doi:10.1016/j.jmb.2007.04.017 J. Mol. Biol. (2007) 369, 1098–1112

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

RRM Proteins Interacting with the Cap Region of Topoisomerase I

Agata M. Trzcińska-Daneluti[1], Adam Górecki[2], Alicja Czubaty[1] Barbara Kowalska-Loth[1], Agnieszka Girstun[1], Magdalena Murawska[1] Bogdan Lesyng[2] and Krzysztof Staroń[1]*

[1]*Institute of Biochemistry Faculty of Biology Warsaw University Miecznikowa 1 02-096 Warsaw, Poland*

[2]*Department of Biophysics Warsaw University Zwirki i Wigury 93 02-089 Warsaw, Poland*

**Corresponding author*

RNA recognition motif (RRM) domains bind both nucleic acids and proteins. Several proteins that contain two closely spaced RRM domains were previously found in protein complexes formed by the cap region of human topoisomerase I, a nuclear enzyme responsible for DNA relaxation or phosphorylation of SR splicing proteins. To obtain molecular insight into specific interactions between the RRM proteins and the cap region of topo I we examined their binary interactions using the yeast two-hybrid system. The interactions were established for hnRNP A1, p54^{nrb} and SF2/ASF, but not for hnRNP L or HuR. To identify the amino acid pattern responsible for binding, experimental mutagenesis was employed and computational modelling of these processes was carried out. These studies revealed that two RRM domains and six residues of the consensus sequence are required for the binding to the cap region. On the basis of the above data, a structural model for the hnRNP A1–topoisomerase I complex was proposed. The main component of the hnRNP A1 binding site is a hydrophobic pocket on the β-surface of the first RRM domain, similar to that described for Y14 protein interacting with Mago. We demonstrated that the interaction between RRM domains and the cap region was important for the kinase reaction catalyzed by topoisomerase I. Together with the previously described inhibitory effect of RRM domains of SF2/ASF on DNA cleavage, the above suggests that the binding of RRM proteins could regulate the activity of topoisomerase I.

Keywords: RRM proteins; topoisomerase I; proteins interaction

Introduction

The RNA recognition motif (RRM) domain is a common eukaryotic RNA-binding motif identified by two short conserved sequences called RNP-2 and RNP-1, separated by 25–35 residues.[1] They form a four-stranded β-sheet backed by two α-helices, which makes a surface with side-chains exposed for the base-binding. The surface is used for a sequence-specific recognition of RNA or DNA, as in the case of hnRNP A1 that binds both types of nucleic acid.[2]

Besides nucleic acid binding, RRM domains are involved in diverse protein–protein interactions. RRM domains recognize either another RRM domain, as for *Drosophila* sex-lethal protein,[3] or a different structure of the protein partner. The latter is the case for complexes formed by Y14 and Mago,[4–6] PABP and Paip2,[7] SF2/ASF and p32,[8] and SF2/ASF and viral e4orf4.[9] Two kinds of interface involved in the protein interactions have been identified for RRMs. The first is a pocket between two α-helices at the back of the surface built by β-strands, which can be penetrated by ligand-W.[10] This structure is universal for complexes formed by RRMs called U2AF homology motifs (UHMs), present in several proteins involved in pre-mRNA processing, as well as DNA repair and signal transduction.[10] The other

Abbreviations used: hnRNP, heterogeneous nuclear ribonucleoprotein; MD, molecular dynamics; PABP, poly (A) binding protein; PB, Poisson–Boltzmann; PSF, polypyrimidine tract-binding protein (PTB)-associated splicing factor; RRM, RNA recognition motif; *SASA*, solvent-accessible surface area; SF2/ASF, splicing factor 2/alternative splicing factor; topo I, human topoisomerase I; TS10, single-stranded $(TAAGGGT)_{10}$ oligonucleotide; UP1, unwinding protein 1; Y2H, yeast two-hybrid system.

E-mail address of the corresponding author: staron@biol.uw.edu.pl

interface is the RRM β-sheet surface used by Y14 in the binding of two α-helices of protein Mago.[4–6] It is not known whether the structure of the complex formed by Y14 and Mago is unique for this protein pair or common for interactions employing RRMs.

In previous work,[11] we found that SF2/ASF used its RRMs to bind human topoisomerase I (topo I). topo I is the main enzyme responsible for DNA relaxation,[12,13] and it is a protein kinase specific for serine/arginine motifs present in essential splicing factors, including SF2/ASF.[14] Interaction of topo I with the kinase substrate SF2/ASF inhibits the DNA cleavage that is a key step in the relaxation reaction.[15,16] topo I is a single polypeptide of 765 amino acid residues, and composed of four domains; the N-terminal domain, the core domain, the linker domain and the C-terminal domain.[12] The core domain is further divided into three subdomains that form two distinct lobes in the three-dimensional structure of topo I: the cap region, containing subdomains I and II, and subdomain III. The cap region, comprising residues 215–433 of topo I, is involved directly in DNA binding,[17] and it is one of two sites of interaction of topo I with SF2/ASF, which binds directly to the tandem of closely spaced RRMs.[11]

Earlier, we showed that besides SF2/ASF, several other proteins that contain a tandem of RRMs are present in complexes formed by the cap region when incubated with the HeLa nuclear extract.[18] If the identified proteins interact directly with the cap region, two interesting hypotheses emerge. First, because the proteins do not have the UHM signature,[10] a novel mode of interaction with a protein partner would be possible for this group of the RRM proteins. Next, because a putative region of the interaction for the RRM proteins is also involved in the binding of the substrates for both topo I activities,[11,17] the regulatory role for the RRM proteins could be considered. The purpose of this study was to test the above hypotheses. We show here, using a yeast two-hybrid system (Y2H) and site-directed mutagenesis, that interaction of the RRM proteins with the cap region requires a specific amino acid pattern present in the first of two closely spaced RRM domains. On the basis of the above, we built a structural model for the complex that pointed to the hydrophobic pocket formed on the β-sheet surface as the main component of the binding site. We show that a similar pocket is used by Y14 to recognize the Mago protein, suggesting a common mode of interaction. We demonstrate that binding of the RRM proteins influences the kinase activity of topo I, suggesting a regulatory role.

Results

Binary interactions between RRM proteins and the cap region

The selection of proteins to be tested in Y2H was based on the list of ten RRM proteins that had been found in the complexes formed by the cap region with proteins present in the nuclear extract from HeLa cells.[18] Because of the high level of sequence identity between RRM domains of hnRNP A1, hnRNP A2/B1 and hnRNP A3,[19] and between $p54^{nrb}$ and the C terminus of PSF,[20] we selected only one protein from each group, hnRNP A1 and $p54^{nrb}$, and we excluded nucleolin and hnRNP R from the analysis. Our earlier experiments performed for SF2/ASF indicated that two closely spaced RRM domains were required for binding to the cap region.[11] Therefore, four RRM domains present in nucleolin and three RRM domains present in hnRNP R did not allow for a simple recognition of which pair was responsible for the interaction. Finally, the initial analysis was performed for five proteins: hnRNP A1, hnRNP L, HuR, $p54^{nrb}$ and SF2/ASF.

All proteins selected to be tested were expressed from pACT2 vector in the yeast strain PJ69-4a (Figure 1(a)). The interaction between Fos and Jun served as a positive control (not shown). The analysis performed for growth on both –His and –Ade selection plates, as well as for β-galactosidase specific activity, provided coherent results (Figure 1(b)). It revealed interaction of the cap region with hnRNP A1, $p54^{nrb}$ and SF2/ASF, but excluded interaction with hnRNP L and HuR. When the specific activity of β-galactosidase was examined, the highest level of activity was found for the interaction between the cap region and hnRNP A1.

Proteins selected in this work for the Y2H test contain two closely spaced RRM domains, which is the motif required for the interaction between SF2/ASF and the cap region,[11] as we found previously. To determine whether the interaction of RRM proteins with the cap region of topo I needs two RRM domains, we performed Y2H analysis using fragments of hnRNP A1, including RRM1, RRM2 or both RRM domains (2×RRM). The assay confirmed that both RRM domains were required for the interaction (Figure 1(b)).

Residues of RRM domains potentially responsible for the interaction

We assumed that residues responsible for the interaction between the RRM proteins and the cap region should be selected from among those that were different for interacting and non-interacting proteins. To identify such residues, amino acid sequences were compared. An alignment of the interacting proteins was used to reveal close amino acid similarities and identities that were found for 46 residues of the proteins (Figure 2). Next, similar alignments were carried out for hnRNP A1 and each non-interacting protein. The latter alignments yielded 23 residues that were different for the interacting and non-interacting proteins, and could be considered as possible sites for the interaction.

Because of the known 3D structure of hnRNP A1,[21,22] this protein was selected for all further computational approaches and experimental mutagenesis. For the sake of simplicity, the numbering of

