgs of DNA) was chosen for overnight digestion with EcoR1 (New England Biolabs) endonuclease at water bath temperature of 37° C. After purifying the digested products, the two DNA fragments were subjected to 1% agarose gel electrophoresis. A 1.5 kb fragment (TFIIIA) was identified and recovered from the gel with DEAE paper electroelution(Carter Bancroft's lab protocol). The TFIIIA construct was then recovered from the DEAE paper, and ethanol-precipitated .

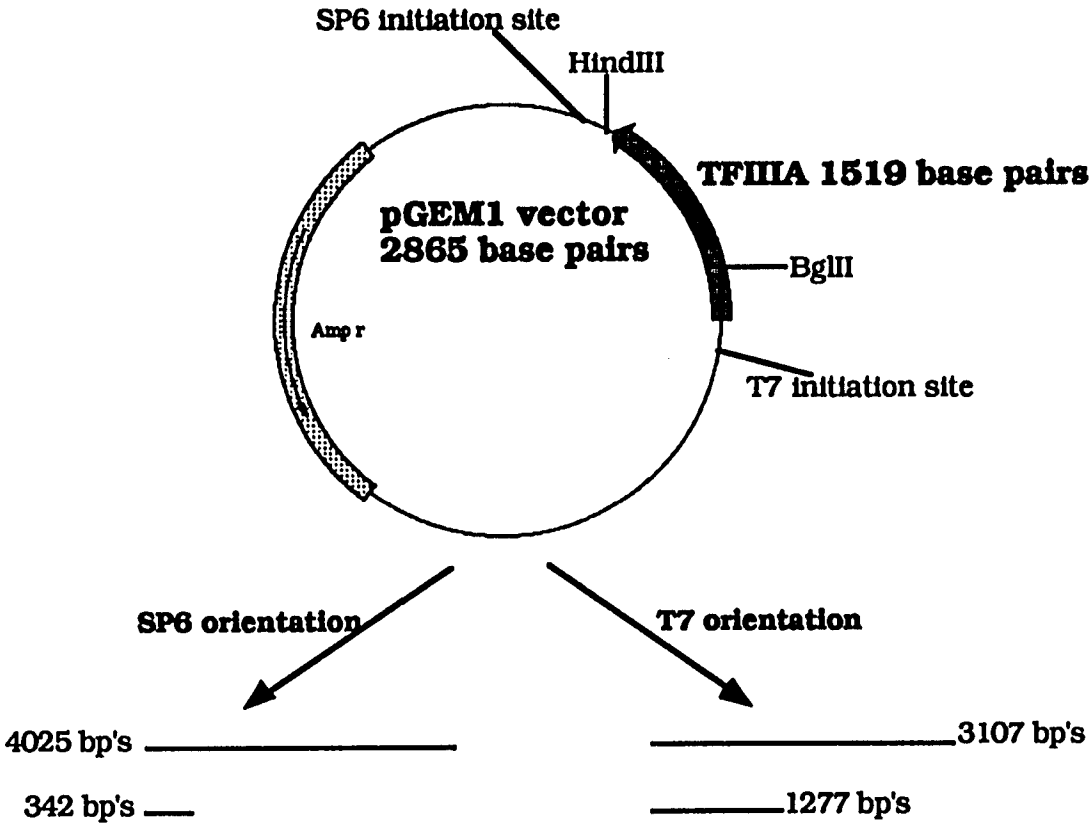
10 μgs of pGEM1 vector (Promega) containing both T7 and SP6 *in vitro* transcription initiation sites was linearized with EcoR1 endonuclease for 6 hrs. at 37° C and phosphatase (Boehringer

Manneheim)[79] treated to increase the ligation efficiency. After the phosphatase treatment, the linearized plasmid was gel-purified (1% agarose) and recovered with DEAE paper. The TFIIIA insert and the pGEM1 vector were added to a ligation reaction[79] and subsequently transformed into MC1000 *E. coli* competent cells. 10 colonies were selected from the LB-amp plates (overnight incubation at 37°C) and each clone was inoculated into 10 mls of TB-amp (20 μg/ml) for overnight incubation at 37°C. After the plasmid miniprep procedure was performed on each clone, each plasmid preparation was subjected to EcoR1 digestion for 6 hrs. at 37° C. The DNA from each sample was electrophoresed in 1% agarose in order to determine which clones contained the insert. Six clones were positive.

20 μgs of DNA from each of the six samples was digested with HindIII (New England Biolab) and BglII (New England Biolab) overnight at 37° C. TFIIIA contains a unique BglII restriction site near the 5' end (nucleic acid base position 342) and no HindIII site. However, pGEM1 vector has a unique HindIII site 3' to the EcoR1 site in the polylinker. Thus a double restriction reaction would be able to distinguish the two orientations of the insert, SP6 or T7, (see Figure 33).

Figure 33

Determination of insert orientation in pGEM1 vector



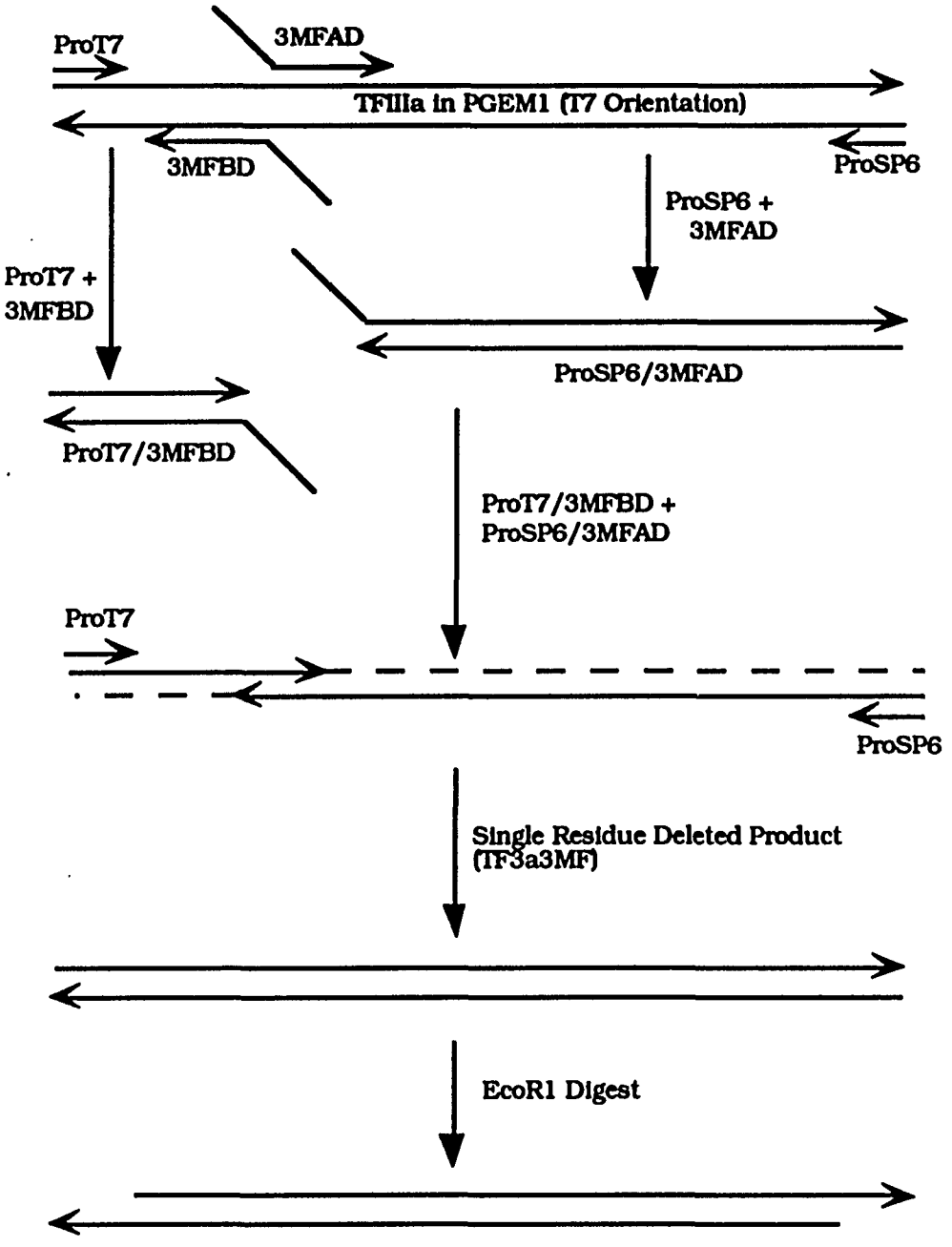
Each sample was electrophoresed (1% agarose gel) and the orientation of the insert was determined. Two clones were selected, one in T7 orientation and the other in SP6, and inoculated in 1 liter TB-ampicillin (20 μg/ml) for 16 hrs incubation at 37° C. Each clone was then harvested for its plasmid and maxiprep[80] procedure was performed. The DNA was sequenced with Applied Biosystems (Model 373A, version 1.0.2).

PCR (Polymerase Chain Reaction) Deletion Mutant (TFIIIAmf)

A site-directed single residue deletion mutant of TFIIIA was made with the extension overlap method[77]. The phenylalanine at amino acid position 94 was deleted to form the mutant protein TFIIIA(iiimf). The strategy for PCR mutation is described in Figure 34.

Figure 34

Single deletion mutation strategy with PCR



PCR oligonucleotide primers (Table 13) were synthesized by GENSET (France). The T7 orientation pGEM1 with TFIIIA insert plasmid was used as the template DNA.

Table 13. The DNA Sequence of PCR Primers

PCR Primer	Sequence
PROT7	5'-TAATACGACTCACTATAGGG-3'
PROSP6	5'-GATTTAGGTGACACTATAG-3'
3MFBD	5'-TCTGTTGTGCTTCTTCAT-3'
3MFAD	5'-AAGCACAACAGATTCCAT-3'

The PCR reaction was conducted in a thermocycler (Perkin Elmer Cetus) for 25 cycles. The cycling conditions were as follows : melting at 94°C for 60 sec., annealing at 37°C for 60 sec., and primer extension at 72°C for 90 sec. Reactions were performed in 25 µl volume with AmpliTaq (Perkin Elmer Cetus) as prescribed by the manufacturer. PCR fragments were purified with 1% low melting agarose gel electrophoresis and recovered with GENECLEAN II Kit(Bio101, La Jolla, CA).

The final PCR product, TFIIIA(iiimf), was digested by EcoR1 for 6 hrs. at 37°C and gel electrophoresis (1% agarose) purified for ligation reaction with the phosphatase treatedpGEM1 vector described previously. DH5α *E. coli* competent cells (Life Technologies) were transformed according to the manufacturer's protocol and plated onto LB-amp plates for overnight incubation at 37°C. Five colonies were selected and inoculated into 10 ml TB-amp for overnight incubation. Three of the five colonies contained the TFIIIA insert (EcoR1 digestion) and their orientations were determined from a double digestion by HindIII and BglII. A T7 orientation clone was selected and grown in 1 liter TB-amp for 16 hrs. at 37°C. The cells were harvested and their plasmids were isolated the maxiprep procedure.

In vitro transcription/translation

50-µg aliquots of plasmid, both TFIIIA (wildtype) and TFIIIA(iiimf) mutant, were linearized with a variety of restriction endonucleases to form varying length transcripts for translation, see Table 14.

Table 14. Different lengths of mRNA can be transcribed. The endonucleases linearize the *in vitro* transcription vector at different sites of the insert resulting in transcripts of varying lengths. Translation of these inserts result in proteins with different numbers of zinc fingers.

Restriction Endonuclease	Base Pair Site	Corresponding Finger Length
HindIII	1518	9
Bsu36I	1121	9
XhoI	712	7
Sau96I	608	6
DraI	494	4
BglII	343	3

The linearized plasmids were transcribed with the MEGAscript Kit T7 RNA polymerase kit (Ambion) according to the manufacturer's protocol. The RNA transcripts were translated either with wheat germ extract (Bancroft's lab protocol) or nuclease-treated rabbit reticulolysate system (Promega). Both *in vitro* systems were -MET (all amino acids were present except for methionine); thus ^{35}S labeled methionine (New England Nuclear) was used to titrate the amount of transcript required for optimal translation. Once the amount of transcript was determined, non-radiolabeled methionine was used according to protocol. Yield was determined by ß emission counting after TCA (Trichloroacetic Acid) precipitation. Radiolabeled translated products were subjected to PAGE (polyacrylamide gel electrophoresis) in 15% SDS (sodium dodecylsulfate) denaturing gels[79] at 30 mA for approximately 3 hours. The gel was dried with EASY BREEZE gel drier (Hoefer Scientific instruments) for exposure to Kodak XAR-5 film at -80° C.

