

Ewings family oncoproteins: drunk, disorderly and in search of partners

Kevin A. W. Lee¹

¹*Department of Biology, Hong Kong University of Science and Technology, Kowloon, Hong Kong, SAR China*

Cell Research (2007) **17**:286-288. doi: 10.1038/cr.2007.22; published online 10 April 2007

Aberrant chromosomal translocations involving the Ewings sarcoma (EWS) oncogene produce the EWS-fusion-protein family (EFPs, Figure 1A) that causes a wide variety of human cancers [1]. Extensive analysis of chromosomal breakpoints [2] has long established that, generally, EFPs contain at least the N-terminal 264 residues of EWS [the EWS-activation-domain (EAD)]. The C-terminal fusion partner is a cellular transcription factor that contributes, minimally, a sequence-specific DNA-binding domain which determines the tumor phenotype. Finally, EFPs are potent transcriptional activators but also interact with other proteins involved in different aspects of mRNA biogenesis, indicating that EFPs induce tumorigenesis by perturbing gene expression.

For many reasons the mechanism of EAD action has remained quite mysterious. One major barrier to progress is presented by the primary structure of the EAD (Figure 1B) comprising ~250 residues and (almost exclusively) 30 copies of a degenerate hexapeptide repeat (DHR, consensus SYGQQS) with the Tyr being absolutely and the first Gln being strongly conserved (Figure 1B). Because the intact EAD is required for full activity and considering its extended and repetitive sequence, it had not been feasible (until our study) to subject the EAD to systematic mutagenesis and to assess EAD structure/function relationships.

The advance in technology (total gene synthesis) highlighted here [3] enabled us to overcome a long standing struggle but a brief comment on the history is fitting. We previously attempted to study the EAD by using various tricks. We were able to show that multiple cis-linked

EAD sub-regions (~40 residues long) could create strong transcriptional activation domains [4] and that similarly small regions could synergise very effectively, in trans, on a promoter containing multiple activator binding sites [4]. The above findings provided evidence that the EAD harbors highly reiterated and flexible functional elements. However, whether the above properties of minimal synthetic activators could be extrapolated to the native EAD or to the biological functions of EFPs was a question that remained up in the air.

Technology comes to the rescue

The advent of commercial gene synthesis enabled us to alter the entire EAD primary sequence at will and determine the functional consequences. We examined transcriptional activation by EWS/ATF1 and cellular transformation by EWS/Flt1 and made several notable findings in both cases [3]. First, multiple Tyr residues are specifically required. Second, an EAD in which every Tyr (a total of 38) was replaced by Phe retained function, establishing that an aromatic side chain is sufficient. Third, sequence inversions between adjacent Tyr residues also retained function indicating that overall sequence composition (and not the DHR) confers EAD activity. In summary, the particular amino acid composition of the EAD creates an enabling structure (due to its enrichment of Pro, Ser, Gln and Gly) with several critical Tyr residues dispersed in a polar/neutral environment favoring hydrogen bonding interactions and flexibility (Figure 1C).

Besides scrutinising function, how could we probe EAD structure more directly? Another major hurdle was presented by the fact that (consistent with the functional data) native EAD structure is largely unfolded and is not

Correspondence: Kevin A. W. Lee
Tel: +852 2358 8636; Fax: +852 2358 1559
E-mail: bokaw@ust.hk