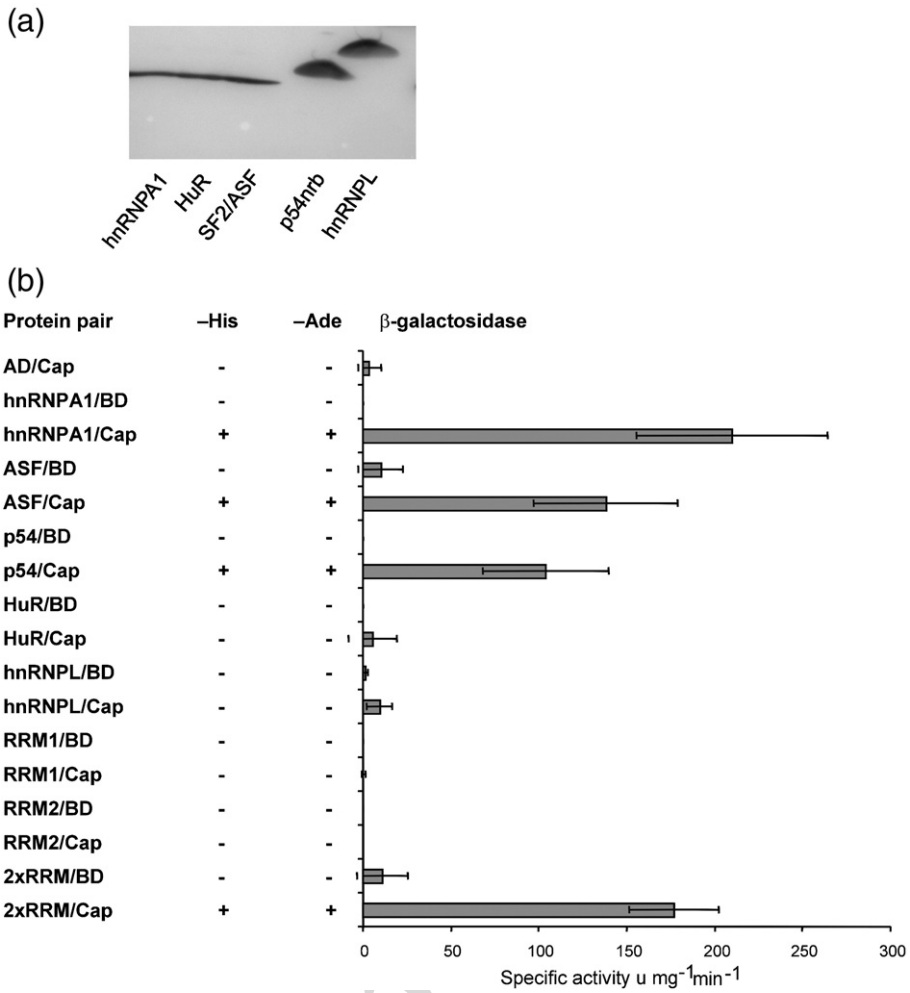

Figure 1. Interaction of the RRM proteins with the cap region of topo I in Y2H system. (a) Expression of the tested proteins in PJ69-4a yeast strain. (b) A representation of growth on plates lacking: histidine (–His), adenine (–Ade) and of β-galactosidase reporter specific activities. BD, binding domain; Cap, the cap region of topo I; RRM1, first RRM domain of hnRNP A1; RRM2, second RRM domain of hnRNP A1; 2xRRM, two RRM domains of hnRNP A1.

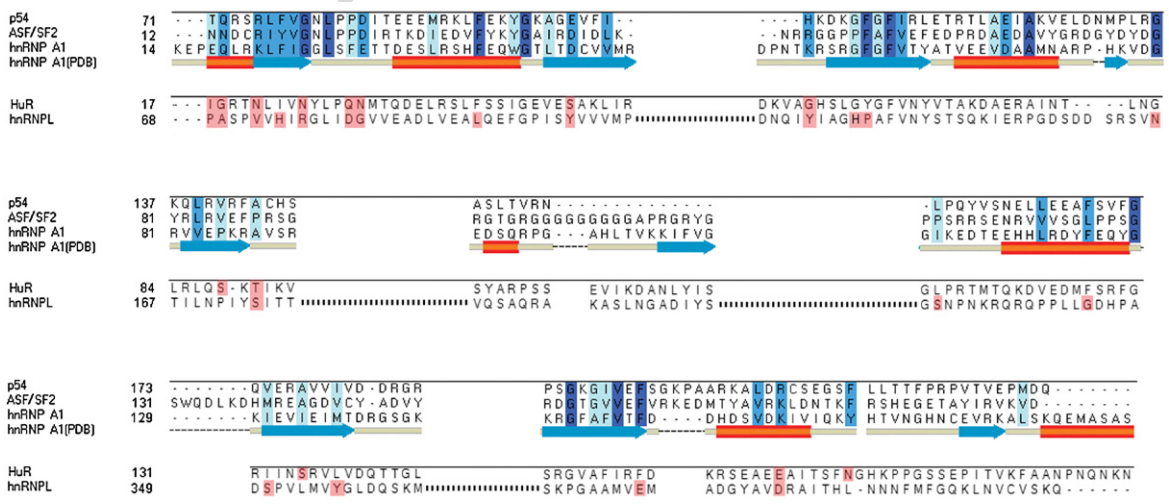

Figure 2. The step-by-step alignment of RRM domains of the interacting and non-interacting proteins. hnRNP A1 [PDB], secondary structure for the RRM domains of hnRNP A1. Residues identical with or similar to the interacting proteins are highlighted in blue. Residues different for the interacting proteins and at least one of the non-interacting proteins (differential residues) are highlighted in red.

residues present in hnRNP A1 (SwissProt P09651) are used in the work for a site identification.

The further procedure employed to reveal RRM residues potentially responsible for the interaction of hnRNP A1, $p54^{nrb}$ and SF2/ASF with the cap region included a search for pockets and cavities on hnRNP A1, computational alanine scanning and site-directed mutagenesis carried out for selected residues.

In the first step, a map of binding pockets and cavities of hnRNP A1 was generated both with the hydrophobic binding free energy-based (Q-SITE Finder[23]) and solvent accessibility surface (CASTp[24]) approaches (Table 1). Next, we used the RosettaInterface[25] computational alanine scanning approach for docking clusters of the topo I – hnRNP A1 complex. We retained a cutoff of $\Delta\Delta G_{bind} > 1$ kcal·mol^{-1} to identify hotspot residue mutations that were predicted to destabilize the interface when mutated to alanine. The RosettaInterface hotspot residues as well as the RRM domains alignment data were loaded into a map of the binding pockets and cavities of the hnRNP A1 protein (Table 1). It resulted in a list of seven residues (E10, Q11, K14, F16, F22, D41 and F56)

Table 1. Potential binding sites of the hnRNP A1 protein determined using the CASTp and Q-SITE Finder applications

Binding site	Residues
A. *Solvent accessibility surface approach*	
P1	S21, **F22**, R81
P2	V162, Q164, V176
P3	L20, T24, S53
P4	K86, D156, D159
P5	E34, T38, L39
P6	R96, P97, L180
P7	F107, E175, R177
P8	R121, D122, E125
P9	K105, A179, L180, S181
P10	L120, **F124**, **I130**, **I133**, V150
P11	**E10**, **D41**, T60
P12	D26, R30, C42
P13	E84, K86, T102
P14	V132, E134, **F149**, V150, T151
P15	L15, **F33**, A70, M71, **P85**
P16	L101, V103, A179, L180, E184
P17	**D41**, F58, V59, T60
P18	D68, M71, N72, R74, R87, D156, **K160**
P19	G110, **I111**, K112, T115, V169, N170, N173, C174
P20	**I111**, K112, E113, T115, I135, **M136**, K144, **G146**, F147
P21	E92, D93, V162, Q164, K165, C174, E175, V176, R177
P22	F107, **M136**, T137, R139, R145, F147, **F149**, A179, L180, K182, M185
P23	**F22**, V64, D68, R87, Y127, **K160**, I163, Q164
P24	**Q11**, L12, **K14**, V67, R87, **A88**, V89, S90, D93, R96, P97, G98, A99, L101, PV103, V162, I163, R177, K178, L180
P25	**E10**, **Q11**, **K14**, **F16**, **D41**, V43, M45, R54, **F56**, F58, K86, R87, **A88**, V89, S90, R91, S94, Q95, L101, T102, V103, D159, V162, I163, K178
B. *Hydrophobic binding free energy-based approach*	
S1	**E10**, **Q11**, L12, R13, **K14**, L15, **F16**, I17, **G18**, L20, **F33**, L39, V43, **F56**, G57, F58, V59, T60, Y61, A62, T63, V64, E65, E66, V67, D68, A70, M71, V78, V83, E84, P85, K86, R87, A88, I163
S2	**E10**, R13, **K14**, L15, I17, L29, R30, F33, E34, W36, G37, T38, L39, T40, **D41**, C42, V43, F58, V59, T60, Y61, A62, V67
S3	**A88**, V89, S90, D93, S94, R96, P97, G98, A99, L101, V103, V162, I163, V176, R177, K178, A179, L180, S181, E184, M185, S187, A188
S4	I106, L120, R123, **F124**, E125, G128, K129, **I130**, E131, V132, **I133**, V150, T151, **F152**
S5	L101, T102, V103, K104, K105, I106, F107, V150, T151, **F152**, D154, D156, S157, V158, D159, **K160**, **I161**, V162, V176, R177, K178, A179
S6	R96, K105, F107, F147, **F149**, R177, K178, A179, L180, S181, K182, E184, M185
S7	L15, **F33**, W36, A70, M71, A73, R74, P75, K77, V78, V83, **P85**
S8	**Q11**, **K14**, **F16**, F58, **A88**, V89, S90, R91, S94
S9	**K14**, I17, L20, T24, T25, D26, L29, R30, **D41**, C42, V43, V44, M45, **F56**, G57, F58, V59
S10	I106, Y123, **F124**, Y127, **F152**, S157, V158, D159, **K160**, **I161**, V162, Q164, V176

Residues different for the interacting proteins and at least one of the non-interacting proteins, are in bold. Residues that are predicted to destabilize the interface significantly when mutated to alanine ($\Delta\Delta G_{bind} > 1$ kcal·mol^{-1}) are highlighted in grey.

that (i) were different between interacting and non-interacting proteins, (ii) were included in the binding pockets of hnRNP A1, and (iii) destabilized the interface between hnRNP A1 and topo I significantly when substituted for alanine. All of the identified residues form the largest pocket of hnRNP A1 (Table 1, P25 and S1).

In the final step, we performed experimental site-directed mutagenesis of selected residues to alanine and searched for interactions of the mutants with the cap region in the Y2H assay. The Y2H assay indicated that five hnRNP A1 mutants, E10A, Q11A, K14A, F16A and F56A, did not interact with the cap region (Figure 3). Mutants F22A and, to lesser extent, D41A, exhibited partially lowered activity of β-galactosidase; however, both mutants grew on –Ade and –His plates (not shown).

