Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2014, 6(12):481-492

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Computational mapping reveals novel druggable sites in signal recognition particle of *E. coli*

Malini Veerasamy

Department of Immunology, National Institute for Research in Tuberculosis, Indian Council of Medical Research, Mayor V. R, Ramanathan Rd, Chetput, Chennai, India

---

ABSTRACT

The signal recognition particle protein is a GTPase and is essential for co-translational translocation of inner-membrane proteins in E. coli. Since it has a human homologue and highly conserved active site selective inhibition of this protein is difficult to achieve. With the recent computational fragment based method, FTMAP, it is possible to accurately identify alternative drug binding sites on protein surfaces that can be used for screening lead compounds. In this study, we have identified the allosteric druggable sites of signal recognition particle of E. coli using FTMAP. Initial amino acid sequence analysis showed E. coli protein to be 32% identical to human counterpart that includes the four characteristic GTPase motifs and the 'ALLEADV' motif. Mapping of crystal structures of E. coli protein revealed six pockets of which two were orthosteric and four were non-orthosteric sites distributed in the inter-domain, protein-protein and solvent interface regions and structurally conserved in T. aquaticus protein. Since these pockets are formed by residues poorly conserved in human protein and present in allosteric space of protein, selective allosteric inhibition is possible, thus qualifying as potential drug target. This study is significant as it reiterates the druggability of prokaryotic essential proteins with human homologues.

Keywords: signal recognition particle, FTMAP, druggability, allosteric sites, protein-protein interaction, drug targets

---

INTRODUCTION

Indiscriminate use of antibiotics has resulted in the spread of multidrug resistant bacteria complicating disease control [1]. Ironically, with the increase in antibiotic resistance, the choice of antibiotics to treat infections is diminishing. This situation necessitates alternative approaches to drug discovery [2]. One strategy is to include essential proteins of prokaryotes that have human homologue [3,4]. But, achieving selectivity without causing side effects is difficult in these proteins. For example, selective inhibition could not be achieved for conserved proteins such as kinases and GPCRs which have been major therapeutic targets for several decades [5]. As an alternative approach, inhibition through alternative binding sites (also called non-orthosteric binding/allosteric sites) have been proposed for these proteins and is now a recent focus in the field of drug discovery [6]. The highly successful allosteric inhibitor, imatinib (Gleevac; Novartis) for kinases and the recent lead molecules designed for G protein coupled receptors [7] exemplify this strategy. However, identification of non-orthosteric binding sites is a time consuming and expensive procedure and is a prerequisite for discovering lead molecules. Two experimental methods are widely used for assessing druggability of protein targets (i) NMR (Nuclear Magnetic Resonance) based method where N^{15}-labeled protein is screened against a library of small probe compounds [8] (ii) multiple-solvent crystal structure (MSCS), an X-ray based method where protein structures are resolved in the presence of different

organic probe molecules in aqueous solutions and screened for probe binding [9]. The MSCS methods were further translated into computational methods which were used for predicting the druggable pockets [10–13]. Among these the FTMAP approach based on fast Fourier transform correlation was able to reproduce MSCS experiments with accuracy (16).

In light of the above developments, we chose to explore the druggability of an important and ubiquitous protein, signal recognition particle (SRP) [16], of *E. coli*. SRP is essential in *E. coli* for viability [17] and is required for the co-translational translocation of inner membrane proteins in *E. coli* [19,20] thus providing rational for exploring it as a drug target.

SRP is a ribonucleoprotein and is conserved across all kingdom of life [21]. In *E .coli* it consists of a single protein (48 KDa) (also called Fifty four homologue, Ffh) complexed with 4.5S RNA [22]. In eukaryotes, it is more complex consisting of a protein (54 KDa and a 7S RNA and six subunits (24). During the co-translational translocation SRP binds to the signal peptide of the nascent chain that emerges from the translating ribosome [24], subsequently it forms a heterodimer with its cognate receptor SR [25], following which it transports the signal peptide-nascent chain-ribosome complexes to the translocation apparatus at the membrane, finally reciprocal GTPase hydrolysis takes place between SRP-SR leading to release of the signal-peptide-RNC cargo to the translocon and dissociation of SRP and its receptor in their GDP bound forms [26]. Both the SRP protein and SR are GTPases and structural homologues [27,28].

SRP is a three domain protein consisting of an N terminal domain and the GTPase G domain together forming the 'NG' domains [29] and a C-terminal methionine-rich 'M' domain, that binds to signal peptide and 4.5S RNA [30,31] (Fig 1A, M domain and 4.5S RNA not shown). Another region unique for SRP family of GTPase is the IBD (Insertion Box Domain) in G domain. The GTPase catalytic centre in G domain, is formed by four characteristic GTPase motifs GTGKTTT (G1), DTRFAG (G2), DTAGR (G3) and TKVD (G4) which bind to different regions of the GTP molecule (Fig 1A)[29]. The catalytic face of G domain along with highly conserved 'ALLEADV' motif of N-domain and the 'DARGG' motif at N/G interface are major SRP-SR interface forming regions of the heterodimer (Fig 1B) [32]. The heterodimerization is a crucial GTPase priming step for SRP and mutations of SRP-SR interface residues are deleterious to protein function [33]. Surprisingly, the mutation of residues of 'DARGG' motif is lethal to *E .coli* the reason for which is attributed to inability of SRP mutant protein to form heterodimers and not due to loss of GTPase activity. This biochemical evidence along with structural data [34] established that the N/G interface is flexible and provides freedom of movement for N-domain during heterodimer formation. Thus, these evidences indicate that protein undergoes allosteric modulation during SRP-SR dimerization, the inhibition of which is deleterious to the survival of bacteria thus providing strong reasons for identifying allosteric binding sites in this protein.

Despite the availability of the above biochemical, mutational studies and crystal structures of *E. coli* and *T. aquaitcus* SRP, to our knowledge there is no report of specific inhibitor for this protein. The only known ligands are the GTP or non–hydrolysable GTP analogues such as GMPPCP and GMPPNP that bind to the known orthosteric sites and cannot be used for selective inhibition. Thus, identifying allosteric sites would not only help in screening lead molecules but in the process also discover several SRP specific inhibitors. In this work, we used the FTMAP algorithm to identify the non-orthosteric druggable sites in crystal structure of *E. coli*, which can be exploited in the structure-based design of allosteric ligands.

EXPERIMENTAL SECTION

The X-ray structure of the SRP protein of *E. coli* and T. *aquaticus* were taken from Protein Data Bank and were mapped using the FTMAP algorithm from the website (http://ftmap.bu.edu/param). This method consists of four steps as follows [13].

