Published in final edited form as:

J Mol Biol. 2008 March 28; 377(3): 636–646.

Identification and Structural Characterization of a CBP/p300-Binding Domain from the ETS Family Transcription Factor GABP α

Hyun-Seo Kang, Mary L. Nelson[†], Cameron D. Mackereth, Manuela Schärpf, Barbara J. Graves[†], and Lawrence P. McIntosh^{*}

Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, BC, Canada V6T 1Z3. Department of Chemistry, and The Michael Smith Laboratory, University of British Columbia, Vancouver, BC, Canada V6T 1Z3.

[†]*Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA 84112*

Summary

Using NMR spectroscopy we identified and characterized a previously unrecognized structured domain near the N-terminus (residues 35–121) of the ETS family transcription factor GABP α . The monomeric domain folds as a 5-stranded β -sheet crossed by a distorted helix. Although globally resembling ubiquitin, the GABP α fragment differs in its secondary structure topology and thus appears to represent a new protein fold that we term the OST (On-SighT) domain. The surface of the GABP α OST domain contains two predominant clusters of negatively-charged residues, suggestive of electrostatically-driven interactions with positively-charged partner proteins. Following a best candidate approach to identify such a partner, we demonstrated through NMR-monitored titrations and GST-pulldown assays that the OST domain binds to the CH1 and CH3 domains of the co-activator histone acetyltransferase CBP/p300. This provides a direct structural link between GABP and a central component of the transcriptional machinery.

Keywords

Transcription factor; protein-protein interaction; ubiquitin; NMR spectroscopy

GA-binding protein (GABP), also known as E4 transcription factor 1 (E4TF-1) and nuclear respiratory factor 2 (NRF-2), is a widely-expressed member of the ETS transcription factor family with roles in gene regulation related to embryogenesis, immune system development, cell cycle progression, neuromuscular junctions, protein synthesis, cellular respiration, and viral pathogenicity.^{1,2,3,4,5,6,7} A distinguishing feature of GABP amongst ETS family members is that it recognizes tandem promoter sites as a heterotetrameric $\alpha\beta_2\alpha$ complex with the DNA-binding and transactivation activities provided by separate polypeptide chains (Fig. 1(a)). The GABP α subunit is composed of a C-terminal winged helix-turn-helix ETS domain,

^{*}Corresponding author: Lawrence P. McIntosh Department of Biochemistry and Molecular Biology Life Sciences Centre 2350 Health Sciences Mall University of British Columbia Vancouver, B.C., Canada, V6T 1Z3 E-mail: mcintosh@chem.ubc.ca Phone: 604–822–3341 Fax: 604–822–5227.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

which binds preferentially to purine-rich regions based on the sequence 5'-GGAA/T-3',⁸ and a central PNT domain. The helical bundle PNT domain⁹ appears to mediate protein-protein interactions with the co-activator histone acetyltransferase CBP/p300.¹⁰ Full length GABP β 1 and closely related GABP β 2 are composed of N-terminal ankyrin repeats (ANK), a transactivation domain (TAD), and a C-terminal leucine-zipper domain (LZ).¹ The LZ domain, which serves for homo/heterodimerization of the GABP β homologs, is absent in two of four alternatively spliced GABP β 1 isoforms.¹ As revealed by the crystal structure of an ETS domain/ANK/DNA ternary complex, GABP β ANK binds GABP α via the ETS domain and a flanking C-terminal sequence, and in doing so indirectly increases core promoter affinity.^{8,11}

GABP α has been reported to interact with several partner transcription factors, including ATF1,¹² Sp1 and Sp3,¹³ and C/EBP α ,¹⁴ as well as CBP/p300^{10,15,16} and histone deacetylase HDAC1.⁷ Although the molecular bases for these interactions remain unknown, those involving ATF1 and CBP/p300 have been mapped to the region of this protein preceding its ETS domain. In addition to the PNT domain, this region also includes a poorly characterized ~170 amino acid sequence that is conserved only among GABP α orthologs, and to a lesser extent, Elg homologs from the ETS family (Supplemental Figure S1). To provide a structural framework for establishing the functional role(s) of this sequence, we investigated by NMR spectroscopy a series of GABP α N-terminal truncation fragments, thereby discovering that residues 35–121 encompass a previously unrecognized folded domain. We determined that this monomeric domain is composed of a five-stranded β -sheet crossed by a distorted helix. Although globally similar to ubiquitin, the GABP α fragment has a distinct secondary structure topology, and thus represents a new fold that we term the OST (On-SighT) domain. Through in vitro binding studies, we also demonstrated that the isolated GABP α OST domain interacts with the cysteine-histidine rich CH1 and CH3 domains of CBP/p300, whereas the PNT domain only binds the latter.

Identification of the OST domain

A new structured domain in the previously uncharacterized N-terminal region of GABP α was evident from the ¹H-¹⁵N HSQC spectrum of GABP α ^{1–320} (Figure 1(b)). In addition to the expected ¹H^N-¹⁵N signals corresponding to the PNT domain,⁹ numerous additional well-dispersed peaks were also observed. The same peaks were present in the spectrum GABP α ^{1–169}, which lacks the PNT domain (not shown). Through partial backbone assignments of this latter truncation fragment, the new domain was localized approximately to residues 35–121, whereas the flanking residues 1–34 and 122–169 exhibited chemical shifts and ¹⁵N relaxation properties diagnostic of a random coil polypeptide. Allowing for a few additional N- and C-terminal residues, GABP α ^{35–121} was expressed and found to yield an excellent quality ¹H-¹⁵N HSQC spectrum. This spectrum, indicative of a well-folded domain termed OST, accounted for all of the additional peaks first observed with GABP α ^{1–320}.

