

Helical conformation of the AD1 peptide in the AML1-ETO–E-protein complex

Horea Porumb *

Université Paris 13, Bobigny, France

and

LBPA, UMR CNRS 8113, ENS de Cachan, Cachan, France

Abstract. The AML1-ETO fusion protein is responsible for 15% of the acute myeloid leukaemias due to its interference with transcription activation by E-proteins. The eTAFH region of the AML1-ETO competes for the AD1 domain of E-protein, thereby preventing wild-type transcription activation. The structural details concerning the eTAFH-bound AD1 domain are not known. We have previously shown that the eTAFH–AD1 interaction is strong ($K_d = 28$ nM, H. Porumb, *Spectroscopy* **22** (2008), 251–260). The secondary structure prediction algorithms are undecided as to the conformation of bound AD1 peptide. Here we demonstrate by circular dichroism that the bound AD1 peptide is fully helical. This will facilitate modeling of the interaction and launches a challenge as to using synthetic peptides to out compete the eTAFH–E-protein interaction.

Keywords: Circular dichroism, difference spectrum, dissociation constants, fusion protein, helical conformation, leukemia, peptides, secondary structure predictions, transcription activation, transcription repression

Abbreviations

bHLH: Basic helix–loop–helix

CD: Circular dichroism

CBF: Coactivator protein cyclic AMP response element-binding protein

ETO: ‘Eight–twenty-one’ protein

K_d : Dissociation constant

NHR1: (*Drosophila*) nervy-homology region 1 domain

TAFH: TATA-binding protein (TBP)-associated factor (TAF) homology

TFIID: Transcription factor IID

1. Introduction

E-proteins are a family of DNA-binding transcription factors involved in regulating cell growth, differentiation, and apoptosis. In particular, they are active in B- and T-cell differentiation. E-proteins are members of the basic helix–loop–helix (bHLH) family, recognize the consensus binding site (E-box) CANNTG, and participate in regulating lineage-specific gene expression through the formation of heterodimers with other bHLH proteins [12].

*Corresponding author: Horea Porumb, LBPA, UMR CNRS 8113, ENS de Cachan, 61 Avenue du Président Wilson, 94235 Cachan, France. Tel.: +33 1 47 40 68 92; Fax: +33 1 47 40 76 71; E-mail: hporumb@hotmail.com.

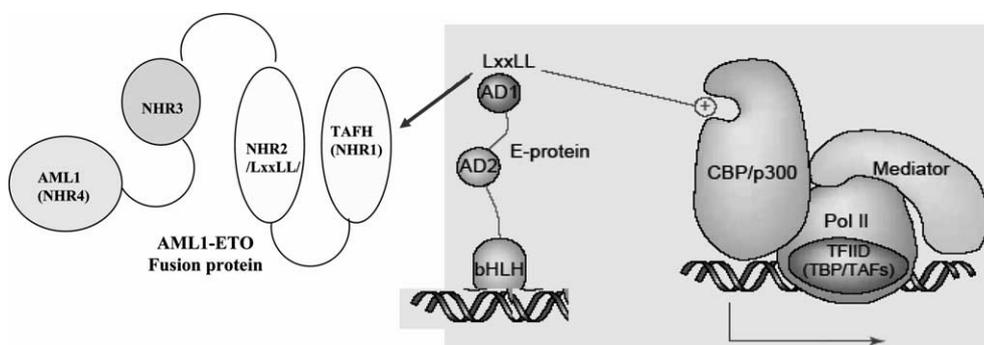


Fig. 1. Wild-type transcription activation and AML repression. E-protein (in the centre) binds DNA using a basic helix–loop–helix domain and recruits co-activators CBP/p300 via two activation domains, AD1 and AD2 (right). The eTAFH region of AML1-ETO fusion protein (left) binds to the LxxLL motif of the E-protein AD1 domain, thereby preventing CBP/p300 to bind and to subsequently activate transcription [14,25]. A possible intramolecular interaction between eTAFH and NHR2 regions of AML1-ETO was suggested by our previous work [17].

Normally, E-proteins interact with the KIX domain of the histone acetyltransferase p300/CREB binding protein [1], leading to the formation of a (wild-type) complex involving CBP/p300, which causes histone modifications that facilitate transcription (Fig. 1, redrawn after [14]). Incidentally, p300 is known to acetylate p53 in response to DNA damage [5].