(i) The FTMAP algorithm scans the surface of SRP with 16 different small organic probe molecules (ethanol, isopropanol, isobutanol, acetone, acetaldehyde, dimethyl ether, cyclohexane, ethane, acetonitrile, urea, methylamine, phenol, benzaldehyde, benzene, acetamide and N,N-dimethylformamide) using the fast Fourier transform correlation. These probes have varying hydrophobicity and hydrogen bonding capacity. For each probe billions of different binding conformations are tried and the energy calculated for each. From these the best 2000 docked

positions are chosen for energy minimization and clustered. Of these the lowest six are taken as the hotspots based on the binding energy.
(ii) The six low-energy clusters of different probe types are once again clustered, which are called consensus sites (CS).
(iii) The CSs are ranked by the number of probes clustered in it and the one having the maximum number of probes is designated the highest rank and also considered the most druggable site. In this work, the 10 largest CSs have been analyzed – namely CS1 (largest CS) through CS10 (smallest CS). The FTMAP server produces three files as output: (i) PDB co-ordinates of protein structure with their corresponding probes. (ii) statistics on non-bonded interactions made by probes to specific residues of the protein (iii) statistics of hydrogen-bonded interactions made by probes with each residue of the protein

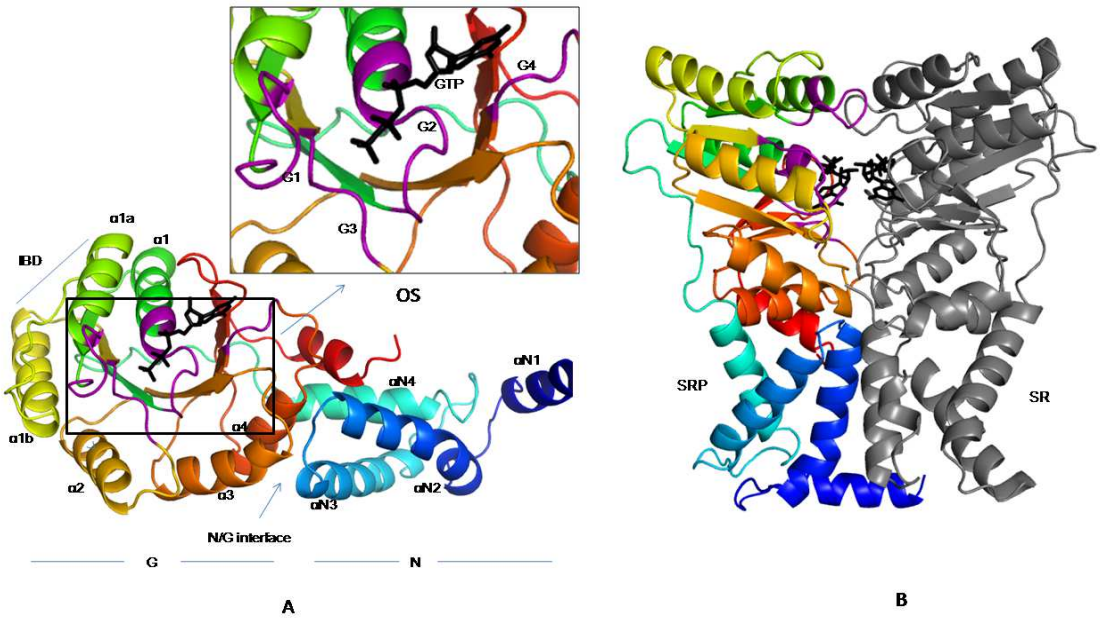

Figure 1 Cartoon diagram of SRP NG domain (A) Schematics of SRP NG domain (cartoon, rainbow) with GTP (black sticks) bound in the orthosteric site (OS). The N domain (blue shades) and G domain (green to red) are shown. The four conserved GTPase motifs G1, G2, G3 and G4 characteristic of SRP is highlighted as purple colour (box). Inset shows the orthosteric site (OS) with bound GTP and the four GTPase motifs G1 (GTGKTTT), G2 (DTFRAG), G3 (DTAGR) and G4 (TKVD) motifs. The IBD region (in G domain is unique for SRP family of GTPases and carries the G2 motif. '*' indicates the 'ALLEADV' motif. The α4 helix between the N and G domain forms the N/G interface (B) Cartoon representation of SRP (rainbow)/SR (grey cartoon) heterodimer complex. The two bound GMPPCP molecules (black sticks) is present in the SRP/SR interface and the four GTPase motifs of SRP (purple regions) are shown

RESULTS AND DISCUSSION

Sequence analysis of *E. coli* and *H. sapiens* SRP

The amino acid sequence (full length) of signal recognition particle protein (SRP) of *E. coli* and *H. sapiens* were compared (Fig 2) using Jalview software [35] to identify the identical and non identical regions of the protein. *E. coli* SRP is only 32% identical to SRP of *H. sapiens.* Out of the four characteristic GTPase motifs the residues of G1, G3 and G4 motif were identical. The residues of G2 motif which bind to the β-phosphate of GTP were poorly conserved. In addition, the residues of characteristic 'ALLEADV' motif were also identical but the residues of 'DARGG' were poorly conserved. Although the conserved nature of SRP is well known [21] we did this analysis to highlight the less conserved regions in *E. coli* which is speculated to form the non-OS druggable sites in the protein.

Mapping non-orthosteric sites in *E. coli SRP*

The crystal structure of *E. coli* SRP (PDBID 2XXA, chain A) (Table 1) was taken from Protein Data Bank (PDB) and used for mapping. Results were viewed using PyMOL [36]. In 2XXA, SRP (A chain) protein is co-crystallized with its receptor (B chain) and 4.5S RNA. The GMPPCP molecules are bound to their respective orthosteric site (OS). All the three domains N, G, and M domain are present in SRP. The mapping results of only the NG domain (amino acid residues 1- 296) of SRP (2XXA, chain A) are discussed in this paper and are summarized in Fig 3A

which shows the global distribution of the 6 predicted consensus sites (CS). In structure 2XXA, two neighbouring CSs overlap with fragments of the orthosteric ligand, GMPPCP shown as black stick (Fig 3A) and four are distributed in sub-regions of inter-domain interface, SRP-SR interface and between the helixes of N domain (Fig 3B). For 2XXA, the OS CSs are ranked 3rd and 6th (Fig 3A). The top two highest ranking non-orthosteric consensus sites were found in the SRP-SR interface in G-domain next to the catalytic centre and in the N-domain (Fig 3B). The CS corresponding to non-orthosteric sites were designated as S1, S2, S3 and S4 in the order of their CS ranking for future analysis; wherein, S1 corresponds to CS1, S2 to CS2, S3 to CS4 and S4 to CS5.

