doi 10.18699/bgrs2024-5.2-22

Phylogenetic analysis and molecular dynamic simulations of CsqR, a transcription factor from *Escherichia coli*

Rybina A.A.[1]*, Glushak R.A.[2], Bessonova T.A.[3, 5], Dakhnovets A.I.[1], Rudenko A.Y.[4], Ozhiganov R.M.[4], Kaznadzey A.D.[5], Tutukina M.N.[1, 3, 5], Gelfand M.S.[1]

[1] *Skolkovo Institute of Science and Technology, Moscow, Russia*

[2] *Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia*

[3] *Institute of Cell Biophysics, Federal Research Center "Pushchino Scientific Center for Biological Research RAS", Pushchino, Russia*

[4] *Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russia*

[5] *Institute for Information Transmission Problems, RAS, Moscow, Russia*

* *rybinaann@gmail.com*

Key words: phylogeny; sulfoquinovose; molecular docking; molecular dynamic simulations

Motivation and Aim: CsqR is a local transcription factor from *Escherichia coli* that regulates expression of *yih* genes responsible for the degradation of sulfosugar sulfoquinovose (SQ) [1, 2]. We recently demonstrated that lactose could potentially regulate the *yih* genes [2]. SQ and its derivatives sulfoquinovosyl glycerol (SQG) and sulforhamnose (SR) were reported to act as putative effectors of CsqR [3]. Preliminary experimental findings from our laboratory suggested the presence of an alternative truncated variant of CsqR. Here, we aimed to study the evolution of CsqR and estimate patterns of its binding to potential effectors SQ, SR, SQG, and lactose.

Methods and Algorithms: Phylogenetic tree of CsqR protein homologs was constructed to study evolutionary patterns of CsqR. Production of the CsqR protein forms was validated in the western-blot experiment. To model interactions of CsqR with candidate effectors, predicted protein structures of CsqR were optimized in 2 μs molecular dynamic simulations and then subjected to blind molecular docking with potential ligands SQ, SR, SQG, and lactose.

Results: Phylogenetic analysis revealed the presence of CsqR homologs in both the Actinobacteria and Proteobacteria phyla. The structure of the tree suggested that *csqR* underwent duplication at some point. Notably, CsqR homologs of Enterobacteriales species had a highly conserved Met25 residue. We suggested the existence of two alternative variants of CsqR: a full-length version (CsqR-l) and a shorter variant lacking 24 N-terminal residues (CsqR-s). Western blot analysis showed that CsqR was produced in two protein forms, 28.5 kDa (CsqR-l) and 26 kDa (CsqR-s), the latter starting translation at Met25. Interestingly, CsqR-s exhibited significant activation during growth with sulfoquinovose as the sole carbon source, displacing CsqR-l during the stationary phase of growth on a rich medium. According to molecular dynamic simulations, CsqR-s might have two potential structural arrangements, with the interdomain linker appearing either as a disordered loop or an α-helix. This helix allowed the N-terminal domain to rotate in a hinge-like motion, resulting in the transition of CsqR-s between two conformations referred to as "open" and "compact". Molecular docking suggests that CsqR-s may have a greater ability to discriminate between putative ligands and other compounds compared to CsqR-l, indicating potential differences in ligand binding affinity or specificity between the two variants

Conclusion: Bacteria rarely possess multiple forms of a single protein, with only only a handful of cases reported to date. CsqR might be an interesting inclusion in this group, poised for additional experimental investigation.
Funding: The study is supported by the RSF via grant 24-14-00276.

References

1. Denger K. et al. Sulphoglycolysis in *Escherichia coli* K-12 closes a gap in the biogeochemical sulphur cycle. *Nature.* 2014;507(7490):114-117. doi 10.1038/nature12947
2. Kaznadzey A. et al. The genes of the sulphoquinovose catabolism in *Escherichia coli* are also associated with a previously unknown pathway of lactose degradation. *Sci Rep*. 2018;8(1):3177. doi 10.1038/s41598-018-21534-3
3. Shimada T. et al. Regulatory role of CsqR (YihW) in transcription of the genes for catabolism of the anionic sugar sulfoquinovose (SQ) in *Escherichia coli* K-12. *Microbiology.* 2019;165:78-89. doi 10.1099/mic.0.000740x