Gel retardation

A 69 base probe 5'-CCGATCTCGTCTGATCTCGAAGCCAAGCAGGGTCGGGCTGGTTAGTACTTGGATGGGGACCGCCT-3' was synthesized along with its complementary strand. Each strand of DNA was urea-gel-purified and then the two complementary strands were annealed[79]. The 69mer probe was subsequently labeled with ^{32}P-GTP (New England Nuclear) taking advantage of the two cytosine overhangs at each end.

The translated product was immediately added to binding reactions which contained 20 mM-Hepes (pH7.5), 12% glycerol, 50 mM-KCl, 1 mM-$MgCl_2$, 5 mM-DTT, 1 µg of dAdT or dIdC (non-specific DNA), and 50 µM-$ZnCl_2$ in a volume of 20 µl[54]. The reaction was allowed to incubate at room temperature for 20 minutes. 1 ng of radiolabeled probe was then added to the mixture and incubation was allowed to continue for 30 minutes at 4°C. Samples were subjected to PAGE (polyacrylamide gel electrophoresis) in 10% non-denaturing polyacrylamide gels (19:1, acrylamide: bis) containing 90 mM Tris-borate (pH 8.3) at 200 volts for 4 hours at 4°C. The running buffer was also at 90mM Tris-borate and the gel was pre-electrophoresed for 1 hour. Gels later were exposed to either Kodak XAR-5 film at -80° C with intensifying screen or phosphoimager (Molecular Dynamics, Sunnyvale, Ca) for quantification.

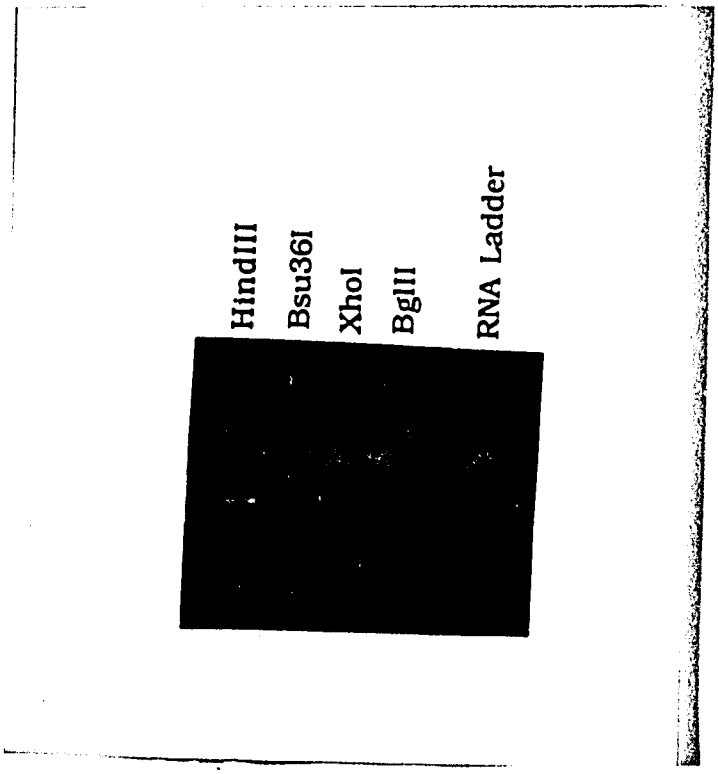
Results

In vitro transcription/translation

Simply by using different restriction enzymes to linearize the TFIIIA plasmid, RNA transcripts of varying lengths can be transcribed by the *in vitro* transcription system as long as the T7 RNA initiation site is retained in the fragmented DNA, see Figure 35.

Figure 35

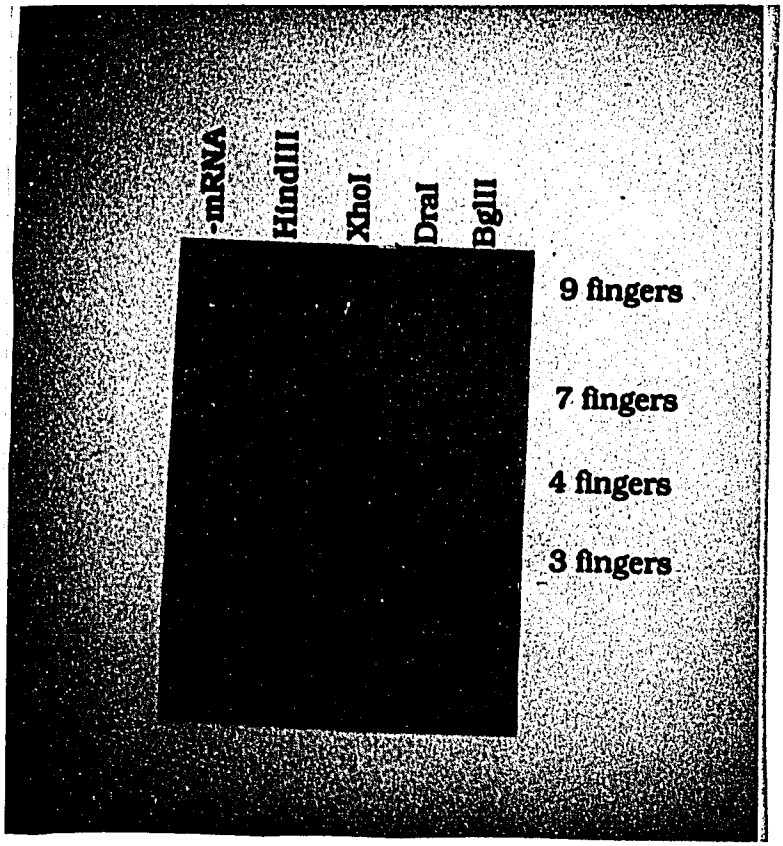
RNA transcripts after *in vitro* transcription reaction were run on a 1% agarose RNA gel. The restriction enzyme used for the linearization of the plasmid is described for each lane.



Because the DNA-binding domain begins at the N terminus and represents approximately 80% of the entire TFIIIA protein (39.7 kD), varying the length of transcripts by an informed selection of restriction endonucleases for linearization resulted in truncated TFIIIA proteins. Thus proteins with varying numbers of zinc fingers (in sequential order) can be translated, see Figure 36.

Figure 36

15% SDS denaturing protein gel of the *in vitro* (wheat germ extract) translated products. The restriction enzyme used for the linearization of the plasmid is identified for each lane. A "mock" translation (-mRNA) as a control was performed simultaneously with other translation reactions. The numbers of zinc fingers calculated from the length of the original transcript are identified on the right.



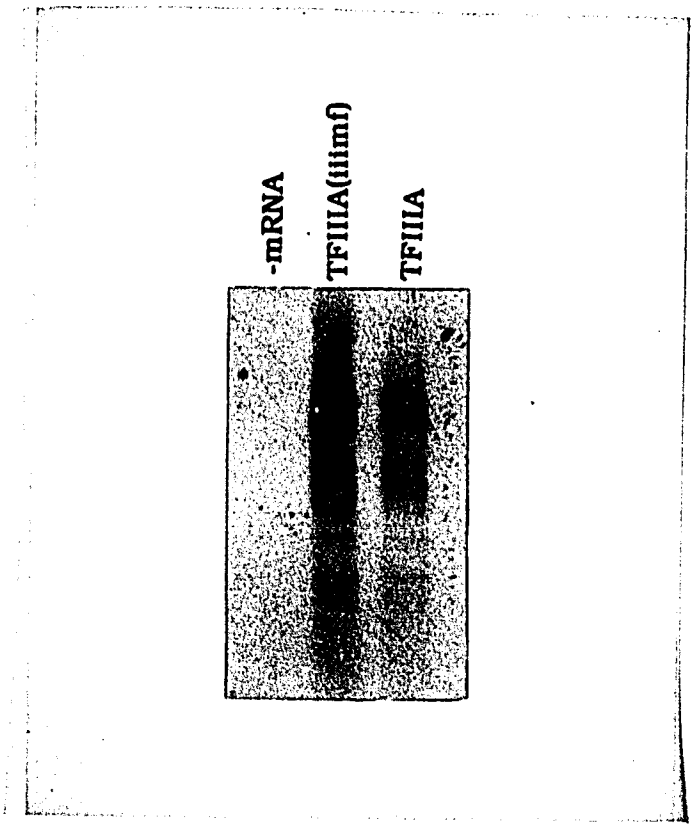
For example, linearization of the plasmid with XhoI restriction enzyme resulted in a transcript of 712 base pairs which accounts for zinc fingers 1 through 7. Premature termination during the translation process also produced minor forms of further truncated proteins.

A single residue (phenylalanine 94) in the spacing region HX4H of the third zinc finger of TFIIIA was deleted with the extension overlap method. Such a single-residue deletion in this region converts a wildtype "even" zinc finger into a mutant "odd." According to our findings, this will reduce the flexibility of this finger. As to the effect of such a deletion on DNA binding, an inspection of the co-crystal structure of Zif268 suggested that residues at this position (+1 residue from the first histidine ligand) do not participate in direct contact with the DNA. Moreover, the last proposed sequence-specific contact is -1 position away from the first histidine ligand. Because of the structural constraint of an α helix (3.4 residues/turn), the wildtype phenyalanine would be positioned away from the major groove of the DNA if this zinc finger bound in a Zif268-like manner. Finally the aromatic sidechain of phenyalanine is unlikely to form any strong sequence-specific DNA contacts (e.g., hydrogen bonding, electrostatic interaction, etc.). Thus phenylalanine at this position should play a minimal role in direct DNA interaction.

Figure 37. shows the result of *in vitro* translation of the TFIIIA(iiimf) "even"-to-"odd" mutant and wildtype TFIIIA. Both plasmids, TFIIIA and TFIIIA(iiimf), were linearized with HindIII and *in vitro* transcribed to produce the full-length transcript. The resultant transcipt was subjected to *in vitro* wheat germ translation.

Figure 37

In vitro translated product of TFIIIA(iiimf) RNA transcript, the single amino acid deletion (phenylalanine) between the HisX4His spacing region of the "even" third zinc finger, and TFIIIA wildtype transcript. Both plasmids were linearized with HindIII thus producing the full-length finger protein.



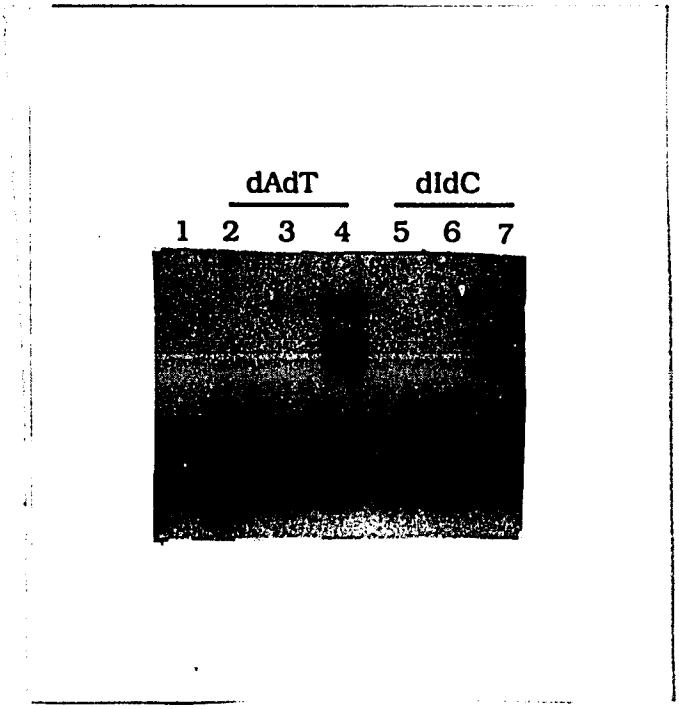
The protein gel was not able to resolve a single residue difference in molecular weight between the two proteins (39.719 kD for TFIIIA and 39.572 kD for TFIIIA(iiimf). However, since both bands co-migrated to the same position, it is likely that the full length mutant and wildtype proteins were translated.