On the other hand, AML1 (or RUNX1) is a DNA-binding transcription factor. The 8;21(q22;q22) reciprocal chromosomal translocation is one of the most common chromosome abnormalities identified in leukemia. The AML1-ETO fusion protein resulting from this translocation is responsible for 15% of the acute myeloid leukaemias [9,10], usually in the presence of a second class of genetic alterations [24]. It is thought that AML1-ETO recruits transcription activators such as the E-protein and inhibits the normal E-protein-mediated transactivation of transcription by physically displacing the wild-type coactivator proteins that usually bind to the latter [14,22,25].

The N-terminal domain of AML1-ETO, ETO-TAFH or eTAFH, is related in sequence to TAF4, a component of TFIID. The structure of eTAFH is comparable to that of yeast Rpb4 and of *Escherichia coli* sigma, this being the first TAF-related protein with structural similarity to the multisubunit RNA polymerases [22].

AML1-ETO has been shown to interact with E-protein via the eTAFH domain of the former [25], which binds to the E-protein (N-terminal) activation domain, AD1 (Fig. 1). The AD1 domain [15] contains 99 amino acids incorporating an LxxLL motif. Peptide AD1, representing a 17 amino acids fragment of AD1 (positions 12–28), is the object of the present investigation.

The 3D solution (NMR) structure of eTAFH [22] and of the AD1-bound state of the human ETO-eTAFH domain (eTAFH) are known [15], but the molecular details concerning the bound AD1 segment could not be determined. A helical conformation of AD1 [15] was inferred from sequence conservation and secondary structure predictions based on analogies [1,6]. In this work we have undertaken the elucidation by circular dichroism of the structure adopted by peptide AD1 upon binding to eTAFH, hoping that the information obtained will facilitate the modeling of the local structure of the eTAFH–AD1 complex, unclear at the time of writing. The possibility of targeting eTAFH by synthetic peptides or peptidomimetics is also discussed.

2. Results

The CD spectrum of the AD1 peptide, with its negative band at wavelengths below 200 nm, shows that free AD1 is unstructured in solution (Fig. 2A). Protein eTAFH alone shows an important helical content (of the order of 60%; typical negative bands at 208 nm and 222 nm), as indeed confirmed by its known structure [22].

The difference CD spectrum of AD1 (eTAFH–AD1 complex minus eTAFH), when expressed in terms of molar differential absorptivity per AD1 residue, displays the features at 208 and 222 nm characteristic for a helix, which shows that, when bound to eTAFH, peptide AD1 adopts a helical conformation. The molar differential absorptivity per residue at 222 nm, $\epsilon\Delta_{222} \sim 10 \text{ M}^{-1} \text{ cm}^{-1}$ is close to the largest theoretically possible value. This value indicates that the bound AD1 peptide has in fact become fully helical.

