J Mol Model (2010) 16:1003–1009
DOI 10.1007/s00894-009-0602-2

ORIGINAL PAPER

Homology modelling of a sensor histidine kinase from *Aeromonas hydrophila*

Mobashar Hussain Urf Turabe Fazil · Sunil Kumar · Naidu Subbarao · Haushila Prasad Pandey · Durg Vijai Singh

Received: 13 May 2009 / Accepted: 30 September 2009 / Published online: 29 October 2009

Abstract *Aeromonas hydrophila* has been implicated in extra-intestinal infection and diarrhoea in humans. Targetting unique effectors of bacterial pathogens is considered a powerful strategy for drug design against bacterial variations to drug resistance. The two-component bacterial system involving sensor histidine kinase (SHK) and its response regulators is considered a lucrative target for drug design. This is the first report describing a three-dimensional (3D) structure for SHK of *A. hydrophila*. The model was constructed by homology modelling using the X-ray structure of PleD—a response regulator—in conjunction with cdiGMP (PDB code 1W25) and HemAT sensor domain (PDB code 1OR4)—a globin coupled sensor. A combination of homology modelling methodology and molecular dynamics (MD) simulations was applied to obtain a reasonable structure to understand the dynamic behaviour of SHK. Homology modelling was performed using MODELLER9v2 software. The structure was relaxed to eliminate bad atomic contacts. The final model obtained by molecular mechanics and dynamics methods was assessed using PROCHECK and VERIFY 3D graph, which confirmed that the final refined model is reliable. Until complete biochemical and structural data of SHK are determined by experimental means, this model can serve as a valuable reference for characterising the protein and could be explored for drug targetting by design of suitable inhibitors.

Keywords Homology modelling · Sensor histidine kinase · *A. hydrophila* · Structural bioinformatics

Abbreviations

SHK	Sensor histidine kinase
EM	Energy minimisation
BLAST	Basic local alignment search tool
MDS	Molecular dynamics simulation

Introduction

Aeromonas hydrophila, a Gram-negative rod-shaped bacterium, is an autochthonous inhabitant of aquatic environments. This organism has been associated with intestinal and extra-intestinal infections in humans [1, 2]. Moreover, this organism has also been included in the contaminant candidate list by the United States Environmental Protection Agency [3]. *A. hydrophila* is known to produce several virulence factors such as haemolysin, haemaglutinnin, cytolysin and enterotoxin, and also exhibits resistance to normal human serum [4–9]. Little is known about metabolic activity in *A.*

Mobashar Hussain Urf Turabe Fazil and Sunil Kumar contributed equally to this work.

M. H. U. T. Fazil · D. V. Singh (✉)
Infectious Disease Biology, Institute of Life Sciences,
Nalco Square,
Bhubaneswar 751023, India
e-mail: durg.singh@gmail.com

S. Kumar
Bioinformatics Centre, Institute of Life Sciences,
Nalco Square,
Bhubaneswar 751023, India

N. Subbarao
Centre for Computational Biology and Bio-informatics,
School of Information Technology, Jawaharlal Nehru University,
New Delhi 110067, India

H. P. Pandey
Department of Biochemistry, Faculty of Science,
Banaras Hindu University,
Varanasi 221005, India

hydrophila or the factors that facilitate its ecological interactions with other prokaryotes or its host [10].

The ability of the bacterium to gauge parameters such as ionic strength, concentration of nutrients and debilitating compounds is a primary requisite for the organism to inhabit a given environment. Sensor histidine kinase (SHK) proteins are involved in signal transduction pathways in bacteria. These proteins, in conjunction with a number of response regulators, help to transduce signals from the extracellular environment to intracellular stimuli [11]. *A. hydrophila* is predicted to host about 31 SHKs paired with cognate regulators [12]. It can be assumed that one or more of these SHKs are involved in *A. hydrophila* virulence and pathogenesis by sensing and responding to signals present in the host environment.

SHKs act through sensing signals and responding through response regulators, either paired or as stand-alone proteins, to bring about changes in gene expression. The sense and response activities are carried out by a number of input and output domains in SHK. Various examples of signal transduction and sensory input domains include the HATPase_c (histidine kinase type ATPase catalytic) domain, MA domain (methyl accepting chemotaxis), HPt (histidine phospho transfer) domain, PAS (a domain identified in PER, ARNT and SIM proteins of *Drosophila*), and GAF (a domain common in cGMP phosphodiesterase, adenylyl cyclase, FhlA), etc. [13].

