Molecular modelling: a tool for understanding the cytoplasm–nucleus–nucleolus shuttling of the ceramide kinase-like protein

L Ayadi, I benrebeh, H Ayadi, S Masmoudi

Citation

L Ayadi, I benrebeh, H Ayadi, S Masmoudi. *Molecular modelling: a tool for understanding the cytoplasm–nucleus–nucleolus shuttling of the ceramide kinase-like protein*. The Internet Journal of Bioengineering. 2008 Volume 4 Number 2.

Abstract

Ceramide kinase like (CERKL) is a protein implicated in the pathology of retinitis pigmentosa. In this study, the 3D predicted model of CERKL was determined by molecular modelling techniques. Our results predicted the presence in the CERKL protein of an N-terminal Pleckstrin Homology (PH) domain. Presence of a nuclear localization signal within the PH domain may suggest the role played by this domain in the CERKL cytoplasm–nucleus shuttling. We also predicted a putative nucleolar localization signal (NoLS) sequence within the Diacylglycerol Kinase (DAGK) C-terminal domain. We suggest that this NoLS sequence located in an exposed helix retains the CERKL protein in the nucleolus. Absence of the predicted NoLS sequence in the R257X and R379X variants may explain their location in the nucleus and their absence from nucleoli. These findings have important implications for our current understanding of the cytoplasm–nucleus–nucleolus CERKL protein shuttling.

INTRODUCTION

Ceramide kinase like (CERKL) is a protein implicated in a form of autosomal recessive retinitis pigmentosa (RP), RP26 [1]. CERKL is a homolog of Ceramide kinase (CERK) sharing 50% of similarity. CERK is an enzyme that phosphorylates ceramide to form ceramide 1-phosphate (C1P) [2]. CERK and C1P have been implicated in the regulation of apoptosis [3], neurotransmitter secretion [4], and inflammation [5]. Surprisingly, CERKL did not phosphorylate ceramide and no target substrate is known for this enzyme [2]. Furthermore, wild type CERKL has been localized in different cell compartments [6]. The variant of CERKL missing 205 a.a. at the N-terminus (Δ1–205) was excluded from the nucleus and the nucleolus and was localized in the cytoplasm. Inversely, the variant of CERKL missing 154 a.a. at the C-terminus (R379X) was found in the nucleus but did not localize to nucleolus nor did a CERKL variant with a point mutation in the putative ATP binding site. The naturally occurring CERKL mutant (R257X), linked to the pathology of RP is also accumulated in the nucleus but was not associated with nucleoli.

In this study, we have derived a putative 3D model for the N-terminal and C-terminal of the CERKL protein based on molecular modelling techniques. Pleckstrin Homology (PH) domain and a putative nucleolar localization signal (NoLS) were predicted. Our results gave a rational explanation for the different puzzling localizations of CERKL in the cytoplasm-nucleus-nucleolus.

MATERIAL AND METHODS

The wild type of the CERKL protein was obtained from GenBank at the National Center for Biotechnology Information (accession number AJ640141). When the entire CERKL protein sequence was analyzed for comparison, no significantly matching sequence/structure was found, therefore the N-terminal and C-terminal domains of CERKL protein were modelled separately. In an attempt to predict the 3D model of the two domains, we used secondary prediction and fold recognition methods (PSIPRED [7], MGENTHREADER [8], 3D-PSSM [9], FUGUE [10] and 123D+[11]). Fold recognition method results were retained if E-value < 0.001 for the MGENTHREADER, E-value below 0.1 for 3D-PSSM (80% certainty), Z-score >= 6.0 for the FUGUE (99% confidence), and Z-score > 5 for 123D+ (high significant).

The predicted model of the N-terminal CERKL (amino acids 1 to 164) was built using the program MODELLER [12] based both on a template and sequence structure alignment

obtained from 3D-PSSM. Due to the lack of homology the 43 first amino acids were not considered.

The predicted model of the C-terminal CERKL (amino acids from 165 to 532) of the wild type was generated using the FUGUE alignment. A cycle of side chain fixing /energy minimization was repeated until no abnormality was detected in the models. The quality of the predicted models was examined using the PROCHECK program [13]. Structural analysis, visualization and figures generation was done using Swiss-PDB Viewer [14].

