Published in final edited form as:
J Struct Funct Genomics. 2014 December ; 15(4): 201–207. doi:10.1007/s10969-014-9184-z.

Solution NMR Structures of Homeodomains from Human Proteins ALX4, ZHX1, and CASP8AP2 Contribute to the Structural Coverage of the Human Cancer Protein Interaction Network

Xianzhong Xu,
Department of Chemistry, The State University of New York at Buffalo, and Northeast Structural Genomics Consortium, Buffalo, NY 14260, USA

Surya VSRK Pulavarti,
Department of Chemistry, The State University of New York at Buffalo, and Northeast Structural Genomics Consortium, Buffalo, NY 14260, USA

Alexander Eletsky,
Department of Chemistry, The State University of New York at Buffalo, and Northeast Structural Genomics Consortium, Buffalo, NY 14260, USA

Yuanpeng Janet Huang,
Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

Thomas B. Acton,
Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

Rong Xiao,
Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

John K. Everett,
Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

Gaetano T. Montelione, and
Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

Correspondence to: Thomas Szyperski, szypersk@buffalo.edu.

Department of Biochemistry and Molecular Biology, Robert Wood Johnson Medical School, Rutgers, The State University of New Jersey, Piscataway NJ 08854, USA

Thomas Szyperski

Department of Chemistry, The State University of New York at Buffalo, and Northeast Structural Genomics Consortium, Buffalo, NY 14260, USA

Thomas Szyperski: szypersk@buffalo.edu

Abstract

High-quality solution NMR structures of three homeodomains from human proteins ALX4, ZHX1 and CASP8AP2 were solved. These domains were chosen as targets of a biomedical theme project pursued by the Northeast Structural Genomics Consortium. This project focuses on increasing the structural coverage of human proteins associated with cancer.

Keywords

PF00046; transcription factor; homeodomain; structural genomics

Introduction

The human homeodomains comprising (i) residues 209–280 of “homeobox protein aristaless-like 4” (ALX4) (UniProtKB accession number: Q9H161), (ii) residues 462–532 (‘homeodomain 2’) of the “zinc fingers and homeoboxes protein 1” (ZHX1) (Q9UKY1), and (iii) the C-terminal domain comprising residues 1916–1982 of “caspase 8-associated protein 2” (CASP8AP2) (Q9UKL3) belong to the SCOP [1] ‘homeodomain-like’ superfamily SSF46689. ALX4(209–280) and ZHX1(462–532) belong to the very large Pfam [2] protein domain family *Homeobox* PF00046 which currently contains 25,115 members from eukaryotic organisms participating in 290 unique domain organizations (architectures). In contrast, CASP8AP2(1916–1982) has not yet been assigned to a Pfam family.

ALX4, a transcription activator, contains a single homeodomain and belongs to the ‘paired-class’ of homeodomain proteins which bind to palindromic DNA target sequences as homo- or heterodimers [3]. It plays an important role in craniofacial development [4, 5] so that loss of ALX4 expression results in both craniofacial and epidermal defects [6, 7]: mutant proteins R218Q [8], Q246Stop [8], and R272P [5] either completely or partially abolish DNA binding ability and transcriptional activation, which cause a rare genetic disorder, parietal foramina 2, manifested by abnormal skull bone development. In particular, the premature termination of transcription of mutant R265Stop [6] leads to ‘frontonasal dysplasia 2’ which is also associated with mental retardation [8]. Furthermore, hyper methylation of ALX4 is correlated with lung [9], bladder [10], gastric [11], and colorectal cancers [12–14], and reduced expression of ALX4 has been suggested as a biomarker for breast cancer[15].

ZHX1, likewise a transcriptional repressor [16, 17], contains five homeodomains and is ubiquitously expressed [16]. ZHX1 interacts with nuclear factor Y subunit A (NFYA) and DNA methyl transferase 3B (DNMT3B) for its repression activity [18, 19]. Changes in

expression profiles of rat ZHX1 ortholog have been associated with glomerular disease [20, 21]. In addition to the five homeodomains, ZHX1, which also contains of two N terminal C2H2 zincfingers [22], forms homodimers via homeodomain 1 [23] and can also form heterodimers with ZHX3 [24]. Thus far, the solution NMR structure of a polypeptide segment containing the two zinc fingers (PDB:2GHF) [25] as well as X-ray structures of homeodomains 3 (2ECB) and 4 (3NAR) [26] were solved. ZHX1(462–532), the construct chosen for the present study, contains homeodomain 2 which shares very low sequence identity (< 20 %) with the other four homeodomains.