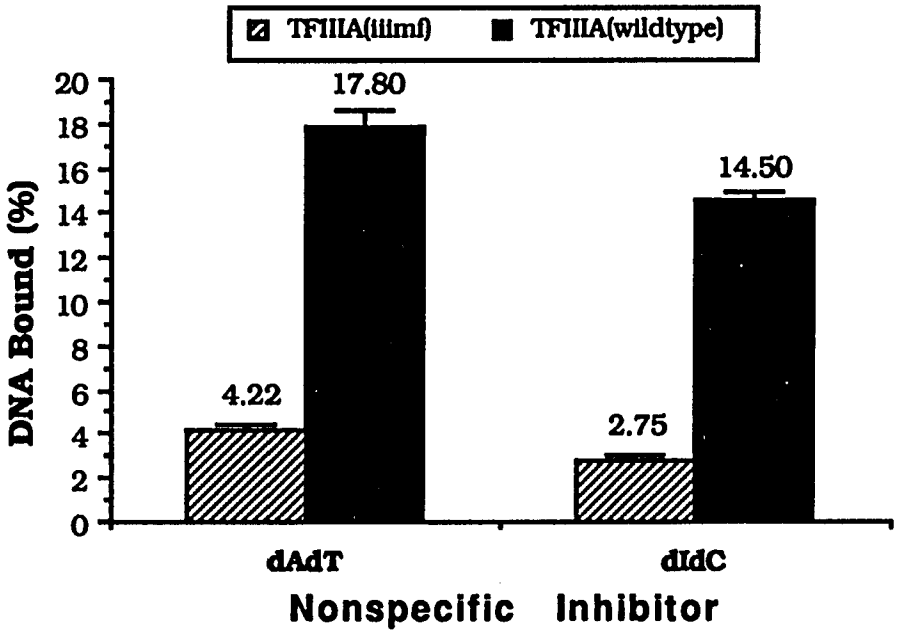
Gel Retardation

Immediately after the cell free translation reaction, the translated products were added to the DNA-binding reactions and loaded onto the non-denaturing electrophoresis gel for retardation assay to assess binding. If the protein were bound to the DNA probe, radioactivity would be observed at an additional band with an apparent higher molecular weight. Figure 38 indicates that the mutant protein did not bind with the same affinity as the wildtype.

Figure 38

Gel retardation of *in vitro* translated products of TFIIIA(iiimf) and TFIIIA: lane 1) 69mer probe only, lanes 2, 5) -mRNA, lanes 3, 6) TFIIIA(iiimf), lanes 4, 7) TFIIIA wildtype. A 10% non-denaturing polyacrylamide gel was used. 1 μg of nonspecific dAdT was added to samples in lanes 2-4, and 1 μg nonspecific dIdC was added to samples in lanes 5-7. The amount of nonspecific DNA represented 1000 fold excess relative probe used (1ng). The amount of bound probe was quantitated by phosphoimaging (below).


dAdT
dIdC
1 2 3 4 5 6 7


TFIIIA(iiimf)
TFIIIA(wildtype)
DNA Bound (%)
20
18
16
14
12
10
8
6
4
2
0
17.80
14.50
4.22
2.75
dAdT
dIdC
Nonspecific Inhibitor

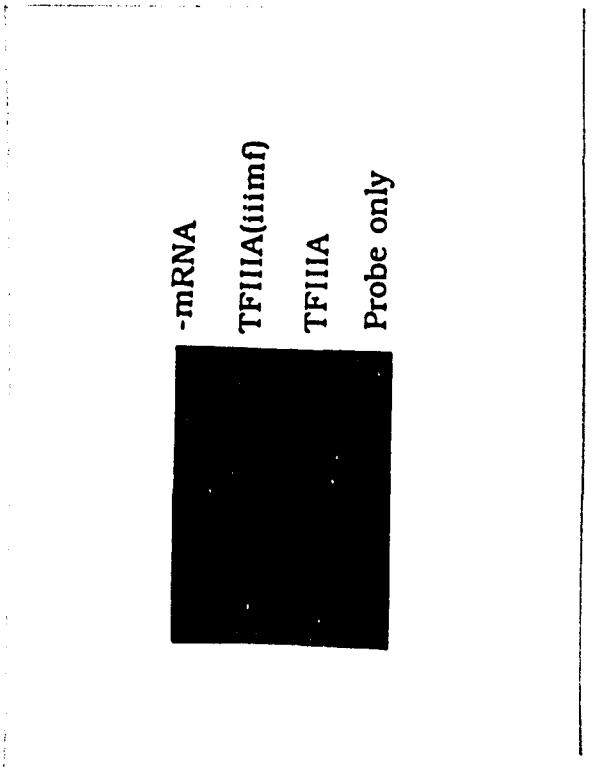
The gel-shift of the wildtype protein (lanes 4 and 7) yielded two band shifts, a lower more diffuse band and an upper single band. In contrast, the mutant protein had only the upper band. The lower band is presumably the result of premature termination or inefficient translation of the full length protein observed in the protein gel of the *in vitro* translated products (Figure 37). This resulted in the binding of inadvertantly truncated wildtype TFIIIA proteins with different numbers of zinc fingers to the probe. The lower molecular weight of these proteins may have produced the more diffuse lower band. Nevertheless, all binding affinity analyses were performed on the upper band only.

The TFIIIA(iiimf) mutant had either 4.21 or 5.27 fold (poly•dAdT or poly•dIdC) lower affinity than the wildtype TFIIIA. Both non-specific competitor DNA's were in excess (1000X) suggesting that the binding to the 69mer probe is specific and that the binding to poly•dIdC showed overall lower affinity. Because TFIIIA is a GC-rich binding protein, poly•dIdC was better able to compete for the TFIIIA binding site than poly•dAdT.

In order to determine if the lower shifted band observed in Figure 38 was due to a truncated TFIIIA wildtype protein, a seven finger TFIIIA protein was translated. Both the wildtype and mutant TFIIIA(iiimf) plasmids were linearized with the restriction endonuclease XhoI. The two truncated transcripts were translated *in vitro* and Figure 39 shows the result of non-denaturing binding gel electrophoresis with the radiolabelled probe.

Figure 39

Binding of the seven zinc finger construct. Linearization of the plasmid with XhoI produces a transcript which encompasses the first seven fingers in sequential order : lane 1) -mRNA, lane 2) TFIIIA(iiimf) mutant translation, lane 3) TFIIIA wildtype translation, lane 4) probe only.

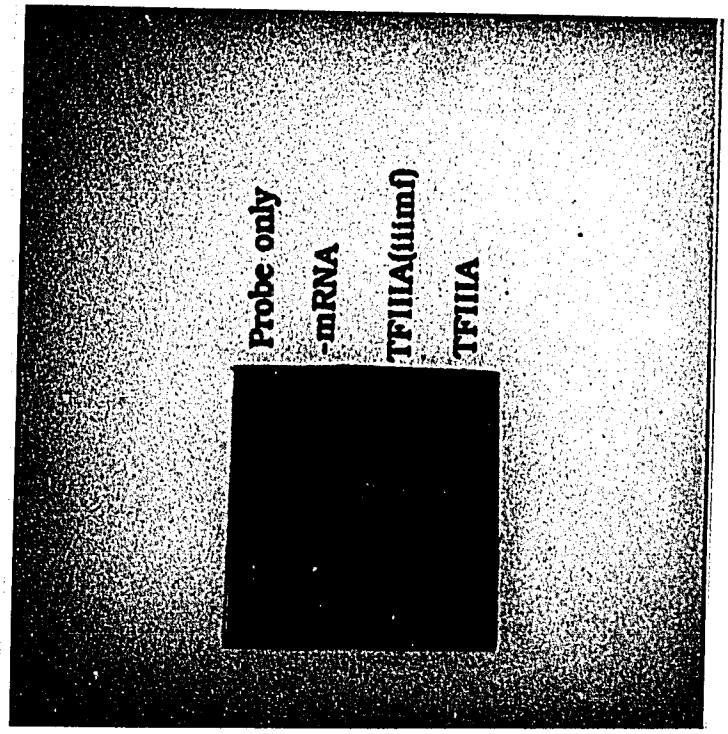


Compared to Figure 38, only one band shift was observed for the wildtype protein (lane 4). This result supports the hypothesis that the lower band of the wildtype protein shift observed in Figure 38 represented truncated protein bound to DNA. However, the mutant TFIIIA(iiimf) 7 zinc finger construct created from the XhoI digestion did not result in any gel shift. The non-specific DNA was able to compete more easily with the 5S DNA probe perhaps due to TFIIIA(iiimf)'s inability to form sufficient DNA contacts in a seven finger construct.

Figure 40 compares the binding of the 3 amino terminal zinc fingers of the wiltype and mutant proteins. If flexibility is an important property of the third zinc finger of TFIIIA because it "allows" for the proper alignment of the subsequent zinc fingers of TFIIIA (fingers 4 through 9), then a three finger mutant construct would not be affected by the mutation. Thus in contrast to Figure 38 and Figure 39, the hypothesis predicts similar binding of a mutant and wildtype stretch of the first three zinc fingers of the TFIIIA protein.

Figure 40

Binding of the three N-terminal fingers. Linearization of the plasmid with BglII produces a transcript which encompasses the first three fingers only. 5mM $MgCl_2$ was used in the binding and buffer solution. Lane 1) probe only, lane 2) -mRNA translation, lane 3) TFIIIA(iiimf) mutant, lane 4) TFIIIA wildtype



Both the mutant and wildtype 3 finger proteins produced slight but similar gel shifts. According to the hypothesis, a single deletion mutation in the spacing region of the third "even" zinc finger would affect the proper positioning of the zinc fingers which follow the "even" finger and not its own (third finger) interaction with DNA. In a 3 zinc finger construct, there are no fingers following the mutant third finger; therefore, the affinity to DNA of the mutant and wildtype proteins would be similar.

Discussion and Conclusions

Recent experimental data of the TFIIIA-5S complex suggest that the binding of the 9 zinc fingers to the ICR occurs in a complex and interdependent manner (sec. 1.4). The complete 9 zinc finger-DNA interaction of TFIIIA represents a departure from the only available co-crystal structure of zinc fingers with DNA, the Zif268 protein, where all three fingers bind in a similar orientation. The preponderance of the footprinting data suggest that the three N terminal fingers bind in a Zif268-like manner. Each finger makes specific base-pair contacts in the major groove of the DNA while the entire group of 3 fingers "wrap" around the central axis of the DNA, following the major groove. In contrast, fingers 4-6 were proposed to bind parallel to the central axis of the DNA which would require finger 4 to cross the minor groove of the DNA.

However, an inspection of the primary sequence of the fourth finger would <u>not</u> have hinted that it would depart away from the "wrap

around" orientation of the first three fingers towards an interaction with the minor groove. The finger 4 sequence adheres to the canonical (CC/HH) zinc finger motif much like fingers 1 and 2 . Based upon the abundance of structural data on a variety of CC/HH zinc fingers (sec. 1.2.2), the (**Y/FXCX$_{2,4}$CX$_3$F/YX$_5$LX$_2$HX$_3$H**) motif results in structurally similar zinc fingers. Furthermore, analyses of crystallographic data of structurally similar DNA-binding motifs such as the helix-turn-helix motif bound to DNA suggest that the structural conservation is extended to a conservation of the mode of interaction with DNA[1]. Thus a structural interpretation of the TFIIIA-5S complex is insufficient by itself to account for the model proposed by recent experimental data.

However, the molecular dynamics simulations of fingers 3 and 4 (Chapter 2) suggest that these two fingers are differentiated by their dynamic property, the flexibility of the molecular structure. The greater flexibility of the third finger, especially in the C-terminal part of the helix, may "allow" finger 4 to reorient itself from the major groove and to a position crossing the minor groove. The proper positioning of finger 4 is likely to be required for the rest of the zinc fingers to "find" their respective binding site[52]. Thus flexibility of the third zinc finger is necessary for the specific interaction of TFIIIA with the ICR.

This property of flexibility in finger 3 was removed in a molecular dynamics simulation in which a single residue between the two conserved histidine ligands (sec. 3.3.2) was deleted.

In an analogous experiment designed to probe the hypothesis of the role of flexibility in the TFIIIA-5S complex, phenylalanine 94 was deleted from TFIIIA with the extension overlap method; thus converting an "even" third finger into an "odd" one, TFIIIA(iiimf). The deletion of a single residue in this spacing region, which is not proposed from the co-crystal data to interact directly with DNA, resulted in a 4 to 5 times lower overall affinity from the gel retardation experiments. According to the working hypothesis obtained from the computational simulations, the TFIIIA(iiimf) construct has a reduced flexibility. Consequently finger 4 is improperly positioned in the TFIIIA(iiimf) mutant, affecting the positioning of subsequent zinc fingers of TFIIIA. Thus a single deletion mutation in the "even" spacing region of finger 3 of the nine zinc finger TFIIIA(iiimf) protein "improperly" positions fingers 4 through 9 to the their respective binding sites resulting in significantly reduced affinity. However, since the mutation affects the proper positioning of finger 4 though 9 relative to the DNA, the binding of the first three zinc fingers of the mutant construct would remain unaltered.

CHAPTER 5

Model of the TFIIIA-5S Complex

Conformation of the 5S DNA

The ICR region of the 5S gene was constructed in a B-type DNA conformation. Earlier studies [49] had proposed that the 5S gene adopted an A-type conformation from an analysis of the DNAse I cleavage sites and the distribution of guanines; however Gottesfeld et al [81] more recently reported that the 5S DNA is B-like. They synthesized a 54-mer oligonucleotide corresponding to the TFIIIA-binding region of the DNA (+44 to +97) and compared its circular dichroism spectrum to reference spectra of A and B-form DNA. The spectrum of TFIIIA shared major features with the B-DNA from calf thymus, most notably the magnitude of the major positive peak. Furthermore, the spectrum remained relatively unchanged upon the addition of TFIIIA protein suggesting that protein binding did not dramatically alter the conformation of the DNA. The only high resolution co-crystal of a zinc finger-DNA complex[1] also indicated that the DNA was in B-like conformation.

