Worm 3:4, e991240; October/November/December 2014; Published with license by Taylor & Francis Group, LLC

Splicing factor SUP-12 and the molecular complexity of apparent cooperativity

Cameron D Mackereth*

Inserm U869; University of Bordeaux; Institut Européen de Chimie et Biologie Pessac; France

Keywords: alternative splicing, ASD-1, *cis* element, *egl-15*, protein, RNA, structure, SUP-12, *trans* factors

Abbreviations: ITC, isothermal titration calorimetry; NMR, nuclear magnetic resonance; RRM, RNA recognition motif.

© Cameron D Mackereth

*Correspondence to: Cameron Mackereth; Email: c.mackereth@iecb.u-bordeaux.fr

Submitted: 11/08/2014

Accepted: 11/20/2014

http://dx.doi.org/10.4161/21624054.2014.991240 This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial License (http://creativecommons.org/ licenses/by-nc/3.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

The splicing factor SUP-12 from C. L elegans, in combination with either ASD-1 or FOX-1 from the Fox-1 (RBFOX) family, is required for generating a muscle-specific isoform of the fibroblast growth factor receptor EGL-15. Biophysical techniques have revealed the sequence preference for the RNA Recognition Motif (RRM) domain from SUP-12 as well as the structural details of the RNA-bound complex. Detailed genetics have identified a requisite need for the presence of both SUP-12 and ASD-1/FOX-1 to regulate the alternative splicing event, prompting speculation of a cooperative mechanism between these proteins on binding RNA. In contrast, the interplay between SUP-12 and ASD-1 suggests that although the RRM domains from each protein are in direct contact on the egl-15 pre-mRNA, there is no simple contribution of binding cooperativity. Evidence for an independent binding mechanism by SUP-12 and ASD-1 will be discussed, including a model in which both positive and negative contributions are balanced during complex assembly. The ability to monitor tissue-specific alternative splicing in live nematodes will continue to provide a powerful method to test in vivo models derived from mechanistic atomic-level investigation.

Interactions between cellular biomolecules such as proteins and nucleic acids are critical for life, and the precise regulation of these interactions in a time- and location-dependent manner allows for normal organism development. Alternative splicing can

contribute to this process through the production of development- and tissuespecific protein isoforms. The spliceosome generally interacts with target premRNA with the help of constitutive protein trans factors that recognize RNA cis elements in order to define the intron boundaries and dictate the series of exons in the mature mRNA. To regulate this process, additional splicing factors also bind to the pre-mRNA in order to enhance or inhibit specific 5'and 3' intron splice sites. The mechanisms by which the final splice patterns are established are in general poorly understood, but predominantly involve cooperative or competitive interactions with the constitutive splicing machinery to alter spliceosome recruitment, premRNA motif availability or the catalytic efficiency.

Several splicing factors have already been identified in C. elegans that display limited expression patterns and lead to cell type-specific changes in splicing regulation.¹ In many cases, multiple factors are linked to a specific splicing event, including both positive and negative regulators that together dictate the final pattern of exon retention in the mature mRNA.² Structures of RNA-binding domains from splicing proteins shed light on the basis for RNA motif preference and this information can explain the phenotypic effects of known genetic variations in the protein factors or RNA elements. The molecular details also help in the strategic design of mutants to selectively perturb the contribution of a splicing factor in a given splicing event. Less known are the atomic details that underlie the ability of these splicing factors to define a splicing pattern; auxiliary domains, competition and

cooperativity all likely play a role in the final outcome.

egl-15 as a model for muscle-specific alternative splicing

The gene egl-15 encodes the only known fibroblast growth factor receptor in C. elegans.³ The constitutive expression of egl-15 includes exon 5B and the resulting protein binds to the fibroblast growth factor LET-756 (Fig. 1).⁴ During development, there is a regulated switch to an alternate EGL-15 isoform that replaces exon 5B with the mutually exclusive exon 5A. This switch is possible due to the presence of the splicing factor SUP-12 and at least one member of the Fox-1 (RBFOX) family that includes ASD-1 and FOX-1.5 These proteins bind to RNA sequences upstream of exon 5B and cause exon skipping such that a downstream 3' splice site is used and exon 5A is retained in the mature mRNA. By replacing exon 5B with 5A, the EGL-15(5A) isoform displays an altered binding specificity and recognizes a different fibroblast growth factor, EGL-17.4 Muscle cells such as the sex myoblasts that produce EGL-15(5A) are attracted to a gradient of EGL-17,

guiding the dramatic migration of this cell pair from a posterior location to the required vulval position midway along the worm body.^{6,7} If this splice isoform is selectively abolished, there is improper myoblast guidance and the mutant worms display impaired egg-laying.