CASP8AP2, which may function as either a transcriptional activator or repressor, contains one homeodomain, and is involved in apoptosis, cell cycle progression through S-phase and activation of histone expression [27–29]. CASP8AP2 is a prognostic marker for acute lymphoblastic leukemia (ALL), as expression levels of CASP8AP2 are correlated to cells undergoing apoptosis in ALL cells [30, 31]. CASP8AP2 interacts with a nuclear protein, ataxia-telangiectasia locus (NPAT), an activator of histone transcription [32] and nuclear receptor coactivator 2 (NCOA2), an activator in glucocorticoid receptor activation [33]. CASP8AP2 C-terminal deletion mutants (involving residues 1403–1962 and 1916–1962 for NPAT; 1709–1982 for NCOA2) do not interact with NPAT and NCOA2 [32, 33].

ALX4(209–280), ZHX1(462–532) and CASP8AP2(1916–1982) were selected by the Northeast Structural Genomics (NESG) Consortium (http://www.nesg.org; target IDs HR4490C, HR7907F, and HR8150A, respectively) for structure determination with the aim to provide extensive structural coverage for human cancer-associated proteins and protein complexes constituting the ‘Human Cancer Protein Interaction Network’ (HCPIN) [34]. Here we report the high-quality solution NMR structures of ALX4(209–280), ZHX1(462–532), and CASP8AP2(1916–1982).

Materials and methods

ALX4(209–280), ZHX1(462–532), and CASP8AP2(1916–1982) were cloned, expressed, and purified following protocols [35–37] established by the NESG (see Supplementary Material and http://www.nmr2.buffalo.edu/nesg.wiki for details). The corresponding pET expression vectors [NESG HR4490C-209–280-NHT.2, HR7907F-462–532-AV6HT.2 and HR8150A-1916–1982-AV6HT.3] have been deposited in the PSI Materials Repository (http://psimr.asu.edu/). Protein samples for ALX4(209–280), ZHX1(462–532), and CASP8AP2(1916–1982), were prepared at 0.9. 0.4 and 0.8 mM concentrations, respectively, in 90% H_2O/10% D_2O containing 10 mM Tris-HCl, 100 mM NaCl, 10 mM DTT, 50 M DSS, 0.02% NaN_3 at pH 6.5 for ALX4 (209–280), or 20 mM MES, 100 mM NaCl, 10 mM DTT, 5 mM $CaCl_2$, 50 M DSS, 0.02% NaN_3 at pH 7.5 for ZHX1(462–532) and CASP8AP2(1916–1982), respectively. Additional [5% ^{13}C; *U*-^{15}N]-labeled samples enabled stereospecific assignments of the methyl groups of Val and Leu residues [38]. Isotropic overall rotational correlation times of ~7, ~ 6, and ~5 ns for ALX4(209–280), ZHX1(462–532), and CASP8AP2(1916–1982), respectively, were inferred from average ^{15}N spin relaxation times (see Supplementary Material and http://www.nmr2.buffalo.edu/nesg.wiki), indicating that all three proteins are monomeric in solution. This finding was confirmed with

analytical gel-filtration (Agilent Technologies) and static light scattering (Wyatt Technology Co.) (Figs. S1 – S3).

NMR data were acquired at 25 °C on Varian INOVA 600 or 750 MHz spectrometers equipped with cryogenic $^{1}H\{^{13}C,^{15}N\}$ probes. The total measurement times for ALX4(209–280), ZHX1(462–532), and CASP8AP2(1916–1982) were ~9, ~8, and ~7 days, respectively. Nearly complete sequence-specific ^{1}H, ^{15}N and ^{13}C resonance assignments (Table I; Figs. S4, S5, and S6) were obtained using G-matrix Fourier transform (GFT) triple resonance experiments for targets ALX4(209–280) and CASP8AP2(1916–1982), and with conventional triple-resonance NMR experiments for ZHX1(462–532) (Supplementary Material) using the automated resonance assignment program AutoAssign 2.3.0 [39, 40], followed by manual assignment of side-chain resonances. Chemical shifts, NOESY peak lists, and time domain NMR data were deposited in the BioMagResBank [accession numbers 18805, 18714, and 18352 for ALX4(209–280), ZHX1(462–532), and CASP8AP2(1916–1982), respectively].

Structure calculations were performed using standardized methods of the NESG consortium [41, 42] (http://www.nmr2.buffalo.edu/nesg.wiki). The process included consensus analysis of automated NOESY cross peak assignments provided by the programs CYANA [43, 44] and AutoStructure 2.2.1 [45] based on ^{1}H-^{1}H NOE-derived upper limit distance constraints, and analysis of backbone dihedral angle constraints derived from chemical shifts using the program TALOS+ [46] for residues located in well-defined regular structure elements. Stereospecific assignments of methylene protons were performed with the GLOMSA module of CYANA and the final structure calculations were performed with CYANA followed by refinement of selected conformers in an “explicit water bath” using the program CNS [47]. Structure validation of the resulting 20 refined conformers was performed with the Protein Structure Validation Software (PSVS) server 1.4 [48] and the agreement of structures and NOESY peak lists was verified using the AutoStructure/RPF 2.2.1 package[41].

