Volume 283, number 1, 23–26 FEBS 09730 May 1991
© 1991 Federation of European Biochemical Societies 00145793/91/$3.50
ADONIS 001457939100439W

Determination of DNA binding specificities of mutated zinc finger domains

Hans-Jürgen Thiesen and Christian Bach

Basel Institute for Immunology, Grenzacherstraße 487, Postfach CH-4005 Basel, Switzerland

Received 13 March 1991

By substituting non conserved amino acids present in the postulated α-helical region of zinc finger domains, we demonstrated that Cys_2/His_2 type zinc finger domains could be targeted to new DNA binding sites. The putative α-helical region of the second SP1 zinc finger (RSDELQRH) was replaced by amino acids (KSSALISH) occurring in analogous zinc finger positions of human zinc finger protein Kox 29. The DNA binding specificity of the FPLC purified chimaeric protein (SP1-Kox 29) was determined by use of the target detection assay (TDA). Chimaeric protein SP1-Kox 29 was subjected to randomized oligonucleotides (GGG NNNN GGC) that were designed on the basis that each SP1 zinc finger interacts with 3–4 nucleotides concerning its cognate target site GGG GCGG GGC. By this analysis the DNA binding specificity of SP1-Kox 29 was shown to have switched from the cognate SP1 binding site to GGG GGTG GGC. Structure-function analysis of this type should facilitate the determination of DNA binding specificities for any individual zinc finger of interest.

Human transcription factor SP1; Zinc finger protein domains; DNA-protein interactions; Zinc finger specific recognition code

1. INTRODUCTION

The classical Cys_2/His_2 zinc finger repeats first identified in *Xenopus* transcription factor IIIa [1,2] and in the *Drosophila* gap gene Krüppel [3] have been shown to be conserved in several hundred genes analysed from yeast to man [4–9]. The zinc finger domains ($YxCx_2Cx_3Fx_5Lx_2Hx_3HTGEKP$) are stabilized by zinc, tetrahedrally coordinating the 2 conserved cysteine and histidine residues [10]. A particular three-dimensional protein conformation seems to be required for achieving sequence-specific DNA contacts [10–15]. Interactive computer modelling strategies indicated the existence of an α-helical region in the zinc finger domain [11,12]. By nuclear magnetic resonance (NMR) analysis an α-helical structure was detected from amino acid position 16 to 24 including the conserved leucine and the conserved histidine [13,14] (for assignment of amino acid positions, see reference [12]). Therefore, in comparison to α-helical regions analysed in helix-turn-helix motifs [16,17] and in homeo-domains [18], it seems likely that the sequence specificity of zinc finger proteins is determined by specific amino acids that are positioned in the putative α-helical region and probably contact DNA via the major groove.

Analysis of mutations in ADR1 [19] and in *Drosophila* Krüppel [20] indicates that sequence-specific DNA binding is impaired once evolutionary conserved amino acid residues, such as cysteine, histidine or amino acid residues of the histidine/cysteine finger link, have been mutated. It is thought that in most of these cases the general conformation of the zinc finger is distorted, e.g. if one of the amino acids involved in coordinating zinc is replaced, the rigidity and structural constraints of the zinc finger are probably altered [20]. The formation of the postulated hydrophobic core in the putative finger tip are most likely effected once the conserved amino acids phenylalanine and leucine of the zinc finger consensus ($YxCx_2Cx_3Fx_5Lx_2Hx_3HTGEKP$) are mutated [21].

In order to deduce specific rules for zinc finger specific DNA-protein interactions, the chimaeric protein SP1-Kox 29 was generated by substituting the amino acid region (KSSALISH), present in Kox 29, for RSDELQRH in the second SP1 finger. By applying the target detection assay (TDA) the analysis of DNA recognition done with chimaeric SP1-Kox 29 protein reveals that SP1-Kox 29 has gained a new DNA target specificity.