Structure and dynamics of the OST domain

The structural ensemble of the OST domain was determined using an extensive set of NMR-derived restraints recorded for GABP α ^{35–121} (Figure 2, Table 1). The GABP α fragment adopts a compact monomeric fold comprised of a five-stranded mixed parallel/anti-parallel β -sheet (S₁; 36–43, S₂; 67–70, S₃; 73–74, S₄; 91–99, S₅; 107–115) crossed by a distorted helix (48–59). Strand S₅ contains a classic β -bulge centred at Leu111-Glu112. Chemical shift, NOE, and ³J_{NH-H α} coupling measurements demonstrate that the single helix in the OST domain begins with ₃10 conformation and extends to Leu59 as an α -helix, despite being kinked due to Pro57 (Supplemental Figure S2).

Complementing the structural analysis, the global and internal backbone dynamics of GABP α ^{35–121} were studied by ¹⁵N heteronuclear relaxation experiments. The isotropic tumbling time, τ_c , of the fragment was calculated to be 5.9 ± 0.04 ns at 30 °C from the T₁ and T₂ relaxation values of well-ordered amides in the OST domain.¹⁷ This is consistent with that expected for a globular, monomeric 10.2 kDa protein.¹⁸ The internal backbone dynamics of GABP α ^{35–121} were fit using the anisotropic Lipari-Szabo model-free approach (Supplemental Figure S3).^{19,20,21} The structured core of the protein, corresponding to the minimal OST domain (residues 36–115), is well ordered, and no significant changes in backbone mobility were found for the bulge in S₅ or within the distorted helix. In contrast, besides the N- and C-terminal tails, there are three relatively flexible loop regions: residues 61–63 between the helix and strand S₂, residues 87–91 between strands S₃ and S₄, and residues 101–104 between strands S₄ and S₅. Among these, the exposed S₄/S₅ loop has the highest degree of flexibility on the psec-nsec timescale, as revealed by reduced heteronuclear NOE values and model free order parameters. This region also has relatively high backbone RMS deviations in the structural ensemble (Figure 2 and Supplemental Figure S3).

Structural comparison of the OST domain with ubiquitin

To obtain possible insights into the function of the OST domain, searches for proteins with similar three dimensional structures were performed using DALI and VAST.^{22,23} Significant matches were found only with ubiquitin (1UBI.pdb, Z = 4.0 for 61 residues aligned) or ubiquitin-like proteins (Yeast ubiquitin-like protein Hub1, 1M94.pdb, Z = 4.0 for 57 residues; hypothetical *Arabidopsis thaliana* protein At1g16640, 1YEL.pdb, Z = 3.0 for 64 residues; mouse D7Wsu128e protein, 1V86.pdb, Z = 3.0 for 62 residues). Indeed, despite their disparate sequences (18% identity using ClustalW²⁴), the global fold of ubiquitin is strikingly similar to that of the OST domain with five stranded β -sheet crossed by a helix (Figure 3). However, upon closer inspection, several key differences exist between the two proteins. First, the linear order of their secondary structure elements is permuted such that ubiquitin has a helix between β -strands S₂ and S₃, whereas in the OST domain, the helix lies between S₁ and S₂. Second, relative to the remainder of the β -sheet, the central β -strand S₅ in ubiquitin has the opposite orientation as that in GABP α . Third, unlike ubiquitin with a regular α -helix, the helix in the OST domain begins with a 3_{10} conformation (Ile48 to Asn50), and ends at Leu59 as a distorted α -helix (Supplemental Figure S2). Finally, the OST domain has loops corresponding to the two short 3_{10} helices in ubiquitin, and its flexible S₄/S₅ loop matches the C-terminal tail of the latter. Based on these differences, and the lack of similarity to any other proteins of known structure, the OST domain appears to represent a new fold.

Surface properties of the OST domain

The lack of any structural similarities beyond that of the ubiquitin-like proteins precluded obvious clues to the function of the GABP α OST domain. Furthermore, no additional insights were provided by the ProFunc server, which attempts to identify the biochemical role of a protein from its three-dimensional structure.²⁵ Following the hypothesis that the OST domain mediates macromolecular interactions necessary for the regulation of transcription by GABP, we examined its surface properties. The amino acid composition of GABP α ^{35–121} has a relatively low predicted pI ~ 4.5 , and thus at neutral pH, is net negatively-charged. Most surface exposed Asp and Glu residues localize in two clusters on the opposite sides of the OST domain (Figure 2(d)). The largest of these is formed by carboxylate sidechains from the loops between S₁/H (Asp43, Glu46), S₃/S₄ (Asp76, Asp78, Asp83, Asp89), and S₄/S₅ (Glu105), and at the start of S₂ (Glu67). Although less prominent, the second cluster arises from residues in S₅ (Glu112) and at the C-terminal end (Glu118 and Glu121) of GABP α ^{35–121}. In contrast, no substantial regions of positively-charged or hydrophobic residues are present on the surface of the OST domain. This suggests that the GABP α OST domain may interact with positively-

charged partner proteins, or other biological molecules, at least in part through electrostatic interactions.

The GABP α OST domain interacts with the CBP CH1 and CH3 domains

Several previous studies have demonstrated that GABP α interacts with the general transcriptional co-activator CBP/p300.^{7,10,15,16} Thus, following a best candidate approach, we sought to determine if the OST domain and/or PNT domain contribute to this interaction by using pulldown assays with GST-tagged fragments corresponding to various regions of mCBP. As shown in Figure 4(f), the OST domain interacted with mCBP fragments containing the CH1 (or TAZ1; residues 340–439) and CH3 (or ZZ-TAZ2; residues 1680–1850) domains. Binding was not observed to other tested regions containing the nuclear receptor (residues 1–101) or KIX (residues 576–679) domains. The isolated PNT domain only bound to the CH3 domain (Figure 4(g)). Consistent with these results, GABP α ^{1–320}, containing both the OST and PNT domains, also bound to CH1 and CH3 containing fragments of mCBP (Figure 4(h)). Binding of the OST domain to the mCBP fragment containing the CH1 domain was significantly reduced upon increasing the concentration of KCl from 180 to 300 mM (not shown), indicating an electrostatic contribute towards their association.