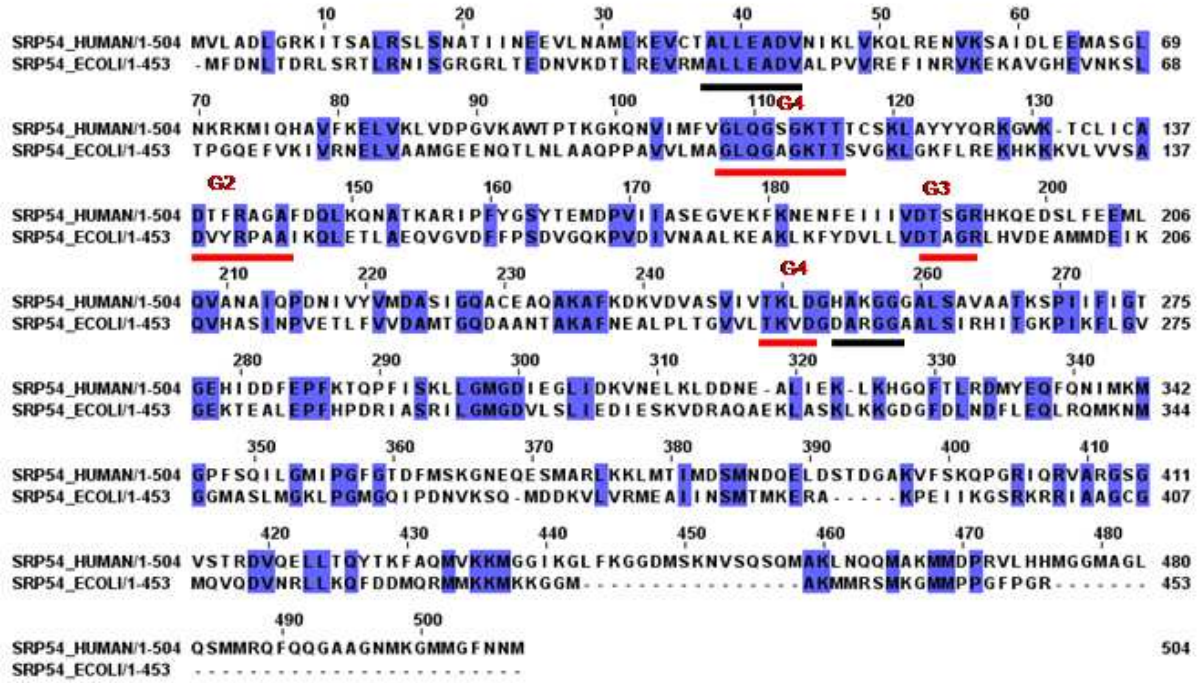

Figure 2 Amino acid sequence alignment of SRP of *Homo sapiens* and *Escherichia coli*. Alignment of amino acid sequences of *H. sapiens* and *E. coli* showing identical residues highlighted as blue. The four characteristic GTPase motifs G1 (GTGKTTT), G2 (DTFRAG), G3 (DTAGR) and G4 (TKVD) are underlined red. The 'ALLEADV' motif and 'DRAGG' motifs are underlined black. The poorly conserved residues (not highlighted) are characteristic of the family of organism they represent. Picture was created using Jalview

Structural conservation of non-orthosteric sites in *E. coli* SRP and *T. aquaticus SRP*

In order to verify whether the non-orthosteric sites were conserved in *E. coli* SRP we mapped the crystal structure of *T. aquaticus* SRP which is 45% identical to *E. coli* SRP. There are 18 structures of *T. aquaticus* SRP available in PDB database out of which four structures 1LS1, 2CO3, 1JPN and 1RJ9 representing four different conformers (Table1) were mapped (see Supplement Fig S1 and Table S1 for mapping results). The CS of each of the *T. aquaticus* structure was manually compared with that of *E. coli* SRP and the sites commonly occurring in structures of both the organism were identified. Fig 4 illustrates the general distribution of consensus sites in *T. aquaitcus* with the commonly occurring sites discriminated as red spheres. The non-OS sites conserved in both *E. coli* and *T. aquaticus* and their respective cluster population are described in Table 2. The S3 site at the N/G interface was the most frequently occurring sites followed by S1 site both within the top five rankings of CS (Table 2). Also S1 typically can accommodate all the 16 different probe types in all the conformations. In contrast the S4 that occurs in G domain near the Insertion Box Domain (IBD) region was found only in 1LS1 structure of *T. aquaticus*.

Residues of non-orthosteric sites of *E. coli* SRP are not conserved in *H. sapiens*

It is important to know whether the residues that line non-OS sites of 2XXA are not conserved in *H. sapiens* to qualify as a drug target. Although the sequence analysis (Fig 2) sufficiently proves the variability, deeper verification by directly comparing the residues of the non-OS sites with equivalent residues of *H. sapiens* SRP would provide stronger evidence. The interacting residues of the four non-OS sites of 2XXA are summarized in Table 3. The key residues are those that form maximum non-bonded interactions with the probe. Since the structure of *H. sapiens* was not available we resorted to comparing residues of non-OS sites from 2XXA with equivalent

residues in the sequence of *H. sapiens* SRP and also *T. aquaticus* SRP (Fig 5). The residues within 4 Aº distance from the probe clusters in each non-OS of 2XXA were taken for comparison. The residues of non-OS sites S3 and S4 of 2XXA are poorly conserved in human SRP. The key interacting residues of non-OS sites S1 and S2 were conserved in human SRP but the surrounding residues are poorly conserved.

