Published online 15 January 2010 *Nucleic Acids Research, 2010, Vol. 38, No. 9* ***2975–2989***
doi:10.1093/nar/gkp1230

Mapping of interaction sites of the Schizosaccharomyces pombe protein Translin with nucleic acids and proteins: a combined molecular genetics and bioinformatics study

Elad Eliahoo[1], Ron Ben Yosef[1], Laura Pérez-Cano[2], Juan Fernández-Recio[2], Fabian Glaser[3] and Haim Manor[1,*]

[1]Department of Biology, Technion-Israel Institute of Technology, Haifa 32000, Israel, [2]Life Sciences Department, Barcelona Supercomputing Center (BSC), Jordi Girona 29, Barcelona, 08034, Spain and [3]Bioinformatics Knowledge Unit, The Lorry I. Lokey Interdisciplinary Center for Life Sciences and Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel

Received December 3, 2009; Revised December 22, 2009; Accepted December 23, 2009

ABSTRACT

Translin is a single-stranded RNA- and DNA-binding protein, which has been highly conserved in eukaryotes, from man to *Schizosaccharomyces pombe*. TRAX is a Translin paralog associated with Translin, which has coevolved with it. We generated structural models of the *S. pombe* Translin (spTranslin), based on the solved 3D structure of the human ortholog. Using several bioinformatics computation tools, we identified in the equatorial part of the protein a putative nucleic acids interaction surface, which includes many polar and positively charged residues, mostly arginines, surrounding a shallow cavity. Experimental verification of the bioinformatics predictions was obtained by assays of nucleic acids binding to amino acid substitution variants made in this region. Bioinformatics combined with yeast two-hybrid assays and proteomic analyses of deletion variants, also identified at the top of the spTranslin structure a region required for interaction with spTRAX, and for spTranslin dimerization. In addition, bioinformatics predicted the presence of a second protein-protein interaction site at the bottom of the spTranslin structure. Similar nucleic acid and protein interaction sites were also predicted for the human Translin. Thus, our results appear to generally apply to the Translin family of proteins, and are expected to contribute to a further elucidation of their functions.

INTRODUCTION

Translin is a single-stranded DNA and RNA binding protein, which has a high affinity for G-rich sequences. It was first discovered in extracts of human and mouse cells and found to constitute a multimeric complex (1–3). Subsequent cloning of the human Translin gene and a study of the recombinant protein by electron microscopy revealed that it forms octameric rings consisting of eight identical subunits; moreover, these octamers were found to be the structures that actually bind single-stranded nucleic acids (4–7). TRAX is a Translin paralog, which does not bind nucleic acids. However, it forms specific complexes with Translin (8) and these complexes may also bind nucleic acids (9,10).

Both Translin and TRAX are highly conserved in many higher and lower eukaryotes, an indication of their participation in biologically significant functions. Based on the selective affinity of the human Translin towards single-stranded DNA sequences found near chromosomal translocation sites, it has been suggested that it might be involved in translocation events (4). The subsequent observation that Translin has even higher affinities towards single-stranded $d(GT)_n$ and $d(TTAGGG)_n$ repeats led to the suggestion that it might play a role in the metabolism of $d(GT)_n.d(AC)_n$ microsatellites and telomeres (11). The initial studies of the mouse Translin focused on its preferential binding to the 3′-untranslated regions of mRNAs found in testis and brain (3). Based on these and subsequent studies, it has been suggested that complexes including Translin and TRAX may be involved in the control of mRNA translation and transport (12–17). Recent studies performed in mouse and *Drosophila* have also provided evidence for involvement

*To whom correspondence should be addressed. Tel: 97248293456; Fax: 97248225153; Email: manor@tx.technion.ac.il

of Translin–TRAX complexes in RNA interference (18,19). Thus, Translin and TRAX appear to be multifunctional proteins.

The 3D structures of the individual subunits of the mouse and the human Translins, both of which consist of 228 amino acids, were solved by X-ray crystallography at resolutions of 2.65 and 2.2 Å, respectively, and were found to be virtually identical (20,21). Both crystal structures contain seven α helices that constitute >70% of the amino acid residues. The crystals had tetramers as asymmetric units. By use of symmetry operations, it has been deduced that the actual crystallized complexes are those of octamers. However, the deduced structures were not consistent with the assembly of the subunits into octameric rings, as seen in the electron microscopic studies (5).

Translin and TRAX were found to be conserved in the fission yeast *Schizosaccharomyces pombe*, but not in the budding yeast *Saccharomyces cerevisiae*. Bearing in mind that the fission yeast system is most suitable for genetic and biochemical analysis of protein function, we and others have initiated studies of the *S. pombe* Translin (spTranslin) and the *S. pombe* TRAX (spTRAX). Deletion of the genes encoding either of the two proteins, or both of them, indicated that these genes are not essential for growth of the fission yeast cells (22,23). Studies of the nucleic acid binding characteristics of recombinant spTranslin revealed that, like the human recombinant Translin, octamers of the yeast protein, but not smaller complexes, bind G-rich nucleic acids. However, unlike the human Translin, which binds single-stranded G-rich RNA and DNA with approximately equal affinities, the yeast protein has a much higher affinity for RNA sequences than for the corresponding sequences of single-stranded DNA (22). These data suggested that spTranslin is primarily engaged in functions related to RNA metabolism.

To gain a further insight into the interactions between spTranslin and nucleic acids, we used several bioinformatics techniques to obtain a structural model of the protein and to search for potential nucleic acids and protein binding surfaces on this structure. In parallel, we generated a series of amino acid substitution mutants of spTranslin and used these mutants to experimentally test the bioinformatics predictions. This combined approach led us to identify and map, for the first time, a nucleic acid binding region on the protein surface. Here we report these studies, as well as data obtained by a similar approach on interactions between individual spTranslin subunits, and between spTranslin and spTRAX. We also show that the information gained in these studies most likely applies to the human Translin and TRAX as well.

MATERIALS AND METHODS

Materials

Oligodeoxyribonucleotides, oligoribonucleotides and [γ-^{32}P]ATP were purchased from Sigma, Dharmacon and New England Nuclear, respectively. The plasmid pREP41-NTAP was kindly sent to us by Dr K.L. Gould of the Howard Hughes Medical Institute, Department of Cell and Developmental Biology, Vanderbilt University School of Medicine (24).

Site-directed mutagenesis

Mutations in plasmids encoding spTranslin were generated by using the QuikChange site-directed mutagenesis kit [Stratagene, La Jolla, CA), as described in ref. (25)]. The sequence of the whole spTranslin gene was determined in all mutants to ensure that mutations only occurred in the desired locations.

Expression and purification of wild-type and mutant spTranslins synthesized in *E. coli* cells

The ORFs encoding the recombinant proteins were all inserted in frame in the bacterial expression vector pQE31 (QIAGEN), such that when expressed, they were tagged with N-terminal hexahistidine. The proteins were affinity-purified in Ni–agarose columns, as described (11).

Gel mobility shift assays and calculations of apparent Kd values

Gel shift assays of binding of wild-type or variant spTranslins to single-stranded RNA or DNA probes were carried out as previously described (11,22). In each assay, the concentration of the nucleic acid probe was kept constant, while the concentration of the protein was varied. The data were plotted as γ = bound probe/total probe versus P_{tot}, the total protein concentration. The data obtained for most of the substitution variants were fitted to the equation $\gamma = b \times P_{tot}/(K_d + P_{tot})$, from which estimates of apparent K_d values were extracted. However, calculations of K_d values by this equation were meaningless for assays in which γ was smaller than 0.30 at the maximal protein concentration of 200 nM and in which γ did not reach a plateau level. Moreover, the measurements of γ were rather inaccurate at these low levels. Hence, in these assays, the apparent K_d values were estimated to be $K_d > 1/2 \times 200 = 10^2$ nM.

Propagation of *S. pombe* cells and DNA transformation into the cells

The procedures used for propagation of *S. pombe* cultures were as previously described (22). Details on the DNA transformation procedure are presented in the Supplementary data.

Expression of N-terminal TAP-tagged wild-type and mutant spTranslins in *S. pombe* cells

We cloned the wild-type and mutant spTranslin ORFs into the plasmid pREP41-NTAP downstream of, and in frame with, the TAP-tag encoding segment (24). Thereby, we created plasmids which expressed N-terminal TAP-tagged spTranslins under the control of the nmt1 promoter. These plasmids were transfected into *S. pombe* cells, in which the endogeneous spTranslin gene had been deleted (22). In the absence of thiamine, the TAP-tagged spTranslins were overexpressed in these cells at a rate that was estimated to be at least 50 times

higher than the expression rate of the endogeneous protein in a wild-type strain of *S. pombe*.

Two-step purification and proteomic analysis of TAP-tagged spTranslin and associated proteins

The preparation of *S. pombe* cell extracts, as well as the two-step purification procedure and proteomic analysis, are described in the Supplementary Data.

Yeast two-hybrid assays

The cDNAs encoding the wild-type spTRAX, the wild-type spTranslin and the spTranslin variant proteins, were cloned into the pGAD-C and pGBD-C vectors (26). Pairs of binding domain and activation domain constructs were cotransfected into the *S. cerevisiae* strain Y1183. Protein–protein interactions were detected by growth on agar plates containing SD medium, as described (22).

Western blotting

Western blotting was carried out as previously described (22).

Bioinformatics methods and algorithms

Modeling. We used 11 techniques to produce 18 models of spTranslin based on the human Translin homologue. The techniques are denoted: FUGUE, HHpred, HHSEARCH, I-TASSER, PAINT, PPA_I, PPA_II, PROSPECT2, SAMT02, SP3 and SPARKS2, and are described in ref. (27). We also tried several improvements to the models, using the MODLOOP (28) and SCWRL3 (29) engines for loop modeling and rotamer prediction, respectively. All these algorithms yielded similar models, with RMSD <0.8 Å (for CA atoms) measured for any pair. Discarding very similar models, we chose a final list of 14 models for further analysis (see Supplementary Figure S2).