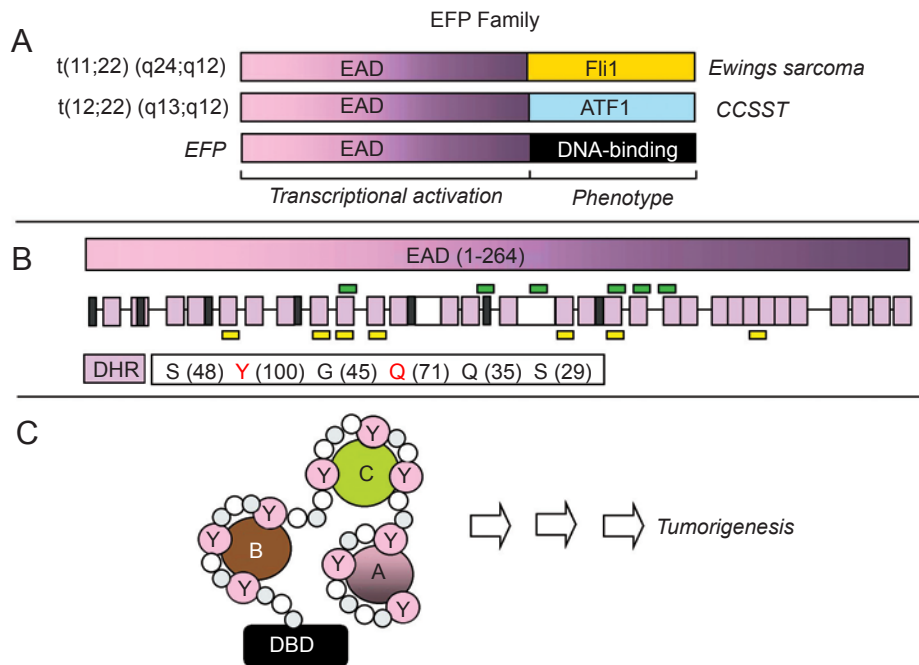


Figure 1 (A) Structure of EFPs. EWS/Fli1 and EWS/ATF1 are associated with Ewings sarcoma and clear cell sarcoma of soft tissue (CCSST) respectively. The extended EFP family (and associated cancers) includes EWS/WT1 (desmoplastic small round cell tumor), EWS/CHOP (myxoid liposarcoma), EWS/CHN (myxoid chondrosarcoma) and EWS/ZSG (small round cell tumor). All EFPs contain at least EWS residues 1-264, referred to as the EWS-activation-domain (EAD) and EWS fusion partners are all transcription factors that contribute a sequence-specific DNA-binding domain to the EFP. **(B)** Primary structure of the EAD. There are multiple degenerate hexapeptide repeats (DHRs, purple boxes) with the consensus sequence SYGQQS and 7 other Tyr residues (dark grey). Several SH2 binding sites (YxxP, yellow boxes) and SH3 binding sites (PxxP, green boxes) are indicated. Spaces between DHRs are generally only a few residues except in two cases (open white boxes) of 12 and 25 residues. The sequence degeneracy of DHRs present in the EAD is indicated, with numbers in parenthesis being % occurrence; absolutely conserved Tyr (position 2) and well conserved Gln (position 4) are highlighted in red. **(C)** A speculative model, suggested by our study, shows the EAD as a flexible string-like structure with several Tyr residues (pink circles) making contact with many different proteins (A-C) as discussed in the text.

amenable to analysis by classical methods such as crystallography or NMR. Very briefly the answer was to exploit a computational method [predictor of natural disordered regions (PONDR)] for the study of intrinsically disordered proteins (IDPs) [5]. PONDR typically predicts overall order/disorder and propensity for target-induced localised order, referred to as molecular recognition fragments/features (MoRFs). By employing a particularly sensitive PONDR algorithm (VL3, access provided by Molecular Kinetics, Inc. <http://www.pondr.com>) we established a robust correlation between EAD down-mutations and disorder propensity. Thus a combination of functional and computational approaches provided a robust analysis of a highly extended, repetitive and disordered protein region (the EAD).

A likely mechanism of EFP action

What are the mechanistic insights gained from our study? For the functions of EWS/ATF1 and EWS/Fli1 examined, the EAD works effectively in the absence of Tyr phosphorylation. Given the numerous Tyr kinase interaction motifs in the EAD (Figure 1B) it is not surprising, however, that Tyr phosphorylation reportedly acts on some EFPs [6]. Thus it must be stressed that our findings [3] do not preclude effects of Tyr phosphorylation on EFPs (including EWS/Fli1 and EWS/ATF1) under circumstances that we did not examine. What is equally clear, nonetheless, is that some EFP functions are highly effective under conditions in which the EAD is not (cannot be) phosphorylated on Tyr. It will be of great interest to explore how broadly

Tyr phosphorylation may impact the wider EFP family and associated malignancies.

