J Mol Model (2011) 17:1109–1119
DOI 10.1007/s00894-010-0792-7

ORIGINAL PAPER

Study of interactions between *Mycobacterium tuberculosis* proteins: SigK and anti-SigK

Vasavi Malkhed · Bargavi Gudlur · Bhargavi Kondagari · Ramasree Dulapalli · Uma Vuruputuri

Received: 19 October 2009 / Accepted: 23 June 2010 / Published online: 31 July 2010
© Springer-Verlag 2010

Abstract The development of novel antituberculosis therapeutic molecules is a global health concern. Complex gene expression in *Mycobacterium tuberculosis* is mediated mainly by various sigma factors. The SigK protein binds to RNA polymerase, facilitating the expression of genes encoding the antigenic proteins mpt70 and mpt83. The anti-SigK protein is a negative regulator of SigK and inhibits the initiation of transcription. This study focuses on the interactions between SigK and the N-terminal domain of anti-SigK. The 3D structures of SigK (187 residues) and the N-terminal domain of anti-SigK (92 residues) are elucidated, using the crystal structures of the A and B chains of sigma E and anti-sigma ChrR of *Rhodobacter spheroides* (PDB code: 2Q1Z) as templates, respectively. Molecular dynamic simulations were performed for the SigK and anti-SigK proteins to refine their structures. The predicted active sites of SigK and anti-SigK and the results of protein–protein docking studies revealed the residues that are important for binding. The models generated and the binding site residues identified in this work throw new light on the interactions between the sigma K and anti-sigma K proteins, which should further aid the modulation of antigenic protein production in *Mycobacterium tuberculosis*.

Keywords *Mycobacterium tuberculosis* · ECF subfamily · Anti-sigma K factor rskA · RNA polymerase sigma factor sigma K · Homology modeling · Molecular dynamics (MD) simulation · Docking

V. Malkhed · B. Gudlur · B. Kondagari · R. Dulapalli · U. Vuruputuri (✉)
Department of Chemistry, Nizam College, Osmania University, Basheerbagh,
Hyderabad 500001 Andhra Pradesh, India
e-mail: vuma@osmania.ac.in

V. Malkhed
e-mail: mvasavi2003@gmail.com

Introduction

Mycobacterium tuberculosis (MTB) is the causative agent of a disease that is major public health problem—tuberculosis (TB). This disease currently causes three million deaths annually, although this figure is rapidly increasing every year. According to the WHO, without a coordinated effort to control it, TB will infect an estimated one billion people by 2020, killing 70 million worldwide [1–4]. Despite advances in chemotherapy and the BCG (bacillus Calmette–Guérin) vaccine, TB remains a significant contagious disease. The development of multi-drug-resistant tuberculosis strains (MDR-TB) and co-infection with human immunodeficiency virus (HIV) as well as its high capacity to survive in different host environments present a challenge to existing therapeutic strategies [5, 6]. The present long-term goal of drug designers is not just to develop compounds that kill the *Mycobacterium*, but rather to design and develop compounds that can also be used as co-therapies to improve and preserve the efficacy of traditional TB treatments [7].

Gene expression in response to specific adaptations, stress stimuli and environmental changes is the main method that *Mycobacterium* uses to survive in the host. Transcription initiation under various conditions is controlled by the RNA polymerase sigma subunit. RNA polymerase sigma factor aids in the specific promoter recognition of DNA during transcription initiation [8–10]. Environmental changes in the periplasm are communicated to the genetic machinery of the cell through one or two membranes and the cytoplasm. The mechanism of action associated with sigma factors involves a two-component

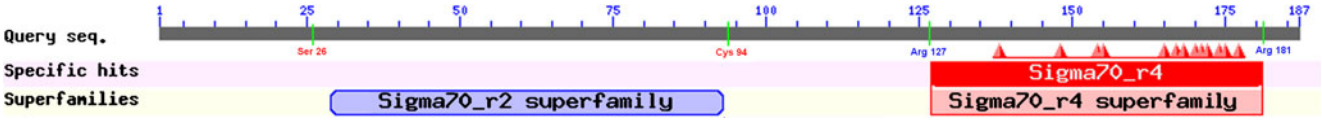


Fig. 1 Conserved domains in the sigma K protein sequence

transmembrane signaling path. Anti-sigma factor inhibits transcription initiation by interrupting the binding of sigma to RNA polymerase and is a negative regulator of transcription. The *Mycobacterium tuberculosis* genome contains 13 sigma factors, ten of which belong to the ECF subfamily [11, 12]. Sigma factor K (SigK) belongs to the ECF subfamily and regulates the expression of genes encoding the antigenic proteins mpt70 and mpt83. Anti-sigma K is the negative regulator of these antigenic proteins. mpt70 and mpt83 are expressed at different levels in different species: low in *Mycobacterium tuberculosis* and high in the bacille Calmette–Guérin (BCG) strain of *Mycobacterium bovine*. This variation is due to the lack of the functional SigK repressor anti-sigma K in *Mycobacterium bovine*, meaning that the expression of mpb70 and mpb83 is high [13, 14]. Studying the interactions between the SigK and the anti-SigK proteins in *Mycobacterium tuberculosis*, homology modeling, and docking studies will all aid the development of novel targets that could improve the treatment of TB.

Materials and methods

Homology modeling

FASTA sequences of *Mycobacterium tuberculosis* sigma K and anti-sigma K protein sequences with accession codes of O53730 and O53729 were retrieved from the Expert Protein Analysis System (ExPASy Swiss-Prot/TrEMBL) server. Transmembrane secondary structure topology was predicted using the SPLIT 4.0 server [15] and was also taken from ExPASy. The sequences were subjected to the Basic Local Alignment Search Tool (BLAST) [16], the PHYRE server [17] and the PSI-BLAST module of the JPred3 server [18] by setting each server to search the Protein Data Bank. Sequence identity, E-value (a statistical measure) and secondary structure similarity were used as constraints during the selection of templates. Pairwise alignment was carried out with ClustalW, using the GONNET matrix [19, 20] to define the conserved regions: identities, similarities and differences between the target and the template. The automated alignment was inspected manually to ensure the presence of conserved motifs and to minimize the number of gaps and insertions. The initial homology models were generated using MODELLER 9v2 [21], and the model with the lowest value of the Modeller objective function was retrieved to further refine structurally variable regions (SVRs). The loop regions were modeled with the aid of "Build Loop" and "Scan Loop" in Swiss-PdbViewer [22], which counts the clashes (bad contacts and H-bonds), provides energy information by performing a partial implementation of the GROMOS96 43b1 force field in vacuo, and computes the mean force potential *PP*.