Many SHKs harbour a unique catalytic domain, characterised by the presence of the amino acid sequence G-G-D-E-F—the GGDEF domain [14]. This catalytic domain of SHKs is involved in the synthesis of cyclic diguanylate (CdiGMP), which acts as a second messenger to stimuli and leads to physiological responses [15]. CdiGMP regulates many diverse functions, including developmental transitions, social behaviour, adhesion, biofilm formation and virulence [16].

Fluoroquinolones have long been used for treatment for *Aeromonas* infections. However, *Aeromonas* strains showing resistance to nalidixic acid are known [17]. Tests of susceptibility of *A. hydrophila* to antimicrobial drugs revealed an increasing number of resistant phenotypes, suggesting acquisition of drug resistance through horizontal gene transfer [18]. These observations indicate the need to examine novel avenues of drug design against this pathogen.

At present, rational drug design involves identification of factors that inhibit virulence of pathogens without inhibiting their growth. This approach will lead to decreased selective pressure for development of drug resistance in the bacterium [19]. SHKs are recognised as a favourite choice for drug design against drug resistant bacteria because of their presence in the bacterial signal transduction system, which has no counterpart in the human host. SHKs in other enteropathogenic bacteria have been used successfully to evaluate the effectiveness of microbial inhibitors [20].

The preferred method for effective drug design is to build a three-dimensional (3D) structure and screen for drugs interacting with this structure. Protein function can be understood in greater detail using 3D structural information. Using X-ray crystallography, electron microscopy, diffraction and NMR spectroscopy, the actual 3D structure of a protein is obtained. However, no information is available on the crystal structure of SHKs. This may be due to the intrinsic difficulty of preparing high quality crystals of SHK. In the absence of experimental data, models based on the known 3D structure of a homologous protein are the only reliable method of obtaining structural information. Therefore, in the present study, we built a 3D-structural model of SHK (AHA_3297) based on the known 3D-structures of stalked-cell differentiation controlling protein from *Caulobacter vibrioides* (1W25) and the HemAT sensor domain from *Bacillus subtilis* (1OR4) using homology methodology. This is the first report describing a 3D structure of SHK, AHA_3297 in *A. hydrophila*, which was earlier designated as a potential drug target for *A. hydrophila* [21].

Materials and methods

Template search and sequence alignment

In the present investigation, the protein sequence of SHK (AHA_3297) of *A. hydrophila* was retrieved from the NCBI database (YP_857788.1). A BLASTP [22] search with default parameters was performed against the Brook Heaven Protein Data Bank (PDB) [23] to find suitable templates for homology modelling. There was no single template available in PDB that could be used to model SHK, therefore, two templates were selected. Based on the maximum identity with high score and lower e-value, stalked-cell differentiation controlling protein from *C. vibrioides* at 2.7 Å resolution (PDB code: 1W25) and the crystal structure of the HemAT sensor domain from *B. subtilis* at 2.15 Å resolution (PDB code: 1OR4) were used as templates, showing similarity at the C- and N-terminus, respectively, for homology modelling. Sequence identity was 31% (similarity 50%) and 27% (similarity 52%) between the target and templates 1W25 and 1OR4, respectively. The ClustalW (http://www.ebi.ac.uk/clustalw) [24] program was used for sequence alignment.

3D structure generation

The academic version of MODELLER (http//:www.salilab.org/modeller) [25] was used for 3D structure generation based on the information obtained from sequence alignment. In general, the homology modelling method is based on the assumption that the structure of an unknown protein

will be similar to the known structures of some reference proteins. Out of 20 models generated by MODELLER, the one with the best G-score of PROCHECK [26], and with the best VERIFY3D [27] profile was subjected to energy minimisation (EM). Using the parameters of a distance-dependent dielectric constant $\in$=1.0 and non-binding cutoff of 14 Å, CHARMm [28], force field and CHARM-all-atom charges, a steepest descent algorithm was used initially to remove close van der Waals contacts, followed by conjugate gradient minimisation until the energy was stable in sequential repetitions. All hydrogen atoms were included in the calculation. The EM was initiated with main chain of the core, and then all core side chains were subjected to the same procedure. All calculations were performed using ACCELRYS DS Modelling 2.0 (Accelrys, San Diego, CA) software suite. During these steps, the quality of the initial model was improved. VERIFY3D (a structure evaluation server) was used to check the residue profiles of the 3D models obtained. STRIDE [29] was used in prediction of secondary structure of the modelled SHK. PROCHECK analysis was performed to assess the stereo-chemical qualities of the 3D models.