The molecular interplay between splicing factors SUP-12 and ASD-1 was recently revealed through the structural investigation of these proteins bound to a segment of *egl-15* pre-mRNA.^{8,9} These studies highlight the steps required to determine the atomic basis of the interaction between adjacent bound splicing factors, and in this case reveals a complex balance between both favorable and unfavourable contacts between SUP-12 and ASD-1.

Atomic details of RNA binding: divide and conquer

Both SUP-12 and ASD-1 (or FOX-1) are required for the alternative splicing of *egl-15* in muscle cells (Fig. 1).⁵ These proteins each contain a single RNA Recognition Motif (RRM) domain and bind to a specific RNA sequence motif: SUP-12 recognizes a core (G)-G-U-G-U-G motif,^{5,10}

whereas the Fox-1 (RBFOX) family is selective for a (U)-G-C-A-U-G sequence.¹¹ The absence of SUP-12 or both ASD-1/FOX-1 prevents the skipping of exon 5B, as do mutations in their corresponding RNA motifs in the *egl-15* premRNA.^{5,12}

A first step toward understanding the molecular basis of alternative splicing is to characterize at the atomic level the manner in which the trans factors bind to their cis element RNA sequences. A divide and conquer approach is generally used for these analyses based on the fact that most RNA-binding proteins harbor a domain that is clearly associated with binding RNA. This domain can typically be expressed in isolation from the full-length protein without affecting RNA-binding properties and these minimal systems are optimal for high-resolution studies aimed at defining the atomic details of ligand binding. Such investigation is important in order to understand the key protein residues involved in RNA-binding so that clever mutations can be designed that surgically eliminate aspects of RNA-binding in the full-length proteins without affecting other functions. In addition,



Figure 1. Alternative splicing of *egl-15* pre-mRNA. Mutually exclusive exons 5B and 5A result in either a constitutive EGL-15(5B) protein that binds LET-756 or the muscle-specific EGL-15(5A) that responds to EGL-17. *C. elegans* SUP-12 (UniProt # O45189/H2L051), ASD-1 (UniProt # G5EEW7/Q86G94) and FOX-1 (UniProt # Q10572) each contain a single RNA recognition motif (RRM) domain that interacts with the *egl-15* pre-mRNA upstream of exon 5B. In addition, the 3 proteins have regions enriched in alanine and glutamine (AQ-rich).

important elements of the RNA motifs are also revealed by the high-resolution structures, helping to define complete RNA motif patterns for improved bioinformatics target searches or to predict the precise effect on affinity when the sequence is altered.

The structure of the RRM domain from human Fox-1 bound to U-G-C-A-U-G-U RNA was determined by NMR spectroscopy.¹³ Numerous contacts between the protein and bases explain and dictate the RNA sequence specificity, and conservation of key residues indicates a similar mode of binding for C. elegans ASD-1 and FOX-1. For ASD-1, the binding details have also been confirmed by structural studies.9 The RRM domain from SUP-12 in complex with RNA has also been characterized at the structural level, bound to the RNA ligands G-G-U-G-U-G-C or G-U-G-U-G-C.^{8,9,14} Extensive mutagenesis on both the protein and RNA were used to identify critical residues involved in protein-RNA association and binding was further quantified by using the in vitro biophysical technique of isothermal titration calorimetry (ITC).¹⁵ As a result, precise affinity measurements (K_D) have been obtained for the binding of the wild-type SUP-12 RRM domain to short RNA ligands based on the parent egl-15 sequence, as well as for SUP-12 mutants originating from genetic screens or designed on the basis of structural studies.