Molecular dynamics

Molecular dynamics simulation was performed under periodic boundary conditions using NAMD2.7b2 [23] at constant molecular number, pressure and temperature (NPT). The CHARMM27 [24, 25] force field was applied to the proteins and solvated with a 10 Å layer of water in all directions. The all-atom model of the SigK protein consisted of 2934 atoms and had a net charge of +1; one chloride counterion was added to make the system electrically neutral. The anti-SigK protein had four negative charges and was neutralized by adding four sodium ions. The sodium and chloride ions were placed 8 Å from the protein at a physiological concentration of 0.15 $moll^{-1}$. The

Table 1 Comparison of the template search results (E-values and PDB codes) obtained from various servers for the SigK protein

Search no.	Name of the protein database search server	Parameter(s) considered for template selection	E-value	PDB code of protein
1.	BLAST	Sequence position specificity	3×10^{-20}	2q1z A
			7×10^{-19}	2z2s A
2.	JPred3	Secondary structure prediction, solvent accessibility and coiled-coil region prediction	9×10^{-20}	2q1z A
			9×10^{-20}	2z2s A
3.	PHYRE	Protein fold recognition (threading)	5.1×10^{-29}	2q1z A
			4.3×10^{-28}	1or7 A

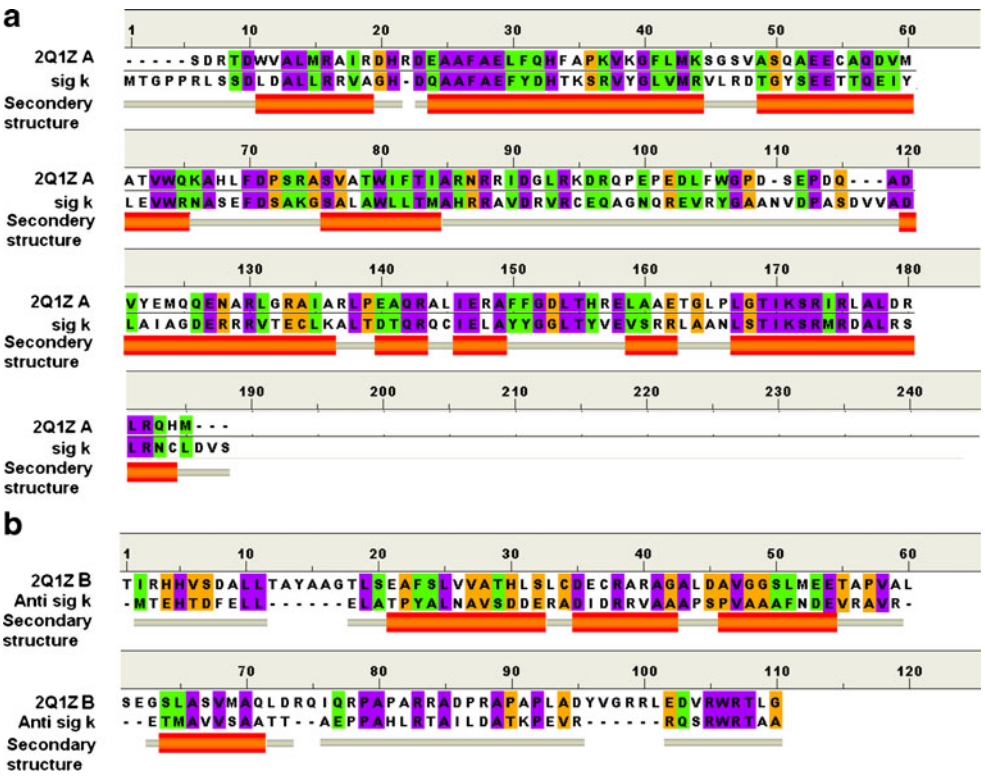


Fig. 2 **a** Sequence alignment of the *Mycobacterium tuberculosis* SigK protein with the template protein *Rhodobacter sphaeroides* sigma E (PDB ID: 2Q1Z, A chain). **b** Sequence alignment of *Mycobacterium tuberculosis* anti-SigK protein with cognate *Rhodobacter sphaeroides* anti-sigma CHrR (PDB ID: 2Q1Z, B chain). Identical residues are shown in *violet*, strongly similar residues in *green*, and weakly similar in *mustard yellow*. The secondary structures of SigK in **a** and anti-SigK in **b** are shown as *orange bars* for helices and *gray lines* for loop regions

protein simulation was performed at 300 K, as the proteins are stable at this temperature [26]. The pressure was held at 1 bar using a Langevin piston coupling algorithm [27]. The integrated time step for the simulation was set to 2 fs and the DCD frequency to 100. Nonbonded van der Waals interactions were treated with a switching function of 10 Å and reached zero at a distance of 12 Å. Long-range electrostatic forces between the atoms were treated with a Particle mesh Ewald (PME) algorithm, as PME is an efficient full electrostatics method with periodic boundary conditions [28, 29]. Initially 1500 equilibration steps and 25,000 production steps were performed in the simulation. VMD 1.8.7 was used to prepare the input files and analyze the results of the simulation [30].

The stereochemical quality of the resulting protein 3D model was validated using the PROCHECK suite [31]. The native fold, pairwise energy, surface energy and Z-score of the modeled structure were calculated using the ProSA II server [32].

Protein–protein docking

Cavities in the proteins that could act as active sites for binding were located using the CASTp [33] and Q-Site [34]

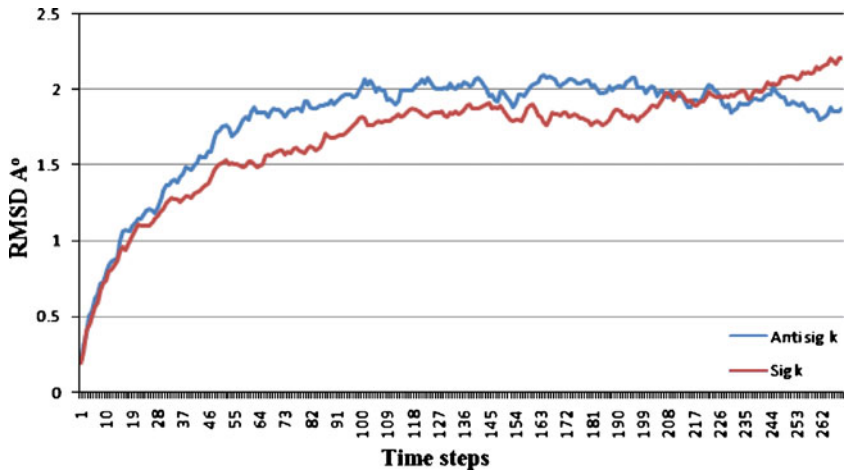


Fig. 3 Plot of MD simulation time step versus RMSD for the SigK and anti-SigK proteins. Each time step in the graph represents 200 fs