Molecular dynamics simulation

A combination of homology modelling and molecular dynamics (MD) simulation was applied to obtain a reasonable structure that would allow exploration of its dynamic behaviour. The structure was further relaxed to eliminate bad atomic contacts and subsequently solvated with water. The whole system was subjected to EM. A MD simulation was carried out for 100 ps of simulation at 300 K after global minimisation using GROMACS 3.1.1 (http://www.gromacs.org). Root mean squares deviation (RMSD) values of the final conformation of the modelled protein are also presented. An SGI ALtix server running the IRIX operating system was used to perform MD simulation on GROMACS.

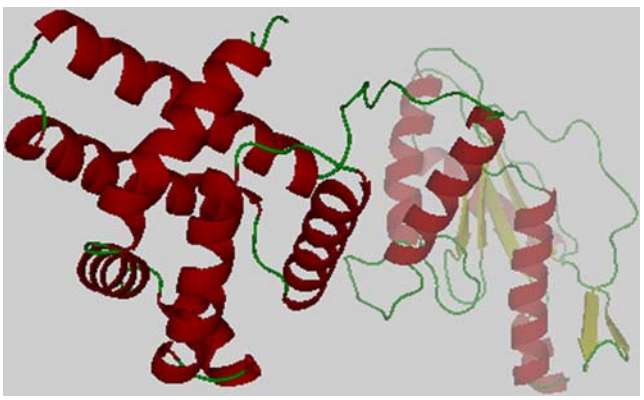

Fig. 1 Ribbon representation of modelled sensor histidine kinase (SHK). The α-helices and β-sheets are shown as *helices* and *ribbons*, respectively. The rest of the structure is depicted as *loops*. The figure was prepared using PyMol

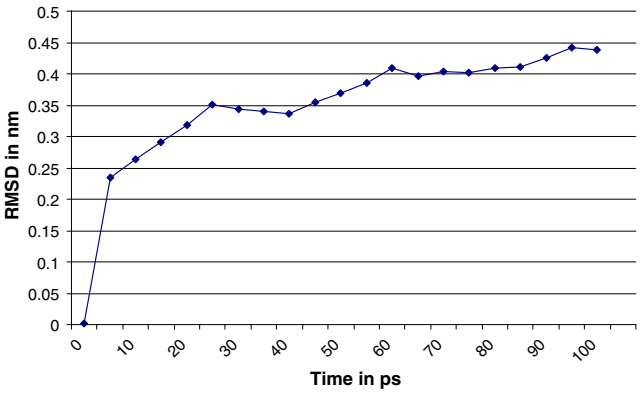

Fig. 2 Calculated root mean square deviation (RMSD) graph using GROMACS software. *x-axis* Time (ps), *y-axis* RMSD

Prediction of transmembrane helices

Different servers (DAS, HMMTOP, TMHMM, TMMPRED) were used to predict transmembrane helices in the SHK protein sequence.

Results and discussion

Model building

No high resolution structure of any SHK conjugate with its response regulator protein has yet been determined experimentally [30]. Therefore, we built a model following a homology modelling protocol. A BLASTP search with default parameters was performed against PDB to find

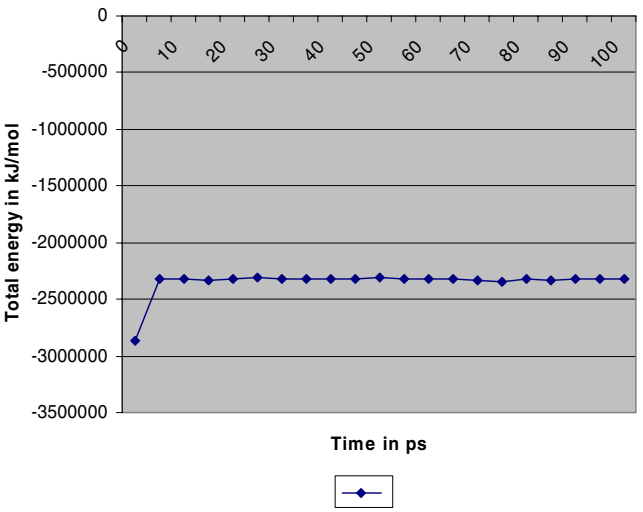

Fig. 3 Calculated energy vs time plot using GROMACS software. *x-axis* Time (ps), *y-axis* energy

