19 Siede, W. *et al*. (1996) The *Saccharomyces cerevisiae* Ku autoantigen homologue affects radiosensitivity only in the absence of homologous recombination. *Genetics* 142, 91–102
20 Khanna, K.K. and Jackson, S.P. (2001) DNA double-strand breaks: signaling, repair and the cancer connection. *Nat. Genet.* 27, 247–254
21 Kurz, E.U. and Lees-Miller, S.P. (2004) DNA damage-induced activation of ATM and ATM-dependent signaling pathways. *DNA Repair (Amst.)* 3, 889–900
22 Fukushima, T. *et al*. (2001) Genetic analysis of the DNA-dependent protein kinase reveals an inhibitory role of Ku in late S-G2 phase DNA double-strand break repair. *J. Biol. Chem.* 276, 44413–44418
23 Burma, S. and Chen, D.J. (2004) Role of DNA-PK in the cellular response to DNA double-strand breaks. *DNA Repair (Amst.)* 3, 909–918
24 Nelms, B.E. *et al*. (1998) *In situ* visualization of DNA double-strand break repair in human fibroblasts. *Science* 280, 590–592
25 Maser, R.S. *et al*. (1997) hMre11 and hRad50 nuclear foci are induced during the normal cellular response to DNA double-strand breaks. *Mol. Cell. Biol.* 17, 6087–6096
26 Takata, M. *et al*. (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J.* 17, 5497–5508

0968-0004/$ - see front matter
doi:10.1016/j.tibs.2004.12.004

The LxxLL motif: a multifunctional binding sequence in transcriptional regulation

Michael J. Plevin, Morgon M. Mills and Mitsuhiko Ikura

Division of Molecular and Structural Biology, Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, Toronto, Ontario M5G 2M9, Canada

LxxLL motifs participate in many protein–protein interactions associated with different aspects of transcriptional regulation. These motifs are present in many transcription factors and cofactors, mediating interactions that can activate or repress transcription. Several recently reported 3D structures of protein–LxxLL motif complexes and an intriguing novel interaction implicated in leukaemia have further highlighted the diversity and regulatory importance of this seemingly simple motif.

Introduction

Protein–protein interactions involved in cell signalling, cell adhesion and regulation of transcription or translation often use short peptide-recognition motifs. Examples include the polyproline α-helix, which interacts with Src-homology 3 (SH3) domains [1], and the array of peptide sequences that bind calmodulin [2]. The Leu-Xaa-Xaa-Leu-Leu (LxxLL) motif was originally observed in cofactor proteins that interact with hormone-activated nuclear receptors [3]. Functionally active examples of LxxLL motifs have also been documented in proteins that do not directly interact with nuclear receptors, including several transcription factors [4,5], calcium response element-binding protein (CREB)-binding protein (CBP) and p300 [6], and mediator subunits. Two recent 3D structures by Zor *et al.* [7] and Razeto *et al.* [8] indicate that there are differences in the binding mode of the LxxLL motif with nuclear-receptor- and non-nuclear-receptor-based complexes. Furthermore, a LxxLL motif has recently been implicated in a leukaemia-associated transcription-silencing mechanism that exhibits parallels with the role of LxxLL motifs in transcriptional regulation by activated nuclear receptors [9].

Corresponding author: Ikura, M. (mikura@uhnres.utoronto.ca).
Available online 24 December 2004

Nuclear receptor–cofactor interactions use LxxLL motifs

The LxxLL sequence was originally identified in proteins that bind the activation function-2 (AF-2) region of nuclear-receptor ligand-binding domains (LBDs) [3,6]. Subsequently, these conserved motifs were shown to have a key role in nuclear-receptor regulation with many nuclear-receptor-binding proteins, including the p160 family of co-activators (NCoA-1, 2 and 3), CBP/p300 and co-repressors such as receptor-interacting protein-140 (RIP140), which contains several functional examples.

