Prediction of Human Nucleoplasmin-3 Protein Structure and Active Sites

Waseem El-Huneidi and Mohammed Hamad

Department of Basic Sciences, College of Science and Health Professions, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia

Abstract: Nucleoplasmins are group of proteins involved in histone binding and play a crucial role in nucleosome assembly. There are four members of this family, NPM1, NPM2, NPM3 and NPM like proteins. NPM3 was the least investigated member of this family and it is the most recently discovered one, this study used in silico tools to predict the structure of Nucleoplasmin-3, a 3D model was predicted and the secondary structure was shown that 30.34% of amino acids presented in alpha helix, 23.60% of amino acids in extended turn, 42.13% in random coils and 3.93% of amino acids in beta turns. In addition, the predicted model was shown to have 31 active binding sites. In conclusion, the results of this study provided a 3D model for nucleoplasmin-3 with potential actives sites which can be used in screening for nucleoplasmin-3 regulators

Key words: In Silico • Nucleoplasmin-3 • I-TASSER • Castp • RAMPAGE

INTRODUCTION

Nucleoplasmins (NPM) are group of proteins that bind to histones and facilitate the assembly of nucleosomes. These proteins are classified into four groups according to their amino acid sequences: NPM1, NPM2, NPM3 and invertebrate NPM proteins [1, 2]. Human NPM1 is 294 amino acids encoded in NPM1 gene, while NPM2 and NPM3 consist of 214 and 178 amino acids respectively, those proteins are encoded in NPM gene family. The tertiary structure of NPM1 and NPM2 were elucidated by X-ray crystallography [3, 4].

Among NPM proteins, NPM1 was the most studies member of the NPM family mainly because it is found to be upregulated in the tumor cells, furthermore, several variants of NPM1 were associated with cancer [5]. The least amount of data is available for nucleoplasmin-3 which is the most recently discovered nucleoplasmin family member [6, 7]. NPM3 is mainly found in the nucleoli and to maintain this localization it needs an active rRNA transcription [8, 9]. It has been reported that NPM3 plays a rolein NPM1 regulation through the binding of NPM3 to NPM1, which will create a complex that incapacitate NPM1 to perform its activities in ribosomal biogenesis [9] In addition, NPM3 expression in mammalian oocytes has been correlated to parental chromatin decondensation [10].

As long as there is no adequate information related to the structure, active site, ligand binding sites of human nucleoplasmin-3. Consequently, this study was performed to deals with the identification of active site, ligand binding sites, in a predicted protein model of human nucleoplasmin-3.

MATERIALS AND METHODS

Data Set and Structure Prediction: The amino acid sequence of human nucleoplasmin-3 was obtained from NCBI database (www.ncbi.nlm.nih.gov/protein). The I-TASSER simulation was used to predict the 3D structure of nucleoplasmin-3, this program uses multiple threading alignment model. A correlation between the C-score and the TM-score are also included in the prediction methodology, with a correlation coefficient of 0.91. [11, 12].
Protein Structure Validation: The predicted model was validated using RAMPAGE which is a ramachandran diagram plots in which phi versus phi-dihedral was calculated for each residue of the human nucleoplasmin-3. The output was either amino acids are in favored, allowed or outlier regions relying on density dependent smoothing for non-glycine, non-Proline and non-preproline residues with B<30 for 500 high resolution protein structures [13].

Prediction of Protein Secondary Structure: SOPMA program (Self-Optimized Prediction Method) was used to predict the secondary structure of human nucleoplasmin-3, the prediction was based on determining the potential role of each amino acid.[http://npsapbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma. html] [14].

Prediction of Active Site: To predict the active sites of the protein, “CASTp” tool was used, in which the prediction is based on the comparison of different potential sites with reported sources [15].

RESULTS AND DISCUSSION

The protein structure of human nucleoplasmin-3 was predicted based on I-TASSER program which build a 3D models on the base of multiple threading alignment by Lometes and Illterative Tasser simulations [11]. The human nucleoplasmin-3 model (Figure 1) had a C-score of -2.39 and the estimated TM score is 0.43 ± 0.14 and RSMD(A) of 10.5 ± 4.6, number of decoys is1586 and the cluster density is 0.0265. Structure prediction by I-TASSER rely on template proteins with known structures obtained from databases. The prediction procedure is by matching the query sequence against a non-redundant sequence database.

The predicted nucleplasmin-3 model was validated using RAMPAGE ramachandran plot of non-glycine and non-proline residues. The predicted structure showed that 84.7% of the total amino acids were presented in most favored regions and the other 6.8 % and 8.5 % of amino acids were presented in allowed and outlier regions respectively.

The analysis of the secondary structure of the protein is used to predict if a certain amino acid is part of any types of secondary structures (Helices, coils, sheets...etc.). The prediction obtained from SOPMA showed that 30.34% of amino acids presented in alpha helix, 23.60% of amino acids in extended turn, 42.13% in random coils and 3.93% of amino acids in beta turns (Figure 2).

Nucleoplasmin model was predicted to have 31 binding pockets CASTp software with ideal parameters. All 31 pockets were characterized to find out its residues around probe radius of 1.4Å and among them, the largest active site has an area of 1018.3Å and volume of 1674.3Å. The green color (Figure 3;A and B) shows the largest active site position in the build protein which lies between amino acid 53 and 178.

![Fig. 1: The 3D model of human nucleoplasmin](image)

![Fig. 2: The secondary structure prediction obtained from SOPMA program (h is helix, e is extended strand, t is Beta turn and c is random coil)](image)
In this study, we proposed a valid and stable 3D model of human nucleoplasmin-3 whose structure is not experimentally elucidated yet. Further analysis provides information about secondary structure and its active sites. On the basis of the findings, it could be concluded that further characterization of nucleoplasmin-3 will be important as nucleosome related protein. This study can be used in broad screening on regulators of nucleoplasmin, also the data could be used to assist in drug design project and can be further applied in future studies.

REFERENCES


