4 DISCUSSION

S. flexneri, a Gram-negative pathogenic bacterium presents global human health problem. It is responsible for significant morbidity and mortality especially in children. In the virulence of many Gram-negative pathogens, Type III secretion systems (TTSSs) have a central function. TTSS mediate the secretion and translocation of bacterial effectors into the cytoplasm or membrane of eukaryotic cells [175]. A key feature of TTSS is the presence of cytosolic molecular chaperones for storage and efficient secretion of many effectors. Chaperone loss reduces the secretion of their specific substrate protein(s). Though chaperones share little amino acid sequence homology between themselves, they possess low molecular mass and acidic pI. Chaperones have been classified according to effector binding into three categories. The chaperones of class I and III have been studied extensively but there is conspicuously lack of information regarding class II chaperones. Understanding the mechanism of *S. flexneri* invasion requires addressing the role of chaperones.

4.1 Crystallizing a functional chaperone

Hitherto chaperones belonging to class II have proved either recalcitrant to crystallization [99, 175] or were of poor quality. Only recently i.e., in the first quarter of 2008 the X-ray crystal structure of a chaperone from this class, SycD from *Yersinia* was published [99]. The structure is based on truncated, core domain of chaperone, SycD lacking 25 of the total 168 residues highlights the challenge faced in obtaining crystals of full-length and functional protein. From my own experience which was also reported by others [99, 175, 176], full-length SycD failed to crystallize despite extensive screening. Moreover, attempts to cocrystallize SycD in the presence of its substrates i.e., either YopB or YopD met with no success. The reported SycD crystals were obtained on truncating 20 N-terminal and five C-terminal residues [99]. These crystals diffracted at best to 3.8 Å. Further, only upon application of reductive methylation of lysine residues on the truncated SycD did the quality of crystal improve. This involved modification of 22 methylation sites in SycD21-163.

In the quest to crystallize a functional class II chaperone, IpgC from *Shigella flexneri* was chosen. Though full length IpgC readily crystallized on improving the crystallization conditions, it diffracted at best to 3.5 Å. Moreover, SeMet crystals grown to obtain phases were instable. The cause of instability was attributed to the

presence of two methionines in the close vicinity at N-terminus. Though stable SeMet crystals were obtained upon removal of the methionines, they diffracted at best to 3.4 Å. As a next strategy to stabilize the N-terminus, a short *Strep*-tag II peptide [150] which doubled as purification tag and provided a stable secondary structure was introduced. Crystals from this construct diffracted poorly to a resolution of 5 Å. Finally, upon removal of only four C-terminal residues i.e., Asp-Ile-Lys-Glu having flexible side chains, IpgC yielded crystals with improved diffraction quality. This construct was fully functional in restoring the secretion of effectors, epithelial cell invasion and macrophage cytotoxicity to the wild type level in a nonpolar *ipgC* null mutant *Shigella* (Fig. 3.7). The structure of a functional and nearly full length IpgC gave exceptional insights in interpreting the function of class II chaperone.

4.2 IpgC reveals TPR-like repeats

The TPR domain is a 34 amino acids consensus motif that is found in tandem repeats of varying number in different proteins [167, 177]. Crystallographic structure analysis of TPR-containing proteins revealed that TPR motif adopts a helix-turn-helix arrangement [99, 168, 174]. Class II chaperones are predicted to contain TPR-like repeats [169] (Fig. 6.1). TPR motifs have been identified in various different organisms, ranging from bacteria to humans. The adjacent TPR motifs pack in a parallel fashion, resulting in a spiral of repeating antiparallel α- helices. The functions of TPR containing protein include cell cycle control [178], transcription and splicing events [179], protein transport especially protein import [180], regulatory phosphate turnover [181], and protein folding [182]. Sequence alignment of the chaperones belonging to class II reveals consensus TPR motifs defined by a pattern of small and large amino acids (Fig. 6.1 and 3.9). The structure of IpgC showed 3.5 TPR repeats that folds into an all α-helical structure. IpgC contains three complete and an incomplete TPR domain at amino acid positions 33-65, 70-99, 104-133 and 137-151. The IpgC structure demonstrates the presence of conserved residues at canonical positions (Fig. 6.1). The TPR core domain in IpgC is capped by an N-terminal helix joined to the rest of the molecule by a loop at position 22-32. The reported structure of SycD was crystallized from a construct that lacked the N-terminal helix [99].

