Predicted interaction of human Ribosomal Protein S15 with Fragile X Mental Retardation Protein

Edmund Ui-Hang Sim¹, Xiang-Ru Ma¹, Stella Li-Li Chan¹, Choon-Weng Lee², Kumaran Narayanan³,⁴

¹Department of Molecular Biology, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia. ²Institute of Biological Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia. ³School of Science, Monash University, Bandar Sunway, 46150 Selangor, Malaysia. ⁴Department of Genetics and Genomics Sciences, Mount Sinai School of Medicine New York, NY10029, USA.

ARTICLE INFO

Article history:
Received on: 12/01/2016
Revised on: 06/03/2016
Accepted on: 28/03/2016
Available online: 21/04/2016

Key words:
Ribosomal protein S15, FMRP, protein interaction, functional prediction, molecular docking

ABSTRACT

In addition to the central role of ribosome biogenesis, the human ribosomal protein S15 (RPS15) has extra-ribosomal roles that include its association with a congenital disease and a few types of cancer. However, current knowledge of its functions in the context of extra-ribosomal activities remains fragmented. An approach to gain insights into the interaction between RPS15 and possible protein partners is via Bioinformatics strategies. Based on the sequence-to-structure-to-function paradigm, structural data of a protein can be computationally analysed to derive logical interacting partners. This method can include three-dimensional model construction, structural neighbour prediction, and molecular docking analysis. By using this approach, we have constructed RPS15 3D-structural models that have allowed the prediction of 23 structural neighbours. Of these, two that are from human origin were further analysed and only one have logical prospect of binary protein-protein interactions. Further analysis of this structural neighbour revealed 7 candidate docking partners. From these, our molecular docking analysis demonstrated two most logical dock models of interactions between RPS15 with two different domains of the Fragile X Mental Retardation Protein 1 (FMRP1) protein. Hence, we have provided in silico evidence of de novo protein-protein interaction between RPS15 and the Fragile X Mental Retardation Protein 1 (FMRP1).

1. INTRODUCTION

Human ribosomal protein, S15 (RPS15) is a component of the 40S ribosome subunit, and is a member of the S19P family of ribosomal proteins. The RPS15 gene was originally identified as RIG, a human homologue of rig (rat insulinoma gene) [1, 2]. Analysis of RIG in human insulinoma tissues revealed elevated expression in tumors relative to normal pancreatic or regenerating islets [3]. Similar observation of activated expression in cancer tissues has also been explained for human esophageal and colon cancers [4]. The mammalian RPS15 gene sequence shows high homology (>88%) among different species, and the encoded protein is also highly conserved among mammalian species by having similarity in several functional domains, namely the two cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP)-dependent kinase phosphorylation, one ribosomal protein S19 signature, 4 N-myristoylation, and 5 casein kinase C phosphorylation sites [5]. During the processing of yeast pre-40S subunit, RPS15 has been shown to interact or bind directly with ribosome assembly factors such as the low temperature viability (Ltv1), and Rlght Open reading frame kinase 2 (RioK2 or Rio2) proteins [6]. In the case of Rio family of protein kinases, depletion or reduction of RPS15 significantly decreased the level of human RioK2 and -K3 [7], suggesting direct interaction between RPS15 and these factors during 40S ribosome biogenesis. Besides its involvement in ribosome biogenesis, the association of RPS15 with diseases has been reported, albeit of mechanism(s) that is largely undefined. For example, like many other RP genes, RPS15 has been suspected to be associated with Diamond Blackfan Anemia (DBA) [8].

* Corresponding Author
Dr. Edmund Ui-Hang Sim, Department of Molecular Biology Faculty of Resource Science and Technology, Universiti Malaysia Sarawak 94300 Kota Samarahan Malaysia. Tel: +60-82-583199 and/or -583041 Fax: +60-82-583160, E-mail: uh_sim@frst.unimas.my

© 2016 Edmund Ui-Hang Sim et al. This is an open access article distributed under the terms of the Creative Commons Attribution License -NonCommercial-ShareAlike Unported License (http://creativecommons.org/licenses/by-nc-sa/3.0/).
However, the link between RPS15 and DBA is weak. Evidence of genetic lesion is relatively rare where only single missense mutation was found in less than 1% of cases studied, and the mutation is neither correlated with DBA pathogenicity nor effect carcinogenesis in the affected individual [8]. This coincide with findings from studies using Zebrafish model where tumours are absent in RPS15 mutants [9, 10]. In a study involving hepatocellular carcinoma (HCC, liver cancer) cells, RPS15 expression was observed to be significantly different between hepatitis B virus (HBV) - and hepatitis C virus (HCV) – associated HCC cells [11].

