Pakistan J. Zool., vol. 48(3), pp. 639-644, 2016.

In Silico Analysis of Missense Substitutions in *RB1* Gene and Their Effect on Metabolic Pathways

Saeeda Kalsoom,[1,2]* Khushnooda Ramzan,[3,4] Honainah Ishaq Tahir,[5] Ali Raza Awan,[1] Aftab Ahmed Anjum[6] and Muhammad Wasim[1]

[1]*Institute of Biochemistry and Biotechnology, University of Veterinary and Animal Sciences, Lahore, Pakistan*
[2]*Institute of Molecular Biology and Biotechnology. The University of Lahore, Pakistan*
[3]*Department of Genetics, King Faisal Specialist Hospital & Research Centre, Riyadh, Saudi Arabia*
[4]*BioTech Labs, Chemhouse, Lahore, Pakistan*
[5]*Department of Medical Education, King Khalid University Hospital, King Saud University, Riyadh, Saudi Arabia*
[6]*Department of Microbiology, University of Veterinary and Animal Sciences, Lahore, Pakistan.*

Saeeda Kalsoom and Khushnooda Ramzan contributed equally to this work

Article Information
Received 22 May 2015
Revised 18 October 2015
Accepted 25 October 2015
Available online 14 March 2016

Authors' Contributions
SK and KR designed the study and wrote the article. HIT studied protein modeling. ARA and AAA helped in data interpretation. MW supervised the study.

Key words
Retinoblastoma, *RB1* gene, homology modeling, Rb1 pathway, PyMol.

ABSTRACT

Retinoblastoma (RB) is a neonatal intraocular tumor caused by biallelic inactivation of RB1 gene encoding 928 amino acids long nuclear phosphoprotein (pRB), which contains three important regions: N-terminal domain, central AB-box and C-terminal domain. Homology modeling was performed for three dimensional structure prediction of pRB, to predict the potential consequences of three missense mutations (p.Thr307Ile, p.Leu688Pro, p.Asp316Ile) identified by direct sequencing in Pakistani RB patients. Keeping in view the importance of molecular diagnosis, we performed in silico protein analysis using PROSA and project HOPE to predict the possible structural changes in the mutant RB protein. Our analysis concludes that although these mutations do not drastically affect the binding affinity of the protein with its potential legends but seem to be disrupting the normal RB1/E2F pathway leading to deregulation of cell growth control.

INTRODUCTION

Retinoblastoma (RB; MIM# 180200) is an intraocular tumor of retina that mostly develops during early infancy with leukocoria as first clinical sign. Hereditary mutation (40%) in RB1 gene (Genebank accession no.L11910) disrupt the synthesis of functional nuclear phosphoprotein (pRB) and usually causes bilateral and multifocal tumor while somatic mutations (60%) in developing retinal cells cause unilateral tumor. The average age of onset is 19.5 months, which in case of bilateral and unilateral RB is reported as 15 and 24 months, respectively (Lohmann and Gallie, 2013). The worldwide incidence rate among live births was reported as 1:20,000 (Kivelä, 2009). RB1 was reported as first tumor suppressor gene (Weinberg *et al.*, 1986) which play important role in regulation of the cell cycle progression, cell differentiation and apoptosis (Weinberg, 1995).

The main components of RB pathway involves five family proteins; retinoblastoma tumor suppressor RB-family of pocket proteins (p110, p107, p130), E2F-family of transcription factors (heterodimers of E2F1-8 with DP1-2), cyclins (D and E-type), cyclin-dependent protein kinases (cdk4, cdk6) and cyclin-dependent kinase inhibitor CDKIs (p16Ink4a, p15Ink4b, p18Ink4c and p19Ink4d). The functional activity of these proteins are well deliberated as the hypophosphorylated pRB has the ability to restrict cell cycle progression at the G1/S transition of the cell cycle by binding E2F transcription factors, which is thought to be essential for its tumor suppression function. The synthesis of Cyclin D starts during G1 and helps in cell cycle transition from G1 to S phase. Cyclin-dependent kinases (CDKs) contains kinase domain with little kinase activity but on binding with Cyclin, the Cyclin/CDK become an active complex and cause the hyperphosphorylation of pRB that ultimately involve in regulation of transcription and mRNA processing while Cyclin/CDK complex is itself regulated by cell cycle inhibitors of the INK4 (inhibitors of CDK4) and CIP/KIP (CDK interacting protein/kinase inhibitory protein) families, especially p16Ink4a and p21Cip1 (Dyer and Bremner, 2005; Sage, 2012).

