Journal of Molecular Structure 1177 (2019) 177–185

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: http://www.elsevier.com/locate/molstruc

Structural analysis of misfolding equilibrative nucleoside transporter 3in H-syndrome

Nahid Askari [a, *], Elham Rezvannejad [a], Sodaif Darvish Moghadam [b], Sara Shafieipour [c]

[a] Research Department of Biotechnology, Institute of Sciences and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran
[b] Gastroenterology and Hepatology Research Center, Kerman University of Medical Sciences, Kerman, Iran
[c] Fellowship of Gastroenterology Taleghani Hospital, Gastroenterology, and Liver Disease Institute, Shahid Beheshti University, Tehran, Iran

ARTICLE INFO

Article history:
Received 28 May 2018
Received in revised form
12 September 2018
Accepted 18 September 2018
Available online 19 September 2018

Keywords:
Equilibrative nucleoside transporter 3
Misfolding
I-TASSER
PyMOL

ABSTRACT

Recent findings interpret that misfolding, aggregation and accumulation of proteins consequence in some disease. Evidence from different practices intensely support this hypothesis and demonstrate that a common therapy for these pernicious disorders might be possible. However, H syndrome is related to the recessive mutations in *SLC29A3*, encoding the equilibrative nucleoside transporter Equilibrative nucleoside transporter 3 (hENT3) which is expressed in mitochondria. The aim of this study is to investigate the misfolding and aggregation of ENT3 protein regarding in H syndrome, which can be a potential target for therapeutic interference in this disorder. Therefore, the 3D structure of the ENT3 protein was assumed using the I-TASSER online server and the best-predicted structure with the maximum confidence score (C-Score) was selected. The reliability and quality of the structure were evaluated by Z-score. However, the structural model of the selected mutant was accomplished using the I-TASSER server. The comparison with the corresponding of non-mutated structure and structural difference between mutant and wild-type structures was demonstrated by PyMOL.

In the present study, computational models combined with experimental approaches provide new insights into the ENT3 which could lead to expanding new classes of nucleobase analogs for treatment and quintessential therapeutic strategies for H syndrome and some other diseases like cancer and viral infection.

1. Introduction

Proteins are molecules that control the most vital cellular functions. Thus, a protein must first fold into its correct three-dimensional structure, in addition to assuming complex tertiary and even quaternary conformations. However, the process is quite complicated and susceptible to errors but many aspects of folding being distinct to the biophysical properties of the protein itself [1]. Proteins consist of the perplexing arrangement of interior folds that collapse into a final stable structure and, for many proteins, only a modest free energy gain (generally only −3 to −7 kcal/mol) [2] is associated with the relevant folding of a protein compared with its numerous potential misfolded states. Therefore, a number of misfolded proteins complicated in disease contain mutations that destabilize the correct fold and/or stabilize a misfolded state. Different chemical natures of protein compartments cause various protein-folding problems. For instance, in eukaryotic cells, protein folding must occur in many different organelles such as peroxisomes and mitochondria [3].

Retaining protein homeostasis is required for normal cell function and its viability.

Protein misfolded has occurred in many diseases of human beings and all other organisms. During diseases, the normally folded proteins, misfold and gain functions such as infectivity, toxicity and loses function. In designing towards conventional therapeutics it is necessary to know more on the miss functions of proteins based on their structures [4].

One of the most credible ways to characterize proteins with an unknown activity is experimental determination and a number of computational approaches have been developed for prediction of protein function [5]. Web portals have been used to contribute information about protein structures [6].

* Corresponding author.
E-mail address: n.askari@kgut.ac.ir (N. Askari).

https://doi.org/10.1016/j.molstruc.2018.09.051
0022-2860/

Previous studies demonstrated that mutations in hENT3 make a wide range of human genetic disorders. Vered Molho-Pessach et al. [34] reported five mutations in H syndrome. On the other hand [7], consider that mutations in *SLC29A3* perceive in Familial Histiocytosis Syndrome and Familial Rosai-Dorfman Disease. Therefore, it's assumed that hENT3 related disorders have some overlapping clinical symptoms.

Accumulating, the assessment of SLC29A3 mutations as the molecular basis in H syndrome, Familial Histiocytosis Syndrome and Familial Rosai-Dorfman certify a direct link between links these disorders to SLC29A3-associated phenotypes.

Mutations on membrane proteins may lead to small structural variations. Prediction of such structural variations can help to further understand the related bio-activities of membrane proteins [8]. It was investigated that, hENT3 is involved in protein sorting at the multivesicular body; Epsin-3 has a role in the trafficking of clathrin between the Golgi network and endosomes. It is involved in the recruitment of clathrin to the Golgi network and endosomes to form clathrin-coated vesicles. The beginning and the end of transmembrane helices play an important role in transferring through the membrane and intracellular activation pathways. hENT3 is involved in protein sorting at the multivesicular body because it binds to membranes and, in association with vacuolar protein sorting-associated protein 27.