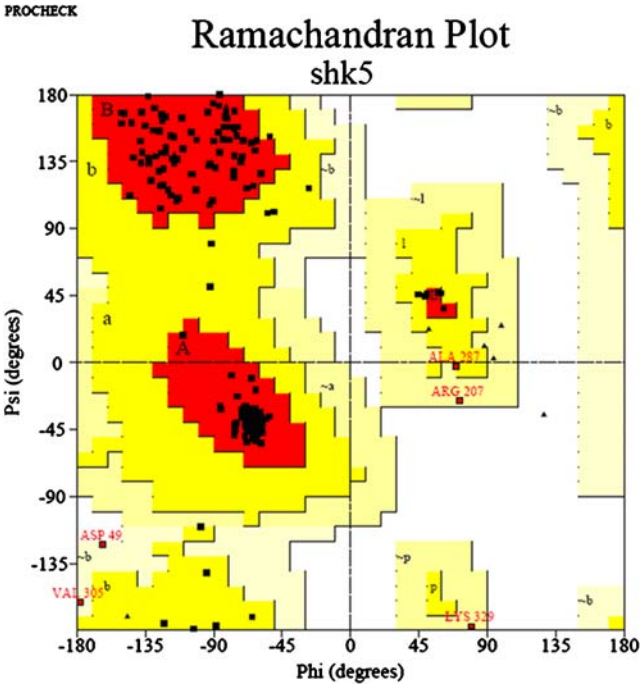

Fig. 4 Ramachandran map of SHK protein. The plot calculation on the three-dimensional (3D) model of SHK protein was calculated with the PROCHECK program

suitable templates for homology modelling. The N- and C-terminal region of the SHK protein was then modelled using the templates 1OR4 and 1W25, respectively. Sequence alignment was performed using the program ClustalW. The alignment was characterised by some insertions and deletions in the loop regions. Since the first 23 residues from the N terminal end had no corresponding equivalent region in 1OR4, modelling was carried out from the 24th residue of the target protein. Modelling was followed by a rigorous refinement of the model by means of EM. The final stable structure of SHK obtained is shown in Fig. 1.

MD simulation

Based on intrinsic dynamics, structural stability and improved relaxation of the modelled structure, the energy and RMSD (Figs. 2, 3) of the energy-minimised structure, which is an important criteria for the convergence of free MD simulation, were calculated. The targeted MD simulation provided better conformation for the 11-residue linker fragment connecting both N and C terminal domains at the lowest energy.

Protein structure validation

To validate the homology modelled SHK structure, a Ramachandran plot was drawn and the structure was analysed by PROCHECK, a well known protein structure checking program. It was found that the phi/psi angles of 93.8% of residues fell in the most favoured regions, 4.6% residues fell in the additional allowed regions, and 1.6% fell in generously allowed regions; none of the residues fell in the disallowed conformations (Fig. 4). The overall PROCHECK G-factor for the homology modelled structure after MD simulation was −0.06. This score indicates that the modelled structure is acceptable because the recommended value is greater than − 0.50. A decrease in the overall G factor was observed after MD simulation. These observations thus indicate that an increase in the number of bad dihedral angles of the modelled structure had occurred. This may be due to MD simulation causing an unfavourable dihedral angle, allowing the protein to overcome high energy barriers. The structural superimposition of C^{α} trace of the model on the two template structures 1OR4 and 1W25 (Fig. 5) resulted in RMSDs of 0.4 Å and 2.0 Å, respectively, for the N-terminal and C-terminal domains using the program Combinatorial Extension (CE; http://cl.sdsc.edu/ce.html).