ConSurf: calculation of residue conservation. The ConSurf server allows the identification of functionally important regions on the surface of a protein of known 3D structure, based on the phylogenetic relations between its close sequence homologs (30).

Optimal protein–RNA area (OPRA): RNA-binding predictor. OPRA is a protein–RNA predictor based entirely on pre-calculated statistical propensities from existing PDB complexes. Patches found on the protein surface are generally indicative of RNA interactions (31).

Patch finder plus (PFP): calculation of positive electrostatics patches. The PFP method is designed to map the largest continuous positive patch on a protein surface. Patches of positive charges are characteristic of protein–nucleic acid interfaces (32).

Optimal docking area (ODA): protein–protein interface predictor. ODA predicts areas of putative protein–protein interaction on a protein surface based on atomic solvation parameters. Areas of low docking surface energy represent regions likely to be buried after interaction with other proteins (33).

Scores obtained from 14 structural models. We calculated a single PFP, OPRA and ODA score for each amino acid in spTranslin, using the 14 models mentioned above. For PFP, a score was assigned according to the number of appearances of each residue on any electrostatic patch in any of the models: score 0 (zero appearances), score 1 (one appearance), score 2 (two appearances), score 3 (three or four appearances), score 4 (five or six appearances). For OPRA, we divided the average scores obtained for each residue, as follows: score 0 (OPRA average value >0), score 1 (OPRA average value from 0 to −0.2), Score 2 (from −0.2 to −0.4), score 3 (from −0.4 to −0.8), score 4 (from −0.8 to −1.2). For ODA, we divided the average scores obtained for each residue, as follows: score 0 (ODA average value > 0), score 1 (ODA average value from 0 to −2), score 2 (from −2 to −5), score 3 (from −5 to −9), score 4 (−9 and lower). In the case of ConSurf, the conservation values are calculated on the basis of the sequence alignment and are, therefore, the same for all models. In this work, only residues with scores 3 and 4 are defined as positive predictions for the OPRA, ODA and PFP methods, while residues with scores 8 and 9 for ConSurf are considered highly conserved.

RESULTS

Initial mapping of RNA binding sites in spTranslin based on sequence alignment data

Figure 1 presents the spTranslin sequence and the degree of conservation for each of the amino acid residues. The conservation was calculated by the ConSeq server (34), with a multiple sequence alignment of 36 unique Translin orthologues (see the Supplementary Figure S1). Also indicated in Figure 1 are two underlined regions that we designate here as mammalian basic region I (MBRI) and II (MBRII). These spTranslin regions align with two basic motifs in the human and the mouse Translins that were reported to be involved in nucleic acids binding (35,36). It can be seen that while the average conservation of MBRII, particularly the YYKYN motif (residues 93–97), is above the general average for the protein, the MBRI region is highly variable, with the exception of E68.

We began our mapping of the RNA binding sites in spTranslin by generating single amino acid substitution mutants in MBRII and MBRI. The wild-type and the mutated proteins, which were tagged with a his-tag, were synthesized in *E. coli* cells and purified by affinity chromatography in imidazole columns. Binding of the proteins to oligonucleotides was assayed by the gel shift technique, as previously described (22). Figure 2 shows binding data obtained in assays of wild-type spTranslin (panel A) and of two substitution variants within MBRII, W100A and R102G (panels B and C). The probe was the ^{32}P-labeled oligoribonucleotide $(GU)_{12}$; this oligo has a relatively high affinity for spTranslin and does not contain secondary-structure features, or unusual structures that may affect the binding (11,22). In the binding mixtures,

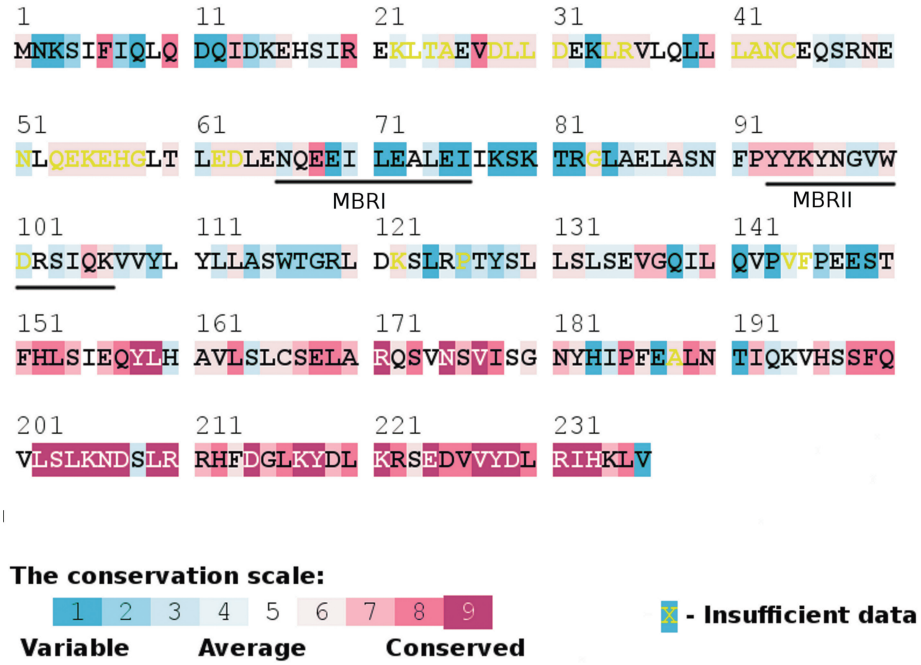

Figure 1. spTranslin sequence and residue conservation based on a multiple sequence alignment of 36 Translin orthologues. The alignment used to calculate the conservation scores is presented in Supplementary Figure S1 of the Supplementary Data.

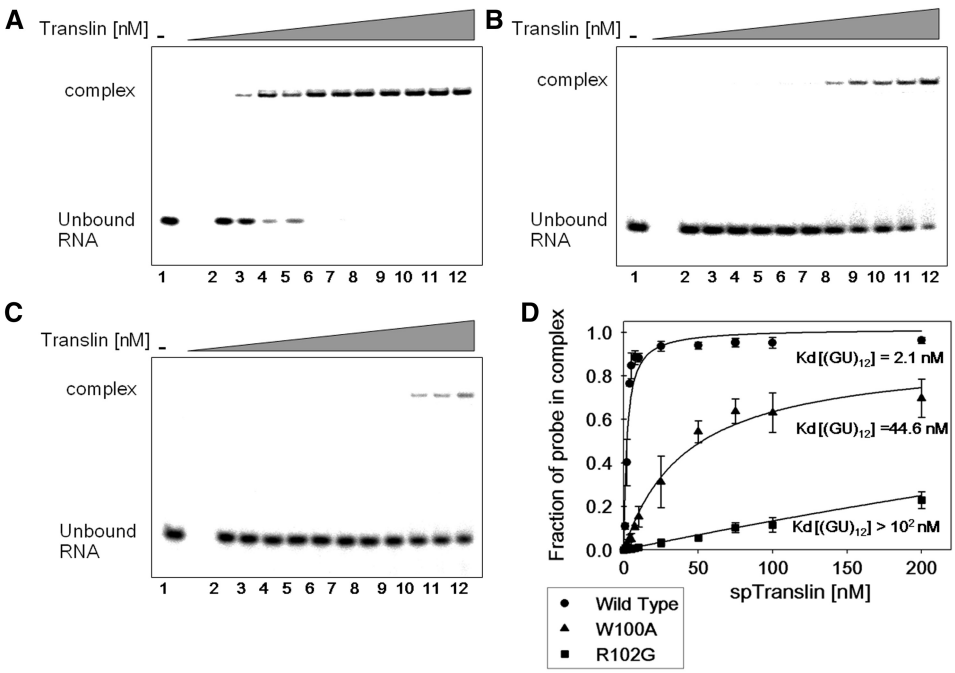

Figure 2. Binding of wild-type and variant spTranslins to the $(GU)_{12}$ (RNA) probe. The oligoribonucleotide $(GU)_{12}$ was end-labeled with $^{32}PO_4$. The labeled probe at a concentration of 0.01 nM was incubated with increasing concentrations of the wild-type spTranslin and of each of the variant spTranslin proteins encoded by the indicated substitution mutants. The complexes were resolved from the unbound RNA probe by polyacrylamide gel electrophoresis, as described in 'Materials and Methods' section. (**A–C**) The gel profiles obtained in the assays of the wild-type, W100A and R102G proteins, respectively. (**D**) Plots of the data shown in these gels and of data obtained in three additional independent experiments, with the standard deviations. The average K_d values calculated from these data are also shown. Lane 1: no protein. Protein concentrations in the assays shown in lanes 2–12 were (in nM): 1, 2, 4, 5, 7.5, 10, 25, 50, 75, 100 and 200.