2. METHODS AND MATERIALS

2.1. *Oligonucleotides*

Oligonucleotides for site-directed mutagenesis and for electrophoretic mobility shift assays (EMSA) were synthesized on 380A Applied Biosystems DNA Synthesizer. The oligonucleotide 1892 used in EMSA contains one SP1 binding site TTGGGGCGGGGCTT surrounded by cassette sequences [22], which contain the appropriate primer annealing sites for primer A and primer B [23,24]. For the TDA analysis randomized cassette oligonucleotide 2736 was generated [25] resulting in TGGGGNNNNGGCTT. For site-directed

Correspondence address: H.-J. Thiesen, Basel Institute for Immunology, Grenzacherstrasse 487, Postfach, CH-4005 Basel, Switzerland. Fax: (41) (61) 6811304.

mutagenesis oligonucleotide 2744 (AAGTCGTCTGCCCTAATTA-GTCACAAACGTACACACACAGGTGAGAAG) and oligonucleotide 2745 (GTGACTAATTAGGGCAGACGACTTTGTGAAGCG-TTTCCCACAGTATGA) were synthesized encoding lysine (K) at zinc finger position 15, serine (S) at position 17, alanine (A) at position 18, isoleucine (I) at position 20 and serine at position 21 (for assignment of amino acid positions, see reference [12]). The oligonucleotides 393 (GTAAAACGACGGCCAGTG) and 392 (AA-ACAGCTATGACCATG), which are universal primers of Bluescript plasmid (Stratagene), have been used together with the oligonucleotides 2744 and 2745 in PCR mutagenesis [26]. Oligonucleotide 1956 (CAGCCCGGGAGATCTGCCACCTGCATGAC) introduces a *Bgl*II site at the 3' end of the SP1 fragment in pB-516c.

2.2. *Site-directed mutagenesis*

The *Bam*HI-*Bgl*II fragment derived from $pP_{st}Sp1$-516c [27], encoding 3 zinc fingers of the human transcription factor SP1 [5], was cloned into the *Bam*HI site of Bluescript (Stratagene) to yield pB-516c. Two polymerase chain reactions (PCR) were performed using oligonucleotide pairs 393/2745 and 2744/392 together with pB-516c generating Sp1 fragments A and B (Fig. 1). Each fragment harbours the introduced mutations either at the 3' or 5' site. They were isolated from a 6% polyacrylamide gel [25]. The complete SP1 fragment encoding the desired mutations and a restored *Bgl*II site was generated by performing a second PCR using primer 393, 1956 on SP1 fragments A and B. The PCR product was extracted with phenol/chloroform, digested with *Bam*HI and *Bgl*II, gel purified and cloned into pAR 3039 [28] to yield pAR-SP1-Kox 29. Standard PCR conditions were applied [25]. Introduced mutations were verified by dideoxy sequence analysis.

2.3. E. coli *expression*

Mutated Sp1 protein was expressed and purified according to the procedure described for the analogous wild type SP1 protein [24]. Mutated SP1 protein was diluted 1:10 in buffer A (8 M urea, 20 mM MES pH 5.0, 2 mM EDTA), subjected to FPLC Mono S chromatography and eluted with an increasing salt gradient of buffer B (1 M NaCl, 8 M urea, 20 mM MES pH 5.0, 2 mM EDTA). Peak fractions were collected and analysed together with recently purified SP1 [23,24] on 15% polyacrylamide-SDS gel. Fractions containing the mutated SP1 protein were pooled. Protein concentrations were determined by the method of Bradford [30] to be 0.5 mg/ml.

2.4. *Target detection assay (TDA)*

Double-stranded randomized cassette oligonucleotides 2736 were generated by PCR [23] and labelled using TAQ polymerase [24]. Because of the reduced complexity of utilized oligonucleotide 2736 ($n = 4$) the TDA selection [23] was simplified by performing only 3 consecutive electrophoretic mobility shift assays (EMSA). Chimaeric SP1-Kox 29 (15 ng) protein was incubated with 10 μl labelled oligonucleotide 2736 (10000–20000 Cerenkov cpm) in a 30 μl standard electrophoretic mobility shift assay (EMSA). The reaction mixture consisted of 10 μl of 3 × band shift buffer (15 mM NaCl, 150 mM KCl, 36 mM HEPES pH 7.9, 36% glycerol, 5 mM $MgCl_2$, 300 μM $ZnCl_2$), 6 μl H_2O, 3 μl DTT (10 mM) and 10 μl labelled oligonucleotide 2736 (10000–20000 Cerenkov cpm), 1 μl protein (15 ng). SP1-Kox 29 protein was diluted by addition of H_2O. The band shift reactions were incubated for 30 min at R.T. and loaded onto a 6% polyacrylamide band shift gel (acrylamide/bisacrylamide 30%:0.8%) containing 100 μM $ZnCl_2$ and 0.25 × Tris-borate electrophoresis buffer (TBE) [24]. The gel electrophoresis was run in 0.25 × TBE at 10 V/cm and a current not exceeding 20 mA. The gels were dried onto Whatman 3MM paper and exposed to Kodak XAR. The retarded band shifts were eluted, amplified, subjected to 2 further rounds of TDA selection and finally cloned into the *Eco*RV site of Bluescript (Stratagene) [24].