Conformation of the zinc finger

The TFIIIA "even" and "odd" zinc fingers were constructed by molecular replacement of the MD-averaged "even" or "odd" fingers as

described in section 2.2.1. Fingers 1, 2, 5, 7, and 9 were constructed from the TFIIIA(iv) MD averaged structure, while finger 6 and 8 were constructed from the MD-averaged structure of TFIIIA(iii) .

5.1.3. General Features of the Model

A molecular model of the interaction between the nine fingers of TFIIIA and the ICR in accordance with aspects of the "modified wrapping around" model was constructed. As shown in Figure 41, the first three zinc fingers were positioned to bind to the 5S DNA in a manner analogous to the Zif268 co-crystal.

Figure 41

Zif268 co-crystal

Each individual zinc finger would contact 3 base pairs with the general sequential orientation of DNA in 3' to 5' direction relative to the N to C terminal orientation of the protein. Thus, the first finger of TFIIIA is positioned to bind to the most 3' end of the recognition sequence essentially by wrapping around the DNA in the major groove. The next set of three zinc fingers, fingers 4, 5 and 6, are positioned to bind in novel orientation (parallel to the central axis of the DNA), and TFIIIA fingers 7, 8 and 9 would again have a binding orientation similar to that of Zif268 structure.

The fingers positioned to bind as in the Zif268 complex are expected to form the type of contacts inferred from the consistency of specific amino acid positions of the zinc finger contacting the base pair of the DNA observed in all three zinc fingers of the co-crystal complex. The relatively straight-forward scheme for protein-DNA interaction is illustrated in Figure 42 which shows the residues in the first three TFIIIA zinc fingers which would make specific base contacts if Pabo's XYZ scheme is applied[82].

Figure 42

Proposed residues of the TFIIIA and their respective base pair contacts.X, Y, and Z correspond to the three amino acids in the helix which bind specifically to the DNA base pair. (**X** is +2 residues from the conserved aromatic residue in the second ß-strand, **Y** is -1 residue from the conserved leucine in the α helix, and **Z** is -1 residue from the first invariant histidine.) If the general code between histidines, arginines, and lysines to guanine is conserved among all zinc fingers, then the DNA bases they should contact are listed underneath each amino acid where **G** is guanine and **n** represaents any nucleotide.

Finger	1	2	3
Positions	XYZ	XYZ	XYZ
Residues	KKA	SHR	TNK
Bind	GGn	nGG	nnG

If recognition occurs on the noncoding strand in the 5S gene, a possible binding site in the ICR for the first three zinc fingers would start at position +87 at the 3' region and end at position +79 for the first three zinc fingers. The corresponding DNA has the sequence 3'-GGT AGG TTC-5'. Fingers 1 and 2 should bind to base pairs (87-85) and (84-82) respectively, see Figure 43.

Figure 43

Three N terminus zinc fingers and their DNA contacts

Finger	1	2	3
Positions	XYZ	XYZ	XYZ
Residues	KKA	SHR	TNK
Bind	GGT	AGG	TTC

The lysine at position 92 of the third finger may bind with the coding base, guanine at ICR base pair 79. The MD simulations of the third "even" zinc finger (Chapter 2) and correlation analysis (Chapter 3) indicated that a residue at this position had significant fluctuations due to the flexibility of the zinc cluster. Thus the long sidechain of lysine in a flexible helix may bind to the coding base rather than to the usual non-coding one. Clemens et al[62] reported that this coding guanine base was protected from methylation interference. The positioning of the first three amino-terminal zinc fingers based on an inspection of the primary sequence and the guidelines from the co-crystal correlate well with experimental data derived from a variety of sources, see Table 15.

Table 15. Summary of Experimental Data of the TFIIIA-5S Complex

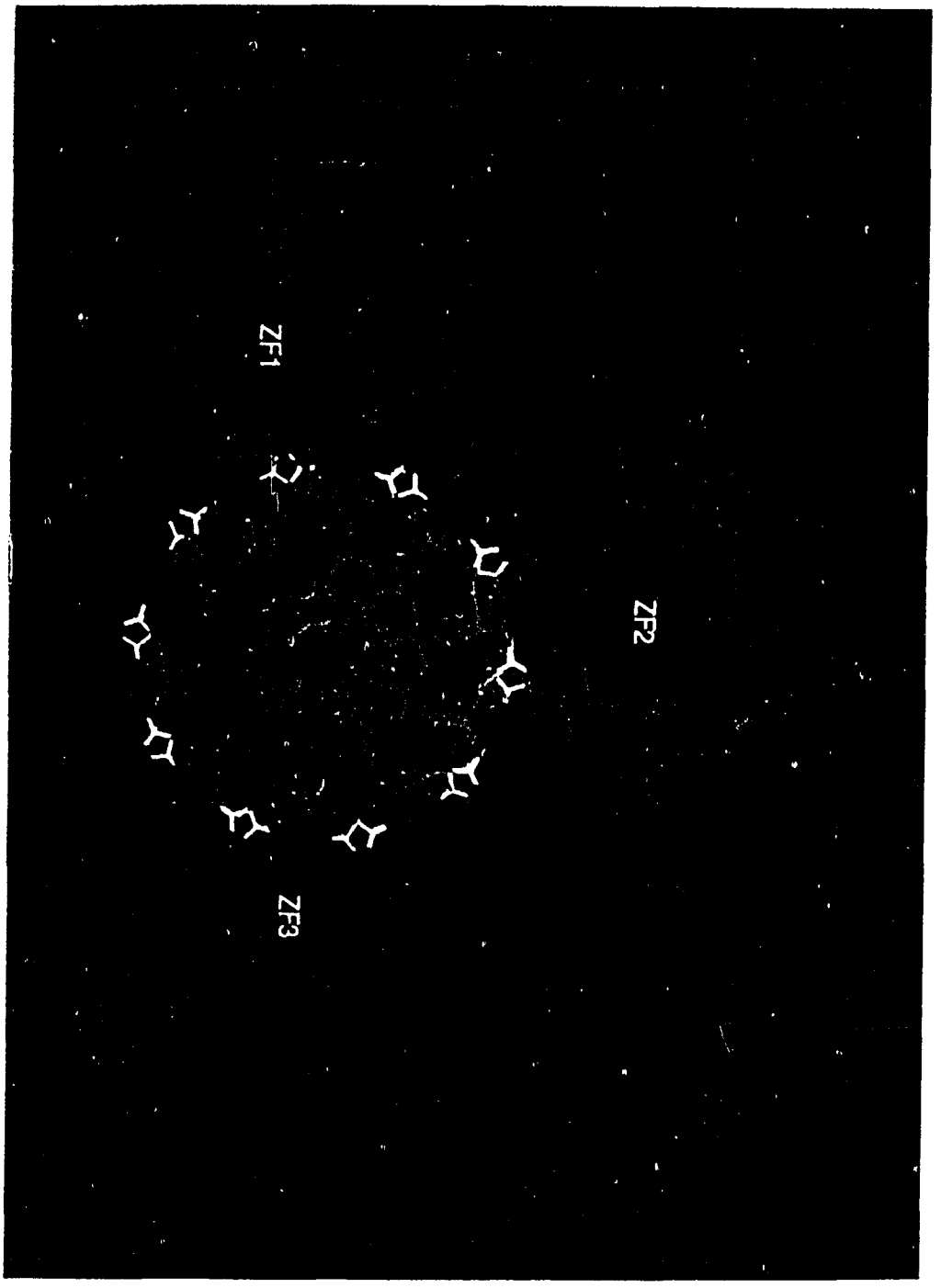
Investigators	Experimental Design	Conclusions
Fairall et al[60]	Analyses of earlier DNAse I footprinting study including the steric interactions of DNAse I with DNA	• TFIIIA binds to the 5S gene in 3 principal regions of the noncoding strand: bp's +78>+88 (C Box, bp's +65>+71 (IE Box), and bp's +49>+58 (A Box) • Protein crosses the minor groove at two positions: bp's +73>+76 and bp's +60>+63
Hayes et al[61]	Missing nucleotide experiments	• C box can be further delineated into two regions +79>+86 and +89>+92 • +92>+78 is the contact region for the amino terminus of TFIIA. • Protein crosses the minor groove at position +78 • Linker region and possibly finger 4 crosses the minor groove • Protein crosses the minor groove at position +65 and forms continuous contact to position +51
Clemens et al[62]	DNAse I footprinting, methylation interference and primer extension with differential TFIIIA zinc finger contructs	• Zinc finger 1-3 (ZF1-3) binds to +95>+77 of the non-coding strand • Protein crosses the minor groove at positions +77>+74 and +64>+60 • ZF5 binds at position +71>69 • ZF3 binds at position +79 • ZF4 crosses the minor groove at positions +77>+73 • Bp's +60>+50 represent the A Box
Hayes et al[63]	Hydroxyl radical footprinting with differential TFIIIA ZF constructs from Clemens	• C Box represent positions +90>+79 • ZF4-6 binds parallel to the central axis of the DNA • ZF4 crosses the minor groove at position +79

5.1.4. Construction of the Model

The Zif268 co-crystal complex was used as a molecular template to find an initial position of the zinc fingers relative to the DNA. A 25 base-pair oligonucleotide of the ICR (+90 to +65) was constructed in a B-DNA conformation. An eleven-base-pair segment (+87 to +79) was superimposed upon the Zif268 DNA with corresponding regions of finger interactions aligned (e.g., position +87 to +85 was superimposed upon the three base pairs of Zif268 to which the first finger bound). Each individual TFIIIA zinc finger was then superimposed with its analogous zinc finger in the Zif268 complex. Thus an initial model of the first three zinc fingers was constructed which approximated the contacts observed in the co-crystal complex. Subsequently, each individual finger was slightly adjusted to optimize the base pair contacts as designated by the scheme described earlier (Figure 42). The first three TFIIIA zinc fingers retained the general features of the Zif268 fingers: all base pair contacts are made in the major groove at the XYZ positions, each finger is separated by approximately 30°, and the α helix is tipped into the major groove, see Figure 44. One arginine in finger two forms two hydrogen bonds to the guanine base as observed in the Zif268 complex.

Figure 44

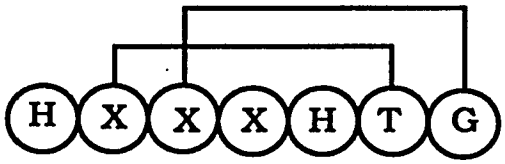
Model of the three N terminal TFIIIA zinc fingers with DNA. The three N terminal fingers are labeled ZF1) finger 1, ZF2) finger 2, and ZF3) finger 3. The linker is shown in red.



Because the N-to-C helical axis of the zinc finger points away from the central axis of the DNA helix, a tight β II' turn[65] formed by the residues that span individual zinc fingers (the linkers) is required to reintroduce the next zinc finger back towards the major groove of the DNA. Two mainchain hydrogen bonds involving the residues between the two conserved histidine ligands and the linker help stabilize the sharp turn.

Figure 45

Two mainchain hydrogen bonds between the linker and the HX3H spacing are found in between fingers 1 and 2 and between 2 and 3 of Zif268[83]. The TGEK linker is a conserved sequence among zinc fingers. The amino acids are represented by their single letter code: H) histidine, T) threonine, G) glycine, X) any amino acid.

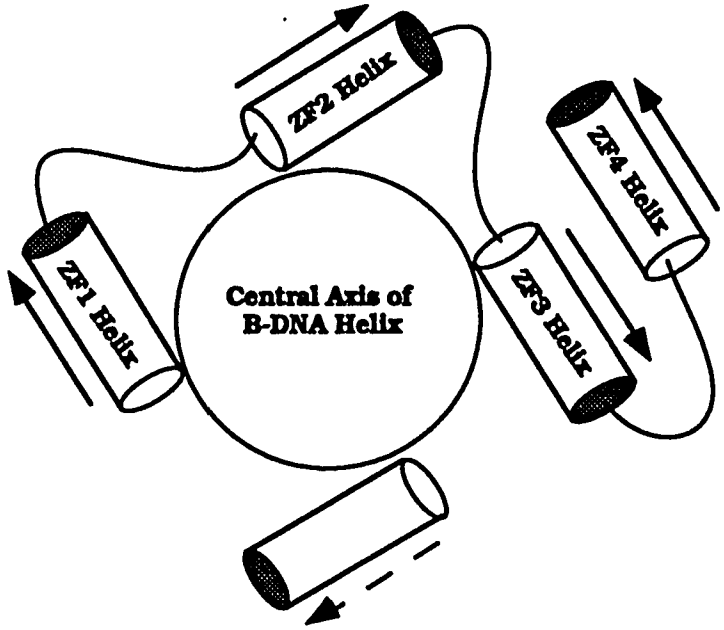


The structural consistency of the linker region (RMS deviation between the two linkers of Zif268 is only 0.15 Å) precisely re-aligns the next finger to follow the major groove by turning the following finger back toward the DNA.