Table 1. Experimental data bioinformatics predictions regarding nuleic acids binding to sp Translin

Substitution	Corresponding amino acida in human Tsn	Helix	Region	K_d [$(GU)_{12}$] (nM)[a]	K_d [$d(GT)_{15}$] (nM)[a]	Conservation score[b]	PFP (cumulative score)	OPRA (Kcal/mol)[c]
Wild-type	–	–	–	2.1 ± 0.2	16.5 ± 1.7	–	–	–
H17A	Q18	1		1.8 ± 0.3	16.6 ± 0.6	5	0	−0.14
R20G	R21	1		$>10^2$	$>10^2$	8	2	−0.14
E21A	E22	1		1.8 ± 0.7	4.1 ± 0.2	5	1	−0.19
R35G	R36	1		$>10^2$	$>10^2$	6	4	−0.17
Q38A	L39	1		4.3 ± 0.5	69.9 ± 10.2	5	3	−0.03
Q67A	R57	2	MBRI	3.1 ± 1.6	14.8 ± 0.9	6	0	0.15
E68A		2	MBRI	2.1 ± 0.5	12.6 ± 1.9	8	0	0.75
E72A	K60	2	MBRI	2.8 ± 0.1	16.9 ± 0.4	1	0	0.49
E75A	E63	2	MBRI	3.5 ± 0.5	15.8 ± 1.6	1	0	0.32
K80A	V68	2		2.2 ± 0.6	15.0 ± 3.2	1	4	0.08
Y93A	Y84	loop2-3	MBRII	3.2 ± 1.3	16.7 ± 4.9	7	3	−0.1
Y94A	Y85	loop2-3	MBRII	24.8 ± 8.7	92.4 ± 3.6	7	3	−0.31
Y94S	Y85	loop2-3	MBRII	30.4 ± 11.4	80.4 ± 15.8	7	3	−0.31
K95A	R86	loop2-3	MBRII	$>10^2$	$>10^2$	7	4	−0.33
Y96A	F87	loop2-3	MBRII	14.2 ± 2.8	35.5 ± 7.1	6	3	−0.22
Y96S	F87	loop2-3	MBRII	17.2 ± 3.2	43.3 ± 2.8	6	3	−0.22
N97A	H88	loop2-3	MBRII	29.5 ± 8.7	50.0 ± 7.3	6	1	−0.42
W100A	W91	loop2-3	MBRII	44.6 ± 11.6	$>10^2$	6	3	−0.16
D101A	-	3	MBRII	2.4 ± 0.4	9.1 ± 5.1	3	0	−0.34
R102G	R92	3	MBRII	$>10^2$	$>10^2$	5	4	−0.37
S103A	V94	3	MBRII	3.5 ± 0.6	16.3 ± 3.4	3	3	−0.15
Q105A	Q96	3	MBRII	$>10^2$	$>10^2$	7	3	−0.4
K106A	R97	3	MBRII	$>10^2$	$>10^2$	6	4	−0.48
E168A	E150	5		2.4 ± 0.9	2.6 ± 1.1	8	1	−0.14
R171G	R153	5		73.4 ± 7.6	$>10^2$	9	0	−0.17
R210G	R192	7		$>10^2$	$>10^2$	9	2	−0.82
R211G	K193	7		$>10^2$	$>10^2$	7	3	−1.07
K221G	K203	7		2.2 ± 0.5	18.5 ± 3.9	9	0	−0.57

[a]See 'Materials and Method' section for details of the K_d calculations.
[b]Calculated from multiple sequence alignment.
[c]Average score for 14 models.

the concentration of the protein varied, while the concentration of the probe was kept constant. Panel D shows a graphic representation of these data combined with data obtained in other independent assays. Evidently, unlike the wild-type protein, the variant proteins did not completely titrate the probe at the concentrations used for these assays. Nevertheless, apparent K_d values were calculated, as described in 'Materials and Methods' section. It can be seen that the W100A and the R102G substitutions caused a 21- and >48-fold increase, respectively, in the apparent K_d value. These results provide a clear evidence for the involvement of these amino acid residues in binding the RNA probe. Table 1 presents a summary of these data and similar data obtained for other substitutions. Each assay was carried out at least in duplicate and the average K_d's and standard deviations are indicated. It can be seen that, in addition to the W100A and the R102G mutations, substitutions of seven other amino acids included in MBRII with either alanine, or glycine, or serine, also caused substantial increases in the K_d values. The largest K_d increases were observed for the protein variants in which the amino acids K95, R102, Q105 and K106 have been substituted. On the other hand, substitutions of Q67, E68, E72 and E75, which map within MBRI, did not cause significant increases in the K_d values, and hence, are apparently not involved in RNA binding.

Modeling and functional residues prediction

Protein structures contain more information than sequences. Hence, we built and analyzed comparative based models of the spTranslin structure based on its human homolog. Then, we employed these models in an attempt to locate the position of the spTranslin–nucleic acids interaction region on the structure of the protein.

Comparative modeling of spTranslin. We have previously built a sequence-dependent homology-based model of the spTranslin subunit, using the solved structure of the human Translin as the template (22). In the present work, we have undertaken to explore more deeply the modeling of Translin, using several techniques (see 'Materials and methods' section). We created a total of 18 models with 11 different techniques, from which we selected for further analysis 14 models having the best ModFold server scores (37) (Supplementary Figure S2). All 14 models are similar to our previously published model (22), having seven α helical regions, which we designated as α1 to α7. Model5.pdb generated by the I-TASSER server (38) received the highest ModFold score (see Supplementary Figure S2).

Prediction of conserved regions. Figure 3A and E shows the model5.pdb of spTranslin color-coded by ConSurf

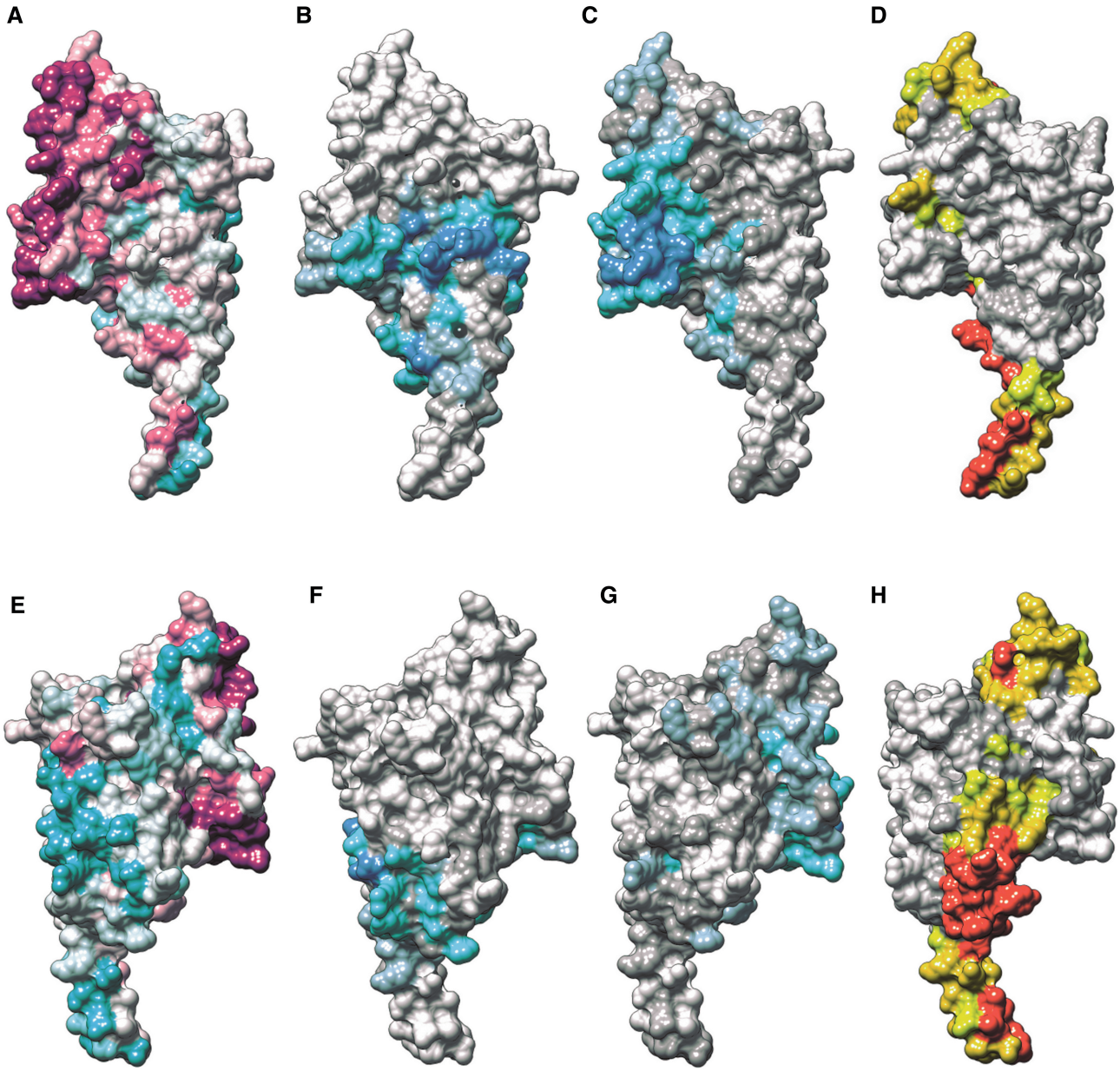

Figure 3. Three-dimensional-structural model of the spTranslin monomer (model5.pdb) displaying scores calculated by ConSurf, PFP, OPRA and ODA. The first row shows the front face of spTranslin. The second row shows the back face of spTranslin. (**A** and **E**) The residues conservation calculated by ConSurf. The conservation scale is shown in Figure 1. (**B** and **F**) The cumulative positive electrostatic patch found by PFP. (**C** and **G**) The average OPRA scores for RNA binding. In panels B, F, C and G, the color code is white / grey / light blue / cyan / blue for OPRA scores 0–4, respectively. (**D** and **H**) average ODA scores for protein–protein binding prediction. Color code in these panels is white /grey/yellow/orange/red for ODA scores 0–4, respectively. In panels B, F, C, G, D, H, residues with scores 3 and 4 are predicted to interact with nucleic acids (PFP), RNA (OPRA) and protein partners (ODA). In panels A and E, ConSurf scores 8 and 9 are considered highly conserved.

evolutionary conservation scores (obtained as in Figure 1). Panel A shows one face of the protein designated here as the front face. ConSurf predicts a highly conserved cluster of amino acids on the left upper corner of this face (in maroon color). In particular, the helix α7 and the loop connecting it with the helix α6 include the majority of the highly conserved residues in spTranslin. Hence, this region of the protein is predicted to have an important functional role. Additionally, the front face presents a second much smaller conserved region including several amino acids on helices α1 and α3, which may also have a functional role. Panel E shows the opposite face of the protein, the rear face, which has a significantly lower level of conservation (cyan colors). It should be noted, however, that while conserved regions in proteins are likely to be functionally significant (39,40), their conservation per se does not provide information on the type of functions which they perform. To predict function we used several additional computational tools, as described in the next sections.