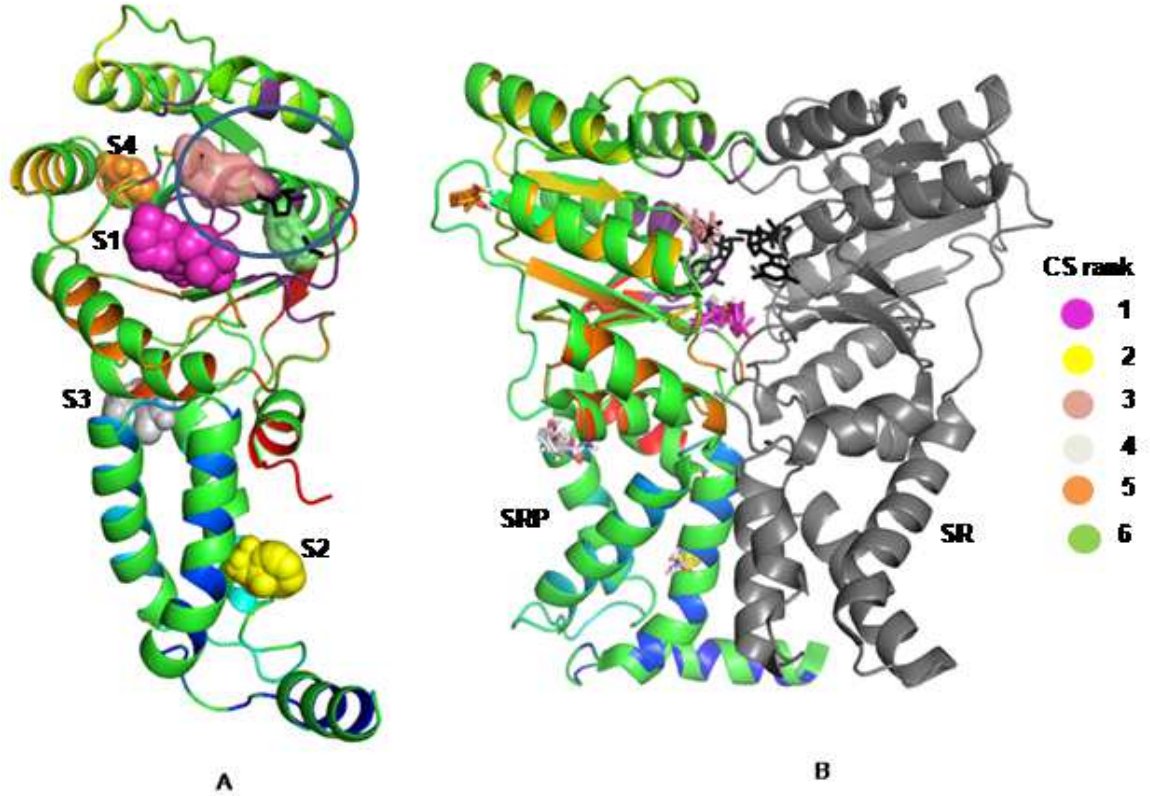

Figure 3 Mapping results and general distribution of consensus sites of *E. coli* SRP. (A) Schematic of SRP NG domain showing the general location of the predicted consensus sites (CS). The results of mapping are superimposed on the NG domain of 2XXA chain A (rainbow colour). The OS substrate (GMPPCP) is represented as stick (black). The protein is shown as green cartoon and 6 predicted consensus sites of 2XXA are depicted as probe clusters. Non-OS clusters shown as spheres and OS clusters represented as surface; colors reflect the rank (numbers of probe clusters per site); see legend for rank. The CS ranked 1 (pink), 2 (yellow), 4 (white) and 5 (orange) are non-orthosteric sites designated as S1, S2, S3 and S4 respectively (B) General location of the CS in SRP with respect to heterodimer conformation. CS are shown as probe clusters represented as sticks; colors reflect the rank (numbers of probe clusters per site) of the CS in SRP (rainbow cartoon); see legend for rank. CS1 (pink stick) is located below the active site with bound GMPPCP molecules in the SRP-SR interface. Images were created using PyMOL

Properties of non-OS sites of *E. coli* SRP

The residues of the four non-OS site of *E. coli* SRP were further analyzed to understand the nature of the cavity. The details of the four non-OS sites S1, S2, S3 and S4 of 2XXA are described below.

(i) S1: SRP- SR interface- P loop

Site 1 lies in the SRP-SR interface sharing residues with the P-loop, G3 and G4 motif of the orthosteric site (Fig 6A,B, Table 3). This pocket is smaller than the orthosteric site and the key interacting residues are L108, Q109 of 'P-loop' (Fig 6B). The residues that form the hydrogen bonds with the probes are Q109 and G110 which are also on P-loop. The K249 of this cavity is a highly conserved residue of the G4 motif involved in binding of the guanine base of GTP. Except for V221 in the β3 sheet of G domain all the residues of this cavity lie in the flexible loop region. The probes in this cavity appear scattered at the mouth of the cavity (Fig 6B).

(ii) S2: N-domain - αN4- αN2

Site 2 is formed within the N domain and lies between helix 2 and helix 4 (Fig *7*A) of SRP. This void lies near the 'ALLEADV' motif and does not occur in any of the 4 SRP structures of *T. aquaticus*. The key non-bonded residue is T30 which is conserved in both *E .coli* and *T. aquaticus*. The key hydrogen bonding residues are R14 and G71.

Unlike S1, in this cavity the probes are aligned to the centre and all the 16 types of probes are incorporated (Fig 7B). Although present near 'ALLEADV' motif it is not present in the interface region and is solvent exposed (Fig 3B).

(iii) S3: N/G domain interface - α4 helix

S3 is found in the interface of N and G domain, in the hydrophobic cavity adjacent to the highly conserved 'DARGG' (Fig 8A). Within this cavity, the probes tend to bind to the wall of the mouth (Fig 8B). The residues of this cavity are contributed by the αN4 of N domain and the short α-helix (α4) in the N/G domain interface. Both the key non-bonded and hydrogen bonding residues are A85 and H264 and is conserved in both *E. coli* and *T. aquaticus*. This void is largely populated by hydrophobic residues confirming the hydrophobic environment of the pocket and can bind to 13 of the 16 types of probes. This site is present in all the conformations of *T. aquaticus*.

(iv) S4: G –domain – IBD domain

S4 (Fig 9) lies on the exterior surface of the G domain facing the solvent. The residues forming this cavity are located within the IBD domain of the G domain (Fig 9A). Unlike other sites which are lined by conserved residues this region is unique for *E. coli SRP* and none of the residues are conserved in *T. aquaticus* and thus can be used for designing drugs specific for *E. coli*. The key hydrogen bonding residues are N213 and P100 both found in the loop region. The key non-bonded residues are the N213 and P99 also in the loop. The cavity can accommodate six probe types and is centered against the wall of the shallow cavity.

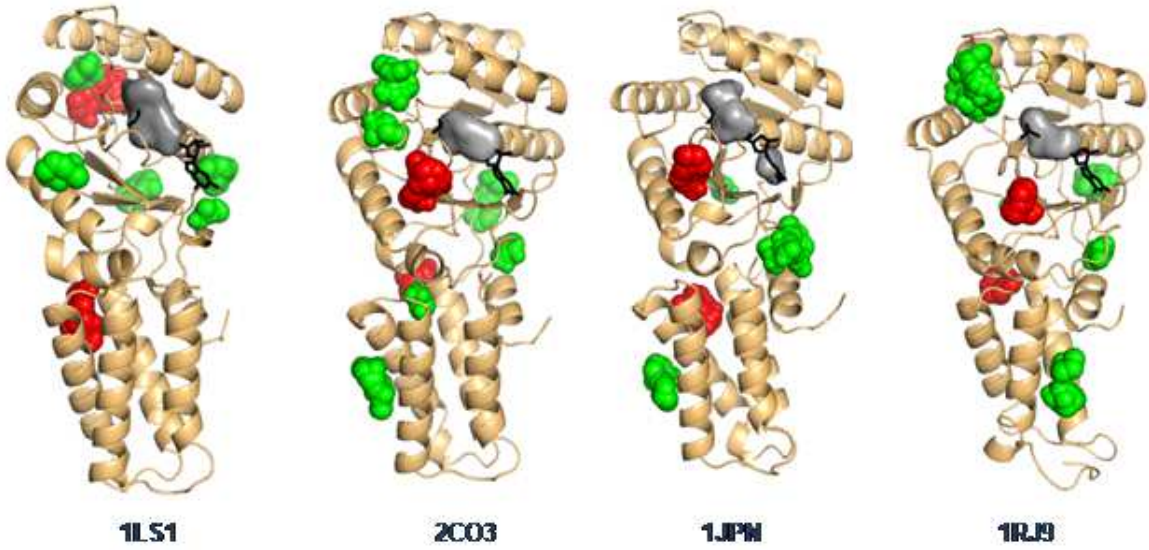

Figure 4 Mapping results of four different conformers of *T. aquaticus* SRP. Schematics of SRP NG domain of *T. aquaticus* showing the general distribution of top 10 CSs of crystal structures 1LS1, 2CO3, 1JPN and 1RJ9. The protein is shown as orange cartoon, the orthosteric site (OS) substrate GMPPCP is represented as stick (black). The predicted CSs are shown as probe clusters (represented as spheres and surface); red spheres indicate non-OS CS conserved in 2XXA; CS overlapping with OS are shown as grey surface. Images were created using PyMOL

	S1	S2	S3	S4
*Escherichia coli**	LQGALVDMTK	RTLEVGQFVM	RAELRHG	QPPAKNV
Homo sapiens	LQGSHVDSIK	RMLEVRKIQM	RLPLVAK	GKQNKQD
Thermus aquaticus	LQGSLVDMTK	RTLEIAEILM	RAGLRHG	KDRNRGD