After M phase either a cell enters in another round of cell cycle or exits in resting phase (G0; differentiation, senescence or apoptosis) which is all well regulated. On initiation of cell cycle the hypo- and hyper-phosphorylated form of pRB control further progression of cell cycle through G1/S phase by crossing the check point (R point). At early G1 phase, the Cyclin/CDK

* Corresponding author: saeeda.kalsoom@imbb.uol.edu.pk.
0030-9923/2016/0003-0639 $ 8.00/0

complex hypo-phosphorylate pRB, which block cell cycle progression into S-phase by sequestering the E2F family of transcription factors. These factors are responsible for the transcription of genes required for DNA replication, apoptotic pathway and cell cycle regulation (Sherr, 1996; Weinberg, 1995). The blockage of cell cycle progression at G1 is unblocked on receiving external signal i.e. growth hormones that directly target Cyclin E/CDK2 complex which on activation hyper-phosphorylate the pRb at G1. The hyper-phosphorylated pRB is unable to sequester E2F transcription factors, which then get released and available for the transcription of all the required genes for cell cycle regulation, particularly Cyclin E gene product that directly bind with cdc2 and rapidly starts DNA replication (Dyer and Cepko, 2001;. Knudsen and Wang, 2010).

Retinoblastoma-associated protein pRB constitutes of three domains, N-terminal (111–228 amino acid), A-B box (373–786 amino acid) and C-terminal (768–927 amino acid). The RB1 sequence was used for homology modeling, which is a technique to predict the structure of a target protein based on its sequence and considering a homolog structure as a template. Protein structure helps in the assessment of protein folding and the change in the interaction between different proteins. Thus the tertiary structure of pRB was predicted to evaluate the changes in folding and protein stability due to mutations. The mutated pRB protein structure was based on previously found two known missense mutations c.920C>T→p.Thr307Ile (exon 9) and c.2063T>C→p.Leu688Pro (exon 20) in three bilateral RB patients. While one novel missense mutation c.947A>T→Asp316Ileu (exon 10) was also found in a unilateral sporadic patient, using direct sequencing method. By using different protein modeling software we did protein analysis and predicted the potential consequences of the missense mutations on 3D protein structure of mutated pRB.

MATERIALS AND METHODS

3D protein structure prediction

In humans, RB1 gene encodes a protein pRB of 928 amino acid residues (Uniprot P06400). The sequence of RB1 was taken from NCBI and the potential PDB templates for the generation of the structure were found by performing Blastp (http://blast.ncbi.nlm.nih.gov/) (Johnson *et al.*, 2008). Sequence alignment was performed using ClustalW (http://www.ebi.ac.uk/clustalw) (Larkin *et al*., 2007). The tertiary structure of the normal and mutant pRB, were predicted through Modeller (Sali *et al.*, 1995).

Protein analysis

Various tools were used to analyze protein structure. Ramachandran plot was conducted using PROCHECK (Laskowski *et al.*, 1993) which assess the conformational quality of the polypeptide backbone and side chains. The structural reliability of models was evaluated via ERRAT (Colovos and Yeates, 1993), Verify3D (Eisenberg *et al*., 1997), Z-Score and Energy Plot (Wiederstein and Sippl, 2007). Project Hope (http://www.cmbi.ru.nl/hope/home) was used to analyze the change affecting the protein. Both the normal and mutated pRB structures were compared and analyzed for their stability and pathological character using I-Mutant (http://folding.biofold.org/cgi-bin/i-mutant2.0.cgi) and PMut (http://mmb2.pcb.ub.es:8080/PMut/) softwares respectively (Capriotti and Casadio, 2005; Ferrer-Costa *et al*., 2005). Molecular graphics were performed using VMD Software (Humphrey *et al*., 1996) and PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC)

RESULTS

Structure alignment of pRB (Uniprot P06400) showed 96% similarity with PDB template (4ELJ_A) and a score of 98.1707 was predicted by ClustalW, thus it was selected on the basis of high alignment score. Initially five models were generated for the normal tertiary structure of pRB through Modeller 9v12. Best modelled structure was selected on the basis of low energy level (Fig. 1A). Mutated structures for two already reported mutations p.Thr307Ile and p.Leu688Pro were predicted in a similar way using Modeller (Fig. 1B).

