Molecular docking uncovers TSPY binds more efficiently with eEF1A2 compared to eEF1A1

Deepak Panwar, Leena Rawal and Sher Ali*

Molecular Genetics Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110067, India

Communicated by Ramaswamy H. Sarma

(Received 7 May 2014; accepted 4 August 2014)

Testis-specific protein, Y-encoded (TSPY) binds to eukaryotic translation elongation factor 1 alpha (eEF1A) at its SET/NAP domain that is essential for the elongation during protein synthesis implicated with normal spermatogenesis. The eEF1A exists in two forms, eEF1A1 (alpha 1) and eEF1A2 (alpha 2), encoded by separate loci. Despite critical interplay of the TSPY and eEF1A proteins, literature remained silent on the residues playing significant roles during such interactions. We deduced 3D structures of TSPY and eEF1A variants by comparative modeling (Modeller 9.13) and assessed protein–protein interactions employing HADDOCK docking. Pairwise alignment using EMBOSS Needle for eEF1A1 and eEF1A2 proteins revealed high degree (~92%) of homology. Efficient binding of TSPY with eEF1A2 as compared to eEF1A1 was observed, in spite of the occurrence of significant structural similarities between the two variants. We also detected strong interactions of domain III followed by domains II and I of both eEF1A variants with TSPY. In the process, seven interacting residues of TSPY’s NAP domain namely, Asp 175, Glu 176, Asp 179, Tyr 183, Asp 240, Glu 244, and Tyr 246 common to both eEF1A variants were detected. Additionally, six lysine residues observed in eEF1A2 suggest their possible role in TSPY–eEF1A2 complex formation essential for germ cell development and spermatogenesis. Thus, more efficient binding of TSPY with eEF1A2 as compared to that of eEF1A1 established autonomous functioning of these two variants. Studies on mutated protein following similar approach would uncover the causative obstruction, between the interacting partners leading to deeper understanding on the structure–function relationship.

Keywords: comparative modeling; 3D structure; binding sites; protein docking; protein–protein interactions

Introduction

Protein–protein interactions (PPI) are regulated by the organizational complexities of protein maintaining biological functions such as signal transduction, gene expression, and cell cycle regulation. Understanding 3D structural topology and domains organization help in deciphering the functional domains of the protein. Despite the availability of structural information on protein–protein complexes in the database, predicting PPI by computational methods is one of the most challenging and resource intensive problems (Elcock, Sept, & McCammon, 2001). We opted to focus on a Y-linked gene known to play significant role in the process of spermatogenesis. The Y chromosome accounts for less than 2% and is the smallest one in the human genome (Vogt, 2005). This chromosome has an essential role in male sex development, spermatogenesis and regulation of fertility. Thus far, a total of 27 genes encoding proteins have been mapped (Skaletsky et al., 2003). One such protein is human testis-specific protein, Y-encoded (TSPY) consisting of 308 amino acid (33 kDa) located proximally on the Y chromosome (Yp11.2) close to putative gonadoblastoma locus (GBY) (Salo et al., 1995; Schnieders et al., 1996). TSPY is a member of proto-oncogene SET and NAP superfamily (Vogel et al., 1998). Members of this family harbor a highly conserved SET/NAP domain participating in various cellular functions, thereby regulating gene expression and cell cycle progression (Li & Lau, 2008; Oram, Liu, Lee, Chan, & Lau, 2006). In particular, proteins with SET/NAP domain have been demonstrated to be histone chaperones and are interactive partners of various transcription factors. TSPY, specifically, plays a significant role in controlling spermatogonial renewal, accurate transition of meiotic prophase during spermatogenesis, cell proliferation by abbreviating the cell cycle checkpoints, and expediting a rapid transition through the G2/M stages even under unfavorable conditions (Honecker et al., 2004; Oram et al., 2006; Schnieders et al., 1996). In humans, TSPY is mainly expressed in fetal and adult testis with limited expression in gonocytes and prepermatogonia during early development stages (Honecker et al., 2004; Lau, Li, & Kido, 2011). TSPY is upregulated in testicular carcinoma (Lau, Chou, Iezzoni, Alonzo, & Komuves, 2000), seminomas, and melanomas (Gallagher et al., 2005) but shows reduced expression in prostate cancer (Lau, Lau, & Komuves, 2003) and hepatocellular carcinoma (Yin et al., 2005).

*Corresponding author. Email: alisher@nii.ac.in

© 2014 Taylor & Francis
TSPY interacts with eukaryotic translation elongation factor 1 alpha (eEF1A), a regulatory protein monitoring cell growth and expression of the oncogenic genes (Kido, & Lau 2008). The eEF1A is a GTP-binding protein, one of the most abundant protein synthesis factors. It exists in two variants, eEF1A1 and eEF1A2, encoded by two separate genes located on human chromosome 6q14 and chromosome 20q13.3, respectively (Lund, Knudsen, Vissing, Clark, & Tommerup, 1996). These variants have similar functions during protein synthesis (Tomlinson et al., 2005). However, small differences in the amino acids between the two eEF1A variants may affect their functions during protein synthesis. Deregulation or misexpression of translational machinery with eEF1A may lead to development of several types of cancers where translational control is seriously compromised. The eEF1A1 and eEF1A2 are oncogenes expressed in breast and lung adenocarcinomas and germ cell tumors (Chen & Madura, 2005; Kulkarni et al., 2007; Tomlinson et al., 2005). While role of eEF1A1 in cancer development is less clear, eEF1A2 is an established oncogene (Anand et al., 2002). Previous studies revealed various post-translational modifications in both the variants leading to alterations in their potential for oncogenic transformations (Dever, Costello, Owens, Rosenberry, & Merrick, 1989; Rosenberry et al., 1989).