Many proteins that interact with AF-2 regions have multiple LxxLL motifs. For example, NCoA proteins each contain three well-defined LxxLL motifs, also called nuclear-receptor boxes (Figure 1a). RIP140 also contains multiple motifs [3] (Figure 1a). Analysis of these LxxLL motifs has revealed a preference for a hydrophobic residue at position −1 [10] (Figure 1a). In most cases, motifs lacking a non-polar −1 residue bind with lower affinity to AF-2 regions. For example, the general co-activator CBP/p300 has three LxxLL motifs that interact with nuclear receptors, although at lower affinity compared with NCoA proteins [10]. Each motif in CBP/p300 is preceded by a polar residue (Figure 1b). A phage display analysis has demonstrated the role of residues that flank the LxxLL motif by identifying subclasses of motif that bind the oestrogen nuclear receptor with different affinities [11] (Figure 1a). Different ligands can also alter cofactor-binding specificities of nuclear receptors,

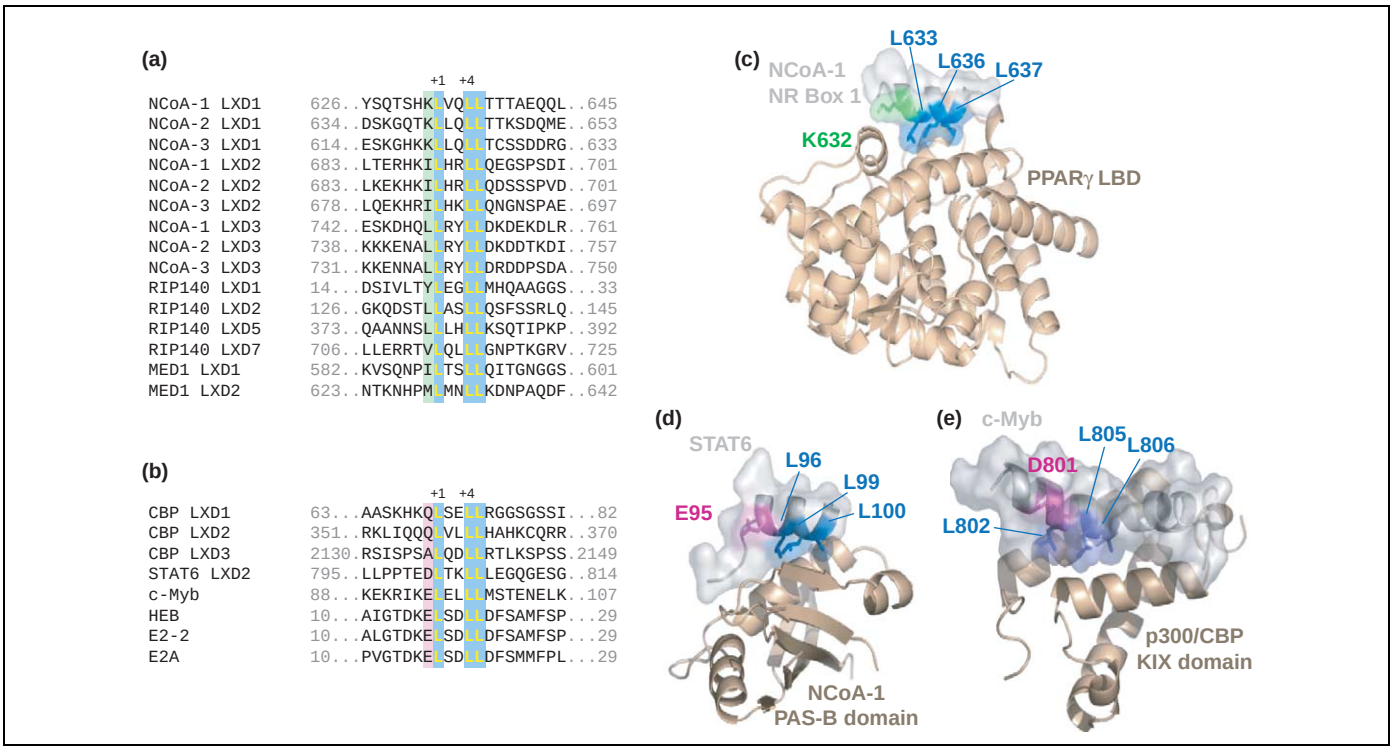

Figure 1. Importance of flanking residues in LxxLL-motif interactions. The chemical nature of the residue at −1 position can determine the binding orientation of the LxxLL motifs. Numbering of residues in LxxLL motifs designates the first conserved leucine as +1. **(a)** Amino acid sequence alignment of a selection of proteins with nuclear-receptor-binding LxxLL motifs reveals a residue with hydrophobic character in the −1 position. The +1, +4 and +5 leucine residues (blue) and the −1 residue (green) are highlighted. **(b)** Sequence alignment of a second subset of LxxLL motifs that are not directly implicated in nuclear-receptor binding shows the tendency for a small or polar residue at position −1 (pink). **(c)** The 3D structure of a subunit of PPARγ (gold ribbons) in complex with NCoA-1 nuclear-receptor box 1 (white surface) [12] shows that the four side-chain CH_2 groups of lysine at −1 partake in non-polar interactions with helix 3 of the PPARγ LBD. The −1 residue is similarly involved in the 3D structures of other LBD–co-activator peptide interactions. The conserved leucine residues and −1 position are represented as sticks, and labelled and coloured as described in (a). The 3D structures of PAS-B domain of NCoA-1 with STAT6 **(d)** [8] and the CBP/p300 KIX domain bound to c-Myb **(e)** [7] reveal that the hydrophilic −1 residue is exposed to the solvent and does not participate in the interface. The 3D structures in (d) and (e) highlight the conserved leucine residues (blue) and the hydrophilic residue at −1 (pink).

generating a complex array of competing cofactor proteins with differential nuclear-receptor selectivity that is dependent on the type of nuclear receptor, the agonist and the LxxLL motif involved (Figure 2a).