RESULTS

N-TERMINAL CERKL DOMAIN

Fold recognition methods were used in an attempt to identify structural homologs to N-terminal CERKL domain. Twenty hits were generated by 3D-PSSM, seven of them are PH domains. The PH domain DYN1_Human (PDB entry: 1dyn [15]) showed the best sequence identity with a value of 27% and E-value of approximately 0.08 which implies a 90% certainty of the N-terminal CERKL domain to be highly likely a PH domain. More PH domains were also found with the 123D+ server. Indeed this later, identified the PLC-delta-1 PH domain (PDB entry: 1mai [16]) as the second ranked fold with the best sequence identity of 20% and a Z score of 4.7. In addition, PSIPRED secondary structure prediction of the N-terminal CERKL domain matches with a PH domain structure Figure 1A.

The predicted model was built using the crystal structure of the PH domain from human dynamin (PDB entry: 1dyn). Analysis of the Ramachandran plot of the N-terminal PH domain showed that 79.8% (85.6 % for the 1dyn) of the residues lie in the most favorable regions, 11.9% in the additional allowed regions, 6.4% generously allowed regions and 1.8% in the disallowed region. Visualization and analysis of the model using Swiss-PDB Viewer showed that the predicted model correlated with the structural features of PH domains super fold: a seven stranded β-sheet with a flanking C-terminal α-helix (Figure 1B).

C-TERMINAL CERKL DOMAIN

Fold-recognition methods MGENTHREADER and FUGUE identified the YegS (PDB entry 2bon [17]) as the best template for the sequence of the C-terminal domain with an E value of 7e-05 (MGENTHREADER) and Z score of 27.8 (FUGUE) respectively. YegS is a cytosolic protein in Escherichia coli annotated as a putative diacylglycerol kinase (DGK) because it shares sequence similarity to the catalytic domain of the first DGK [18]. Due to the remote homology with the template, several insertions were found in the predicted model of the C-terminal domain. Nevertheless, PSIPRED secondary structure prediction of the C-terminal domain of the CERKL protein correlated well with the secondary structure of the YegS with the exception of fragments Gly327-Ala332 and Val350-Lys358 Figure 2A. So, we believed that the presence of insertions does not compromise the validity of the model, since the secondary structure elements seem to be well modelled. None of these insertions were considered in the predicted model. The predicted C-terminal CERKL domain was then analyzed using the PROCHECK program. The Ramachandran plot shows that 87.4 % of the residues lie in the most favorable regions, 9.8% in the additional allowed regions, 2.4 % in generously allowed regions and 0.4 % in the disallowed regions.

Structural analysis of the predicted model showed that the C-terminal domain is further divided into two lobes (two sub domains) Figure 2B. The N-terminal sub domain is formed by residues (Arg165 to Val304) and (Leu518 to Lys532) and comprises five helices (α2-α6) and six strands (β8-β12 and β25). This sub domain resembles the Rossman fold nucleotide binding domain found in YegS, in NAD kinase [19] and PFK family [20] members of the lipid kinase family. Two insertions (Leu193-Leu196) and (Ala277-Ala290) were found within the helix α2 and in the loop between β4 and β5 respectively. These solvent exposed insertions were not modelled since their absence from the predicted model did not compromise the integrity of the Rossman fold.

The beta-sandwich fold of C–terminal sub domain (Asp331-Arg543) is formed by twelve beta strands (β13-β24) and two helices (α7 and α8) Figure 2B. This sub domain was found in the YegS protein and in NAD kinases [19]. The beta-sandwich sub domain comprises six insertions: (Lys314-Leu316), (Ala332-Met340), (Phe369-Trp392), (Cys409-Cys411), (Tyr449-Lys453), and (Pro482-V499). All these insertions are located at the surface of the protein which did not compromise the folding of the sub domain.

The conserved GGDG motif present in the YegS as well as in NAD kinase and sphingosine kinase is also present in the CERKL protein (231-GGDGS-235) and superimposes well in the structure.

We identified the KHLK motif (residues from 444 to 447) in

the wild type CERKL protein. Such motif, R/K(R/K)X(R/K) (where the R or K are occasionally replaced by H) was identified as the NoLS in a number of nucleolar proteins [21], [22], [23], [24], and [25]. Furthermore in the CERKL protein, this motif is located in a solvent exposed helix (α8) Figure 2B.