The specific binding of the OST and CH1 domains was confirmed by reciprocal ¹H-¹⁵N HSQC monitored titration experiments. In each case, addition of the unlabeled partner led to a progressive decrease in the signal intensities of selected backbone or sidechain amides in the ¹⁵N-labeled domain (Figure 4(a-b)). Such binding in the intermediate-slow exchange regime on the chemical shift timescale is generally indicative of a dissociation constant K_d in the micromolar or lower range. However, the absence of new signals from the resulting complex upon saturation is suggestive of conformational exchange broadening and/or formation of species exceeding a 1:1 stoichiometry. Mapping of the intensity perturbations onto the tertiary structures of the OST and CH1 domains provides a qualitative identification of their respective binding interfaces (Figure 4(c-d)). For the OST domain, this interface partially overlaps its predominant negatively-charged surface patch and includes the flexible S₄/S₅ loop. In the case of the CH1 domain, the interface abuts a positively-charged region of its surface and partially overlaps the known binding sites for the transactivation domains of HIF-1 α and CITED2.^{26,27,28,29} It is also noteworthy that, at least for the OST domain, residues showing spectral perturbations are located over a rather large region of its surface. This could result indirectly through conformational changes upon CH1 binding, or might reflect multiple binding sites leading to higher order complex formation. Further structural, thermodynamic, and mutational studies of the interactions of the OST and CH1 domains are needed to resolve these possibilities. In contrast to the OST domain, the ¹H-¹⁵N HSQC spectra of the PNT domain remained completely unperturbed upon titration with the CH1 domain, confirming the lack of any significant interaction between these species (Figure 4(e)). Also, neither the OST or PNT domain bound to the isolated ZZ domain³⁰ of mCBP in NMR-monitored titrations (not shown), suggesting that their association with the CH3 fragment is dependent upon the TAZ2 domain.³¹ Studies to test this hypothesis are underway.

Concluding Remarks

Through deletion studies followed by NMR spectroscopic analyses, we have identified and structurally characterized a previously unrecognized domain in GABP α . Similar to the ETS and PNT domains, the OST domain adopts a stable folded structure when isolated. Importantly, the near exact superimposition of signals from the corresponding amides in the ¹H-¹⁵N HSQC spectra of GABP α ^{35–121} and GABP α ^{1–320} (Figure 1(b)) demonstrates that the OST domain retains this independent structure within the context of larger GABP α fragments and that it does not interact intramolecularly with the PNT domain or any flanking residues. Similarly, no

intermolecular interactions between the isolated OST and PNT domains were detectable through NMR-monitored titrations (not shown). Thus, at least in the context of GABP α ^{1–320}, the OST domain behaves as a “bead-on-a-string” (Figure 1(c)). Although not tested explicitly due to challenges in obtaining suitable samples for NMR analysis, this conclusion likely extends to the full length protein with the independently folded ETS domain. Also, algorithms for identifying disordered protein regions, such as PrDOS,³² predict that the sequences flanking the three domains of GABP α are unstructured. Interestingly, it has been reported that, in the absence of DNA, GABP α exists as a stable heterodimer, and that deletion of the N-terminal two-thirds of GABP α , along with serine substitutions for four cysteines in the remaining ETS domain-containing C-terminal fragment, enhances heterotetramer formation.^{33,34} This raises the possibility of additional GABP α interactions, beyond that between the ANK and ETS domains, which may involve the OST, PNT or other yet unidentified segments of these two proteins. Nevertheless, the “beads-on-a-string” model of GABP α provides great flexibility for the assembly of a complex network of transcriptional regulatory proteins at target promoter sites.

The OST domain globally resembles ubiquitin, yet differs significantly in its secondary structural topology. Furthermore, based on sequence similarities, this domain can only be recognized in GABP α orthologs including *Drosophila* Elg. Thus, the OST domain defines a new protein fold, possibly with functions in mediating protein-protein interactions specific to this subset of the ETS transcription factor family. An inspection of the surface properties of the OST domain revealed two predominant clusters of negatively-charged Asp and Glu residues, suggestive of electrostatically-driven binding to positively-charged partner proteins.

Following a best-candidate approach, we demonstrated by GST-pulldown assays that the OST domain of GABP α binds to fragments of the co-activator CBP/p300 containing the CH1 and CH3 domains, whereas the PNT domain only interacts with the latter. NMR-monitored titrations confirmed that the OST domain, but not PNT domain, specifically binds the isolated CH1 domain of mCBP. In addition to serving as a scaffold for assembly of specific transcriptional factors, CBP/p300 recruits the basal transcription factors and functions as a histone acetyltransferase to direct chromatin remodelling. Consistent with the results of Figure 4, binding of full length GABP α to fragments of p300 spanning residues 1–596 and 1572–2370, but not 744–1571, was reported based on in vitro GST-pulldown studies; these contain the CH1, CH3, and CH2 domains, respectively.¹⁵ However, more recently published GST-pulldown experiments indicated that the GABP α PNT domain binds to the CH3 domain of p300, whereas residues 1–143 (encompassing the OST domain) bind weakly to both the CH2 and CH3 domains; no interactions with the CH1 domain were detected.¹⁰ Resolving the exact functions of these GABP α regions will require unbiased searches for interacting proteins and additional in vitro and in vivo analyses, carried out with the knowledge of the newly-discovered OST domain. For example, co-immunoprecipitation measurements revealed that upon neuregulin-activation, presumably through MAPK or JNK signalling cascades, GABP binds p300, whereas in the unactivated state, it binds HDAC1 at N box promoters.⁷

The CH1 and CH3 domains of CBP/p300 have been implicated in binding a remarkable array of diverse transcription factors, suggesting a significant degree of structural plasticity for intermolecular complex formation. Both of the net positively-charged CH1 (or TAZ1) and CH3 (or ZZ-TAZ2) domains fold as helical bundles with zinc ions chelated via cysteine and histidine residues.^{30,31} To date, structures of the isolated CH1 domain of CBP/p300 in complex with peptides corresponding to the transactivation domains of CITED2 and HIF-1 α have been determined using NMR spectroscopy.^{26,27,28,29,35} In each case, binding induces folding of the otherwise disordered peptides to predominantly helical structures that wrap around extended surface regions of the CH1 domain. In contrast, the OST domain is a β -sheet module with a well-defined structure. Similarly, phosphorylation-dependent binding of Ets-1

and Ets-2 to the CH1 and CH3 domains of CBP/p300 is mediated by their α -helical PNT domains.³⁶ We are currently undertaking NMR spectroscopic studies to determine the structural bases for the interactions of these ETS family members with this central component of the transcriptional machinery.

