Structure of a new nucleic-acid-binding motif in eukaryotic transcriptional elongation factor TFIIIS

Weiss, Michael A.; Qian, Xiuqu; Jeon, Choon Ju; Yoon, Ho Sup; Agarwal, Kan

1993

Qian, X., Jeon, C., Yoon, H., Agarwal, K., & Weiss, M.A. (1993). Structure of a new nucleic-acid-binding motif in eukaryotic transcriptional elongation factor TFIIIS. *Nature*, 365, 277-279.

<https://hdl.handle.net/10356/94482>

<https://doi.org/10.1038/365277a0>

© 1993 Nature Publishing Group. This is the author created version of a work that has been peer reviewed and accepted for publication by Nature, Nature Publishing Group. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: <http://dx.doi.org/10.1038/365277a0>

Downloaded on 26 Apr 2023 14:45:49 SGT

Structure of a new nucleic-acid-binding motif in eukaryotic transcriptional elongation factor TFIIS

Xinqu Qian^{*†}, ChoonJu Jeon[‡], HoSup Yoon[‡], Kan Agarwal^{‡ §} & Michael

A. Weiss^{*†}

^{*} Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA

[†]Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

Departments of [‡] Biochemistry and Molecular Biology and

[§] Chemistry, University of Chicago, Chicago, Illinois 60637, USA

TRANSCRIPTIONAL elongation involves dynamic interactions among RNA polymerase and single-stranded and double-stranded nucleic acids in the ternary complex¹⁻⁴. In prokaryotes its regulation provides an important mechanism of genetic control¹. Analogous eukaryotic mechanisms are not well understood⁵, but may control expression of proto-oncogenes^{6,7} and viruses, including the human immunodeficiency virus HIV-1 (ref. 8). The highly conserved eukaryotic transcriptional elongation factor TFIIS⁹ enables RNA polymerase II (RNAPII) to read through pause or termination sites, nucleosomes and sequence-specific DNA-binding proteins¹⁰⁻¹⁴. Two distinct domains of human TFIIS, which bind RNAPII and nucleic acids, regulate read-through¹⁰ and possibly nascent transcript cleavage¹¹⁻¹⁵. Here we describe the three-dimensional NMR¹⁶ structure of a Cys₄ nucleic-acid-binding domain from human TFIIS^{9,10}. Unlike previously characterized zinc modules¹⁷⁻²¹, which contain an α -helix, this structure consists of a three-stranded β -sheet. Analogous Cys₄ structural motifs may occur in other proteins involved in DNA or RNA transactions²²⁻²⁴, including RNAPII itself²⁵. This new structure, designated the Zn ribbon, extends the repertoire of Zn-mediated peptide architectures²⁶ and highlights the growing recognition of the β -sheet as a motif of nucleic-acid recognition^{27,28}.

The nucleic-acid-binding domain of TFIIS (C-terminal residues 231-280) contains a novel Cys₄ Zn²⁺-binding site that is required for activity¹⁰. In its absence, TFIIS fragments

retain RNAPII binding but inhibit (rather than stimulate) elongation. Nucleic-acid binding is not exhibited by the intact protein but is ‘unmasked’ by deletion of the RNAPII-binding domain. Unmasking of a cryptic nucleic-acid-binding domain by deletion of adjoining sequences has also been described in sigma factor σ^{70} , a subunit of *Escherichia coli* RNAP required for specific initiation²⁹; occluding sequences in the native protein presumably reorganize on binding to RNAP core subunits. An analogous model is proposed for TFIIS¹⁰ and supported by identification of point mutations in the C-terminal domain that reduce factor-stimulated transcriptional elongation and weaken nucleic-acid binding by the isolated Cys₄ domain (C.-J.J. and K.A., unpublished results). Assembly-dependent exposure of protein surfaces may limit nonspecific nucleic-acid binding by components of the basal transcriptional machinery^{10,29}.