A high resolution crystal structure of *Saccharomyces cerevisiae* eEF1A revealed various functionally important sequences spread across three structural domains of the protein (Andersen et al., 2000). Three-dimensional (3D) structures of human eEF1A1 and eEF1A2 modeled using the yeast counterpart, as template have shown highly conserved nature of these residues linked to the translation related functions of eEF1A (Soares, Barlow, Newbery, Porteous, & Abbott, 2009). Three consensus GTP-binding motifs were uncovered in domain I, whereas domains II and III are implicated in delivering aminoacylated-tRNAs to the ribosome besides acting as an initiator (Dever, Glyniás, & Merrick, 1987; Slobin, 1980). Both the eEF1A variants bind with TSPY’s C-terminus domain and regulate protein synthesis (Kido, & Lau 2008). However, no reports regarding the TSPY structural analysis and binding residues involved in such interactions having clinical implications are available.

In the present study, by employing bioinformatic tools, we deduced 3D structures of TSPY and eEF1A variants and uncovered protein–protein interactions between the two proteins along with their binding sites playing significant roles in germ cell development and spermatogenesis. This base line data is envisaged to be useful for comparison of the mutant elongation factors with that of normal ones elucidating precisely the effect of changed amino acids affecting the 3D structure.

Materials and methods

Target sequences, secondary structures, and domain analysis

The amino acid sequence of *Homo sapiens* TSPY protein (Accession No: AAB51693.1) was retrieved from NCBI protein database. The sequences of *Homo sapiens* eEF1A1 (SwissProt Accession No: P68104) and eEF1A2 proteins (SwissProt Accession No: Q05639) were obtained from Swiss-Prot database. To identify regions of similarity/variation between eEF1A1 and eEF1A2, pairwise sequence alignment was assessed through EMBOSS Needle (Rice, Longden, & Bleasby, 2000). SOPMA was used for calculating the secondary structural features of the protein sequences (http://npsa-pbil.ibcp.fr) (Geourjon & Deleage, 1995). Domain analysis was assessed using CDD (Conserved Domain Database) (Marchler-Bauer et al., 2011).

Three-dimensional structure modeling and evaluation

PSI-BLAST against Protein Data Bank (PDB) was carried out to identify their homologous structures based on the maximum identity with high score and lower e-value (Altschul et al., 1997). Template with 35% similarity (PDB ID: 2e50, crystal structure of SET/TAF-1beta/INHAT chain A, Resolution: 3.30 Å) with TSPY and yeast elongation factor (PDB ID: 1F60, crystal structure of the yeast elongation factor complex eEF1A:eEF1BA chain A, Resolution: 1.67 Å) with similarity of 80% for both eEF1A1 and eEF1A2, were selected to build 3D structures using Modeller 9.13 (Sali & Blundell, 1993). Models showing the closest Cα RMSD (root-mean-square deviation) with respect to their templates upon superposition were selected for structural analysis and docking studies. Stereiochemical properties were assessed by analyzing Ramachandran plot using PROCHECK through SAVES server (http://nihserver.mbi.ucla.edu/SAVES/) (Laskowski, MacArthur, Moss, & Thornton, 1993). The coarse packing qualities of the models were confirmed using WHATIF server (http://swift.cmbi.ru.nl/servers/html/index.html) (Vriend, 1990). ProSA verified the protein structure from X-ray analysis, NMR spectroscopy and other theoretical calculations (Wiederstein & Sippl, 2007). The modeled structures were confirmed by Protein Quality Predictor server (ProQ) (Wallner & Elofsson, 2006).

Protein docking

Binding site residues were retrieved by CASTp server (Joe Dundas et al., 2006) and selected residues were used as constraints under protein docking, through HADDOCK web server (Dominguez, Boelens, & Bonvin, 2003). HADDOCK integrates restraints from mutagenesis and
NMR chemical shift data in protein docked complexes which were then converted into a series of ambiguous interaction restraints (AIR). HADDOCK run resulted in 200 final protein docked complexes, which were then analyzed in the context of RMSD to the native structure. The HADDOCK Score was calculated based on sum of the van der Waals energy ($E_{vdw}$), electrostatic energy ($E_{elec}$), desolvation energy, energy from restraint violations ($E_{AIR}$), and the buried surface area (HADDOCK Score = $E_{vdw} + E_{elec} + E_{AIR}$). After subjection to simulation and water refinement the HADDOCK score was again calculated i.e. HADDOCK Score = $1.0E_{vdw} + .2E_{elec} + .1E_{dist} + 1.0E_{AIR}$. Where $E_{dist}$ is the distance restraints energy contribution that includes both unambiguous interaction restraints and AIRs. The best docked complexes with lowest intermolecular energies were selected for further studies (Dominguez et al., 2003; van Dijk, Boelens, & Bonvin, 2005; van Dijk, 2006). Final docked complexes, TSPY–eEF1A1 and TSPY–eEF1A2 were analyzed using Protein Interactions Calculator (PIC) to confirm the interacting residues within 5 Å (Tina, Bhadra, & Srinivasan, 2007). All the structures and docked complexes were visualized using PyMol server (http://www.pymol.org; DeLano Scientific, San Carlos, CA, USA).

Results

Amino acid variations and secondary structure

Human eEF1A1 and eEF1A2 showed 92% sequence identity and 98% homology at the amino acid level. Pairwise alignment identified amino acid variations at 35 positions between eEF1A1 and eEF1A2 proteins (Figure 1). TSPY was found to have 46.10, 10.71, 2.92, and 40.26% of α helices, extended strand, β turns, and random coils, respectively. Whereas, eEF1A1 and eEF1A2 encompasses 32.25 and 28.08% α helices, 21.21 and 22.25% extended strand, 8.01 and 9.07% β turns,

![Figure 1. Pairwise sequence alignment between human eEF1A1 and eEF1A2. The two sequences showed 92% identity and 98% homology at the amino acid level. Identical residues are shown in gray background and variant amino acid differences are in yellow. One dot denotes semi-conservative amino acid residues and two dots, conserved substitutions.](http://www.example.com/figure1.png)
and 40.26 and 38.53% random coils, respectively. A stretch of 181 residues (108–288 position) pertaining to nucleosome assembly protein (NAP) domain was predicted in TSPY protein. Both eEF1A variants encompass three domains. Domain I (residues 9–238), domain II (residues 242–331), and domain III (residues 335–438) are made up almost entirely of the beta-strands; each domain contains two beta-sheets forming a beta-barrel (Figure 2).