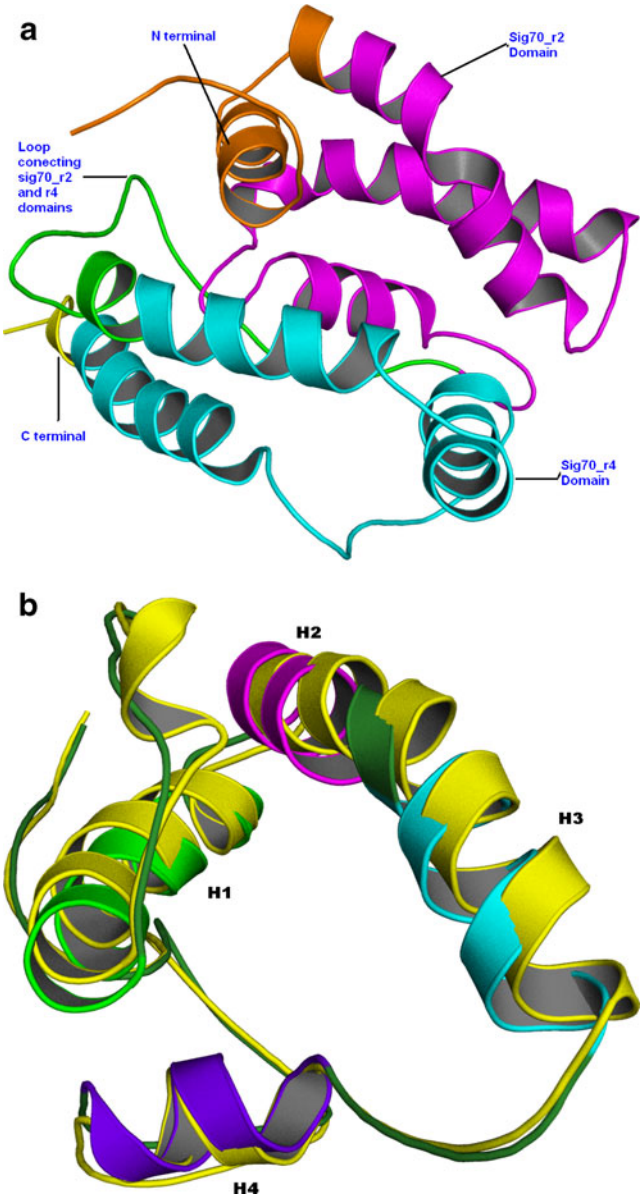
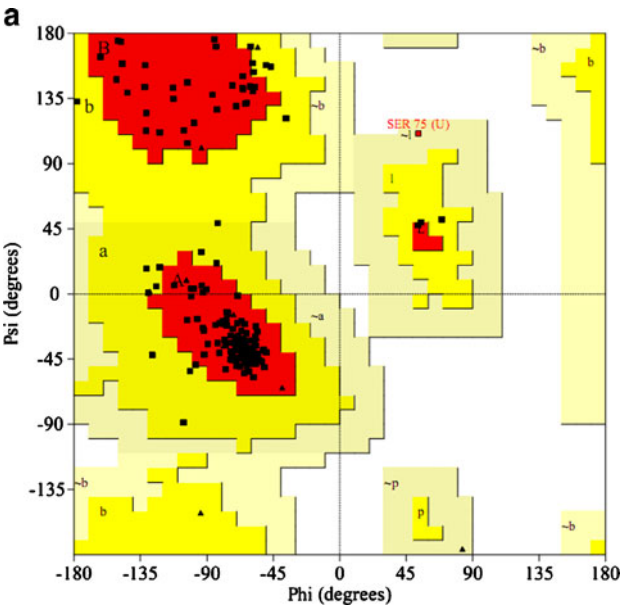
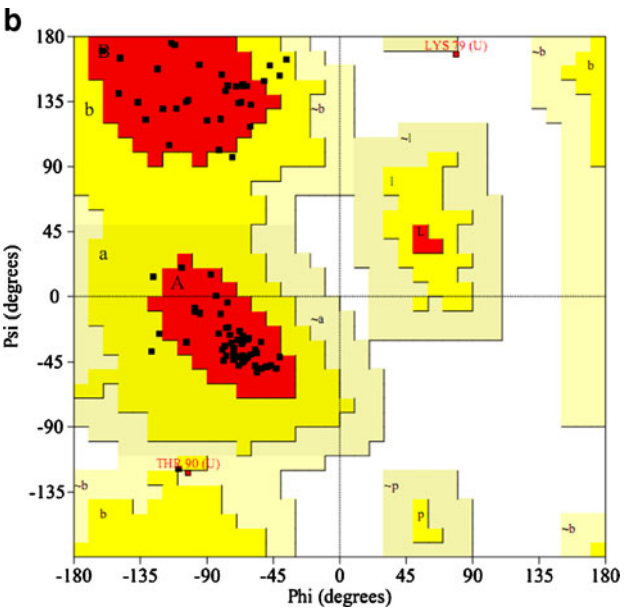


Fig. 4 **a** 3D structure of the SigK protein (after molecular dynamics simulation). Conserved regions of the group 4 sigma factor subfamily (extracytoplasmic function, ECF) are shown in different colors: *orange*, N terminal; *magenta*, Sig70_r2 domain; *cyan*, Sig70_r4 domain; *green*, the loop connecting these two regions; *yellow*, C terminal. **b** 3D structure of the anti-SigK protein cytoplasmic domain before (*yellow*) and after molecular dynamics (MD) simulation. Four conserved helices of the group 4 anti-sigma domain after MD simulation are shown in different colors: *green*, H1; *magenta*, H2; *cyan*, H3; *purple*, H4



Legend : Ramachandran plot statistics (Summary) for Sig k protein

Residues in most favoured regions [A, B, L]	154	89.5%
Residues in additional allowed regions [a, b, l, p]	17	9.9%
Residues in generously allowed regions [~a, ~b, ~l, ~p]	1	00.6%
Residues in disallowed regions		00.0%
Number of non-glycine and non-proline residues	172	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	10	
Number of proline residues	3	
Total number of residues	187	



Legend : Ramachandran plot statistics (Summary) for Anti sig k protein

Residues in most favoured regions [A, B, L]	74	88.1%
Residues in additional allowed regions [a, b, l, p]	8	9.5%
Residues in generously allowed regions [~a, ~b, ~l, ~p]	1	1.2%
Residues in disallowed regions	1	1.2%
Number of non-glycine and non-proline residues	84	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	0	
Number of proline residues	6	
Total number of residues	92	