The novel mutation p.Asp316Ile was predicted with the help of Modeller 9.v10. The structure was chosen according to the MOLPDF score (18109.44727) and DOPE score of -75272.22656. molpdf is the standard MODELLER scoring function and is simply the sum of all the restraints. DOPE or Discrete Optimized Protein Energy, is a statistical potential used to assess homology models in protein structure prediction. The best structure is chosen on the basis of the lowest molpdf and DOPE score (Fig. 2).

Structures were then analyzed using various tools. Ramachandran plot for the normal pRB showed 91.7% residues in the favorable regions. Verify_3D predicted the score of 79.20%. ERRAT gave the strong value of 80.000. The Z-score of the model was -12.2 lying in somewhat comfortable region. Similarly for the mutated pRB structure for (p.307 and p.688), Ramachandran plot showed 92.0% residues in the favorable regions.

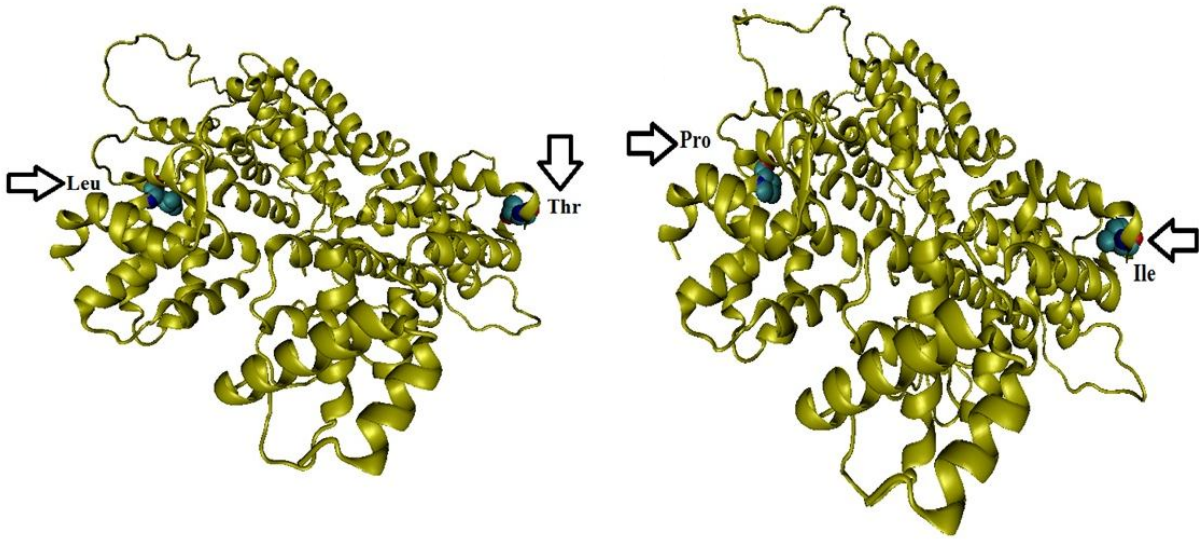

Fig. 1. Structure analysis for the Retinoblastoma protein (pRB). (A) Visualization of normal pRB having wild type Thr307 and Leu688, predicted using MODELLER 9v12. (B) Mutant pRB structure having Ileu307 and Pro688 residues at their respective positions in patients. The respective amino acids are shown in blue spheres showing their side chain. The images have been produced by VMD software.