Quality of the predicted 3D structure

Models with RMSD of .39, .13, and .11 Å for TSPY, eEF1A, and eEF1A2, respectively, were selected for structural analysis and docking studies (Figure 3). Structural localization of amino acid variations between eEF1A1 and eEF1A2 is shown in Figure 4. Ramachandran dihedral statistics for TSPY revealed a total of 88.1, 8.1, 2.7, and 1.1% residues in most favored, additionally allowed, generously allowed, and

Figure 2. Domain view of human elongation factor eEF1A and TSPY proteins assessed through Conserved Domain Database (CDD). (A) TSPY protein consist of 181 amino acids (108–288 position) long NAP domain which plays significant role in nuclear translocation affecting transcriptional programs in spermatogenesis. (B) Elongation factor eEF1A1 and eEF1A2 consist of three main structural domains: domain I is a GTP-dependent binding domain and domains II and III are responsible for binding of aminoacyl-tRNAs to ribosomes for initiation of protein synthesis.

Figure 3. Three-dimensional models of human TSPY, eEF1A1, and eEF1A2 predicted by Modeller. (A) Cartoon representation of predicted model of human TSPY protein, NAP domain shown in magenta color. (B) Predicted model of human eEF1A1, red, yellow, and gray color represent helix, beta-sheets and loops, respectively. (C) Predicted model of human eEF1A2, green, magenta, and blue color represent alpha helix, beta-sheets and loops, respectively.
disallowed regions, respectively. However, eEF1A1 and eEF1A2 were found to bear 95.2 and 93.9% residues in favorable regions, 3.7 and 4.8% residues in the additionally allowed ones, .5 and 1.1% residues in the generously allowed regions, and .5 and .3% in the disallowed regions, respectively. The Ramachandran plots show the excellent geometry of the models (SI Figure 1). The overall G-scores were −.13, −.06, and

Figure 4. Schematic representation of amino acid variations mapped onto 3D models of human eEF1A1 and eEF1A2 in surface mode. (A) 3D model of eEF1A1 rotated by 180° about the y-axis in surface representation depicting variant amino acid residues in red color (domains I, II, and III are in yellow, pale green, and marine blue, respectively). (B) 3D model of eEF1A2 rotated by 180° about the y-axis in surface representation depicting variant amino acid residues shown in red color (domains I, II, and III colored cyan, magenta, and green, respectively). All figures were visualized using PyMol.

Table 1a. Active site residues of TSPY, eEF1A1, and eEF1A2 proteins predicted by CASTp server.

| Protein | Active residues |
|---------|-----------------|
| TSPY    | 134 Arg, 138 Glu, 141 Arg, 142 Lys, His 144, Leu 145, Arg148, Glu 176, Glu 188, Val 189, Glu 190, Glu 191, Lys 193, 237 Trp, 258 Phe, 259 Asn, 260 Trp, 261 Phe, 277 Cys, 281 Trp, 282 Arg |
| eEF1A1  | 346 Ile, 347 Leu, 348 Asn, 362 Asp, 364 His, 365 Thr, 366 Ala, 367 His, 368 Ile, 381 Arg, 396 Ser, 397 Gly, 408 Ile, 409 Pro, 410 Met, 411 Cys, 415 Phe, 419 Pro, 420 Pro, 423 Arg, 425 Ala, 427 Arg, 430 Arg, 431 Gln, 432 Thr |
| eEF1A2  | 345 Ile, 346 Ile, 347 Leu, 348 Asn, 362 Asp, 364 His, 365 Thr, 368 Ile, 381 Arg, 397 Gly, 415 Phe, 419 Pro, 420 Pro, 423 Arg, 408 Lys, 409 Pro, 410 Met, 411 Cys, 425 Ala, 426 Val, 427 Arg, 430 Arg, 431 Gln, 432 Thr |

Table 1b. Amino acid residues of proteins, submitted to HADDOCK web server for protein–protein docking analysis.

| Protein | eEF1A1 | eEF1A2 |
|---------|--------|--------|
| TSPY    | 176,237,258,259,260,261 | 343,362,381,397,420,432 | 343,365,381,399,415,438 |
| eEF1A1  |        |        |        |
| eEF1A2  |        |        |        |
Table 2. Statistical analysis for HADDOCK generated TSPY-eEF1A1 and TSPY-eEF1A2 docked complexes.