When the TFIIS nucleic-acid-binding domain was expressed in *E. coli* with and without N-terminal His₆ sequences, nucleic-acid binding was similar in each case (Fig. 1). The peptides were monomeric, contained one Zn²⁺ ion and gave similar circular dichroism and ¹H-nuclear magnetic resonance (NMR) spectra (see Fig. 1 caption). The His₆ extended domain was investigated by three-dimensional NMR¹⁶, as described elsewhere²⁴. Its structure (peptide residues 5-50) was calculated by distance-geometry/restrained molecular dynamics (DG/RMD) based on 934 restraints (Fig. 2). The structure is elongated: a disordered loop is observed at one end (residues 27-34) and an internal Zn-binding site at the other (red arrow in Fig. 2*b*). The structure is similar to that of other Zn modules¹⁷⁻²⁰; in that Zn²⁺ coordination stabilizes a globular mini-domain with well defined secondary structure, nonpolar interior and polar or charged surface. The topology is otherwise different from previously characterized Zn modules^{17 21,26}, defining a new motif. Its central feature is a three-stranded antiparallel β -sheet (Fig. 2*b*); β -strands consist of residues 20-25, 35-40 and 45-48. Each C-X₂-C element defines a β -turn. Figure 2*d* shows the DG/RMD ensemble; representative side chains are labelled in Fig. 2*c*. Inspection of sequence databases identifies analogous motifs in the 16K subunit of yeast RNAPII (ref. 25), the RNAPII transcriptional initiation factor TFIIE (ref. 22), T-phage DNA primases²³ and other proteins involved in RNA or DNA transactions²⁴.

The α -helix and β -sheet provide complementary mechanisms for nucleic-acid binding^{27,28}. Such mechanisms are well characterized in sequence-specific recognition of double-stranded DNA. Less well understood are dynamic protein-nucleic-acid interactions involving non-canonical nucleic-acid structures, as must occur as reaction

intermediates in transcription, replication and recombination. Although its mechanism of nucleic-acid binding is not known, the TFIIS domain binds in gel-retardation experiments to single- and double-stranded DNA, RNA and DNA-RNA heteroduplex under high-salt conditions¹⁰, that is, the range of nucleic acids found in an elongation bubble. Intact TFIIS stimulates nascent transcript cleavage and antitermination of RNAPII complexes arrested at various template sequences¹¹⁻¹⁴. In accord with this broad range of biological targets, the TFIIS domain binds random nucleic-acid sequences but has preferential affinity for oligopyrimidine single strands¹⁰. As template DNA sequences in RNAPII pause and termination sites are frequently A-rich^{2,5-8,10,30}, potential interactions between the Zn ribbon and U-rich single-stranded RNA and RNA-DNA hetero-oligomers will be of particular interest. The TFIIS RNAPII-binding domain may contribute additional specificity to the interaction between the intact protein and the stalled elongation complex¹⁰.

The structure of the TFIIS nucleic-acid binding domain thus defines a novel Zn-dependent architecture. Designated the Zn ribbon, this structure consists of a three-stranded β -sheet and disordered loop. Analogous Zn-binding sites are predicted in a variety of other proteins which interact with non-canonical nucleic-acid structures. The present study provides a foundation for analysis of the interaction between TFIIS and oligonucleotide models of an elongation bubble⁴. Because of the presumed role of TFIIS in viral and proto-oncogene expression, its structure may also provide a target for antiviral and antineoplastic drug design.

ACKNOWLEDGEMENTS. We thank M. Izban, D. Luse and D. Reines for communication of results before publication, S. N. Gozani for database search and discussion, J. P. Lee for assistance with NMR measurements, A. Brunger, T. Havel and M. Nilges for computer programs and advice, Molecular Simulations, Inc. (Waltham, Massachusetts) for the program *Quanta* used to generate Fig. 2*d*, and K. Struhl and C. Richardson for discussion. Coordinates and NMR-derived restraint list have been deposited in the Brookhaven Data Bank. NMR spectra were obtained at Harvard Medical School. This work was supported in part by grants from the Council for Tobacco Research (M.A.W.) and NIH (K.A.). C.J.J. and H.-S.Y. contributed equally to this work.