Figure 5 Conservation of residues of non-OS sites of *E. coli* SRP. The residues of non-orthosteric sites of *E. coli* SRP (PDB ID 2XXA) aligned with equivalent residues of *T. aquaticus* and *H. sapiens* SRP sequence. Residues conserved in *E. coli* and *H. sapiens* SRP are shown in red

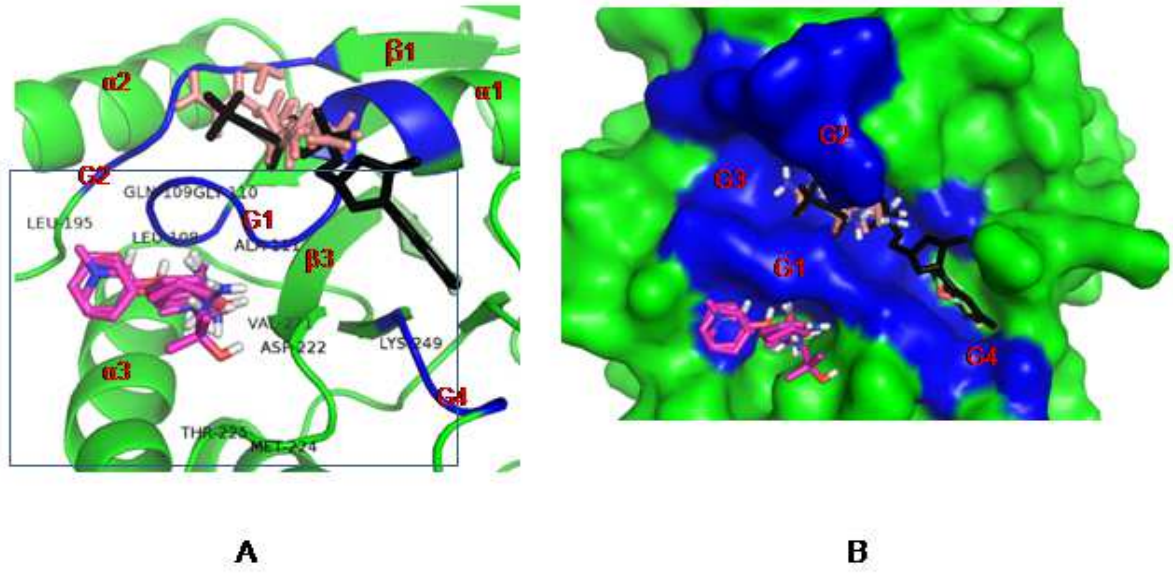

Figure 6 Non-orthosteric site S1: SRP-SR interface – P - loop. A close-up view of the probes in the non-orthosteric (OS) site S1 identified in mapping of 2XXA. Protein backbone is shown as a green cartoon. GMPPCP ligand is shown as black sticks. S1 (CS1) probes are shown as pink sticks. The orthosteric CS3 and CS6 are shown as salmon and green probe sticks overlapping the GMPPCP ligand. The interacting residues of S1 are labeled on the protein backbone. The G1, G2, G3 and G4 motifs are highlighted as blue. The S1 site shares residues with G1 motif and G4 motif (coloured blue) (B) Close up view of S1 with protein shown as green surface and the probes (pink sticks) scattered at the mouth of cavity

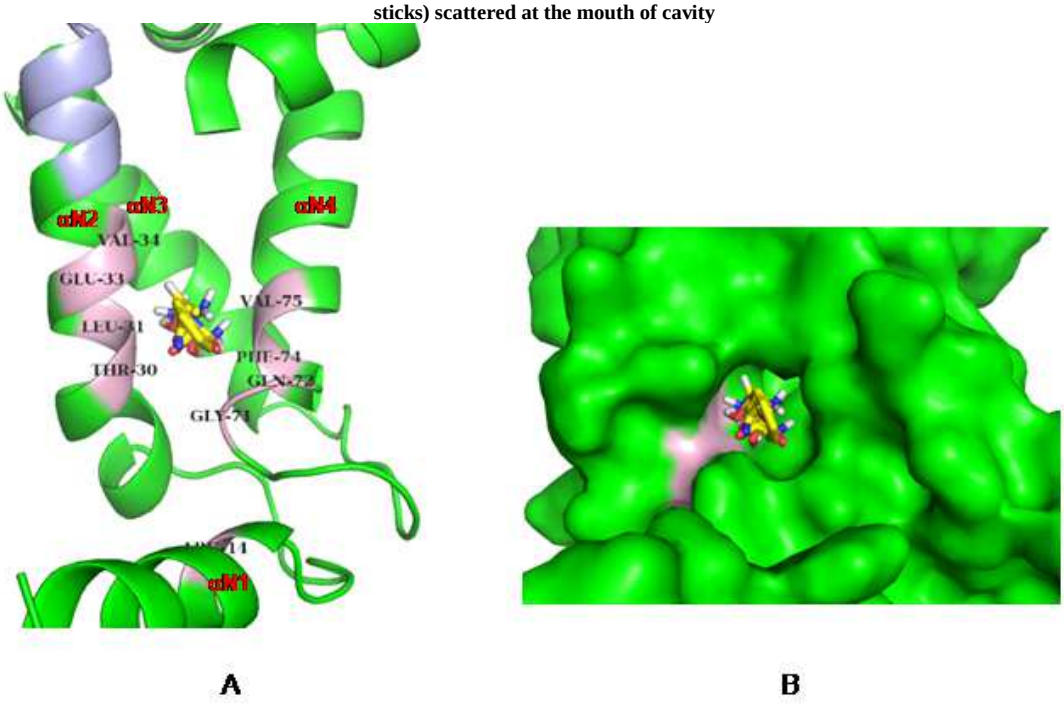

Figure 7 Non-orthosteric Site S2: N-domain helix 2 and 3. A close-up view of the probes in the non-orthosteric (OS) site S2, identified by mapping of the 2XXA structure of SRP. Protein backbone is shown as green cartoon. Probes are shown as yellow sticks. The interacting residues and the secondary structure of the protein are labeled on the protein backbone and highlighted as light pink (B) A close up view of probes in S2. Protein backbone shown as green surface with the probes (yellow sticks) centered at the cavity and key residue T30 coloured light pink