From the microarray data of a large pool of differentially expressed genes between primary tumour biopsies of nasopharyngeal carcinoma (NPC) and non-cancer nasopharyngeal tissues, RPS15 was found to be significantly up-regulated in NPC samples [12]. Despite the proven association of RPS15 with HCC and NPC, its molecular effect(s) or roles in the tumorigenic event of these two cancer types is yet to be unravelled. The understanding of interaction between RPS15 and cancer-associated factors is also largely incomplete, and there has only been one study so far [13], that demonstrated a possibility of RPS15 in regulating the tumour suppressor, p53, via binding to Mdm2 and MdmX. Whether this in vitro finding provides an explanation to part of RPS15’s role in carcinogenesis remains to be verified.

Knowledge on interacting partners of RPS15 in pathogenesis is limited and fragmented. To gain pertinent insight into its role in molecular interaction and functional complexes, the approach of bioinformatics and in silico analysis is plausible. In theory, this can be achieved by deriving the functions of proteins based on the computational data analysis of their sequence, structure and association with other proteins [14]. This form of analysis is based on the sequence-to-structure-to-function paradigm.

Firstly, it involves the construction of three dimensional (3D) structural models of studied proteins, which in turn, will be used as templates for identifying structural neighbours. Secondly, following the assumption that close homologs or analogues almost always interact in the same way [15], knowledge of known interactions between structural neighbours and other molecules/partners would be used to derive possible interacting partners of proteins under study. Finally, simulated molecular (protein-protein) docking assay of the 3D-models of targeted protein with its candidate interacting partners will reveal predicted de novo functional complexes.

The docking analysis involves the computational construction of structural complexes by the logical fitting of two or more reliable 3D models – taking into account the conceptual basis that proteins associate at their interacting sites based on surface complementarity and electrostatics [16]. Herein we provide new theoretical insight into the activities of RPS15 in our report on in silico evidence of interaction between RPS15 and the Fragile X Mental Retardation Protein 1 (FMRP1).

3. METHODS

3.1 Overview of methodology

Our analytical procedures involve 5 stages, namely (i), identification of suitable structural templates of RPS15 using standard protein Basic Local Alignment Search Tool (protein BLAST) and ClustalX program; (ii), Three dimensional (3D) protein modeling using SWISS-MODEL workspace; (iii), identification of protein structural neighbors using Vector Alignment Search Tool (VAST) program; (iv), identification of candidate partners for each structural neighbor from IntAct database; and (v), protein-protein docking analysis via ClusPro 2.0 server.

3.2 Identification of structural templates

Amino acid sequence of human RPS15 protein (reference no. NP_001009.1) was retrieved from the National Center for Biological Information (NCBI) protein database. The search for templates was performed by entering this sequence data into the ‘protein BLAST’ site to search against non-redundant protein database. The Position-Specific Iterated BLAST (PSI-BLAST) algorithm was selected for this analysis, and two PSI-BLAST iterations were performed. Amongst the hits, only templates with structures of expected value (E-value) better than threshold for high sequence identity (at least 99%) and structure resolution were selected for further analysis.

Target-template alignment data was obtained by the multiple sequence alignment of amino acid sequences of BLAST-retrieved templates (PDB ID: 2ZKQ_S and 4KZX_P), target protein (protein accession no. NP_001009.1), and experimental data. The experimental data was from translated RPS15 transcript sequences of 10 NPC tissue samples. Alignment was performed on the ClustalX program, and Jalview workbench (Version 2) [17, 18] was used to illustrate the analysis output.