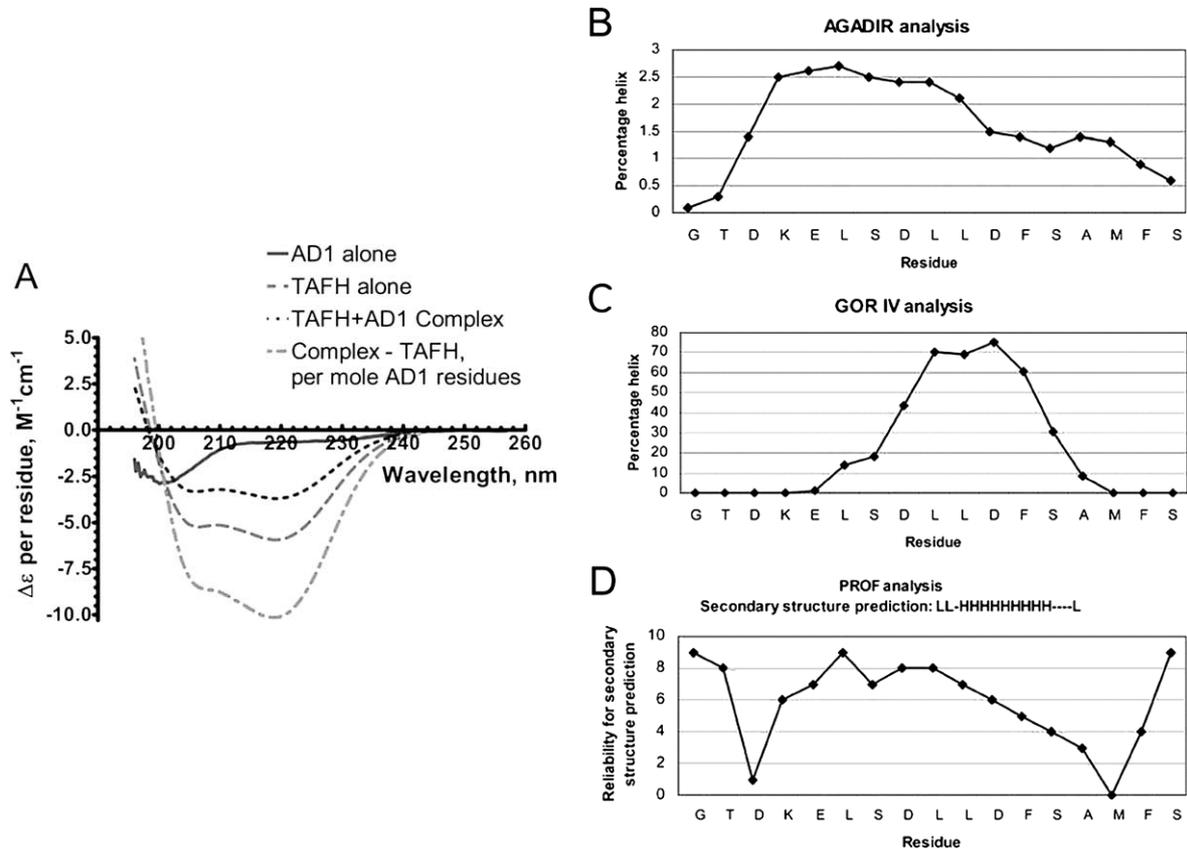


Fig. 2. Elucidation of the secondary structure of peptide AD1. (A) Circular dichroism difference spectral analysis of the interaction of peptide AD1 with protein eTAFH (the color code is indicated). (B) AGADIR prediction of the helix propensity of peptide AD1 in solution. (C) GOR IV prediction of the helicity of peptide AD1 within a protein environment. (D) PROF analysis of peptide AD1, predicting as secondary structure 71% helix (H) vs. 29% loop (L), disposed as: LL-HHHHHHHHHH---L, where “-” means undecided feature. The plotted curve indicates the reliability of the prediction at each residue (scale from 1 to 9).

3. Discussion

Our laboratory has been involved in developing spectroscopic tools for quantifying ligand–macromolecule interactions [16] and for fingerprinting (mostly) secondary structural features present in nucleic acids and proteins [7,8,18–20].

In a recent work, we looked by optical spectroscopic techniques (CD, ANS fluorescence and intrinsic fluorescence) at protein eTAFH and at peptide NHR2, issued from AML1-ETO, and at peptide AD1, issued from E-protein, respectively [17]. We described the nature of the oligomerisation process involving the NHR2 region of AML1-ETO. We have also shown that, separately from binding to AD1, the eTAFH domain of AML1-ETO might recognize the LxxLL motif of the NHR2 domain, the first likely interaction of eTAFH thus being intramolecular, within the AML1-ETO protein itself. Furthermore, a possibly accidental, weak interaction of NHR2 with AD1 was also revealed. We obtained dissociation constants for the self-association of peptide NHR2 ($K_d = 2.8\text{--}4\ \mu\text{M}$), for the binding of peptide NHR2 to eTAFH ($K_d = 1.28\ \text{nM}$) and to AD1 ($K_d = 240\ \text{nM}$) and, last but not least, for the eTAFH–AD1 interaction that concerns this work ($K_d = 28\ \text{nM}$).