However, detailed protection studies of the TFIIIA-5S DNA complex indicated a short region of protection at the minor groove near position +78 suggesting that the fourth zinc finger (ZF4) and/or ZF3-ZF4 linker cross the minor groove in this vicinity (Table 15). In order to span the minor groove, the third-to-fourth finger linker (NIKIC) which connect the C terminus of ZF3 to the N terminus of ZF4 must reverse the directionality commonly found in "odd" zinc fingers. The minor groove of bp's +78 to +76 lie below the major groove where ZF3 binds. Instead of being reintroduced back into the major groove, ZF4 must bind in an opposite orientation, (see Figure 46).

Figure 46

The graphic depiction of the orientation of the zinc finger recognition helices (ZF1-4) around the central axis of the DNA. Solid arrows indicate the directionality of the helices which are connected by curved line representing the linkers. Dashed arrow represent the orientation of finger 4 if the conformation of the conserved TGEK linker (found between finger 1 and 2, and 2 and 3) were applied to linker region between finger 3 and 4.



Thus the linker region of ZF3 shifts away from the "wrapping around" model and places the next zinc finger (ZF4) in a different orientation, in the minor groove. This becomes possible by virtue of the special property uncovered in the molecular dynmics simulation: the flexibility of the "even" zinc fingers. As a result of the wider exploration of conformational space by the linker structure, the hydrogen bonds between the HX4H spacing region to the linker residues are no longer observed. In contrast to the "odd" zinc finger, no hydrogen bonds between the linker to the HX4H are found in the 5 MD (10 psec)-averaged structures of the "even" TFIIIA(iii) zinc finger. Therefore the lack of strong intermolecular interactions to stabilize the linker now allows it to assume conformations unavailable for "odd" zinc fingers.

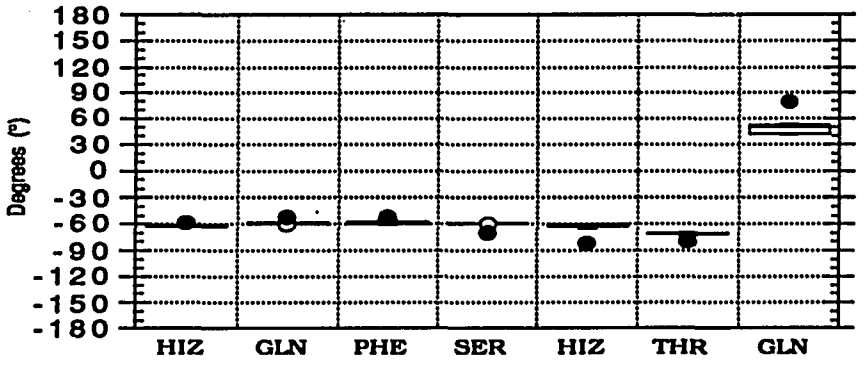
Structural consistency observed in the "odd" TFIIIA zinc finger was extendable to the two linker residues as well as the the mainchain hydrogen bonding between its linker and the HX3H spacing, because the dynamics simulations indicated a conserved 3-D structural pattern. In contrast, the lack of "finger-to-linker" hydrogen bonding in "even" zinc finger linkers provided for a greater variability in the structures of these regions thus allowing them the possibility of positioning the following zinc finger (ZF4) into the minor groove.

The conformational freedom of the key linker region in the "even" zinc finger is evident from quantitative analysis of the values of ϕ and φ torsional angles from the five 10 psec MD averaged structures of TFIIIA(iii) and TFIIIA(iv), shown in Figures 47a, b and 48a, b.

Figure 47

Φ dihedral angles of the HX3,4H spacing and two linkers from the 5 MD time averaged structures (10 psec): a) TFIIIA(iv) "odd" zinc finger and b) TFIIIA(iii) "even" zinc finger. Boxes represent the range of dihedral angles. Closed circles are the average values from the zif268 co-crystal complex zinc fingers. Open circles represent large deviations from the mean value.

a)



b)

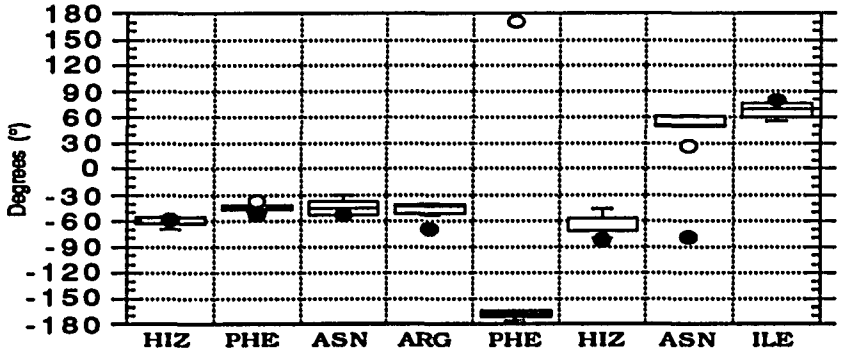
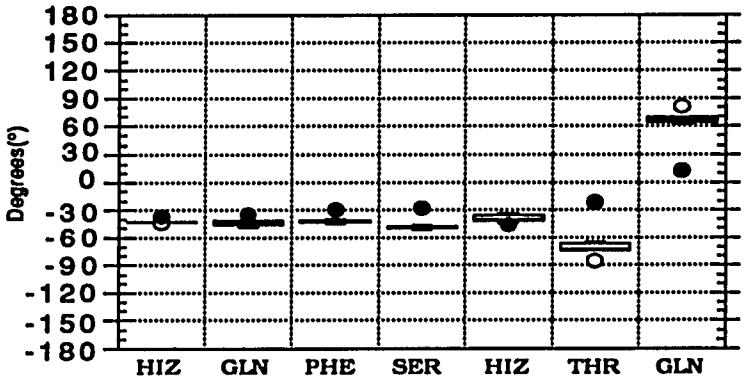


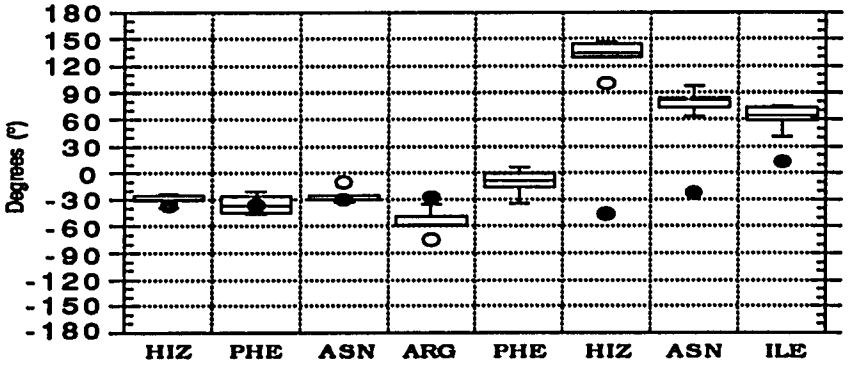
Figure 48

Ψ dihedral angles of the HX3,4H spacing and two linkers from the 5 MD time averaged structures (10 psec): a) TFIIIA(iv) "odd" zinc finger and b) TFIIIA(iii) "even" zinc finger. Boxes represent the range of dihedral angles. Closed circles are the average values from the zif268 co-crystal complex zinc fingers. Open circles represent large deviations from the mean value.

a)



b)



Clearly, the "even" TFIIIA(iii) had a greater range of torsional angles and a larger deviation from the Zif268 values suggesting it had a greater possibility of altering the directionality of the helix.

Alignment for the rest of the fingers

Finger 4 is modeled spanning the minor groove at base pairs 79 to 77. If the convention depicted in Figure 40 is applied to finger with the X and Y positions contacting instead to the DNA backbone, then the sidechain of lysine at position 118 would form a favorable electrostatic with the phosphate of base pair position 79 and the sidechain of glutamine 121 would interact with the phosphate group on the coding strand at position 78. Thus this model proposes that finger 4 makes only contacts to the phosphate backbone.

As a consequence of finger 4 crossing the minor groove and forming an orientation approximately parallel to the central axis of the DNA, finger 5 is positioned back into the major groove at base pair 73 to 71, TGG. Finger 5 is positioned in the major groove in a Zif268-like manner, with the its XYZ amino acids contacting base pairs shown in Figure 49.

Figure 49

The proposed amino acids and their respective base contacts. The leucline sidechain was modeled to form favorable interaction with the methyl of thymine. The arginines were modeled as typical arginine to guanine contacts observed in the Zif268 crystal structure.

Finger	5
Positions	XYZ
Residues	LRR
Bind	TGG

Because only one residue separates fingers 5 and 6, finger 6 is not able to follow along in the major groove in a zif268-like manner (sec. 5.1.4). Instead, the directionality of the helix of the fifth zinc finger positions helix 6 out of the major groove. Furthermore, it is questionable whether the structure of the sixth zinc finger of TFIIIA adopts a canonical zinc-finger folding because only one residue separates the conserved cysteine and the conserved aromatic residue (3 residue spacing in typical CC/HH zinc fingers). The difference in spacing may result in a disrupted hydrophobic core since the phenyalanine would not be in position to form favorable hydrophobic interactions with the conserved leucine[39]. Thus the overall folding of finger 6 may not be similar to the canonical zinc finger and the possible unusual conformation of finger 6 may be responsible for interaction with RNA. Theunissen et al[84] recently demonstrated through internal deletions of single TFIIIA zinc fingers that zinc finger 6 is the only zinc finger in TFIIIA responsible for sequence-specific recognition of the 5S RNA. Therefore, structure and positioning of the sixth zinc finger is uncertain. This relatively large protection region of 5-7 base pairs (IE Box) where only one zinc finger (finger 5) is proposed to bind in the major groove may require conformational changes in the DNA[56] or a possible non-TFIIIA-like conformation of finger 6 may account for the experimental data. Figure 48 shows the first six fingers of TFIIIA with DNA.

Figure 50

The first six zinc fingers of TFIIIA with DNA. ZF1-6 with DNA: Green) finger 1, Blue) finger 2, Red) finger 3, Yellow) finger 4, Orange) finger 5 and White) finger 6. Note that fingers 4 and 6 span the minor groove.

Fingers 7 and 8 are modeled in the Zif268-like orientation beginning with base pairs positions 62 to 57, see Figure 51. Zinc finger 9 is positioned out of the major groove in a similar orientation to finger 1-3. The last group of zinc fingers is shown in Figure 52. The alignment of all nine zinc fingers with respect to the DNA is shown in Figure 53.

Figure 51

The proposed amino acids and their respective base contacts.

Finger	7	8
Positions	XYZ	XYZ
Residues	HYD	TNS
Bind	GGG	ACG

Figure 52

TFIIIA fingers 7, 8, and 9 with DNA: ZF7) seventh zinc finger, ZF8) eight zinc finger, and ZF9) ninth zinc finger.