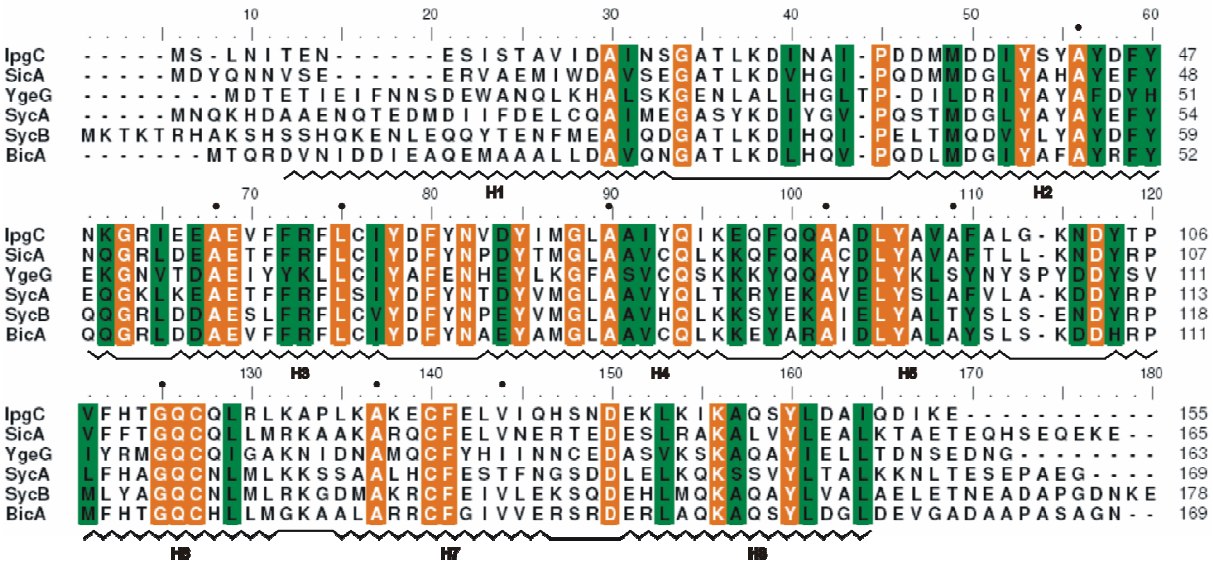


Fig. 6.1: Sequence alignment of class II chaperones. The sequence alignment highlights the conservation of amino acids at canonical positions in the TPR motifs. Identical residues are highlighted in orange and very similar residues are marked in green. The line above the sequence indicates the predicted secondary structure with zigzag line representing helix and straight line indicating a loop. The residues at the canonical positions are marked by a dot on top of the alignment. The chaperones used the alignment are IpgC (*Shigella flexneri*), SicA (*Salmonella typhimurium*), YgeG (*Escherichia coli*), SycA (*Yersinia enterocolitica*), SycB (*Yersinia enterocolitica*) and BicA (*Burkholderia mallei*)

4.3 Polypeptide binding groove of the TPR domain

The uniform angular and spatial arrangement of neighbouring α-helices from the TPR motifs create a groove. This groove provided by the super-helical structure provides a surface suitable for the recognition and accommodation of the target substrate proteins [168]. In fact, association of the peptide with the groove was shown in the HOP structure in complexes with peptides from Hsp70 and Hsp90 [174]. These are the only X-ray structures of a TPR with a bound peptide available so far [174]. No structure of TPR bound to the peptide is available from prokaryotes.

The currently known TPR structures have revealed that the positively-charged pocket on the concave surface of a TPR domain is conserved for binding C-terminal peptide segments of interacting proteins. Apparently, in the HOP structure in complexes with peptides from Hsp70 and Hsp90, the concave surface of a TPR domain is lined with positively-charged residues (Lys229, Asn233, Asn264, and Lys301) [174]. Besides, in the complex structure between PEX5 and PTS1 peptide, the general interaction between the two structures is accomplished via the conserved Asn ladder (Asn489, Asn497, Asn524, and Asn531) [183]. In addition, Lys31, Asn35, Asn66, and Lys96 of the CHIP TPR domain contribute to the binding of Hsp90 peptide [184]. Moreover,

asparagines in the concave surface of TPR domains are found in other TPR-containing proteins such as protein phosphate 5 (PP5) [185], PilF [186], and Tom70 [187], which may be involved in binding their respective peptides.

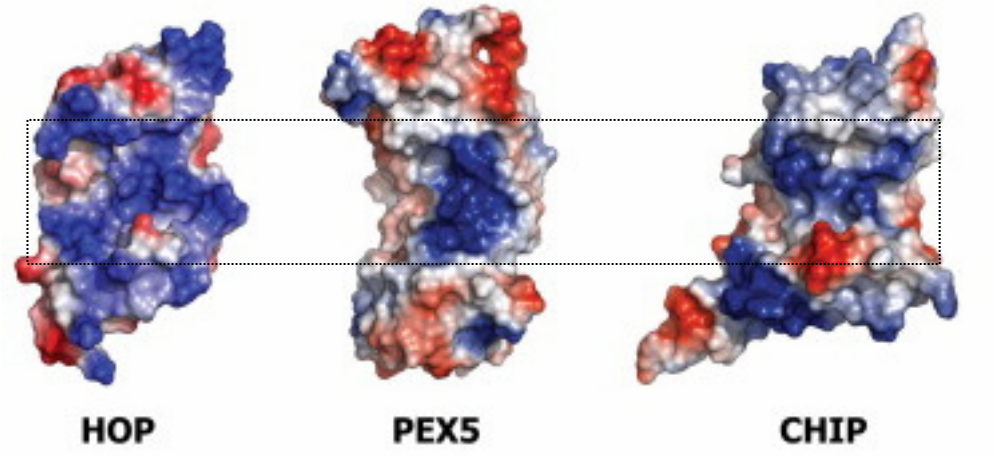


Fig. 6.2: Surface charge presentation of TPR domains in peptide-binding site (boxed). Surface representation of the TPR1 domain of HOP (PDB ID: 1ELW), PEX5 (TPR region of human pex5; PDB ID: 1FCH), and CHIP (PDB ID: 2C2L), colored based on electrostatic potential. Blue and red represent regions of positive and negative electrostatic potential, respectively