2.5. *Electrophoretic Mobility Shift Assay (EMSA)*

Double-stranded oligonucleotide 1892 was generated by PCR amplification using primer A and primer B [22] and labelled as previously described [24]. Cloned fragments derived from TDA selected oligonucleotide 2736 were amplified by PCR and labelled. By performing EMSA analysis, the fragment N4CB4 present in plasmid pN4CB4 was identified to bind selectively to SP1-Kox 29 in contrast to oligonucleotide 1892. The binding site in pN4CB4 was determined by dideoxy sequence analysis.

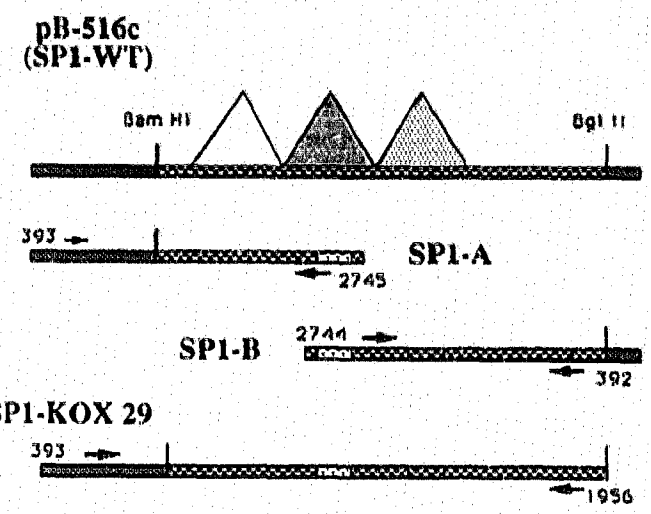

Fig. 1. PCR based site-directed mutagenesis [26]. The 3 SP1 zinc fingers are represented by 3 pyramids. The second finger was mutated by using oligonucleotides 2744 and 2745.

3. RESULTS AND DISCUSSION

The modular organization of zinc finger proteins supports the notion that certain rules for zinc finger specific DNA protein interactions can be deduced by a careful structure-function analysis of individual zinc finger domains [9]. Recently, non conserved amino acids present in the putative alpha helical regions of Kox-20 [31] and of SP1 [29] have been identified to alter the binding specificities of zinc finger domains. However, this type of analysis is hampered whenever zinc finger mutants are generated whose cognate DNA target sites cannot be predicted. Thus, the analysis of DNA binding affinities is dependent on the availability of corresponding DNA target sites. In order to overcome this limitation, we introduced the target detection assay (TDA) [23,24]. The TDA method only requires a mixture of randomized oligonucleotides [22] and a purified DNA binding protein for identifying putative DNA target sites.

In this study, the chimaeric protein SP1-Kox 29 was employed to demonstrate that sequence-specific DNA recognition of zinc finger proteins can be altered by substituting amino acid residues in the putative alpha helical zinc finger region. The helical region of the second SP1 finger, RSDELQRH, was replaced by the amino acids KSSALISH, which are present at identical positions in one of the fingers in Kox 29, a human zinc finger protein [9]. Oligonucleotides with mutations in the second finger of SP1 were made by performing

PCR based site-directed mutagenesis on pB-516c (Fig. 1). The non conserved amino acids, arginine (R), aspartic acid (D), glutamic acid (E), glutamine (Q) and arginine (R) at zinc finger positions 15,17,18,20 and 21, were replaced by lysine (K), serine (S), alanine (A), isoleucine (I) and serine (S) (for assignment, see reference [12]). Recombinant mutated SP1-Kox 29 protein was purified by FPLC Mono S chromatography (Fig. 2A). Peak fraction 5 containing mutated SP1-Kox 29 protein was identified by sodium dodecyl sulfate (SDS)-polyacrylamide electrophoresis [32] (Fig. 2B).