Results and discussion

High-quality NMR structures of ZHX1(462–532), ALX4(209–280), and CASP8AP2(1916–1982) were obtained (Fig. 1a, Fig. S7, Table I) and their coordinates were deposited in the Protein Data Bank (PDB) [49], respectively, on 09/14/2012 (accession code 2LY9), 10/24/2012 (2M0C), and 03/27/2012 (2LR8). The three structures exhibit the well-known [50] fold of homeodomains (Fig. 1a) consisting of three α-helices (Table I) and are quite similar: the root mean square deviations (RMSDs) calculated for the mean coordinates of the backbone heavy atoms N, C, and C’ of the three -helices are 0.9, 1.8, and 1.6 Å, respectively, for ALX4(209–280) and ZHX1(462–532), ALX4(209–280) and CASP8AP2(1916–1982), and ZHX1(462–532) and CASP8AP2(1916–1982).

A structure based sequence alignment using the DALI server [51] reveals that ALX4(209–280) shares 26% sequence identity (over 61 aligned residues) and 19% sequence identity (over 58 residues) with ZHX1(462–532) and CASP8AP2(1916–1982), respectively. ZHX1(462–532) shares 11% with CASP8AP2(1916–1982) (over 56 residues) (Fig.1b). A

search of the PDB for structurally similar proteins using the DALI server [51] yielded, as expected for structures of homeodomains [50], a very large number of highly significant hits with Z-scores > 4.0 [*i.e.* 244, 229, and 468 for ALX4(209–280), ZHX1(462–532), and CASP8AP2(1916–1982), respectively]. Remarkably, however, the highest scoring hit (Z-score 9.0) for ALX4(209–280) was a structure comprising residues 87 to 144 of the aristaless homeodomain protein [52] (Q06453), Al(87–144), from *Drosophila melanogaster* bound to DNA (PDB ID: 3LNQ: RMSD of 1.9 Å for the C atoms of 58 structurally aligned residues). The corresponding structure based sequence alignment reveals a high sequence identity of 78% for the two domains (Fig. 1b). Like ALX4, Al belongs to the 'paired-class' of homeodomain proteins which bind to palindromic DNA target sequences as homoor heterodimers [3]. Specifically, ALX4(209–280) binds to the palindromic repeat target sequence comprising 5'- TAAT-3' [53] preferably with three intervening nucleotides (5' – TAAT NNN ATTA-3'), while Al binds to a consensus DNA sequence of 5'-(T/C)TAATTAA(T/A)(T/A)G-3'[54]. Since the sequence alignment (Fig. 1b) shows that all residues interacting with the DNA duplex in the structure of the Al(87–144)-DNA complex [52] are conserved in ALX4(209–280), it appears quite likely that ALX4(209–280) binds to DNA in a very similar manner [52] (Fig. 1c). This finding suggests that this particular homeodomain-DNA interaction motif evolved before mammals evolved. Consistently, the sequence of human ALX4(209–280) is entirely conserved in all known mammalian genomes (see Supporting Information). Even though highly significant DALI hits were obtained also for ZHX1(462–532) [likewise Al(87–144) with Z-score = 8.8, RMSD – 1.5 Å, sequence identity 28%] , and CASP8AP2(1916–1982) [e.g., mouse 2610100B20RIK gene product homologous to Myb DNA binding protein; 1UG2; Z-score = 6.9; sequence identity: 43%], similar insights into the corresponding DNA interactions could not be derived because atomic resolution structures of DNA-complexes of homologues are not available.

Structural insights into Al(87–144)-DNA interactions (Fig. 1c) allow one to hypothesize that the ALX4 mutation R218Q [8] results in a malignant phenotype because the ALX4-DNA interaction is weakened and corresponding transcriptional activation impeded: the side chain of Arg 218 is expected to form hydrogen bonds with bases A and T in the minor groove which are removed in mutant protein. In contrast, mutant protein R272P [5] might result in a malignant phenotype because ALX4 homo- and/or heterodimer formation is affected. Arg 272 is located in the C-terminal flexible tail which is not expected to interact with DNA. However, this polypeptide segment plays a role for the formation of ALX4 dimer formation [3, 55] and the prolyl residue might restrict the conformational flexibility of this segment possibly required for protein-protein complex formation. Note that in the Q246Stop mutant the homeodomain is not formed, which abolishes DNA binding [8].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health, grant number: U54 GM094597 (T.S. and G.T.M.).