SUP-12 affinity for RNA appears to be independent of ASD-1

In the egl-15 pre-mRNA, the Fox-1 and SUP-12 RNA motifs are directly adjacent, to the degree that a guanine is actually shared by both motifs (Fig. 1). This situation causes the RRM domains from the bound ASD-1 and SUP-12 to come into close contact and a solution structure of this arrangement has been determined.9 Given the requirement for both proteins in the splicing switch from exon 5B to 5A in muscle cells, and the close proximity of the RNA-binding domains from each protein on the egl-15 pre-mRNA, it is reasonable to imagine that simple binding cooperativity could aid in the assembly of ASD-1 and SUP-12 and provide a basis for target specificity. As will be shown

below, affinity measurements do not support this simple scenario.

Binding cooperativity with respect to RNA-binding proteins results in the increased affinity of a protein for RNA due to the prior binding of an additional component. The recent description of Drosophila Sxl and Unr binding to the msl2 mRNA provides an excellent example of this effect since the binding affinity of the first cold shock domain of Unr increases 1000-fold when the adjacent Sxl is also bound.¹⁶ In the egl-15 model system, cooperativity would manifest itself as the increased affinity of SUP-12 for RNA when ASD-1 is also bound to the same pre-RNA molecule. Based on the series of affinity studies reported by Amrane et al.,⁸ the isolated RRM domain from SUP-12 binds to its minimal RNA ligand with a dissociation constant (K_D) of 69 nM (Fig. 2A). To look for cooperativity, a measurement was then made with a longer piece of RNA containing the adjacent Fox-1 motif pre-bound with the RRM domain from ASD-1 (Fig. 2B). Surprisingly, the affinity of SUP-12 for RNA was unchanged from the measurement on the shorter ligand (68 nM compared to 69 nM) thus effectively ruling out a cooperative effect on binding. Although the binding conditions are different, the ITC studies from Kuwasako et al. also demonstrate a similar affinity of SUP-12 toward a minimal G-U-G-U-G-C motif (K_D of 27 nM) or with the longer U-G-C-A-U-G-G-U-G-U-G-C RNA and pre-bound ASD-1 (K_D of 42 nM).⁹

The apparent independence in RNAbinding by SUP-12 in the presence of ASD-1 suggests that an increased separation between their corresponding RNA motifs should not affect the binding behavior. The addition of 2 bases indeed has a minimal effect on affinity with a K_D of 94 nM (Fig. 2C) and NMR experiments demonstrate a significant mobility between the RRM domains on this ligand.⁸ Another consequence of binding independence is the prediction that SUP-12 constructs with mutations in critical RNA-binding residues will display a consistent affinity in the absence or presence of ASD-1. A previous genetic screen identified a SUP-12 mutant, yb1010, where Gly113 is replaced with glutamate.¹² The

G113E mutant has reduced binding to the short ligand and, as predicted, displays a similarly reduced affinity in the presence of ASD-1 and the longer RNA (**Figs. 2D**, E).⁸ As a second example, the Y44A mutation has further reduced RNA binding, but has comparable affinity without or with the adjacent ASD-1 (**Figs. 2F,G**).

A balance of favorable and unfavourable elements

The regions of the RNA-bound ASD-1 and SUP-12 that are in contact with each other have been identified from NMR spectroscopy studies. When bound to egl-15 pre-mRNA, a negatively charged region on the ASD-1 RRM domain associates with a positively charged basic region on SUP-12 (Fig. 3).^{8,9} This charge complementarity is a favorable arrangement and should in theory stabilize an interaction between ASD-1 and SUP-12. Accordingly, a mutation to disrupt one of the charged surfaces, for example the double mutation D128A/E130A on ASD-1, results in a correspondingly lower affinity the adjacent binding SUP-12 for (Fig. 2H). Other mutations that disrupt the charge complementarity (such as E103R on ASD-1, or R103E and R103M on SUP-12) produce a similar decrease in binding.^{8,9} It therefore appears that the charged surfaces indeed contribute an expected stabilizing force within the complex.