| S.No. | Cluster | HADDOCK score | Cluster Size | RMSD from overall lowest-energy structure (Å) | Van der Waals energy | Electrostatic energy | Desolvation energy | Restraints violation energy | Buried surface area (Å²) | Z-Score |
|-------|---------|---------------|--------------|-----------------------------------------------|----------------------|---------------------|---------------------|----------------------------|--------------------------|---------|
|       |         |               |              |                                               |                      |                     |                     |                            |                          |         |
| A) TSPY–eEF1A1 docking complex |
| 1    | 1       | -129.5 ± 8.6  | 69           | 1.2 ± 0.7                                     | -63.6 ± 0.4          | -541.8 ± 60.6       | 33.7 ± 7.7          | 87.6 ± 37.6                 | 1861.2 ± 17.6             | -1.8    |
| 2    | 3       | -106.6 ± 8.1  | 9            | 13.3 ± 0.5                                    | -62.8 ± 10.3         | -544.6 ± 68.9       | 54.8 ± 11.4         | 103.3 ± 11.7                | 2297.3 ± 116.1            | -1.1    |
| 3    | 4       | -101.2 ± 7.5  | 8            | 22.8 ± 0.1                                    | -54.3 ± 10.6         | -532.8 ± 57.7       | 45.1 ± 9.9          | 145.5 ± 32.9                | 2140.0 ± 34.5             | -0.9    |
| 4    | 5       | -86.4 ± 10.0  | 8            | 20.8 ± 0.3                                    | -49.6 ± 4.8          | -478.3 ± 48.7       | 50.3 ± 10.0         | 85.5 ± 43.5                 | 1895.0 ± 123.0            | -0.4    |
| 5    | 6       | -72.4 ± 9.7   | 6            | 5.0 ± 0.6                                     | -45.7 ± 2.7          | -434.2 ± 45.9       | 38.2 ± 5.0          | 219.6 ± 29.9                | 1832.0 ± 110.8            | 0.0     |
| 6    | 2       | -71.3 ± 10.4  | 11           | 17.2 ± 0.3                                    | -36.8 ± 6.2          | -461.3 ± 29.9       | 44.0 ± 4.9          | 137.3 ± 19.6                | 1692.1 ± 125.6            | 0.0     |
| 7    | 9       | -49.9 ± 14.2  | 4            | 19.4 ± 1.0                                    | -37.3 ± 5.4          | -358.6 ± 26.0       | 52.2 ± 8.9          | 68.7 ± 50.5                 | 1519.9 ± 222.6            | 0.7     |
| 8    | 7       | -46.4 ± 5.3   | 5            | 18.2 ± 0.5                                    | -44.8 ± 9.9          | -289.2 ± 36.7       | 48.3 ± 1.8          | 80.0 ± 33.1                 | 1265.1 ± 128.5            | 0.8     |
| 9    | 10      | -41.8 ± 33.6  | 4            | 4.3 ± 1.6                                     | -31.5 ± 10.9         | -471.4 ± 70.2       | 72.4 ± 16.5         | 116.2 ± 37.2                | 1528.2 ± 99.6             | 1.0     |
| 10   | 8       | -20.3 ± 19.4  | 5            | 24.8 ± 0.2                                    | -25.0 ± 4.5          | -278.7 ± 81.3       | 52.2 ± 12.7         | 82.5 ± 17.4                 | 1286.5 ± 270.3            | 1.6     |
| B) TSPY–eEF1A2 docking complex |
| 1    | 9       | -114.4 ± 25.6 | 4            | 0.9 ± 0.6                                     | -31.9 ± 19.9         | -891.8 ± 133.1      | 80.3 ± 19.8         | 154.9 ± 15.4                | 2324.4 ± 280.7            | -1.9    |
| 2    | 1       | -91.1 ± 3.5   | 19           | 22.8 ± 1.0                                    | -46.8 ± 4.3          | -614.8 ± 36.8       | 58.6 ± 11.0         | 200.5 ± 46.6                | 2261.0 ± 144.3            | -0.9    |
| 3    | 3       | -87.9 ± 7.9   | 11           | 25.3 ± 0.3                                    | -45.4 ± 7.2          | -682.1 ± 34.9       | 77.6 ± 4.9          | 163.2 ± 34.6                | 2173.3 ± 122.9            | -0.7    |
| 4    | 5       | -83.3 ± 17.0  | 9            | 29.8 ± 0.2                                    | -53.1 ± 9.4          | -599.1 ± 72.5       | 70.7 ± 11.9         | 188.8 ± 75.6                | 2200.4 ± 150.5            | -0.5    |
| 5    | 7       | -81.3 ± 10.0  | 6            | 23.0 ± 0.4                                    | -45.9 ± 8.3          | -608.5 ± 75.2       | 76.0 ± 12.2         | 103.0 ± 42.8                | 2150.9 ± 78.9             | -0.4    |
| 6    | 4       | -66.7 ± 17.0  | 10           | 27.0 ± 0.2                                    | -22.6 ± 1.1          | -582.5 ± 48.9       | 56.0 ± 8.0          | 164.2 ± 22.5                | 1554.5 ± 94.8             | 0.2     |
| 7    | 6       | -52.2 ± 11.6  | 6            | 22.2 ± 0.1                                    | -36.4 ± 12.6         | -518.5 ± 22.2       | 73.0 ± 10.8         | 149.7 ± 28.8                | 1940.5 ± 154.8            | 0.9     |
| 8    | 2       | -50.9 ± 9.7   | 13           | 24.6 ± 0.3                                    | -44.7 ± 8.0          | -405.5 ± 33.2       | 64.1 ± 12.9         | 108.9 ± 66.8                | 1795.2 ± 188.7            | 0.7     |
| 9    | 10      | -50.0 ± 14.8  | 4            | 24.5 ± 0.2                                    | -22.9 ± 4.2          | -498.7 ± 43.6       | 59.7 ± 11.1         | 129.7 ± 15.7                | 1618.2 ± 65.4             | 1.0     |
| 10   | 8       | -39.9 ± 11.2  | 5            | 24.5 ± 0.7                                    | -49.6 ± 4.9          | -264.2 ± 61.6       | 46.9 ± 15.0         | 155.6 ± 43.5                | 1668.3 ± 128.1            | 1.4     |

aThe HADDOCK score = 1.0Evdw + 0.2Eelec + 1.0Edesol + 0.1EAir
bNon-bonded interactions were calculated with the Optimized Potentials for Liquid Simulations (OPLS) force field using an 8.5Å cut-off.
−.05 for TSPY, eEF1A1, and eEF1A2, respectively, indicating that the predicted models were acceptable. Coarse packing quality of atoms in models gave the scores of −1.009 (TSPY), −.775 (eEF1A1), and −.769 (eEF1A2), suggesting that the models generated were of good quality. ProSA predicted Z-scores values were −6.21, −10.07, and −9.91 for TSPY, eEF1A1, and eEF1A2, respectively, evidencing highly reliable structures. Additionally, the energy plots showed the local model quality based on plotting energies as a function of amino acid sequence position (SI Figure 2).