Prediction of nucleic acid interaction regions. We used two different computational methods to predict putative regions of nucleic acid binding on the surface of spTranslin. The first method, designated PFP, is based on the premise that large positive electrostatic regions on protein surfaces are likely to interact with nucleic acids (32). We used the PFP server to find the largest continuous positive electrostatic patch on the surface of each of the 14 homology-based models of spTranslin (41). As shown in Supplementary Figure S3, different models produced quite different results. Presumably, this variation is due to model-specific errors. Assuming that the patch obtained for each individual model represents only a portion of the real patch, we then calculated an accumulative score for each residue in the 14 models (see 'Materials and methods' section) and present these scores on the model5.pdb produced by I-TASSER. The cumulative patch, which includes a total of 93 residues, is shown in Figure 3B (front face) and F (rear face). We surmise that the residues appearing in patches found in several models (scored 3 or 4 in Figure 3) have the highest probability of being involved in nucleic acid interactions. Comparison of Figure 3A and B reveals that most residues predicted to be involved in nucleic acid binding are not highly conserved and some of them are even moderately variable. This observation is in line with previous studies on nucleic acid interaction surfaces in proteins (32,42). We shall show in the experimental part below that many of the highest score residues in the cumulative patch of spTranslin do, in fact, play a role in nucleic acid binding.

We have also used a second method called OPRA to identify putative nucleic acid binding residues on the models of spTranslin (31). OPRA is a knowledge-based method that predicts protein–RNA interfaces based on pre-calculated statistical propensities for residues that form protein–RNA interfaces. In this case too, a patch was calculated for each of the 14 models (Supplementary Figure S4). Then, an average patch was calculated and is shown on the model 5.pdb produced by I-TASSER (Figure 3C and G). Combining the results obtained by OPRA and PFP (Figure 3B, C, F and G), it can be seen that the strongest predictions (in blue) for putative nucleic binding areas lie on the equatorial region of the front face of the spTranslin structure. However, whereas the PFP highest scored residues are mainly found on the moderately conserved right side of the equatorial region (Figure 3B), the OPRA highest scored residues (Figure 3C) reside on the more highly conserved left side of the equatorial region. Thus, the combined predictions of OPRA and PFP suggest that many residues in the equatorial region of the front face are involved in nucleic acid binding.

A more detailed inspection of the OPRA and PFP results reveals that the patches predicted by both algorithms are unusually rich in arginine residues (see also Figures 4 and 8). The OPRA highest scores, which peak at the left corner of the patch, encompass three arginine residues: R210 and R211 (score 4) and R222 (score 3). The PFP-predicted highest scores, which peak on the right side of the equatorial patch, include residues R35 and R102 (score 4) and R211 (score 3). The four highest scored

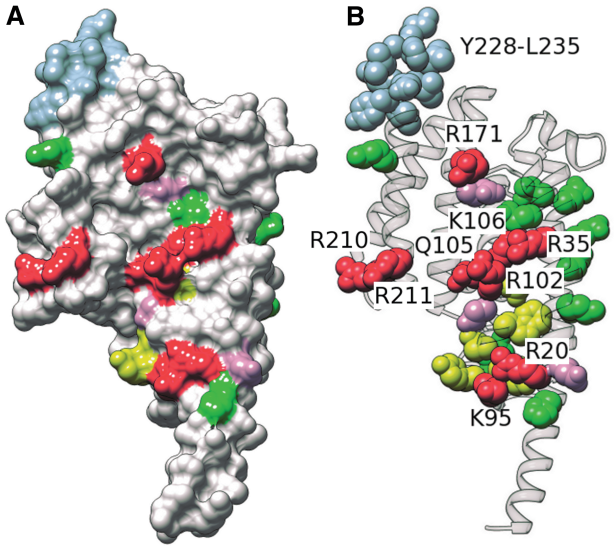

Figure 4. Space filling (**A**) and ribbon representation (**B**) of the spTranslin front face displaying the substitution and truncation mutations generated and analyzed in the present study. Substitution of residues stained in red caused the strongest reduction in the affinities towards the RNA or DNA probe (see Table 1 for details). Substitution of residues stained in yellow and green caused partial and no inhibition, respectively, in binding to the probes. Substitution of residues colored violet caused an increase in the affinity for single-stranded DNA. The light blue amino acid residues at the carboxy terminus were missing from the truncation variant designated Y228stop (see Table 2 and the text).

arginines in both methods, R35, R102, R210 and R211 were, in fact, experimentally observed to play a significant role in nucleic acid binding (see the next section of the Results and Table 1). It should also be noted that there is an overlap of both predictions in the lower part of their respective patches. This overlap includes residuesY94 and K95 (in light blue and cyan), which have lower overall scores, but were also shown experimentally to be involved in nucleic acid binding. Finally, it is noteworthy that the two methods are subject to different types of errors. The PFP predictions are highly dependent upon the model structures, and are, therefore, subject to modeling errors. OPRA is much less dependent on the specific model geometries and is, therefore, less sensitive to modeling errors (see Supplementary Figures S3 and S4). However, the OPRA scores are based on limited statistics of existing non-redundant PDB complexes of protein–RNA, while proteins related to Translin may be underrepresented in the OPRA statistics. Thus, the strength of our predictions stems from the use of these two different bioinformatics methods and 14 models. Indeed, the predictions seem to agree well with experimental data presented in the section 'Experimental verification of the bioinformatics predictions regarding the functional regions in spTranslin'.

Prediction of protein-protein interaction regions. We have previously reported a series of yeast two-hybrid assays, which indicated that, like the human and the mouse Translin (8,36), spTranslin can homodimerize and can also form heterodimers with spTRAX (22). In addition

to these experimentally established protein–protein interactions, spTranslin and spTRAX may interact with other proteins, as expected from their presumed multifunctional nature (see the 'Introduction' section). It is, therefore, of interest to identify putative protein interaction sites in spTranslin. To this end, we used ODA, a method for the prediction of sites involved in protein–protein interactions. Figure 3D and H shows averaged ODA results for the 14 models, while ODA calculations for each model are presented in Supplementary Figure S5. As can be seen, ODA identifies two regions on spTranslin, which are likely to be engaged in protein–protein interactions. The strongest signal (red and orange colors) is found at the bottom of the structure, encompassing residues from the α1 helix (N-terminus), the loop connecting α2 and α3 helices and the 144–154 loop. This region is moderately conserved (compare to Figure 3A). A second smaller, but quite significant ODA signal is found at the highly conserved region around the top of the spTranslin model and includes several residues from the helices α7, α6 and α5. Interestingly, the ODA and the OPRA/PFP predictions cover non-overlapping areas on the front face of the spTranslin structure.

Experimental verification of the bioinformatics predictions regarding the functional regions in spTranslin

Mapping the nucleic acid binding region. In the section 'Initial mapping of RNA binding sites in spTranslin based on sequence alignment data', we described the substitution mutations made in spTranslin on the basis of the sequence alignment data, and determined the affinities of the variant proteins towards RNA. Based on the bioinformatics predictions presented in section 'Modeling and functional residues prediction', we made additional substitutions and assayed the binding of the variant proteins to the RNA probe, $(GU)_{12}$, using the gel shift procedure (Figure 2). The substituted amino acids were mainly positively charged residues, Arg or Lys (see, e.g. the highly conserved residues R210, R211, which are predicted by OPRA to bind RNA, as seen in Figure 3, but are not included in MBRII or MBRI). Figure 4A displays all the substitutions that can be seen on the front face of the protein. The color codes illustrate the effects of these substitutions on binding of the variant proteins to the RNA probe. The residues stained in red are those whose substitution strongly inhibited the binding of the probe. The residues stained in yellow and in green, respectively, are those whose substitution caused a partial inhibition, or no significant inhibition of the binding, respectively. Figure 4B shows a ribbon representation of the spTranslin structure, on which the substituted residues are clearly seen relative to the seven α helices of the protein and the loops between them. By inspection of panels A and B, it is also possible to see which mutated residues are buried in the protein structure and which residues are exposed on the faces of the protein.