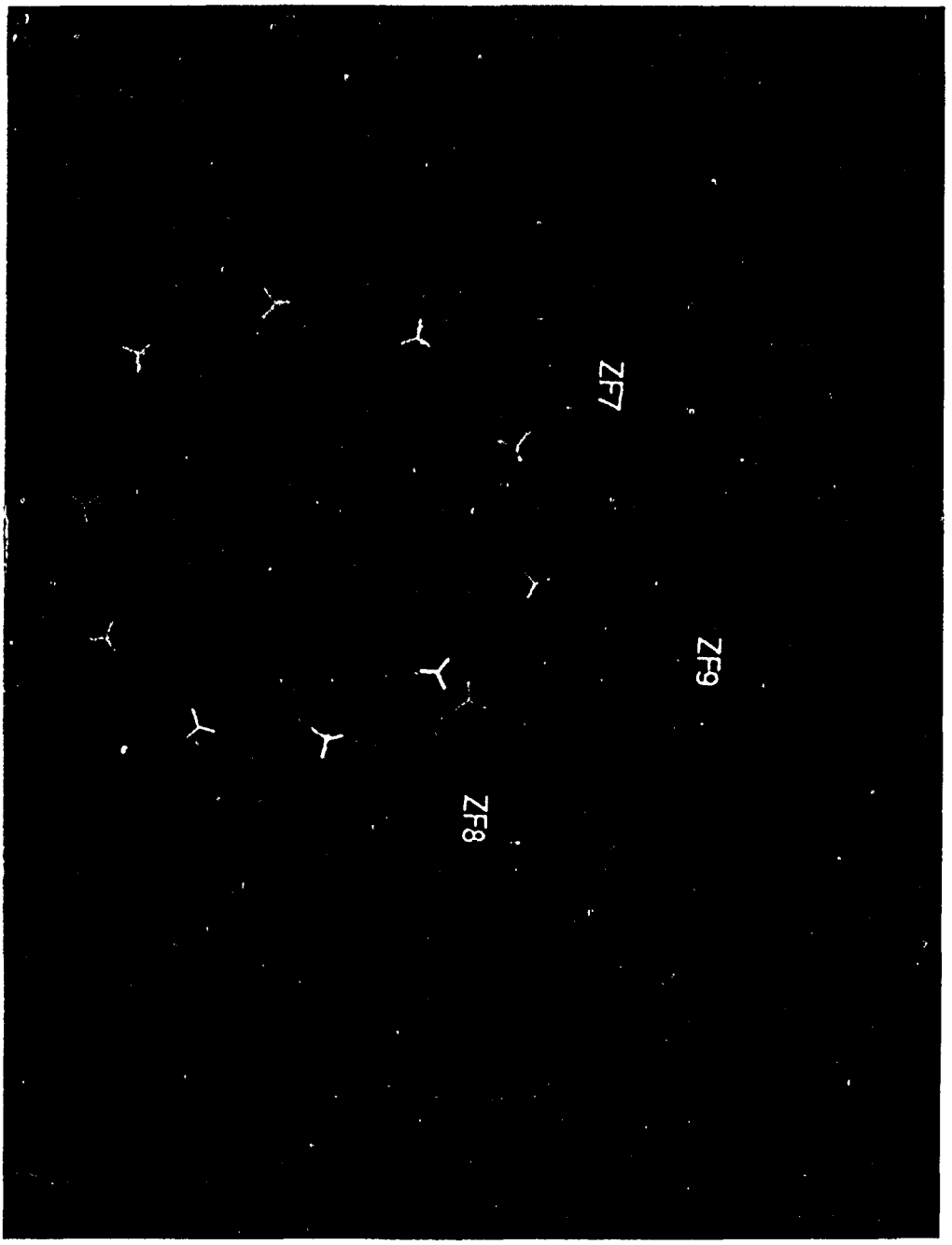
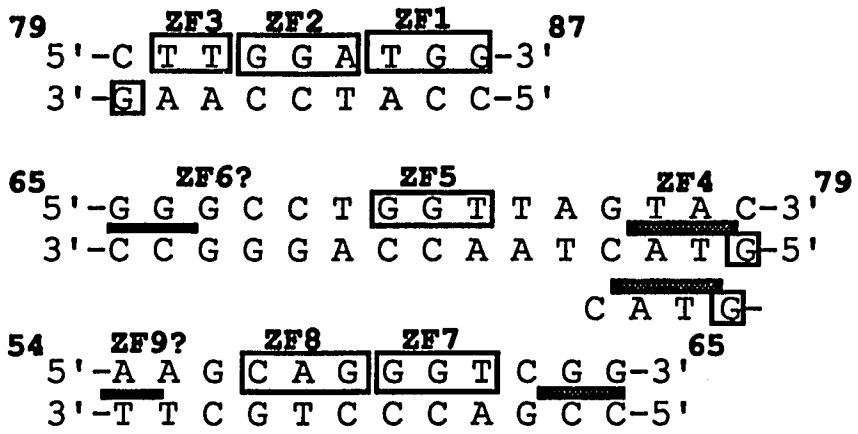


Figure 53

Model of the TFIIIA-5S complex. Base pairs 87 to 54 are divided into three promoter regions (A Box, IE Box and C Box). The base pair contacts by the TFIIIA zinc finger (ZF) are boxed. Grey boxes between the complementary base pairs represent proposed phosphate contacts in the minor groove.



6. CONCLUSION

The simulation of an "even" zinc finger (**HX4H**) such as the third zinc finger of TFIIIA have shown that "even" zinc fingers may form a subclass of the classical zinc fingers because of their increased structural flexibility compared to "odd" zinc fingers (**HX3H**). The molecular dynamic simulations described the mechanistic details of the dynamic properties of "even" zinc fingers as a consequence of the four residue spacing between the invariant histidines. The *disruption* of a seemingly complementary arrangement of a helical structure for zinc metal binding and for DNA recognition in "even" zinc fingers results in a greater fluctuation of the zinc metal cluster, a more variable hydrogen bonding network in the helix and a more flexible overall zinc finger.

Theoretical studies have identified protein flexibility, which was not readily apparent from an inspection of the primary sequence of TFIIIA or an analysis of the co-crystal complex, as a possible component for sequence-specific recognition. Furthermore the molecular dynamics experiments led to a testable prediction - conversion of the wildtype "even" zinc finger to an "odd" finger would alter binding affinity without a disruption to the sequence-specific contacts of the "even" finger. The gel retardation studies have confirmed that a single amino acid deletion mutant in the third zinc finger of TFIIIA can reduce the overall affinity of the entire protein. Thus theoretical methods in conjunction with appropriate

experiments can provide a powerful means in which to probe biomolecular interactions.

Structural modeling of the TFIIIA-5S complex revealed the necessity for a flexible zinc finger. In order for TFIIIA to achieve different modes of interaction with the DNA with essentially equivalent zinc-finger structures, a dynamic property of flexibility was required. The flexibility of the "even" third zinc finger of TFIIIA changes the directionality of the subsequent zinc finger. Instead of binding in the major groove, flexibility "allowed" the fourth zinc finger to achieve a novel mode of interaction first proposed by the detailed footprinting experiments.

The binding of long stretches of DNA by multi-zinc finger proteins presents challenges for sequence-specific recognition because of the high entropic costs in prepositioning the numerous DNA contact sites of TFIIIA before complexation. Therefore DNA-binding by TFIIIA probably occurs in a series of steps with each step dependent upon the previous one. For example, the three N-terminal zinc fingers would be the first group of fingers to bind to their cognate DNA site. The high entropic cost for this initial step would require a large enthalpic component for favorable binding free energy. Clemens et al [62]have proposed from footprinting competition studies that the first three zinc fingers of TFIIIA contribute to the majority of the free energy of binding. Subsequent groups of zinc fingers would have a lower entropic component due to the binding by the previous group of zinc fingers.

Because the first three fingers interact with almost a complete turn of B-conformation DNA (10.5 base pairs), the next group of fingers must bind in a novel fashion. If fingers 4, 5, and 6 simply followed the major groove in a Zif268-like manner, then TFIIIA would have "wrapped" around the DNA almost twice. This would require that TFIIIA be in a nearly unfolded state or in a dramatically different conformation before complexation; and upon binding, it would fold around the DNA. These requirements would make a complete "wrapping around" model energetically unfavorable and detailed footprinting studies support this notion. However, the possibility of a variety of DNA-binding modes within the TFIIIA DNA domain is restricted because of the structural stability and consistency of the zinc-finger motif. Thus the role of flexibility in "even" zinc fingers of TFIIIA allows for favorable energetic interaction of the protein with DNA from specific recognition.

REFERENCES

[1] Carl O. Pabo, and Robert T. Sauer, "Transcription factors structural families and principles of DNA recognition," Richardson, C. C. (Ed.). Annual Review Of Biochemistry 61 (1992): 1054-1088.

[2] Daniela Rhodes, and Aaron Klug, "Zinc fingers," Scientific American February (1993): 56-65.

[3] Jeremy M. Berg, "Zinc-finger proteins," Current Opinion in Structural Biology 3.1 (1993): 11-16.

[4] Cynthia Wolberger, "Transcription factor structure and DNA-binding," Current Opinion in Structural Biology 3.1 (1993): 3-10.

[5] Nadrian C. Seeman, John M. Rosenberg, and Alexander Rich, "Sequence specific recognition of double helical nucleic-acids by proteins," Proceedings of the National Academy of Sciences of the United States of America. 73.3 (1976): 804-808.

[6] E. N. Baker, and R. E. Hubbard, "Hydrogen bonding in globular proteins," Progress in Biophysical and Molecular Biology 44.2 (1984): 98-179.

[7] Z. Otwinowski, et al., "Crystal structure of *trp* repressor/operator complex at atomic resolution," Nature 335.22 (1988): 321-329.

[8] R. P. Wharton, and Mark Ptashne, "A new-specificity mutant of 434 repressor that define an amino acid-base pair contact," Nature 326.6116 (1987): 888-891.

[9] Robert Ivarie, "Thymine methyls and DNA-protein interaction," Nuc. Acids Res. 15.23 (1987): 9975-9983.

[10] G. B. Koudelka, Steven C. Harrison, and Mark Ptashne, "Effect of non-contacted bases on the affinity of 434 operator for 434 repressor and cro," Nature 326.6116 (1987): 886-888.

[11] G. B. Koudelka, et al., "Dna twisting and the affinity of bacteriophage 434 operator for bacteriophage 434 repressor," Proceedings Of The National Academy Of Sciences Of The United States Of America 85.13 (1988): 4633-4637.

[12] Marc R. Gartenberg, and Donald M. Crothers, "DNA sequence determinants of CAP-induced bending and protein affinity," Nature 333.30 (1988): 824-829.

[13] Peter H. von Hippel, and Otto G. Berg, "On the specificity of DNA-protein interactions," Proceedings of the National Academy of Sciences of the United States of America 83 (1986): 1608-1612.

[14] D. R. Engelke, et al., "Specific interaction of a purified transcription factor with and internal control region of 5S RNA genes," Cell 19.3 (1980): 717-728.

[15] Klug, "'Zinc Fingers': a novel protein motif for nucleic acid recognition," TIBS .Dec. 1987 (1987): 464-469.

[16] Robert Kaptein, "Zinc fingers," Current Opinion in Structural Biology 1 (1991): 63-70.

[17] Bert L. Vallee, Joseph E. Coleman, and David S. Auld, "Zinc fingers, zinc clusters, and zinc twists in DNA-binding protein domains," Proceedings from the National Academy of Science 88 (1991): 999-1003.

[18] Jay S. Hanas, et al., "Xenopus transcription factor A requires zinc for binding to the 5S RNA gene," Journal of Biological Chemistry 258.23 (1983): 14120-14125.

[19] S. Sakonju, David Bogenhagen, and Donald D. Brown, "A control region in the center of the 5S RNA gene directs specific initiation of transcription III. The 3' border of the region," Cell 19 (1980): 13-25.

[20] David F. Bogenhagen, S. Sakonju, and Donald Brown, "A control region in the center of the 5S RNA gene directs specific initiation of transcription III. The 3' border of the region," Cell 19 (1980): 27-35.

[21] D. D. Brown, "The role of stable complexes that repress and activate Genes," Cell 37 (1984): 359-365.

[22] Ann M. Ginsberg, Balas O. King, and Robert G. Roeder, "Xenopus 5s gene transcription factor TFIIIA characterization of a complementary dna clone and measurement of RNA levels throughout development," Cell 39 (1984): 479-489.

[23] J. Miller, A.D. McLaughlin, and Aaron Klug, "Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes," EMBO (European Molecular Biology Organization) Journal 4.6 (1985): 1609-1614.

[24] Alan D. Frankel, Jeremy M. Berg, and Carl O. Pabo, "Metal-dependent folding of a single zinc finger from transcription factor IIIA," Proceedings of The National Academy of Sciences of The United States of America. 84.14 (1987): 4841-4845.

[25] Grace Parraga, et al., "Zinc-Dependent Structure of a Single-finger Domain of Yeast ADR1," Science 241.16 (1988): 1489-1492.

[26] G. P. Diakun, Louise Fairall, and Aaron Klug, "EXAFS study of the zinc-binding sites in the protein transcription factor IIIA," Nature 324.18 (1986): 698-703.

[27] Jeremy M Berg, "On the metal ion specificity of the "Zinc Finger" Protein," Journal of the American Chemical Society 111.10 (1989): 3759-3761.

[28] Norbert Redemann, Ulrike Gaul, and Herbert Jackle, "Disruption of a putative Cys-Zinc interaction eliminates the biological activity of the Kruppel finger protein," Nature 332.3 (1988): 90-92.

[29] Grace Parraga, et al., "Spectroscopic studies of wild-type and mutant zinc finger peptides determinants of domain folding and structure," Proceedings Of The National Academy Of Sciences Of The United States Of America 87.1 (1990): 137-141.

[30] Jeremy Berg, "Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins," Proceedings Of The National Academy Of Sciences Of The United States Of America. 85.1 (1988): 99-102.

[31] Toby J. Gibson, et al., "A model for the tertiary structure of the 28 residue DNA-binding motif ('zinc finger') common to many eukaryotic transcriptional regulatory proteins," Protein Engineering 2.3 (1988): 209-218.

[32] Min S. Lee, et al., "Three- dimensional solution structure of a single zinc finger DNA-Binding domain," Science 245 (1989): 635 - 637.