Abbreviations

ALX4	Aristaless-like 4
ZHX1	Zinc fingers and homeoboxes protein 1
CASP8AP2	Caspase 8 associated protein 2
Al	Aristaless
HCPIN	Human Cancer Pathway Interaction Network
DSS	4,4-dimethyl-4-silapentane-1-sulfonate sodium salt
DTT	Dithiothreitol
NESG	Northeast Structural Genomics Consortium
NOE	Nuclear Overhauser effect
PDB	Protein Data Bank
RMSD	Root mean square deviation

References

1. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol. 1995; 247:536–540. [PubMed: 7723011]
2. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer ELL, Eddy SR, Bateman A. The Pfam protein families database. Nucleic Acids Res. 2010; 38:D211–D222. [PubMed: 19920124]
3. Wilson D, Sheng G, Lecuit T, Dostatni N, Desplan C. Cooperative dimerization of paired class homeo domains on DNA. Genes Dev. 1993; 7:2120–2134. [PubMed: 7901121]
4. Qu S, Tucker SC, Zhao Q, deCrombrugghe B, Wisdom R. Physical and genetic interactions between Alx4 and Cart1. Development. 1999; 126:359–369. [PubMed: 9847249]
5. Wuyts W, Cleiren E, Homfray T, Rasore-Quartino A, Vanhoenacker F, Van HW. The ALX4 homeobox gene is mutated in patients with ossification defetcs of the skull (foramina parietalia permagna,OMIM 1685000). J Med Genet. 2000; 37:916–920. [PubMed: 11106354]
6. Kayserili H, Uz E, Niessen C, Vargel I, Alanay Y, Tuncbilek G, Yigit G, Uyguner O, Candan S, Okur H, Kaygin S, Balci S, Mavili E, Alikasifoglu M, Haase I, Wollnik B, Akarsu NA. ALX4 dysfunction disrupts craniofacial and epidermal development. Hum Molec Genet. 2009; 18:4357–4366. [PubMed: 19692347]
7. Joshi PA, Chang H, Hamel PA. Loss of Alx4, a stromally-restricted homeodomain protein, impairs mammary epithelial morphogenesis. Dev Biol. 2006; 297:284–294. [PubMed: 16916507]
8. Mavrogiannis LA, Antonopoulou I, Baxova A, Kutilek S, Kim CA, Sugayama SM, Salamanca A, Wall SA, Morriss-Kay GM, Wilkie AOM. Haploinsufficiency of the human homeobox gene ALX4 causes skull ossification defects. Nature Genet. 2001; 27:17–18. [PubMed: 11137991]
9. Liu WB, Han F, Du XH, Jiang X, Li YH, Liu Y, Chen HQ, Ao L, Cui ZH, Cao J, Liu JY. Epigenetic silencing of Aristaless-like homeobox-4, a potential tumor suppressor gene associated with lung cancer. Int J Cancer. 2013
10. Yu J, Zhu T, Wang Z, Zhang H, Qian Z, Xu H, Gao B, Wang W, Gu L, Meng J, Wang J, Feng X, Li Y, Yao X, Zhu J. A novel set of DNA methylation markers in urine sediments for sensitive/specific detection of bladder cancer. Clin Cancer Res. 2007; 13:7296–7304. [PubMed: 18094410]
11. Chen HY, Zhu BH, Zhang CH, Yang DJ, Peng JJ, Chen JH, Liu FK, He YL. High CpG island methylator phenotype is associated with lymph node metastasis and prognosis in gastric cancer. Cancer Sci. 2012; 103:73–79. [PubMed: 22017425]