As already noted in the 'Introduction' section, we have previously shown that spTranslin binds single-stranded DNA, as well as RNA, though its affinity to DNA is much lower (22). Therefore, we also examined the effects of many of the substitutions on DNA binding. Figure 5

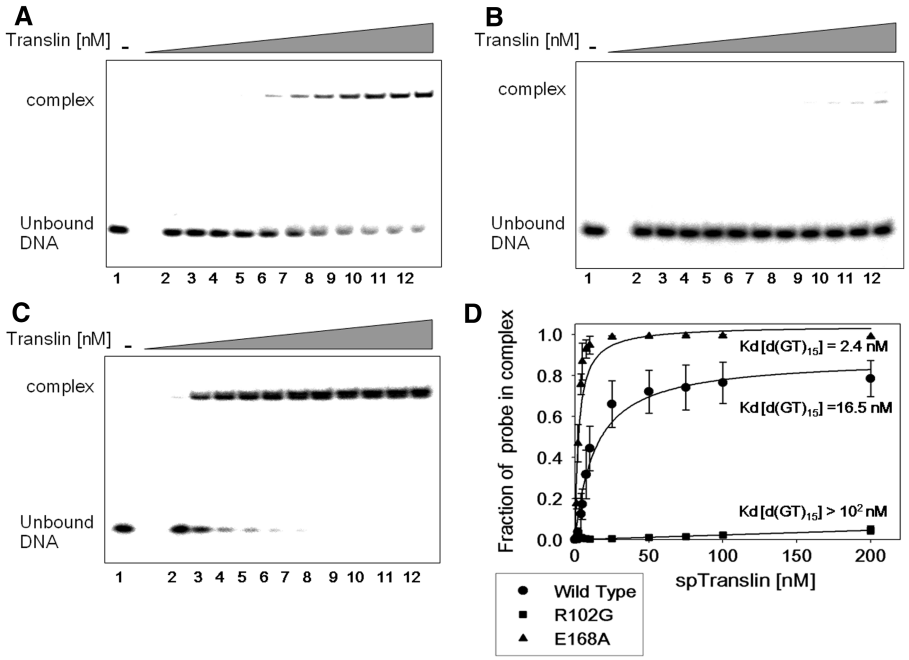

Figure 5. Binding of wild-type and variant spTranslins to the single-stranded $d(GT)_{15}$ (DNA) probe. The oligodeoxyribonucleotide $d(GT)_{15}$ was end-labeled with $^{32}PO_4$. The labeled probe at a concentration of 0.01 nM was incubated with increasing concentrations of the wild-type spTranslin and each of the variant spTranslin proteins encoded by the indicated substitution mutants. The complexes were resolved from the unbound single-stranded DNA probe by polyacrylamide gel electrophoresis, as described in 'Materials and Methods' section. (**A–C**) The gel profiles of the wild-type, R102G and E168A variants, respectively. (**D**) Plots of the data presented in these gels and the data obtained in additional two independent experiments. The calculated K_d values are also indicated. Lane 1: No protein. Protein concentrations in the assays presented in Lanes 2–12 were the same as in Figure 2.

shows the results of gel shift binding assays of wild-type spTranslin and two substitution variants, performed with a $d(GT)_{15}$ oligodeoxyribonucleotide probe. It can be seen that the K_d value of the wild-type spTranslin for the DNA probe was seven times larger than the K_d value for the RNA probe. This difference in the K_d values is about the same as the previously observed difference (22). It is also seen in Figure 5 that one of the substitutions, R102G, almost completely eliminated the binding to the DNA probe, as it did for the RNA probe (Table 1). However, another substitution, E168A, caused a 7-fold increase in the affinity to the DNA probe, such that it became almost as large as the affinity of the wild-type protein to the RNA probe. As shown in Table 1, this mutation did not significantly alter the affinity of the protein to the RNA probe.

The entire collection of substitution mutations that we have made in spTranslin are listed in Table 1 along with the experimentally determined apparent K_d values for binding RNA and DNA, as well as the conservation, PFP and OPRA scores. Based on the bioinformatics predictions and the experimental data, as summarized in Table 1, we draw the following conclusions:

- Many substitutions that affect the binding of single-stranded RNA and DNA to spTranslin map within the motif MBRII shown in Figure 1, but none map within MBRI. This observation is compatible with the predictions based on the spTranslin structure; for the amino acid residues that constitute MBRI are situated on the rear of the structure, and far away from the PFP or OPRA patches. Moreover, the 11 residues in MBRI have an average conservation degree of four, which indicates a variable region. On the other hand, the 13 residues that constitute MBRII are situated on the front of the structure and have an average conservation value of almost six, which indicates a moderately conserved region. Additionally, all MBRII residues are found at either the OPRA, or PFP patch; eight of these residues have high scores, a strong indication of their involvement in nucleic acid binding.
- There are five residues, R20, R35, R171, R210 and R211, which are not included in MBRII, but whose substitution strongly affects binding of spTranslin to both the RNA and the DNA probes. In the 3D structure of the protein, residues R20 and R35 are adjacent to amino acids included in MBRII. R210 and R211 are highly conserved and are predicted to interact with RNA by OPRA. R211, as well as R35, are predicted by PFP to bind nucleic acids. However, the residues R20 and R171 are below our cutoff points in both predictions. These and other apparent discrepancies between the bioinformatics predictions and the experimental results are further discussed below.
- Substitutions of the amino acids Y94, Y96, N97 and W100, which are buried within the spTranslin structure (visible in Figure 4B, but not in Figure 4A) partially affect the binding of RNA and DNA to the protein (yellow residues). These results might be due to changes in the structure of the protein, rather than to direct effects on nucleic acid binding.
- The highest degree of inhibition of RNA and DNA binding resulted from substitutions of the basic arginine and lysine residues found within, or near the PFP and the OPRA patches shown in Figure 3B and C. These amino acids surround a shallow cavity on the protein surface, which is also present in the human structure [see section 'Bioinformatics predictions of the locations of the nucleic acid and protein–protein binding regions on the human Translin (hTranslin)']. The possible involvement of this cavity in binding of nucleic acids is discussed below (see the 'Discussion' section).
- Substitutions of some (basic) lysine residues, such as K80 and K221, with alanine did not substantially affect RNA or DNA binding despite their being included in the PFP and OPRA predictions, respectively. Thus, even though the bioinformatics predictions generally agree with the experimental results on the involvement of amino acids in nucleic acid binding, there are exceptions. Such exceptions might be due to inaccuracies in the modeled structure of spTranslin and/or in the bioinformatics prediction methodologies.
- Most substitutions that caused a reduction in RNA binding also led to reduction in DNA binding. However, the substitutions E168A, E21A and possibly also D101A, which did not substantially affect RNA binding, caused an increase in DNA binding (the three residues are stained in violet in Figure 4). A plausible explanation of these results will be presented in the 'Discussion' section.

Taken together, the data listed in Table 1 indicate that most residues of spTranslin, which play a role in nucleic acid binding, are clustered in and around the PFP and the OPRA patches on the front face of the protein (Figure 3B and C). Many of these residues are basic, mostly arginines, or polar (Figure 8).

Experimental verification of the involvement of the conserved carboxy terminus of spTranslin in protein–protein interactions. As indicated in section 'Modeling and functional residues prediction' above, the highly conserved C-terminus of spTranslin is predicted to be involved in protein–protein interactions. To experimentally test this prediction, we generated two mutants with stop codons inserted after the amino acid residues with position numbers 204 and 228, respectively. As a result, it was expected that these mutant genes would encode truncated proteins which would lack the terminal 32, or 8 residues. We cloned these mutant spTranslin genes, as well as the wild-type spTranslin gene, and the substitution mutant K106A, into a plasmid containing the TAP-tag (24,43). The resulting constructs were transfected into *S. pombe* cells, in which the endogeneous spTranslin gene had been deleted. Then, we performed a two-step TAP-tag affinity purification of each of the tagged spTranslin proteins. SDS–polyacrylamide gel electrophoresis of the purified fractions, followed by a western analysis, revealed that the truncated proteins

Table 2. Association of spTRAX with TAP-tagged wild-type and variant sp Translins[a]

Protein	Mass (k Da)	TAP–sp Translin (WT)		TAP–sp Translin (K106A)		TAP–sp Translin (S204stop)[b]		TAP–sp Translin (Y2228stop)[b]		TAP alone	
		No. of peptides	Percentage of coverage	No. of peptides	Percentage of coverage	No. of peptides	Percentage of coverage	No. of peptides	Percentage of coverage	No. of peptides	Percentage of coverage
Sp Translin	27.3	16	61	22	56.8	11	47	8	29.7	–	–
Sp TRAX	26.2	11	52	7	47.7	–	–	–	–	–	–

[a]Determined by two-step affinity purification, followed by proteomic analysis.
[b]deletion variants generated by substitution of the codons for the specified residues with stop codons.

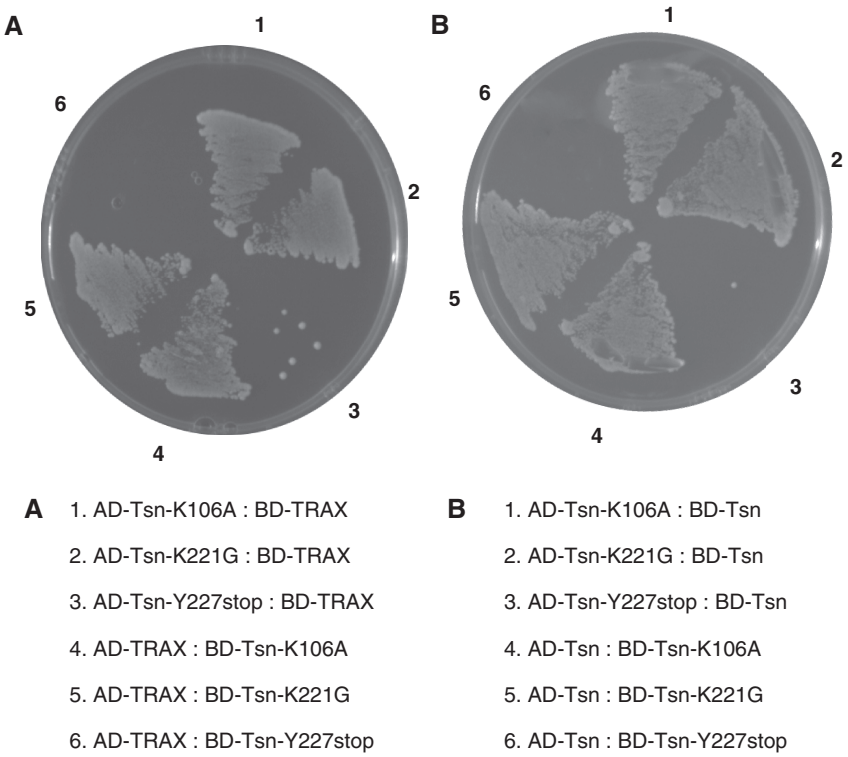

Figure 6. Yeast two-hybrid assays of heterodimerization of wild-type spTranslin and spTranslin variants with spTRAX (**A**), and of homodimerization of the corresponding spTranslin monomers (**B**). K106A and K221G are substitution variants. Tsn-Y227stop is a truncation variant, in which eight amino-acid residues at the carboxy-terminus are missing (see the text for further details). AD, activation domain; BD, binding domain.