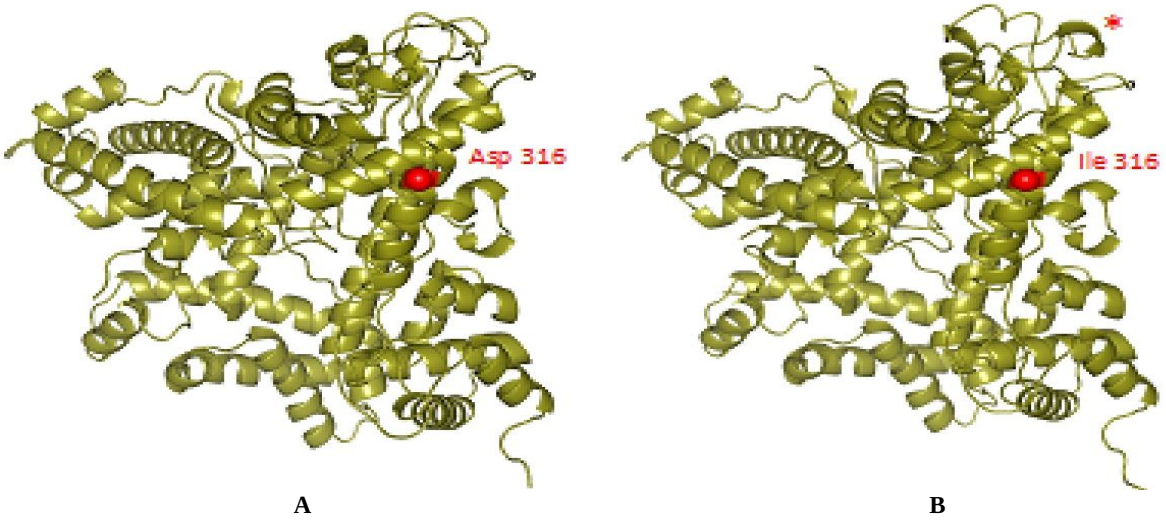

Fig. 2. Predicted structure of normal and mutated pRB: (A) Preferred predicted structure of normal pRB having wild type Asp316 (B) Mutant protein structure having Ileu at the respective position with red sphere. An extra helix is shown by "*". As the mutation is present in the helix, it will disturb the correct folding of pRB and which will be damaging to the protein. The images have been created using PyMOL.

Verify_3D predicted the score of 82.94% and ERRAT gave the strong value of 78.636. The Z-score of the model was -12.4 which is in the acceptable region. Energy plots were also plotted and gave reasonable energy values for both the structures. Protein structure analysis for the normal pRB protein is shown in Figure 3. For the mutant protein structure p.Asp316Ile, Ramachandran plot showed 92.2% residues in the favorable regions. Verify_3D predicted the score of 62.35% and ERRAT gave the value of 50.672. The Z-score of the model was -3.29 which is in the acceptable region. Energy plots were also plotted and gave reasonable energy values for both the structures.

