Journal of Molecular Graphics and Modelling 29 (2010) 573–580

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Computational approach to ensure the stability of the favorable ATP binding site in *E. coli* Hfq

Prettina Lazar, Songmi Kim, Yuno Lee, Keun Woo Lee*

Division of Applied Life Science (BK21 Program), Environmental Biotechnology National Core Research Center (EB-NCRC), Plant Molecular Biology and Biotechnology Research Center (PMBBRC), Gyeongsang National University (GNU), 900 Gazwa-dong, Jinju 660-701, Republic of Korea

ARTICLE INFO

Article history:
Received 10 September 2010
Received in revised form 1 November 2010
Accepted 5 November 2010
Available online 11 November 2010

Keywords:
Host factor protein-Hfq
ATP
Oligoribonucleotide
RNA
Post-transcriptional regulation
Molecular dynamics simulation
Mutation
Aromatic stacking
Destabilization

ABSTRACT

Bacterial Hfq is a highly conserved thermostable protein of about 10 kDa. The Hfq protein was discovered in 1968 as an *E. coli* host factor that was essential for replication of the bacteriophage Qβ. It is now clear that Hfq has many important physiological roles. In *E. coli*, Hfq mutants show a multiple stress response related phenotypes. Hfq is now known to regulate the translation of two major stress transcription factors RpoS and RpoE in Enterobacteria and mediates its plieotrophic effects through several mechanisms. It interacts with regulatory sRNA and facilitates their antisense interaction with their targets. It also acts independently to modulate mRNA decay and in addition acts as a repressor of mRNA translation. Recent paper from Arluison et al. [9] provided the first evidence indicating that Hfq is an ATP-binding protein. They determined a plausible ATP-binding site in Hfq and tested Hfq's ATP-binding affinity and stoichiometry. Experimental data suggest that the ATP-binding by the Hfq–RNA complex results in its significant destabilization of the protein and the result also proves important role of Tyr25 that flanks the cleft and stabilizes the adenine portion of ATP, possibly via aromatic stacking. In our study, the ATP molecule was docked into the predicted binding cleft using GOLD docking software. The binding nature of ATP and its effect on Hfq–RNA complex was studied using molecular dynamics simulations. Importance of Tyr25 residue was monitored and revealed using mutational study on the modeled systems. Our data and the corresponding results point to one of Hfq functional structural consequences due to ATP binding and Tyr25Ala mutation.

1. Introduction

Hfq is an abundant RNA-binding protein in *E. coli* that is highly conserved and appears to function as a global regulator of gene expression. Intriguingly, Hfq has been reported to have ATPase activity, albeit it is relatively weak [1]. Sequence comparison between Hfq and the heat shock protein ClpB revealed some homology between the ATP binding site of ClpB and the stretch of the Hfq sequence that includes β2. This 'modified Walker A box' is rather open, and how and where ATP binds and the role of the ATPase activity are under investigation. It should be noted that the chaperone activities of Hfq are independent of ATP hydrolysis, thereby leaving the function of the ATPase activity of Hfq an open question [2]. Finally, Hfq was also recently reported to interact with both ribosomal proteins S1 and an RNA polymerase, to exhibit ATPase activity and to affect the polyadenylation of bacterial RNAs. Alternatively, it is also possible that stimulation of PAP I-dependent poly(A) synthesis is correlated with the ATPase activity of Hfq [3,4]. Hfq is known to associate with the ribosomal protein S1 to promote replication of the RNA phage Qβ [5]. Hfq and S1 have also been found to associate with RNA polymerase, and an ATPase activity was found for Hfq, although no catalytic site is obvious from the sequence or crystal structures. If Hfq either has an ATPase activity or associates with an ATPase, such an ATPase might help remodel RNAs as they anneal and could provide external energy for reactions such as strand displacement [6]. Biochemical and genetic evidence suggests that the ATPase activity is an intrinsic activity of Hfq rather than that of an Hfq-associated protein. Hfq purified to apparent homogeneity shows ATPase activity. This finding suggests that at least some representatives of the extended family of eukaryotic Sm-like proteins (required for processes as diverse as pre-mRNA splicing, mRNA degradation and telomere formation) that share sequence homology with Hfq are also ATPases. There is limited sequence homology between Hfq and known ATP-utilizing enzymes, and in fact, Hfq belongs to new class of ATPases named AAA^+, which is a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes [7,8]. The hexameric structure of Hfq is quite similar to those of

* Corresponding author. Tel.: +82 55 751 6276; fax: +82 55 752 7062.
E-mail address: kwlee@gnu.ac.kr (K.W. Lee).

1093-3263/$ – see front matter © 2010 Published by Elsevier Inc.
doi:10.1016/j.jmgm.2010.11.003

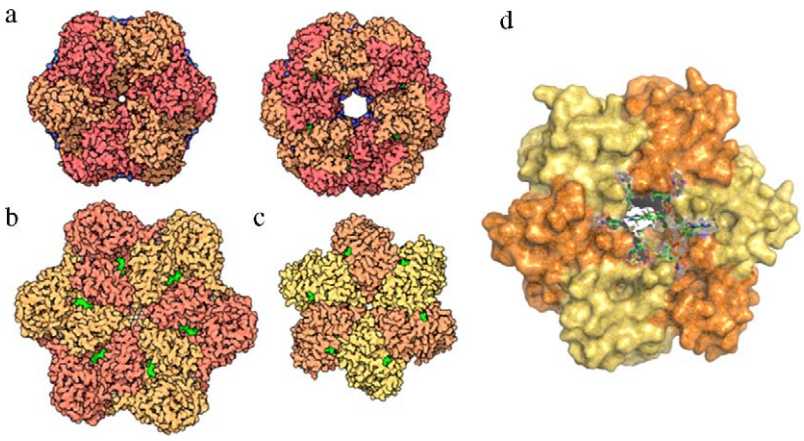

Fig. 1. Structural comparison of Hfq with three known examples of AAA+ ATPases (a) HslUV, (b) p97, (c) T antigen and (d) Hfq.

known AAA+ ATPases like HslUV, p97 and T Antigen, and hence the probable ATP binding site in Hfq must be similar to that of the known AAA+ ATPases (Fig. 1).