Structural model of the hnRNP A1–topo I complex

The experiments described in the preceding section indicated that E10A, Q11A, K14A, F16A and F56A were the residues located within the RRM domain of hnRNP A1 that mediated its interaction with topo I. To build a structural model, we selected docking clusters that were the most consistent with the above experimental results. A detailed model of the complexes was created applying visual modelling with the VMD program,[26] followed with structure relaxation using the NAMD MD package (Figure 4).[27] Two hydrophobic patches were found at the hnRNP A1–topo I interface; a socket of residues F16, F56 and F58 within the first RRM domain of hnRNP A1, and the corresponding M319 within the cap region of topo I (Figure 5(a)). The next components of the binding site

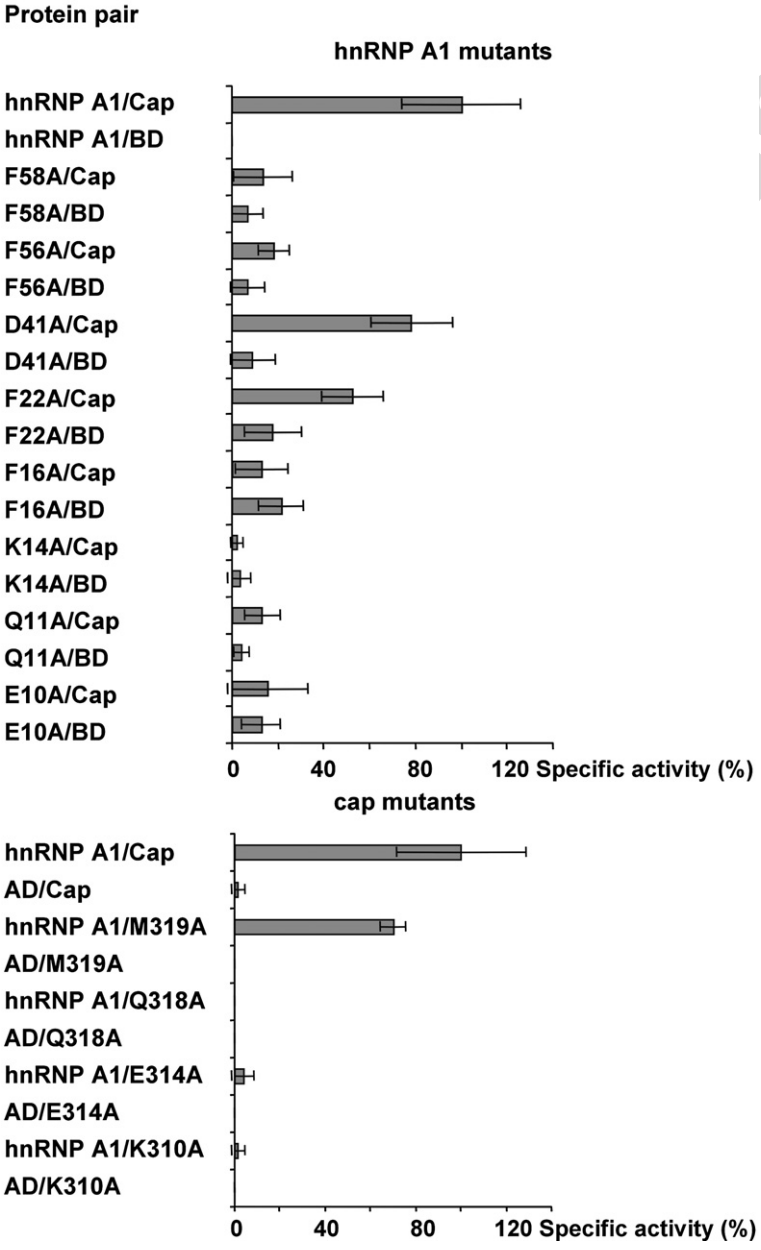

Figure 3. Interaction of alanine mutants of hnRNP A1 with the cap region and the cap mutants with hnRNP A1 in the Y2H system. hnRNP A1 and Cap, wild-type proteins. Other abbreviations as for Figure 1. The activities of β-galactosidase of the mutants are referred to that of hnRNP A1:Cap.

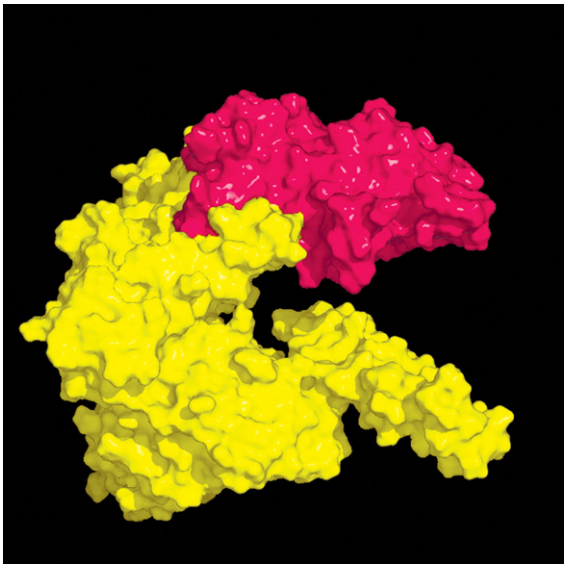

Figure 4. The structural model of the topo I-hnRNP A1 complex. Topo I is coloured yellow, hnRNP A1 pink.

in the topo I–hnRNP A1 complex are two salt-bridges formed between hnRNP A1:E10 and topo I:K310, as well as hnRNP A1:K14 and topo I:E314. (Figure 5(b) and (c)). Also, double hydrogen bonding between hnRNP A1:Q11 and topo I:Q318 (Figure 5(d)) provides some contribution to the complex stabilization. The calculated structure of Q-Q hydrogen bonding is characterized by the following parameters: (i) the distances |NE2–OE1|=2.9 Å, |H–OE1|=1.85 Å, (ii) an angle $<_{NE2,\ H,\ OE1}$=175.7°, in agreement with statistical data.[28] Studies on protein crystal structures indicate that such H-bond interactions between glutamine or asparagine side-chains are common for protein complexes.[29,30] The most important, approximate energy contribution of the interactions to the free energy of the binding are as follows: [topo I:K310]: [hnRNP A1:E10], −9.0 kcal·mol^{-1}; [topo I:E314]: [hnRNP A1:K14], −9.0 kcal·mol^{-1}; [topo I:Q318]: [hnRNP A1:Q11], −3.5 kcal·mol^{-1}; [topo I:M319]: [hnRNP A1:F16, F56, F58], −14.5 kcal·mol^{-1}.

The model excludes F22 and D41 from the binding site of the complex, even though alanine mutants at both residues exhibited partially lowered β-galactosidase activity in the Y2H test (Figure 3). Different

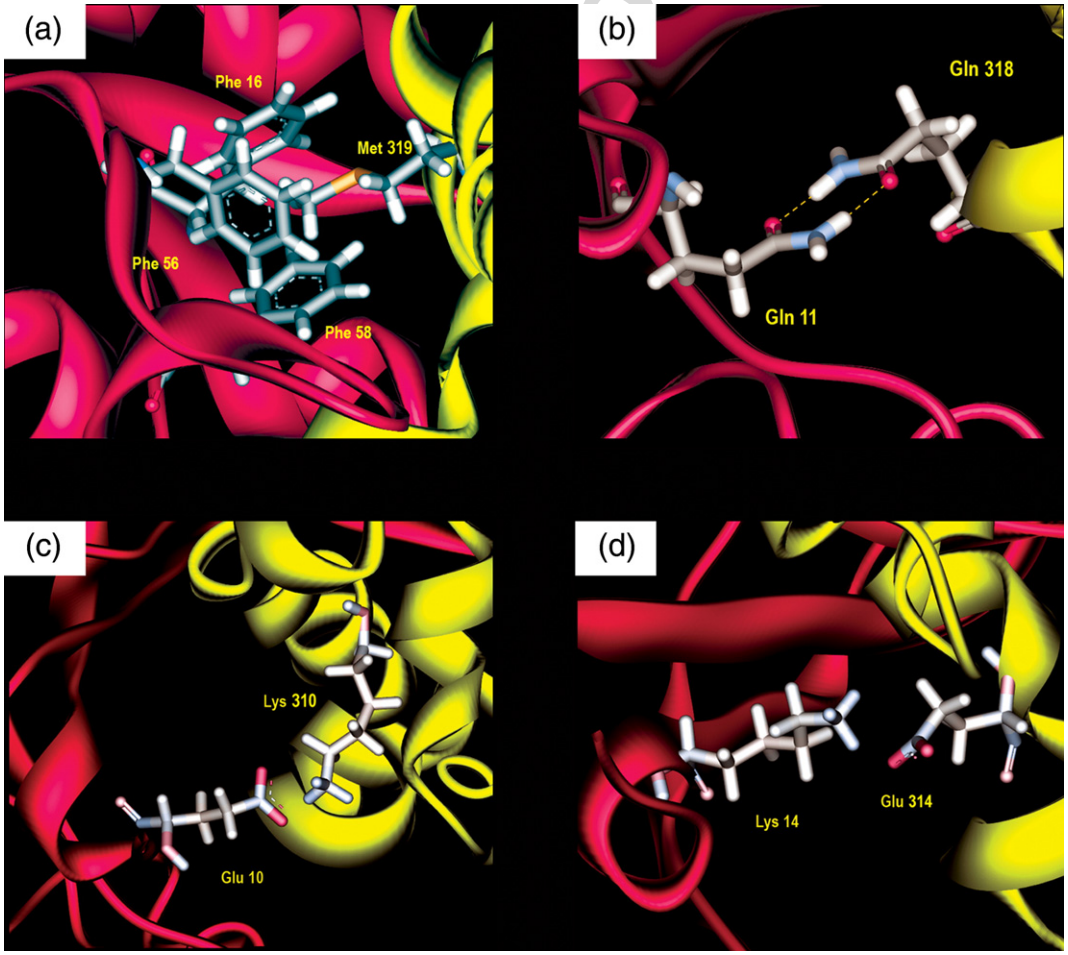

Figure 5. Interactions that stabilize the topoI-hnRNP A1 complex. Topo I is in yellow, hnRNP A1 is in pink. (a) The hydrophobic pocket composed of topo I:M319, hnRNP A1:F16, hnRNP A1:F56 and hnRNP A1:F58. (b) The hydrogen bonding between topo I:Q318 and hnRNP A1:Q11. (c) The charge interaction, topo I:K310–hnRNP A1:E10. (d) The charge interaction topo I:E314–hnRNP A1:K14.

reasons could be considered for each residue. Although D41 is in close proximity of topo I, there is no basic residue in the neighbourhood to create a salt-bridge. However, a weak hydrogen bond with hydrophilic side-chains (e.g. topo I:Q307) could be formed and slightly increase the stability of the complex. This bond would not exist in the D41A mutant. As concerns F22, it is distant from the expected binding site, hydrophobic and exposed to the solvent. Substitution of the large aromatic ring for the small alanine side-chain in the F22A mutant could alter thermal movements of hnRNP A1 protein and a solvent accessibility surface, eventually resulting in a slightly decreased stability of the complex.

According to the model presented here, the interface involved in the interaction with the cap region is limited to the first RRM domain of hnRNP A1, whereas the pull-down experiments with SF2/ASF[11] and the Y2H tests with hnRNP A1 indicate that two RRM domains are required for the binding.

