Protein Phosphatase 1D (PPM1D) Structure Prediction Using Homology Modeling

Santhosh Kumar Nagarajan and Thirumurthy Madhavan

Abstract

Protein phosphatase manganese dependent 1D (PPM1D) is one of the Ser/Thr protein phosphatases belongs to the PP2C family. They play an important role in cancer tumorigenesis of various tumors including neuroblastoma, pancreatic adenocarcinoma, medulloblastoma, breast cancer, prostate cancer and ovarian cancer. Even though PPM1D is involved in the pathophysiology of various tumors, the three dimensional protein structure is still unknown. Hence in the present study, homology modelling of PPM1D was performed. 20 different models were modelled using single- and multiple-template based homology modelling and validated using different techniques. Best models were selected based on the validation. Three models were selected and found to have similar structures. The predicted models may be useful as a tool in studying the pathophysiological role of PPM1D.

Keywords: PPM1D, WIP1, PP2C, Homology Modelling.

1. Introduction

Protein phosphatase, Mg(2+)/Mn(2+) dependent 1D (PPM1D) also known as WIP1 (wild-type p53-induced phosphatase 1) is an enzyme that belongs to the PP2C family of Ser/Thr protein phosphatases[1]. PP2C family members are known to be negative regulators of cell stress response pathways. This phosphatase mediates a feedback regulation of p38-p53 signaling which contributes to the inhibition of growth and the suppression of stress induced apoptosis. PPM1D plays an important role in cancer tumorigenesis[2]. PPM1D overexpression has been reported in various human tumors, including neuroblastoma[3], pancreatic adenocarcinoma[4], medulloblastoma[5], breast cancer[6] and ovarian cancer[7]. PPM1D phosphatase is also form an integral component of ATM-dependent signaling pathway[8]. In a recent study, the chance of identifying PPM1D as a novel biomarker for prostate cancer was studied[9]. Even though PPM1D is involved in the pathophysiology of various tumors, the three dimensional structure is still unknown.

Homology Modeling provides a rapid way of predicting the three-dimensional structure of a protein when only the sequence data of the protein is available. As it requires an enormous amount of time is required to prepare protein for crystallization, the number of protein structures resolved experimentally lags behind the sequence data available[10]. Experimental process such as protein expression, purification and finally crystallization, requires years to perform. Homology modeling can provide as a tool for the experimental procedures in finding the structure of the protein in a rather short time. In this study we have developed different homology models of PPM1D protein and validated them. This predicted model would be useful in studying the pathophysiological role of PPM1D in various cancers and to study the functioning of the protein.

2. Material and Methods

2.1. Template Selection

From the Uniprot database, the amino acid sequence of the human Protein phosphatase 1D (accession No: O15297) was retrieved. Protein BLAST[11] search was performed against PDB[12] to find suitable templates for homology modelling. Three templates were selected based on sequence identity, query coverage and E-value. The selected templates were – 4RAG, 1A6Q and 3FXJ. It has been statistically proven that if the level of
Table 1. The query coverage and identity values of the templates

| PDB ID | Max Score | Total score | Query Coverage | E Value | Identity % |
|--------|-----------|-------------|----------------|---------|------------|
| 4RAG   | 92.4      | 92.4        | 46%            | 6e-20   | 30%        |
| 1A6Q   | 92.0      | 92.0        | 46%            | 1e-19   | 30%        |
| 3FXJ   | 92.0      | 92.0        | 46%            | 1e-19   | 30%        |

Fig. 1. Alignment between the target (PPM1D) and templates (4RAG, 1A6Q and 3FXJ).
sequence identity is above 30%, then up to 90% of the polypeptide conformation tends to be modelled well\cite{10,13,15}. All three templates were having sequence identity of exactly 30%. As the query coverage and identity of the template was low (Table 1), to improve the model accuracy, the multiple template based homology modelling was also performed. Query coverage for all the templates was 46%.

2.2. Homology Modelling

Using EasyModeller 4.0\cite{16}, the three dimensional structures of PPM1D were modelled. EasyModeller 4.0 uses MODELLER 9.12\cite{17} and Python 2.7.1 in the back-end. The predicted models were assessed and validated using the RMSD values. Using RAMPAGE web server, Ramachandran plots for the models were plotted\cite{18}. A Ramachandran plot also known as \( \phi, \psi \) plot, provides a way to visualize backbone dihedral angles \( \psi \) against \( \phi \) of amino acid residues in protein structure. The sterically allowed regions for these angles can be identified using RC plot. Later, validation by Verify3D was carried out. Verify3D determines the compatibility of the predicted model with its own amino acid sequence by assigning a structural class based on its location and environment (alpha, beta, loop, polar, nonpolar etc.) and comparing the results to good structures\cite{19}.