Essence of protein docking

The predicted binding residues obtained by CASTp corresponding to TSPY, eEF1A1, and eEF1A2 proteins having more than 50% solvent accessible area were further analyzed by docking (Tables 1(A) and 1(B)). Ten best clusters generated by docking for each complex (TSPY–eEF1A1 and TSPY–eEF1A2) are shown in Table 2(A) and (B). Clusters with low HADDOCK scores and low RMSD values were considered as the best docked complexes. Model with HADDOCK score of −129.5 ± 8.6 kcal/mol and RMSD value of 1.2 ± 7 Å, was selected for TSPY–eEF1A1 complex (Figure 5(A)). Moreover, TSPY–eEF1A2 chosen complex had a HADDOCK score of −114.4 ± 25.6 kcal/mol and RMSD value of .9 ± .6 Å (Figure 6(A)).

Docked complexes and protein interactions

Analysis of docked complexes (TSPY–eEF1A1/eEF1A2) using PIC server showed the presence of interacting

![Docked model of the TSPY–eEF1A1 complex and its structural mapping of intermolecular interfaces. (A) Cartoon representation of TSPY–eEF1A1 docked complex having low energy score determined by HADDOCK docking. TSPY displayed in blue and eEF1A1 in yellow. (B) A close-up of interacting interface corresponding to TSPY–eEF1A1 complex is represented in surface representation. Interacting residues from TSPY and eEF1A1 are shown in red and green, respectively. (C) Interacting residues between TSPY and eEF1A1 obtained through hydrogen bonding are shown in line representation, TSPY residues (red) and eEF1A1 residues (green) with the atomic distances (magenta dotted lines) along with labeled residues name.](image-url)
residues elucidating extensive H-bonding interactions and interacting interface demonstrating the abundance of polar amino acid residues (Tables 3 and 4). TSPY was shown to interact with eEF1A1 having buried surface area of 1861.2 ± 17.6 Å² (Figure 5(B)). High affinity was seen in TSPY–eEF1A2 complex with 2324.4 ± 280.7 Å² buried surface area (Figure 6(B)). Furthermore, TSPY’s NAP domain was found to mediate interactions between the two proteins. Fourteen amino acid residues of eEF1A1 (2 were part of domain II and 12 belonged to domain III) showed interaction with 12 amino acid residues of TSPY’s NAP domain (Figure 5(C)). Moreover, eEF1A2 had 14 residues from which 6 were placed in domain I and 8 resided in domain III showing active participation in complex formation with 15 residues of NAP domain of TSPY (Figure 6(C)). Further, analysis of the two complexes (TSPY–eEF1A1 and TSPY–eEF1A2) showed the occurrence of seven actively interacting amino acid residues namely Asp 175, Glu 176, Asp 179, Tyr 183, Asp 240, Glu 244, and Tyr 246 (Table 5) common to both the elongation factors eEF1A1 and eEF1A2. Interestingly, six lysine residues of eEF1A2 at positions 129, 179, 180, 378, 385, and 395 were found to have electrostatic interactions with NAP domain of TSPY. Despite having 92% sequence identity and similar 3D structures of both variants of eEF1A, docking analysis revealed different binding modes of TSPY with eEF1A1 and eEF1A2.
Table 3. Interfacial residues in docked complex of TSPY–eEF1A1. Protein–protein main chain/side chain–side chain hydrogen bonds. Chain A represents TSPY protein and chain B represents eEF1A1 protein.

| S.No. | Chain A Distance (Å) Chain B |
|-------|-----------------------------|
| 1     | ARG 247 [N] 2.82 ASP 252 [OD1] |
| 2     | ARG 247 [N] 2.98 ASP 252 [OD2] |
| 3     | ARG 248 [N] 3.16 ASP 252 [OD2] |
| 4     | ARG 247 [HH21] 1.81 TYR 254 [OH] |
| 5     | ASN 208 [HD21] 1.94 ALA 358 [O] |
| 6     | GLU 244 [OE2] 1.72 ARG 321 [HH11] |
| 7     | ASP 175 [OD1] 1.75 ARG 321 [HH22] |
| 8     | TYR 246 [OH] 3.12 ARG 322 [N] |
| 9     | ARG 206 [NH2] 2.93 ALA 355 [O] |
| 10    | ARG 206 [NH2] 3.46 GLY 356 [O] |
| 11    | SER 207 [O] 3.49 ALA 358 [N] |
| 12    | ASP 179 [OD2] 3.58 ALA 369 [O] |
| 13    | TYR 183 [OH] 2.17 LYS 371 [HZ1] |
| 14    | TYR 183 [OH] 3.21 LYS 369 [O] |
| 15    | TYR 183 [OH] 3.42 GLY 407 [N] |
| 16    | ASP 179 [OD1] 3.76 LYS 408 [N] |
| 17    | ASP 175 [OD2] 1.58 LYS 408 [HZ2] |
| 18    | ASP 175 [O] 2.19 LYS 408 [HZ1] |
| 19    | ASP 240 [OD2] 1.67 ARG 427 [HH11] |
| 20    | ASP 240 [O] 3.50 HIS 367 [NE] |

Table 4. Interfacial residues in docked complex of TSPY–eEF1A2. Protein–protein main chain/side chain–side chain hydrogen bonds. Chain A represents TSPY protein and chain B represents eEF1A2 protein.