[33] Mark D. Carr, et al., "NMR and molecular dynamics studies of the mKr2 'zinc finger'," European Journal of Biochemistry 188 (1990): 455-461.

[34] Rachel E. Klevit, Jon R. Herriott, and Suzanna J. Horvath, "Solution structure of a zinc finger domain of yeast ADR1," Proteins Structure Function And Genetics 7.3 (1990): 215-226.

[35] David Neuhaus, et al., "Sequence-specific proton nmr assignments and secondary structure identification for 1 zinc and 2 zinc finger constructs from SWI5 a hydrophobic core involving four invariant residues," Febs (Federation Of European Biochemical Societies) Letters 262.2 (1990): 179-184.

[36] Michel Kochoyan, et al., "Alternating zinc fingers in the human male associated protein ZFY: 2D NMR Structure of an even finger and implications for "jumping-linker" DNA recognition," Biochemistry 30 (1991): 3371-3386.

[37] James G. Omichinski, et al., "High-resolution solution structure of the double CYS-2 HIS-2 zinc finger from the human enhancer binding protein mbp-1," Biochemistry 31.16 (1992): 3907-3917.

[38] Scott Michael, et al., "Metal binding and folding properties of a minimalist CYS-2 HIS-2 zinc finger peptide," Proceedings Of The National Academy Of Sciences Of The United States Of America 89.11 (1992): 4796-4800.

[39] Russel J. Mortishire-Smith, et al., "Structural determinants of CYS-2-HIS-2 zinc fingers," Febs (Federation Of European Biochemical Societies) Letters 296.1 (1992): 11-15.

[40] Robert X. Xu, Suzanna Horvath, and Rachel E. Klevit, "ADR1 a zinc finger peptide exists in two folded conformations," Biochemistry 30.14 (1991): 3365-3371.

[41] Michael Weiss, et al., "Alternating zinc-finger motifs in the human male-associated protein zfy," Biochemistry 29.24 (1990): 5660-5664.

[42] Jeannette Nardelli, et al., "Base sequence discrimination by zinc-finger DNA-binding domains," Nature 349.6305 (1991): 175-178.

[43] Hans-Jurgen Thiesen, and Christian Bach, "Determination of DNA-binding specificities of mutated zinc finger domains," Febs (Federation Of European Biochemical Societies) Letters 283.1 (1991): 23-26.

[44] Hans-Jurgen Thiesen, and Bjorn Schroder, "Amino acid substitutions in the sp1 zinc finger domain alter the DNA-binding

affinity to cognate sp1 target site," Biochemical And Biophysical Research Communications 175.1 (1991): 333-338.

[45] J. R. Desjarlais, and Jeremy M. Berg, "Toward rules relating zinc finger protein sequences and DNA-binding site preferences," Proc Natl Acad Sci U S A 89.16 (1992): 7345-7349.

[46] J. R. Desjarlais, and Jeremy Berg, "Redesigning the DNA-binding specificity of a zinc finger protein a data base-guided approach," Proteins Struct Funct Genet 12.2 (1992): 101-104.

[47] Jeremy M. Berg, "Sp1 and the subfamily of zinc finger proteins with guanine-rich binding sites," Proceedings Of The National Academy Of Sciences Of The United States Of America 89 (1992): 11109-11110.

[48] Daniela Rhodes, "Structural analysis of a triple complex between the histone octamer a Xenopus gene for 5s RNA and transcription factor IIIA," Embo (European Molecular Biology Organization) Journal 4.13A (1985): 3473-3482.

[49] Daniela Rhodes, and Aaron Klug, "An underlying repeat in some transcriptional control sequences corresponding to half a double helical turn of DNA," Cell 46 (1986): 123-132.

[50] Louisa Fairall, Daniela Rhodes, and Aaron Klug, "Mapping of the sites of protection on a 5S RNA Gene by the Xenopus transcription factor IIIA," Journal of Molecular Biology 192 (1986): 577-591.

[51] Kent E. Vrana, et al., "Mapping functional regions of transcription factor tfiiia," Molecular And Cellular Biology 8.4 (1988): 1684-1696.

[52] Jay S. Hanas, et al., "Internal deletion mutants of xenopus transcription factor IIIA," Nucleic Acids Research 17.23 (1989): 9861-9870.

[53] Jorgen Holst Christensen, et al., "Sequence-specific binding of the N-terminal three-finger fragment of *Xenopus* transcription factor IIIA to the internal control region of a 5S RNA gene," Federation of European Biochemical Societies Letters 281.1 (1991): 181-184.

[54] Xiubei Liao, et al., "Specific interaction of the first three zinc fingers of TFIIIA with the internal control region of the Xenopus 5S RNA gene," Journal of Molecular Biology 223 (1992): 857-871.

[55] Mair Churchill, Thomas D. Tullius, and Aaron Klug, "Mode of interaction of the zinc finger protein TFIIIA with a 5s RNA gene of xenopus," Proceedings Of The National Academy Of Sciences Of The United States Of America. 87.14 (1990): 5528-5532.

[56] Jeremy Berg, "Zinc finger domains hypotheses and current knowledge," Engelman, D. M. (Ed.). Annual Review of Biophysics and Biophysical Chemistry 19 (1990).

[57] G. P. Schroth, et al., "Transcription factor IIIA induced bending of the xenopus somatic 5s gene promoter," Nature 340.6233 (1989): 487-488.

[58] Min S. Lee, Joel M. Gottesfeld, and Peter E. Wright, "Zinc is required for folding and binding of a single zinc finger to DNA," Febs (Federation Of European Biochemical Societies) Letters 279.2 (1991): 289-94.

[59] Brian Matthews, "No code for recognition," Nature 335.22 (1988): 294-295.

[60] Louisa Fairall, and Daniela Rhodes, "A new approach to the analysis of Dnase I footprinting data and its application to the tfiiia-5s dna complex," Nucleic Acids Research 20.18 (1992): 4727-4731.

[61] Jeffrey J. Hayes, and Thomas D. Tullius, "Structure of the TFIIIA-5S DNA complex," Journal Of Molecular Biology 227.2 (1992): 407-417.

62 Karen R. Clemens, et al., "Definition of the binding sites of individual zinc fingers in the transcription factor IIIA-5S RNA gene complex," Proceedings Of The National Academy Of Sciences Of The United States Of America 89 (1992): 10822-10826.

63 Jeffrey J. Hayes, and Karen R. Clemens, "Location of contacts between individual zinc fingers of *Xenopus laevis* transcription factor IIIA and the internal control region of the 5S RNA gene," Biochemistry 31 (1992): 11600-11605.

64 Thomas Pieler, Jorg Hamm, and R. G. Roeder, "The 5S Gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing," Cell 48 (1987): 91-100.

65 Thomas E. Creighton, Proteins Structure and Molecular Principles, (New York: W. H. Freeman and Company, 1984) 171-173.

66 David A. Pearlman, et al., "Assisted model building with energy refinement rel. 3a (AMBERrev3a)," 3a ed. (San Francisco: University of California San Francisco, 1991) .

67 Grant Jacobs, "Determination of the base recognition positions of zinc fingers from sequence analysis," EMBO (European Molecular Biology Organization) Journal 11.12 (1992): 4507-4517.

68 Wilfred F. van Gunsteren, "Molecular dynamics studies of proteins," Current Opinion in Structural Biology 3.2 (1993): 277-281.

69 J. Andrew McCammon, and Stephen C. Harvey, Dynamics of Proteins and Nucleic Acids, (Cambridge: Cambridge University Press, 1987) 234.

70 Paul K. Weiner, et al., "A new force field for molecular mechanical simulation of nucleic acids and proteins," Journal of the American Chemical Society 106 (1984): 765-783.

[71] Paul K. Weiner, Peter A. Kollman, and David A. Case, "An all atom force field for simulations of proteins and nucleic acids," Journal of Computational Chemistry 7 (1986): 230-252.

[72] Beth Allyn Krizek, et al., "A consensus zinc finger peptide design high-affinity metal binding a pH-dependent structure and a HIS to CYS sequence variant," Journal Of The American Chemical Society 113.12 (1991): 4518-4523.

[73] Arthur G. Palmer, and David A. Case, "Molecular dynamics analysis of NMR relaxation in a zinc-finger peptide," Journal of American Chemical Society 114 (1992): 9059-9067.

[74] Toshiko Ichiye, and Martin Karplus, "Collective motions in proteins: A covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations," Proteins: Structure, Function, Genetics 11 (1991): 205-217.

[75] J. Andrew McCammon, Bruce R. Gerlin, and Marin Kaplus, "Dynamics of folded proteins," Nature 267.5612 (1977): 585-590.

[76] Jeremy M. Berg, "Zinc fingers and other metal-binding domains," Journal of Biological Chemistry 265.12 (1990): 6513-6516.

[77] Steffan N. Ho, et al., "Site-directed mutagenesis by overlap extension using polymerase chain reaction," Gene 77.77 (1989): 51-59.

[78] Kiyoshi Nagai, et al., "Zinc-finger motifs expressed in *E. coli* and folded *in vitro* direct specific binding to DNA," Nature 232.17 (1988): 284-286.

[79] J. Sambrook, E. F. Fritsch, and T. Maniatis, Molecular Cloning, A Laboratory Manual, 2 ed. , ed. Chris Nolan (New York: Cold Spring Harbor Press, 1989) .

[80] H. C. Birnbohm, and J. Doly, "A rapid alkaline extraction procedure for screening recombinant plasmid DNA," Nucleic Acids Research 7 (1979): 1513-1523.

[81] Joel Gottesfeld, Blanco, and Tennant, "The 5S gene internal control region is B-form both in solution and in a complex with TFIIIA," Nat. 329 (1987): 460-462.

[82] Rachel E. Klevit, "Recognition of DNA by Cys2,His2 zinc fingers," 253 (1990): 1390-1393.

[83] W. Knochel, et al., "Evolutionary conserved modules associated with zinc fingers in *Xenopus laevis*," Proceedings Of The National Academy Of Sciences Of The United States Of America 86 (1989): 6097-6100.

[84] Oliver Theunissen, et al., "RNA and DNA-binding zinc fingers of TFIIIA," 71 (1993): 670-690.

BIBLIOGRAPHY

Baker, E. N. and R. E. Hubbard. "Hydrogen bonding in globular proteins." Progress in Biophysical and Molecular Biology 44.2 (1984): 98-179.

Berg, Jeremy M. "Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins." Proceedings Of The National Academy Of Sciences Of The United States Of America. 85.1 (1988): 99-102.

Berg, Jeremy M. "On the metal ion specificity of the "Zinc Finger" Protein." Journal of the American Chemical Society 111.10 (1989): 3759-3761.

Berg, Jeremy M. "Zinc fingers and other metal-binding domains." Journal of Biological Chemistry 265.12 (1990): 6513-6516.

Berg, Jeremy M. "Zinc finger domains hypotheses and current knowledge." Engelman, D. M. (Ed.). Annual Review of Biophysics and Biophysical Chemistry 19 (1990).

Berg, Jeremy M. "Sp1 and the subfamily of zinc finger proteins with guanine-rich binding sites." Proceedings Of The National Academy Of Sciences Of The United States Of America 89 (1992): 11109-11110.

Berg, Jeremy M. "Zinc-finger proteins." Current Opinion in Structural Biology 3.1 (1993): 11-16.

Birnbohm, H. C. and J. Doly. " A rapid alkaline extraction procedure for screening recombinant plasmid DNA." Nucleic Acids Research 7 (1979): 1513-1523.

Bogenhagen, David F., S. Sakonju, and Donald Brown. "A control region in the center of the 5S RNA gene directs specific initiation of transcription III. The 3' border of the region." Cell 19 (1980): 27-35.

Brown, Donald D. "The role of stable complexes that repress and activate Genes." Cell 37 (1984): 359-365.

Carr, Mark D. et al. "NMR and molecular dynamics studies of the mKr2 'zinc finger'." European Journal of Biochemistry 188 (1990): 455-461.

Christensen, Jorgen Holst et al. "Sequence-specific binding of the N-terminal three-finger fragment of *Xenopus* transcription factor IIIA to the internal control region of a 5S RNA gene." Federation of European Biochemical Societies Letters 281.1 (1991): 181-184.

Churchill, Mair, Thomas D. Tullius, and Aaron Klug. "Mode of interaction of the zinc finger protein TFIIIA with a 5S RNA gene of Xenopus." Proceedings Of The National Academy Of Sciences Of The United States Of America. 87.14 (1990): 5528-5532.

Clemens, Karen R. et al. "Definition of the binding sites of individual zinc fingers in the transcription factor IIIA-5S RNA gene complex." Proceedings Of The National Academy Of Sciences Of The United States Of America 89 (1992): 10822-10826.

Creighton, Thomas E. Proteins Structure and Molecular Principles . (New York: W. H. Freeman and Company, 1984) 171-173.