Fig. 5 Structural alignment of N-terminal and C-terminal domains of the modelled structure with template structures using the program Combinatorial Extension (CE)

```
1OR4:A  39/40    KMVRLGDAELYVLEQLQPLIQENIVNIVDAFYKNLDHESSLMDIINDHSSVDRLKQTLKR
Model:_  1/2     ELLLLDERDFALLASYRPKIEPHIDALVDKFYTLQTGITEIALLIGDADTLTRLRAAQRR

1OR4:A  99/100   HIQEMFAGVIDDEFIEKRNRIASIHLRIGLLPKWYMGAFQELLLSMIDIYEASITNQQEL
Model:_ 61/62    YILDLFSGLYDLEYVNNRLRIGLVHKRIGVEPKLYLAAINTLKGLLIEDIFTQIDHEPDR

1OR4:A  159/160  LKAIKATTKILNLEQQLVLE
Model:_ 121/122  ITMLTALDKLFLFDITLVFE

1W25:A  262/264  ARVKTQIQRKRYTDYLRNNLDHSLELAVTDQLTGLHNRRYMTGQLDSLVKRATLGGDPVS
Model:_ 153/154  SKEKSEIYAKSLEEKVRERTQQLEEMSRIDALTGLLSVRHLQENLTRTLRTCQRRSEPVV

1W25:A  322/324  ALLIDIDFFKKINDTFGHDIGDEVLREFALRLASNVRAIDLPCRYGGEEFVVIMPDTALA
Model:_ 213/214  IAYLDIDDFKQINDSQGHQRGDEILRIVGDCIRKVSRAEDCCFRYGGDEFCLILPNCNE-

1W25:A  382/384  DALRIAERIRMHVSGSPFTVAHGREMLNVTISIGVSATA-GEGDTPEALLKRADEGVYQA
Model:_ 272/273  --EQARDV-----YLQRLTNALAEQISDVTLSIGLVQTGPVQFDEGNSLIRLADERMYAA

1W25:A  441/443  KASGRNAVVGKAA
Model:_ 325/326  KKAMKLVNQNSTH
```