The elucidation of several 3D structures of apo-nuclear receptor LBDs and LBD–co-activator complexes has revealed much about the mechanism of ligand-activated co-activator recruitment and the role of LxxLL motifs [12,13]. Interaction of a LxxLL motif with AF-2 is agonist-dependent. The apo-LxxLL motif is unstructured but, on binding AF-2, forms a short, amphiphatic α helix. The termini of the co-activator α helix are locked in place by a 'charge clamp' involving two conserved AF-2 residues [12]. Leucine residues from the co-activator LxxLL motif line one face of the binding α helix and provide the mainstay of the van der Waals contact (Figure 1c). A co-crystal structure of a NCoA-1 peptide containing two LxxLL motifs bound to homodimeric peroxisome proliferator-activated receptor-γ (PPARγ) reveals that each motif interacts with a different subunit [12]. The same 3D structure shows that the −1 residue is buried in the interface supporting the preference for a hydrophobic residue at this position (Figure 1c). Although the −1 residue in the first motif of NCoA-1 is a lysine, it essentially acts as a hydrophobic residue because the four side-chain methylene groups participate in non-polar interactions with AF-2 (Figure 1c).

LxxLL motifs are not limited to nuclear-receptor signalling

Outside of nuclear-receptor signalling there are numerous examples protein–protein interactions involving LxxLL motifs. Again, many centre on interactions between activated transcription factors and co-activator proteins. Some co-activators are bi-functional, containing LxxLL motifs and also domains that bind them. For example, a PAS-B domain in NCoA-1 interacts with a LxxLL motif in the C-terminal transactivation domain of signal transducer and activator of transcription factor 6 (STAT6), which is a transcription factor involved in regulating the response to interleukin-4 [5]. Unlike the LxxLL motifs of NCoA cofactors, the −1 residue of the STAT6 motif is polar (Figure 1b). The recent 3D structure of this interaction shows PAS-B to adopt a mixed α/β structure with the α-helical STAT6 peptide binding in a shallow hydrophobic groove [8] (Figure 1d). The 3D structure also reveals several mechanistic differences in the LxxLL-motif-binding mode compared with previous, nuclear-receptor-based complexes. For example, PAS-B does not employ a 'charge-clamp' mechanism to tether the LxxLL α helix. Furthermore, the −1 residue does not contribute to the complex interface (Figure 1d). Interestingly, although many of the residues lining the LxxLL-binding site of PAS-B are conserved throughout the NCoA family, only NCoA-1 interacts with STAT6. Razeto *et al.* [8] attribute