The DNA binding specificity of the chimaeric protein SP1-Kox 29 was established by using a modified TDA procedure. In previous TDA studies, the analysis of recombinant wild type SP1 protein confirmed that SP1 protein preferentially selects high affinity binding sites containing GCG nucleotides in the middle of its cognate target site (GGGGCGGGGC) [17,30]. Footprint analysis, methylation interference studies and mutational analysis of SP1 target sites demonstrated that the SP1 target site encompasses 9–10 nucleotides [5]. Recent reports indicated that zinc fingers might interact independently of each other with their target site leading to the postulate that individual zinc fingers recognize about 3 nucleotides [29,31]. Theoretically,

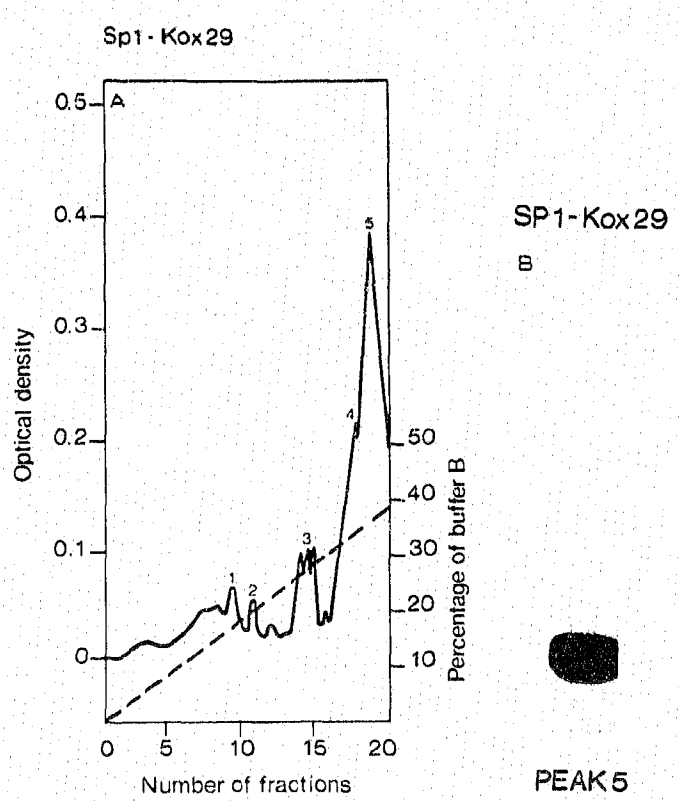

Fig. 2. (A) Elution pattern of chimaeric protein Sp1-Kox 29 by FPLC purification. Proteins bound to FPLC Mono S column were eluted by an increasing gradient of buffer B. (B) 15% SDS polyacrylamide gel electrophoresis [32] of purified SP1-Kox 29. SP1-Kox 29 was eluted in peak fraction 5, demonstrated by Coomassie blue staining of 25 μg purified SP1-Kox 29 protein.

each of the 3 SP1 zinc fingers can be assigned to 3 nucleotides of the target site. Thus, we assume that mutations in the second zinc finger of SP1 should only effect the DNA-protein interactions of the mutated protein to 3 or 4 nucleotides positioned in the middle of the SP1 cognate target site GGG GCGG GGC (Fig. 3). Therefore, we simplified the TDA procedure by reducing the complexity of the utilized randomized oligonucleotides by employing randomized oligonucleotides encoding GGG NNNN GGC. These oligonucleotides harbour only random nucleotides at 4 positions generating a complexity of 256 (4^4) putative target sites (Fig. 3).