12. Ebert MP, Model F, Mooney S, Hale K, Lograsso J, Tonnes-Priddy L, Hoffmann J, Csepregi A, Rocken C, Molnar B, Schulz HU, Malfertheiner P, Lofton-Day C. Aristaless-like homeobox-4 gene methylation is a potential marker for colorectal adenocarcinomas. Gastroenterology. 2006; 131:1418–1430. [PubMed: 17101318]
13. Tanzer M, Balluff B, Distler J, Hale K, Leodolter A, Rocken C, Molnar B, Schmid R, Lofton-Day C, Schuster T, Ebert MP. Performance of epigenetic markers SEPT9 and ALX4 in plasma for detection of colorectal precancerous lesions. PLoS One. 2010; 5:e9061. [PubMed: 20140221]
14. Zou H, Harrington JJ, Shire AM, Rego RL, Wang L, Campbell ME, Oberg AL, Ahlquist DA. Highly methylated genes in colorectal neoplasia: implications for screening. Cancer Epidemiol Biomarkers Prev. 2007; 16:2686–2696. [PubMed: 18086775]
15. Chang H, Mohabir N, Done S, Hamel PA. Loss of ALX4 expression in epithelial cells and adjacent stromal cells in breast cancer. J Clin Pathol. 2009; 62:908–914. [PubMed: 19783719]
16. Yamada K, Printz RL, Osawa H, Granner DK. Human ZHX1: cloning, chromosomal location, and interaction with transcription factor NF-Y. Biochem Biophys Res Commun. 1999; 261:614–621. [PubMed: 10441475]
17. Chen S, Yu X, Lei Q, Ma L, Guo D. The SUMOylation of zinc-fingers and homeoboxes 1 (ZHX1) by Ubc9 regulates its stability and transcriptional repression activity. J Cell Biochem. 2013; 114:2323–2333. [PubMed: 23686912]
18. Kim SH, Park J, Choi MC, Kim HP, Park JH, Jung Y, Lee JH, Oh DY, Im SA, Bang YJ, Kim TY. Zinc-fingers and homeoboxes 1 (ZHX1) binds DNA methyltransferase (DNMT) 3B to enhance DNMT3B-mediated transcriptional repression. Biochem Biophys Res Commun. 2007; 355:318–323. [PubMed: 17303076]
19. Yamada K, Osawa H, Granner DK. Identification of proteins that interact with NF-YA. Febs Lett. 1999; 460:41–45. [PubMed: 10571058]
20. Clement LC, Liu G, Perez-Torres I, Kanwar YS, Avila-Casado C, Chugh SS. Early changes in gene expression that influence the course of primary glomerular disease. Kidney Int. 2007; 72:337–347. [PubMed: 17457373]
21. Liu G, Clement LC, Kanwar YS, Avila-Casado C, Chugh SS. ZHX proteins regulate podocyte gene expression during the development of nephrotic syndrome. J Biol Chem. 2006; 281:39681–39692. [PubMed: 17056598]
22. Iuchi S. Three classes of C2H2 zinc finger proteins. Cell Mol Life Sci. 2001; 58:625–635. [PubMed: 11361095]
23. Yamada K, Kawata H, Matsuura K, Shou Z, Hirano S, Mizutani T, Yazawa T, Yoshino M, Sekiguchi T, Kajitani T, Miyamoto K. Functional analysis and the molecular dissection of zincfingers and homeoboxes 1 (ZHX1). Biochem Biophys Res Commun. 2002; 297:368–374. [PubMed: 12237128]
24. Yamada K, Kawata H, Shou Z, Hirano S, Mizutani T, Yazawa T, Sekiguchi T, Yoshino M, Kajitani T, Miyamoto K. Analysis of zinc-fingers and homeoboxes (ZHX)-1-interacting proteins: molecular cloning and characterization of a member of the ZHX family, ZHX3. Biochem J. 2003; 373:167–178. [PubMed: 12659632]
25. Wienk H, Lammers I, Hotze A, Wu J, Wechselberger RW, Owens R, Stammers DK, Stuart D, Kaptein R, Folkers GE. The tandem zinc-finger region of human ZHX adopts a novel C2H2 zinc finger structure with a C-terminal extension. Biochemistry. 2009; 48:4431–4439. [PubMed: 19348505]
26. Bird L, Ren J, Nettleship J, Folkers G, Owens R, Stammers D. Novel structural features in two ZHX homeodomains derived from a systematic study of single and multiple domains. BMC Structural Biology. 2010; 10:13. [PubMed: 20509910]
27. Kino T, Ichijo T, Chrousos GP. FLASH interacts with p160 coactivator subtypes and differentially suppresses transcriptional activity of steroid hormone receptors. J Steroid Biochem Mol Biol. 2004; 92:357–363. [PubMed: 15698540]
28. Barcaroli D, Bongiorno-Borbone L, Terrinoni A, Hofmann TG, Rossi M, Knight RA, Matera AG, Melino G, De Laurenzi V. FLASH is required for histone transcription and S-phase progression. Proc Natl Acad Sci U S A. 2006; 103:14808–14812. [PubMed: 17003125]