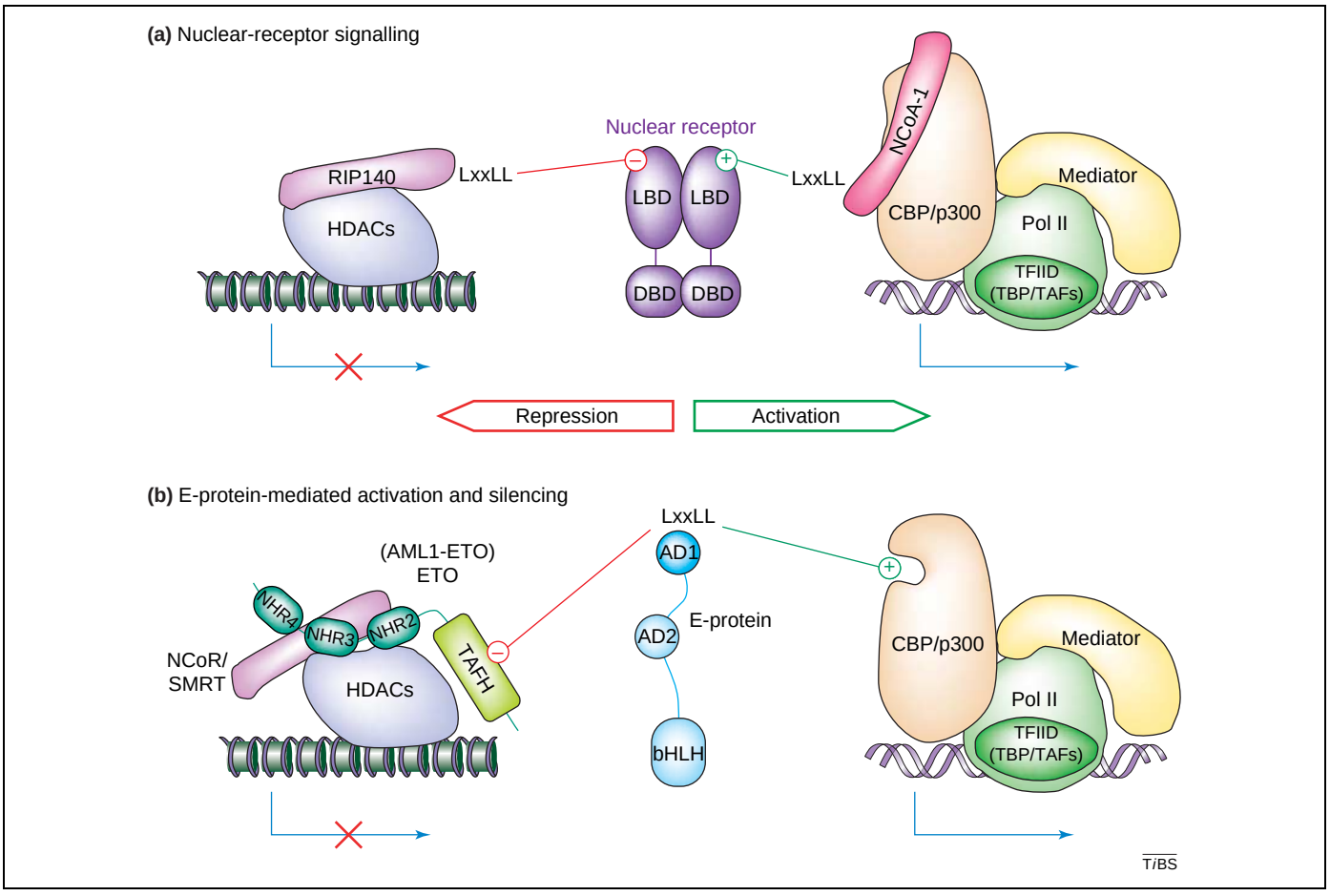

Figure 2. Cofactor-exchange mechanisms in transcriptional regulation. **(a)** Transcriptional regulation by nuclear receptors involves LxxLL-motif-mediated interactions between nuclear receptors and cofactor proteins (left). In the absence of an agonist, nuclear receptors interact with co-repressor proteins [e.g. RIP140, nuclear receptor co-repressor (NCoR) and silencing mediator of retinoid and thyroid hormone receptor (SMRT)] and, indirectly, with histone deacetylases (HDACs), all of which repress basal transcription by producing chromatin. Upon agonist binding, nuclear receptors recruit co-activator proteins that act as interaction platforms that enable the assembly of multi-component protein complexes. These complexes possess intrinsic histone acetyltransferase (HAT) activity and initiate transcription by relieving chromatin repression and recruiting the RNA polymerase II (Pol II) pre-initiation complex (right). **(b)** E-proteins activate transcription by recruiting CBP/p300 and other factors to E-box elements (right). E-proteins bind DNA using a basic helix–loop–helix (bHLH) domain and interact with co-activators, such as CBP/p300, via two activation domains, AD1 and AD2. The aberrant expression of ETO following t(8;21) silences E-protein-driven transactivation (left). The TAFH domain of ETO (or AML1-ETO) binds to the LxxLL motif in E-protein AD1, thereby preventing CBP/p300 binding. Sequences in the C-terminal domain of ETO recruit co-repressor and HDAC proteins culminating in the inhibition of transcription.

this to a 'surface complementarity' between NCoA-1 PAS-B and STAT6.