Interestingly IpgC provides a surface with both hydrophobic and polar properties in the groove (Fig. 3.12 B) facilitating the association of the interacting polypeptide via both hydrophobic and polar interactions. Additional specificity is rendered by the distribution of the amino acids in the groove. Helix H2, forming the border of the cleft on one side contains an assembly of tyrosine ladder consisting of residues 40, 44 and 47 with their hydroxyl groups pointing towards the cleft. At the other side of the cleft bordered by H8, charged residues Lys138, Lys142 and Tyr146 point towards the cleft (Fig. 6.3).

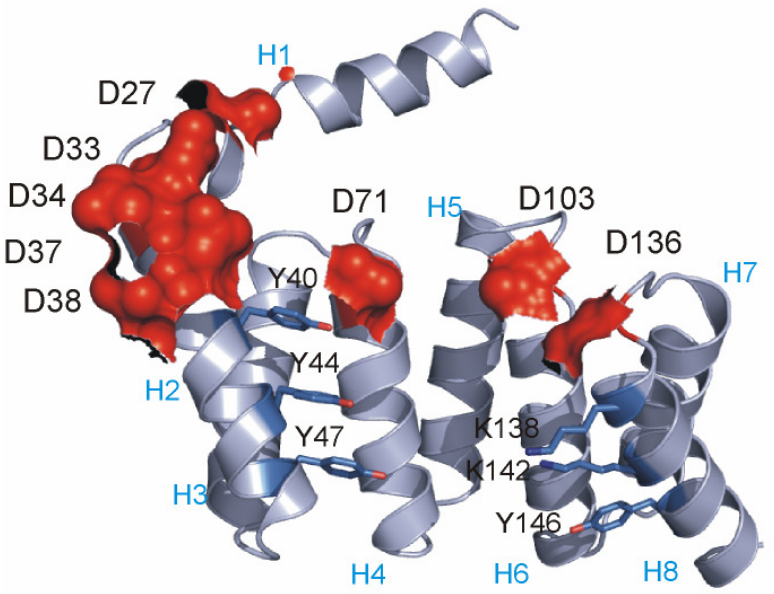


Fig. 6.3: Distribution of amino acids in the groove imparts specificity to the interaction with the target protein. Distinct tyrosine ladder comprising Y40, Y44 and Y47 line up H2 on one side and on the other side charged residues K138, K142 and Y146 line H8. Aspartic acid-array formed by 8 residues form a charged acidic patch on one end.

Furthermore, unique charge distribution at one end of the groove in IpgC forming an 'aspartic acid-array' may provide specificity and directionality to the substrate binding (Fig. 6.3). This is supported by the fact that C-terminus of IpaB CBD makes contact with the acidic patch in crystal structure. Moreover in the purified IpaB-IpgC complex, the C-terminal region following the CBD in IpaB is resistant to proteolytic degradation (Fig. 5.2) suggesting a protective role for the acidic patch. Interestingly, a similar highly negatively charged patch has been proposed to mediate substrate binding in YrrB [188], a TPR containing protein from Gram-positive bacterium, *Bacillus subtilis*. It functions as mediator in complex formation among RNA sulfuration components. The recently solved class II chaperone, SycD also exhibits such a feature [99].

The groove at its narrowest end between H2 and H8 is 10.5 Å. This constricted zone can barely accommodate a secondary structural element like a helix as previously proposed for binding of a peptide in the PP5 [168] and YopD peptide in the modeled structure of SycD [169]. The accommodation of the peptide in an extended conformation becomes evident from the X-ray structures of TPR domain-peptide complexes of Hop with Hsp70 and Hsp90 [174].

4.4 Functional IpgC is a dimer

The structure of chaperones of TTSS currently available show dimeric organization [74, 83-89]. With regard to IpgC, the dimer found in the structure could be corroborated by analytical as well as functional assays. Static light scattering reveals IpgC to be dimer. A remarkably large area i.e., over 30% of the theoretically accessible surface of both copies are involved in dimer contact. The intermolecular contact over this large area is facilitated by extensive hydrophobic interactions supported by eight hydrogen bonds (listed below and in Table 3.3) between helices H3, H4 and H5 of SU A and H1, H2 and H3 of SUB (Fig. 6.4). Highly conserved residues are involved in the intermolecular interaction (Fig. 6.1).