References

1. Yanofsky, C. *J. biol. Chem.* **263**, 609-612 (1988).
2. Rice, G. A., Kane, C. M. & Chamberlin, M. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4245-4249 (1991).
3. von Hippel, P. H. & Yager, T. D. *Science* **255**, 809-812 (1992).
4. Daube, S. S. & von Hippel, P. H. *Science* **258**, 1320-1324 (1992).
5. Rougvie, A. E. & Lis, J. T. *Cell* **54**, 795-804 (1988).
6. Bentley, D. L. & Groudine, M. *Nature* **321**, 702-706 (1986).
7. Reines, D., Chamberlin, M. J. & Kane, C. M. *J. biol. Chem.* **264**, 10799-10809 (1989).
8. Feinberg, M. D., Baltimore, D. & Frankel, A. D. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4045-4049 (1991).
9. Yoo, O.-J. *et al. Nucleic Acids Res.* **19**, 1073-1079 (1991).
10. Agarwal, K. *et al. Biochemistry* **30**, 7842-7851 (1991).
11. Renies, D. *J. biol. Chem.* **267**, 3795-3800 (1992).
12. Izban, M. G. & Luse, D. S. *Genes Dev.* **6**, 1342-1356 (1992).
13. Wang, D. & Hawley, D. K. *Proc. natn. Acad. Sci. U.S.A.* **90**, 843-847 (1993).
14. Reines, D. & Mote, J. Jr *Proc. natn. Acad. Sci. U.S.A.* **90**, 1917-1921 (1993).
15. Kassavetis, G. A. & Geiduschek, E. P. *Science* **259**, 944-945 (1993).
16. Clore, G. M. & Gronenborn, A. M. *Prog. NMR Spectrosc.* **23**, 43-92 (1991).
17. Lee, M.S., Gippert, G. P., Soman, K. V., Case, D. A. & Wright, P. E. *Science* **245**, 635-637 (1989).
18. Hard, T. *et al. Science* **249**, 157-160 (1990).
19. Pavletich, N. P. & Pabo, C. O. *Science* **252**, 809-817 (1991).
20. Luisi, B. F. *et al. Nature* **352**, 497-505 (1991).
21. Marmorstein, R., Carey, M., Ptashne, M. & Harrison, S. C. *Nature* **356**, 408-414 (1992).
22. Ohkuma, Y. *et al. Nature* **354**, 398-401 (1991).
23. Dunn, J. J. & Studier, F. W. *J. molec. Biol.* **166**, 477-535 (1983).
24. Qian, X., Yoon, H.-S., Jeon, C.-J., Agarwal, A. & Weiss, M. A. *Biochemistry* (in the press).
25. Woychik, N.A., Lane, W. S. & Young, R. A. *J. biol. Chem.* **266**, 19053-19055 (1991).
26. Vallee, B. L., Coleman, J. E. & Auld, D. S. *Proc. natn. Acad. Sci. U.S.A.* **88**, 999-1003 (1991).

27. Somers, W. J. & Phillips, S. E. V. *Nature* **359**, 387-393 (1992).
28. Nikolov, D. B. *et al. Nature* **360**, 40-46 (1992).
29. Dombroski, A. J., Waiter, W. A., Record, M. T. Jr, Siegele, D. A. & Gross, C. A. *Cell* **70**,501-512 (1992).
30. Qian, X. & Weiss, M. A. *Biochemistry* **31**, 7463-7476 (1992).

List of Figures

Fig. 1 Binding of single-stranded DNA (ssDNA; 45 bases) by the isolated Zn ribbon and His₆ extended Zn ribbon is demonstrated by gel mobility-shift assay. Lane 1, 5' ³²P-labelled oligodeoxynucleotide alone; lane 2, ³²P-labelled ssDNA and Zn ribbon; and lane 3, ³²P-labelled ssDNA and His₆ extended Zn ribbon. The ssDNA sequence was 5'-GGCCCGTCCCCCTGTGCCCCTCTTTGTTCTGAATTCCTCTGGTTC-3' (designated D1; ref. 10).