3. Results and Discussion

3.1. Model Generation

Using EasyModeller, 5 models are modelled for each of the three templates -- 4RAG, 1A6Q and 3FXJ. Using the three different templates, single and multiple template based approaches were carried out. Five models were modelled for each template. For Multiple template based approach, using the CLUSTALW\cite{20} program, multiple sequence alignment was done to find conserved residues. The alignment of the templates was represented in Fig. 1. Totally 20 models were modelled.

| Model No | Templates Used | RMSD | Ramsachandran Plot | Verify3D |
|----------|----------------|------|---------------------|----------|
|          |                |      | Number of residues in favored region | Number of residues in allowed region | Number of residues in outlier region | (% of the residues had an averaged 3D-1D score >= 0.2) |
| 1        | 4RAG           | 0.295 | 86.6               | 7.6      | 5.8      | 51.40   |
| 2        | 4RAG           | 0.546 | 82.9               | 9.3      | 7.8      | 55.87   |
| 3        | 4RAG           | 0.332 | 85.4               | 9.1      | 5.5      | 51.07   |
| 4        | 4RAG           | 0.385 | 83.4               | 9.6      | 7.0      | 56.53   |
| 5        | 4RAG           | 0.248 | 85.1               | 9.3      | 5.6      | 59.17   |
| 6        | 1A6Q           | 0.245 | 85.2               | 9.8      | 5.0      | 49.75   |
| 7        | 1A6Q           | 0.460 | 82.3               | 11.1     | 6.6      | 59.34   |
| 8        | 1A6Q           | 0.401 | 81.6               | 11.4     | 7.0      | 45.12   |
| 9        | 1A6Q           | 0.414 | 84.9               | 8.1      | 7.0      | 57.02   |
| 10       | 1A6Q           | 0.474 | 84.1               | 11.1     | 4.8      | 48.76   |
| 11       | 3FXJ           | 0.335 | 84.7               | 9.6      | 5.6      | 52.56   |
| 12       | 3FXJ           | 0.362 | 86.6               | 8.3      | 5.1      | 46.94   |
| 13       | 3FXJ           | 0.368 | 85.6               | 8.8      | 5.6      | 49.26   |
| 14       | 3FXJ           | 0.273 | 85.9               | 8.0      | 6.1      | 51.24   |
| 15       | 3FXJ           | 0.269 | 84.6               | 9.8      | 5.6      | 53.88   |
| 16       | 4RAG, 1A6Q, 3FXJ | 0.297 | 89.2               | 6.1      | 4.6      | 53.88   |
| 17       | 4RAG, 1A6Q, 3FXJ | 0.301 | 87.1               | 9.5      | 3.5      | 52.40   |
| 18       | 4RAG, 1A6Q, 3FXJ | 0.224 | 89.6               | 7.3      | 3.2      | 53.06   |
| 19       | 4RAG, 1A6Q, 3FXJ | 0.207 | 87.1               | 8.1      | 4.8      | 51.24   |
| 20       | 4RAG, 1A6Q, 3FXJ | 0.344 | 89.1               | 7.3      | 3.6      | 53.88   |
3.2. Model Validation

Using different validation techniques, the predicted models were validated. Root mean square deviation of all the predicted models with their respective template was calculated. Based on the RMSD values, models 1, 6, 15 and 19 were found to be good models. Ramachandran plot was plotted for the models and the number of residues in favourable, allowed and disallowed region was identified. Verify3D was also performed for all the models. All the statistical values were tabulated in Table 2. Based on the statistical results, models 1, 6, 15 and 19 (one model from each template) were found to be the best models. All the models have similar structure and are found to be reliable based on the validation. The selected models are represented in Fig. 2. RC plot of the selected models were represented in Fig. 3.

![Selected models after model validation. Structures modelled using different templates (a) 4RAG (b) 1A6Q (c) 3FXJ (d) All three templates](image)

**Fig. 2.** Selected models after model validation. Structures modelled using different templates (a) 4RAG (b) 1A6Q (c) 3FXJ (d) All three templates
4. Conclusion

Three dimensional models for Protein phosphatase 1D (PPM1D) were generated using three different approaches. Model numbers 1, 6, 15 and 19 were selected as best, based on their RMS, RC plot, and Verify3D values. The four selected models show similar structures. Based on the model validation, we can also...

Fig. 3. RC plot for selected models modelled from templates (a) 4RAG (b) 1A6Q (c) 3FXJ (d) All three templates
say that all the generated models were reliable. These predicted models would be useful in the studying the pathophysiological role of PPM1D in future. Also, these models may serve as a reliable tool for analysing the structure and function of PPM1D.