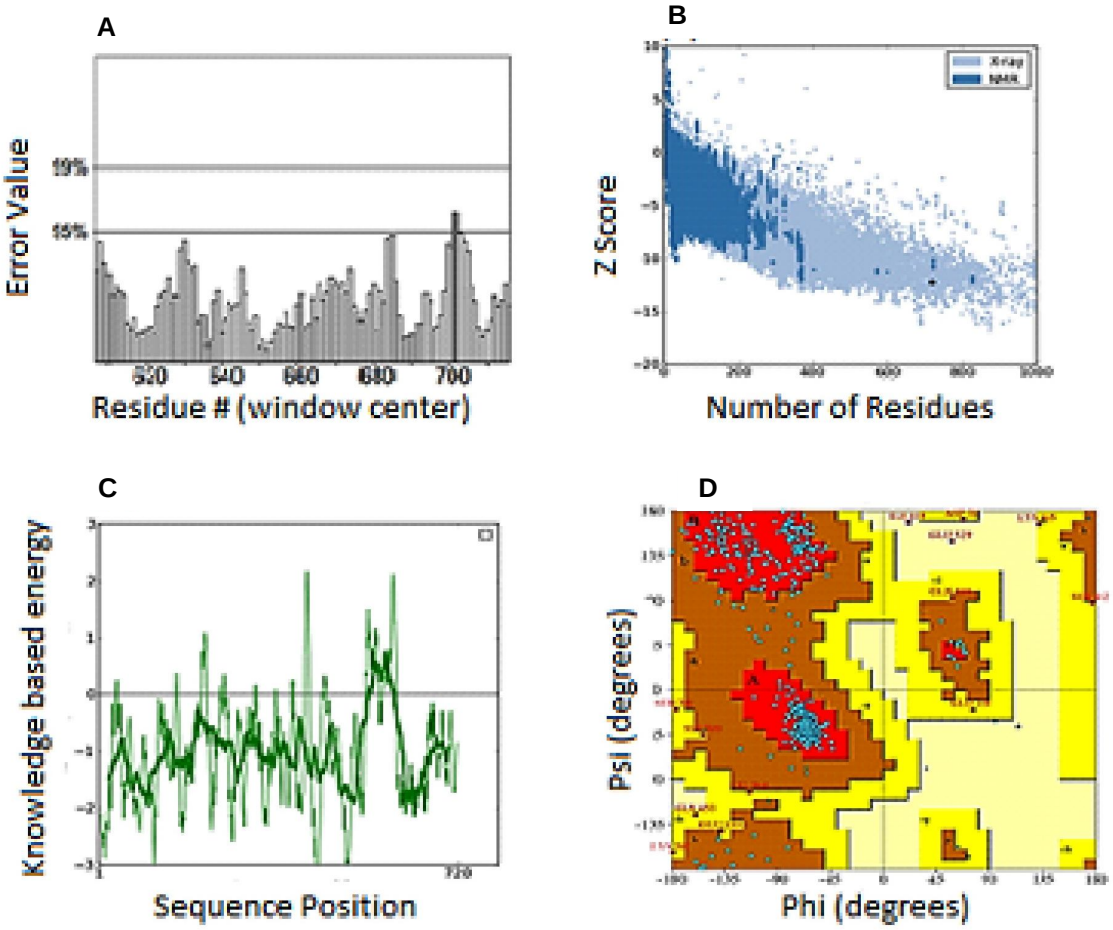

Fig. 3. Protein Structure Analysis for the normal pRB protein: (A) Errat plot, Overall quality factor: 80.000 (B) Z-score plot, Value: -12.2 (C) Local model quality (D) Ramachandran plot: 91.7% residues in favored region, 6.4% in additionally allowed region, 1.6% in generously allowed region.

DISCUSSION

Mutations that affect the pocket domain can simultaneously disrupt many aspects of pRB and have made dissecting interactions within this domain very difficult (Dick, 2007).

Missense mutation p.Thr307Ile sits in an unstructured loop located in between N-terminal and A-box. A-box is the binding site for E2F. Thus this mutation might confer instability and affect the essential binding site and critical functional changes in neighboring domains.

Analysis for the mutation p.Leu688Pro revealed that it is present in highly conserved pocket domain-B that contain LXCXE motif which is also shared by other cellular and RB binding viral proteins. The difference in this domain can affect the E2F inhibition (Burke *et al.*, 2012). The A-domain is essential for proper and stable folding of the B-domain. The A and B interface is also well conserved and act as secondary binding site for other proteins while cyclin fold structural motif is also the part of both A and B domains. Various ligands including E2F1 and SKP2 also bind to the cyclin box fold domain. The p.Leu688Pro mutation is predicted to possibly affect the normal binding of these major ligands (Fig. 4).

For the novel missense mutation p.Asp316Ileu, it was found that although the asparagine residue is not conserved at 316 position, the mutant isoleucine was not found in other homologous sequences. Due to smaller size of the isoleucine residue, an empty space in the protein is caused that might lead to possible rearrangements of the surrounding residues. Any

hydrogen bond made by Asparagine will also be lost, because isoleucine is a hydrophobic residue. In view of its location in the three dimensional structure, these changes may result in loss of interactions and disturb the correct folding of pRB, which will be damaging to the protein. This region binds independently to the other E2F domain (Xiao *et al*., 2003).

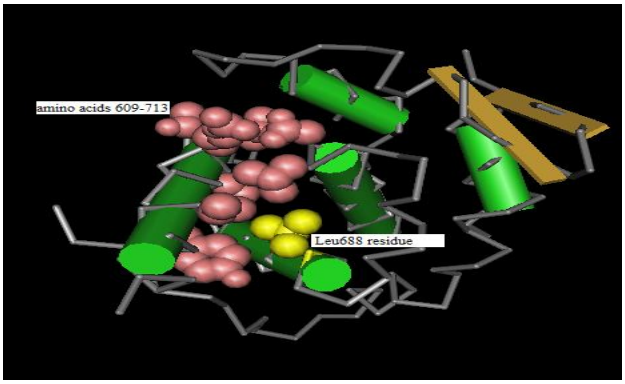

Fig. 4. 3-D spatial structure of pRB: Highlighting LXCXE binding site from amino acids 609-713. The residue Leu688 is also shown. The mutation p.Leu688Pro might possibly affect the normal binding of LXCXE ligand.

According to I-Mutant software, the stability of the protein is decreased by the insertion of these mutations. Nonetheless the pathological character remains neutral as analyzed by the results of PMut. Moreover, the normal and mutated pRB structures were superimposed by using PyMOL in order to highlight their differences. The predicted structure of mutant protein in comparison to the wild type structure showed considerable similarity with the differences of only the point mutations but it showed a decrease in the stability of the mutant protein. Thus our in silico analysis hypothesizes that although the reported mutations do not drastically affect the binding affinity of the protein with its potential ligands but they might have an indirect effect on the physiological characteristics of the protein. Loss of pRB function can lead to different cancers and as these mutations seem to be disrupting the normal RB1/E2F pathway. Presence of these variants in our patients with RB suggests that these mutations might ultimately prevent the mutant RB1 protein from functionally normally and can be responsible for the deregulation of cell growth control.