29. Milovic-Holm K, Krieghoff E, Jensen K, Will H, Hofmann TG. FLASH links the CD95 signaling pathway to the cell nucleus and nuclear bodies. EMBO J. 2007; 26:391–401. [PubMed: 17245429]
30. Jiao Y, Cui L, Gao C, Li W, Zhao X, Liu S, Wu M, Deng G, Li Z. CASP8AP2 is a promising prognostic indicator in pediatric acute lymphoblastic leukemia. Leuk Res. 2012; 36:67–71. [PubMed: 21696825]
31. Flotho C, Coustan-Smith E, Pei D, Iwamoto S, Song G, Cheng C, Pui CH, Downing JR, Campana D. Genes contributing to minimal residual disease in childhood acute lymphoblastic leukemia: prognostic significance of CASP8AP2. Blood. 2006; 108:1050–1057. [PubMed: 16627760]
32. Kiriyama M, Kobayashi Y, Saito M, Ishikawa F, Yonehara S. Interaction of FLASH with arsenite resistance protein-2 is involved in cell cycle progression at S phase. Mol Cell Biol. 2009; 29:4729–4741. [PubMed: 19546234]
33. Kino T, Chrousos GP. Tumor necrosis factor alpha receptor- and Fas-associated FLASH inhibit transcriptional activity of the glucocorticoid receptor by binding to and interfering with its interaction with p160 type nuclear receptor coactivators. J Biol Chem. 2003; 278:3023–3029. [PubMed: 12477726]
34. Huang YJ, Hang D, Lu LJ, Tong L, Gerstein MB, Montelione GT. Targeting the human cancer pathway protein interaction network by structural genomics. Mol Cell Proteomics. 2008; 7:2048–2060. [PubMed: 18487680]
35. Acton TB, Gunsalus KC, Xiao R, Ma LC, Aramini J, Baran MC, Chiang YW, Climent T, Cooper B, Denissova NG, Douglas SM, Everett JK, Ho CK, Macapagal D, Rajan PK, Shastry R, Shih LY, Swapna GVT, Wilson M, Wu M, Gerstein M, Inouye M, Hunt JF, Montelione GT. Robotic cloning and protein production platform of the Northeast Structural Genomics Consortium. Methods Enzymol. 2005; 394:210–243. [PubMed: 15808222]
36. Acton TB, Xiao R, Anderson S, Aramini J, Buchwald WA, Ciccosanti C, Conover K, Everett J, Hamilton K, Huang YJ, Janjua H, Kornhaber G, Lau J, Lee DY, Liu GH, Maglaqui M, Ma LC, Mao L, Patel D, Rossi P, Sahdev S, Shastry R, Swapna GVT, Tang YF, Tong SC, Wang DY, Wang H, Zhao L, Montelione GT. Preparation of protein samples for NMR structure, function, and small-molecule screening studies. Methods Enzymol. 2011; 493:21–60. [PubMed: 21371586]
37. Xiao R, Anderson S, Aramini J, Belote R, Buchwald WA, Ciccosanti C, Conover K, Everett JK, Hamilton K, Huang YJ, Janjua H, Jiang M, Kornhaber GJ, Lee DY, Locke JY, Ma LC, Maglaqui M, Mao L, Mitra S, Patel D, Rossi P, Sahdev S, Sharma S, Shastry R, Swapna GVT, Tong SN, Wang D, Wang H, Zhao L, Montelione GT, Acton TB. The high-throughput protein sample production platform of the Northeast Structural Genomics Consortium. J Struct Biol. 2010; 172:21–33. [PubMed: 20688167]
38. Neri D, Szyperski T, Otting G, Senn H, Wuthrich K. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional 13C labeling. Biochemistry. 1989; 28:7510–7516. [PubMed: 2692701]
39. Moseley HNB, Monleon D, Montelione GT. Automatic determination of protein backbone resonance assignments from triple resonance nuclear magnetic resonance data. Methods Enzymol. 2001; 339:91–108. [PubMed: 11462827]
40. Zimmerman DE, Kulikowski CA, Huang YP, Feng WQ, Tashiro M, Shimotakahara S, Chien CY, Powers R, Montelione GT. Automated analysis of protein NMR assignments using methods from artificial intelligence. J Mol Biol. 1997; 269:592–610. [PubMed: 9217263]
41. Huang YJ, Powers R, Montelione GT. Protein NMR recall, precision, and F-measure scores (RPF scores): structure quality assessment measures based on information retrieval statistics. J Am Chem Soc. 2005; 127:1665–1674. [PubMed: 15701001]
42. Liu G, Shen Y, Atreya HS, Parish D, Shao Y, Sukumaran DK, Xiao R, Yee A, Lemak A, Bhattacharya A, Acton TA, Arrowsmith CH, Montelione GT, Szyperski T. NMR data collection and analysis protocol for high-throughput protein structure determination. Proc Natl Acad Sci USA. 2005; 102:10487–10492. [PubMed: 16027363]
43. Güntert P, Mumenthaler C, Wüthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol. 1997; 273:283–298. [PubMed: 9367762]

44. Herrmann T, Güntert P, Wüthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol. 2002; 319:209–227. [PubMed: 12051947]
45. Huang YJ, Tejero R, Powers R, Montelione GT. A topology-constrained distance network algorithm for protein structure determination from NOESY data. Proteins. 2006; 62:587–603. [PubMed: 16374783]
46. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR. 1999; 13:289–302. [PubMed: 10212987]
47. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr Sect D Biol Crystallogr. 1998; 54:905–921. [PubMed: 9757107]
48. Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. Proteins. 2007; 66:778–795. [PubMed: 17186527]
49. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucleic Acids Res. 2000; 28:235–242. [PubMed: 10592235]
50. Gehring WJ, Affolter M, Burglin T. Homeodomain proteins. Annu Rev Biochem. 1994; 63:487–526. [PubMed: 7979246]
51. Holm L, Sander C. Dali: a network tool for protein structure comparison. Trends Biochem Sci. 1995; 20:478–480. [PubMed: 8578593]
52. Miyazono K, Zhi Y, Takamura Y, Nagata K, Saigo K, Kojima T, Tanokura M. Cooperative DNA-binding and sequence-recognition mechanism of aristaless and claswless. The EMBO Journal. 2010; 29:1613–1623. [PubMed: 20389279]
53. Tucker SC, Wisdom R. Sitespecific heterodimerization by paired-class homeodomain proteins mediates selective transcriptional responses. J Biol Chem. 1999; 274:32325–32332. [PubMed: 10542273]
54. Kojima T, Tsuji T, Saigo K. A concerted action of a paired-type homeobox gene, aristaless, and a homolog of Hox11/tlx homeobox gene, clawless, is essential for the distal tip development of the Drosophila leg. Dev Biol. 2005; 279:434–445. [PubMed: 15733670]
55. Stark MR, Johnson AD. Interaction between two homeodomain proteins is specified by a short C-terminal tail. Nature. 1994; 371:429–432. [PubMed: 8090224]
56. Moseley HNB, Sahota G, Montelione GT. Assignment validation software suite for the evaluation and presentation of protein resonance assignment data. J Biomol NMR. 2004; 28:341–355. [PubMed: 14872126]
57. Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. Proteins. 2007; 66:778–795. [PubMed: 17186527]