3.3 Three-dimensional (3D) protein modeling

The SWISS-MODEL workspace [19-21] was used to construct the three dimensional model of RPS15 based on homology modelling strategy. In this study, RPS15 was comparatively modelled with template structures using data from target-template alignment result. The generated 3D model was then evaluated using Qualitative Model Energy Analysis (QMEAN) [22] and PROCHECK tools [23] for model quality estimation and structural assessment respectively. QMEAN assessment was by referring to the composite scoring function of QMEAN (QMEAN6 score) [24, 25] obtained via structural analysis of constructed model using the QMEAN website. Refined stereochemical quality of RPS15’s 3D structure was assessed using the Ramachandran plot generated from PROCHECK (a program hosted by the SWISS-MODEL workspace) analysis. Models with QMEAN6 score higher than 0.5 and with less bad contacts (residues that are located on disallowed region in Ramachandran plot) were analysed further.
Fig. 1: Result of multiple sequence alignment of template to RPS15 protein sequences. Labels A and C are the sequences of BLAST-retrieved structural templates with PDB accession number and chain code of 2ZKQ_S and 4KZX_P, respectively. B is the target protein (RPS15) sequence (accession no.: NP_001009.1). Sequences T20, T22, T24, T27, T39, T41, T48, T55, T67 and T96 F & R are experimental data from translated RPS15 transcripts sequences of NPC samples. Color reference in alignments is according to ClustalX scheme (orange: glycine G; gold: proline P; blue: small and hydrophobic amino acids A, V, L, I, M, F, W; green: hydroxyl and amine amino acids S, T, N, Q; magenta: negative-charged amino acids D, E; red: positive-charged amino acids R, K; dark-blue: histidine H and tyrosine Y). Numbering of sequence is according to the RPS15 reference protein sequence.
3.4 Identification of structural neighbours and candidate partners

The comparative 3D-model of RPS15 was searched against medium redundancy subset of PDB structures for potential structural neighbours using Vector Alignment Search Tool (VAST) [26] that is hosted by NCBI website. Structural neighbour proteins of human origin were selected and the IntAct database [27, 28] was used to retrieve data of their interactions with candidate partners. In order to narrow down the number of candidate partners for docking, PROSITE scan analysis (via ScanProsite) [29] was performed and only human origin candidate partners that have high sequence pattern similarity to one another were selected.

3.5 Protein-protein docking simulation

Computational docking of the comparative 3D-model of RPS15 with its candidate partners was carried out using the web-based program, ClusPro 2.0 [30, 31]. For our analysis, the coordinate files of RPS15 and candidate partner protein structures were uploaded through ClusPro web interface. The RPS15 and candidate partners’ structures served as receptor and ligands, respectively. Evaluation of the docked complexes was based on output score of cluster and energy. Any docked complex that was bound to other ligands or molecular structures which did not display single protein model-candidate partner interaction was excluded from the analysis. The protein-model-candidate partner complex was later examined on SWISS-PdbViewer v4.1 (a.k.a. DeepView) [32]. For our analysis, the root mean square deviation (RMSD) was also computed to predict the potential interaction sites based on interface contact residues (< 5 Å).

4. RESULTS

4.1 Structural 3D models of RPS15

Following protein BLAST search, two structural templates are selected, namely the s chain of a mammalian ribosomal 40s subunit within a 80s complex (PDB: 2ZKQ_S), and the P chain of rabbit 40s ribosomal subunit in complex with Eif1 (PDB: 4KZX_P). The 2ZKQ_S and 4KZX_P structures have 100% (E-value of 2e-103) and 99% (E-value of 1e-102) identity to the queried sequence, respectively (human RPS15 protein). The 2ZKQ_S structure is from Canis lupus familiaris and has a structural resolution of 8.7Å, whereas 4KZX_P with structural resolution of 7.81Å is from Oryctolagus cuniculus. Multiple sequence alignment of RPS15 target sequence (NP_001009.1), selected structural templates (2ZKQ_S and 4KZX_P) and 10 translated protein sequences from NPC samples shows a highly consensus region (Figure 1). PfamA search assay of the RPS15 target sequence revealed a Ribosomal Protein S19 domain defined by Residues 47-128 with an e-value of 1.2e-35. This domain is located within the highly consensus region of the alignment (Figure 1), and ScanProsite analysis revealed a S19 family signature motif (Residues 98-122) within the domain. Results from this alignment were used for 3D-modeling analysis. The 3D-structural model of RPS15 protein (Figure 2) constructed using the 2ZKQ_S structure as template sequence was selected after model quality and structural assessments. This model gave a QMEAN6 score of 0.662 and Z-score of -0.664 (Table 1).