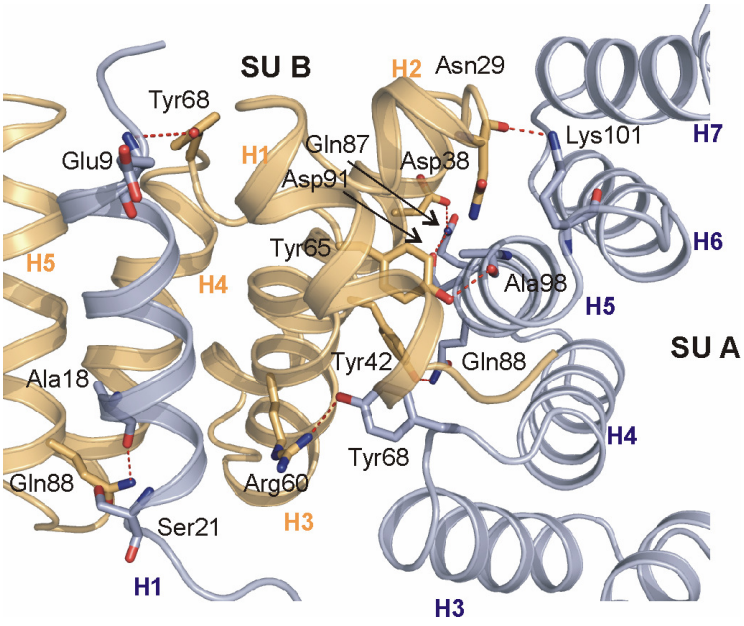


Fig. 6.4: Polar interaction at the dimer interface. Hydrophobic SU interactions are strengthened by H-bonds formed between SU A and B (Glu9/Tyr68, Ala18/Gln88, Ser21/Gln88, Tyr68/Arg60, Gln87/Asp38, Gln88/Tyr42, Ala98/Tyr65, Lys101/Asn29).

4.5 Role of H1 in dimeric interaction

The amphiphilic helix, H1 is crucial for maintaining the dimer conformation of IpgC. As a consequence of asymmetric dimer organization, H1 and the loop connecting it to H2 show different arrangement in the subunits. H1 interacts hydrophobically and is stabilized by hydrogen bonds at the termini. A shallow groove on the convex side of subunit B is formed as a result of interaction of H1 from SU A with H3, H4, and H5 from SU B (Fig. 6.5). Keeping in view that YopD could interact with the convex side of SycD [189], the shallow groove on the convex side of IpgC might have a functional significance in providing additional binding surface.

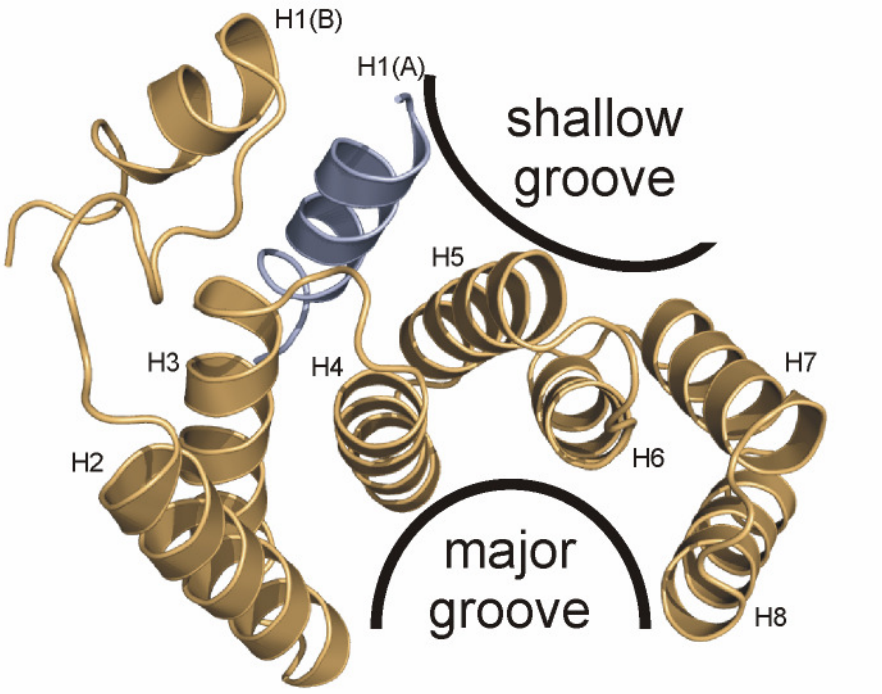


Fig. 6.5: Close up of the interaction between H1 of both subunits. There resultant shallow groove formed by H1 (A), H3 (B), H4 (B) and H5 (B) provides a surface for additional binding on IpgC.

To investigate physiological relevance of H1 in dimerization, *ipgC* null mutant *Shigella* was complemented with H1 truncated IpgC. This strain exhibited loss of invasion function indicating the absolute requirement of H1 for proper function of IpgC. Furthermore, purified H1 truncated IpgC ($IpgC^{\Delta 21}$) is soluble but prone to aggregation as shown by static light scattering experiments. This demonstrates that H1 plays a crucial role in the functional dimerization of IpgC. Moreover dimerization is absolutely required for virulence.

Based on the crystal packing, Büttner *et. al.*, [99] present three different alternative homodimer assemblies of H1 truncated SycD$^{21-163}$ (Fig. 1.9). Of these, the head-to-head dimer was concluded as physiological dimer based on dimer disrupting mutations involving A61E and L65E substitutions (Fig. 6.6). In IpgC due to asymmetric interactions equivalent residues i.e., F61 and Y65 from SU B but not SU A face the dimer interface. The interactions contributed by H1 become obvious when IpgC and SycD dimers are compared to each other (Fig. 6.6). The helices H1 (A) and H1 (B) in IpgC interact extensively with the corresponding other SU, while absence of H1 in SycD results in strongly reduced overall interaction surface.