| S.No. | Chain A Distance (Å) Chain B |
|-------|-----------------------------|
| 1     | GLN 168 [HE21] 1.82 GLU 124 [OE1] |
| 2     | GLU 244 [N] 2.85 ASP 380 [OD2] |
| 3     | TYR 246 [HH1] 1.71 ARG 382 [O] |
| 4     | ALA 164 [O] 1.94 ASN 130 [HD21] |
| 5     | SER 170 [O] 2.67 LYS 129 [N] |
| 6     | SER 170 [O] 3.83 ASN 130 [N] |
| 7     | ALA 171 [O] 1.65 LYS 129 [HZ2] |
| 8     | LEU 172 [O] 2.43 LYS 180 [HZ1] |
| 9     | ASP 175 [OD1] 2.07 ARG 134 [HH22] |
| 10    | ASP 175 [OD1] 1.76 ARG 381 [HH1] |
| 11    | ASP 175 [OD2] 3.38 ARG 398 [N] |
| 12    | ASP 175 [OD2] 3.34 ALA 399 [N] |
| 13    | GLU 176 [OE1] 1.57 ARG 381 [HH22] |
| 14    | GLU 176 [OE2] 2.33 ARG 382 [HE] |
| 15    | GLU 176 [OE2] 1.93 ARG 382 [HH21] |
| 16    | GLU 176 [OE2] 1.78 ARG 381 [HH12] |
| 17    | ASP 179 [OD2] 1.60 LYS 378 [HZ2] |
| 18    | ASP 179 [OD2] 1.55 LYS 395 [HZ2] |
| 19    | TYR 183 [OH] 2.05 LYS 395 [HZ1] |
| 20    | ASP 240 [O] 1.96 LYS 385 [HZ1] |
| 21    | TYR 241 [OH] 2.24 LYS 378 [HZ1] |
| 22    | GLU 244 [OE2] 1.59 ARG 382 [HH12] |
| 23    | ASN 252 [OD1] 1.74 LYS 179 [HZ2] |
| 24    | LEU 255 [O] 1.84 LYS 180 [HZ1] |

Protein–protein side chain–side chain hydrogen bonds

| S.No. | Chain A Distance (Å) Chain B |
|-------|-----------------------------|
| 1     | GLU 244 [OE2] 3.86 ARG 321 [NE] |
| 2     | GLU 244 [OE2] 2.71 ARG 321 [NH1] |
| 3     | ASP 175 [OD1] 2.68 ARG 321 [NH2] |
| 4     | ASP 175 [OD2] 3.99 ARG 321 [NH2] |
| 5     | ASP 240 [OD1] 2.68 HIS 367 [NE2] |
| 6     | ASP 240 [OD2] 2.68 HIS 367 [NE2] |
| 7     | GLU 176 [OE1] 2.97 LYS 408 [N] |
| 8     | GLU 176 [OE2] 2.68 LYS 408 [NZ] |
| 9     | ASP 240 [OD2] 3.81 ARG 427 [NE] |
| 10    | ASP 240 [OD2] 2.66 ARG 427 [NH1] |

Protein–protein side chain–side chain hydrogen bonds

| S.No. | Chain A Distance (Å) Chain B |
|-------|-----------------------------|
| 1     | ASP 175 [OD1] 2.88 ARG 134 [NH2] |
| 2     | ASP 175 [OD1] 2.70 ARG 381 [NH2] |
| 3     | ASP 175 [OD1] 3.27 ARG 381 [NE] |
| 4     | GLU 176 [OE1] 3.43 ARG 381 [NH1] |
| 5     | GLU 176 [OE2] 2.58 ARG 381 [NH2] |
| 6     | GLU 176 [OE2] 3.16 ARG 382 [NE] |
| 7     | GLU 176 [OE2] 2.88 ARG 382 [NH2] |
| 8     | GLU 176 [OE2] 2.79 ARG 381 [NH1] |
| 9     | GLU 176 [OE2] 3.56 ARG 381 [NH2] |
| 10    | ASP 179 [OD1] 3.74 LYS 378 [NZ] |
| 11    | ASP 179 [OD2] 2.64 LYS 378 [NZ] |
| 12    | ASP 179 [OD2] 2.59 LYS 395 [NZ] |
| 13    | GLU 244 [OE2] 2.60 ARG 382 [NH1] |

Discussion

TSPY protein harbors a conserved SET/NAP domain and serves diverse cellular functions, including regulation of transcription, cell cycle, and chromatin assembly/remodeling (Honecker et al., 2004; Lau et al., 2000; Slobin, 1980). Currently, the factors mediating the main physiological functions of TSPY are not known. Interactions between TSPY and eEF1A could potentially enhance protein synthesis associated with cell proliferation and germ cell tumorigenesis (Anand et al., 2002). The present study is the first attempt towards the computational analysis of TSPY protein interaction with eukaryotic elongation factor 1 alpha (eEF1A) employing protein–protein docking approach. As crystal structures were not available in databases, 3D models were predicted by homology modeling to ascertain their functional relevance. Both eEF1A1 and eEF1A2 are 92% identical and predicted to have similar tertiary structures. Our data is an extension of the earlier work showing amino acid variation at 35 positions between the two variants of eEF1A. The eEF1A1 is reported to be highly conserved indicating its mandatory role for critical biological functions. Substantial level of variation within the eEF1A2 suggests its adaptability during the course of evolution across the species (Soares & Abbott, 2013).