In order to interpret the experimental findings, computations of relative binding free energies were carried out. The binding free energies, ΔG_{total}, were approximated with the sum of mean-field, electrostatic Poisson–Boltzmann, ΔE_{elect}, non-polar (hydrophobic) ΔG_{nonp} and entropic, $\Delta G_{entropic}$, contributions. For details of representative theoretical approaches see e.g. Froloff *et al.*[31], Antosiewicz *et al.*[32], Baker *et al.*[33] and Gruziel *et al.*[34] Descriptions of the implemented model and computational details are presented in Materials and Methods.

The electrostatic contributions to the free energy of binding of topo I with hnRNPA1 and with the single RRM domain, are similar. However, the solvent-accessible surface area (*SASA*) decreases more in the first case, which results in stronger stabilization of the [topo I]:[hnRNP A1] complex by the non-polar (hydrophobic) interactions (Table 2). The difference in the binding free energies of [topo I]:[hnRNP A1] and [topo I]:[RRM] is about 22.5 kcal·mol^{-1}.

Experimental verification of the model

The model revealed several residues potentially involved in the interaction that had not been examined for their significance for the binding. First, it shows hnRNP A1:F58, which builds the hydrophobic pocket together with hnRNP A1:F16 and hnRNP A1:F56 (Figure 5(a)). hnRNP A1:F58 was not selected initially for the site-directed mutagenesis or tested in the Y2H assay because it is conserved in the interacting proteins as well as in the non-interacting HuR (Figure 2). The effects of F58 substitution for alanine were tested in the Y2H assay, which confirmed its crucial role in the binding (Figure 3). Next, we tested four residues of topo I identified by the model, forming the RRM-topo I binding site: K310, E314, Q318 and M319. Substitutions of the first three residues for alanine indicated that they are necessary for the interaction (Figure 3). In contrast, the M319A mutant of the cap region exhibited only slightly lowered β-galactosidase activity in the Y2H test. Probably, it resulted from the hydrophobic character of alanine, which partly replaced methionine in the binding process in the hydrophobic pocket.

Table 2. Theoretical, relative free energies of binding ΔG_{nonp} (kcal·mol^{-1})

Protein complexes	ΔE_{elect}	ΔG_{nonp}	$\Delta G_{entropic}$	ΔG_{total}
[Topo I]:[hnRNP A1]	40.3	−61.2	8.0	−12.9
[Topo I]:[RRM1]	38.2	−36.6	8.0	9.6
[Y14]:[Mago]	18.0	−30.3	4.8	−7.5

$\Delta G_{total} \approx \Delta E_{elect} + \Delta G_{nonp} + \Delta G_{entropic} + C$, where C is a constant, ΔE_{elect} is the electrostatic Poisson–Boltzmann binding energy computed with APBS.[35] ΔG_{nonp} and $\Delta G_{entropic}$ are the non-polar (hydrophobic) and entropic terms, respectively. ΔG_{nonp} was computed from the change of the solvent-accessible surface area (Δ_{SASA}) upon complex formation, $\Delta G_{nonp} \approx \gamma\ \Delta_{SASA}$, $\Delta G_{entropic}$ is the protein side-chain conformational entropy loss upon binding. $\Delta G_{entropic} \approx B\ NR_{buried-res}$, where $NR_{buried-res}$ is the number of residues buried upon binding for each protein (ten for the first and second complex, and six for the third one). For parameters and details of the theoretical model see Materials and Methods.

Effects of alanine substitutions on the stability of the hnRNP A1 protein and the hnRNP A1–topo I complex

The strategy used in this work for selection of putative interacting residues was to substitute them for alanine and to test mutated proteins in Y2H. This strategy assumed that alanine substitution of critical residues did not destroy the structure of hnRNP A1, and altered only its interaction with the cap region.

To evaluate a possible effect of alanine substitutions on the stability of the hnRNP A1 protein, we used the existing data on the structure of RRM proteins, as well as experimental data for the hnRNP A1 mutant. Out of six residues involved in the topo I binding, E10, Q11 and K14 are not in the RNP consensus sequences that determine a basic structure of the RRM domain.[1] The natural alanine mutants at these sites can be found among proteins recognized as RRM proteins (Figure 6(a)). The remaining residues F16, F56 and F58 are located in the RNP motifs critical for the RRM structure. The natural substitutions for alanine at sites corresponding to F16 and F58 are A67 and A174 in human U2AF 35 and TAP proteins, respectively (Figure 6(a)), for which the RRM structures have been confirmed experimentally.[1] To examine the effect of the alanine substitution at residue F56, we examined the specific ability of the mutated hnRNP A1 to recognize single-stranded oligonucleotide carrying ten contiguous telomeric repeats (TS10).[2] The experiment used UP1, a shortened derivative of hnRNP A1, instead of the complete protein. UP1 encompasses both RRM domains but is devoid of the C-terminal glycine-rich region, which is largely unstructured and can interfere with interactions of the recombinant hnRNP A1.[2] In the gel mobility-shift assay, UP1 forms specific complexes with TS10 similar to the complete hnRNP A1 (Figure 6(b)).

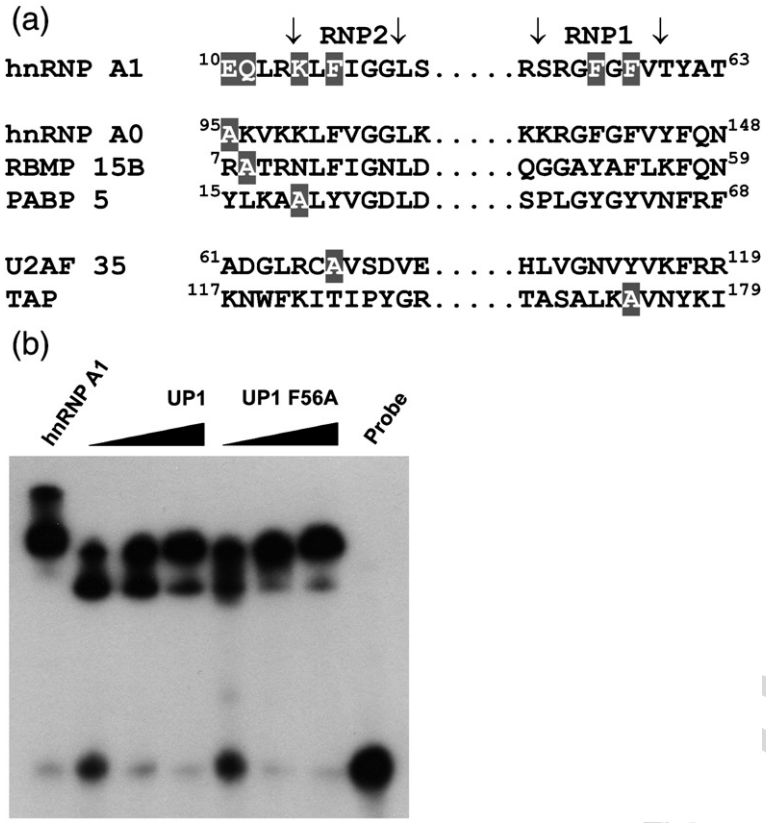

Figure 6. Stability of RRM domains with alanine substitutions used in this work. (a) Natural alanine substitution in RRM proteins at positions corresponding to hnRNP A1: 10, 11, 14, 16 and 58. The alignment of hnRNP A0, RBMP 15B and PABP 5 has been taken from the Prosite documentation PDOC00030; the alignment of U2AF 35 and TAP from Maris *et al.*[1] (b) The gel mobility-shift assay was done with labelled TS10 (approximately 25 fmol) and GST-tagged proteins: hnRNP A1 (3 pmol), UP1 or UP1 F56A. Increasing amounts of UP1 or UP1 F56A (0.16, 0.33 and 1 pmol) were used.

Although the interaction between UP1 and telomeric DNA involves all of the phenylalanine residues present in RNP motifs, F56 forms only weak van der Waals contacts with backbones of the bases that do not seem to be critical for the stability of the complexes.[35] On the other hand, the undisrupted RRM1 is necessary for the interaction.[36] The gel mobility-shift assay (Figure 6(b)) showed that complexes with TS10 were formed by UP1 WT and UP1 F56A at similar concentrations of protein, pointing to a lack of disruption of the RRM1 structure by alanine substitution at F56.

We expected that alanine substitutions in the mutant proteins that made them unable to bind with the cap region should alter the molecular surface of the protein, i.e. the solvent-accessible surface. To examine this, we carried out a comparative visual analysis of the shape of the solvent-accessible surfaces, computed with DS Visualizer, with a probe radius of 1.4 Å. The change of the charged side-chains of hnRNP A1:K14 and hnRNP A1:E10 residues for the uncharged alanine, destroys salt-bridges vital for the stability of the complex. The substitution of hnRNP A1:Q11 makes the complex unstable due to the following effects: (i) change of the solvent-accessible surface and (ii) impossibility of the hydrogen bond recreation (Figure 7(a)). Finally, the substitution of the large hydrophobic aromatic rings for small alanine side-chains at hnRNP A1:F16, hnRNP A1:F56 or hnRNP A1:F58, causes breaking of the hydrophobic pocket and makes the potential topo I:M319 binding impossible (Figure 7(b)–(d)).

Role of the RRM-binding site in the kinase activity

An obvious question is whether the binding of hnRNP A1 to the RRM-binding site of topo I described above has any physiological significance. In previous work, we demonstrated that binding of the tandem RRM domains of SF2/ASF is responsible for inhibition of DNA cleavage catalyzed by topo I.[11] Here, we tested the effect of UP1 and its mutated form UP1 F56A on the kinase activity of topo I.

We demonstrated that substitution of F56 for alanine resulted in substantial reduction of the interaction between recombinant proteins in the pull-down test (Figure 8(a)), similarly to what had been observed for proteins synthetized in yeast in the Y2H assay (Figure 3). Next, we found that UP1 competed with SF2/ASF in binding to topo I and inhibited the kinase reaction (Figure 8(b), lanes 2–4). However, the latter effect was reduced significantly when UP1 F56A was used instead of UP1 WT (Figure 8(b), lanes 5–7).

