Biophysical Chemistry 159 (2011) 82–89

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Human U2B″ protein binding to snRNA stemloops

Sandra G. Williams, Kathleen B. Hall *

Department of Biochemistry and Molecular Biophysics, Washington University Medical School, St Louis, MO 63110, United States

ARTICLE INFO

Article history:
Received 3 April 2011
Received in revised form 9 May 2011
Accepted 9 May 2011
Available online 16 May 2011

Keywords:
RRM
RNA binding protein
U2B″
Protein:RNA thermodynamics

ABSTRACT

The human U2B″ protein is one of the unique proteins that comprise the U2 snRNP, but it is also a representative of the U1A/U2B″ protein family. In the U2 snRNP, it is bound to Stem-Loop IV (SLIV) of the U2 snRNA. We find that *in vitro* it binds not only to human SLIV, but also to Stem-Loop II (SLII) from human U1 snRNA and to *Drosophila* U2 snRNA SLIV. The thermodynamics of these binding interactions show a striking similarity, leading to the conclusion that U2B″ has a relaxed specificity for its RNA targets. The binding properties of U2B″ are distinct from those of human U1A and of *Drosophila* SNF, despite its high homology to those proteins, and so provide important new information on how this protein family has modulated its target preferences.

1. Introduction

The spliceosomal proteins U1A and U2B″ are highly conserved in eukaryotes, where they are components of the U1 and U2 snRNP, respectively. In the snRNPs, U1A protein is thought to bind exclusively to stem-loop II (SLII) of the U1 snRNA and U2B″ to stem-loop IV (SLIV) of the U2 snRNA. These RNA hairpins are highly conserved, and their loop sequences are very similar to each other. Although they are components of snRNPs, the roles of U1A and U2B″ in splicing remain unclear, and in fact, *in vitro* snRNP reconstitution in the absence of U1A has no effect on splicing [1]. Other experiments using mutations of the fly homologue, SNF, that exclude the protein from either the U1 or U2 snRNP resulted in relatively mild phenotypic consequences [2,3]. In contrast, knocking out both U1A and U2B″ in *C. elegans* is embryonic-lethal, as is the SNF knockout in *Drosophila* [4,5]. There is certainly a possibility that these proteins have alternative functions outside of the snRNPs.

U1A, U2B″, and SNF consist of two RNA recognition motifs (RRMs) connected by a variable, flexible linker. The RRM is the most commonly used RNA binding domain in eukaryotes {Maris et al., 2005, FEBS J, 272, 2118-31}, and can be identified by two amino acid sequences (RNP1 and RNP2) that are located in two of the four β strands on its β-sheet {Ghetti et al., 1989, FEBS Lett, 257, 373-6}. The canonical view of RNA–RRM interactions is that single-stranded RNA binds to the β-sheet surface. Favorable electrostatic interactions, hydrogen bonding, and stacking between RNA bases and aromatic residues located in the RNP motifs are regarded as the predominant determinants of RNA binding {Clery et al., 2008, Curr Opin Struct Biol, 18, 290-8}. U1A uses its N-terminal RRM to bind its *in vivo* U1 snRNA target, Stem-Loop II (SLII), with very high affinity and specificity [6,7]. There are crystal and solution structures of U1A RRM1 [8,9], and a co-crystal structure of U1A RRM1 bound to SLII [10] shows how the loop of the RNA hairpin is positioned on the surface of the β-sheet. A co-crystal of U2B″ RRM1 bound to both SLIV and U2A′ (an auxiliary protein) positions the RRM in the middle of the complex, with the RNA bound on the β-sheet surface and U2A′ wrapping around the opposite face of the RRM, predominantly making contacts with RRM α_1[11]. There is a solution structure of SNF RRM1 [12]. Not surprisingly, the structures of these three RRMs are similar to each other and for U1A, the structures of the free and bound proteins are also similar {Nagai et al., 1990, Nature, 348, 515-20}.

The U1A/U2B″ family of proteins provides a valuable opportunity to understand determinants of RNA:protein affinity and specificity. RRM1 of U1A is ~75% identical to the N-terminal RRMs of U2B″ and the *Drosophila* homologue, SNF. While SNF binds to both U1 snRNA SLII and U2 snRNA SLIV [13], there has been significant debate regarding how (and how well) U2B″ binds to its target, U2 snRNA SLIV. Some data report that U2B″ binds both SLII and SLIV [14] , and there are conflicting studies on whether the U2A′ protein, which is present in the U2 snRNP and which binds U2B″, is required for U2B″ to bind SLIV [15]. Regardless, the studies do suggest that RNA binding by U2B″ is much weaker than binding of SLII by U1A.