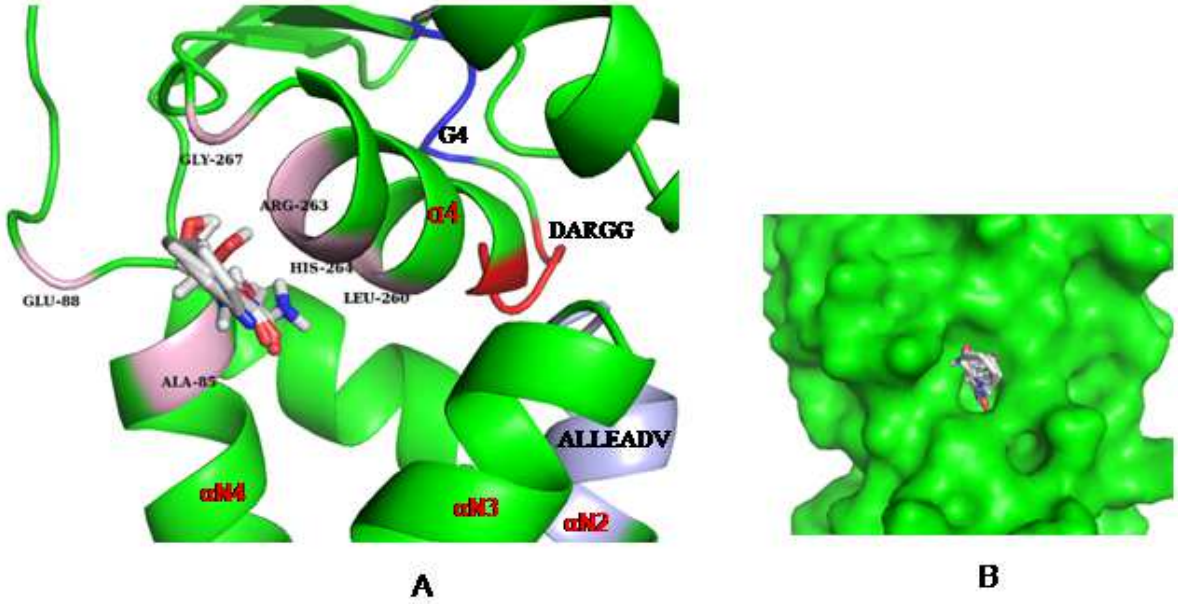

Figure 8 Non-orthosteric Site S3: N/G interface (A) A close-up view of the probes in the non-orthosteric (OS) S3 identified in mapping of the 2XXA structure of SRP. Protein backbone is shown as a green cartoon. Probes are shown as white sticks. The residues interacting with probe is labeled on the protein backbone (B) A close up view of S3 with protein backbone shown as green surface and the probes (white sticks) with only the side chains occupying the centre of cavity

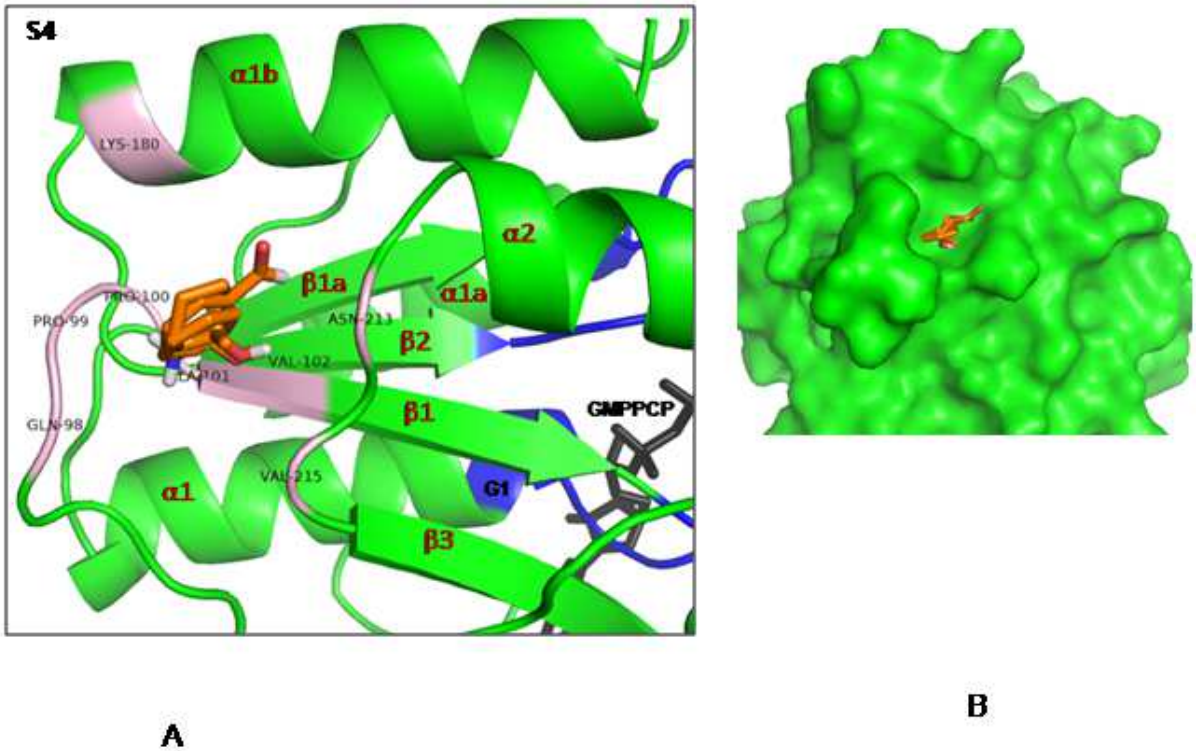

Figure 9 Non-orthosteric Site S4: G domain- α2 - α1b (A) A close-up view of the probes in the non-orthosteric (OS) S4 identified in mapping of the 2XXA structure of SRP. Protein backbone is shown as a green cartoon ribbon. Probes are shown as orange sticks. The interacting residues are labeled on the protein backbone (B) A close up view of S2 with protein shown as green surface and probes represented as orange sticks

Table S1 Summary of mapping results of four representative conformers of *T. aquaticus*

Table 1 List of PDB structures used in this study

PDB ID_chain	ligands*	Description of SRP† protein conformers	Exp. Method	Resolution	Reference
Escherichia coli					
2XXA_A	GMPPCP	Heterodimer, SRP co-crystallized with SR and in complex with 4.5S RNA; GMPPCP is bound to active site	X-ray diffraction	3.94	[42]
Thermus aquaticus					
1LS1_A	-	Heterodimer, SRP co-crystallized with SR and in complex with 4.5S RNA; GMPPCP is bound to active site	X-ray diffraction	1.1	[34]
2C03_A	ED,GDP	Apo monomer, SRP with no substrates bound to active site	X-ray diffraction	1.24	[43]
1JPN_A	GMPPNP	Monomer, SRP with GMPPNP bound to active site	X-ray diffraction	1.97	[44]
1RJ9_A	GMPPCP	Heterodimer, SRP co-crystallized with SR; GMPPCP bound to the active site	X-ray diffraction	1.9	[32]

Table 2 Summary of conserved non-orthosteric sites found in SRP of *E. coli* and *T. aquaticus*

Non-OS in *E. coli*	Location in structure	consensus site rank (cluster population)				
		2XXA A	1LS1 A	2C03 A	1JPN A	1RJ9 A
S1	G domain	1(16)	0	2(13)	2(13)	2(13)
S2	N domain	2(14)	0	0	0	0
S3	N/G interface	3(13)	2(13)	5(8)	6(8)	10(2)
S4	G domain	4(6)	3(13)	0	0	0

Table 3 Interacting residues of the 4 non-orthosteric binding sites of SRP of *E. coli.* Key residues are shown in bold