The DNA binding activity of mutated SP1-Kox 29 protein was assessed by incubating double-stranded randomized oligonucleotides (n = 4) with chimaeric protein Sp1-Kox 29. After performing electrophoretic mobility shift assays (EMSA), the retarded oligonucleotides were isolated, amplified and applied to 2 further TDA rounds using EMSA. The affinity of SP1-Kox 29 to the cognate SP1 binding site was compared with affinities of the TDA selected oligonucleotides that were finally cloned. One fragment N4CB4 was identified to bind selectively SP1-Kox 29 (Fig. 4, lane 1a) as opposed to SP1 binding site GGG GCGG GGC (Fig. 4, lane 1b) which did not generate a visible band shift in EMSA analysis. Sequence of N4CB4 encodes nucleotides N_4 = GGTG. Thus, the structure-function analysis of mutated SP1 identifies the introduced mutations in the second finger as a loss of function (no binding to the cognate SP1 site) as well as a gain of function mutation (new target specificity by binding to GGG GGTG GGC).

The mutations in the second finger of SP1-Kox 29 represents the first example where the binding affinity of SP1 has been changed by site-directed mutagenesis of amino acids that are not highly conserved in the zinc finger consensus [12]. Chimaeric protein SP1-Kox 29 displays a new target specificity. It should be highly interesting in assigning sequence-specific and non sequence-specific DNA interactions to individual amino acids once crystallographic data of zinc fingers are available. Structure-function analysis of the kind reported here will reveal whether a zinc finger specific DNA recognition code can be deduced and later on be

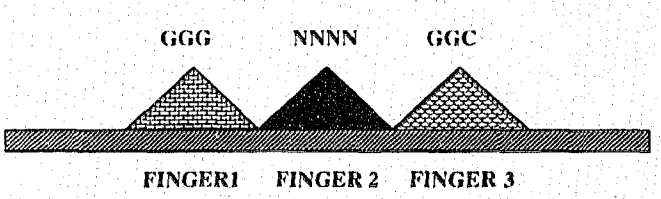

Fig. 3. Model of chimaeric SP1 protein SP1-Kox 29 binding to randomized oligonucleotides (n = 4) positioned in the SP1 cognate target site GGG GCGG GGC. Each finger of SP1 has been assigned to 3 or 4 nucleotides. This model implies that the second finger should always interact with target sites which have been preserved for finger 1 (GGG) and 3 (GGC).

Fig. 4. Electrophoresis mobility shift assay (EMSA) using SP1-Kox 29 protein. Chimaeric protein SP1-Kox 29 (15 ng) was incubated with labelled double stranded oligonucleotide N4CB4 encoding GGG GGTG GGC and with labelled double-stranded oligonucleotide 1892 that harbours one cognate SP1 binding site GGG GCGG GGC. (Lane 1a: ds 2736/N4CB4; Lane 1b: ds 1892.) Protein–DNA complexes were analysed on a 6% polyacrylamide gel.

applied for designing DNA binding domains with desired specificities [9].

The experimental procedure of mutating zinc finger domains and analysing corresponding effects in their DNA target specificities can be applied for studying any zinc finger mutation of interest. It can be estimated that probably several thousand individual zinc finger domains occur in nature representing Cys/His type zinc finger structures. For example, the human genome encodes several hundred zinc finger proteins with multiple zinc finger domains [7,8]. Based on the assumption that each finger recognizes 3–4 nucleotides, less than 256 putative DNA binding sites might be recognized by them. Thus, numerous non identical zinc fingers should be capable of interacting with identical DNA target sites. However, it seems likely that zinc fingers that display similar target specificities may have different affinities for target DNA. In this study, the DNA target site specificities of zinc finger proteins could be altered by replacing amino acid residues in the putative alpha helical zinc finger region and detected by employing the target detection assay (TDA). In ongoing experiments we are currently analysing whether the binding specificity of an individual zinc finger, determined in the background of the SP1 zinc finger configuration, is dependent or independent of the binding specificities of the neighbouring zinc fingers. If the latter assumption should turn out to be the case, it seems likely that in the near future DNA binding sites of zinc finger domains can be predicted from their primary protein structure.

Acknowledgements: Peter Lane and Roland Lauster are thanked for critical reading of the manuscript. R. Tjian is thanked for supplying the plasmid $pP_{\alpha}Sp1\text{-}516e$ [27]. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche Ltd., CH-4005 Basel, Switzerland.