Data Bank accession codes

The atomic coordinates of the GABP α ^(35–121) ensemble (accession code: 2JUO) have been deposited in the Research Collaboratory for Structural Bioinformatics Protein Databank (<http://www.rcsb.org/pdb/>), and the NMR chemical shift list (accession code: 15451) has been deposited in the BioMagResBank (<http://www.bmrb.wisc.edu/>).

Acknowledgements

This research was supported by grants from National Cancer Institute of Canada with funds from the Canadian Cancer Society (to L.P.M.) and the National Institutes of Health grants GM38663 (to B.J.G.) and CA42014 (to the Huntsman Cancer Institute). B.J.G. also acknowledges funding from the U.S. Department of Energy and the Huntsman Cancer Institute. Instrument support was provided by the Canadian Institutes for Health Research, the Protein Engineering Network of Centres of Excellence, the Canadian Foundation for Innovation, the British Columbia Knowledge Development Fund, the UBC Blusson Fund, and the Michael Smith Foundation for Health Research.

Abbreviations

ANK, ankryin repeat
 GST, glutathione S-transferase
 HSQC, heteronuclear single quantum correlation
 LZ, leucine zipper
 NMR, nuclear magnetic resonance
 NOE, nuclear Overhauser effect
 NOESY, nuclear Overhauser effect spectroscopy
 PNT, Pointed
 RDC, residual dipolar coupling
 RMS, root-mean-square
 TAD, transactivation domain. GABP α and CBP fragments are indicated by the residues spanned, e.g. GABP α ^(35–121) corresponds to residues 35–121 of the 454 amino acid protein.

References

1. Rosmarin AG, Resendes KK, Yang ZF, McMillan JN, Fleming SL. GA-binding protein transcription factor: A review of GABP as an integrator of intracellular signaling and protein-protein interactions. *Blood Cells Mol. Diseases* 2004;32:143–154.
2. Hsu T, Schulz RA. Sequence and functional properties of *ets* genes in the model organism *Drosophila*. *Oncogene* 2000;19:6409–6416. [PubMed: 11175357]
3. Risteovski S, O'Leary DA, Thornell AP, Owen MJ, Kola I, Hertzog PJ. The Ets transcription factor GABP alpha is essential for early embryogenesis. *Mol. Cell. Biol* 2004;24:5844–5849. [PubMed: 15199140]
4. Xue HH, Bollenbacher-Reilly J, Wu Z, Spolski R, Jing X, Zhang YC, McCoy JP, Leonard WJ. The transcription factor GABP is a critical regulator of β lymphocyte development. *Immun* 2007;26:421–431.
5. O'Leary DA, Noakes PG, Lavidis NA, Kola I, Hertzog PJ, Risteovski S. Targeting of the Ets factor GABP alpha disrupts neuromuscular junction synaptic function. *Mol. Cell. Biol* 2007;27:3470–3480. [PubMed: 17325042]
6. Yang ZF, Mott S, Rosmarin AG. The Ets transcription factor GABP is required for cell-cycle progression. *Nature Cell Biol* 2007;9:339–U164. [PubMed: 17277770]
7. Ravel-Chapuis A, Vandromme M, Thomas JL, Schaeffer L. Postsynaptic chromatin is under neural control at the neuromuscular junction. *EMBO J* 2007;26:1117–1128. [PubMed: 17304221]