Non-OS in *E. coli*	Interacting residues		secondary structure	location
	E. coli	*T. aquaticus*		
1	L108 Q109 G110 A111 L195 V221 D222 M224 T225 K249	L106 Q107 G108 S109 L218 D219 M220 T221 G222	P - loop	G domain. SRP-SR interface
2	R14 T30 L31 E33 V34 G71 Q72 F74 V75 M297 V300	-	αN2 – αN4	N domain
3	R54 A85 E88 L260 R263 H264 G267	A85 L86 L87 R260 H261 G262	α4 - αN3	N/G domain interface
4	Q98 P99 P100 A101 K180 N213 V215	D97 R98 N99 V172 E173 A174 R175 A180 R181 D182 I183	β1- α1b	G domain, SRP-solvent interface

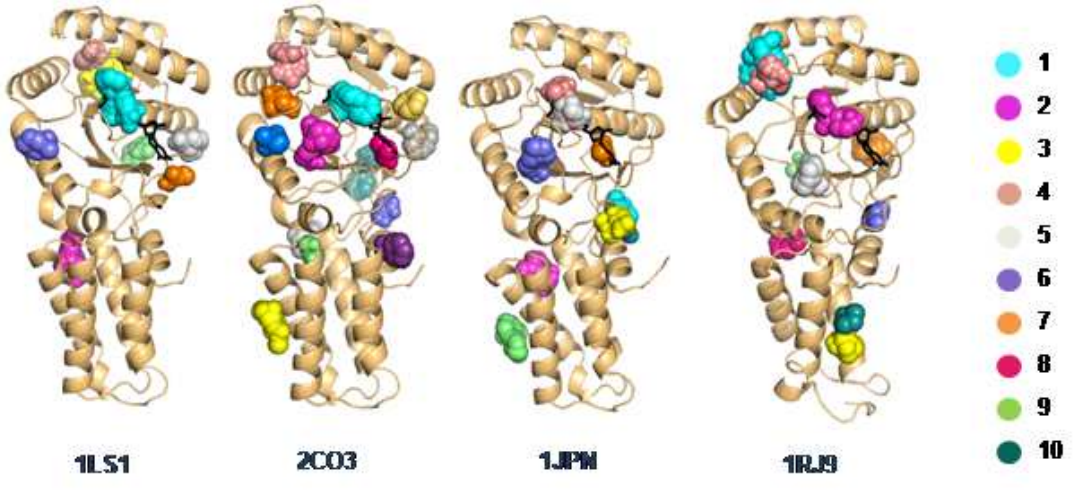

Figure S1 Mapping results and general distribution of consensus sites in representative SRP structures of *T. aquaticus*. (A) Schematic of SRP NG domain showing the general location of the predicted consensus sites (CS) in 1LS1 (chain A), 2CO3 (chain A), 1JPN (chain A) and 1RJ9 (chain A). The OS substrate (GMPPCP) is represented as stick (black). The protein is shown as light orange cartoon and top 10 predicted consensus sites of *T. aquaticus* are depicted as probe clusters represented as spheres ; colors reflect the rank (numbers of probe clusters per site); see legend for rank.. Images were created using PyMOL

Table S2 Summary of mapping results of four representative conformers of *T. aquaticus*

CS Rank	1LS1	2CO3	1JPN	1RJ9
1	26	13	16	26
2	22	13	16	14
3	20	12	15	12
4	10	11	13	11
5	7	8	12	8
6	2	6	8	7
7	2	4	5	7
8	2	4	4	3
9	2	4	3	2
10	0	3	2	2

CONCLUSION

Finding new drug targets and targeting the non-orthosteric drug binding sites of protein and enzymes are the current focus in the field of drug discovery [37]. The non-OS binding sites provide scope for specific targeting by binding to less conserved regions of SRP. The availability of experimental structures of SRPs in free and receptor bound forms has provided opportunity for structure based drug designing using computational methods. In this work, we report the application of a fragment-based algorithm, FTMAP, to map the surface of the *E. coli* SRP for druggable sites distinct from the orthosteric site (OS).

We found four non-OS sites binding sites in *E. coli* which were structurally conserved in *T. aquaticus*, a distant homologue of *E. coli*. Since the residues forming these sites are not conserved in human SRP these sites can be valuable for identifying small molecule inhibitors selective against *E. coli*. Among the four sites, Site 1 is found next to the OS site and is of relatively rigid nature as it is found in four different conformers of SRP. Since all sixteen types of probe types bind to this site, this is a potential druggable site. Also, as this site shares residues with P-loop of G1 motif it has direct influence on the orthosteric site and also on the heterodimer formation. Therefore small molecules binding to this site would compete for the active site residues and prevent the GTP binding or can act as potential protein-protein interaction inhibitors of heterodimer formation.

Unlike the stable S1 site, the S2 non-orthosteric site of N-domain is transient and suggests this pocket may form less frequently and may be associated with particular state of SRP. The dynamic nature of N-domain supports this transient appearance [34]. The role of N-domain in *E. coli* for the efficient signal sequence binding and translocation has been established by both biochemical and genetic analysis [19,38]. Therefore, binding of the small molecules in this site would affect protein translocation or may freeze SRP in one particular conformation preventing the recycling of this molecule necessary for the co-translational translocation process.

S3 represents pocket formed at the N/G domain interface, in the hydrophobic core of the SRP molecule and adjacent to 'DARGG' motif. Finding a binding site in this region was surprising as previous reports confirm that this interface forms a tightly packed environment [33]. Since this site is found in all the crystal structures we believe that this site although being small is a potential drug binding site. Occupying the junctions of two domains, it is possible that small molecules occupying these locations would stabilize inter-domain packing interactions and thus hinder conformational flexibility. This impact is supported by previously reported experimental evidences, where mutation of the two glycine residues of 'DARGG' motif are lethal to *E. coli* [33].

The residues of the S4 are the most poorly conserved among the four sites. Although there is no information available connecting the significance of residues towards the function, we presume that alterations in this region may indirectly affect the movement of 'P loop' (G1 motif) towards the orthosteric site and GTP binding as the residues of this site are on the same beta sheet structure as the origin of 'P-loop'. Thus binding of small molecule inhibitors may act as allosteric inhibitors.

Since all these sites are located in allosteric space of the protein which explains the discrepancy in the occurrence of the consensus sites within the structures of *T. aquaticus* (Table 2), these sites may be used for the selective allosteric modulation [39]. Also, as these non-OS sites were predicted from a single crystal structure it is logical to presume that using new generation computational methods such as molecular dynamics [40] and accelerated molecular dynamics [41] would reveal more sites that can be used for species specific inhibition.

This study is a starting point of the virtual screening of existing drug-like compound libraries for ligands that can bind to this pocket with high affinity. Also, as an alternative approach these probes can be developed and extended (fragment evolution and fragment linking) to design larger molecules. Although SRP is the target protein to begin such screening studies it is expected that additional druggable sites are present on SR (FtsY), which is a homologue to this protein. This study can also be extended to other essential proteins of prokaryotes with human homologues.

Acknowledgements
MV gratefully acknowledges the financial assistance provided by the Indian Council of Medical Research (ICMR) for Senior Research Fellowship.