PROCHECK analysis reveals a Ramachandran plot showing 75.3% of the total residues in core regions and three residues (His 27, Val 48 and Glu 91) ain disallowed high energy regions (Figure 3). Amongst these 3 residues, Val 48 and Glu 91 are located within the ribosomal protein S19 domain (amino acids 47-128). The RPS15 model constructed by SWISS Model is an 88-amino-acid structure which characterized by presence of one α helix and four β strands (Figure 2a). The ribosomal protein S19 domain encompasses almost the entire constructed model (Figure 2b).

![Fig. 2: RPS15 3D-structural protein model. In (a), α-helix structure is indicated in red, whereas β-sheet structure is represented by yellow color (N = N-terminal, C = C-terminal); and in (b), the ribosomal protein S19 domain is indicated in red. These diagrams were rendered by SWISS-PdbViewer v4.1.](image)

<table>
<thead>
<tr>
<th>Scoring function term</th>
<th>Raw score</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-beta interaction energy</td>
<td>-32.24</td>
<td>-0.60</td>
</tr>
<tr>
<td>All-atom pairwise energy</td>
<td>-1636.00</td>
<td>-1.01</td>
</tr>
<tr>
<td>Solvation energy</td>
<td>-10.93</td>
<td>0.63</td>
</tr>
<tr>
<td>Torsion angle energy</td>
<td>-10.55</td>
<td>-1.55</td>
</tr>
<tr>
<td>Secondary structure agreement</td>
<td>81.8%</td>
<td>-0.23</td>
</tr>
<tr>
<td>Solvent accessibility agreement</td>
<td>78.4%</td>
<td>0.45</td>
</tr>
<tr>
<td>QMEAN6 score</td>
<td>0.662</td>
<td>-0.664</td>
</tr>
</tbody>
</table>

4.2 Structural neighbours and candidate partners

The VAST analysis of constructed RPS15 3D-model revealed 23 structural neighbours, in which two structures: the crystal structure of KHI and KH2 domains from Human Fragile X Mental Retardation protein (PDB ID: 2QND); and the crystal structure of human Rab27b bound to Gdp (PDB ID: 2F7S) were selected for further analysis because they are human proteins. Based on IntAct database, 2QND was found to have 36 binary protein-protein interactions whereas 2F7S did not return any known binary interaction. Of the 36 binary interactions identified, protein partners of non-human origin and with ‘spike’ as expansion method were excluded to minimize chances of false positives.
Fig. 3: Results (output data) of Ramachandran Plot and Procheck analysis.

Fig. 4: Protein-protein interaction models of RPS15: (a), RPS15_2BKD dock model; (b), RPS15_2QND dock model; (c) and (d), logical interaction model between S19 signature of RPS15 (red) and Agenet like domain 1 of FMRP1 (green); and (e) and (f), logical interaction model between RPS15 (pink) and the two chains (A and B) of Type 1 KH domain of FMRP1 (green). In (a) and (b), the red region represents ribosomal protein S19 signature of RPS15. Generally, RPS15 is indicated in red whereas candidate interacting partner is in yellow. The wireframe representation (b) illustrates interface contact residues that are close to another chain (<5 Å), hence depicting possible interaction sites. In (c) – (f), contact residues are represented as blue ribbon [in (c) and (e)] and wireframe [in (d) and (f)].
To narrow down our search, PROSITE scan analysis revealed seven candidate partners. These are the x-ray crystal structure of human Stau1 SSM-1bd5 domain-swapped dimer (PDB ID: 4DKK); solution NMR structure of N-terminal domain of Fragile X Mental Retardation protein (PDB ID: 2BKD); solution NMR structure of KH1 from the Fragile X Mental Retardation protein 1 (PDB ID: 2FMR); x-ray crystal structure of KH1-2 domains from human Fragile X Mental Retardation protein (PDB ID: 2QND); solution NMR structures of 1st (PDB ID: 1G47) and 4th (PDB ID: 1NYP) LIM domain of Pinch protein; and solution NMR structure of 4th LIM domain of Pinch protein; and the solution NMR structure of the 3rd LIM domain of a particularly interesting new Cys-His protein (PDB ID: 2COR). These were identified based on most similar PROSITE patterns of 2QND structural neighbour with RPS15 model.