Our in silico results demonstrate TSPY interaction with eEF1A variants, eEF1A1, and eEF1A2. The detected interacting amino acid residues could be the potential ones regulating the structural–functional relationship between TSPY–eEF1A1/A2 complexes thereby affecting the protein synthesis. Our work substantiates the earlier studies demonstrating the stronger binding of TSPY–eEF1A2 complex as compared to TSPY–eEF1A1 complex (Kido & Lau, 2008). Additionally, domain III
of both the eEF1A variants is not the sole interacting domain with TSPY but also domain II from eEF1A1 and domain I from eEF1A2 actively participate in the complex formation. The lysine residues belonging to eEF1A2 are reported to be involved in acetylation, ubiquitination, and methylation events occurring during post-translational modifications (Soares & Abbott, 2013). Therefore, the occurrence of six lysine residues in eEF1A2 interacting with TSPY suggests their possible roles in enhanced protein synthesis and cancer development through post-translational modifications (Lee, Ann & Wang, 1994; Soares & Abbott, 2013). A vast repertoire of phosphorylation events at serine, threonine, and tyrosine residues have been reported for eEF1A but with a greater affinity towards eEF1A2 than eEF1A1 variant (Panasyuk, Nemazanyy, Filonenko, Negrutskii, & El’skaya, 2008).

While the tumorigenicity status of the eEF1A1 is less evident, eEF1A2 has been often implicated as an oncogene (Anand et al., 2002). The eEF1A is found to be in excess over the other translation elongation components. However, it is not yet clear if tumors with upregulated eEF1A2 would have a greater protein synthesis capacity (Tomlinson et al., 2007). Regardless of the high degree of amino acid identity between eEF1A variants, both forms might perform completely distinct non-canonical functions as reported earlier suggesting independent oncogenic roles of eEF1A2 (Anand et al., 2002; Kulkarni et al., 2007).

A deeper insight towards the interaction of eEF1A variants with TSPY and their binding residues would open up newer avenues to understand the mechanism of germ cell development, spermatogenesis, and tumorigenesis, highlighting its biological and clinical relevance. Taken together, bioinformatic approaches coupled with molecular biology prove to be more reliable tools addressing important events such as germ cell development and spermatogenesis.

## Conclusions

We have developed homology models and set out protein docking to probe the interaction of TSPY and eEF1A variants. Despite having a significant sequence homology, both eEF1A1 and eEF1A2 displayed differential binding affinity for TSPY, a testis-specific protein, and probably undergo different regulatory pathways. TSPY appears to regulate protein synthesis by binding specifically to eEF1A2 during spermatogenesis and germ cell tumor growth. Thus, the present study has uncovered a novel TSPY–eEF1A2 mediated pathway of protein synthesis in normal and pathological conditions that may have therapeutic implications.

### List of abbreviations

| Acronym | Description |
|---------|-------------|
| TSPY    | Testis-specific protein, Y-encoded |
| GBY     | Gonadoblastoma locus on the Y chromosome |
| eEF1A   | Eukaryotic translation elongation factor I alpha |
| PPI     | Protein–protein interactions |
| RMSD    | Root-mean-square deviation |

### Supplementary material

The supplementary material for this paper is available online at [http://dx.doi.org/10.1080/07391102.2014.952664](http://dx.doi.org/10.1080/07391102.2014.952664).

### Acknowledgments

SA acknowledges award of the J.C. Bose National Fellowship by DST, New Delhi and equipment donation from Alexander Von Humboldt Foundation, Bonn, Germany. Technical assistance from Khem Singh Negi is acknowledged. We thank Debasish Mohanty for his critical comments on the manuscript. We express our deep sense of appreciation to the anonymous referees for their valuable inputs on our manuscript, which indeed have modified its quality.

### Funding

This work was supported by a core grant from the Department of Biotechnology (DBT) to the National Institute of Immunology, New Delhi and research grants nos. BT/PR11805/MED/12/424/2009, BT/PR14102/AAQ/01/438/2010 from DBT, New Delhi and no.SR/SO/AS-115/2012 from the Department of Science and Technology (DST), New Delhi to SA.

### References

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., ... Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, 25, 3389–3402.
Anand, N., Murthy, S., Amann, G., Wernick, M., Porter, L. A., Cukier, H., ... Lee, J. M. (2002). Protein elongation factor EEF1A2 is a putative oncogene in ovarian cancer. Nature Genetics, 31, 301–305.

Andersen, G. R., Pedersen, L., Valente, L., Chatterjee, I., Kinzy, T. G., Kjeldgaard, M., ... Nyborg, J. (2000). Structural basis for nucleotide exchange and competition with tRNA in the yeast elongation factor complex eEF1A: eEF1B alpha. Molecular Cell, 6, 1261–1266.

Chen, L., & Madura, K. (2005). Increased proteasome activity, ubiquitin-conjugating enzymes, and eEF1A translation factor detected in breast cancer tissue. Cancer Research, 65, 5599–5606.

Dever, T. E., Costello, C. E., Owens, C. L., Rosenberry, T. L., & Merrick, W. C. (1989). Location of seven post-translational modifications in rabbit elongation factor 1 alpha including dimethyllysine, trimethyllysine, and glycerylphosphorylethanolamine. Journal of Biological Chemistry, 264, 20518–20525.

Dever, T. E., Glynias, M. J., & Merrick, W. C. (1987). GTP-binding domain: Three consensus sequence elements with distinct spacing. Proceedings of the National Academy of Sciences of the United States of America, 84, 1814–1818.

Dominguez, C., Boelens, R., & Bonvin, A. (2003). HADDOCK: A protein–protein docking approach based on biochemical or biophysical information. Journal of the American Chemical Society, 125, 1731–1737.

Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., & Liang, J. (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Research, 34, 116–118.

Elcock, A. H., Sept, D., & McCammon, J. A. (2001). Computer simulation of protein–protein interactions. Journal of Physical Chemistry B, 105, 1504–1518.

Gallagher, W. M., Bergin, O. E., Rafferty, M., Kelly, Z. D., Nolan, I. M., Fox, E. J., ... Easty, D. J. (2005). Multiple markers for melanoma progression regulated by DNA methylation: insights from transcriptomic studies. Carcinogenesis, 26, 1856–1867.

Geourjon, C., & Delage, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Computer Applications in the Biosciences, 11, 681–684.