Other RRM proteins that use a similar mode of binding

The model of the hnRNP A1-topo I complex presented here raises the question of whether the described mode of binding to the partner protein is unique to hnRNP A1 or common among the RRM proteins. Since the model originated from the observation that SF2/ASF and p54^{nrb} bound to the

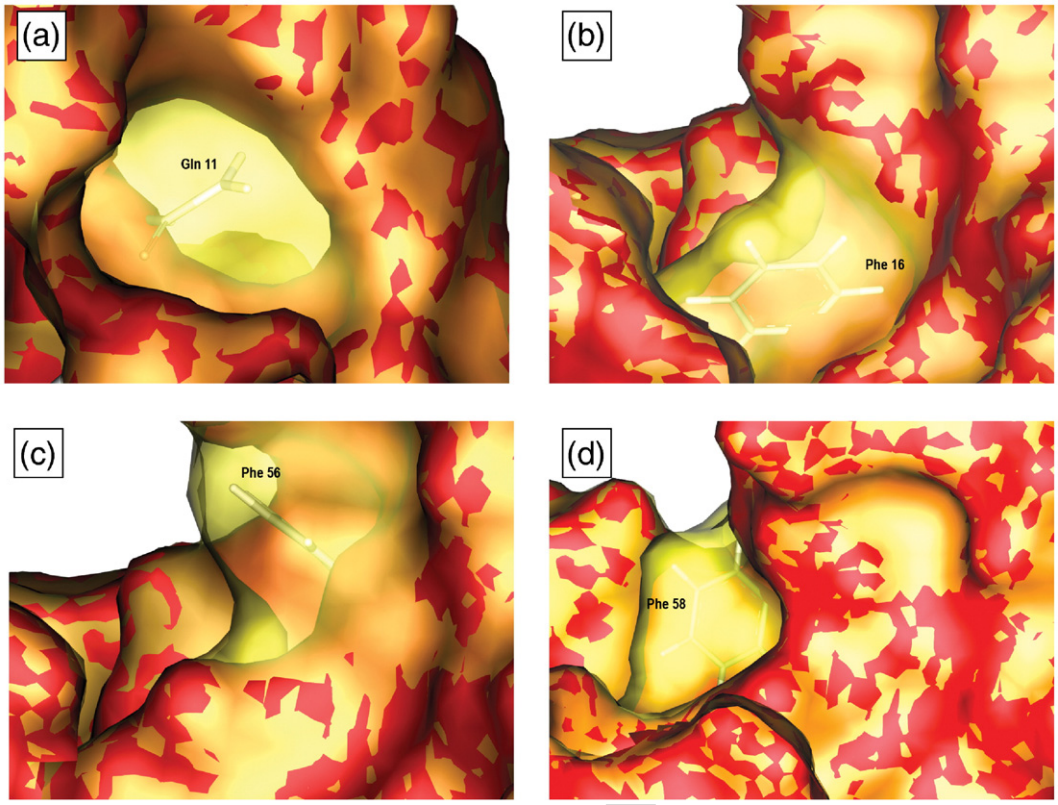

Figure 7. The change of the solvent-accessible surface in four hnRNP A1 mutant proteins. The surfaces of the wild-type (yellow) and the mutant protein (red) were superimposed. (a) Q11A, (b) F16A, (c) F56A,(d) F58A.

cap region in a way similar to that of hnRNP A1, we tested whether these two proteins fulfilled the conditions of the model. In order to predict protein complexes of topo I with SF2/ASF and p54^{nrb}, knowledge-based modelling was employed. The structural models obtained were then superimposed with the structure of RRM1 of hnRNP A1 within the complex with the cap region of topo I (Figure 9(a)). Three aromatic phenylalanine residues of the hydrophobic pocket are structurally conserved in all proteins analysed, except a substitution of hnRNP A1:F16 for Y18 in SF2/ASF. However, SF2/ASF:Y18 can also serve as a pocket wall, as it still contains an aromatic and hydrophobic phenyl ring (Figure 9(b)). Similarly, hnRNP A1:K14 corresponds to SF2/ASF:R16 or p54^{nrb}:R75. Since the basic character of this residue and its geometric position are conserved, it is able to create the salt-bridge. Also Q11 - hydrogen bonding is structurally conserved in all three proteins, since hnRNP A1:Q11 has its counterparts in the form of SF2/ASF:N13 and p54^{nrb}:Q72 residues. Each of these glutamine/asparagine side-chains creates geometrically similar complexes with the same topo I:Q318 residue. Only the charge interaction of hnRNP A1:E10 is not conserved in the proteins examined, since the aforementioned residue is substituted for SF2/ASF:N12 and p54^{nrb}:T71. Nevertheless, both SF2/ASF:N12 and p54^{nrb}:T71 are able to participate in hydrogen bonding with topo I:K310 and topo I:E314, which should stabilize the protein complex.

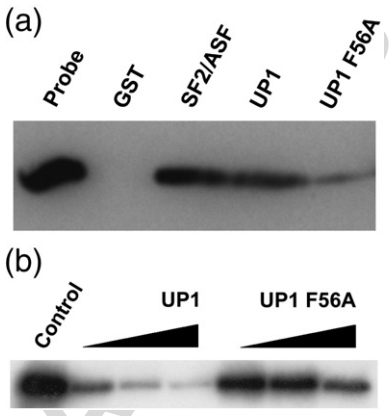

Figure 8. Competition between UP1 and SF2/ASF for the binding site of the kinase substrate. (a) Pull-down assay for binding of proteins to the cap polypeptide. Each sample contained 1 nmol of the indicated GST-tagged protein as the bait and 1 nmol of His-tagged cap as the probe. Western blot was performed with anti-His antibodies. (b) Phosphorylation of SF2/ASF by topo I in the presence of UP1 or UP1 F56A. Each sample contained 12 pmol of His-tagged SF2/ASF, topo I and increasing amounts (24, 48 and 72 pmol) of the competing GST-tagged protein.

To evaluate the significance of the residues listed above for stability of the [SF2/ASF]:[topo I] complex, computational alanine scanning was performed. As shown in Table 3, alanine substitution

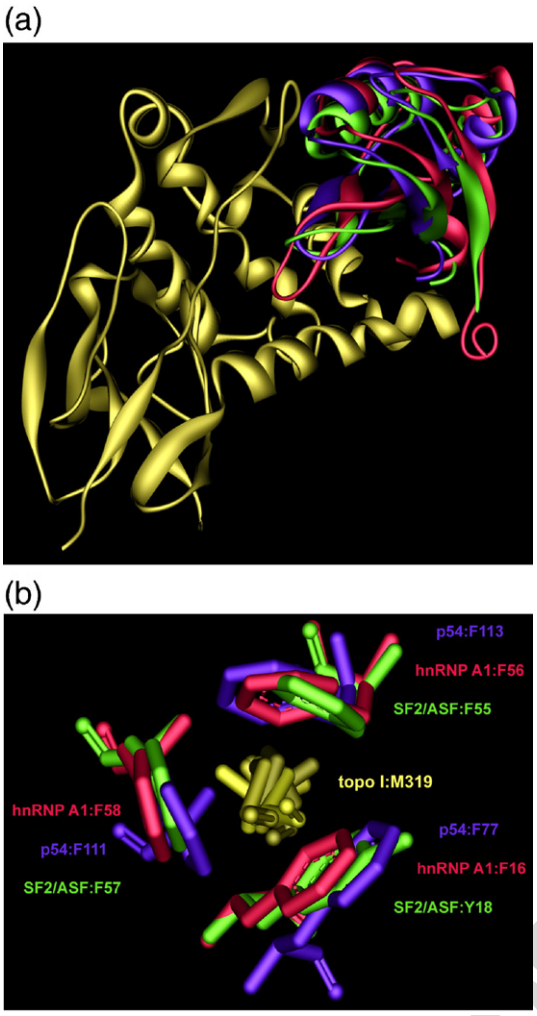

Figure 9. Binding of SF2/ASF and p54^{nrb} to the cap region. (a) Structural superposition of the SF2/ASF and p54^{nrb} models with the hnRNP A1 structure. The model is limited to the cap region of topo I and RRM1 of the ligand proteins. (b) Structural superposition of residues building the hydrophobic pocket on the β-surface. The ligand proteins in (a) and (b) are coloured as follows: SF2/ASF, green, p54^{nrb}, blue, hnRNP A1, pink; topo I is in yellow.

for almost all residues changed the free energy of the complex by ≥1 kcal·mol^{-1}, indicating their significance for the complex stability.

Whereas the interaction surface in the RRM1 domain of SF2/ASF is very similar to that of hnRNP A1, the orientation of the RRM2 is less obvious. In contrast to hnRNP A1, the loop connecting RRM1 and RRM2 domains in SF2/ASF contains a cluster of nine glycine residues (Figure 2), which could make the orientation of both domains more flexible.

Out of several protein–protein interactions described for the RRM domains,[1] only the interaction between RRM protein Y14 and Mago[4–6] uses the β-sheet and resembles the one that was found for topo I and hnRNP A1, subsequently extended to SF2/ASF and p54^{nrb}. Moreover, three hydrophobic residues, F56, Y116 and L118, used by Y14 to bind Mago,[6] are at structural positions similar to F16, F56 and F58 used by hnRNP A1 to bind the cap region (Figure 10(a)). They form a hydrophobic pocket penetrated by Mago: L136 (Figure 10(b)), similar to the hydrophobic pocket on the β-sheet surface of hnRNP A1, which interacts with topo I:M319 (Figure 5(a)). As calculated, alanine substitution for all Y14 residues building the pocket and for Mago:L136 could destabilize the complex (Table 3). One should note that the computed hydrophobic component in the [Y14]:[Mago] interaction is significantly weaker in comparison to the [topo1]:[hnRNP A1] interaction (Table 2). In turn, the electrostatic contribution prefers [Y14]:[Mago], locating finally the latter complex in between [topo I]:[hnRNP A1] and [topo I]:[RRM1].

Discussion

Two basic conclusions emerge from the present study. The first is that several RRM proteins use the β-sheet surface as the interface for interaction with topo I. The other is that the above interaction plays a role in the kinase activity of topo I.

Two of the residues (E10 and Q11) used by RRM1 of hnRNP A1 to bind the cap region are in helix α0 close to the β-sheet surface[22]; the other two (K14 and F16) are in the β1 and (F56 and F58) β3 strands of the sheet. This agrees with the model in which the RRM domains of hnRNP A1 can interact with each other, but the intramolecular interaction occurs through the α-helices and leaves the β-sheet surface free for the binding of nucleic acids or proteins.[22] A key component of the interaction of hnRNP A1 with the cap region is a hydrophobic pocket formed by three phenylalanine residues. A similar pocket is used by RRM of Y14 for the interaction with Mago. On the other hand, the residues involved in the non-hydrophobic interactions are different in hnRNP A1 and in Y14. The above observations suggest that the three residues that produce a hydrophobic pocket

Table 3. Computational alanine scanning for SF2/ASF-topo I and Y14-Mago complexes

SF2/ASF-topo I		Y14-Mago	
Residue	ΔG	Residue	ΔG
SF2:N12	1.32		
SF2:N13	0.94		
SF2:R16	6.22		
SF2:Y18	2.59	Y14:F76	2.60
SF2:F55	1.21	Y14:Y116	2.05
SF2:F57	1.30	Y14:L118	1.44
topo I:K310	0.83		
topo I:E314	3.99		
topo I:Q318	2.80		
topo I:M319	0.41	Mago:L136	1.16

Residues that form a hydrophobic pocket are highlighted in grey. All residues involved in the binding are shown for SF2/ASF-topo I and only those building the hydrophobic interaction for Y14-Mago. The model of the complex described here was used for SF2/ASF-topo I and PDB:1p27 for Y14-Mago. Energy values are in kcal·mol^{-1}.