Fig. 5 Conformational analyses of **a** the SigK protein and **b** the anti-SigK protein using PROCHECK ▶

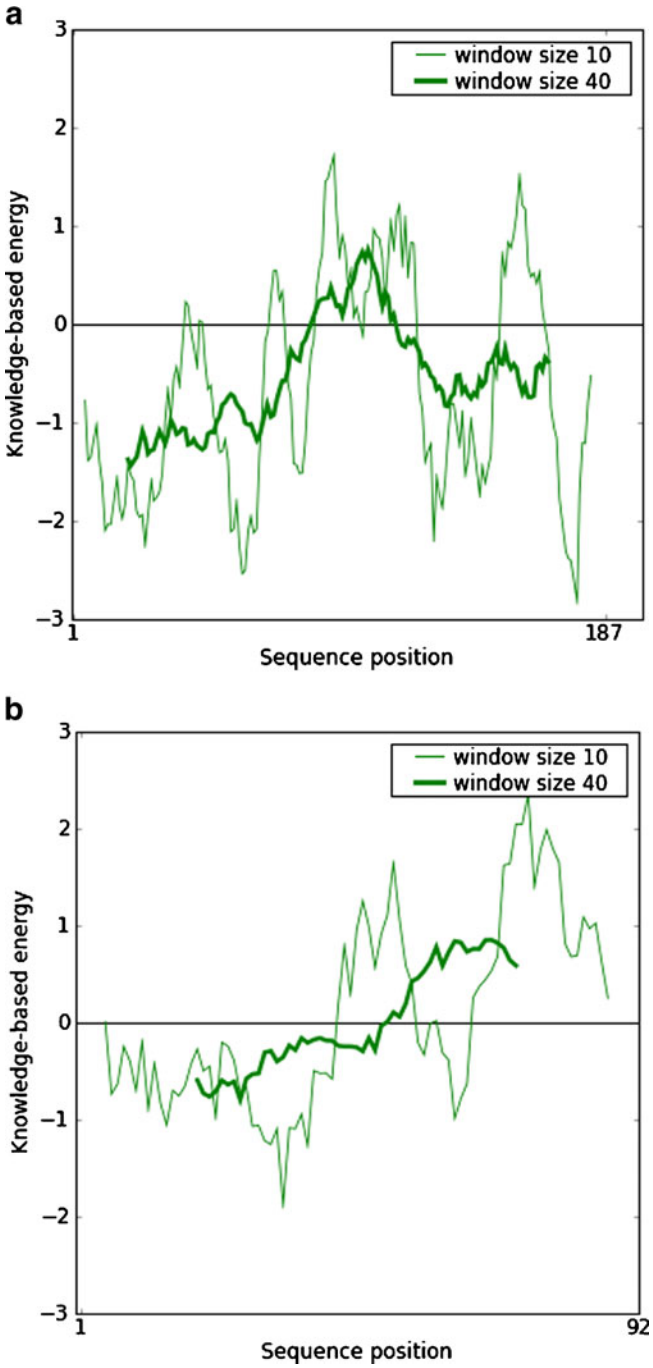


Fig. 6 **a** SigK protein energy profile obtained using ProSA II analysis, with a calculated Z-score value of −3.95. **b** Anti-SigK protein energy profile obtained using ProSA II analysis, with a calculated Z-score value of −1.25

servers, and the conserved domains in the protein family were taken into account. The residues identified in the active site were blocked by applying block.pl in ZDOCK 2.3 [35]. The ZDOCK scoring function considers pairwise shape complementarity with desolvation and electrostatics (PSC + DE + ELEC) at a 6° rotational sampling interval. Initial protein–protein docking studies were carried out using ZDOCK 2.3. Docked poses were created using create.pl, and the 2000 docked complexes generated were ranked using the Z-score. The top-ranked docked complexes were then re-ranked according to energy-based potential using ZRANK [36]. The solvent-accessible surface area (SASA) was calculated, setting the probe radius to 1.40 Å and using 240 grid points per atom, with Accelrys DS Visualizer [37]. The hydrogen bonds in the complexes were generated, visualized and analyzed using Accelrys DS Visualizer and PyMOL [38].

Results and discussion

This section is divided up into discussions of the results from homology modeling, active site prediction, and protein–protein docking studies of SigK and anti-SigK, respectively.

SigK of *Mycobacterium tuberculosis* is a transcription initiation protein that belongs to the sigma 70 family and the ECF subfamily, and has 187 amino-acid residues. Two conserved domains of the SigK protein identified from BLAST are given in Fig. 1. The first domain, the sigma 70_r2 superfamily, ranges from Ser26 to Cys94, while the second, the sigma 70_r4 superfamily, lies between Arg127 and Arg181.

SigK protein sequence homologs in the Protein Data Bank were obtained from the BLAST, PHYRE and Jpred servers, and their corresponding E-values are given in Table 1. A low E-value indicates a high protein sequence identity [16]. The

Table 2 Summary of helices present in the modeled anti-SigK and SigK proteins

Start	End	No. of residues	Length	Amino-acid sequence
Anti-SigK protein				
Thr14	Glu25	12	19.74	TPYALNAVSDDE
Asp28	Ala35	8	11.94	DIDRRVAA
Ser38	Glu47	10	15.84	SPVAAAFNDE
Thr54	Ala61	8	12.73	TMAVVSAA
SigK protein				
Leu11	Val18	8	11.77	LDALLRRV
Gln23	Tyr30	8	12.47	QAAFAEFY
Lys34	Arg43	10	16.15	KSRVYGLVMR
Thr48	Arg64	17	25.57	TGYSEETTQEIYLEVWR
Ala76	Met83	8	12.01	ALAWLLTM
Asp119	Lys135	17	25.87	DLAIAGDERRRVTECLK
Asp139	Arg142	4	6.15	DTQR
Ile145	Ala148	4	6.01	IELA
Val158	Arg161	4	7.17	VSRR
Ser167	Cys183	17	26.70	STIKSRMRDALRSLRNC