4.3 Molecular docking

Our ClusPro 2.0 analysis yielded dock model for each of the seven RPS15-candidate partner (Table 2). The number of clusters for each dock model is 26 for RPS15_4DKK, 30 for RPS15_2BKD, 23 for RPS15_2FMR, 30 for RPS15_2QND, 25 for RPS15_1G47, 30 for RPS15_1NYP, and 30 for RPS15_2COR. For each dock model, the clusters with the highest size and lowest energy scores is considered the likeliest docking scenario (Table 2).

Table 2: ClusPro scores of RPS15-candidate partner dock models.

<table>
<thead>
<tr>
<th>Dock model</th>
<th>Cluster size (number of members)</th>
<th>Center free energy (kcal/mol)</th>
<th>Lowest free energy (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPS15_4DKK</td>
<td>125</td>
<td>-857.7</td>
<td>-867.8</td>
</tr>
<tr>
<td>RPS15_2BKD</td>
<td>126</td>
<td>-812.9</td>
<td>-890.5</td>
</tr>
<tr>
<td>RPS15_2FMR</td>
<td>282</td>
<td>-844.9</td>
<td>-1027.1</td>
</tr>
<tr>
<td>RPS15_2QND</td>
<td>78</td>
<td>-844.1</td>
<td>-944.8</td>
</tr>
<tr>
<td>RPS15_1G47</td>
<td>162</td>
<td>-759.7</td>
<td>-938.1</td>
</tr>
<tr>
<td>RPS15_1NYP</td>
<td>95</td>
<td>-593.4</td>
<td>-858.9</td>
</tr>
<tr>
<td>RPS15_2COR</td>
<td>128</td>
<td>-651.4</td>
<td>-753.6</td>
</tr>
</tbody>
</table>

Amongst the seven RPS15-candidate partner dock models, only RPS15_2BKD and RPS15_2QND show potential protein-protein interaction based on presence of interface contact residues that are close (<5 Å) to another chain (Figure 4a and b). Both 2BKD and 2QND were identified as domains found on Fragile X Mental Retardation Protein 1 (FMRP1). Functional interaction sites based on RPS15_2BKD dock model are between the ribosomal protein S19 signature of RPS15 (residues Lys108, Pro109, Glu110, Met111, Ile112, Gly113, His114, Tyr115, Leu116, and Phe119), and the Agenet-like motif 1 of FMRP1 (residues Leu4, Val5, Phe19, Val20, Lys21, Asp22, Val23, Glu25, Pro 50) (Figure 4c and d). On the other hand, the interaction sites for RPS15_2QND dock model are between the residues outside of the ribosomal protein S19 signature (Arg43, Arg44, Leu45, Asn46, Arg47 and Arg50), and the two chains of Type 1 K-homology (KH) domain of FMRP1, namely in Chain A (residues of Arg75, Asn76, Leu77, Val78, Gly79, Lys80, Val81, Ile82, Gly83, Lys84, Asn85, Gly86, Lys87, Ile89, Gln90, Val93, Asp94, Gly97, Val98, Val99, Arg100, Val101, Arg102, Ile 103, Glu104, Ala105, Glu106, Asn107, Glu108, Lys109, Asn110, Gln113) and Chain B (residues of Arg11, Glu12, Asp13, Gln72, Arg75, Asn76, Val78, Gly79, Ile82, Gly83, Asn85, Gly86, Glu90, Val93, Asp94, Lys95, Gly97, Val98, Val99, Arg100, Val101, Arg102, Ile103, Glu104, Ala105, Glu106, Asn107, Glu108, Lys109, Asn110, Val111, Pro112, Glu113, Glu114, Pro119, Val121) (Figure 4e and f).