The co-crystal structure of U2B″:SLIV showed that many of the interactions between U2B″ and SLIV were also present in the earlier U1A:SLII cocrystal structure [11]. Since both the RNA loops contain identical 5′ sequences (human SLII: A1UUGCA6CUCC and human SLIV A1UUGCA6GUA9CC) and the RNP sequences of the proteins are identical, preservation of these contacts is not unexpected. Stacking of the nucleobases with Tyr10 in RNP2 (β1) and Phe53 in RNP1 (β3) (U2B″ numbering) occurs in both complexes. Despite the

* Corresponding author. Tel.: +1 314 362 4196.
E-mail address: kathleenhal@gmail.com (K.B. Hall).

0301-4622/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2011.05.010

phylogenetic conservation of A9 on the 3′ side of the SLIV RNA loop (and its absence from stem-loop II sequences), the U2B″ co-crystal showed no interactions between the protein and A9. Within the 3′ UA9CC sequence, U8 and C11 do pack against the VALKT amino acids of U2B″ β2 and Loop 3; those amino acids are frequently conserved in U2B″ proteins but are distinct from the corresponding residues of U1A. The authors noted that the loop amino acids R52 and T48 appeared to interact with the phosphate backbone of the RNA near its loop/stem junction [11]. It is worth noting that the RNA stem in the U2B″ co-crystal is distorted from normal A-form, and those 3′ ACC loop nucleotides stack on each other, perhaps due to crystal packing. Considering the similar patterns of interactions between U1A and U2B″, we might anticipate that the proteins have comparable RNA binding affinities and specificities. However, we find that U1A and U2B″ have different RNA binding preferences *in vitro* which could not be anticipated by the apparent similarities in their respective cocrystal structures.

The U1A:RNA interaction is characterized by complicated thermodynamics. More specifically, the U1A:SLII interaction has a large apparent heat capacity ($\Delta C_{P,obs}$) of −3 kcal/mol, and its enthalpy and entropy are both temperature-dependent [7,16]. Interpretation of $\Delta C_{P,obs}$ is made more difficult by the conformational transitions of both RNA and protein upon complex formation. It is reasonable to anticipate that the U2B″:RNA interaction will also involve conformational changes of RNA and protein.

Measurements of thermodynamic pairwise coupling in the U1A protein and in the complex helped to identify a network of amino acid sidechains that span the RNA binding surface, including the conserved Tyr13, Phe56, and Gln54 (numbered as per U1A) on the surface of the β-sheet, residues in Loop 3 and in the C-terminal tail of RRM1 {Kranz and Hall, 1998, J Mol Biol, 275, 465-81}[17]. While the Loop 3 sequences in the two proteins are quite different, the other amino acids that have been implicated in this network are conserved between U1A and U2B″, leading to our expectation that the RNA binding surface of U2B″ will also span the entire face of RRM1. A comprehensive thermodynamic study of U2B″:RNA interactions is essential to describe this protein's binding mechanism and elucidate how and why it differs from the other members of this protein family. Most significantly, this analysis will help us understand how RRMs with such similar sequences and structures have such different RNA binding properties.

In this study, we use full-length wild-type human U2B″ to assess the binding affinities of this protein for several RNA stemloops (hairpins). We also assess the salt and temperature dependence of these interactions. An important result that has implications for the protein's biological functions is that human U2B″ binds U1 snRNA SLII and U2 snRNA SLIV with almost equal affinity. In contrast, human U1A protein effectively binds only SLII. Within this protein family, binding affinities and specificities for RNA sequences have been exquisitely modulated, and biothermodynamics is the only way to compare them.

2. Results

The U2B″ protein has two RNA recognition motifs, separated by a 40 amino acid linker. This linker is much shorter than the corresponding linker of U1A. Sequences of the N-terminal RRMs of human U1A, U2B″, and *Drosophila* SNF are compared in Fig. 1, including some of the linker sequences. A structural depiction is shown that highlights the residues on the RNA binding surface that differ between U1A and U2B″. Differences in β2 are extensive, but the most significant difference with respect to RNA binding is in Loop 3, which contacts the RNA.