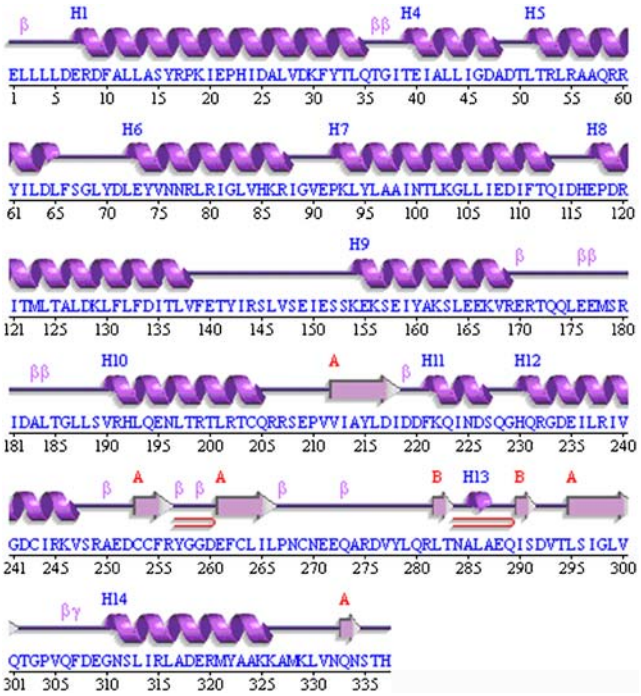

Fig. 6 Secondary structure assignment of modelled SHK protein

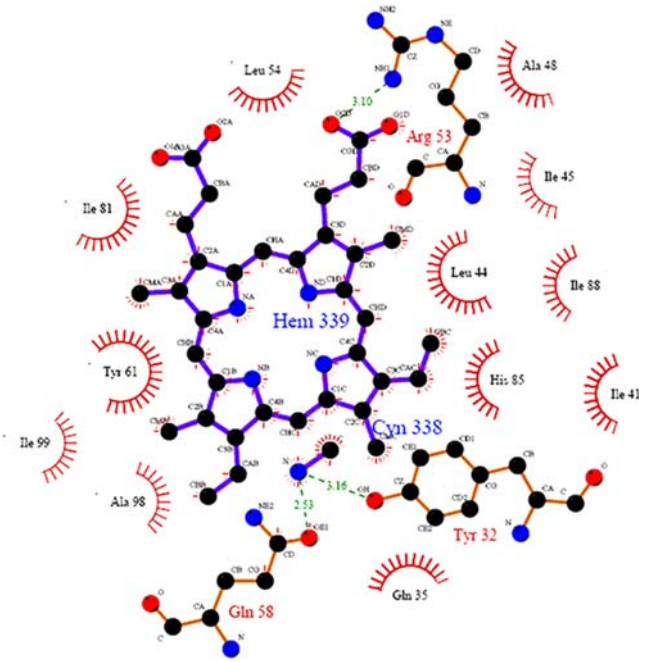

Fig. 8 Ligplot of interaction of haem with SHK protein

STRIDE [29], which uses hydrogen bond energy and main chain dihedral angles to recognise helix, coils and strands, was used to predict the secondary structure (Fig. 6) of the modelled SHK. We analysed SHK sequence using different transmembrane prediction algorithms—DAS [31], HMMTOP [32], TMHMM [33], TMPRED [34]—to predict the transmembrane helices in the protein. These predicting servers did not identify the presence of any transmembrane helices in the SHK protein.

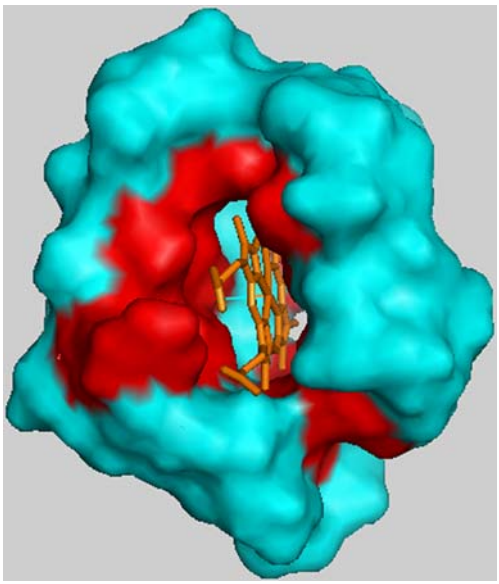

Fig. 7 Interaction of haem with SHK protein

Conserved domain identification and sequence alignments [35] were used, along with a protein BLAST [36, 37], to determine the domains present in the *A. hydrophila* SHK, AHA_3297. We found a sensor globin sensing domain and a diguanylate cyclase domain (DGC) in the amino acid sequences of this SHK.

The globin fold was characterised by the predominant presence of α helices [38]. Various sensor domains, such as HemAT, Dos, and FixL, are known to posses globin folds and are grouped with the family of globin-coupled sensors (GCS). Oligomeric haem-based sensors are known to be mediators of responses to intra- and extra-cellular stimuli. A globin-coupled sensor motif was also identified along with a phosphodiesterase domain, in addition to histidine kinases. A total of 17 amino acid residues were conserved in the sensor globin domain. An amino acid sequence comparison was carried out with the well characterised oxygen sensor binding domain from *B. subtilis*—a representative GCS—and other GCS proteins. The results showed that, of the 17 conserved residues, 8 residues, namely Phe (position 31), Tyr (positions 32, 95), Ile (positions 45, 81, 88), Leu (position 54) and His (position 85) were conserved in AHA_3297. However, substantial residue variation was observed at other positions known to be conserved. The high rate of sequence variability may be attributed to a specific signal sensing mechanism unique to this protein. In the majority of haem-based sensors, except those that contain a thiolate ligand, a conserved His residue is thought to be involved in binding haem through coordination [39]. We aligned the template with the target

structure along with haem using all atom superposition in a structural comparison program [40] to construct a complex of haem plus template. This complex was subjected to EM to study the binding residues involved in haem binding in this protein in order to understand similarities with other known sensor proteins. Tyr 32, Arg 53, and Glu 58 were directly involved in binding of haem through hydrogen bonding, whereas a number of other amino acid residues, including His 85, showed involvement in binding of the haem moiety through a coordination complex (Figs. 7, 8).

The DGC domain carries out the condensation of two GTP to CdiGMP. The DGC or GGDEF domain occurs in various combinations with sensory and/or regulatory domains. The DGC domain consists of a five-stranded central β-sheet surrounded by helices. The G-G-D-E-F signature motif occurs at the central β hairpin. Around ten amino acid residues appear to be conserved in proteins carrying the GGDEF domain [41].