8. Batchelor AH, Piper DE, de la Brousse FC, McKnight SL, Wolberger C. The structure of GABP alpha/beta: An Ets domain ankyrin repeat heterodimer bound to DNA. *Science* 1998;279:1037–1041. [PubMed: 9461436]
9. Mackereth CD, Scharpf M, Gentile LN, MacIntosh SE, Slupsky CM, McIntosh LP. Diversity in structure and function of the Ets family PNT domains. *J. Mol. Biol* 2004;342:1249–1264. [PubMed: 15351649]
10. Resendes KK, Rosmarin AG. GA-binding protein and p300 are essential components of a retinoic acid-induced enhanceosome in myeloid cells. *Mol. Cell. Biol* 2006;26:3060–3070. [PubMed: 16581781]
11. Graves BJ. Transcription - inner workings of a transcription factor partnership. *Science* 1998;279:1000–1002. [PubMed: 9490475]
12. Sawada J, Simizu N, Suzuki F, Sawa C, Goto M, Hasegawa M, Imai T, Watanabe H, Handa H. Synergistic transcriptional activation by hGABP and select members of the activation transcription factor/camp response element-binding protein family. *J. Biol. Chem* 1999;274:35475–35482. [PubMed: 10585419]
13. Galvagni F, Capo S, Oliviero S. Sp1 and Sp3 physically interact and cooperate with GABP for the activation of the utrophin promoter. *J. Mol. Biol* 2001;306:985–996. [PubMed: 11237613]
14. Shimokawa T, Ra C. C/EBP alpha functionally and physically interacts with GABP to activate the human myeloid IgA Fc receptor (Fc alpha R, Cd89) gene promoter. *Blood* 2005;106:2534–2542. [PubMed: 15928042]
15. Bannert N, Avots A, Baier M, Serfling E, Kurth R. GA-binding protein factors, in concert with the coactivator CREB binding protein p300, control the induction of the interleukin 16 promoter in t lymphocytes. *Proc. Natl. Acad. Sci. USA* 1999;96:1541–1546. [PubMed: 9990060]
16. Bush TS, Coeur MS, Resendes KK, Rosmarin AG. GA-binding protein (GABP) and Ap1 are required, along with retinoid receptors, to mediate retinoic acid responsiveness of Cd18 (beta(2) leukocyte integrin): A novel mechanism of transcriptional regulation in myeloid cells. *Blood* 2003;101:311–317. [PubMed: 12485937]
17. Dosset P, Hus JC, Blackledge M, Marion D. Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR* 2000;16:23–28. [PubMed: 10718609]
18. de la Torre JG, Huertas ML, Carrasco B. HydroNMR: Prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Mag. Reson* 2000;147:138–146.
19. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc* 1982;104:4559–4570.
20. Kay LE, Torchia DA, Bax A. Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: Application to Staphylococcal nuclease. *Biochem* 1989;28:8972–8979. [PubMed: 2690953]
21. Mandel AM, Akke M, Palmer AG. Backbone dynamics of *Escherichia coli* Ribonuclease-H1 - correlations with structure and function in an active enzyme. *J. Mol. Biol* 1995;144–163. [PubMed: 7531772]
22. Dietmann S, Park J, Notredame C, Heger A, Lappe M, Holm L. A fully automatic evolutionary classification of protein folds: Dali domain dictionary version 3. *Nuc. Acids Res* 2001;29:55–57.
23. Gibrat JF, Madej T, Spouge JL, Bryant SH. The VAST protein structure comparison method. *Biophys. J* 1997;72:298–298.
24. Higgins DG, Thompson JD, Gibson TJ. Using Clustal for multiple sequence alignments. *Comp. Meth. Macromol. Seq. Anal* 1996;266:383–402.
25. Laskowski RA, Watson JD, Thornton JM. PROFUNC: A server for predicting protein function from 3D structure. *Nucl. Acids Res* 2005;33:W89–W93. [PubMed: 15980588]
26. Dames SA, Martinez-Yamout M, De Guzman RN, Dyson HJ, Wright PE. Structural basis for HIF-1 alpha/CBP recognition in the cellular hypoxic response. *Proc. Natl. Acad. Sci. USA* 2002;99:5271–5276. [PubMed: 11959977]
27. De Guzman RN, Martinez-Yamout MA, Dyson HJ, Wright PE. Interaction of the TAZ1 domain of the CREB-binding protein with the activation domain of CITED2 - regulation by competition

- between intrinsically unstructured ligands for non-identical binding sites. *J. Biol. Chem* 2004;279:3042–3049. [PubMed: 14594809]
28. Freedman SJ, Sun ZYJ, Kung AL, France DS, Wagner G, Eck MJ. Structural basis for negative regulation of hypoxia-inducible factor-1 alpha by CITED2. *Nature Struct. Biol* 2003;10:504–512. [PubMed: 12778114]
 29. Freedman SJ, Sun ZYJ, Poy F, Kung AL, Livingston DM, Wagner G, Eck MJ. Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 alpha. *Proc. Natl. Acad. Sci. USA* 2002;99:5367–5372. [PubMed: 11959990]
 30. Legge GB, Martinez-Yamout MA, Hambly DM, Trinh T, Lee BM, Dyson HJ, Wright PE. ZZ domain of CBP: An unusual zinc finger fold in a protein interaction module. *J. Mol. Biol* 2004;343:1081–1093. [PubMed: 15476823]
 31. De Guzman RN, Liu HY, Martinez-Yamout M, Dyson HJ, Wright PE. Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP. *J. Mol. Biol* 2000;303:243–253. [PubMed: 11023789]
 32. Ishida T, Kinoshita K. ProDos: Prediction of disordered protein regions from amino acid sequence. *Nucl. Acids Res* 2007;W460–464. [PubMed: 17567614]
 33. Chinenov Y, Henzl M, Martin ME. The alpha and beta subunits of the GA-binding protein form a stable heterodimer in solution - revised model of heterotetrameric complex assembly. *J. Biol. Chem* 2000;275:7749–7756. [PubMed: 10713087]
 34. Chinenov Y, Schmidt T, Yang XY, Martin ME. Identification of redoxsensitive cysteines in GA-binding protein-alpha that regulate DNA binding and heterodimerization. *J. Biol. Chem* 1998;273:21430–21430.
 35. De Guzman RN, Wojciak JM, Martinez-Yamout MA, Dyson HJ, Wright PE. CBP/p300 TAZ1 domain forms a structured scaffold for ligand binding. *Biochem* 2005;44:490–497. [PubMed: 15641773]
 36. Foulds CE, Nelson ML, Blaszcak AG, Graves BJ. Ras/mitogenactivated protein kinase signaling activates Ets-1 and Ets-2 by CBP/p300 recruitment. *Mol. Cell. Biol* 2004;24:10954–10964. [PubMed: 15572696]
 37. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRpipe - a multidimensional spectral processing system based on Unix pipes. *J. Biomol. NMR* 1995;6:277–293. [PubMed: 8520220]
 38. Goddard, TD.; Kneller, DG. Sparky 3. Univerisity of California; San Francisco:
 39. Sattler M, Schleucher J, Griesinger C. Heteronuclear multidimensional nmr experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. NMR Spect* 1999;34:93–158.
 40. Yamazaki T, Foreman-Kay JD, Kay L. Two-dimensional NMR experiments for correlating $^{13}\text{C}^\beta$ and $^1\text{H}^\delta/\epsilon$ chemical shifts of aromatic residues in ^{13}C -labelled proteins via scalar couplings. *J. Am. Chem. Soc* 1993;115:11054–11055.
 41. Archer SJ, Ikura M, Torchia DA, Bax A. An alternative 3D-NMR technique for correlating backbone ^{15}N with side-chain H^β -resonances in larger proteins. *J. Magn. Reson* 1991;95:636–641.
 42. McIntosh LP, Brun E, Kay LE. Stereospecific assignment of the NH_2 resonances from the primary amides of asparagine and glutamine side chains in isotopically labeled proteins. *J. Biomol. NMR* 1997;9:306–312.
 43. Neri D, Szyperski T, Otting G, Senn H, Wüthrich 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 ^{13}C labeling. *Biochem* 1989;28:7510–7516. [PubMed: 2692701]
 44. Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM. Overcoming the overlap problem in the assignment of ^1H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^1H - ^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: Application to interleukin 1 beta. *Biochem* 1989;28:6150–6156. [PubMed: 2675964]
 45. Pascal SM, Muhandiram DR, Yamazaki T, Kay JDF, Kay LE. Simultaneous acquisition of ^{15}N and ^{13}C -edited NOE spectra of proteins dissolved in H_2O . *J. Magn. Reson* 1994;B:197–201.