Why then is there no change in affinity for SUP-12 binding RNA in the absence or presence of ASD-1? The most likely explanation is that there is in fact a concurrent destabilizing effect of having ASD-1 and SUP-12 binding cis elements in such a close proximity. This unfavourable aspect can perhaps be seen in the small but consistent reduction in affinity for most of the comparisons upon binding adjacent to ASD-1 (as evident for example in Fig. 2 for the G113E or Y44A SUP-12 mutants or with other mutants in ref.⁸). It may be either important for biological function or merely a coincidence that these 2 opposing forces cancel one another nearly perfectly upon simultaneous binding of the ASD-1 and SUP-12 RRM domains to the egl-15 pre-mRNA.



urements indicate that the affinity is either significantly reduced or too complicated to fit to a simple one-site binding model.^{8,9} Analysis of the sequence reveals a potential stemloop structure, and indeed base-pairing for this RNA is evident by NMR spectroscopy and a melting temperature (T_m) of 32°C has been determined by thermal denaturation monitored by UV spectroscopy.⁸ It is therefore apparent that a competing stem-loop structure complicates or hinders the binding of the first RRM domain toward this longer RNA ligand in vitro. Consistent with this hypothesis, mutation of the RNA sequence to prevent stemloop formation restores high affinity binding by the SUP-12 RRM domain (Fig. 2J). The binding of an RRM domain to the ligand also destroys the stem-loop structure, which is the reason why this complication is not observed when for example ASD-1 RRM domain is pre-bound to the RNA before addition of SUP-12 (Fig. 2B). Whether or not a competition between the stem-loop structure and splicing factor exists in vivo has not been determined. However, the moderate T_m would ensure transient accessibility by SUP-12 or ASD-1/FOX-1, and such a structure would interesting present an mechanism to limit nonspecific binding by other single-stranded RNA-bind-

with the tandem cis ele-

ments (Fig. 2I). ITC meas-

Figure 2. SUP-12 binding to RNA in the absence or presence of the ASD-1 RRM domain. The panels illustrate different ligands (left) to which wild-type or mutant SUP-12 RRM domains are added and are described in the text. Affinity measurements (K_D) by isothermal titration calorimetry (ITC) are from Amrane et al.⁸

A role for RNA?

There is one experimental condition in which the affinity measurements do not

correspond to the above trends: in the absence of ASD-1, the RRM domain of SUP-12 binds poorly to the RNA ligand

ing proteins. It is also important to note that it was the apparently weak binding by SUP-12 (or ASD-1 or FOX-1) to long RNA ligands containing the stem-loop sequence that had led to a proposed molecular cooperativity. It appears that at least *in vitro* this RNA structure prevents simple interpretation of such ITC and electrophoretic mobility gel shift experiments.

The role of auxiliary domains

The majority of the above analyses have used the isolated recombinant RRM domains from the ASD-1, FOX-1 and SUP-12 splicing factors. If this were solely the case, then reasonable caution would be warranted in translating this information to the behavior of not just the full-length proteins but also within live nematodes. However, the use of fluorescent splicing reporter mini-genes has facilitated the transition from in vitro data to in vivo observations at a cell or tissue resolution. In the regulation of egl-15 alternative splicing, a 2-color reporter had been used to identify mutants with impaired splicing patterns.^{5,12} To test the relevance of the invitro findings, RNA sequence mutations based on the biophysical data can be used to form clear predictions on the effect of alternative splicing. The observation of a qualitative and quantitative loss of exon 5B skipping in this live worm assay with full-length proteins reassuringly correlates well with the effects on RNA binding observed for the isolated RRM domains by NMR spectroscopy and ITC.^{8,9}

The fact that the splicing effects with mutant egl-15 sequences can be largely explained by the predicted changes in simple protein-RNA affinity supports a minor role for other interactions between the ASD-1 and SUP-12 proteins. However, detailed in vivo studies with mutant proteins will be required to fully address this issue. In addition, protein regions of splicing factors outside of the RNA-binding domains can have key roles in biological function. Isolated RNA-binding domains in vivo predominantly retain RNA-binding characteristics, but at the same time seldom recapitulate the function of the fulllength proteins. The most common explanation in most cases is that the auxiliary domains harbor a specific biological function, such as a domain required to recruit components of the splicing