CBP/p300 also possesses the capacity to recognize LxxLL motifs. The KIX domain of CBP interacts with a LxxLL motif in the transactivation domain of c-Myb, which, like STAT6, contains a polar −1 residue that is not involved in the interaction (Figure 1e). The c-Myb motif binds to the same interface on KIX as the phosphorylated kinase-inducible domain (pKID) of CREB [7,14] (Figure 1e). Interestingly, no sequence similarity exists between pKID and c-Myb.

E-proteins are another family of transcription factors that interact with CBP/p300 via LxxLL motifs. E-proteins [i.e. E2A, HeLa E-box-binding (HEB) and E2-2] are involved in regulation of cell growth, differentiation and apoptosis [15]. In E-protein transactivation, CBP/p300 is recruited by a LxxLL motif in the N-terminal activation domain 1 (AD1) [9]. This functionally crucial interaction is inhibited in the presence of AML1-ETO, a fusion protein commonly associated with acute myeloid leukaemia (AML) [9]. Zhang and co-workers demonstrated that AML1-ETO specifically targets the E-protein LxxLL motif, thereby, directly preventing the binding of CBP/p300 [9]. AML1-ETO is a fusion protein that results from t(8;21) chromosomal translocations [16]. The protein product of t(8;21) contains the DNA-binding domain of AML1 fused to an unrelated, multidomain protein called eight twenty-one (ETO). The LxxLL-binding activity of AML1-ETO was localized to the TATA-binding protein (TBP)-associated factor (TAF) homology (TAFH; also called *nervy*-homology region 1, NHR1) domain from the fusion partner ETO [9]. The TAFH domain family – which has examples in several TAFs from transcription factor IID (TFIID), other ETO-related proteins and the *Drosophila* protein *nervy* [17,18,19] – had been functionally annotated, albeit loosely, as a protein–protein interaction module. The results of Roeder and colleagues suggest the TAFH domain functions as one of a growing number of non-nuclear-receptor LxxLL-binding domains [9].

Parallels between E-protein silencing and nuclear-receptor regulation

Mechanistic similarities, in addition to differences, exist between 'cofactor-exchange'-mediated silencing of

E-proteins [9] and regulation of nuclear-receptor signalling [20] (Figure 2). Both systems repress basal transcription by promoting histone and chromatin formation rather than by directly acting on transcription machinery. Activation or repression is achieved via interactions with co-activator and co-repressor proteins. These interactions are mediated by short peptide motifs. However, in nuclear-receptor signalling, these interactions are commonly transient and can be modulated by agonist or antagonist binding and subunit exchange. The interaction between ETO and E-proteins is not thought to be ligand-dependent, but might be regulated by mechanisms yet to be determined.

Concluding remarks and future perspectives

The LxxLL sequence is a protein-recognition motif widely used in transcriptional regulation. It is also worth noting that many non-functional examples of LxxLL sequences can be found in proteins that are involved in transcriptional regulation or other cellular processes. Like many short recognition sequences, the 'LxxLL nomenclature' fails to communicate the complexity of the system. Many factors, including α-helical propensity, availability for interaction and the composition of the flanking residues, contribute to determining the binding affinity and permit discrimination between potential binding partners.

Commonly, proteins that interact with LxxLL motifs are required to undergo structural changes, thus providing a mechanism for modulating binding. Indeed, both nuclear receptors and CBP/p300 bind multiple, unrelated target sequences at the same interfaces used for LxxLL-motif interaction, presumably, requiring some plasticity of the 3D structure. Many LxxLL-binding domains, including nuclear receptor [12], KIX [14] and TAFH [9], adopt – or are predicted to adopt – α-helical folds, which are well suited to execute such conformational adaptability. Interestingly, of the published 3D structures of LxxLL-containing complexes, only the PAS-B domain of NCoA-1 adopts a more structurally rigid mixed α/β fold, which perhaps contributes to the high specificity of the STAT6 LxxLL motif for NCoA-1 [8].

The presence of TAFH domains in subunits of TFIID raises the possibility of direct regulation of the pre-initiation RNA polymerase II complex by LxxLL-carrying proteins. Such a system has already been established for certain nuclear receptors that can interact with LxxLL motifs in Med1, a subunit of the mediator complex [21]. Currently, it remains to be established whether other members of the TAFH family are capable of interacting with LxxLL motifs.