46. Slupsky CM, Gentile LN, McIntosh LP. Assigning the NMR spectra of aromatic amino acids in proteins: Analysis of two Ets Pointed domains. *Biochem. Cell Biol* 1998;76:379–390. [PubMed: 9923707]
47. Zwahlen C, Gardner KH, Sarma SP, Horita DA, Byrd RA, Kay LE. An nmr experiment for measuring methyl-methyl NOEs in ^{13}C -labeled proteins with high resolution. *J. Am. Chem. Soc* 1998;120:7617–7625.
48. 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]
49. Bax A, Vuister GW, Grzesiek S, Delaglio F, Wang AC, Tschudin R, Zhu G. Measurement of homo- and heteronuclear J couplings from quantitative J correlation. *Methods. Enzymol* 1994;239:79–105. [PubMed: 7830604]
50. Ottiger M, Delaglio F, Bax A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Mag. Reson* 1998;131:373–378.
51. Ishii Y, Markus MA, Tycko R. Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel. *J. Biomol. NMR* 2001;21:141–151. [PubMed: 11727977]
52. Clore GM, Gronenborn AM, Tjandra N. Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J. Mag. Reson* 1998;131:159–162.
53. Pelton JG, Torchia DA, Meadow ND, Roseman S. Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from *Escherichia coli*, using two-dimensional heteronuclear NMR techniques. *Prot. Sci* 1993;2:543–558.
54. Schubert M, Labudde D, Oschkinat H, Schmieder P. A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on C-13 chemical shift statistics. *J. Biomol. NMR* 2002;24:149–154. [PubMed: 12495031]
55. Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM. Aqua and Procheck-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 1996;8:477–486. [PubMed: 9008363]
56. Hutchinson EG, Thornton JM. Promotif - a program to identify and analyze structural motifs in proteins. *Prot. Sci* 1996;5:212–220.
57. Willard L, Ranjan A, Zhang HY, Monzavi H, Boyko RF, Sykes BD, Wishart DS. VADAR: A web server for quantitative evaluation of protein structure quality. *Nuc. Acids Res* 2003;31:3316–3319.
58. Koradi R, Billeter M, Wüthrich K. MolMol: A program for display and analysis of macromolecular structures. *J. Mol. Graph* 1996;14:51–55. [PubMed: 8744573]
59. Delano WL, Lam JW. PyMol: A communications tool for computational models. *Abs. Papers Am. Chem. Soc* 2005;230:U1371–U1372.
60. Bax A, Kontaxis G, Tjandra N. Dipolar couplings in macromolecular structure determination. *Meth. Enyz* 2001;339:127–164.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Figure 1.

(a) Cartoon of GABP domain organization. Boundaries are predicted or correspond to fragments used for studies of the individual domains. (b) The close superimposition of corresponding peaks in the ^1H - ^{15}N HSQC spectra of ^{15}N -labeled GABP α^{35-121} (black) and GABP α^{1-320} (red) demonstrates that the OST domain adopts an independently folded structure. The additional well-dispersed peaks in the spectrum of GABP α^{1-320} arise from the PNT domain,⁹ whereas those clustered near ~8.0 to 8.6 ppm in ^1H correspond to residues with predominantly random coil conformations. Assignments are given for GABP α^{35-121} , and aliased peaks are denoted with an asterisk. (c) “Beads-on-a-string” model of full-length

GABP α showing the 3 structured domains (PNT, 1SXD.pdb; ETS, 1AWC.pdb) along with all remaining residues in extended conformations.