Figure 1

Figure 1: (A) 3DPSSM sequence alignment of N-terminal CERKL domain with pleckstrin homology (PH) domain from human dynamin (PDB entry: 1dyn) and the correspondence of observed and predicted secondary structure elements. The highlighted residues in the CERKL protein represent the PSIPRED secondary structure predictions (red, alpha helix; blue, beta strand). The same color convention was adopted to represent the location of secondary structure elements in the 1dyn structure. (B) Predicted 3D model of the N-terminal PH domain of CERKL protein with helices in red, strands in blue and loops in grey (the 43 first residues in the alignment (A) were not considered). NLS2 localization is highlighted in pink; tryptophan PH domain signature is highlighted in cyn. The two termini are marked with capital letters

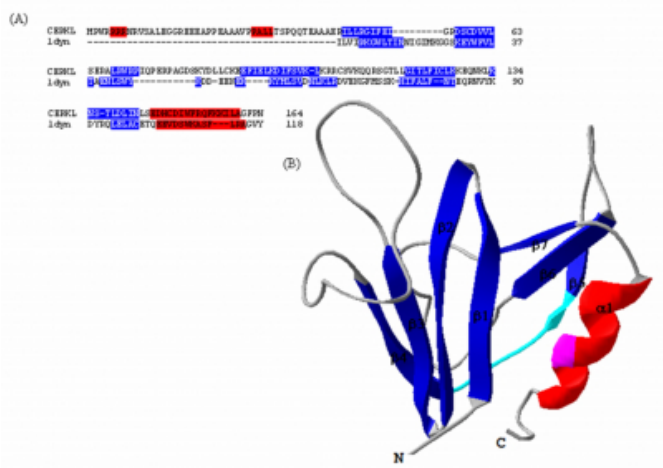
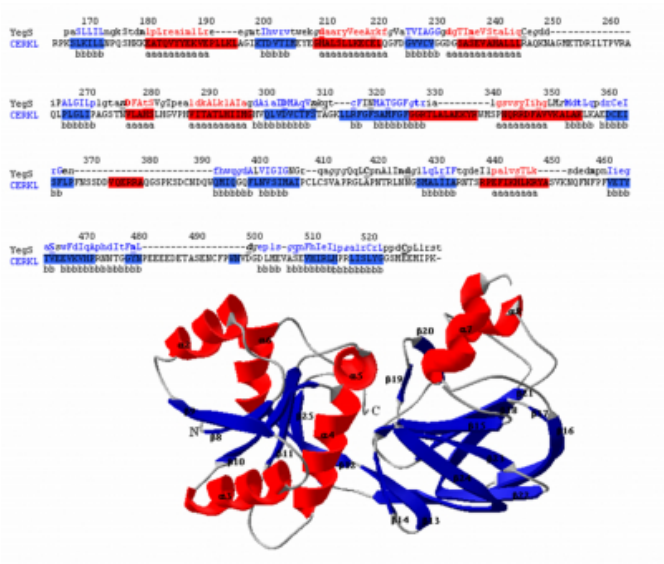


Figure 2

Figure 2: (A) FUGUE sequence alignment of C-terminal domain of CERKL protein with the YegS protein (pdb entry: 2bon). Lowercase, solvent accessible; uppercase, solvent inaccessible; bold, hydrogen bond to main-chain amide; underline, hydrogen bond to main-chain carbonyl; and italic, positive torsion angle. The proposed secondary structure elements in the predicted model are specified (a, ?-helix; b, ?-strand). The highlighted residues in the CERKL protein represent the PSIPRED secondary structure predictions (red, alpha helix; blue, beta strand). (B) Predicted model of the C-terminal domain of CERKL protein with helices in red, strands in blue and loops in grey. The two termini are marked with capital letters.



DISCUSSION

CERKL gene encodes a ceramide kinase that is assumed to be involved in sphingolipid-mediated apoptosis in the retina. In absence of structure solved by x-ray crystallography, our results predicted that the N-terminal CERKL protein is a Pleckstrin Homology (PH) domain. Although the invariant tryptophan PH domain signature residue was identified (Trp151) by alignment with the N-terminal PH domain of the CERK protein [6], others groups failed to detect the N-terminal PH domain of CERKL protein [1] and [6]. PH domains share only limited sequence similarity making them relatively difficult to detect by the conventional methods.

Two putative nuclear localization signal sequences termed NLS1 (1-MPWRRRRNRVSA-12) and NLS2 (100-SVKLKRRCSVKQ-111) were identified in the N-terminal domain of CERKL protein by Inagaki group [2]. They also showed that a mutation in NLS1 did to not alter the subcellular localization of CERKL, while a mutation in NLS2 disrupted its nuclear import. NLS1 did not appear in

our predicted model since the N-terminal 43 residues were not modeled due to the lack of homology. However, NLS2 is located within the PH domain (in the strand 5 and in the loop downstream this latter) Figure 1B.