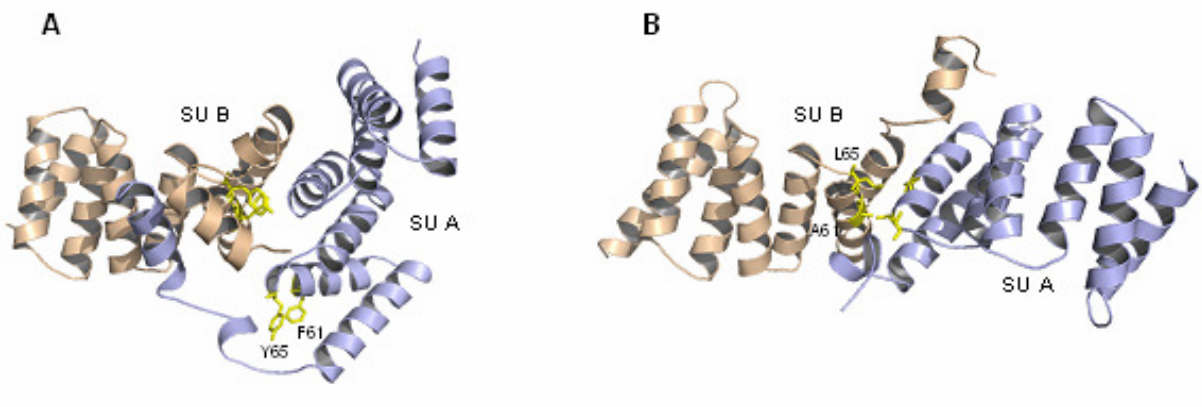


Fig. 6.6: Comparison of IpgC dimer (**A**) and SycD$^{21-163}$ dimer (**B**). SU B of both dimers are superimposed using Secondary Structure Matching server [190]. The algorithm matched 120 residues (residues 32-151 of IpgC and corresponding residues 32-135 and 140-155 of SycD$^{21-163}$) with Cα RMSD of 1.34 Å. Residues A61 and L65 in SycD the corresponding residues F61 and Y65 in IpgC are represented as yellow sticks.

The comparison of superimposed structures of IpgC and SycD highlights the importance of H1 in stabilizing a particular quaternary structure as seen in IpgC. Moreover, the application of reductive methylation on truncated SycD$^{21-163}$ could have influenced the crystal packing. Although different mode of binding exhibited by IpgC and SycD to their substrates in influencing the quaternary structure organization cannot be ruled out. IpgC has been shown to interact independently with IpaB and IpaC [65] while SycD interacts simultaneously with both YopB and YopD [68]. The involvement of H1 in clustering of the helices to generate a grooved surface on the convex side of SU B that can accommodate another polypeptide augurs well with the finding that YopD interacts with the convex side of SycD [189].

4.6 Characterization of the IpaC-IpgC interaction

In order to gain insights into the mechanism of how IpgC interacts with its substrates, characterization of IpaC-IpgC complex was undertaken. IpaC is a 363 amino acid long, 42 kDa protein. IpaC together with IpaB serve as the primary effector for *Shigella* invasion of epithelial cells. IpaC when expressed recombinantly is insoluble and tends to result mainly in inclusion bodies [170], however substantial gain in solubility was achieved when co-expressed with IpgC. Using a yeast two-hybrid system, Page et al. [108], suggested that IpgC interacts with the residues 73-122 of IpaC. Yet in another study by Harrington et al. [116], a region from 50-80 in IpaC was shown to interact with IpgC. For successful crystallization attempts detailed knowledge of the binding region is a prerequisite. Therefore, the chaperone binding region was precisely mapped and corroborated by different techniques.

4.6.1 CBD in IpaC spans from residues 36-68

Limited proteolysis can be used to probe conformational features of proteins and in the elucidation of structure-dynamics-function relationships of protein [191]. The technique was employed to examine the regions of IpaC protected by IpgC from proteolytic degradation. Contrarily, proteolysis as a protein foot-printing tool in mapping the putative chaperone binding region of IpaC in IpaC-IpgC complex did not yield any conclusive lead. IpaC is prone to complete degradation by both thermolysin having low cleavage specificity and trypsin, possessing narrow cleavage specificity. The rapid degradation of IpaC is indicative of unstructured nature of this molecule. This intrinsic property of IpaC necessitates it association with a chaperone to prevent premature degradation in the cytosol of *Shigella*.