REFERENCES

[1] P. Nordmann, T. Naas, N. Fortineau, L. Poirel, *Curr. Opin. Microbiol.* **2007**, 10, 436–40.
[2] D. J. Payne, M. N. Gwynn, D. J. Holmes, M. Rosenberg, *Methods Mol. Biol.* **2004**, 266, 231–259.
[3] T. P. Soares da Costa, W. Tieu, M. Y. Yap, N. R. Pendini, S. W. Polyak, D. Sejer Pedersen, R. Morona, J. D. Turnidge, J. C. Wallace, M. C. J. Wilce, G. W. Booker, A. D. Abell, *J. Biol. Chem.* **2012**, 287, 17823–17832.
[4] R. Zoraghi, R. H. See, P. Axerio-Cilies, N. S. Kumar, H. Gong, A. Moreau, M. Hsing, S. Kaur, R. D. Swayze, L. Worrall, E. Amandoron, T. Lian, L. Jackson, J. Jiang, L. Thorson, C. Labriere, L. Foster, R. C. Brunham, W. R. McMaster, B. B. Finlay, N. C. Strynadka, A. Cherkasov, R. N. Young, N. E. Reiner, *Antimicrob. Agents Chemother.* **2011**, 55, 2042–2053.
[5] L. Wang, B. Martin, R. Brenneman, L. M. Luttrell, S. Maudsley, *J. Pharmacol. Exp. Ther.* **2009**, 331, 340–8.
[6] A. Peracchi, A. Mozzarelli, *Biochim. Biophys. Acta* **2011**, 1814, 922–33.
[7] J. A. Lewis, E. P. Lebois, C. W. Lindsley, *Curr. Opin. Chem. Biol.* **2008**, 12, 269–80.
[8] P. J. Hajduk, J. R. Huth, S. W. Fesik, *J. Med. Chem.* **2005**, 48, 2518–25.
[9] C. Mattos, D. Ringe, *Nat. Biotechnol.* **1996**, 14, 595–9.
[10] A. T. R. Laurie, R. M. Jackson, *Bioinformatics* **2005**, 21, 1908–1916.
[11] F. Glaser, R. J. Morris, R. J. Najmanovich, R. A. Laskowski, J. M. Thornton, *Proteins Struct. Funct. Genet.* **2006**, 62, 479–488.
[12] J. An, M. Totrov, R. Abagyan, *Mol. Cell. Proteomics* **2005**, 4, 752–761.
[13] R. Brenke, D. Kozakov, G.-Y. Chuang, D. Beglov, D. Hall, M. R. Landon, C. Mattos, S. Vajda, *Bioinformatics* **2009**, 25, 621–7.
[14] N. Huang, M. P. Jacobson, *PLoS One* **2010**, 5, e10109.
[15] C. H. Ngan, T. Bohnuud, S. E. Mottarella, D. Beglov, E. A. Villar, D. R. Hall, D. Kozakov, S. Vajda, **2012**, 40, 271–275.
[16] R. J. Keenan, D. M. Freymann, R. M. Stroud, P. Walter, *Annu. Rev. Biochem.* **2001**, 70, 755–775.
[17] G. J. Phillips, T. J. Silhavy, *Nature* **1992**, 359, 744–746.
[18] H. Tian, J. Beckwith, **2002**, DOI 10.1128/JB.184.1.111.
[19] H. Tian, J. Beckwith, *J. Bacteriol.* **2002**, 184, 111–118.
[20] N. D. Ulbrandt, J. A. Newitt, H. D. Bernstein, **1997**, 88, 187–196.
[21] K. Nagai, C. Oubridge, A. Kuglstatter, E. Menichelli, C. Isel, L. Jovine, **2003**, 22, 3479–3485.
[22] M. A. Poritz, H. D. Bernstein, K. Strub, D. Zopf, H. Wilhelm, P. Walter, *Science* **1990**, 250, 1111–1117.
[23] R. Gilmore, G. Blobel, 1983, 35, 677–685.
[24] M. R. Pool, J. Stumm, T. A. Fulga, I. Sinning, B. Dobberstein, *Science* **2002**, 297, 1345–1348.
[25] L. F. Estrozi, D. Boehringer, S.-O. Shan, N. Ban, C. Schaffitzel, *Nat. Struct. Mol. Biol.* **2011**, 18, 88–90.
[26] T. Connolly, P. J. Rapiejko, R. Gilmoret, **1991**, 1171–1173.
[27] R. Kusters, G. Lentzen, E. Eppens, A. van Geel, C. C. van der Weijden, W. Wintermeyer, J. Luirink, *FEBS Lett.* **1995**, 372, 253–8.
[28] J. D. Miller, H. D. Bernstein, P. Walter, *Nature* **1994**, 367, 657–9.
[29] D. M. Freymann, R. J. Keenan, R. M. Stroud, P. Walter, *Nature* **1997**, 385, 361–4.
[30] C. Y. Janda, J. Li, C. Oubridge, H. Hernández, C. V Robinson, K. Nagai, *Nature* **2010**, 465, 507–510.
[31] R. T. Batey, R. P. Rambo, L. Lucast, B. Rha, J. A. Doudna, *Science* **2000**, 287, 1232–1239.
[32] P. F. Egea, S.-O. Shan, J. Napetschnig, D. F. Savage, P. Walter, R. M. Stroud, *Nature* **2004**, 427, 215–21.
[33] Y. Lu, H. Y. Qi, J. B. Hyndman, N. D. Ulbrandt, A. Teplyakov, N. Tomasevic, H. D. Bernstein, *EMBO J.* **2001**, 20, 6724–6734.
[34] U. D. Ramirez, G. Minasov, P. J. Focia, R. M. Stroud, P. Walter, P. Kuhn, D. M. Freymann, *J. Mol. Biol.* **2002**, 320, 783–99.

[35] A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, G. J. Barton, *Bioinformatics* **2009**, 25, 1189–1191.
[36] W. DeLano, *CCP4 Newsl. Protein Crystallogr*. **2002**.
[37] T. L. Blundell, B. L. Sibanda, R. W. Montalvão, S. Brewerton, V. Chelliah, C. L. Worth, N. J. Harmer, O. Davies, D. Burke, *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. **2006**, 361, 413–23.
[38] J. a Newitt, H. D. Bernstein, *Eur. J. Biochem.* **1997**, 245, 720–9.
[39] A. Panjkovich, X. Daura, *BMC Bioinformatics* **2012**, 13, 273.
[40] A. Ivetac, J. McCammon, *Methods Mol. Biol.* **2012**, 819, 3–12.
[41] D. Hamelberg, J. Mongan, J. A. McCammon, *J. Chem. Phys.* **2004**, 120, 11919–11929.
[42] S. F. Ataide, N. Schmitz, K. Shen, A. Ke, S. Shan, J. A. Doudna, N. Ban, *Science* **2011**, 331, 881–6.
[43] U. D. Ramirez, D. M. Freymann, *Acta Crystallogr. D. Biol. Crystallogr*. **2006**, 62, 1520–34.
[44] S. Padmanabhan, D. M. Freymann, *Structure* **2001**, 9, 859–67.

ABBREVIATIONS

GTP	Guanosine triphosphate
GMPPCP	phosphomethyl phosphonic acid guanylate ester
GMPPNP	phosphoaminophosphonic acid-guanylate ester
SRP	Signal Recognition Particle
SR	Signal Recognition Particle Receptor
Ffh	Fifty four homologue
IBD	Insertion Box Domain
OS	Orthosteric sites