Desjarlais, J. R., and Jeremy Berg. "Redesigning the dna-binding specificity of a zinc finger protein a data base-guided approach." Proteins Struct Funct Genet 12.2 (1992): 101-104.

Desjarlais, J. R., and Jeremy M. Berg, "Toward rules relating zinc finger protein sequences and DNA binding site preferences." Proc Natl Acad Sci U S A 89.16 (1992): 7345-7349.

Diakun, G. P., Louise Fairall, and Aaron Klug. "EXAFS study of the zinc-binding sites in the protein transcription factor IIIA." Nature 324.18 (1986): 698-703.

Engelke, D. R., et al. "Specific interaction of a purified transcription factor with and internal control region of 5S RNA genes." Cell 19.3 (1980): 717-728.

Fairall, Louisa, Daniela Rhodes, and Aaron Klug. "Mapping of the Sites of Protection on a 5S RNA Gene by the Xenopus Transcription Factor IIIA." Journal of Molecular Biology 192 (1986): 577-591.

Fairall, Louisa, and Daniela Rhodes. "A new approach to the analysis of dnase i footprinting data and its application to the TFIIIA-5S DNA complex." Nucleic Acids Research 20.18 (1992): 4727-4731.

Frankel, Alan D. , Jeremy M. Berg, and Carl O. Pabo. "Metal-dependent folding of a single zinc finger from transcription factor IIIA," Proceedings of The National Academy of Sciences of The United States of America. 84.14 (1987): 4841-4845.

Gartenberg, Marc R., and Donald M. Crothers. "DNA sequence determinants of CAP-induced bending and protein affinity," Nature 333.30 (1988): 824-829.

Gibson, Toby J. et al. "A model for the tertiary structure of the 28 residue DNA-binding motif ('zinc finger') common to many eukaryotic transcriptional regulatory proteins." Protein Engineering 2.3 (1988): 209-218.

Ginsberg, Ann M., Balas O. King, and Robert G. Roeder, "Xenopus 5S gene transcription factor TFIIIA characterization of a complementary DNA clone and measurement of RNA levels throughout development." Cell 39 (1984): 479-489.

Gottesfeld, Joel, Blanco, and Tennant. "The 5S gene internal control region is B-form both in solution and in a complex with TFIIIA." Nat. 329 (1987): 460-462.

van Gunsteren, Wilfred F. "Molecular dynamics studies of proteins." Current Opinion in Structural Biology 3.2 (1993): 277-281.

Hanas, Jay S. et al. "Xenopus transcription factor requires zinc for binding to the 5S RNA gene." Journal of Biological Chemistry 258.23 (1983): 14120-14125.

Hanas, Jay S. et al. "Internal deletion mutants of Xenopus transcription factor IIIA," Nucleic Acids Research 17.23 (1989): 9861-9870.

Hayes, Jeffrey J., and Karen R. Clemens. "Location of contacts between individual zinc fingers of Xenopus laevis transcription factor IIIA and the internal control region of the 5S RNA gene." Biochemistry 31 (1992): 11600-11605.

Hayes, Jeffrey J. and Thomas D. Tullius. "Structure of the TFIIIA-5S DNA complex." Journal Of Molecular Biology 227.2 (1992): 407-417.

von Hippel, Peter H., and Otto G. Berg. "On the specificity of DNA-protein interactions." Proceedings of the National Academy of Sciences of the United States of America 83 (1986): 1608-1612.

Ho, Steffan N., et al. "Site-directed mutagenesis by overlap extension using polymerase chain reaction." Gene 77.77 (1989): 51-59.

Ichiye, Toshiko , and Martin Karplus, "Collective motions in proteins: A covariance analysis of atomic fluctuations in molecular dynamics

and normal mode simulations." Proteins: Structure, Function, Genetics 11 (1991): 205-217.

Ivarie, Robert. "Thymine methyls and DNA-protein interaction." Nucleic Acids Research. 15.23 (1987): 9975-9983.

Jacobs, Grant. "Determination of the base recognition positions of zinc fingers from sequence analysis." EMBO (European Molecular Biology Organization) Journal 11.12 (1992): 4507-4517.

Kaptein, Robert. "Zinc fingers," Current Opinion in Structural Biology 1 (1991): 63-70.

Klevit, Rachel E., Jon R. Herriott, and Suzanna J. Horvath, "Solution structure of a zinc finger domain of yeast ADR1." Proteins Structure Function And Genetics 7.3 (1990): 215-226.

Klevit, Rachel E. "Recognition of DNA by Cys2,His2 zinc fingers." Science 253 (1990): 1390-1393.

Klug, Aaron "'Zinc Fingers': a novel protein motif for nucleic acid recognition." TIBS .Dec. 1987 (1987): 464-469.

Knochel, W. et al., "Evolutionary conserved modules associated with zinc fingers in *Xenopus laevis*,." Proceedings Of The National Academy Of Sciences Of The United States Of America 86 (1989): 6097-6100.

Kochoyan, Michel et al. "Alternating zinc fingers in the human male associated protein ZFY: 2D NMR Structure of an even finger and implications for "jumping-linker" DNA recognition." Biochemistry 30 (1991): 3371-3386.

Koudelka, G. B., Steven C. Harrison, and Mark Ptashne. "Effect of non-contacted bases on the affinity of 434 operator for 434 repressor and Cro." Nature 326.6116 (1987): 886-888.

Koudelka, G. B., et al., "DNA twisting and the affinity of bacteriophage 434 operator for bacteriophage 434 repressor." Proceedings Of The National Academy Of Sciences Of The United States Of America 85.13 (1988): 4633-4637.

Krizek, Beth Allyn et al. "A consensus zinc finger peptide design high-affinity metal binding a ph-dependent structure and a His to Cys sequence variant." Journal Of The American Chemical Society 113.12 (1991): 4518-4523.

Lee, Min S. et al. "Three- dimensional Solution Structure of a Single Zinc Finger DNA-Binding Domain." Science 245 (1989): 635 - 637.

Lee, Min S., Joel M. Gottesfeld, and Peter E. Wright. "Zinc is required for folding and binding of a single zinc finger to DNA." FEBS (Federation Of European Biochemical Societies) Letters 279.2 (1991): 289-94.

Liao, Xiubei et al. "Specific interaction of the first three zinc fingers of TFIIIA with the internal control region of the Xenopus 5S RNA gene." Journal of Molecular Biology 223 (1992): 857-871.

McCammon, J. Andrew, Bruce R. Gerlin, and Martin Kaplus. "Dynamics of folded proteins." Nature 267.5612 (1977): 585-590.

McCammon, J. Andrew, and Stephen C. Harvey, Dynamics of proteins and nucleic acids. (Cambridge: Cambridge University Press, 1987) 234.

Matthews, Brian. "No code for recognition." Nature 335.22 (1988): 294-295.

Michael,Scott et al. "Metal binding and folding properties of a minimalist cys-2his-2 zinc finger peptide." Proceedings Of The National Academy Of Sciences Of The United States Of America 89.11 (1992): 4796-4800.

Miller, J. A.D. McLaughlin, and Aaron Klug. "Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes," EMBO (European Molecular Biology Organization) Journal 4.6 (1985): 1609-1614.

Mortishire-Smith, Russel J. et al. "Structural determinants of Cys-2-His-2 zinc fingers." FEBS (Federation Of European Biochemical Societies) Letters 296.1 (1992): 11-15.

Nagai, Kiyoshi et al. "Zinc-finger motifs expressed in *E. coli* and folded *in vitro* direct specific binding to DNA." Nature 232.17 (1988): 284-286.

Nardelli, Jeannette et al. "Base sequence discrimination by zinc-finger DNA-binding domains." Nature 349.6305 (1991): 175-178.

Neuhaus, David et al. "Sequence-specific proton nmr assignments and secondary structure identification for 1 zinc and 2 zinc finger constructs from SWI5 a hydrophobic core involving four invariant residues." FEBS (Federation Of European Biochemical Societies) Letters 262.2 (1990): 179-184.

Omichinski, James G. et al. "High-resolution solution structure of the double cys-2his-2 zinc finger from the human enhancer binding protein MBP-1." Biochemistry 31.16 (1992): 3907-3917.

Otwinowski, Z. et al. "Crystal structure of trp repressor/operator complex at atomic resolution." Nature 335.22 (1988): 321-329.

Pabo, Carl O. and Robert T. Sauer. "Transcription factors structural families and principles of DNA recognition." Richardson, C. C. (Ed.). Annual Review Of Biochemistry 61 (1992): 1054-1088.

Palmer, Arthur G. and David A. Case. "Molecular dynamics analysis of NMR relaxation in a zinc-finger peptide." Journal of American Chemical Society 114 (1992): 9059-9067.

Parraga, Grace et al., "Zinc-Dependent structure of a single-finger domain of yeast ADR1." Science 241.16 (1988): 1489-1492.

Parraga, Grace et al. "Spectroscopic studies of wild-type and mutant zinc finger peptides determinants of domain folding and structure." Proceedings Of The National Academy Of Sciences Of The United States Of America 87.1 (1990): 137-141.

Pearlman, David A. et al. "Assisted Model Building with Energy Refinement Rel. 3a (AMBERrev3a)," 3a ed. (San Francisco: University of California San Francisco, 1991) .

Pieler, Thomas, Jorg Hamm, and R. G. Roeder. "The 5S Gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing." Cell 48 (1987): 91-100.

Redemann, Norbert, Ulrike Gaul, and Herbert Jackle,. "Disruption of a putative Cys-Zinc interaction eliminates the biological activity of the Kruppel finger protein." Nature 332.3 (1988): 90-92.

Rhodes, Daniela. "Structural analysis of a triple complex between the histone octamer a xenopus gene for 5S RNA and transcription factor IIIA." Embo (European Molecular Biology Organization) Journal 4.13A (1985): 3473-3482.

Rhodes, Daniela and Aaron Klug. "An underlying repeat in some transcriptional control sequences corresponding to half a double helical turn of DNA." Cell 46 (1986): 123-132.

Rhodes, Daniela and Aaron Klug. "Zinc Fingers." Scientific American February (1993): 56-65.

Sakonju, S. , David Bogenhagen, and Donald D. Brown. "A control region in the center of the 5S RNA gene directs specific initiation of transcription III. The 3' border of the region." Cell 19 (1980): 13-25.

Sambrook, J., E. F. Fritsch, and T. Maniatis. Molecular Cloning, A Laboratory Manual, 2 ed.. ed. Chris Nolan (New York: Cold Spring Harbor Press, 1989) .

Seeman, Nadrian C., John M. Rosenberg, and Alexander Rich. "Sequence specific recognition of double helical nucleic-acids by proteins." Proceedings of the National Academy of Sciences of the United States of America. 73.3 (1976): 804-808.

Schroth, G. P. et al. "Transcription factor iiia induced bending of the xenopus somatic 5S gene promoter." Nature 340.6233 (1989): 487-488.

Theunissen, Oliver et al. "RNA and DNA-binding zinc fingers of TFIIIA," 71 (1993): 670-690.

Thiesen, Hans-Jurgen and Christian Bach. "Determination of DNA binding specificities of mutated zinc finger domains." FEBS (Federation Of European Biochemical Societies) Letters 283.1 (1991): 23-26.

Thiesen, Hans-Jurgen and Bjorn Schroder. "Amino acid substitutions in the SP1 zinc finger domain alter the DNA binding affinity to cognate SP1 target site." Biochemical And Biophysical Research Communications 175.1 (1991): 333-338.

Vallee, Bert L., Joseph E. Coleman, and David S. Auld, "Zinc Fingers, zinc clusters, and zinc twists in DNA-binding protein domains." Proceedings from the National Academy of Science 88 (1991): 999-1003.

Vrana, Kent E. et al."Mapping functional regions of transcription factor TFIIIA." Molecular And Cellular Biology 8.4 (1988): 1684-1696.

Weiner, Paul K. et al. "A new force field for molecular mechanical simulation of nucleic acids and proteins." Journal of the American Chemical Society 106 (1984): 765-783.

Weiner, Paul K. Peter A. Kollman, and David A. Case. "An all atom force field for simulations of proteins and nucleic acids." Journal of Computational Chemistry 7 (1986): 230-252.

Michael Weiss, et al. "Alternating zinc-finger motifs in the human male-associated protein ZFY." Biochemistry 29.24 (1990): 5660-5664.

Wharton, R. P., and Mark Ptashne. "A new-specificity mutant of 434 repressor that define an amino acid-base pair contact." Nature 326.6116 (1987): 888-891.

Wolberger, Cynthia. "Transcription factor structure and DNA-binding." Current Opinion in Structural Biology 3.1 (1993): 3-10.

Xu, Robert X., Suzanna Horvath, and Rachel E. Klevit. "ADR1a a zinc finger peptide exists in two folded conformations." Biochemistry 30.14 (1991): 3365-3371.