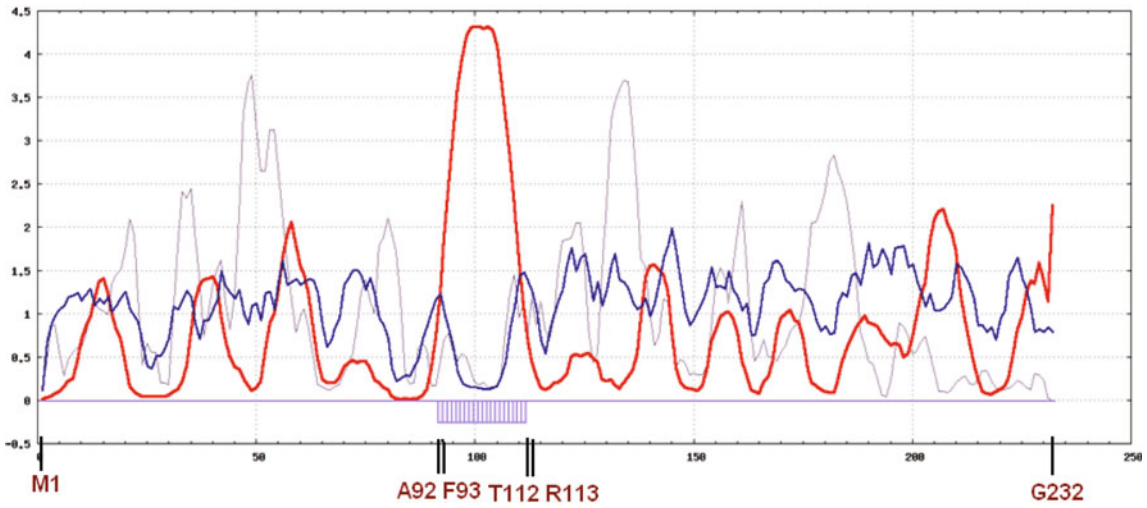


Fig. 7 Secondary structure prediction for the transmembrane protein anti-SigK of *Mycobacterium tuberculosis*, obtained using the method of preference functions for proteins in the SPLIT 4.0 server. *Red line*, transmembrane helix preference (THM index); *blue line*, beta preference (BET index); *gray line*, modified hydrophobic moment index (INDA index); *violet boxes (below abscissa)*, predicted transmembrane helix position (DIG index). Residue numbers with single-letter amino-acid codes are marked below the plot

Table 3 Binding pockets in the SigK and anti-SigK proteins, with the corresponding cavity volumes in Å^3, as defined by the CASTp and Q-Site servers

Cavity	SigK CASTp volume	SigK Q-Site volume	anti-SigK CASTp volume	anti-SigK Q-Site volume
1	1622.6	285	1368.6	171
2	341.6	177	496.3	120
3	312.6	160	98.6	145
4	242.0	145	87.3	155
5	143.5	145	26.7	107
6	130.2	142	32.5	95

Table 4 Intermolecular interactions in the docked complex of the SigK and anti-SigK proteins

Bond	Residue name and atom name	Bond distance in Å
Hydrogen bond 1	ARG101:HH11 - GLU3:OE2	2.213
Hydrogen bond 2	ARG160:HH11 - GLU8:OE1	2.091
Hydrogen bond 3	ARG160:HH21 - GLU8:OE1	2.476
Hydrogen bond 4	SER187:OT1 - ARG87:HH22	2.403
Salt bridge 1	SER185:OD2 - ARG87:NH1	3.534
Salt bridge 2	GLU3:OE2 - ARG101:NH1	2.825
Salt bridge 3	GLU3:OE2 - ARG101:NH2	3.613
Salt bridge 4	ASP6:OD2 - ARG174:NH1	3.927
Salt bridge 5	ASP6:OD2 - ARG174:NH2	3.914
Salt bridge 6	GLU8:OE1 - ARG160:NH1	3.002
Salt bridge 7	GLU8:OE1 - ARG160:NH2	3.292
Salt bridge 8	GLU8:OE2 - ARG160:NH1	3.894

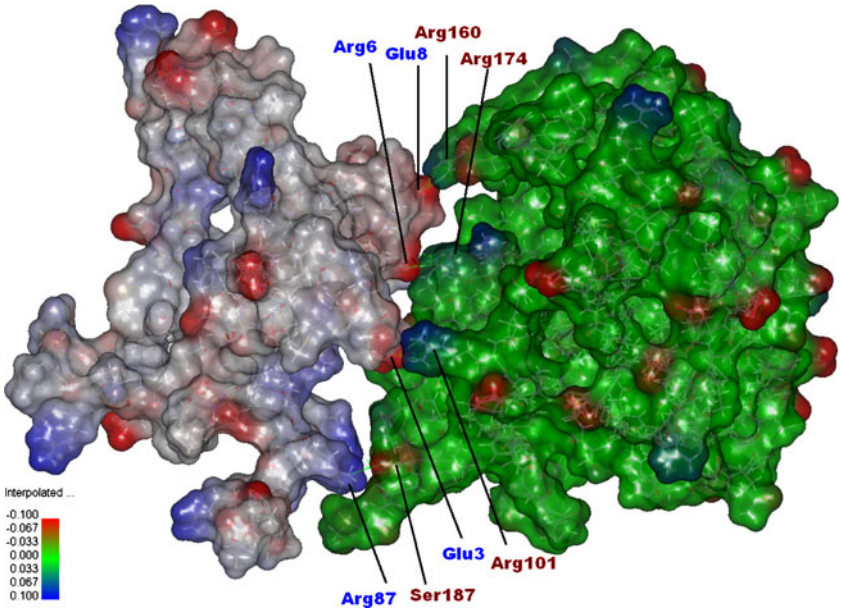


Fig. 8 Anti-SigK (*gray*) docked with SigK (*green*). The figure shows a molecular surface representation of the binding region based on electrostatic potential distributions

A chain of the sigma E protein of *Rhodobacter spheroides* (PDB code: 2Q1Z; 2.4 Å resolution) was chosen as the template protein based on E-values and the maximum sequence alignment. Alignment of the SigK protein with the template protein gives a sequence identity of 32% and a sequence similarity of 52%, as shown in Fig. 2a. Twenty-five initial models were generated using MODELLER 9v2, and the model with the lowest value of the Modeller objective function (719.28) was considered for further refinement. MD simulation of the SigK protein resulted in an average RMSD per residue of 1.85 Å, which is ≤2 Å. This shows that the resulting structure is that of the native structure of the protein [39, 40]. The RMSDs of the SigK protein structure at various time steps with respect to the initial structure are shown in Fig. 3. The 3D model obtained after refinement, loop building and molecular dynamic simulations is shown in Fig. 4a. A Ramachandran contour plot of phi versus psi (backbone dihedral angles) for the SigK protein, along with plot statistics, is shown in Fig. 5a. Among the 187 amino-acid residues, 154 are in the most favored region, 17 residues are in the additionally allowed region, one residue is in the generously allowed region, and none are in the disallowed