REFERENCES

[1] Ginsberg, A.M., King, B.O. and Roeder, R.G. (1984) Cell 39, 479–489.
[2] Miller, J., McLachlan, A.D. and Klug, A. (1985) EMBO J. 4, 1609–1614.
[3] Rosenberg, U.B., Schroeder, C., Priess, A., Kienlin, A., Côte, S., Riede, I. and Jäckle, H. (1986) Nature 319, 336–339.
[4] Nehlin, O.J. and Ronne, H. (1990) EMBO J. 9, 2891–2898.
[5] Chowdhury, K., Deutsch, U. and Gruss, P. (1987) Cell 48, 771–778.
[6] Kadonaga, J.T., Carner, K.R., Masiar, F.R. and Tjian, R. (1987) Cell 51, 1079–1090.
[7] Bellefroid, E.J., Lecocq, P.J., Benhida, A., Poncelet, D.A., Belayew, A. and Martial, J.A. (1989) DNA 8, 377–387.
[8] Bray, P. and Thiesen, H.J. (1990) N. Biol. 2, 325–327.
[9] Thiesen, H.J. (1990) N. Biol. 2, 363–374.
[10] Diakun, G.P., Fairall, L. and Klug, A. (1986) Nature 324, 689–691.
[11] Berg, J. (1988) Proc. Natl. Acad. Sci. USA 85, 99–102.
[12] Gibson, T.J., Postma, J.P.M., Brown, R.S. and Argos, P. (1988) Protein Eng. 2, 209–218.
[13] Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A. and Wright, P.E. (1989) Science 245, 635–637.
[14] Parraga, G., Horvath, S.J., Eisen, A., Taylor, W.E., Hood, L., Young, E.T. and Klevit, R.E. (1988) Science 241, 1489–1492.
[15] Gottesfeld, J.M., Blanco, J. and Tennant, L.K.L. (1987) Nature 329, 460–462.
[16] Wharton, R.P. and Ptashne, M. (1985) Nature 316, 601–605.
[17] Anderson, W.F., Ohlendorf, D.H., Takeda, Y. and Mathews, B.W. (1981) Nature 290, 754–758.
[18] Kissinger, C.R., Liu, B., Martin-Blanco, E., Kornberg, T.B. and Pabo, C.O. (1990) Cell 63, 579–590.
[19] Blumberg, H., Eisen, A., Sledziewski, A., Bader, D. and Young, E.T. (1987) Nature 328, 443–445.
[20] Redemann, N., Gaul, U. and Jäckle, H. (1988) Nature 332, 90–92.
[21] Gaul, U., Redemann, N. and Jäckle, H. (1989) Proc. Natl. Acad. Sci. USA 86, 4599–4603.
[22] Thiesen, H.J. (1990) DNA Protein Eng. Tech. 2, 93–96.
[23] Thiesen, H.J. and Bach, C. (1990) Nucleic Acids Res. 18, 3203–3209.
[24] Thiesen, H.J. (1990) in: Immunological Methods (Lefkovitz, I. and Pernis, B. eds) vol. IV, pp. 61–74, Academic Press, Orlando, FL.
[25] Thiesen, H.J., Casorati, G., Lauster, R. and Wiles, M.V. (1990) in: Immunological Methods (Lefkovitz, I. and Pernis, B. eds) Vol. IV, Academic Press, Orlando, FL, pp. 61–74.
[26] Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Gene 77, 51–59.
[27] Courey, A.J. and Tjian, R. (1988) Cell 55, 887–898.
[28] Rosenberg, A.H., Lade, B.N., Chui, D.S., Lin, S.W., Dunn, J.J. and Studier, F.W. (1987) Gene 56, 125–135.
[29] Thiesen, H.-J. and Schröder, B. (1991) Biochem. Biophys. Res. Commun. (in press).
[30] Bradford, M.M. (1976) Anal. Biochem. 72, 248–254.
[31] Nardelli, J., Gibson, T.J., Vesque, C. and Charnay, P. (1991) Nature 349, 175–178.
[32] Laemmli, U.K. (1970) Nature 227, 680–685.
[33] Letovsky, J. and Dynan, W.S. (1989) Nucleic Acids Res. 17, 2639–2653.