In the pursuit to map the fragments of IpaC that associate with IpgC, a novel methodology was adopted. The strategy involved sequentially combining proteolysis of IpaC-IpgC complex followed by size exclusion chromatography and finally analysing the fragments of IpaC bound to IpgC by mass spectrometry. The results from mass spectrometry were utilised in designing several construct addressing the regions of IpaC that bound to IpgC. These constructs were co-expressed with tagged IpgC and pull-down was performed. By this systematic approach the CBD in IpaC was narrowed down to the N-terminus 126 amino acids which comprised the fragment, $IpaC^{47-75}$, identified by mass spectrometry and also the regions reported in the literature [108, 116]. Similar methodology can be emulated to address such a

situation involving unstructured protein in a protein complex to gain knowledge about the putative binding regions. Next, to specifically and precisely map the boundaries of the CBD, a relatively new technique was adapted [134, 171]. The strategy involved site-specifically incorporating photo-excitable unnatural amino acid, *p*-benzoyl phenylalanine (*p*Bpa). IpaC was extensively screened using this method for region involved in binding to the chaperone. A region from amino acids 36-68 could be precisely mapped as the chaperone binding domain in IpaC. This method not only enabled to validate the binding regions but also to precisely define it. Hence using a combination of techniques including limited proteolysis combined with size exclusion chromatography and mass spectrometry, pull-down assays and photocrosslinking, the chaperone binding region in IpaC was precisely mapped to amino acids 36 to 63. To further gain insights into substrate-chaperone function, this information was exploited for structural studies. Cocrystals of IpaC-IpgC were obtained in many conditions by incubating IpaC36-68 peptide encompassing the residues of chaperone binding domain from amino acid 36-68 together with IpgC. Unfortunately, in none the electron density for the peptide could be traced. This could be due to weak affinity of IpaC to IpgC. To overcome this problem, the residues IpaC 36 to 68 were fused to the C-terminus of the IpgC separated by a 20 amino acid linker. This strategy yielded crystals that diffract X-rays. The crystallization conditions are being optimized currently to improve the size of the crystals. The X-ray structure would not only the reveal mode of binding but also throw light on the role of the chaperone in stabilizing an unstructured IpaC.

4.7 Characterization of the IpaB-IpgC interaction

IpaB is a 580 amino acid long, 62 kDa protein. In the cytoplasm of the *Shigella*, IpaB is bound by its cognate chaperone IpgC and is among the first proteins to be secreted. IpaB is believed to form an extracellular complex with IpaC and IpaD after their secretion across the bacterial membrane and facilitate further effector delivery [44, 65]. IpaB mediates *S. flexneri* entry into the cytoplasm of the eukaryotic host cells, an essential step in the pathogenesis of shigellosis [115]. In addition to its role in invasion, IpaB is both necessary and sufficient to induce pyroptosis a form of apoptosis [103]. Several structural features on IpaB have been mapped and linked to functions [111]. By two-hybrid selection in yeast Page et al. [108] mapped the chaperone binding domain between residues 58 and 72 of IpaB.

4.7.1 Characterizing the CBD in IpaB

Heterologous expression of IpaB alone was only observable either by immunoblot or in the membrane fraction of *E. coli* grown to a higher density and induced for a shorter time [109]. Indeed the expression of *ipaB* alone led to growth retardation in *E. coli*. Cytosolic expression of IpaB was achievable only when co-expressed with IpgC. This underlines the role of IpgC in the stable expression of *ipaB*. Thus, IpgC not only improves the expression level of *ipaB*, but also prevents the toxic effect of IpaB on *E. coli*. In *Shigella*, very low amounts of IpaB were detected in the absence of IpgC.

Protease footprinting on IpaB-IpgC complex yielded a stable core of IpaB. This protease resistance core of IpaB starts from amino acid 51 and is in agreement with the yeast two-hybrid assignment of CBD as $IpaB^{58-72}$ [108]. The result emphasizes the utility of proteolysis in obtaining information about protein core that remains resistant to proteolysis.

To further verify these results and to define the boundaries of the CBD on IpaB, photocrosslinking experiments by site-specific incorporation of photo-excitable unnatural amino acid in IpaB was performed. Applying the information from proteolysis in screening the region involved in binding to the chaperone by photocrosslinking, the amino acids 51-72 on IpaB could be mapped. This method not only validated the finding from proteolysis but also to precisely define it. This CBD will be henceforth referred to as CBD2.

4.7.2 An additional CBD in IpaB (CBD1)

Interestingly a second chaperone binding region (CBD1) in IpaB could be assigned to residues 18-35. The presence of an additional chaperone binding site in the N-terminus was evident as CBD2 deletion mutant, IpaB$^{\Delta 51\text{-}72}$ could be co-purified with IpgC while IpaB$^{\Delta 1\text{-}72}$ could not be expressed even in the presence of IpgC. The association of IpaB$^{\Delta 51\text{-}72}$ lacking the CBD2 but not IpaB$^{\Delta 1\text{-}72}$ with IpgC was indicative of the presence of an additional binding site N-terminus to amino acid 51 in IpaB. Next, to rule out the possibility of local structural rearrangement upon deletion of CBD2 contributing to the association with IpgC, a chimeric IpaB was generated. This chimeric IpaB having a protease cleavage site in place of CBD2 (residues 51-72) could be co-purified with IpgC. Moreover efficient cleavage of the proteolytic site was not observed plausibly due to masking of this site by bound IpgC. Next by employing the site-specific incorporation of photo-excitable amino acid at various positions at the N-terminus of the IpaB a second chaperone binding region (CBD1) between amino acids 15-45 could be mapped.