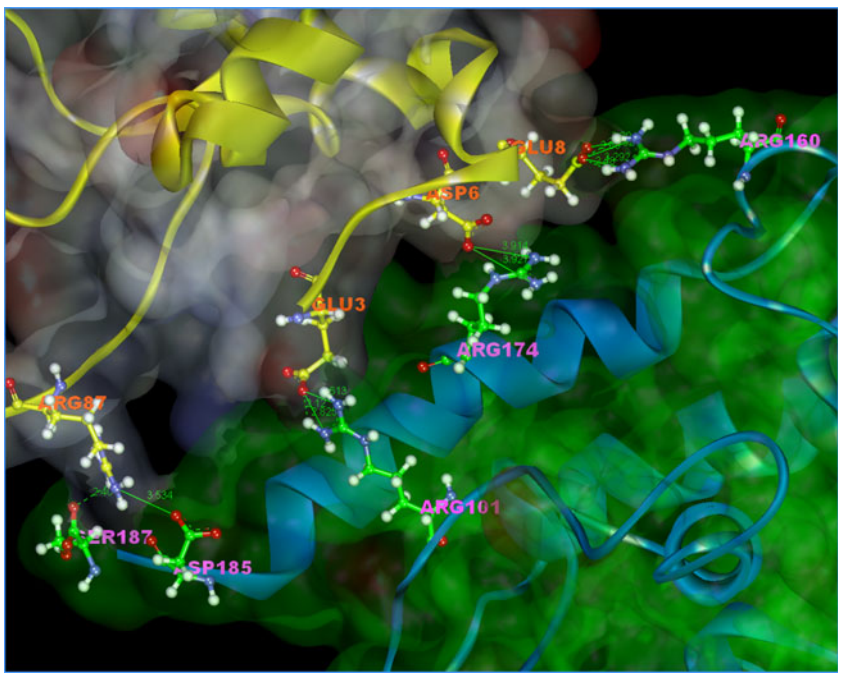


Fig. 9 Close-up of the binding site of the docked complex; amino-acid residues are represented by *balls and sticks*; the SigK protein is shown in *blue* and anti-SigK in *yellow*

region. Five bond angles deviated from their formal values. This shows that the model generated is stereochemically valid. The ProSA II server gave pairwise energy and surface energy values based on the mean force potential (a distance-based pair potential) as a function of amino-acid sequence. The results of the ProSA II analysis of the SigK protein are given in Fig. 6a. Amino-acid residues with negative ProSA energies are more reliable, and most of the amino-acid residues in the SigK protein have negative energies. The overall Z-score was −3.95 for the SigK protein. The 3D model of the SigK protein consisted of 10 helices, 20 helix–helix interactions, 15 beta turns and one gamma turn (Table 2). The nonbonded Π–cation interaction between the Tyr50 and Arg46:NH2 residues added stability to the structure.

Anti-SigK factor, a negative regulator of transcription initiation, is a transmembrane protein with 232 amino-acid residues. The membrane topology obtained from the SPLIT 4.0 server is shown in Fig. 7, and this suggests that the amino-acid residues Met1 to Ala92 are present in the

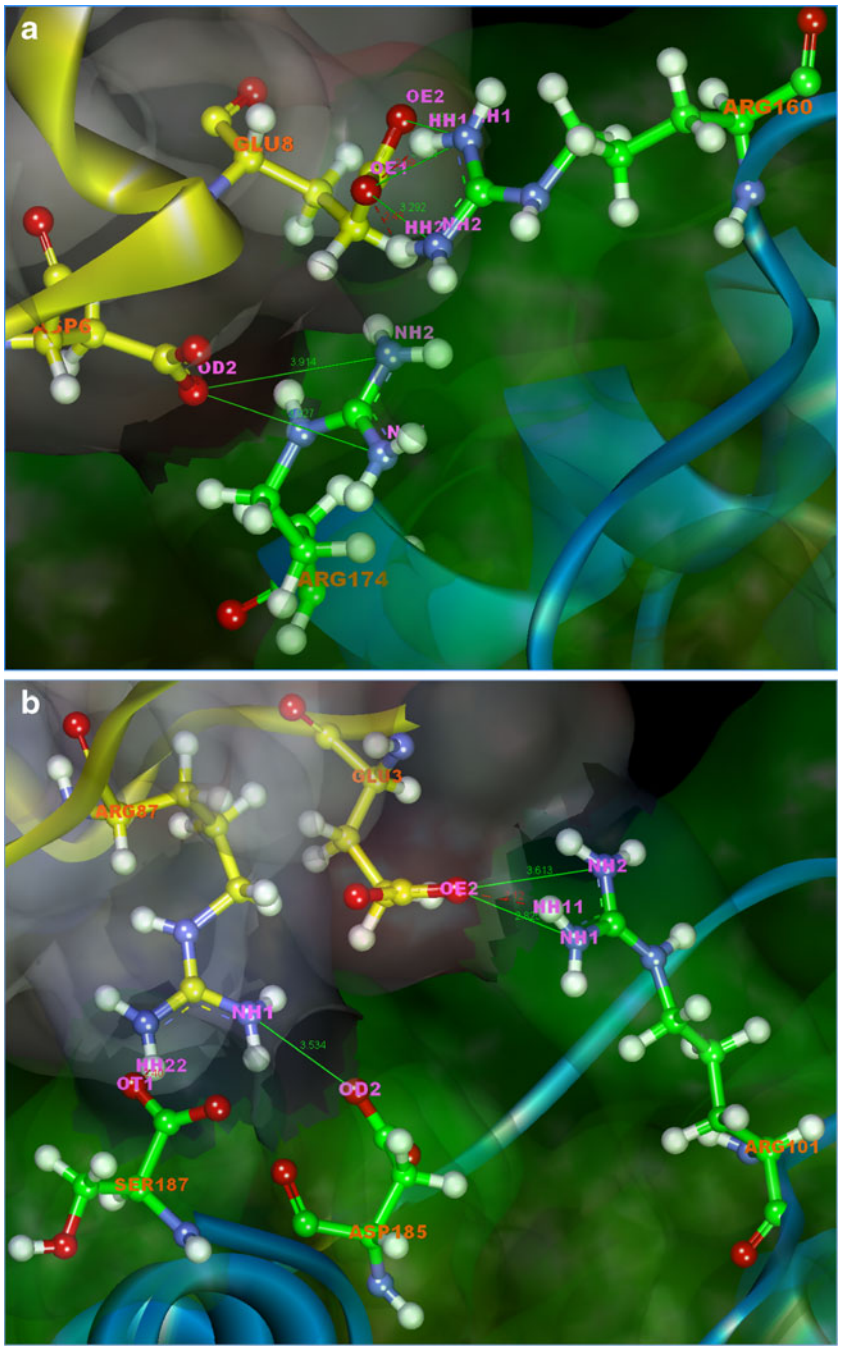


Fig. 10 **a** & **b** Hydrogen bonds in the SigK and anti-SigK docked complexes are represented by *red dashed lines*; salt bridges of 4 Å or less are represented by *green lines* that are labeled with the bridge distances