Using a combination of biochemical and genetic techniques, a plausible ATP-binding site was identified in Hfq and Hfq's ATP-binding affinity and stoichiometry were tested. The results of RNA footprinting and binding analyses suggest that ATP binding by the Hfq–RNA complex results in its significant destabilization. RNA footprinting indicates deprotection of Hfq-bound RNA tracts in the presence of ATP, suggestive of their release by the protein. Moreover, author mentioned that a Y25A mutation resulted in a near-knockout of ATP hydrolysis in the purified protein [9].

Computational approaches like molecular docking simulations, molecular modeling and molecular dynamics simulations were used to look for the most stable and favorable ATP binding site and any significant effects of ATP on the structure or stability of the Hfq–RNA complex. The importance of Tyr25 was also assessed by hexamutant analysis of Hfq. An attempt was made to substantiate the microscopic and atomic details of Arluison's experimental results with our study. These types of studies will address where it is that ATP comfortably binds on Hfq and characterize the influence of Hfq's ATPase activity on the six-fold symmetry and conformation of the Hfq–RNA complex.

2. Materials and methods

2.1. Molecular docking simulation

Binding interactions can be ascertained by docking inhibitors into the active site of a protein. The GOLD 3.01 [10] program was used to find a stable ATP binding mode along the boundary between individual Hfq subunits. It employs genetic algorithm in which information about the conformation of the ligand and hydrogen bonding is encoded into a chromosome. GOLD considers complete ligand flexibility and partial protein flexibility, and the energy functions are partly based on conformational and non-bonded interactions. Several types of scoring functions such as GoldScore, ChemScore and User-defined score are available. The following default genetic algorithm parameters were used: population size, 30; 1.1 for selection, number of islands, 5; number of genetic operations, 100,000; and niche size, 2. A pseudo-atom was created at the center of the interface gap between two adjacent monomers (chain A and chain B) of *E. coli* Hfq, and the active site was defined as 10 Å around it. The GoldScore was adopted to rank the docked conformations of a single ATP molecule between the two subunits of *E. coli* Hfq.

2.2. System set-up using molecular modeling

Using the selected conformation of ATP docked with two Hfq subunits, an Hfq hexamer with six individual ATP molecules docked along the boundaries of pairs of Hfq subunits was modeled and prepared by superimposition. Similarly, a structure of *E. coli* Hfq in the RNA-binding conformation (PDB ID: 1HK9) [11] was adopted from the only available crystal structure of *S. aureus* Hfq (PDB ID: 1KQ2) [12] with AU_5G. In *E. coli*, the oligo-ribonucleotide was replaced with rA7 instead of AU_5G, because *E. coli* Hfq selectively binds to rA7. Four different conformations of wild-type *E. coli* Hfq systems were modeled, including: apo form Hfq, Hfq with bound RNA (rA7), Hfq with six bound ATP molecules, and Hfq bound to both ATP (6 molecules) and RNA (rA7). A hexamutant Hfq with the mutation Tyr25Ala was also modeled to study the importance of Tyr25. This hexamutant was modeled in the same four configurations mentioned above: apo form; with bound RNA (rA7); with bound ATP (6 molecules); and with RNA (rA7) + ATP (6 molecules). All the molecular modeling studies were carried out using Discovery Studio version 2.5 [13].

2.3. Computational details of the molecular dynamics simulations

The GROMACS package [14,15] was used to perform MD simulations, where the protein and water molecules were described by parameters from AMBER99 [16] and TIP3P [17] force fields, respectively. Hydrogen atoms were added and the protonation state of ionizable groups was chosen appropriately at pH 7.0. A cubic box of solvent 12 nm in length was generated to perform the simulations in an aqueous environment. Na^+ or Cl^- counter-ions were added by replacing water molecules to ensure the overall charge neutrality of the simulated system. The particle mesh Ewald (PME) method was applied to accurately determine the long-range electrostatic interactions [18]. A grid spacing of 1.2 Å was used for fast Fourier transform calculations and van der Waals interactions were considered by applying a cutoff of 9 Å. Aconstant temperature and pressure (300 K and 1 bar) was maintained with a Berendsen thermostat [19] and a Parrinello–Rahman [20] barostat. The systems were subjected to the steepest descent energy minimization process with a tolerance of 1000 kJ/mol. The time step for the simulations was set to 2 fs. During the system equilibration process, the protein backbone was frozen and the solvent molecules with counterions were allowed to move for 100 ps under NPT conditions at 300 K. The equilibrated structure was then used for the following 5-ns production runs. Bonds between heavy atoms and corresponding hydrogen atoms were constrained to their equilib-