Presence of a nuclear localization within the PH domain was recently reported in the PIKE PH domain [26]. The cytoplasmic–nuclear shuttling of PIKE was shown to be regulated by the balancing actions of the lipid-binding property of both the split PH domain and the nuclear targeting function of its nuclear localization sequence. We suggest that both NLS2 and PH domain in the CERKL protein are implicated in the regulation of the travelling trough nuclear membrane since they are present in the Wild type, in the C-terminal splice variant R379X and in the naturally occurring mutant R257X. Our suggestions were rationalised by the fact that the N-terminal splice variant (1-205) which lacks the regulatory PH domain (NLS2) was found in the cytoplasm excluded from the nucleus.

Structural analysis of the predicted model showed that the C-terminal domain is further divided into two sub domains: An N-terminal sub domain which resembles the Rossman fold nucleotide binding domain and a twelve stranded beta-sandwich fold which was found in the YegS protein and in NAD kinases. The conserved GGDG motif which has been implicated in binding to phosphate of ATP in 6-phosphofructokinase [27] and involved in nucleotide-binding in sphingosine kinase [28] is present in the CERKL protein. Mutation of the third glycine (G234D) in this motif was found to prevent the localisation of the mutant to nuclei in CERKL protein [6]. A functional ATP-binding site may be a requirement for nuclear import.

The R/K(R/K)X(R/K) motif identified as the NoLS in a number of nucleolar proteins is present in the wild type CERKL protein. This motif is located in a solvent exposed helix (8) allowing a direct interaction with other molecules which may direct or retain the CERKL protein in the nucleolus. This motif is absent from the R379X variant because the 154 missing amino acid residues are located in 17, 18, 19, 20, 21, 22, 23, 25 and more importantly in helix 8 containing the NoLS. Absence of nucleolar localization from the R379X or R257X CERKL mutants may explain their exclusion from the nucleolus and their accumulation in the nucleus.

In conclusion, molecular modelling was used to predict nuclear/nucleolar localizations in the CERKL protein implicated in the regulation of its subcellular localisation. Further studies should be carried out to substantiate our predictions.

References

1. M. Tuson, G. Marfany and R. Gonzalez-Duarte, Mutation of CERKL, a novel human ceramide kinase gene, causes autosomal recessive retinitis pigmentosa (RP26), Am. J. Hum. Genet. 74 (2004), pp. 128–138.
2. Y. Inagaki, S. Mitsutake, Y. Igarashi, Identification of a nuclear localization signal in the retinitis pigmentosa-mutated RP26 protein, ceramide kinase-like protein, Biochem. Biophys. Res. Commun. 343 (2006), pp. 982-987.
3. A. Gómez-Muñoz, J.Y. Kong, B. Salh and U.P. Steinbrecher, Ceramide-1-phosphate blocks apoptosis through inhibition of acid sphingomyelinase in macrophages, J. Lipid Res. 45 (2004), pp. 99–105
4. S.M. Bajjalieh, T.F. Martin and E. Floor, Synaptic vesicle ceramide kinase. A calcium-stimulated lipid kinase that co-purifies with brain synaptic vesicles, J. Biol. Chem. 264 (1989), pp. 14354–14360.
5. B.J. Pettus, A. Bielawska, S. Spiegel, P. Roddy, Y.A. Hannun and C.E. Chalfant, Ceramide kinase mediates cytokine- and calcium ionophore-induced arachidonic acid release, J. Biol. Chem. 278 (2003), pp. 38206–38213.
6. F. Bornancin, D. Mechtcheriakova, S. Stora, C. Graf, A. Wlachos, P. Devay, N. Urtz, T. Baumruker and A. Billich, Characterization of a ceramide kinase-like protein, Biochim. Biophys. Acta. 1687 (2005), pp. 31–43.
7. D.T. Jones, K. Bryson, A. Coleman et al., Prediction of novel and analogous folds using fragment assembly and fold recognition, Proteins 61 (2005), 143-151
8. D.T. Jones, GenTHREADER, an efficient and reliable protein fold recognition method for genomic sequences, J. Mol. Biol. 287 (1999), 797-815
9. L.A. Kelley, R.M. MacCallum and M.J. Sternberg, Enhanced genome annotation using structural profiles in the program 3D-PSSM, J. Mol. Biol. 299 (2000), pp. 499–520.
10. J. Shi, T.L. Blundell and K. Mizuguchi, FUGUE: sequence–structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J. Mol. Biol. 310 (2001), pp. 243–257
11. N.N. Alexandrov, R. Nussinov, R.M. Zimmer, Fast protein fold recognition via sequence to structure alignment and contact capacity potentials. Pac Symp Biocomput. (1996),pp.53–72.
12. A. Sali, T. L. Blundell, Comparative protein modeling by satisfaction of spatial restraints. J. Mol. Biol. 234 (1993), 779-815.
13. R.A. Laskowski, M.W. MacArthur, D.S. Moss and J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Cryst. 26 (1993), 283-291.
14. Guex, N., and M. C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis. 18 (1997) 2714-2723.
15. K.M. Ferguson, M.A. Lemmon, J. Schlessinger, P.B. Sigler, Crystal structure at 2.2 A resolution of the pleckstrin homology domain from human dynamin.
16. 199-209.
17. K.M.Ferguson, M.A. Lemmon, J. Schlessinger, P.B. Sigler, Structure of the high affinity complex of inositol trisphosphate with a phospholipase C pleckstrin homology domain. Cell 83 (1995), pp. 1037-1046.