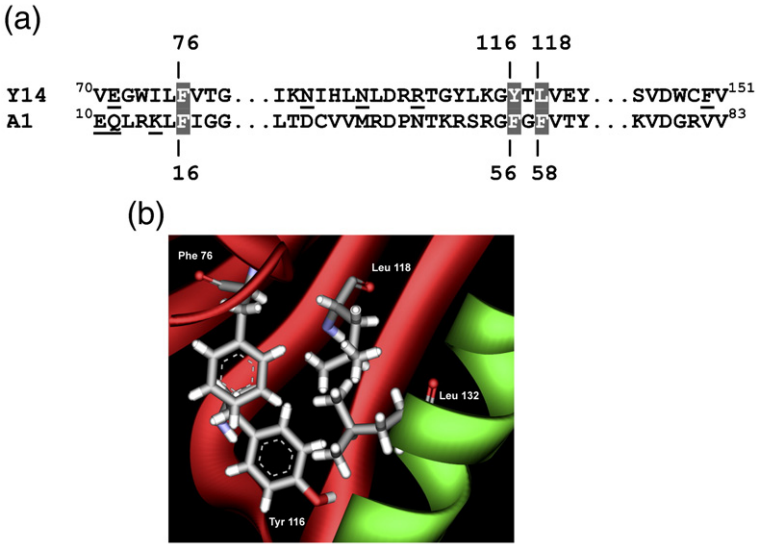

Figure 10. Comparison of hydrophobic pockets used by hnRNP A1 and Y14 for interaction with partner proteins. (a) Alignment of Y14 and hnRNP A1 fragments including residues building hydrophobic pockets. Residues building a hydrophobic pocket are highlighted in grey. Other residues involved in the interaction with the partner protein are underlined. (b) The hydrophobic pocket composed of Y14:F76, Y14:Y116, Y14:L118 and Mago:L136.

could be involved in several interactions of different RRM proteins, whereas the specificity of the interactions could be determined by additional binding residues. It remains to be determined whether this model of the interaction is common among RRM proteins or is limited to the proteins described here. It should be stressed that the aromatic or hydrophobic residues at positions corresponding to hnRNP A1 F16, F56 and F58 are often found in RRM domains that interact with proteins.

The consensus sequence established in this work for the RRM domains required for binding the cap region is indexed as in hnRNP A1: polar residues at positions 10 and 11, a basic residue at position 14, and phenylalanine or tyrosine at positions 16, 56 and 58. Besides the proteins identified here as interacting with the cap region, only a few others fulfil the sequence consensus. This group includes proteins initially considered to be very similar either to hnRNP A1 (hnRNP A2/B1 and hnRNP A3) or p54^{nrb} (PSF) and, additionally, to paraspeckle protein 1 (PSP1). The latter protein shares about 50% sequence identity with PSF and p54^{nrb}, and forms a complex with p54^{nrb}.[37] Thus, when looking for the possible partners that interact with the cap region of topo I through their RRM domains, three groups of proteins should be considered: SF2/ASF, hnRNP A/B proteins as well as a group including PSF, p54^{nrb}, and PSP1. Out of these proteins, only PSP1 was not found in the complexes formed *in vitro* by the cap polypeptide.[18] On the other hand, several observations suggest a possible *in vivo* interaction between the remaining proteins and topo I. SF2/ASF is in contact with topo I during spliceosome assembly,[38,39] PSF and p54^{nrb} copurify with topo I.[40] As concerns hnRNP A1 and topo I, both proteins preferentially bind telomere DNA.[41,42]

The analysis of the protein partners raises the question of why four RRM proteins, HuR, hnRNP L, hnRNP R and nucleolin, were found in the complexes formed by the cap polypeptide,[18] although they did not fulfil the binding consensus. We found that HuR, hnRNP L and hnRNP R (not shown) did not interact directly with the cap region in the Y2H system. Thus, an indirect interaction should be considered. Such a scenario is possible because all four proteins were shown to interact directly with several other proteins present in the complexes formed by the cap polypeptide (see the Human Protein Reference Database).

When looking for a possible physiological role for the interaction between RRM domains and the cap region of topo I it should be noted that the interaction influences both activities of topo I and thus could be considered as a regulatory factor. Firstly, it has been found that SF2/ASF inhibits the DNA cleavage catalyzed by topo I[11,16] because of binding of its RRM tandem domains.[15] Next, we observed in this work that the interaction between the RRM tandem and topo I is vital for the kinase activity. Detailed mechanism that explains the role of the interaction in the inhibition of the DNA cleavage and phosphorylation activity of topo I requires further clarification. Despite this, the results presented here and in our previous paper,[11] point to the interesting possibility that the RRM proteins interacting with the cap region, SF2/ASF and hnRNP A1, could subtly regulate the relaxation and the kinase activities of topo I. This idea could be relevant to the reported changed levels of both SF2/ASF and hnRNP A1 in neoplastic cells.[43–45]

It should be noted that the interaction described here implies a competition between topo I and nucleic acid binding to RRM proteins. Out of six hnRNP A1 residues involved in the interaction with the cap region, five (Q11, K14, F16, F56 and F58) also contact telomeric DNA.[22] Although there is no evidence that such a competition for hnRNP A1 occurs between topo I and telomeric DNA, some observations suggest that other RRM proteins could use the same β-sheet surface for the interaction with both nucleic acids and proteins. Interaction of RRM domains of PABP with Paip2 is accompanied by inhibition of poly(A) RNA binding,[7,12] and interaction of p32 with SF2/ASF blocks functions of the latter protein as a splicing factor.[8]

Materials and Methods

Yeast two-hybrid assay

Constructs prepared using plasmids pACT2 (Gal4AD) and pAS2-1 (Gal4BD) (Clontech) were transformed into the yeast strain PJ69-4a.[46] hnRNP A1, p54^{nrb} and HuR sequences were amplified from cDNA generated using the Invitrogen system with total RNA isolated from HeLa cells and subcloned into the pACT2 vector. The RRM1 (residues 1–96 according to SwissProt P09651), RRM2 (residues 104–183) and 2×RRM (residues 1–183) fragments were generated by PCR from the pACT-hnRNP A1 template. The primers used are shown in Table 4. Mutant clones were sequenced to exclude possible errors introduced during PCR. pACT-SF2, pACT-Jun and pAS-Fos2 were obtained as described.[11] The hnRNP L cDNA was a kind gift from Dr Shouhong Guang, University of Wisconsin. The cap cDNA was subcloned directly into the pAS2-1 vector. Interacting clones were selected for growth on plates lacking histidine or adenine and the level of β-galactosidase activity was measured as described by Clontech. Expression of the fusion proteins in yeast was confirmed by Western blotting using anti-HA monoclonal antibodies (Santa Cruz).

Site-directed mutagenesis

All mutations corresponding to single amino acid substitutions were introduced using DpnI-mediated site-directed mutagenesis with the pACT-hnRNP A1 or pAS-cap plasmids as a template, using the QuikChange® site-directed mutagenesis kit, as described by the manufacturer (Stratagene). The primers used are shown in Table 4. Mutant clones were sequenced to exclude possible errors introduced during PCR.

Pull-down assay

The cDNA for UP1 and UP1 F56A were amplified from pACT vectors carrying wild-type hnRNP A1 or hnRNP A1 F56A, cloned into pGEX-4T-1 and expressed into *Escherichia coli* BL21(DE3). Expression of His-tagged cap, purification and the pull-down assay were performed as described.[11]

Electrophoretic mobility-shift assay

hnRNP A1 cDNA was cloned into pGEX-4A vector and expressed as described above. UP1 and UP1F56A were additionally purified on mono Q and hnRNP A1 on heparin-agarose columns. The mobility-shift assay was done as described.[2] The TS10 oligonucleotide used in the assay consisted of ten repeats of telomeric sequence $(TTAGGG)_{10}$.

Kinase assay

The topo I kinase activity was measured using SF2/ASF and [γ-^{32}P]ATP (5000 Ci·mmol^{-1}; Amersham) as substrates as described.[14] His-tagged UP1 and UP1 F56A cDNA were cloned into the pQE30 vector and expressed as described.[11] Different amounts of His-tagged UP1 or UP1 F56A had been added before addition of topo I. The proteins were analysed using SDS/10% polyacrylamide gels, dried and exposed for 12 h using Rentgen XS-1 (Foton) films.

In silico analyses

The multiple sequence alignment of RRM domains of the proteins examined was generated using CLUSTALX with default parameters.[47] Manual adjustments were introduced on the basis of pairwise alignments of non-interacting proteins and hnRNP A1 and the known 3D structure of hnRNP A1 (PDB 1u1k), as well as the results of secondary structure prediction and tertiary fold recognition for other proteins. A map of binding pockets and cavities of the hnRNP A1 protein was generated with the Q-SITE Finder[23] and CASTp[24] applications. RosettaInterface[25] was used for alanine scanning. *Ab initio* docking calculations were carried out using the GRAMM and DOT programs as described.[18]

Refinement of the hnRNP A1- topo I complex and estimation of binding free energy contributions

Docking clusters of the topo I-hnRNP A1 complex obtained from the GRAMM and DOT *ab initio* calculations that were the most consistent with the site-directed mutagenesis data were selected for further analysis. The selected structural models constitute a subset of the hnRNP A1 residues that were found to be important for the binding. The final hnRNP A1–topo I structural model was created by applying visual modelling using the VMD program.[26] The topo I surface was analyzed in order to find residues of topo I that are complementary for the hnRNP A1 sites. Regarding criteria for the complementarity, the presence of hydrophobic clusters and the ability to create salt-bridges and hydrogen bonds were applied. The predicted structural complex was protonated at pH 7.0 and solvated in the water box. Na^+ and Cl^- were added to neutralize the total charge. Topology was generated with VMD 1.8.4 plug-ins, and the CHARMM 27 force-field was applied for further simulations.