The CD spectrum of peptide AD1 shows that the helical content of the free peptide is negligible (Fig. 2A). This is consistent with the AGADIR prediction (helical content below 3%, Fig. 2B), as indeed this algorithm predicts helix propensities for free polypeptide chains in solution. On the other hand, both secondary structure prediction algorithms GOR IV and PROF are rather “uncertain” in so far as the helical content of the peptide in a protein environment is concerned. GOR IV algorithm, which predicts the helical content of a peptide segment within the protein context, tells that peptide AD1 would be structured to a large, but not full extent, its helical content reaching 80% in the middle of the sequence (Fig. 2C). Likewise, with not very high reliability (Fig. 2D), the PROF algorithm predicts 71% helix vs. 29% loop.

The present CD results show that upon interacting with the eTAFH partner, the structuring effect on peptide AD1 is even greater than that predicted by GOR IV and PROF algorithms for the situation where the peptide were buried in its own protein environment. Indeed, the molar differential absorptivity per bound AD1 residue at 222 nm reached the limiting value of $10\ \text{M}^{-1}\ \text{cm}^{-1}$ [3,26] (Fig. 2A). We inferred that, when bound to eTAFH, peptide AD1 changed from random to a fully helical conformation. The small K_d value of the eTAFH–peptide AD1 complex (28 nM) indicates that the interaction is strong. It is remarkable that upon binding, in spite of the major unfavorable entropic term associated with the passage of peptide AD1 from coil to helix, the K_d of the interaction is still so low. We expect that, when in the full protein context, this interaction be even stronger.

The case described in this work recalls a well known situation, where a 68-residue polypeptide from yeast, called IA3, modulates through active-site occupation the activity of its target enzyme, aspartic proteinase. Alike with peptide AD1, the free IA3 peptide is essentially unstructured. Upon contact with the active site of the enzyme, residues 2–32 of peptide IA3 adopt an amphipathic alpha-helical conformation. The charged Lys-18 and Asp-22 on the otherwise hydrophobic face of the amphipathic helix are key selectivity-determining residues within the inhibitor, whereas Ala-213 is the dominant selectivity-governing feature in the proteinase. An extensive water-mediated hydrogen bond network involving Lys-18 and Asp-22 of the peptide and the catalytic Asp-32 and Tyr-75 of the active site locked the inhibitor inside the enzyme with an unprecedented specificity [13,23].

Since a three-dimensional structure of eTAFH is available, the structural information obtained from the present CD spectroscopic experiments should facilitate the modeling of the eTAFH–AD1 complex.

The modeling approach would also allow analysis of the inferred importance of the LxxLL motif for the interaction between AD1 and eTAFH. This development is outside the purpose of the present work.

Given the ubiquity of the LxxLL motif, deeper insights are needed (and will undoubtedly become available) in order to distinguish the factors ensuring the specificity of the interactions involved in wild-type transcription vs. AML repression (Fig. 1). The induction of full helicity into the LxxLL-containing AD1 peptide upon binding eTAFH, together with the understanding of the factors ensuring the specificity of this interaction, raise the exciting possibility of using synthetic helical LxxLL-containing peptides or peptidomimetics as competitors targeted to the AML1-ETO protein, aiming to prevent the formation of the AML1-ETO-E-protein complex.

4. Materials and methods

4.1. Proteins

The eTAFH protein of AML1-ETO used here was a 105 amino acid region of 11.873 kDa [2,15]. It was provided by M. Plevin and it belongs to the lot that served to the elucidation of its NMR structure (details on how the protein was obtained and purified have been described [15]). The protein was dissolved in the working buffer (15 mM Tris pH 7.5, 150 mM NaCl) at about 1.2 mM. Its concentration was subsequently measured spectrophotometrically, based on a molar absorptivity value of $1280 \text{ M}^{-1} \text{ cm}^{-1}$, calculated from the amino acid composition (the first two amino acids belong to the tag):

GSARQLSKLKRFLTTLQQFGNDISPEIGERVRTLVLGLVNSTLTIEEFHSKLQEAT

NFPLRPFVIPFLKANLPLLQRELLHAARLAKQNPAQYLAQHEQLLLDAS.

Peptide AD1 (residues 12–28 of the N-terminal activation domain, AD1, of E-Protein, GTDKELS DLLDFSAMFS) was a 17 amino acid, 1.876 kDa peptide obtained by chemical synthesis. As it had no chromophore suitable for UV absorption, it was put into solution by weight. The overall accuracy in handling the concentrations is expected to be better than $\pm 5\%$.