Figure 3. Balance between charge complementarity from the acidic region on ASD-1 and the basic region of SUP-12 with the steric-based instability due to the overlapping RNA motifs

machinery. SUP-12, ASD-1 and FOX-1 proteins all harbor extensive alanine and glutamine-rich (AQ-rich) regions (Fig. 1). These low-complexity regions may mediate critical protein-protein interactions or help with localization.¹⁷ At the cellular level it is therefore possible that the dependence on having both SUP-12 and a Fox-1 family protein may reside in presenting functional auxiliary domains on each protein that serve different but equally important biological roles. This functional cooperativity would be entirely independent of a requirement for cooperativity at the level of RNA binding. It is however possible that the alanine and glutamine-rich regions of each protein interact with each other and thus provide an auxiliary means to connect the 2 proteins and increase the overall affinity of the heterodimer for RNA. The correspondence between the in vitro and in vivo data argue against a significant change in RNA-binding properties by such an intermolecular tether. A detailed study is however hindered by the substantial in vitro insolubility of constructs containing the glutamine and alanine-rich regions.18

Concluding Remarks

Investigating the interplay between SUP-12 and ASD-1 on the *egl-15* pre-

mRNA provides an informative study of the molecular complexity that forms the basis of a functional link at the cellular level. In this case, disadvantageous elements of complex formation (likely due to the overlapping RNA motifs) are almost exactly balanced by the favorable interaction of charge complementarity. What then is the benefit of a net zero effect on affinity? It is possible that this mechanism leads to specificity not through increased affinity but by preventing the binding of other RNA-binding proteins that could, for example, otherwise interact with the GU-rich sequences (for example UNC-75 from the CELF family).¹⁹ Prior binding of the pre-mRNA by ASD-1 would prevent binding of any of these other GUbinding proteins, since only SUP-12 displays the proper shape and charge complementarity to bind the adjacent RNA motif without reduced affinity. To address this and other open questions will require studies such as the measurement of splicing factor occupancy on wild-type and mutant egl-15 pre-mRNA. Investigation into the possible interplay between SUP-12 and ASD-2 in the regulation of unc-60 alternative splicing would provide an informative comparison.²⁰ Co-regulation appears to be an increasingly common finding in C. elegans splicing² and additional complexity will likely be revealed from the careful study of other alternative splicing events at both the atomic and organism levels.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

I thank D. Dupuy and A. Innis for discussion and critical reading of the manuscript.

Funding

The work was supported by Inserm, Institut Européen de Chimie et Biologie and the Aquitaine region.

References

- Zahler AM. Pre-mRNA splicing and its regulation in Caenorhabditis elegans. WormBook 2012:1-21; PMID:22467343
- Barberan-Soler S, Medina P, Estella J, Williams J, Zahler AM. Co-regulation of alternative splicing by diverse splicing factors in Caenorhabditis elegans. Nucleic Acids Res 2011; 39:666-74; PMID:20805248; http:// dx.doi.org/10.1093/nar/gkq767
- Birnbaum D, Popovici C, Roubin R. A pair as a minimum: the two fibroblast growth factors of the nematode Caenorhabditis elegans. Dev Dyn 2005; 232:247-55; PMID:15614779; http://dx.doi.org/10.1002/dvdy.20219
- Goodman SJ, Branda CS, Robinson MK, Burdine RD, Stern MJ. Alternative splicing affecting a novel domain in the C. elegans EGL-15 FGF receptor confers functional specificity. Development 2003; 130:3757-66; PMID:12835392; http://dx.doi.org/10.1242/dev.00604
- 5. Kuroyanagi H, Ohno G, Mitani S, Hagiwara M. The Fox-1 family and SUP-12 coordinately regulate tissue-

specific alternative splicing in vivo. Mol Cell Biol 2007; 27:8612-21; PMID:17923701; http://dx.doi.org/ 10.1128/MCB.01508-07