ACKNOWLEDGMENTS

The authors thank all RB affected families/patients for their participation and cooperation. We sincerely acknowledge Prof. Dr. Syed Ali Haider (Late), senior surgeon and ophthalmologist at Lahore General Hospital, for his assistance and valuable help throughout the progress of this research. Additionally we acknowledge all collaborating institutes for their valuable input.

Conflict of interest statement

Authors declare no conflict of interest.

REFERENCES

Burke, J.R., Hura, G.L. and Rubin, S.M., 2012. Structures of inactive retinoblastoma protein reveal multiple mechanisms for cell cycle control. *Genes Dev*., **26**: 1156-1166.

Capriottie, F.P. and Casadio, R., 2005. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucl. Acids Res*., **33**: 306-310.

Colovos, C. and Yeates, T.O., 1993. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci*., **9**: 1511-1519.

Dick, F.A., 2007. Structure-function analysis of the retinoblastoma tumor suppressor protein - is the whole a sum of its parts? *Cell Div*., **2:** 26.

Dyer, M.A. and Cepko, C.L., 2001. Regulating proliferation during retinal development. *Nat. Rev. Neurosci*., **2**: 333-342.

Dyer, M.A. and Bremner, R., 2005. The search for the retinoblastoma cell of origin. *Nat. Rev. Cancer*, **5**: 91-101.

Eisenberg, D., Luthy, R. and Bowie, J.U., 1997. VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol*., **277**: 396-404.

Ferrer-Costa, C., Gelpi, J.L., Zamakola, L., Parraga, I., Dela Cruz, X. and Orozco, M., 2005. PMUT: a web-based tool for the annotation of pathological mutations on proteins. *Bioinformatics*, **21**: 3176-3178.

Humphrey, W., Dalke, A. and Schulten, K., 1996. VMD: visual molecular dynamics. *J. Mol. Graph*., **14**: 33-38.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., Mcginnis, S. and Madden, T.L., 2008. NCBI BLAST: a better web interface. *Nucl. Acids Res.*, **36**: W5-9.

Kivela, T., 2009. The epidemiological challenge of the most frequent eye cancer: retinoblastoma, an issue of birth and death. *Br. J. Ophthalmol.*, **93**: 1129-1131.

Knudsen, E.S. and Wang, J.Y., 2010. Targeting the RB-pathway in cancer therapy. *Clin. Cancer Res*., **16**: 1094-1099.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., Mcgettigan, P.A., Mcwilliam, H., valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*., **23**: 2947-2948.

Laskowski, R.A., Macarthur, M.W., Moss, D.S. and Thornton, J.M.,1993. PROCHECK: a program to check the

stereochemical quality of protein structures. *J. appl. Cryst.*, **26**: 283-291.

Lohmann, D.R. and Gallie, B.L., 2013. *Retinoblastoma. Gene Reviews*. Available from: http://wwwncbinlmnihgov/books/NBK1452/#retinoblastomaMolecular_Genetics

Sage, J., 2012. The retinoblastoma tumor suppressor and stem cell biology. *Genes Dev*., **26**: 1409-1420.

Sali, A., Potterton, L., Yuan, F., Van Vlijmen, H. and Karplus, M., 1995. Evaluation of comparative protein modeling by MODELLER. *Proteins*, **23**: 318-326.

Sherr, C.J., 1996. Cancer cell cycles. *Science*, **274**: 1672-1677.

Weinberg, R.A., 1995. The retinoblastoma protein and cell cycle control. *Cell*, **81**: 323-330.

Weinberg, R.A., Rapaport, J.M. and Albert, D.M., 1986. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature*, **323**: 643.

Wiederstein, M. and Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucl. Acids Res*., **35**: W407-410.

Xiao, B., Spencer, J., Clements, A., Ali-Khan, N., Mittnacht, S., Broceno, C., Burghammer, M., Perrakis, A., Marmorstein, R. and Gamblin, S.J., 2003. Crystal structure of the retinoblastoma tumor suppressor protein bound to E2F and the molecular basis of its regulation. *Proc. natl. Acad. Sci. USA*, **100:** 2363-2368.