4.7.3 *In vivo* relevance of CBD deletions

To determine the effect of chaperone binding domain on the secretion, *ipaB*$^{\Delta 1\text{-}50}$ (CBD1 deletion), *ipaB*$^{\Delta 51\text{-}72}$ (CBD2 deletion) as well as *ipaB*$^{\Delta 1\text{-}72}$ (total CBD deletion) mutants were complemented in *ipaB* null mutant and tested for secretion and invasion. Stable expression of IpaB$^{\Delta 1\text{-}50}$ and IpaB$^{\Delta 51\text{-}72}$ but not IpaB$^{\Delta 1\text{-}72}$ was observed indicating that presence either of the two CBDs is essential. Only IpaB$^{\Delta 51\text{-}72}$ was secreted. Lack of secretion of IpaB$^{\Delta 1\text{-}50}$ could be attributed to the absence of N-terminal signal sequence. In the invasion assay only IpaB$^{\Delta 51\text{-}72}$ showed invasiveness similar to wild type *Shigella*. This proves the prerequisite for the presence of either of the CBDs in the stable expression by associating with IpgC.

Next to check the dispensability of IpgC, *ipaB/ipgC* double mutant was generated and complemented with the above mentioned mutants. None were stably maintained in the cytosol and were non-invasive. Invasiveness could be restored only when complemented by both wild type IpaB and IpgC. This proves the absolute requirement of IpgC for stability of IpaB in the cytosol of the *Shigella*.

4.8 The IpaB-IpgC structure

IpgC was crystallized in complex with CBD1 (IpaB18-35) and CBD2 (IpaB51-72) from IpaB. The structure revealed the presence of CBD2 bound in the major groove of IpgC. The peptide-bound and apo-IpgC structures superpose well, with an overall root mean square deviation of 0.5 Å. The SUs in the complex are tilted by about 15° and shifted by 5 Å relative to each other (Fig. 6.7). It is unclear whether this difference is the result exclusively of substrate binding or is influenced by crystal packing.

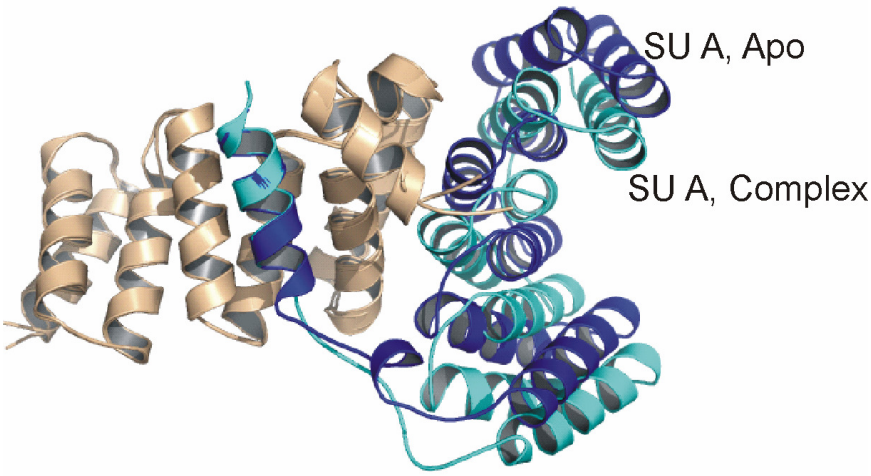


Fig. 6.7: Superimposition of the apo- and complexed IpgC. Monomer B (brown) was set as a reference. The TPR domains of subunits A (blue and cyan for apo and complex, respectively) are displaced.

Moreover the IpaB16-72 peptide-bound and apo-IpgC structures remain unchanged when superposed. Hence, the superimposability of both apo- and complexed IpgC indicates conservation and rigidness of overall structural architecture.

4.8 Polypeptide binding to groove of the TPR domain

The super-helical TPR structure forming a continuous helical groove accommodates the IpaB peptide in an extended conformation. The extended conformation of the peptide in the groove is also seen in the only other crystal structures of TPR domain-peptide complexes available namely the TPRs of the adaptor protein Hop with molecular chaperones Hsp70 and Hsp90 [174]. The surface provided by the groove can barely accommodate a helical peptide as shown previously by modeling a helix in the peptide-binding groove of PP5 [168] and SycD [169]. Additionally the peptide is

bound in a similar antiparallel orientation in the grooves of both IpgC and Hop. The last 13 residues in the 22 residue ipab51-72 peptide are ordered in the IpgC complex and thus visible in the final electron density map (Fig. 6.8).