It remains to be established to what extent the biological effect of EFPs is conferred by their trans-activation capacity. I therefore speculate on the latter without implying, necessarily, an impact on oncogenic function. The EAD resembles normal transcriptional activation domains (TADs) in that it lacks a native structure [7] (PONDR predicts that TADs are 73-94% disordered) and is highly flexible. Thus TADs may provide instructive models for the EAD and such models [8] invoke multiple weak contacts or a “molecular Velcro” involving TADs and their partners. Considering such a framework, we speculated [3] that the length and malleability of the EAD may account for potent trans-activation via recruitment of multiple different proteins (Figure 1C). A precedent for the above scenario is provided by HMGA (an architectural transcription factor [TF]) which is highly unstructured [7] and binds a vast array of proteins, many of which are themselves TFs. The EAD is known to contact a number of TFs (including the co-activator CBP, multiple TAFs and the Pol II sub-units rpb7/rpb5) but the functional significance of these contacts has not been clarified in mammalian cells. The human proteome interaction network also places the EAD (by association with EWS) at the hub of a disease-associated protein network via interactions with ~100 different proteins [9]. The malleable nature of the EAD most likely allows EFPs to mesh with a large number of proteins, many of which could participate in transcriptional and/or pathogenic effects of EFPs.

The overall sequence of EWS (including the EAD) is highly similar in mammals and frogs, with 70% dispersed conservation including most Tyr residues, phosphorylation sites and SH2/SH3 binding motifs (Figure 1B). Since the above EAD features presumably play important roles for normal EWS and yet are (at least in certain respects) dispensable for EFPs, there appears to be a very fine distinction, with respect to the EAD *per se*, between EWS and EFP function. The fact that EFPs are implicated in tumor

maintenance [10] indicates that such a distinction might ultimately allow the development of EFP-specific inhibitors with therapeutic potential. Unique EFP-specific interactions with other cellular proteins [10] and the absolute tumor specificity of EFPs, augur well for the above possibility. Future progress will rest on identification of physiologically relevant EAD-binding proteins (Figure 1C, proteins A-C) and the EAD mutants described [3] should serve as invaluable tools in this quest.

References

1. Arvand A, Denny CT. Biology of EWS/ETS fusions in Ewing's family tumors. *Oncogene* 2001; **20**:5747-54.
2. Zucman J, Melot T, Desmaze C, *et al.* Combinatorial generation of variable fusion proteins in the Ewing family of tumours. *EMBO J* 1993; **12**:4481-87.
3. Ng KP, Potikyan G, Savene RO, Denny CT, Uversky VN, Lee KAW. Multiple aromatic side chains within a disordered structure are critical for transcription and transforming activity of EWS family oncoproteins. *Proc Natl Acad Sci USA* 2007; **104**:479-84.
4. Feng L, Lee KAW. A repetitive element containing a critical tyrosine residue is required for transcriptional activation by the EWS/ATF1 oncogene. *Oncogene* 2001; **20**:4161-68.
5. Uversky VN, Oldfield CJ, Dunker AK. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J Mol Recognit* 2005; **18**:343-84.
6. Kim J, Lee JM, Branton PE, Pelletier J. Modification of EWS/WT1 functional properties by phosphorylation. *Proc Natl Acad Sci USA* 1999; **96**:14300-05.
7. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. Intrinsic disorder in transcription factors. *Biochemistry* 2006; **45**:6873-88.
8. Melcher K. The strength of acidic activation domains correlates with their affinity for both transcriptional and non-transcriptional proteins. *J Mol Biol* 2000; **301**:1097-112.
9. Rual JF, Venkatesan K, Hao T, *et al.* Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 2005; **437**:1173-78.
10. Kovar H, Aryee D, Zoubek A. The Ewing family of tumors and the search for the Achilles' heel. *Curr Opin Oncol* 1999; **11**:275-84.