H syndrome caused by mutations in the *SLC29A3* gene. That is an autosomal recessive genodermatosis disease with multisystem involvement. H syndrome is identified by hyperpigmentation, cutaneous changes of progressive sclerosis and hypertrichosis that follow a specific pattern with multiple systemic demonstrations. H syndrome as an anomaly discloses to the major clinical findings of hyperpigmentation, hypertrichosis, hepatosplenomegaly, heart anomalies, hearing loss, hypogonadism. Additionally, laboratory test results showed systemic manifestations, with characteristic cutaneous findings accompanying systemic inherited histiocytosis.

The human SLC29 family of proteins is recognized as equilibrative nucleoside transporters (ENTs). They belong to the eukaryotic ENT family of equilibrative and concentrative nucleoside and nucleobase transporters. ENTs are polytopic imperative membrane proteins which transport the nucleosides and some therapeutic analogs. Equilibrative nucleoside transporter 3 protein is one of ENTs members of equilibrative transporter which is arbitrated both efflux and influx of nucleosides across the membrane. Equilibrative nucleoside transporters act in the extracellular space by ecto-ATPases and nucleotidases uses [9].

Our aim is to understand the structure of ENT3 leading to H syndrome, relevant to the development of new therapeutic strategies in the future.

Our model seems to show well the differences between the wild-type and mutant form of the protein existing beside it consents with the laboratory results of patients. Based on our computational results and good agreement with all available information on this protein structure, we hope that experimentalists will find this problem challenging in H syndrome and will eventually confirm our findings.

2. Materials and methods

[10] illustrated that SLC29A3 had a heterozygous point mutation in exon 6, (at positions 53) which was a nucleotide transition c.1309G > A resulting in the missense amino acid substitution P.Glycine 437 Arginine. On the other hand, they revealed, there was a mutation in exon 3 (at positions 437) G155 > A mutation, changing Threonine to Alanine.

After detection the mutations in the eligible sequence, to study whether the mutations in the gene can either have an effect or modify the structure of the product of a gene, the amino acid sequences of the proteins were retrieved from National Centre for Biotechnology Information database (NCBI) and aligned using ClustalW software [11] to determine the appropriate sequence for protein structure prediction. By using the sequence similarity model, the structural homologs for retrieved sequences was documented from the available structures present in the protein data bank (PDB). Additionally, There was no protein structure (PDB) encoded by the SLC29A3 gene in PDB.

However, the structure of the protein was auspicated by fold recognition methodology using i-TASSER prediction server. The structures of the protein accomplished using I-TASSER servers were then approved by SAVes server [12].

I-TASSER simulations develop tens of thousands conformations that denominated decoys, for each target sequence. SPICKER program [13] is used to cluster all the decoys based on the pair-wise structure similarity, and indicate up to five models which correspond to the five largest structure clusters in I-TASSER. Based on the Monte Carlo theory, the largest clusters correspond to the states of the lowest free energy or largest partition function and thus have the highest confidence. In I-TASSER the confidence score for estimating the quality of the predicted model was quantitatively measured by C-score.

The C-score is computed based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. It is typically in the range of −5 to 2, where a C-score of higher value signifies a model with a high confidence and vice-versa.

The C-score of the I-TASSER models is described as

$$\mathrm{C-score} = \ln\left(\frac{M}{M_{tot}}\cdot\frac{1}{RMSD}\cdot\frac{\prod_{i=1}^{4} z(i)}{\prod_{i=1}^{4} z_{0(i)}}\right)$$

where M is the multiplicity of structures in the SPICKER cluster; M_{tot} is the total number of the I-TASSER structure decoys used in the clustering; RMSD is the average RMSD of the decoys to the cluster centroid; Z(i) is the highest Z-score (the energy to mean in the unit of standard deviation) of the templates by the ith PPA threading program and Z0(i) is a program-specified Z-score cutoff for distinguishing between good and bad templates [14].

Z(i)/Z0(i) is the normalized Z -score of the best template gene and I-TASSER adjusts a normalized B-factor with the Z-score-based transformation. C-score defined in Equation 1 is correlated with the quality of the assumed models (with a Pearson correlation coefficient >0.9 to the TM-score relative to the native) [15]. TM-score is a sequence length-independent metric for adjusting structure similarity with a value in the range (0, 1).

In the top 5 models which were provided by I-TASSER, the first model had a higher C-score and a better quality.

However, the C-score has a strong correlation with the quality of the final models, which has been used to quantitatively estimate the RMSD of the final models. Therefore, the C-score of the first model was used for further analysis.

A typical secondary structure prediction using I-TASSER accommodated three states: alpha helix (H), beta strand (S) and coil (C), with confidence scores for each residue. The predicted secondary structure was used for estimating the secondary structure of the protein.

Normalized B-factor with the *Z*-score-based transformation was computed using I-TASSER. The normalized B-factor is predicted by ResQ by the combination of template-based assignment and