4.2. Circular dichroism of TAFH-bound AD1

CD spectra were recorded on a Jasco J-810 dichrograph, in a 0.1 mm demountable cell at 25°C in 0.1 nm steps in the wavelength range 190–260 nm, with 1 nm bandwidth and 4 s integration. Spectra of eTAFH protein at $30 \mu\text{M}$ ($3.15 \times 10^4 \text{ M}$ residue concentration) and of a stoichiometric amount of AD1 peptide, alone and complexed to eTAFH, were recorded. The spectrum of the eTAFH protein was then subtracted from that of the complex, on the assumption, supported by NMR data [15], that the protein did not change structure upon AD1 binding. The difference spectrum was then converted into units of molar differential absorptivity per AD1 residue. Recall that the α -helix content of peptides can be approximated by the relation: $P_\alpha = -[\Delta\epsilon_{222} \times 10]$ (P_α : percentage of α helix; $\Delta\epsilon_{222}$: differential molar absorptivity per residue at 222 nm) [15,26].

4.3. Secondary structure predictions

The algorithms for the secondary structure predictions AGADIR [11], GOR IV [4] and PROF [21] were available on the web respectively from: <http://www.embl-heidelberg.de/serices/serrona/agadir-start.html>, <http://pbil.univ-lyon1.fr/> (Pôle Bioinformatique Lyonnais, France) and <http://www.predictprotein.org/main.php>.

Acknowledgements

The author acknowledges with gratitude the collaboration with, and provision of excellent facilities by M. Ikura and M. Plevin, while being a visiting scientist in the Division of Signaling Biology, Ontario Cancer Institute, University Health Network, Toronto, ON, Canada. M. Plevin provided the author with the samples of protein eTAFH and of peptide AD1. The hospitality of Jasco, Easton, MD, is gratefully recognized.

References

- [1] R. Bayly, T. Murase, B.D. Hyndman, R. Savage, S. Nurmohamed, K. Munro, R. Casselman, S.P. Smith and D.P. LeBrun, Critical role for a single leucine residue in leukemia induction by E2A-PBX1, *Mol. Cell. Biol.* **26** (2006), 6442–6452.
- [2] J.N. Davis, L. McGhee and S. Meyers, The ETO (MTG8) gene family, *Gene* **303** (2003), 1–10.
- [3] G.D. Fasman, *Circular Dichroism and the Conformational Analysis of Biomolecules*, Plenum Press, New York, 1996.
- [4] J. Garnier, J.F. Gibrat and B. Robson, GOR method for predicting protein secondary structure from amino acid sequence, *Methods Enzymol.* **266** (1996), 540–553.
- [5] S.R. Grossman, p300/CBP/p53 interaction and regulation of the p53 response, *Eur. J. Biochem.* **268** (2001), 2773–2778.
- [6] Y. Liu, M.D. Cheney, J.J. Gaudet, M. Chruszcz, S.M. Lukasik, D. Sugiyama, J. Lary, J. Cole, Z. Dauter, W. Minor, N.A. Speck and J.H. Bushweller, The tetramer structure of the Nervy homology two domain, NHR2, is critical for AML1/ETO's activity, *Cancer Cell* **9** (2006), 249–260.
- [7] R.G. Maroun, S. Gayet, M.S. Benleulmi, H. Porumb, L. Zargarian, H. Merad, H. Leh, J.F. Mouscadet, F. Troalen and S. Fermandjian, Peptide inhibitors of HIV-1 integrase dissociate the enzyme oligomers, *Biochemistry* **40** (2001), 13840–13848.
- [8] H. Merad, H. Porumb, L. Zargarian, B. Rene, Z. Hobaika, R.G. Maroun, O. Mauffret and S. Fermandjian, An unusual helix turn helix motif in the catalytic core of HIV-1 integrase binds viral DNA and LEDGF, *PLoS One* **4** (2009), e4081.
- [9] K. Mrozek and C.D. Bloomfield, Clinical significance of the most common chromosome translocations in adult acute myeloid leukemia, *J. Natl. Cancer Inst. Monogr.* **39** (2008), 52–57.
- [10] K. Mrozek, G. Marcucci, P. Paschka and C.D. Bloomfield, Advances in molecular genetics and treatment of core-binding factor acute myeloid leukemia, *Curr. Opin. Oncol.* **20** (2008), 711–718.
- [11] V. Munoz and L. Serrano, Development of the multiple sequence approximation within the AGADIR model of alpha-helix formation: comparison with Zimm–Bragg and Lifson–Roig formalisms, *Biopolymers* **41** (1997), 495–509.
- [12] C. Murre, Helix–loop–helix proteins and lymphocyte development, *Nat. Immunol.* **6** (2005), 1079–1086.
- [13] L.H. Phylip, W.E. Lees, B.G. Brownsey, D. Bur, B.M. Dunn, J.R. Winther, A. Gustchina, M. Li, T. Copeland, A. Wlodawer and J. Kay, The potency and specificity of the interaction between the IA3 inhibitor and its target aspartic proteinase from *Saccharomyces cerevisiae*, *J. Biol. Chem.* **276** (2001), 2023–2030.
- [14] M.J. Plevin, M.M. Mills and M. Ikura, The LxxLL motif: a multifunctional binding sequence in transcriptional regulation, *Trends Biochem. Sci.* **30** (2005), 66–69.
- [15] M.J. Plevin, J. Zhang, C. Guo, R.G. Roeder and M. Ikura, The acute myeloid leukemia fusion protein AML1-ETO targets E proteins via a paired amphipathic helix-like TBP-associated factor homology domain, *Proc. Natl. Acad. Sci. USA* **103** (2006), 10242–10247.
- [16] H. Porumb, The solution spectroscopy of drugs and the drug-nucleic acid interactions, *Prog. Biophys. Mol. Biol.* **34** (1978), 175–195.
- [17] H. Porumb, Optical spectroscopy investigation of peptides issued from the AML1-ETO-E-protein complex relevant to acute myeloid leukemia, *Spectroscopy* **22** (2008), 251–260.