The structure equilibrations were carried out applying MD simulations with the NAMD package.[27] The simulations were carried out using the Langevin dynamics model with a constant temperature of 310 K and a pressure of 1013.25 hPa. In some simulations, additional harmonic constraints for distances or dihedrals were applied, i.e. to stabilize some long-term effects resulting from hydrophobic interactions (phenylalanine hydrophobic pocket), or for interactions that were not reproduced precisely by the applied force-field (glutamine–glutamine double hydrogen bonds). The contribution of the glutamine–glutamine double hydrogen bonding was computed quantum-mechanically with the DMol3 program (Accelrys). The salt-bridge energies were computed by applying a semi-empirical formula that accounts for the screening of the electric field.[48] Hydrophobic interactions were based on the Solvent Accessibility Surface Area model.[34,49]

Homology modelling and superposition of the SF2/ASF, p54^{nrb} models

Knowledge-based modelling was employed to predict the structures of the SF2/ASF and p54^{nrb} proteins. The known 3D structures of related family members (PDB template files u1k, 1l3k, and 1pgz) were found and the reference structure was aligned with the target sequence by using the Accelrys Discovery Studio Protein Modeling tool. The refined sequence–structure alignment was used to construct a 3D model with a successive model evaluation. The structures were verified with a DS Verify Structure tool with a good scoring and further optimized by MD (NAMD). The optimized SF2/ASF and p54^{nrb}

Table 4. Primers used in this study

Protein	Primers	Restriction endonucleases
A. *Yeast two-hybrid*		
hnRNP A1	5′-CTTGGATCCTCATGTCTAAGTCAGAGTCT-3′ 5′-TGACGAATTCTTAAAATCTTCTGCCACT-3′	EcoRI/BamHI
SF2/ASF	5′-AAAGGATCCTCATGTCGGGAGGTGGTGTGAT-3′ 5′-AAAGAATTCGTTATGTACGAGAGCGAGATCTGC-3′	EcoRI/BamHI
p54	5′-ACGGGATCCTCATGCAGAGTAATAAAACTTT-3′ 5′-ACGAATTCTTAGTATCGGCGACGTTT-3′	EcoRI/BamHI
HuR	5′-AGCGAATTCTCATGTCTAATGGTTATGA-3′ 5′-TTTCTCGAGTTTATTTGTGGGACTTGTT-3′	EcoRI/XhoI
hnRNP L	5′-AAAGGATCCTAATGTCGCGGAGGCTGCTG-3′ 5′-AAAGAATTCTTAGGAGGCGTGCTGAGCA-3′	EcoRI/BamHI
RRM1	5′-CTTGGATCCTCATGTCTAAGTCAGAGTCT-3′ 5′-TATGAATTCTCATCTTTGAGAATCTTCTCT-3′	EcoRI/BamHI
RRM2	5′-AAAGGATCCTCAAGATATTTGTTGGTGGCA-3′ 5′-TGTGAATTCTCATTGCTTTGACAGGGCTTT-3′	EcoRI/BamHI
2xRRM	5′-CTTGGATCCTCATGTCTAAGTCAGAGTCT-3′ 5′-TGTGAATTCTCATTGCTTTGACAGGGCTTT-3′	EcoRI/BamHI
Fos2	5′-AAAGAATTCTCTCCAGAAGAAGAAGAGA-3′ 5′-AAAGGATCCCTTACTCTTCTGGGAAGCCCA-3′	EcoRI/BamHI
Jun	5′-AAAGGATCCTCCGGATCGCCCGGCTCGA-3′ 5′-AAAGAATTCGTTACTAGTGGTTCATGACTTTCTG-3′	EcoRI/BamHI
B. *Pull-down assay*		
UP1/UP1F56A	5′–CGTAGGATCCATGTCTAAGTCAGAGTCT–3′ 5′–TACGGTCGACTCATCGACCTCTTTGGCTGGAT–3′	BamHI/Sal I
C. *Mutagenesis*		
hnRNP A1 residues		
E10A	5′-GAGTCTCCTAAAGAGCCCGCACAGCTGAGGAAGCTCTTC-3′ 5′-GAAGAGCTTCCTCAGCTGTGCGGGCTCTTTAGGAGACTC-3′	
Q11A	5′-TCTCCTAAAGAGCCCGAAGCGCTGAGGAAGCTCTTCAT-3′ 5′-ATGAAGAGCTTCCTCAGCGCTTCGGGCTCTTTAGGAGA-3′	
K14A	5′-GAGCCCGAACAGCTGAGGGCGCTCTTCATTGGAGGGTT-3′ 5′-AACCCTCCAATGAAGAGCGCCCTCAGCTGTTCGGGCTC-3′	
F16A	5′-ACAGCTGAGGAAGCTCGCCATTGGAGGGTTGAGC-3′ 5′-GCTCAACCCTCCAATGGCGAGCTTCCTCAGCTGT-3′	
F22A	5′-TTCATTGGAGGGTTGAGCGCTGAAACAACTGATGAGAG-3′ 5′-CTCTCATCAGTTGTTTCAGCGCTCAACCCTCCAATGAA-3′	
D41A	5′-GGGGAACGCTCACGGCCTGTGTGGTAATGAG-3′ 5′-CTCATTACCACACAGGCCGTGAGCGTTCCCC-3′	
F56A	5′-ACCAAGCGCTCTAGGGGCGCTGGGTTTGTCACATATGC-3′ 5′-GCATATGTGACAAACCCAGCGCCCCTAGAGCGCTTGGT-3′	
F58A	5′-TCCAGGGGCTTTGGGGCTGGGTTTGTCACATAT-3′ 5′-ATATGTGACAAACCCAGCCCCAAAGCCCCTGGA-3′	
D. *topo I residues*		
K310A	5′-GAGCCAGTATTTCGCAGCCCAGACGGAAGCTC-3′ 5′-GAGCTTCCGTCTGGGCTGCGAAATACTGGCTC-3′	
E314A	5′-CAAAGCCCAGACGGCAGCTCGGAAACAGATG-3′ 5′-CATCTGTTTCCGAGCTGCCGTCTGGGCTTTG-3′	
Q318A	5′-GACGGAAGCTCGGAAAGCGATGAGCAAGGAAGAG-3′ 5′-CTCTTCCTTGCTCATCGCTTTCCGAGCTTCCGTC-3′	
M319A	5′-GGAAGCTCGGAAACAGGCGAGCAAGGAAGAGAAAC-3′ 5′-GTTTCTCTTCCTTGCTCGCCTGTTTCCGAGCTTCC-3′	

models were superimposed with the hnRNP A1 structure within the complex with topo I. The docked structures were equilibrated by applying MD with constraints preserving hydrophobic and hydrogen bond interactions, which might not be reproduced sufficiently precisely by the force-field. The equilibrated complexes were visualized with DS Visualizer and further analysed by superimposing and comparing the binding sites of the examined proteins.

Estimation of the binding free energy

For an overview of methods and applications used in this study see e.g. Froloff *et al.*,[31] Antosiewicz *et al.*[32] and Gruziel *et al.*[34]

The binding free energies in water were computed for the following complexes: (i) topo I with the hnRNP A1 protein; (ii) topo I with the RRM1 fragment; and (iii) Mago with the Y14 protein (PDB 1p27). Before the binding free energy calculations, the designed geometries of the complexes were relaxed using the energy minimization and MD simulations with the scalable molecular dynamics (NAMD†)[27] and Visual Molecular Dynamics. (VMD‡),[26] simulation packages. Following the study by Trylska *et al.*,[51] the total binding free energy (ΔG_{total}) was approximated as the sum of the following contributions:

$$\Delta G_{\text{total}} \approx \Delta E_{\text{elect}} + \Delta G_{\text{nomp}} + \Delta G_{\text{entropic}} + C$$

† http://www.ks.uiuc.edu/Research/namd/
‡ http://www.ks.uiuc.edu/Research/vmd/

where ΔE_{elect} is the electrostatic Poisson–Boltzmann binding energy in water, ΔG_{nonp} and $\Delta G_{entropic}$ and are non-polar (hydrophobic) and entropic terms, respectively, and C is a constant.

The electrostatic binding energy ΔE_{elect} of a complex was computed as the difference of the energy of the complex minus the energies of the monomers, with all energies computed using the grid of the complex, with fixed origin and orientation. The energies were computed using the Adaptive Poisson-Boltzmann Solver (APBS§)[33] software for evaluating the electrostatic properties of nanoscale biomolecular systems. The following parameters were assumed: 0.1 M salt concentration, $T = 300$ K, the protein dielectric constant $\varepsilon_p = 4$, and the solvent dielectric constant $\varepsilon_s = 80$. Other parameters were set to default values.

The non-polar binding free energy ΔG_{nonp} was estimated multiplying the change of the solvent-accessible surface area (Δ_{SASA}) upon complex formation by an average tension coefficient, γ, of 19 cal·mol^{-1}·Å^{-2}, which was applied in studies of large conformational changes in calmodulin by Yang *et al.*[50] Δ_{SASA} was computed using the Accelrys Discovery Studio Visualizer‖.

$\Delta G_{entropic}$ is the protein side-chain conformational entropy loss upon binding. $\Delta G_{entropic} \approx B\ NR_{buried-res}$, where $NR_{buried-res}$ is the number of residues buried upon binding for each protein. The B coefficient was assumed to be 0.8 kcal/mol per residue buried.[51]

Acknowledgements

We thank Dr. Shouhong Guang, University of Wisconsin, for kind gift of the hnRNP L cDNA and Arkadiusz Miciałkiewicz, Institute of Biochemistry and Biophysics, Warsaw, for substantial help with yeast maintenance. We thank also Dr Jan Antosiewicz, Dr Joanna Trylska and Piotr Kmiec for discussions on the electrostatics and free energy calculations. This work was supported by the Polish Committee for Scientifc Research Grant N301 053 32/1969. Computations were partially carried out at ICM, Warsaw University. Theoretical studies were also supported by CoE BioExploratorium. A.M.T. and A.G. thank for the PhD grants of Faculty of Biology and Faculty of Physics, Warsaw University.