had the expected faster mobilities (not shown). Finally, we used mass spectrometry to analyze the proteins in the purified fractions (44) (see the Supplementary data for details). As Table 2 shows, spTRAX was clearly copurified with the wild-type spTranslin and with the K106A variant. However, no spTRAX was found to copurify with either truncation mutant of spTranslin. A similar procedure was performed with *S. pombe* cells expressing the TAP-tag alone. As expected, neither spTranslin nor spTRAX were detected in these cells. We conclude that both truncations eliminated the association between spTranslin and spTRAX. The K106A substitution causes a severe reduction in the affinity of spTranslin to RNA and DNA, as shown in Table 1. Therefore, the ability of this variant K106A protein to bind spTRAX as efficiently as the wild-type protein (Table 2) indicated that the interaction between the two proteins is not dependent upon the ability of spTranslin to also bind nucleic acids. Although not shown here, up to 200 additional proteins were also identified in the purified spTranslin fraction in each individual TAP-tag assay. However, so far we have not been able to unequivocally demonstrate that the association with any protein other than spTRAX is specific and that it is lost in the assays performed with the variant spTranslins.

We have also used the yeast two-hybrid technique to test the ability of the spTranslin truncation variants, as well as the K106A and K221G substitution variants, to bind spTRAX. Figure 6A shows reciprocal two-hybrid assays performed with the wild-type and the variant spTranslins being the bait and spTRAX being the catch, or vice versa. It can be seen that the interaction between

the two proteins has been eliminated by the truncations, but not by the substitutions. Figure 6B shows similar assays of the ability of spTranslin to dimerize. Evidently, the truncations, but not the substitutions, eliminated the dimerization of spTranslin. Hence, the carboxy-terminus of the protein is required for the interactions between spTranslin subunits, as well as the interactions between spTranslin and spTRAX. Both interactions do not depend on the ability of spTranslin to bind nucleic acids. Thus, two different experimental tests provide support for the bioinformatics prediction that the carboxy-terminal region of spTranslin is involved in protein–protein interactions. The evolutionary requirement for these interactions is apparently very strict, as indicated by the high degree of conservation of this region of the protein. So far, we have not subjected to similar experimental tests the N-terminal part of the protein, which was also predicted by our bioinformatics analysis to be involved in protein–protein interactions.

Bioinformatics predictions of the locations of the nucleic acid and protein–protein binding regions on the human Translin (hTranslin)

To assess the generality of the results of our combined molecular genetics and bioinformatics study of spTranslin, we used the bioinformatics tools described above to predict the likely positions of the nucleic acid and protein–protein binding regions in the hTranslin template structure [PDB 1j1j (21)] These predictions are expected to be more accurate than those made for spTranslin, because they are based on the high-resolution crystallographic structure of hTranlin, rather than on comparative models.

Prediction of the location of the nucleic acid binding region on hTranslin. Figure 7A and B present PFP and OPRA predictions for the position of the nucleic acid binding region on the front face of the hTranslin PDB structure (1j1j). The PFP calculation (Figure 7A) reveals two large positive electrostatic patches on this face of the protein. The larger patch (in blue) corresponds approximately in its position and shape to the joined patches predicted by PFP and OPRA in spTranslin (compare Figure 7A to Figure 3B and C). It is also evident that the back face of hTranslin, like that of spTranslin, lacks large electrostatic patches (compare Figure 7D to Figure 3F). The OPRA tool produces similar results (Figure 7B and E). Hence, despite the sequence differences and modeling errors, it appears that the nucleic acid binding regions of hTranslin and spTranslin map at corresponding positions in their 3D structures. Furthermore, like the spTranslin, the putative nucleic acid binding region of the human Translin includes many positively charged amino acids, mostly arginine residues that surround a shallow cavity (see Figure 8 and the 'Discussion' section).

The protein–protein interaction regions on hTranslin. The ODA calculation predicts that protein–protein interaction regions reside at the top and bottom extensions of hTranslin, as they do in spTranslin (Figure 7C and F, and Figure 3D and H). However, in the human ortholog the N-term patch signal is much stronger. Also, both proteins form octamers, which are their actual nucleic acid binding configurations. Thus, it appears likely that their protein–protein interaction regions are similar.

DISCUSSION

We used several bioinformatics tools to delineate, for the first time, a region in the 3D structure of the spTranslin monomer, which is likely to bind single-stranded nucleic acids. We then tested the effects of single amino acid substitutions in this and other regions on the binding and compared our experimental results with the bioinformatics predictions. This combined approach provided a consistent view of the interactions between nucleic acids and spTranslin, which probably applies to the human Translin and to other eukaryotic Translins as well. In this regard, we note that even though there is only a limited conservation along the primary amino acid sequences within the delineated nucleic acid binding regions of the Translins [Supplementary Figure S1, Figure 1 and ref. (22)], there are structural features of Translin that are more highly conserved. In particular, there is a shallow cavity in the equatorial part of the predicted nucleic acid binding region in both the yeast and the human Translins (Figure 8, encompassed by blue lines). In both Translins, the cavity is surrounded by positively charged amino acids, mostly arginines (in red in Figure 8). In spTranslin, substitution of each of these arginine residues with a non-polar amino acid caused a substantial reduction in the affinity of the protein towards single-stranded RNA and DNA (Figure 4). Lowering of the affinity was also observed by substitutions of other amino acids within and around the cavity. Taken together, these data suggest that the positively charged amino acids may anchor the negatively charged nucleic acid backbone, while a number of nucleic acid bases and sugar groups could interact with amino acid residues found within the cavity. This view is supported by the observation that Lys and especially Arg residues favor contacts with nucleic acids through the phosphate moiety (45). It is noteworthy that while the spTranslin and the hTranslin cavities occupy similar positions within their putative nucleic acid binding regions, the yeast cavity is quite hydrophilic and the human cavity is quite hydrophobic (Figure 8). Moreover, even though the size and the shape of the cavity vary in the various structural models of spTranslin, the hydrophilic nature of the cavity is predicted in all of them. We suggest that these striking differences at such a critical site for nucleic acid binding in the Translins of the two species may contribute to the differences that we observed in their affinity for single-stranded DNA (22). This suggestion is supported by the observation that the affinity of spTranslin towards single-stranded DNA (which is less hydrophilic than RNA) was increased by the three substitutions E168A, E21A and D101A that reduced the net negative charge and the hydrophilicity of the nucleic acid binding region on spTranslin.

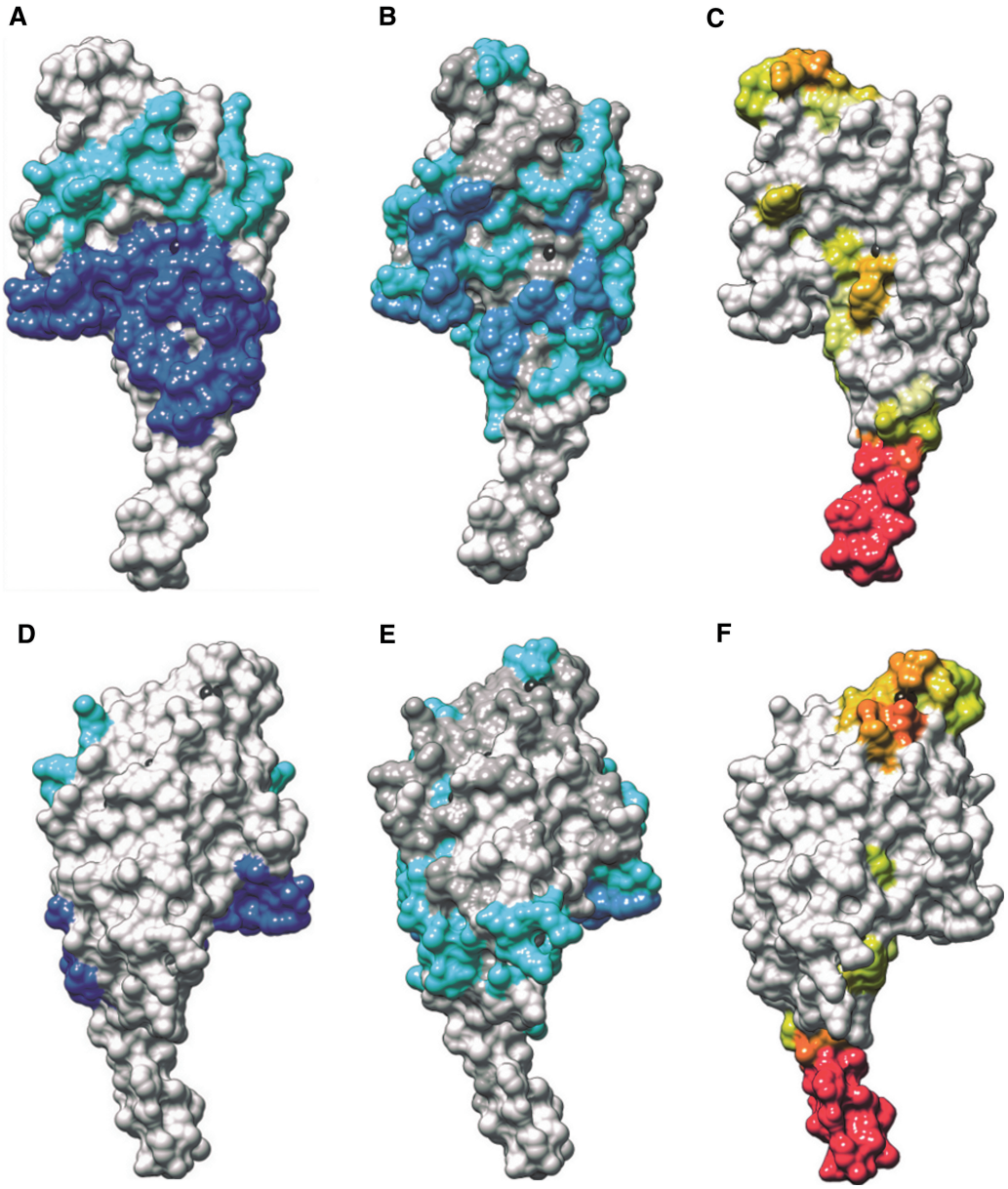

Figure 7. Front (first row) and rear (second row) faces of the 3D structure of the hTranslin monomer (PDB 1j1j, chain A), displaying results of the PFP, OPRA and ODA calculations. (**A** and **D**) The two largest positively charged electrostatic patches calculated by the PFP server, colored blue and cyan, respectively. (**B** and **E**) The OPRA scores for RNA binding prediction. In panels A, B, D and E, residues colored in blue and cyan are likely to interact with RNA and DNA probes. (**C** and **F**) ODA scores for protein-protein binding prediction. Residues colored red and orange are likely to interact with other proteins. Color-codes are as in Figure 3.