cytoplasmic region, Phe93 to Thr112 fall in the transmembrane region, and the residues Arg113 to Gln232 reside in the extracytoplasmic region. The N-terminal cytoplasmic domain of anti-SigK plays an important role in interactions with the SigK protein. Alignment between the first 92 residues of anti-SigK and the first 110 residues of the cognate anti-sigma ChrR protein of *Rhodobacter sphaeroides* (PDB code: 2Q1Z, B chain) gave a sequence identity of about 23.91%. The model with the lowest value of the Modeller objective function (445.40) was considered for further refinement. The average RMSD of the anti-SigK protein was 1.79 Å, and an RMSD plot for the MD simulation is shown in Fig. 3. The backbone dihedral angle distribution of the anti-SigK cytoplasmic domain after MD simulation, along with plot statistics, is given in Fig. 5b. The energy profile from ProSA II for the modeled anti-SigK protein is presented in Fig. 6b, and the calculated Z-score was −1.25. The predicted anti-SigK protein structure contains four helices, four helix–helix interactions, four beta turns and one gamma turn. The four helices are conserved in the group 4 anti-sigma domain of the N-terminal region, as listed in Table 2 and illustrated in Fig. 4b. The anti-SigK protein contains a single Π–cation interaction between the His4 and Met1:N residues, which add to the stability of the protein. The amino-acid residues His4, Leu9, Leu10, Leu18, Ala56, Val 58 and Ala60 in the anti-SigK protein are conserved [41]. The specific FecR iron-binding motif (TrpX3AspX2His) and the ZAS Zn^{2+}-binding motif (Cys/His X23-26HisX3CysX2Cys) are absent in the anti-SigK protein as neither iron nor zinc are required for the binding of the anti-SigK protein to SigK.

The active sites of the SigK and anti-SigK proteins were predicted using cavity prediction servers and the characteristics of the template. The cavity volumes predicted for SigK and anti-SigK were obtained from CASTp and Q-Site predictions and are listed in Table 3. The cavities 1, 2, 4, 5 and 6 predicted using Q-Site share the amino acids from Asp12 to Leu120. The sites 1, 2, 4 and 5 predicted by the CASTp server all contain Leu7 to Glu157 amino-acid residues. Cavity 6 from the CASTp server and cavity 5 from the Q-Site server contained Val41, Arg46, Tyr50, Thr54, Arg87, Ala88, Val89, and Asp90 in common, while Ser51 and Glu53 were also reported by the Q-Site server; these two cavities had nearly the same volumes of 143.5 and 142 $Å^3$, respectively. The predicted SigK protein cavities lie in the region of amino-acid residues 10–95. The active site for the crystallized template protein 2Q1Z (A and B chains) has been identified experimentally [41]. The active site contains 26 hydrogen bonds and 320 nonbonding interactions. The amino-acid region ranging from 13 to 95 was taken to be the active site in the SigK protein, based on the predicted cavities and the binding site of the template. In the anti-SigK protein, the N-terminal cytoplasmic domain was considered to be the active site for binding.

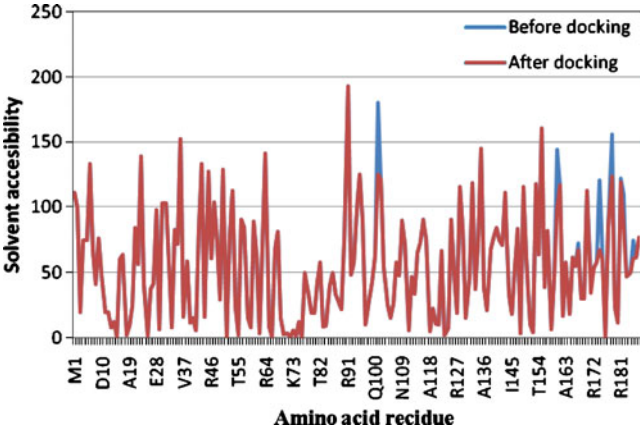


Fig. 11 Solvent-accessible surface area (SASA) in the SigK protein before and after docking

The protein–protein docking results between the SigK and anti-SigK N-terminal proteins, obtained using ZDOCK 2.3, reveal that four sets of amino acids are involved in the formation of the complex. The main interactions that orient both of the proteins into the complementary regions of the binding pockets are four hydrogen bonds and eight salt bridges, as listed in Table 4. The electrostatic potential surfaces of the docked complex, as presented in Fig. 8, show the complementary positive and negative charge distributions. The docked complex and the interactions between SigK and anti-SigK are depicted in Figs. 9 and 10. The residues Glu3, Asp6, Glu8, and Arg87 of anti-SigK bind to Arg101, Arg174, Arg160 and Ser187 of SigK, respectively. The results of an analysis of the solvent-accessible surface area (SASA) before and after the docking of the SigK and anti-SigK proteins are given in Fig. 11 and Fig. 12. These plots show that the residues Thr2, Glu3, Asp6, Phe7, Glu8, Tyr16, Arg84, and Arg87 in the anti-SigK protein and Arg101, Arg160, Arg174, Arg178, and Ser187 in SigK have lower SASA values

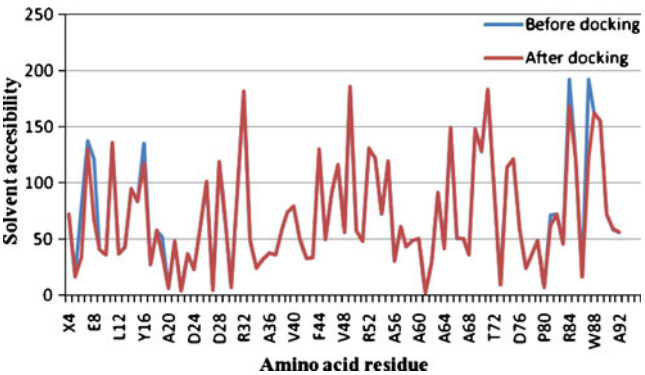


Fig. 12 Solvent-accessible surface area (SASA) in the anti-SigK protein before and after docking

after docking than before docking. All of the residues highlighted in the SASA analysis are either in close proximity to the binding site or involved in bond formation. Further studies aimed at designing new chemical entities based on these interacting residues are underway; the ultimate objective is to use these entities to inhibit the anti-SigK protein and thus facilitate the participation of the SigK protein in transcription, allowing the production of antigenic proteins in *Mycobacterium tuberculosis*.