Loop 3 of U1A, (SRSLKMRG) is a site that contributes to RNA:protein specificity; in both U1A:SLII and U2B″:SLIV co-crystals, the protein Loop 3 protrudes into the RNA loop where it is juxtaposed with the nucleotides on the bottom of the RNA loop, and so splays the RNA open. The sequence of this important protein loop is different in U2B″ (LKTMKMRG) and *Drosophila* SNF (LKTLKMRG). For U1A, Loop 3 makes contacts with the RNA, including a hydrogen bond between the R47 amide and the backbone at the RNA loop-closing G. In the U2B″:SLIV co-crystal, Loop 3 contacts with the U–U appear to be minimal.

The RNA also needs to be considered in a discussion of U1A/U2B″ binding. In the cocrystal of U2B″ RRM1, Stemloop IV of U2 snRNA, and the U2A′ auxiliary protein, the RRM contacts the 5′ AUUGCAG sequence of the RNA loop. These interactions appear similar to those formed in the U1A:SLII complex, as expected based on the RNA sequence (AUUGCAC in SLII). One significant difference between SLII and SLIV is the loop-closing base pair, which in SLII is a C:G but in SLIV is a noncanonical U:U pair. In *Drosophila*, the U–U sequence has become U–G, and this difference appears to be important for recognition of the RNA by SNF {Williams and Hall, 2010, Biochemistry, 49, 4571-82}. The adjacent base pair in the stem is a G:C in human and fly SLIV (Fig. 2), and it may be that this is the effective loop-closing base pair. Here, we refer to the U–U as either the loop-closing base pair or as inserted nucleotides at the bottom of the RNA loop. The difference in the loop-closing basepair may well be important for discrimination between the RNAs by the proteins.

2.1. RNA Binding to U2B″

Stemloop IV, the binding site for U2B″, is located at the 3′ end of U2 snRNA. In vertebrates, the stem of this hairpin has eleven base pairs and an asymmetric internal A bulge in the middle of the stem (Fig. 2). In most vertebrates, the loop sequence is conserved: 5′gUA1UUG-C5AGUAC10CUc, where the loop-closing base pair is indicated in lower case. Much of this sequence is shared with U1 snRNA SLII: cA1UUGC5ACUCC10g. Notable differences include the 'inserted' U/U pair at the bottom of the SLIV loop, the C7-to-G7 substitution, and the longer SLIV loop. These phylogenetic differences could identify unique contacts with the corresponding protein, or indicate sites that are insensitive to mutation.

Binding of U2B″ to several RNA sequences suggests some features recognized by the protein. As summarized in Table 1, the human SLII and SLIV RNAs are bound with equal affinity by the protein at room temperature in 100 mM KCl, 1 mM $MgCl_2$, 10 mM sodium cacodylate pH 7.4. This rather surprising result indicates that the presence of the inserted U–U nucleotides does not limit protein binding. Indeed, the protein binds with equal affinity to *Drosophila* SLIV RNA with its U–G pair. In contrast, the human U1A protein does not detectably bind to the human SLIV RNA in nitrocellulose filter binding assays, but it is able to weakly bind to *Drosophila* SLIV ($5 \pm 3 \times 10^{-7}$ M in 100 mM KCl, 12 mM $MgCl_2$, 10 mM sodium cacodylate pH 7, 22°C). The U–U pair in SLIV appears to function as a discriminator for U1A, but not for U2B″, pointing to a fundamental difference in their mechanisms of RNA recognition. Additionally, a C7-to-G7 substitution in SLII leads to a 10-fold weaker affinity of the U1A:RNA interaction ($\Delta\Delta G^\circ =$ 1.4 kcal/mol) [6]. In contrast, while the binding affinity of U2B″ for SLII is significantly lower than the affinity of U1A for SLII, the protein tolerates the G7 in SLIV without a further loss of affinity ($\Delta\Delta G^\circ = 0$), indicating different binding mechanisms.

Although the stem of SLIV in U2 snRNA contains an internal asymmetric bulge, here we use perfect duplexes for the hairpins. We find that stems of length 6 and 9 base pairs are bound with equal affinity by the protein. In the co-crystal, the RNA stem was a perfect 6 base pair duplex that appeared to make contacts with α_1 of U2B″ [11]. We posit that this contact arose through crystal packing that led to a buckling of the base pairs that unwound the stem, together with a twist of the RNA loop.

A significant difference between SLII and SLIV is the insertion of A9 on the 3′ side of the loop. While interactions with the protein were not observed in the cocrystal structure, this A9 could be used as a point of specific contact for U2B″ recognition. To observe the structural