Of the ten residues conserved in the GGDEF domain, all were present and conserved in SHK. *A. hydrophila* SHK, AHA_3297 contained a conventional and conserved G-G-D-E-F sequence at positions 258–262, which was present at the β hairpin bend of the central β sheet. The catalytically important Asp residues of the adenylate cyclase domain, as observed in all known GGDEF domains, occurred at positions 226 and 234 [42]. The presence of haem-binding amino acid residues at the N-terminus showed that this model is reliable, indicating that the sensing domain belongs to the family of haem-based GCS. This implies that this SHK may be involved in the intra-cellular signalling mechanism that senses intra-cellular concentrations of signals through the haem-based sensing domain. The high degree of variability in amino acid sequences among proteins of the same family indicate that there may be a protein specific evolution to perceive unique signals or a unique mechanism of signal perception in SHKs [13]. As the SHK contains a GGDEF domain and does not possess a region of homology with phophodiesterase domain, it can be predicted that it acts mainly through regulation of cyclic-diGMP levels. This molecular model can be used for characterisation, structural analysis and drug design against this SHK of *A. hydrophila*. Further biochemical and structural characterisation of this protein will be required to assign its precise functions.

Conclusions

Emergence of antibiotic resistance in pathogenic bacterial strains is a common phenomenon. Therefore there is a definite need to look at atypical drugs and drug targets in order to outscore the bacterial pathogen. SHK is considered to be one of the novel targets that can be exploited to counteract the increasing menace of multi-drug resistance in bacteria. Due to their precise inhibition of signalling cascades, which lead to only low selective pressure on the organism, causing a slowdown in the process of the evolution of resistance, SHKs are one of the targets of choice. The validation of a 3D structure of the *A. hydrophila* SHK is a positive step towards the rational drug design to combat this emerging human pathogen.

Acknowledgements The study was supported by funds contributed by the Department of Biotechnology to the Institute of Life Sciences, Bhubaneswar, India. A Senior Research Fellowship to M.H.U.T.F. awarded by Indian Council of Medical Research, New Delhi is kindly acknowledged. S.K. was supported by Distributed Information Sub-centre of Institute of Life Sciences, Bhubaneswar. We thank Dr. Sanjeev Kumar Singh and Mr. Sunil Tripathi for their help with the molecular dynamics simulation study.

References

1. Altwegg M, Geiss HK (1989) *Aeromonas* as a human pathogen. Crit Rev Microbiol 16:253–286
2. Janda JM, Duffey PS (1988) Mesophilic aeromonads in human disease: current taxonomy, laboratory identification, and infectious disease spectrum. Rev Infect Dis 10:980–997
3. USEPA (1998) Announcement of the drinking water contaminant candidate list: Notice Federal Register. 63:10274–10287
4. Annapurna E, Sanyal SC (1977) Enterotoxicity of *Aeromonas hydrophila*. J Med Microbiol 10:317–323
5. Burke V, Robinson J, Atkinson HM, Gracey M (1982) Biochemical characteristics of enterotoxigenic *Aeromonas* spp. J Clin Microbiol 15:48–52
6. Chakraborty T, Montenegro MA, Sanyal SC, Helmuth R, Bulling E, Timmis KN (1984) Cloning of enterotoxin gene from *Aeromonas hydrophila* provides conclusive evidence of production of a cytotonic enterotoxin. Infect Immun 46:435–441
7. Singh DV, Sanyal SC (1992) Production of haemolysis and its correlation with enterotoxicity in *Aeromonas* spp. J Med Microbiol 37:262–267
8. Singh DV, Sanyal SC (1992) Biochemical characteristics and enterotoxicity of *Aeromonas* species isolated from man and environment. J Diarrhoeal Dis Res 10:231–234
9. Singh DV, Sanyal SC (1993) Haemagglutinating activity, serum sensitivity and enterotoxigenicity of *Aeromonas* spp. J Med Microbiol 38:49–53
10. Khalil AH, Mansour EH (2003) Toxicity of crude extracellular products of *Aeromonas hydrophila* in tilapia, Tilapianilotica. Lett Appl Microbiol 25:269–273
11. Stock AM, Robinson VL, Goudreau PN (2000) Two-component signal transduction. Annu Rev Biochem 69:183–215
12. Seshadri R, Joseph SW, Chopra AK, Sha J, Shaw J, Graf J, Haft D, Wu M, Ren Q, Rosovitz MJ, Madupu R, Tallon L, Kim M, Jin S, Vuong H, Stine OC, Ali A, Horneman AJ, Heidelberg JF (2006) Genome sequence of *Aeromonas hydrophila* ATCC 7966T: jack of all trades. J Bacteriol 188:8272–8282
13. Mascher T, Helmann JD, Unden G (2006) Stimulus perception in bacterial signal-transducing histidine kinases. Microbiol Mol Biol Rev 70:910–938
14. Volz K (1993) Structural conservation in the CheY superfamily. Biochemistry 32:11741–11753