References

1. Maris, Ch., Dominguez, C. & Allain, F. H.-T. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* **272**, 2118–2131.
2. Fiset, S. & Chabot, B. (2001). hnRNP A1 may interact simultaneously with telomeric DNA and the human telomerase RNA *in vitro*. *Nucl. Acids Res.* **29**, 2268–2275.
3. Samuels, M., Deshpande, G. & Schedl, P. (1998). Activities of the Sex-lethal protein in RNA binding and protein:protein interactions. *Nucl. Acids Res.* **26**, 2625–2637.
4. Shi, H. & Xu, R.-M. (2003). Crystal structure of the *Drosophila* Mago nashi–Y14 complex. *Genes Dev.* **17**, 971–976.
5. Lau, C-K., Diem, M. D., Dreyfuss, G. & Van Duyne, G. D. (2003). Structure of the Y14-Magoh core of the exon junction complex. *Curr. Biol.* **13**, 933–941.
6. Fribourg, S., Gatfield, D., Izaurralde, E. & Conti, E. (2003). A novel mode of RBD-protein recognition in the Y14-Mago complex. *Nature Struct. Biol.* **10**, 433–439.
7. Khaleghpour, K., Kahvejian, A., De Crescenzo, G., Roy, G., Svitkin, Y. V., Imataka, H. *et al.* (2001). Dual interactions of the translational repressor Paip2 with poly(A) binding protein. *Mol. Cell. Biol.* **21**, 5200–5213.
8. Petersen-Mahrt, S. K., Estmer, C., Ohrmalm, C., Matthews, D. A., Russell, W. C. & Akusjarvi, G. (1999). The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation. *EMBO J.* **18**, 1014–1024.
9. Estmer Nilsson, C., Petersen-Mahrt, S., Durot, C., Shtrichman, R., Krainer, A. R., Kleinberger, T. *et al.* (2001). The adenovirus E4-ORF4 splicing enhancer protein interacts with a subset of phosphorylated SR proteins. *EMBO J.* **20**, 864–871.
10. Kielkopf, C. L., Lücke, S. & Green, M. R. (2004). U2AF homology motifs: protein recognition in the RRM world. *Genes Dev.* **18**, 1513–1526.
11. Kowalska-Loth, B., Girstun, A., Trzcińska, A. M., Piekiełko-Witkowska, A. & Staroń, K. (2005). SF2/ASF protein binds to the cap region of human topoisomerase I through two RRM domains. *Biochem. Biophys. Res. Commun.* **331**, 398–403.
12. Leppard, J. B. & Champoux, J. J. (2005). Human DNA topoisomerase I: relaxation, roles, and damage control. *Chromosoma*, **114**, 75–85.
13. Corbett, K. D. & Berger, J. M. (2004). Structure, molecular mechanisms, and evolutionary relationships in DNA topoisomerases. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 95–118.
14. Rossi, F., Labourier, E., Forné, T., Divita, G., Derancourt, J., Riou, J. F. *et al.* (1996). Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. *Nature*, **381**, 80–82.
15. Kowalska-Loth, B., Girstun, A., Piekiełko, A. & Staroń, K. (2002). SF2/ASF protein inhibits camptothecin-induced DNA cleavage by human topoisomerase I. *Eur. J. Biochem.* **269**, 3504–3510.
16. Andersen, F. F., Tange, T.Ø., Sinnathamby, T., Olsen, J. R., Andersen, K. E., Westergaard, O. *et al.* (2002). The RNA splicing factor ASF/SF2 inhibits human topoisomerase I mediated DNA relaxation. *J. Mol. Biol.* **322**, 677–686.
17. Yang, Z. & Champoux, J. J. (2002). Reconstitution of enzymatic activity by the association of the cap and catalytic domains of human topoisomerase I. *J. Biol. Chem.* **277**, 30815–30823.
18. Czubaty, A., Girstun, A., Kowalska-Loth, B., Trzcińska, A. M., Purta, E., Winczura, A. *et al.* (2005). Proteomic analysis of complexes formed by human topoisomerase I. *Biochim. Biophys. Acta*, **1749**, 133–141.
19. Ma, A. S., Moran-Jones, K., Shan, J., Trent, P., Munro, T. P., Snee, M. J. *et al.* (2002). Heterogeneous nuclear ribonucleoprotein A3, a novel RNA trafficking response element-binding protein. *J. Biol. Chem.* **277**, 18010–18020.
20. Shav-Tal, Y. & Zipori, D. (2002). PSF and p54(nrb)/NonO-multi-functional nuclear proteins. *FEBS Lett.* **531**, 109–114.
21. Ding, J., Hayashi, M. K., Zhang, Y., Manche, L., Krainer, A. R. & Xu, R. M. (1999). Crystal structure of

§ http://www.apbs.sourceforge.net/
‖ http://www.accelrys.com/

the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev.* **13**, 1102–1115.
22. Myers, J. C. & Shamoo, Y. (2004). Human UP1 as a model for understanding purine recognition in the family of proteins containing the RNA recognition motif (RRM). *J. Mol. Biol.* **342**, 743–756.
23. Laurie, A. T. & Jackson, R. M. (2005). Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics*, **21**, 1908–1916.
24. Binkowski, T. A., Naghibzadeh, S. & Liang, J. (2003). CASTp: computed atlas of surface topography of proteins. *Nucl. Acids Res.* **31**, 3352–3355.
25. Kortemme, T., Kim, D. E. & Baker, D. (2004). Computational alanine scanning of protein-protein interfaces. *Sci. STKE*, **2004**, pl2.
26. Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD - Visual Molecular Dynamics. *J. Molec. Graphics*, **14**, 33–38.
27. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E. *et al.* (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802.
28. McDonald, I. K. & Thornton, J. M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793.
29. Stapley, B. J. & Doig, A. J. (1997). Hydrogen bonding interactions between glutamine and asparagine in α-helical peptides. *J. Mol. Biol.* **272**, 465–473.
30. Shimoni, L. & Glusker, J. P. (1995). Hydrogen bonding motifs of protein side chains: descriptions of binding of arginine and amide groups. *Protein Sci.* **4**, 67–74.
31. Froloff, N., Windemuth, A. & Honig, B. (1997). On the calculation of binding free energies using continuum methods: application to MHC class I protein-peptide interactions. *Protein Sci.* **6**, 1293–1301.
32. Antosiewicz, J., Błachut-Okrasinska, E., Grycuk, T. & Lesyng, B. (2000). A correlation between protonation equilibria in biomolecular systems and their shapes: studies using a Poisson-Boltzmann model. In *Free Boundary Problems. Theory and Applications. Mathematical Sciences and Applications* (Kenmochi, N., ed.), vol. 14, pp. 11–17. Gakuto International Series, Tokyo.
33. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
34. Gruziel, M., Kmiec, P., Trylska, J. & Lesyng, B. (2007). Selected microscopic and mezoscopic modelling tools and models - an overview. In *Molecular Materials with Specific Interactions: Modeling and Design* (Sokalski, A., ed.), pp. 203–224, Springer, London.
35. Ding, J., Mariko, K., Hayashi, M. K., Zhang, Y., Manche, L., Krainer, A. R. *et al.* (1999). Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev.* **13**, 1102–1115.
36. Dallaire, F., Dupuis, S., Fiset, S. & Chabot, B. (2000). Heterogeneous nuclear ribonucleoprotein A1 and UP1 protect mammalian telomeric repeats and modulate telomere replication *in vitro*. *J. Biol. Chem.* **275**, 14509–14516.
37. Fox, A. H., Bond, C. S. & Lamond, A. I. (2005). P54nrb forms a heterodimer with PSP1 that localizes to paraspeckles in an RNA-dependent manner. *Mol. Biol. Cell.* **16**, 5304–5315.
38. Soret, J., Gabut, M., Dupon, C., Kohlhagen, G., Stévenin, J., Pommier, Y. *et al.* (2003). Altered serine/arginine-rich protein phosphorylation and exonic enhancer-dependent splicing in mammalian cells lacking topoisomerase I. *Cancer Res.* **63**, 8203–8211.
39. Pilch, B., Allemand, E., Facompré, M., Bailly, C., Riou, J. F., Soret, J. *et al.* (2001). Specific inhibition of serine- and arginine-rich splicing factors phosphorylation, spliceosome assembly, and splicing by the antitumor drug NB-5061. *Cancer Res.* **61**, 6876–6884.
40. Straub, T., Grue, P., Uhse, A., Lisby, M., Knudsen, B. R., Tange, T.Ø. *et al.* (1998). The RNA-splicing factor PSF/p54nrb controls DNA-topoisomerase I activity by a direct interaction. *J. Biol. Chem.* **273**, 26261–26264.
41. Dallaire, F., Dupuis, S., Fiset, S. & Chabot, B. (2000). Heterogeneous nuclear ribonucleoprotein A1 and UP1 protect mammalian telomeric repeats and modulate telomere replication *in vitro*. *J. Biol. Chem.* **275**, 14509–14516.
42. Kang, M. R., Muller, M. T. & Chung, I. K. (2004). Telomeric DNA damage by topoisomerase I. A possible mechanism for cell killing by camptothecin. *J. Biol. Chem.* **279**, 12535–12541.
43. Ghigna, C., Moroni, M., Porta, G., Riva, S. & Biamonti, G. (1998). Altered expression of heterogeneous nuclear ribonucleoproteins and SR factors in human colon adenocarcinomas. *Cancer Res.* **58**, 5818–5824.
44. Maeda, T. & Furukawa, S. (2001). Transformation-associated changes in gene expression of alternative splicing regulatory factors in mouse fibroblast cells. *Oncol. Rep.* **8**, 563–566.
45. Zerbe, L. K., Pino, I., Pio, R., Cosper, P. F., Dwyer-Nield, L. D., Meyer, A. M. *et al.* (2004). Relative amounts of antagonistic splicing factors, hnRNP A1 and ASF/SF2, change during neoplastic lung growth: implications for pre-mRNA processing. *Mol. Carcinogen.* **41**, 187–196.
46. James, P., Halladay, J. & Craig, E. A. (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics*, **144**, 1425–1436.
47. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* **25**, 4876–4882.
48. Leach, A. R. (2001). *Molecular Modelling. Principles and Applications.* Pearson Education Limited, Essex, UK.
49. Horton, N. & Lewis, M. (1992). Calculation of the free energy of association for protein complexes. *Protein Sci.* **1**, 169–181.
50. Yang, Ch., Jas, G. S. & Kuczera, K. (2001). Structure and dynamics of calcium-activated calmodulin in solution. *J. Biomol. Struct. Dynam.* **19**, 247–271.
51. Trylska, J., McCammon, J. A. & Brookes, Ch. L., 3rd (2005). Exploring assembly energetics of the 30S ribosomal subunit using an implicit solvent approach. *J. Am. Chem. Soc.* **127**, 11125–11133.

Edited by P. J. Hagerman

(Received 15 February 2007; accepted 4 April 2007)
Available online 12 April 2007