As mentioned in the 'Introduction' section, previous studies revealed that the octameric ring configuration of the yeast and the human Translins, rather than their free individual subunits, is the structural entity that actually binds single-stranded nucleic acids. It appears likely that an oligonucleotide that binds the octamer contacts more than one subunit. In this case, the minimal length of oligonucleotides that can bind the octamer would provide a measure of the distance between the contact sites in the subunits. In human Translin, we determined that this minimal length is 11-nucleotide residues (11). The symmetrical arrangement of the subunits in the octameric ring should favor binding of oligonucleotides consisting of regular repeats, such as the $(GU)_n$ and $d(GT)_n$; this could be one reason for the relatively high affinities of these oligonucleotides to both the human and the *S. pombe* Translins (11,22).

We have also shown in the present study that the highly conserved C-terminal region in spTranslin is required for association of individual spTranslin subunits (Figure 6). Hence, this segment is likely to be involved in formation of the octamer. The involvement of the C-terminal region of hTranslin in formation of octamers has also been previously deduced by analysis of the access of differentially located cysteine residues to chemical modification (7). Furthermore, studies of the interactions

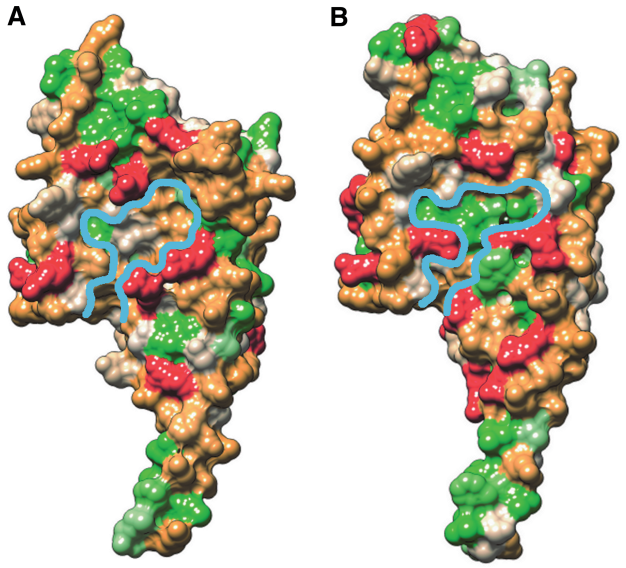

Figure 8. Front face of the 3D structural model of the spTranslin model 5.pdb (**A**) and of the hTranslin crystal structure (**B**). Residues are colored according to the hydrophobicity scale of Kyte and Doolittle, (46), from brown (most hydrophilic residues) to green (most hydrophobic residues). The central shallow cavities in both cases are surrounded by a blue line. Arginine residues, which have the highest hydrophilic score, are shown in red.

between fluorescent probes attached to these cysteines indicated that the subunits in the octameric ring are arranged in a tail-to-tail configuration (7). It should be noted, however, that the cysteines are not conserved in spTranslin. A priori, it was conceivable that nucleic acids may bind either the outside or the inside surface of the octameric ring. In fact, binding inside the ring was indicated by electron microscopic imaging of complexes of the human Translin with single-stranded DNA and RNA (5). This view is also consistent with our previous biochemical data, which indicated that even though the human Translin is able to bind single-stranded DNA extensions of double-stranded DNA molecules, it is unable to bind a stretch of single-stranded DNA flanked on both sides by double-stranded DNA (11). Apparently, to be bound, the single-stranded DNA must be inserted into the inner part of the ring, and this is hindered by the double-stranded DNA, whose diameter is too large to fit into the ring.

The experimental data discussed above regarding the binding of nucleic acids to spTranslin octamers were obtained in assays which were carried out with recombinant Translin preparations that did not contain TRAX. However, it has been reported that Translin/TRAX complexes purified from human cell extracts also bound nucleic acids (10). The size and the subunit composition of these complexes have not been determined in those studies. Our earlier study, though, revealed that native complexes capable of binding single-stranded DNA, which were prepared from human cell nuclei, had a molecular mass of ~190 kDa (1). Clearly, if this were a mixed complex, it would have consisted of at least six molecules of Translin and TRAX. Such a complex may contain adjoining Translin subunits, but cannot contain adjoining TRAX subunits, since the yeast two-hybrid assays indicated that a TRAX subunit cannot directly interact with another TRAX subunit (22). It should also be noted that the same C-terminal segment found to be required for Translin dimerization, is also required for binding of TRAX to Translin (Figure 6). Thus, the Translin–TRAX binding surface appears to overlap with the Translin–Translin binding surface.

Recent studies have indicated that in *Drosophila*, Translin–TRAX complexes are involved in siRNA metabolism (19). In particular, the TRAX subunit was reported to have a ribonuclease activity that appears to activate RISC complexes by removing siRNA passenger strand cleavage products. The ribonuclease activity of TRAX depends on specific glutamate and aspartate residues, which have been conserved in TRAX proteins from many eukaryotes, including man and *S. pombe* (19). The role of Translin in this function of Translin–TRAX complexes has not been elucidated. However, it appears likely that its RNA binding capability may contribute to the function. It is also conceivable that, in addition to the mixed Translin–TRAX complexes, eukaryotic cells assemble Translin octamers that lack TRAX, which may also play a role in siRNA metabolism, or in a different pathway, such as mRNA transport (17). In either case, our data on the Translin–nucleic acid interactions and on the interactions of Translin with TRAX should be helpful for a further elucidation of the functions of these highly conserved proteins in eukaryotic cells.

CONCLUSIONS

We present in this article, for the first time, a systematic mapping of nucleic acids and protein-binding surfaces on the 3D structure of Translin. The strength of our data stems from the combined use of powerful bioinformatics computation tools and molecular genetics techniques. A putative nucleic acid interaction surface was identified in the equatorial part of the protein front face. This interaction surface, which is only moderately conserved, includes polar and positively charged amino acids, mostly arginines, surrounding a shallow cavity. Additionally, two protein–protein interaction surfaces were identified at the top and the bottom of the protein structure. One protein interaction surface comprises amino acids that are close to the C-terminus and resides in a highly conserved region. It was shown experimentally to be required for interaction with TRAX and for interaction between Translin subunits. Both interactions were found not to depend on the ability of spTranslin to bind nucleic acids. Another protein–protein interaction surface was predicted at the bottom N-terminal region, which is only moderately conserved.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Dr K.L. Gould of the Howard Hughes Medical Institute, Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, for sending us the plasmid pREP41-NTAP. They also thank Dr Yael Mandel-Gutfreund of the Department of Biology, Technion, for helpful discussions and suggestions.

FUNDING

This work was supported by the Israel Science Foundation (Grants 378/03 and 967/08 to H.M.), by Philip Morris USA Inc. (to H.M.), and by the Spanish Plan Nacional I+D+i (grant BIO2008-02882 to J.F.-R.). L.P.-C. was supported by an FPU fellowship from the Spanish Ministry of Science.

Conflict of interest statement. None declared.