Methods: Genes encoding GABP α^{35-121} , GABP α^{1-169} , and GABP α^{1-320} were PCR-amplified from the full-length murine GABP α DNA (Genbank 34328119) and inserted in the vector pET28a (Novagen) for expression with an N-terminal His₆-tag and thrombin cleavage site. A codon for non-native Trp122 was added to the GABP α^{35-121} gene to facilitate concentration measurements by absorbance spectroscopy. His₆-GABP $\alpha^{168-256}$ was previously described.⁹ The GABP constructs were expressed in *E. coli* BL21(λ DE3) cells grown at 37 °C in LB or minimal M9T medium containing 1 g/l ¹⁵NH₄Cl for uniform ¹⁵N-labeling, 1 g/l ¹⁵NH₄Cl and 3 g/l ¹³C₆-glucose for uniform ¹³C/¹⁵N-labeling, and 0.3 g/l ¹³C₆-glucose and 2.7 g/l ¹²C₆-glucose for non-random 10% ¹³C-labeling (Stable Spectral Isotopes). After induction with 1 mM IPTG at OD₆₀₀ ~ 1 and growth for 4 hours at 37 °C, the cells were harvested by centrifugation. The cell pellets were resuspended in buffer A (5 mM imidazole, 50 mM HEPES (pH 7.5), 500 mM NaCl, 5 % glycerol) and lysed by passage through a French press three times at 10,000 psi, followed by 15 min of sonication on ice. The lysate was cleared by centrifugation, and the supernatant loaded onto a Ni-NTA column (GE Healthcare), pre-equilibrated with buffer A. The column was washed with 10 column volumes of buffer A plus 60 mM imidazole, and then the bound proteins were eluted with buffer A plus 250 mM imidazole. Following SDS-PAGE analysis, the appropriate fractions were pooled and dialyzed overnight in 20 mM sodium phosphate (pH 7.2) and 20 mM NaCl with a few crystals of thrombin (Roche) for the cleavage of the N-terminal His₆-tag. Thrombin and the cleaved His₆-tag were removed by incubation with *p*-aminobenzamidine beads (Sigma) and TALON metal affinity resin (BD Biosciences) at room temperature with mild shaking for 15 min, followed by centrifugation. The purified proteins, with an N-terminal Gly-Ser-His remaining from the cleavage site, were concentrated to ~ 1 to 1.5 mM in NMR sample buffer (20 mM sodium phosphate (pH 7.0), 50 mM NaCl, 2 mM DTT, ~10 % D₂O) using a 5 kDa cut-off Amicon Ultrafiltration device (Millipore). Spectra were recorded at 30 °C using Varian 500 MHz Unity, 600 MHz Inova, or 800 MHz Inova NMR spectrometers equipped with a triple resonance gradient probes, and analyzed using NMRpipe³⁷ and Sparky.³⁸ Main-chain and aliphatic side-chain ¹H, ¹³C, and ¹⁵N resonances were assigned using standard triple resonance correlation experiments.³⁹ Resonances from aromatic side-chains were assigned using ¹H-¹³C HSQC, C β H δ and C β H ϵ experiments.⁴⁰ Stereospecific assignments of prochiral H ^{β} , ^{β} signals were obtained from HNHB and short mixing time (30 msec) ¹⁵N-TOCSY-HSQC spectra,⁴¹ of Asn and Gln ¹⁵NH₂ groups from a EZ-HMQC-NH₂ spectrum,⁴² and of Val and Leu methyl groups from a constant time ¹H-¹³C HSQC spectrum of the non-randomly 10% ¹³C-labeled protein.⁴³

Figure 2.

(a) Ribbon diagram of the lowest-energy NMR-derived structure of the GABPα OST domain ("front" view), with the β-sheet, helix, and two distinct loops (S₃/S₄ and S₄/S₅) shown in yellow, red, and cyan, respectively. Superimposed backbone (b) and hydrophobic sidechain (c) atoms from the structural ensemble of GABPα³⁵⁻¹²¹, with selected residues labelled for reference. (d) Two clusters of negatively-charged Asp and Glu residues (red) are observed on the "front" (left) and "back" (right) of the OST domain. In contrast, no extended clusters of (d) positively-charged Arg, His, or Lys residues (blue) or (e) exposed hydrophobic residues (green) are found on its surface (polar residues in grey).

Methods: The tertiary structure of OST was calculated with ARIA(v1.2)/CNS utilizing distance, dihedral angle, and orientation restraints (Table 1). Interproton distance restraints were obtained from 3D ¹⁵N- NOESY-HSQC (600 MHz, 150 msec mixing time),⁴⁴ simultaneous aliphatic ¹³C- and amide ¹⁵N-NOESY-HSQC (800 MHz, 125 msec),⁴⁵ aromatic ¹³C- NOESY-HSQC (500 MHz, 125 msec),⁴⁶ and constant time methyl-methyl and amide-methyl-NOESY spectra (600 MHz, 140 msec),⁴⁷ followed by the automatic and iterative NOE assignments in ARIA. Mainchain dihedral angles were derived from ¹H, ¹³C, and ¹⁵N chemical shifts using TALOS.⁴⁸ The predicted dihedral angles were only used when the Φ values agreed with ³J_{NH-Hα} coupling constants obtained from a HNHA spectrum.⁴⁹ The χ₁ angle restraints for residues with prochiral H^β protons were determined in concert with the stereospecific assignments, and those of the Val, Thr, and Ile residues were determined from ³J_{NCα} and ³J_{C'Cα} coupling constants measured using N-Cα and C'-Cα spin echo experiments.⁴⁹ For residues not showing ³J couplings indicative of rotamer averaging, the χ₁ angles were set to ± 60° or 180° based on a staggered rotamer model. ¹H-¹⁵N RDC

orientational restraints were measured from ^1H - ^{15}N IPAP-HSQC spectra using protein aligned with stretched gels^{50, 51} and incorporated at iteration 4 with the SANI routine in ARIA, with default energy constants of 0.2 and 1 kcal mol⁻¹ Hz⁻² for the first and second simulated annealing cooling stages, respectively. Values of the alignment tensor ($R = 0.3$, $D_a = 5.3$ Hz) were estimated by the histogram method, followed by a grid search for minimal SANI violations.⁵² All His imidazole sidechains were determined to be in the neutral N $^{\epsilon 2}$ tautomeric state from long-range ^1H - ^{15}N correlation measurements.⁵³ All Xaa-Pro amides were constrained to the *trans* conformation based on a chemical shift analysis with POP.⁵⁴ In the final set of 10 water-refined structures, there were no violations greater than 0.5 Å and 5° for distance and angle restraints, respectively. The structural ensemble was analyzed using Procheck-NMR⁵⁵ and secondary structures determined according to Promotif⁵⁶ and Vadar.⁵⁷ Structural figures were generated using MOLMOL⁵⁸ and PyMOL.⁵⁹

Figure 3.

Structure comparisons of the GABP α OST domain and ubiquitin (1UBQ.pdb). (a) The global folds of the two proteins are clearly similar. (b) However, secondary structure topology diagrams show that the direction of β -strand S5 in the OST domain is opposite from that of the corresponding strand in ubiquitin. (c) The sequential orders of secondary structural elements are also different. In addition, ubiquitin has a regular α -helix and 2 short 3_{10} -helices, whereas the OST domain has a distorted α -helix preceded by a 3_{10} -helix.

Figure 4.