The observations of Zhang and co-workers [9] raise the exciting possibility of using modified LxxLL-containing peptides as targeted inhibitors of the AML1-ETO–E-protein complex or other LxxLL interactions involved in disease. Potentially, such molecules would not only provide substantial insight into the functional importance of regulatory interactions involving these motifs, but might also represent promising therapeutic strategies. However, given the continuing uncertainty over the role of flanking sequences in determining binding specificities, further biochemical and structural characterization of LxxLL-motif interactions are required before such avenues are opened.

Acknowledgements

M.J.P. is a recipient of Canadian Institutes of Health Research (CIHR) post-doctoral research fellowship. M.I. is a Canadian Research Chair in Cancer Structural Biology. This work was supported by Canadian Institutes of Health Research (CIHR) grants to M.I.

References

1 Kuriyan, J. and Cowburn, D. (1997) Modular peptide recognition domains in eukaryotic signaling. *Annu. Rev. Biophys. Biomol. Struct.* 26, 259–288
2 Hoeflich, K.P. and Ikura, M. (2002) Calmodulin in action: diversity in target recognition and activation mechanisms. *Cell* 108, 739–742
3 Heery, D.M. *et al*. (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* 387, 733–736
4 Parker, D. *et al*. (1999) Role of secondary structure in discrimination between constitutive and inducible activators. *Mol. Cell. Biol.* 19, 5601–5607
5 Litterst, C.M. and Pfitzner, E. (2002) An LxxLL motif in the transactivation domain of STAT6 mediates recruitment of NCoA-1/SRC-1. *J. Biol. Chem.* 277, 36052–36060
6 Torchia, J. *et al*. (1997) The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. *Nature* 387, 677–684
7 Zor, T. *et al*. (2004) Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb. *J. Mol. Biol.* 337, 521–534
8 Razeto, A. *et al*. (2004) Structure of the NCoA-1/SRC-1 PAS-B domain bound to the LxxLL motif of the STAT6 transactivation domain. *J. Mol. Biol.* 336, 319–329
9 Zhang, J. *et al*. (2004) E protein silencing by the leukemogenic AML1-ETO fusion protein. *Science* 305, 1286–1289
10 Heery, D.M. *et al*. (2001) Core LxxLL motif sequences in CREB-binding protein, SRC1, and RIP140 define affinity and selectivity for steroid and retinoid receptors. *J. Biol. Chem.* 276, 6695–6702
11 Chang, C. *et al*. (1999) Dissection of the LxxLL nuclear receptor-coactivator interaction motif using combinatorial peptide libraries: discovery of peptide antagonists of estrogen receptors α and β. *Mol. Cell. Biol.* 19, 8226–8239
12 Nolte, R.T. *et al*. (1998) Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-γ. *Nature* 395, 137–143
13 Shiau, A.K. *et al*. (1998) The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* 95, 927–937
14 Radhakrishnan, I. *et al*. (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell* 91, 741–752
15 Quong, M.W. *et al*. (2002) E protein function in lymphocyte development. *Annu. Rev. Immunol.* 20, 301–322
16 Licht, J.D. (2001) AML1 and the AML1-ETO fusion protein in the pathogenesis of t(8;21) AML. *Oncogene* 20, 5660–5679
17 Kokubo, T. *et al*. (1993) The *Drosophila* 110-kDa transcription factor TFIID subunit directly interacts with the N-terminal region of the 230-kDa subunit. *Proc. Natl. Acad. Sci. U. S. A.* 90, 5896–5900
18 Kitabayashi, I. *et al*. (1998) The AML1-MTG8 leukemic fusion protein forms a complex with a novel member of the MTG8(ETO/CDR) family, MTGR1. *Mol. Cell. Biol.* 18, 846–858
19 Davis, J.N. *et al*. (1999) ETO-2, a new member of the ETO-family of nuclear proteins. *Oncogene* 18, 1375–1383
20 Torchia, J. *et al*. (1998) Co-activators and co-repressors in the integration of transcriptional responses. *Curr. Opin. Cell Biol.* 10, 373–383
21 Rachez, C. *et al*. (2000) The DRIP complex and SRC-1/p160 coactivators share similar nuclear receptor binding determinants but constitute functionally distinct complexes. *Mol. Cell. Biol.* 20, 2718–2726

0968-0004/$ - see front matter
doi:10.1016/j.tibs.2004.12.001