- DeVore DL, Horvitz HR, Stern MJ. An FGF receptor signaling pathway is required for the normal cell migrations of the sex myoblasts in C. elegans hermaphrodites. Cell 1995; 83:611-20; PMID:7585964; http://dx.doi. org/10.1016/0092-8674(95)90101-9
- Lo TW, Branda CS, Huang P, Sasson IE, Goodman SJ, Stern MJ. Different isoforms of the C. elegans FGF receptor are required for attraction and repulsion of the migrating sex myoblasts. Dev Biol 2008; 318:268-75; PMID:18455716; http://dx.doi.org/10.1016/j.ydbio. 2008.03.026
- Amrane S, Rebora K, Zniber I, Dupuy D, Mackereth CD. Backbone-independent nucleic acid binding by splicing factor SUP-12 reveals key aspects of molecular recognition. Nat Commun 2014; 5:4595; PMID:25183497; http://dx.doi.org/10.1038/ncomms5595
- Kuwasako K, Takahashi M, Unzai S, Tsuda K, Yoshikawa S, He F, Kobayashi N, Guntert P, Shirouzu M, Ito T, et al. RBFOX and SUP-12 sandwich a G base to cooperatively regulate tissue-specific splicing. Nat Struct Mol Biol 2014; 21:778-86; PMID:25132178; http://dx.doi.org/10.1038/nsmb.2870
- Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS, Li X, Gueroussov S, Albu M, Zheng H, Yang A, et al. A compendium of RNA-binding motifs for decoding gene regulation. Nature 2013; 499:172-7; PMID:23846655; http://dx.doi.org/10.1038/nature12311
- Kuroyanagi H. Fox-1 family of RNA-binding proteins. Cell Mol Life Sci 2009; 66:3895-907; PMID:19688295; http://dx.doi.org/10.1007/s00018-009-0120-5
- Kuroyanagi H, Kobayashi T, Mitani S, Hagiwara M. Transgenic alternative-splicing reporters reveal tissue-specific expression profiles and regulation mechanisms in vivo. Nat Methods 2006; 3:909-15; PMID:17060915; http://dx.doi.org/10.1038/nmeth944
- Auweter SD, Fasan R, Reymond L, Underwood JG, Black DL, Pitsch S, Allain FH. Molecular basis of RNA recognition by the human alternative splicing factor

Fox-1. EMBO J 2006; 25:163-73; PMID:16362037; http://dx.doi.org/10.1038/sj.emboj.7600918

- Amrane S, Mackereth CD. Protein chemical shift assignments of the unbound and RNA-bound forms of the alternative splicing factor SUP-12 from C. elegans. Biomol NMR Assign 2014; 8:109-12; PMID:23334698; http://dx.doi.org/10.1007/s12104-013-9463-9
- Brown A. Analysis of cooperativity by isothermal titration calorimetry. Int J Mol Sci 2009; 10:3457-77; PMID:20111687; http://dx.doi.org/10.3390/ijms 10083457
- Hennig J, Militti C, Popowicz GM, Wang I, Sonntag M, Geerlof A, Gabel F, Gebauer F, Sattler M. Structural basis for the assembly of the SxI-Unr translation regulatory complex. Nature 2014; PMID:25209665
- Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, Mirzaei H, Goldsmith EJ, Longgood J, Pei J, et al. Cellfree formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell 2012; 149:753-67; PMID:22579281; http:// dx.doi.org/10.1016/j.cell.2012.04.017
- Anyanful A, Ono K, Johnsen RC, Ly H, Jensen V, Baillie DL, Ono S. The RNA-binding protein SUP-12 controls muscle-specific splicing of the ADF/cofilin premRNA in C. elegans. J Cell Biol 2004; 167:639-47; PMID:15545320; http://dx.doi.org/10.1083/jcb.2004 07085
- Kuroyanagi H, Watanabe Y, Suzuki Y, Hagiwara M. Position-dependent and neuron-specific splicing regulation by the CELF family RNA-binding protein UNC-75 in Caenorhabditis elegans. Nucleic Acids Res 2013; 41:4015-25; PMID:23416545; http://dx.doi.org/ 10.1093/nar/gkt097
- Ohno G, Ono K, Togo M, Watanabe Y, Ono S, Hagiwara M, Kuroyanagi H. Muscle-specific splicing factors ASD-2 and SUP-12 cooperatively switch alternative pre-mRNA processing patterns of the ADF/cofilin gene in Caenorhabditis elegans. PLoS Genet 2012; 8: e1002991; PMID:23071450