References

[1] X. Lu, T. A. Nguyen, S. H. Moon, Y. Darlington, M. Sommer, and L. A. Donehower, “The type 2C phosphatase Wip1: an oncogenic regulator of tumor suppressor and DNA damage response pathways”, Cancer Metast. Rev., Vol. 27, pp. 123-135, 2008.
[2] H. Yang, X. Y. Gao, P. Li, and T. S. Jiang, “PPM1D overexpression predicts poor prognosis in non-small cell lung cancer”, Tumor Biol., Vol. 36, pp. 2179-2184, 2015.
[3] F. Saito-Ohara, I. Imoto, J. Inoue, H. Hosoi, A. Nakagawara, T. Sugimoto, J. Inazawa, “PPM1D is a potential target for 17q gain in neuroblastoma”, Cancer Res., Vol. 63, pp. 1876-1883, 2003.
[4] P. Loukopoulos, T. Shibata, H. Katoh, A. Kokubu, M. Sakamoto, K. Yamazaki, T. Kosuge, Y. Kanai, F. Hosoda, I. Imoto, M. Ohki, J. Inazawa, and S. Hirohashi, “Genome-wide array-based comparative genomic hybridization analysis of pancreatic adenocarcinoma: identification of genetic indicators that predict patient outcome”, Cancer Sci., vol. 98, pp. 392-400, 2007.
[5] R. C. Castellino, M. D. Bortoli, X. Lu, S. H. Moon, T. A. Nguyen, M. A. Shepard, P. H. Rao, L. A. Donehower, and J. Y. Kim, “Medulloblastomas overexpress the p53-inactivating oncogene WIP1/PPM1D”, J. Neuro-oncol., Vol. 86, pp. 245-256, 2008.
[6] M. B. Lambros, R. Natrajan, F. C. Geyer, M. A. Lopez-Garcia, K. J. Dedes, K. Savage, M. Lacroix-Triki, R. L. Jones, C. J. Lord, S. Linardopoulos, A. Ashworth, and J. S. Reis-Filho, “PPM1D gene amplification and overexpression in breast cancer: a qRT-PCR and chromogenic in situ hybridization study”, Modern Pathol., Vol. 23, pp. 1334-1345, 2010.
[7] J. Parssinen, E. L. Alaromo, S. Khan, R. Karhu, M. Vihinen, and A. Kallioniemi, “Identification of differentially expressed genes after PPM1D silencing in breast cancer”, Cancer Lett., Vol. 259, pp. 61-70, 2008.
[8] S. Shreeram, O. N. Demidov, W. K. Hee, H. Yamaguchi, N. Onishi, C. Kek, O. N. Timofeev, C. Dudgeon, A. J. Fornace, C. W. Anderson, Y. Minami, E. Appella, and D. V. Bulavin, “Wip1 phosphatase modulates ATM-dependent signaling pathways”, Mol. Cell, Vol. 23, pp. 757-764, 2006.
[9] L. Jiao, D. Shen, G. Liu, J. Jia, J. Geng, H. Wang, and Y. Sun, “PPM1D as a novel biomarker for prostate cancer after radical prostatectomy”, Anticancer Res., Vol. 34, No. 2919-2925, 2014.
[10] G. C. Baker, J. J. Smith, and D. A. Cowan, “Review and re-analysis of domain-specific 16S primers”, J. Microbiol. Meth., Vol. 55, pp. 541-555, 2003.
[11] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, “Basic local alignment search tool”, J. Mol. Biol., Vol. 215, pp. 403-410, 1990.
[12] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, “The protein data bank”, Nucleic Acids Res., Vol. 28, pp. 235-242, 2000.
[13] M. Shalini and T. Madhavan, “Homology modeling of CCR 4: Novel therapeutic target and preferential maker for Th2 cells”, J. Chosun Natural Sci., Vol. 7, pp. 234-240, 2014.
[14] B. Sathy and T. Madhavan, “Homology modeling of cysteinyl leukotriene1 receptor”, J. Chosun Natural Sci., Vol. 8, pp. 13-18, 2015.
[15] S. K. Nagarajan, and T. Madhavan, “3D structure prediction of thromboxane A2 receptor by homology modeling”, J. Chosun Natural Sci., Vol. 8, pp. 75-79, 2015.
[16] B. K. Kantal, P. Aparoy, and P. Reddanna, “EasyModeller: A graphical interface to MODELLER”, BMC Research Notes, Vol. 3, pp. 226, 2010.
[17] N. Eswar, M. A. Marti-Renom, B. Webb, M. S. Madhusudhan, D. Eramian, M. Shen, U. Pieper, and A. Sali, “Comparative protein structure modeling With MODELLER”, Current Protocols in Bioinformatics, New York: John Wiley & Sons, Inc., pp. 5.6.1-5.6.30, 2006.
[18] S. C. Lovell, I. W. Davis, W. B. Arendall III, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson, “Structure validation by Cz geometry: φ, ψ and Cβ deviation”, Proteins., Vol. 50, pp. 437-450, 2002.
[19] J. U. Bowie, R. Lüthy, D. Eisenberg, “A method to identify protein sequences that fold into a known three-dimensional structure”, Science, Vol. 253, pp. 164-170, 1991.
[20] J. D. Thompson, D. G. Higgins, and T. J. Gibson, “CLUSTAL W: improving the sensitivity of progressive sequence weighting, position-specific gap penalties and weight matrix choice”, Nucleic Acids Res., Vol. 22, pp. 4673-4680, 1994.