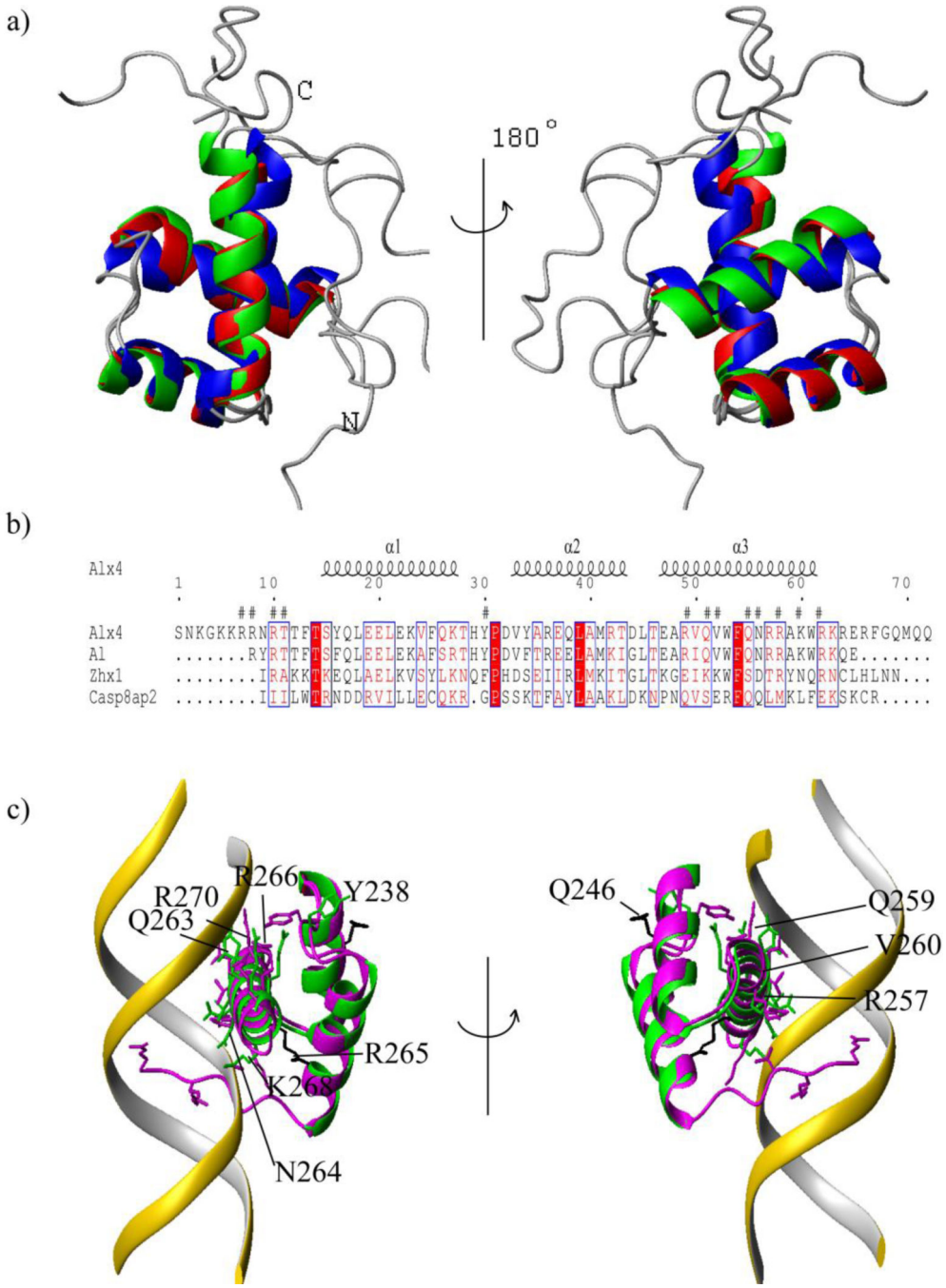

Fig. 1.