15. Paul R, Weiser S, Amiot NC, Chan C, Schirmer T, Giese B, Jenal U (2004) Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. Genes Dev 18:715–727
16. Ryan RP, Fouhy Y, Lucey JF, Dow JM (2006) Cyclic di-GMP signaling in bacteria: recent advances and new puzzles. J Bacteriol 188:8327–8334
17. Vila J, Marco F, Soler L, Chacon M, Figueras MJ (2002) In vitro antimicrobial susceptibility of clinical isolates of *Aeromonas caviae*, *Aeromonas hydrophila* and *Aeromonas veronii* biotype *sobria*. J Antimicrob Chemother 49:701–702
18. Huddleston JR, Zak JC, Jeter RM (2006) Antimicrobial susceptibilities of *Aeromonas* spp. isolated from environmental sources. Appl Environ Microbiol 72:7036–7042
19. Barett JF, Hoch JA (1998) Two-component signal transduction as a target for microbial anti-infective therapy. Antimicrob Agents Chemother 42:1529–1536
20. Rasko DA, Moreira CG, de Li R, Reading NC, Ritchie JM, Waldor MK, Williams N, Taussig R, Wei S, Roth M, Hughes DT, Huntley JF, Fina MW, Falck JR, Sperandio V (2008) Targeting QseC signaling and virulence for antibiotic development. Science 321:1078–1080
21. Sharma V, Gupta P, Dixit A (2008) In silico identification of putative drug targets from different metabolic pathways of *Aeromonas hydrophila*. In Silico Biol 8:331–338
22. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
23. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–242
24. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680
25. Sali A, Blundell TL (1993) Comparative protein modeling by satisfaction of spatial restraints. J Mol Biol 234:779–815
26. Laskoswki RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291
27. Eisenberg D, Luthy R, Bowie JU (1997) VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol 277:396–404
28. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1993) CHARMM: a program for macromolecular energy minimization and dynamics calculations. J Comput Chem 4:187–217
29. Frishman D, Argos P (1995) Knowledge-based protein secondary structure assignment. Proteins 23:566–579
30. Skerker JM, Perchuk BS, Siryaporn A, Lubin EA, Ashenberg O, Goulian M, Laub MT (2008) Rewiring the specificity of two-component signal transduction systems. Cell 133:1043–1054
31. Cserzo M, Wallin E, Simon I, von Heijne G, Elofsson A (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng 10:673–676
32. Tusnady GE, Simon I (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. J Mol Biol 283:489–506
33. Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580
34. Hofmann K, Stoffel W (1993) TMbase—a database of membrane spanning proteins segments. Biol Chem Hoppe Seyler 374(166–171):166–171
35. Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao L, He S, Hurwitz DI, Jackson JD, Ke Z, Krylov D, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Thanki N, Yamashita RA, Yin JJ, Zhang D, Bryant SH (2007) CDD: a conserved domain database for interactive domain family analysis. Nucleic Acids Res 35:D237–D240
36. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402
37. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schaffer AA, Yu YK (2005) Protein database searches using compositionally adjusted substitution matrices. FEBS J 272:5101–5109
38. Hou S, Freitas T, Larsen RW, Piatibratov M, Sivozhelezov V, Yamamoto A, Meleshkevitch EA, Zimmer M, Ordal GW, Alam M (2001) Globin-coupled sensors: a class of heme-containing sensors in Archaea and Bacteria. Proc Natl Acad Sci USA 98:9353–9358
39. Rodgers KR (1999) Heme-based sensors in biological systems. Curr Opin Chem Biol 3:158–167
40. Subbarao N, Haneef I (1991) Defining topological equivalences in macromolecules. Protein Eng 4:877–884
41. Chan C, Paul R, Samoray D, Amiot NC, Giese B, Jenal U, Schirmer T (2004) Structural basis of activity and allosteric control of diguanylate cyclase. Proc Natl Acad Sci USA 101:17084–17089
42. Pei J, Grishin NV (2001) GGDEF domain is homologous to adenylyl cyclase. Proteins 42:210–216