18. H.M.Bakali, M.D. Herman, K.A. Johnson, A.A Kelly, A. Wieslander, B.M. Hallberg, P. Nordlund, Crystal Structure of Yegs, a Homologue to the Mammalian Diacylglycerol Kinases, Reveals a Novel Regulatory Metal Binding Site. J.Biol.Chem.
282 (2007) pp. 19644-19652.
19. F. Sakane, K. Yamada, H. Kanoh, C. Yokoyama, T. Tanabe, Porcine diacylglycerol kinase sequence has zinc finger and E–F hand motifs, Nature 344 (1990), pp. 345–348.
20. J.Liu, Y. Lou, H. Yokota, P.D. Adams, R. Kim, S.H. Kim, Crystal Structures of an NAD Kinase from Archaeoglobus fulgidus in Complex with ATP, NAD, or NADP, Journal of Molecular Biology 2005, 354 (2005), pp. 289-303.
21. Y.Shirakihara,P.R. Evans. Crystal structure of the complex of phosphofructokinase from Escherichia coli with its reaction products, J Mol Biol. 204(1988), pp. 973–994.
22. C. Schmidt, E. Lipsius,J. Cruppa, Nuclear and nucleolar targeting of human ribosomal protein S6. Mol Biol Cell 6(1995), pp. 1875–1885.
23. W.M. Michael, G. Dreyfuss, Distinct domains in ribosomal protein L5 mediate 5 S rRNA binding and nucleolar localization. J Biol Chem 271(1996), pp. 11571–11574.
24. S. Horke,K. Reumann, M. Schweizer,H. Will,T. Heise, Nuclear trafficking of La protein depends on a newly identified nucleolar localization signal and the ability to bind RNA. J Biol Chem 279 (2004), pp. 26563–26570
25. T. Annilo, A. Karis, S. Hoth, T. Rikk, J. Kruppa,A. Metspalu, Nuclear import and nucleolar accumulation of the human ribosomal protein S7 depends on both a minimal nuclear localization sequence and an adjacent basic region. Biochem Biophys Res Commun, 249 (1998) pp. 759–766.
26. Z. Hao, X. Li, T. Qiao, R. Du, G. Zhang, D. Fan, Subcellular Localization of CIAPIN1, Journal of Histochemistry and Cytochemistry , 54 (2006) pp. 1437-1444.
27. J. Yan, W. Wen, L.N. Chan, M. Zhang, Split Pleckstrin Homology Doma in-Mediated Cytoplasmic–Nuclear Localization of PI3-Kinase Enhancer GTPase, J. Mol. Biol. 378 (2008), pp. 425–435.
28. Y. Shirakihara and P.R. Evans, Crystal structure of the complex of phosphofructokinase from Escherichia coli with its reaction products, J. Mol. Biol. 204 (1988), pp. 973–994.
29. S.M. Pitson, P.A. Moretti, J.R. Zebol, R. Zareie, C.K. Derian, A.L. Darrow, J. Qi, R.J. D'Andrea, C.J. Bagley, M.A. Vadas and B.W. Wattenberg, The nucleotide-binding site of human sphingosine kinase 1, J. Biol. Chem. 277 (2002), pp. 49545–49553.

Author Information

Leila Ayadi
Unité Cibles pour le Diagnostic et la Thérapie, Centre de Biotechnologie de Sfax, Tunisie.

Imen benrebeh
Unité Cibles pour le Diagnostic et la Thérapie, Centre de Biotechnologie de Sfax, Tunisie.

Hammadi Ayadi
Unité Cibles pour le Diagnostic et la Thérapie, Centre de Biotechnologie de Sfax, Tunisie.

Saber Masmoudi
Unité Cibles pour le Diagnostic et la Thérapie, Centre de Biotechnologie de Sfax, Tunisie.