Acknowledgments V.M. is thankful to the University Grants Commission, New Delhi, for their financial support of this work under the major research project scheme (UGC project F. No. 34-302/2008(SR), dated: 24-12-2008). V.M., B.G., B.K. and R.D. are thankful to the Principal of Nizam College for providing the facilities to carry out this work.

References

1. Digby F, Warner VM (2006) Tuberculosis chemotherapy: the influence of bacillary stress and damage response pathways on drug efficacy. Clin Microbiol Rev 19:558–570
2. Munro SA, Lewin SA, Smith HJ, EngeL ME, Fretheim A, Volmink J (2007) Patient adherence to tuberculosis treatment: a systematic review of qualitative research. PLoS Med 4:1230–1245
3. de Souza MV (2006) Promising drugs against tuberculosis. Rec Pat Anti-infect Drug Discov 1:33-44
4. Jossy VDB, Gibson SK, Elton RK, Martin JB, Rob EA (2009) New drugs against tuberculosis problems: progress and evaluation of agents in clinical development. Antimicrob Agents Chemother 53:849–862
5. OBrien RJ, Nunn PP (2001) The need for new drugs against tuberculosis. Obstacles, opportunities and next steps. Am J Respir Crit Care Med 163:1055–1058
6. Zhang Y (2005) The magic bullets and tuberculosis drug targets. Annu Rev Pharmacol Toxicol 45:529–564
7. Peter AS, Floyd ER (2007) Combating bacteria and drug resistance by inhibiting mechanisms of persistence and adaptation. Nat Chem Biol 3:549–556
8. Douglas FB, Stephen JWB (2004) The regulation of bacterial transcription initiation. Nat Rev Microbiol 2:1–9
9. Murali DB, Seyed EH (2004) The extracytoplasmic function sigma factors: role in bacterial pathogenesis. Infect Genet Evol 4:301–308
10. Dagmar H, Lenka H, Jan K (2008) Cascade of extracytoplasmic function sigma factors in *Mycobacterium tuberculosis*: identification of a σ^J-dependent promoter upstream of *sigI*. FEMS Microbiol Lett 280:120–126
11. Benjamin EB, Susan KB (2008) Signaling mechanisms for activation of extracytoplasmic function (ECF) sigma factors. Biochim Biophys Acta 1778:1930–1945
12. Kit LB, Kelly TH (1995) The role of anti-sigma factors in gene regulation. Mol Microbiol 16:397–404
13. Bottouli SS, Serge M, Arnold SK, Marcel AB (2006) Mutations in *Mycobacterium tuberculosis Rv0444c*, the gene encoding anti-SigK, explain high level expression of MPB70 and MPB83 in *Mycobacterium bovis*. Mol Microbiol 62:1251–1263
14. Danielle C, Serge M, David ALS, Harald GW, Marcel AB (2005) Reduced expression of antigenic proteins MPB70 and MPB83 in *Mycobacterium bovis* BCG strain due to a start codon mutation in *sigK*. Mol Microbiol 56:1302–1313
15. Juretic D, Zoranic L, Zucic D (2002) Basic charge clusters and predictions of membrane protein topology. J Chem Inf Comput Sci 42:620–632
16. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. J Mol Biol 215:403–410
17. Kelley LA, Sternberg MJE (2009) Protein structure prediction on the web: a case study using the Phyre server. Nat Protoc 4:363–371
18. Christian C, Jonathan DB, Geoffrey JB (2008) The Jpred 3 secondary structure prediction server. Nucleic Acids Res 36:197–201
19. 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
20. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2. Bioinformatics 23:2947–2948
21. Sali A, Pottertone L, Yuan F, Van VH, Karplus M (1995) Evaluation of comparative protein modeling by MODELLER. Proteins 23:318–326
22. Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18:14–23
23. James CP, Rosemary B, Wei W, James G, Emad T, Elizabeth V, Christophe C, Robert DS, Laxmikant K, Klaus S (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26:1781–1802
24. MacKerell AD, Feig MJ, Brooks CLIII (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J Comput Chem 25:1400–1415
25. MacKerell JAD, Bashford D, Bellott M, Dunbrack JRL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, McCD J, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher IIIWE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz KJ, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102:3586–3616
26. Giovanni S, Alan RF (2008) High temperature unfolding simulations of the TRPZ1 peptide. Biophys J 94:4444–4453
27. Feller SE, Zhang Y, Pastor RW, Brooks RW (1995) Constant pressure molecular dynamics simulations: the Langevin piston method. J Chem Phys 103:4613–4621
28. Deserno M, Holm C (1998) How to mesh up Ewald sums. I. A theoretical and numerical comparison of various particle mesh routines. J Chem Phys 109:7678–7693
29. Essmann U, Perera L, Berkowitz ML, Darden TA, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys 103:8577–8593
30. Humphrey W, Dalke A, Schulten K (1996) VMD—Visual Molecular Dynamics. J Mol Graph 14:33–38
31. Laskowsky 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
32. Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:407–441
33. Liang J, Edelsbrunner H, Woodward C (1998) Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. Protein Sci 7:1884–1897

34. Laurie AT, Jackson RM (2005) Q-SiteFinder: an energy-based method for the prediction of protein–ligand binding sites. Bioinformatics 21:1908–1916
35. Rong CLL, Zhiping W (2003) ZDOCK: an initial-stage protein docking algorithm. Proteins 52:80–87
36. Pierce B, Weng Z (2007) ZRANK: reranking protein docking predictions with an optimized energy function. Proteins 67:1078–1086
37. Accelrys Software Inc. (2007) Accelrys Discovery Studio Visualiser v 2.5.1.1967. Accelrys Software Inc., San Diego
38. DeLano WL (2002) The PyMOL molecular graphics system. Delano Scientific, Palo Alto (see http://www.pymol.org)
39. Voelz VA, Shell MS, Dill KA (2009) Predicting peptide structures in native proteins from physical simulations of fragments. PLoS Comput Biol 5(2):e1000281. doi:10.1371/journal.pcbi.1000281
40. Ravi DS, Andrew ML, Pradeep KS, Rajnee SJ (2009) High temperature unfolding of *Bacillus anthracis* amidase-03 by molecular dynamics simulations. Bioinformation 10:430–434
41. Elizabeth AC, Roger G, Jennifer RA, Sheng W, Lionel L, Kalyan D, Heidi JS, Timothy JD, Seth AD (2007) A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria. Mol Cell 27:793–805