METHODS. The domain (TFIIS residues 231-280; refs 9,10) was expressed as a 55-residue peptide containing 5 additional residues at the N terminus (MARAP; numbered -4 to 0) and as a 61-residue His₆ fusion peptide²⁴ containing 11 additional residues at the N terminus (MRHHHHHPMA; numbered -10 to 50). Peptide residues 1-50 correspond to protein residues 231-250 (ref. 9). The 55-residue peptide was purified as a 1:1 Zn²⁺ - complex by the method described for intact protein¹⁰; the His₆ extended peptide was purified as a 1:1 Zn²⁺ complex by nickel-affinity chromatography (Invitrogen, San Diego)²⁴. Thiolate coordination of one Zn²⁺ is demonstrated in the intact protein and fragments by atomic absorption and extended X-ray fine structure (EXAFS)¹⁰. In gel mobility-shift assays, 2 nM ³²P-5'-labelled oligodeoxyribonucleotide was incubated with 200 nM peptide at room temperature for 30 min in 200 mM KCl, 2 mM MgCl₂, 1mM dithiothreitol, 25 mM HEPES (pH 7.5), 0.1 mM ethylenediaminetetraacetic acid, 100 µg ml⁻¹ bovine serumalbumin, 0.05% NP-40 and 10% glycerol (vol/vol).

Fig. 2 *a*, Sequence of TFIIS domain^{9,10} (single-letter amino-acid code); cysteines in the Zn-binding site are shown in black. *b*, Ribbon model (stereo pair). The position of Zn is indicated by a red arrow; β -strands are labelled β_1 - β_3 . The model was generated using the program *Ribbon* (J. P. Priestle). *c*, DG/RMD model with selected side chains labelled. *d*, DG/RMD ensemble (stereo pair) aligned as in *c*. The main chain is shown in red and selected side chains in green. The N-terminal Sy in each C-X₂-C element (C12 and C40) forms bifurcating sulphur-amide hydrogen bonds²⁴; in the Zn-finger such bonds involve both Sy atoms^{17,19,30}. The interior of the TFIIS domain contains edge-to-face aromatic packing (involving W47, F10 and F49) reminiscent of that in the Zn-finger^{17,19}. DNA or RNA contact residues have not been determined; addition of ssDNA (5'-GAGATGACTCATCTC) leads to broadening of

resonances in the β -strands but not in the N-terminal extension or central loop.

METHODS. NMR spectra (three-dimensional TOCSY-NOESY, ^{15}N -edited three-dimensional NOESY and standard two-dimensional experiments) were obtained at 500 MHz in 50 mM deuterated Tris-HCl (pH 6.0) and 10 mM deuterated dithiothreitol at a peptide: Zn stoichiometry of 1:1 as described^{24,30}. No structural information was obtained for residues N-terminal to T5, which seem to be disordered. The structure of the central loop (residues 27-34) was also not well defined. DG/RMD structures (residues 5-50) were calculated using programs *DGII* (T. A. Havel) and *X-PLOR* (A. T. Brunger) based on 637 upper-bound distance restraints (162 intraresidue, 216 sequential, 48 medium-range and 211 long-range), 236 lower-bound restraints, 35 dihedral angle restraints (20ϕ and $15\chi_1$), 16 distance restraints related to hydrogen bonds (two restraints per hydrogen bond) and 10 distance restraints related to Zn coordination. Hydrogen-bond restraints were introduced based on slowly exchanging amide resonances in D_2O (ref. 30). Average violations of upper-bound distance and dihedral restraints were 0.028 Å and 1.2°, respectively. Excluding disordered regions, root-mean-square deviations were 0.40 Å and 0.89 Å for backbone and side-chain atoms, respectively. Assignment of ligands: Although the peptide contains six cysteines, only four are invariant among TFIIS homologues (Cys 12, Cys15, Cys 40 and Cys 43)⁹; mutation of the two non-conserved cysteines (Cys 20→Val and Cys 50→Ser) are functionally conservative and do not affect Zn^{2+} binding (C.-J.J. and K.A., unpublished results). In DG/RMD structures calculated in the absence of Zn-related restraints, the conserved cysteines define a unique tetrahedral Zn^{2+} -binding site. Inclusion of Cys 20 or Cys 50 in potential Zn-binding sites leads to significant restraint violations²⁴.

Fig. 1

Fig. 2