5. DISCUSSION

We revealed in silico evidence of protein-protein interactions between RPS15 and different domains of FMRP1. The first dock model (RPS15_2BKD) shows interaction between RPS15 and the Agenet-like Motif 1 region of FMRP1. Located in the N-terminus domain of FMRP (NDF), this motif harbours a highly conserved region of 134 amino acids rich in β structure [33]. Functional significance of NDF region is seen in bio-molecular interactions such as protein-RNA, protein-protein, and dimerization context [33, 34]. It consists of two Agenet-like/Tudor motif repeats. Agenet-like 1/NDF1 and 2/NDF2, followed by a short α helix [34]. Although structurally similar, the dynamic and functional properties of these two repeats vary slightly where the second repeat is more flexible and could interact with methylated lysine and a FMRP nuclear partner, 82-FIP [34]. Our findings, however, demonstrate that the interaction between that RPS15 and FMRP1 in this location does not involve the second repeat (Agenet-like/NDF 2), but instead only with the first repeat (Agenet-like 1/NDF 1).

The second dock model (RPS15_2QND), indicates interaction of RPS15 with the KH domain 1 of FMRP. KH domains are nucleic acid binding motifs with tandem arrays of 2-16 repeats that bind RNA or ssDNA, and are also found in proteins associated with transcriptional and translational regulation [35]. There are two homologous KH motifs in FMRP – the first one is folded whereas the second remains unfolded even when extended at both N- and C-terminal [36]. Only the first KH motif has been demonstrated to bind RNA [36].

It is interesting to note that our data predicted the phenomenon of RPS15’s association with two different sites in the FMRP1 protein. However, we argue that RPS15 protein most likely interacts with FMRP1 via the Agenet-like motif 1 of the N-terminus domain and less likely via the KH1 motif. Based on our analysis, RPS15 interacts with Agenet-like motif via residues that lie within the ribosomal protein S19 signature motif whereas its interaction with KH1 motif is via residues not residing in the S19 signature. Furthermore, the Agenet-like 1 (NDF1) repeat has 5 beta strands matching that of SMN Tudor domain, of which binds to the Sm D1 and D3 proteins [37]. This shows that the Agenet-like 1 (NDF1) repeat is capable of protein-protein interaction – a scenario yet to be proven for the KH1 motif of FMRP1. It should also be noted that of the 4 heteroatom groups in the S15_2QND dock model, one of these falls within the KH1 motif and lack proper CONNECT card information. The lack of this information
will hinder DeepView from generating bonds and force it to make guesses of molecular structure based on atomic coordinates, thus may contribute to false positives. Nevertheless, while the likelihood of RPS15’s interaction with the Agenet-like motif I site of FMRP1 is theoretically plausible, the functional significance of this interaction remains to be empirically explored.

The FMRP1 protein is encoded by the *FMR1* gene, and its absence is a causative factor of mental retardation in the congenital condition of Fragile X Syndrome [38, 39]. The molecular mechanism of this is largely unclear, and whether there is an involvement of RPS15 in this is difficult to speculate. One of the possible roles of FMRP1 is in the negative regulation of translation [40, 41], where it strongly inhibits translation of various mRNAs [40]. It has also been reported to be associated with polyribosomes as a component of the mRNP complex that transports mRNA from nucleus to cytoplasm [42, 43]. On the other hand, RPS15 is known to be involved in ribosome biogenesis and function [6]. Therefore, could the regulation of ribosome-mediated translation include the sequestration of negative translation regulators such as FMRP1 by RPS15? The answer to this suspicion is yet to be experimentally investigated. Nevertheless, the findings herein represent conceptual insights to the role(s) of RPS15 in the modulation of translation activities, of which are pertinent for the design of functional studies pertaining to the molecular roles of RPS15.

6. CONCLUSIONS

Structural analysis coupled with molecular docking simulation via computational strategy has allowed the prediction of protein-protein interaction between RPS15 and two domains of FMRP1, suggesting a resultant effect of RPS15-mediated negative regulation of translation.

7. ACKNOWLEDGEMENTS

This study is funded by the Ministry of Education Malaysia via the Research Acculturation Collaborative Effort Program [Grant code: RACE/a(1)/883/2012(1)], and the University of Malaya via its Research Partner Program (Grant code: UM.TNC2/IPP/PPGP/261/PRPUM).

8. REFERENCES


23. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of