REFERENCES

1. Aharoni,A., Baran,N. and Manor,H. (1993) Characterization of a multisubunit human protein which selectively binds single stranded d(GA)n and d(GT)n sequence repeats in DNA. *Nucleic Acids Res*., **21**, 5221–5228.
2. Kasai,M., Aoki,K., Matsuo,Y., Minowada,J., Maziarz,R.T. and Strominger,J.L. (1994) Recombination hotspot associated factors specifically recognize novel target sequences at the site of interchromosomal rearrangements in T-ALL patients with t(8;14)(q24;q11) and t(1;14)(p32;q11). *Int. Immunol*., **6**, 1017–1025.
3. Han,J.R., Gu,W. and Hecht,N.B. (1995) Testis-brain RNA-binding protein, a testicular translational regulatory RNA-binding protein, is present in the brain and binds to the 3′ untranslated regions of transported brain mRNAs. *Biol. Reprod*., **53**, 707–717.
4. Aoki,K., Suzuki,K., Sugano,T., Tasaka,T., Nakahara,K., Kuge,O., Omori,A. and Kasai,M. (1995) A novel gene, Translin, encodes a recombination hotspot binding protein associated with chromosomal translocations. *Nat. Genet*., **10**, 167–174.
5. VanLoock,M.S., Yu,X., Kasai,M. and Egelman,E.H. (2001) Electron microscopic studies of the translin octameric ring. *J. Struct. Biol*., **135**, 58–66.
6. Lee,S.P., Fuior,E., Lewis,M.S. and Han,M.K. (2001) Analytical ultracentrifugation studies of translin: analysis of protein-DNA interactions using a single-stranded fluorogenic oligonucleotide. *Biochemistry*, **40**, 14081–14088.
7. Han,M.K., Lin,P., Paek,D., Harvey,J.J., Fuior,E. and Knutson,J.R. (2002) Fluorescence studies of pyrene maleimide-labeled translin: excimer fluorescence indicates subunits associate in a tail-to-tail configuration to form octamer. *Biochemistry*, **41**, 3468–3476.
8. Aoki,K., Ishida,R. and Kasai,M. (1997) Isolation and characterization of a cDNA encoding a Translin-like protein, TRAX. *FEBS Lett*., **401**, 109–112.
9. Chennathukuzhi,V.M., Kurihara,Y., Bray,J.D., Yang,J. and Hecht,N.B. (2001) Altering the GTP binding site of the DNA/RNA-binding protein, Translin/TB-RBP, decreases RNA binding and may create a dominant negative phenotype. *Nucleic Acids Res*., **29**, 4433–4440.
10. Finkenstadt,P.M., Jeon,M. and Baraban,J.M. (2002) Trax is a component of the Translin-containing RNA binding complex. *J. Neurochem*., **83**, 202–210.
11. Jacob,E., Pucshansky,L., Zeruya,E., Baran,N. and Manor,H. (2004) The human protein translin specifically binds single-stranded microsatellite repeats, d(GT)n, and G-strand telomeric repeats, d(TTAGGG)n: a study of the binding parameters. *J. Mol. Biol*., **344**, 939–950.
12. Wu,X.Q. and Hecht,N.B. (2000) Mouse testis brain ribonucleic acid-binding protein/translin colocalizes with microtubules and is immunoprecipitated with messenger ribonucleic acids encoding myelin basic protein, alpha calmodulin kinase II, and protamines 1 and 2. *Biol. Reprod*., **62**, 720–725.
13. Finkenstadt,P.M., Kang,W.S., Jeon,M., Taira,E., Tang,W. and Baraban,J.M. (2000) Somatodendritic localization of Translin, a component of the Translin/Trax RNA binding complex. *J. Neurochem*., **75**, 1754–1762.
14. Morales,C.R., Lefrancois,S., Chennathukuzhi,V., El Alfy,M., Wu,X.Q., Yang,J.X., Gerton,G.L. and Hecht,N.B. (2002) A TB-RBP and Ter ATPase complex accompanies specific mRNAs from nuclei through the nuclear pores and into intercellular bridges in mouse male germ cells. *Dev. Biol*., **246**, 480–494.
15. Chennathukuzhi,V.M., Lefrancois,S., Morales,C.R., Syed,V. and Hecht,N.B. (2001) Elevated levels of the polyadenylation factor CstF 64 enhance formation of the 1kB testis brain RNA-binding protein (TB-RBP) mRNA in male germ cells. *Mol. Reprod. Dev*., **58**, 460–469.
16. Yang,J.X., Chennathukuzhi,V., Miki,K., O'Brien,D.A. and Hecht,N.B. (2003) Mouse testis brain RNA-binding protein/translin selectively binds to the messenger RNA of the fibrous sheath protein glyceraldehyde 3-phosphate dehydrogenase-S and suppresses its translation in vitro. *Biol. Reprod*., **68**, 853–859.
17. Chiaruttini,C., Vicario,A., Li,Z., Baj,G., Braiuca,P., Wu,Y., Lee,F.S., Gardossi,L., Baraban,J.M. and Tongiorgi,E. (2009) Dendritic trafficking of BDNF mRNA is mediated by translin and blocked by the G196A (Val66Met) mutation. *Proc. Natl Acad. Sci. USA*, **106**, 16481–16486.
18. Wang,J., Boja,E.S., Oubrahim,H. and Chock,P.B. (2004) Testis brain ribonucleic acid-binding protein/translin possesses both single-stranded and double-stranded ribonuclease activities. *Biochemistry*, **43**, 13424–13431.
19. Liu,Y., Ye,X., Jiang,F., Liang,C., Chen,D., Peng,J., Kinch,L.N., Grishin,N.V. and Liu,Q. (2009) C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. *Science*, **325**, 750–753.
20. Pascal,J.M., Hart,P.J., Hecht,N.B. and Robertus,J.D. (2002) Crystal structure of TB-RBP, a novel RNA-binding and regulating protein. *J. Mol. Biol*., **319**, 1049–1057.
21. Sugiura,I., Sasaki,C., Hasegawa,T., Kohno,T., Sugio,S., Moriyama,H., Kasai,M. and Matsuzaki,T. (2004) Structure of human translin at 2.2 A resolution. *Acta Crystallogr. D. Biol. Crystallogr*., **60**, 674–679.
22. Laufman,O., Ben Yosef,R., Adir,N. and Manor,H. (2005) Cloning and characterization of the Schizosaccharomyces pombe homologs of the human protein Translin and the Translin-associated protein TRAX. *Nucleic Acids Res*., **33**, 4128–4139.
23. Jaendling,A., Ramayah,S., Pryce,D.W. and McFarlane,R.J. (2008) Functional characterisation of the Schizosaccharomyces pombe homologue of the leukaemia-associated translocation breakpoint binding protein translin and its binding partner, TRAX. *Biochim. Biophys. Acta*, **1783**, 203–213.
24. Tasto,J.J., Carnahan,R.H., McDonald,W.H. and Gould,K.L. (2001) Vectors and gene targeting modules for tandem affinity purification in Schizosaccharomyces pombe. *Yeast*, **18**, 657–662.
25. Romi,E., Baran,N., Gantman,M., Shmoish,M., Min,B., Collins,K. and Manor,H. (2007) High-resolution physical and functional mapping of the template adjacent DNA binding site in catalytically active telomerase. *Proc. Natl Acad. Sci. USA*, **104**, 8791–8796.
26. James,P., Halladay,J. and Craig,E.A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics*, **144**, 1425–1436.
27. Wu,S. and Zhang,Y. (2007) LOMETS: a local meta-threading-server for protein structure prediction. *Nucleic Acids Res*., **35**, 3375–3382.
28. Fiser,A., Do,R.K. and Sali,A. (2000) Modeling of loops in protein structures. *Protein Sci*., **9**, 1753–1773.
29. Wang,Q., Canutescu,A.A. and Dunbrack,R.L. Jr (2008) SCWRL and MolIDE: computer programs for side-chain conformation prediction and homology modeling. *Nat. Protoc*., **3**, 1832–1847.
30. Landau,M., Mayrose,I., Rosenberg,Y., Glaser,F., Martz,E., Pupko,T. and Ben-Tal,N. (2005) ConSurf 2005: the projection of

evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.*, **33**, W299–W302.
31. Perez-Cano,L. and Fernandez-Recio,J. (2009) Optimal protein-RNA area, OPRA: a propensity-based method to identify RNA-binding sites on proteins. *Proteins*, **78**, 25–35.
32. Shazman,S. and Mandel-Gutfreund,Y. (2008) Classifying RNA-binding proteins based on electrostatic properties. *PLoS. Comput. Biol.*, **4**, e1000146.
33. Fernandez-Recio,J., Totrov,M., Skorodumov,C. and Abagyan,R. (2005) Optimal docking area: a new method for predicting protein-protein interaction sites. *Proteins*, **58**, 134–143.
34. Berezin,C., Glaser,F., Rosenberg,J., Paz,I., Pupko,T., Fariselli,P., Casadio,R. and Ben-Tal,N. (2004) ConSeq: the identification of functionally and structurally important residues in protein sequences. *Bioinformatics.*, **20**, 1322–1324.
35. Aoki,K., Suzuki,K., Ishida,R. and Kasai,M. (1999) The DNA binding activity of Translin is mediated by a basic region in the ring-shaped structure conserved in evolution. *FEBS Lett.*, **443**, 363–366.
36. Chennathukuzhi,V.M., Kurihara,Y., Bray,J.D. and Hecht,N.B. (2001) Trax (translin-associated factor X), a primarily cytoplasmic protein, inhibits the binding of TB-RBP (translin) to RNA. *J. Biol. Chem.*, **276**, 13256–13263.
37. McGuffin,L.J. (2008) The ModFOLD server for the quality assessment of protein structural models. *Bioinformatics.*, **24**, 586–587.
38. Zhang,Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC. Bioinformatics.*, **9**, 40.
39. Glaser,F., Pupko,T., Paz,I., Bell,R.E., Bechor-Shental,D., Martz,E. and Ben-Tal,N. (2003) ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics.*, **19**, 163–164.
40. Pupko,T., Bell,R.E., Mayrose,I., Glaser,F. and Ben-Tal,N. (2002) Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. *Bioinformatics.*, **18(Suppl 1)**, S71–S77.
41. Stawiski,E.W., Gregoret,L.M. and Mandel-Gutfreund,Y. (2003) Annotating nucleic acid-binding function based on protein structure. *J. Mol. Biol.*, **326**, 1065–1079.
42. Jones,S., Shanahan,H.P., Berman,H.M. and Thornton,J.M. (2003) Using electrostatic potentials to predict DNA-binding sites on DNA-binding proteins. *Nucleic Acids Res.*, **31**, 7189–7198.
43. Rigaut,G., Shevchenko,A., Rutz,B., Wilm,M., Mann,M. and Seraphin,B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.*, **17**, 1030–1032.
44. Steen,H. and Mann,M. (2004) The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.*, **5**, 699–711.
45. Ellis,J.J., Broom,M. and Jones,S. (2007) Protein-RNA interactions: structural analysis and functional classes. *Proteins*, **66**, 903–911.
46. Kyte,J. and Doolittle,R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.*, **157**, 105–132.