Honecker, F., Stoop, H., de Krijger, R. R., Chris Lau, Y. F., Arnemann, J., Vogel, T., Werner, M., & Blum, W. (2004). Pathobiological implications of the expression of markers of testicular carcinoma in situ by fetal germ cells. Journal of Pathology, 203, 849–857.

Kido, T., & Lau, Y. F. (2008). The human Y-encoded testis-specific protein interacts functionally with eukaryotic translation elongation factor eEF1A, a putative oncogene. International Journal of Cancer, 123, 1573–1585.

Kulkarni, G., Turbin, D. A., Amiri, A., Jeganathan, S., Andrade-Navarro, M. A., Wu, T. D., ... Lee, J. M. (2007). Expression of protein elongation factor eEF1A2 predicts favorable outcome in breast cancer. Breast Cancer Research and Treatment, 102, 31–41.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). Procheck – A program to check the stereochemical quality of protein structures. Journal of Applied Crystallography, 26, 283–291.

Lau, Y., Chou, P., Iezzoni, J., Alonzo, J., & Komuves, L. G. (2000). Expression of a candidate gene for the gonadoblastoma locus in gonadoblastoma and testicular seminoma. Cytogenetics and Cell Genetics, 91, 160–164.

Lau, Y. F., Lau, H. W., & Komuves, L. G. (2003). Expression pattern of a gonadoblastoma candidate gene suggests a role of the Y chromosome in prostate cancer. Cytogenetic and Genome Research, 101, 250–260.

Lau, Y. F., Li, Y., & Kido, T. (2011). Role of the Y-located putative gonadoblastoma gene in human spermatogenesis. Systems Biology in Reproductive Medicine, 57, 27–34.

Lee, S., Ann, D. K., & Wang, E. (1994). Cloning of human brain cDNAs coding for S1, the second member of the mammalian elongation factor-1 alpha gene family: analysis of a possible evolutionary pathway. Biochemical and Biophysical Research Communications, 203, 1371–1377.

Li, Y., & Lau, Y. F. (2008). TSPY and its X-encoded homologue interact with cyclin B but exert contrasting functions on cyclin-dependent kinase 1 activities. Oncogene, 27, 6141–6150.

Lund, A., Knudsen, S. M., Vissing, H., Clark, B., & Tommerup, N. (1996). Assignment of human elongation factor 1 alpha genes: eEF1A1 maps to chromosome 6q14 and eEF1A2 to 20q13.3. Genomics, 36, 359–361.

Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., ... Bryant, S. H. (2011). Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Research, 39, D225–D229.

Oram, S. W., Liu, X. X., Lee, T. L., Chan, W. Y., & Lau, Y. F. (2006). TSPY potentiates cell proliferation and tumorigenesis by promoting cell cycle progression in HEla and NIH3T3 cells. BMC Cancer, 6, 154.

Panasyuk, G., Nemazanyy, I., Filonenko, V., Negrutskii, B., & El'skaya, A. V. (2008). A2 isoform of mammalian translation factor eEF1A displays increased tyrosine phosphorylation and ability to interact with different signalling molecules. International Journal of Biochemistry & Cell Biology, 40, 63–71.

Rice, P., Longden, I., & Bleasby, A. (2000). EMBOSS: The European molecular biology open software suite. Trends in Genetics, 16, 276–277.

Rosenberry, T. L., Krall, J. A., Dever, T. E., Haas, R., Louvard, D., & Merrick, W. C. (1989). Biosynthetic incorporation of [3H] ethanolamine into protein synthesis elongation factor 1 alpha reveals a new post-translational protein modification. Journal of Biological Chemistry, 264, 7096–7099.

Šali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. Journal of Molecular Biology, 234, 779–815.

Salo, P., Kaariainen, H., Petrovic, V., Peltomaki, P., Page, D. C., & de la Chapelle, A. (1995). Molecular mapping of the putative gonadoblastoma locus on the Y chromosome. Genes, Chromosomes and Cancer, 14, 210–214.

Schneider, F., Dork, T., Arnesmann, J., Vogel, T., Werner, M., & Schmidtke, J. (1996). Testis-specific protein, Y-encoded (TSPY) expression in testicular tissues. Human Molecular Genetics, 5, 1801–1807.

Skaletsky, H., Tomok, K., Patrick, J. M., Holland, S. C., LaDeana, H., Laura, G. B., ... David, C. P. (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature, 423, 825–837.

Slobin, L. I. (1980). The role of eucaryotic factor Tu in protein synthesis. The measurement of the elongation factor Tu content of rabbit reticulocytes and other mammalian cells by a sensitive radioimmunoassay. European Journal of Biochemistry, 110, 555–563.
Molecular docking

van Dijk, A. D., & Bonvin, A. M. (2006). Solvated docking: introducing water into the modeling of biomolecular complexes. *Bioinformatics*, 22, 2340–2347.

Vogel, T., Dittrich, O., Mehrain, Y., Dechend, F., Schnieders, F., & Schmidtke, J. (1998). Murine and human TSPYL genes: Novel members of the TSPY-SET-NAP1L1 family. *Cytogenetics and Cell Genetics*, 81, 265–270.

Vogt, P. H. (2005). Azoospermia factor (AZF) in Yq11: Towards a molecular understanding of its function for human male fertility and spermatogenesis. *Reproductive BioMedicine Online*, 10, 81–93.

Vriend, G. (1990). *WhatIF* – A molecular modeling and drug design program. *Journal of Molecular Graphics*, 8, 52–56.

Wallner, B., & Elofsson, A. (2006). Identification of correct regions in protein models using structural, alignment, and consensus information. *Protein Science*, 15, 900–913.

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, 35, W407–W410.

Yin, Y. H., Li, Y. Y., Qiao, H., Wang, H. C., Yang, X. A., Zhang, H. G., … Chen, W. F. (2005). TSPY is a cancer testis antigen expressed in human hepatocellular carcinoma. *British Journal of Cancer*, 93, 458–463.