The GABP α OST domain binds to the mCBP CH1 and CH3 domains. (a,c) Superimposed ^1H - ^{15}N HSQC spectra of ^{15}N -labeled OST domain before (black, 0.1 mM) and after (red, 0.09 mM) addition of unlabeled CH1 (0.09 mM final). The top quartile of residues with backbone or sidechain amides showing intensity decreases due to CH1 domain binding are mapped in green on to the surface of the OST domain. (b,d) Superimposed ^1H - ^{15}N HSQC spectra of ^{15}N -labeled CH1 before (black, 0.1 mM) and after (red, 0.05 mM) addition of unlabeled OST domain (0.05 mM final). The top quartile of residues showing intensity decrease due to OST domain binding are mapped in green on to the surface of the CH1 domain (1U2N.pdb). (e) Superimposed ^1H - ^{15}N HSQC spectra of ^{15}N -labeled PNT domain before (black, 0.1 mM) and after (red, 0.1 mM) addition of unlabeled CH1 (0.1 mM final). The absence

of any spectral changes indicates that no measurable binding occurs between the PNT and CH1 domains. These conclusions are supported by pulldown assays, involving 200–300 pmol of the indicated GST-tagged mCBP fragments, pre-bound to glutathione-sepharose beads (20 μ L), and 5 μ M of FLAG-HMK-tagged (f) GABP α^{35-121} (OST domain), (g) GABP $\alpha^{168-256}$ (PNT domain), and (h) GABP α^{1-320} (OST and PNT domains). Following bead collection and washing, binding was analyzed by Western blotting using an anti-FLAG antibody.

Approximate molecular weight marker positions on the 4–15% SDS-PAGE gels are shown. The slower migrating band in the GABP α^{35-121} sample is attributed to disulfide-linked dimers.

(i) Cartoon of the domain structure of CBP.

NMR Methods: The gene encoding the His₆-tagged CH1 domain (mCBP^{340–439}) was PCR cloned from the full length CBP gene (Genbank 70995311) into the pET28b (Novagen) vector for expression in *E. coli* BL21(λ DE3) cells. Cells were grown at 30 °C in ZnSO₄ supplemented (150 μ M) LB or ¹⁵N-minimal M9T medium to OD₆₀₀ ~ 1, followed by induction with 1 mM IPTG and growth for overnight at 16 °C. Purification was as in Figure 1, except that buffer A was modified to 150 μ M ZnSO₄, 5 mM imidazole, 50 mM Tris (pH 7.5), 500 mM NaCl, and 5 % glycerol. The purified CHI (mCBP^{340–439}), OST (GABP α^{35-121}), and PNT (GABP $\alpha^{168-256}$) domains were dialyzed into NMR buffer (20 mM Tris (pH 6.9), 50 mM NaCl, 2 mM DTT, ~10 % D₂O) and concentrated by ultrafiltration. ¹H-¹⁵N HSQC spectra were recorded at 25 °C.

GST-pulldown Methods: A DNA cassette encoding the FLAG tag and heart muscle kinase (HMK) target sequences was inserted into pET28a *Nde*I site of the GABP α constructs to generate clones for His₆-FLAG-HMK-GABP α^{35-121} , His₆-FLAG-HMK-GABP $\alpha^{168-256}$, and His₆-FLAG-HMK-GABP α^{1-320} . Tagged fragments were purified as described for GABP α^{35-121} in Figure 1. GST pulldown assays were conducted in 20 mM Tris pH 7.9, 10% glycerol, 180 mM KCl, 0.2 mM EDTA, 10 μ M ZnSO₄, 0.5% NP40, 0.5 mM PMSF, 2–10 mM DTT, 0.1 mg/mL BSA, with Glutathione-Sepharose 4B beads (GE Healthcare). After incubation for 16 hr at 4 °C with constant rotation, the beads were collected by centrifugation and washed with binding buffer. The bound proteins were eluted using 3x SDS sample buffer, separated by SDS-PAGE, transferred to PVDF membranes, blotted with a 1:20,000 dilution of FLAG M2 antibody (Sigma), and visualized by ECL Plus (GE Healthcare). Panels (f), (g), and (h) are representative immunoblots from experiments repeated four times. An SDS-PAGE gel of the protein samples used in these assays is provided in Supplemental Figure S4.

Table 1NMR restraints and structural statistics for the GABPa^{35–121} ensemble

Summary of restraints		
NOEs		
Unambiguous		2745
Ambiguous		923
Total		3668
Dihedral angles		
ϕ, ψ, χ^1		51, 51, 34
Residual dipolar couplings ($^1\text{H}^{\text{N}} - ^{15}\text{N}$)		52
Deviation from restraints		
NOEs (Å)		0.027±0.001
Dihedral angles (deg.)		0.85±0.06
Residual dipolar couplings (Hz)		1.81±0.02
Deviation from idealized geometry		
Bond lengths (Å)		0.005±0.000
Bond angles (deg.)		0.73±0.01
Improper angles (deg.)		2.29±0.08
Residues in allowed region of Ramachandran plot		98.1%
Mean energies ^a , kcal·mol ⁻¹		
E _{vdw}		-168±20
E _{bonds}		35±2
E _{angles}		487±4
E _{impr}		138±11
E _{NOE}		121±7
E _{cdih}		7±1
E _{sani}		171±6
Residual dipolar couplings ^c		
Q value		0.186 ± 0.013
Q' value		0.194 ± 0.015
RMSD from average structure, Å		
	Backbone	Heavy atoms
Structured residues ^b	0.15±0.02	0.41±0.02
All residues	0.59±0.09	0.97±0.06

^a Final ARIA/CNS energies for van der Waals (vdw), bonds, angles, NOE restraints, dihedral restraints (cdih), and residual dipolar coupling restraints (sani).

^b Structured residues: 36–43, 48–59, 67–70, 73–74, 91–99, 107–115

^c $Q = \text{rms}(\text{D}^{\text{obs}} - \text{D}^{\text{calc}}) / \text{rms}(\text{D}^{\text{obs}})$ and $Q' = \text{rms}(\text{D}^{\text{obs}} - \text{D}^{\text{calc}}) / (\text{D}_a^2[4 + 3R^2]/5)^{1/2}$, as defined in ref ⁶⁰.