a: Overlay of the ribbon representations of ALX4 (209–280) (green), ZHX1 (462–532) (red), and CASP8AP2 (1916–1982) (blue). **b**: Structure based sequence alignment of ALX4(209–280), Al(87–144), ZHX1(462–532) and CASP8AP2(1916–1982) Residues involved in protein-DNA interactions in the Al homeodomain DNA complex (PDB ID: 3LNQ) are indicated with the # symbol. **c**: Superposition of ALX4(209–280) (green) onto Al(87–144) (pink) bound to double-stranded DNA. Side chains involved in protein-DNA interactions are shown as stick representations. The corresponding interacting residues in

ALX4 are also shown (except for those located in the termini). Mutations of ALX4 residues Q246 and R265 (shown in black) have been implicated in parietal foramina 2 and frontonasal dysplasia 2, respectively [8].

Table I

Statistics for NMR Structures of ALX4(209–280), ZHX1(462–532), and CASP8AP2(1916–1982)

	ALX4 (209–280)	ZHX1 (462–532)	CASP8AP2 (1916–1982)
Completeness of resonance assignments[b]			
Backbone (%)	74.0	98.9	94.3
Side chain (%)	70.8	93.2	93.2
Aromatic (%)	100	94.4	78.3
Stereospecific methyl (%)	100	100	100
Conformationally-restricting constraints[c]			
Distance constraints			
Total	841	875	892
intra-residue ($i = j$)	188	278	180
sequential ($\|i-j\| = 1$)	216	204	245
medium range ($1 < \|i - j\| < 5$)	232	217	271
long range ($\|i - j\| \geq 5$)	205	176	196
Dihedral angle constraints	80	62	64
Hydrogen bond constraints	0	0	0
No. of long range constraints per residue	3.7	2.7	3.1
Residual constraint violations[c]			
Average no. of distance violations per structure:			
0.1 – 0.2 Å	2.7	2.7	5.8
0.2 – 0.5 Å	0.6	0.7	1.8
> 0.5 Å	0	0	0
Average no. of dihedral angle violations per structure:			
1 – 10°	0.7	1.6	3.9
> 10°	0	0	0
Model Quality[c]			
RMSD backbone atoms (Å)[d]	0.5	0.6	0.6
RMSD heavy atoms (Å)[d]	1.0	1.1	1.1
RMS deviation for bond lengths (Å)	0.022	0.022	0.022

	ALX4 (209–280)		ZHX1 (462–532)		CASP8AP2 (1916–1982)	
RMS deviation for bond angles (°)		1.1		1.1		1.1
MolProbity Ramachandran statistics[c,d]						
Most favored regions (%)		99.3		99.3		94.9
Allowed regions (%)		0.7		0.7		5.1
Disallowed regions (%)		0.0		0.00		0.0
Global quality scores (Raw / *Z*-score)[c]						
Verify3D	0.21	−4.01	0.16	−4.82	0.15	−4.98
ProsaII	0.65	0.00	0.34	−1.28	0.50	−0.62
Procheck (phi-psi)[d]	0.37	1.77	0.48	2.20	0.18	1.02
Procheck (all)[d]	0.28	1.66	0.36	2.13	0.12	0.71
MolProbity clash score	5.87	0.52	7.06	0.31	9.98	−0.19
RPF Scores[e]						
Recall / Precision	0.987	0.894	0.989	0.888	0.930	0.867
F-measure / DP-score	0.938	0.804	0.936	0.773	0.897	0.750
BMRB accession number		18805		18714		18352
PDB ID		2M0C		2LY9		2LR8

[a] Structural statistics computed for the ensemble of 20 deposited structures.

[b] Computed using AVS software [56] from the expected number of resonances, excluding: highly exchangeable protons (N-terminal, Lys, and Arg amino groups, hydroxyls of Ser, Thr, Tyr), carboxyls of Asp and Glu, and non-protonated aromatic carbons. The comparatively low completeness of assignment for ALX4 is due to the missing H^N signals while aliphatic signals were detected at the flexible tails.

[c] Calculated using PSVS 1.4 [57]. Average distance violations were calculated using the sum over r^{-6}.

[d] Based on ordered residue ranges [S(phi) + S(psi) > 1.8], Residues 222–269 for ALX4, 472–489 and 492–519 for ZHX1, and 1928–1977 for CASP8AP2. Three α helices for each protein: I (473–485), II (491–501), and III (505–518) for ZHX1(462–532), I (223–235), II (241–251), and III (255–269) for ALX4(209–280), and I (1931–1943), II (1948–1958), and III (1962–1977) for CASP8AP2(1916–1982).

[e] RPF scores [41] reflecting the goodness-of-fit of the final ensemble of structures (including disordered residues) to the NOESY data and resonance assignments.