- [18] H. Porumb, H. Gousset and E. Taillandier, Parallel and antiparallel triple helices with G, A-containing third strands, *Electrophoresis* **20** (1999), 511–513.
- [19] H. Porumb, M. Monnot and S. Femandjian, Circular dichroism signatures of features simultaneously present in structured guanine-rich oligonucleotides: a combined spectroscopic and electrophoretic approach, *Electrophoresis* **23** (2002), 1013–1020.
- [20] H. Porumb, L. Zargarian, H. Merad, R. Maroun, O. Mauffret, F. Troalen and S. Femandjian, Circular dichroism and fluorescence of a tyrosine side-chain residue monitors the concentration-dependent equilibrium between U-shaped and coiled-coil conformations of a peptide derived from the catalytic core of HIV-1 integrase, *Biochim. Biophys. Acta* **1699** (2004), 77–86.
- [21] B. Rost and C. Sander, Prediction of protein secondary structure at better than 70% accuracy, *J. Mol. Biol.* **232** (1993), 584–599.
- [22] Y. Wei, S. Liu, J. Lausen, C. Woodrell, S. Cho, N. Biris, N. Kobayashi, Y. Wei, S. Yokoyama and M.H. Werner, A TAF4-homology domain from the corepressor ETO is a docking platform for positive and negative regulators of transcription, *Nat. Struct. Mol. Biol.* **14** (2007), 653–661.
- [23] T.J. Winterburn, D.M. Wyatt, L.H. Phylip, D. Bur, R.J. Harrison, C. Berry and J. Kay, Key features determining the specificity of aspartic proteinase inhibition by the helix-forming IA3 polypeptide, *J. Biol. Chem.* **282** (2007), 6508–6516.
- [24] Y. Yuan, L. Zhou, T. Miyamoto, H. Iwasaki, N. Harakawa, C.J. Hetherington, S.A. Burel, E. Lagasse, I.L. Weissman, K. Akashi and D.E. Zhang, AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations, *Proc. Natl. Acad. Sci. USA* **98** (2001), 10398–10403.
- [25] J. Zhang, M. Kalkum, S. Yamamura, B.T. Chait and R.G. Roeder, E protein silencing by the leukemogenic AML1-ETO fusion protein, *Science* **305** (2004), 1286–1289.
- [26] L. Zhong and W.C. Johnson Jr., Environment affects amino acid preference for secondary structure, *Proc. Natl. Acad. Sci. USA* **89** (1992), 4462–4465.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