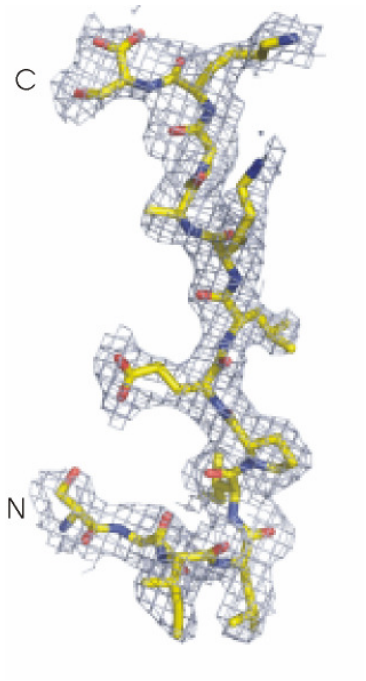
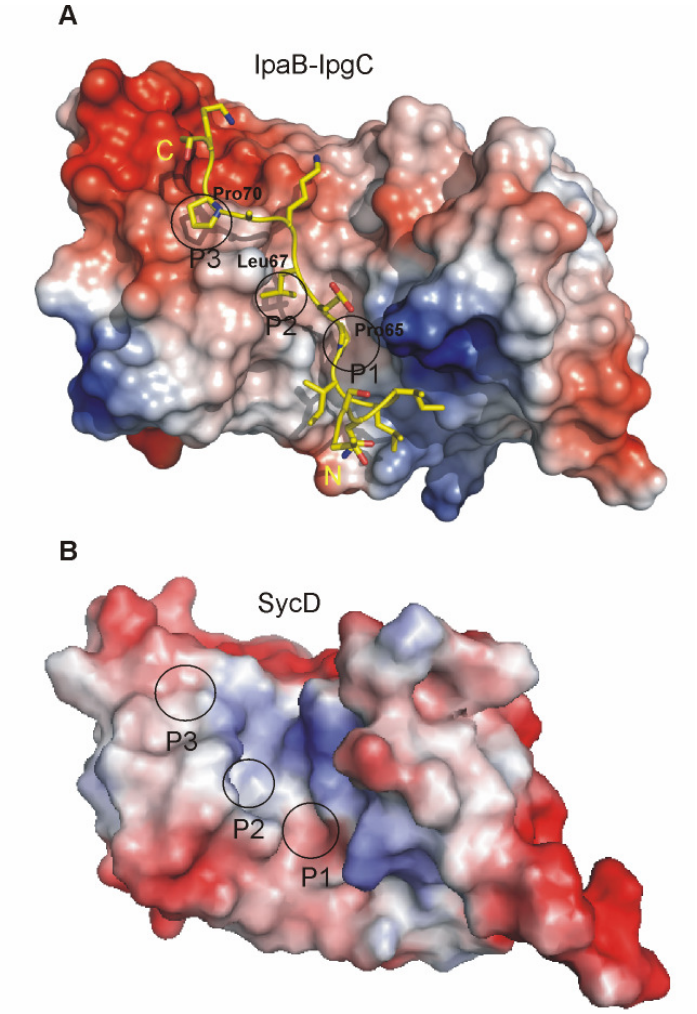


Fig. 6.8: Composite omit map around the CBD found in contact with monomer B. The simulated annealing 2Fo-Fo electron density map of IpgC-bound peptide in SU B, calculated without including the peptide, shows that the last 13 IpaB51-72 peptide residues are very well defined in the groove of SU B of IpgC

The binding of the IpaB peptide in an extended conformation allows specific as well as extensive interaction with the surface of the groove presented by IpgC. Indeed, the amphipathic surface in the groove (Fig. 3.12 B) facilitates three types of interactions with the bound peptide. First there are three specific salt bridges, second, three strong but not necessarily sequence specific hydrogen bonds and finally, hydrophobic and van der Waals interactions. The salt bridge interactions involve aspartates (Asp37, Asp71) from the acidic patch and Lys142 from the charged residue assembly on H8 of IpgC which interact with Lys71, Lys68 and Glu68 of the peptide. The Tyr40 and Tyr47 from the tyrosine ladder participate in the hydrogen bond interaction with the peptide backbone. The hydrophobic and van der Waals interactions support the interaction of the peptide with three discrete pockets labelled P1, P2 and P3 on the surface of IpgC. These pockets are formed by highly conserved residues that could be mapped in SycD as well (Fig. 6.9).



Pockets	IpgC	SycD
P1	Tyr47	Tyr47
	Ala78	Ala78
	Gln81	Gln81
	Tyr93	Tyr93
	His109	His109
	Ile82	Ala82
P2	Tyr44	Phe44
	Tyr47	Tyr47
	Phe59	Phe59
	Met74	Leu74
	Gly75	Gly75
	Ala78	Ala78
P3	Tyr40	Tyr40
	Ser41	Ser41
	Tyr44	Phe44
	Asp45	Asn45
	Asp37	Glu37

Fig. 6.9: Conserved residues line the three discrete pockets. A, Ligand bound IpgC with the pockets P1, P2 and P3 marked in circles. B, SycD shows similar po comparison of residues lining the pockets in the groove from IpgC and SycD reveal identical residues. Table on the right lists the conserved residues lining the pocket from IpgC and SycD.

The pockets P1, P2 and P3, interact with side chains of residues Pro65, Leu67 and Pro70 in the IpaB peptide. Interestingly, protein binding in the eukaryotic adaptor protein Hop [174] is also mediated through pockets similar to P1 and P2 to the molecular chaperones Hsp70 and Hsp90. Any mutations in the residues forming the pockets have been demonstrated to abolish binding of SycD to it substrates [189].

4.9 A conserved chaperone binding motif

The pockets formed mostly by conserved residues impose a selection criterion on the interacting amino acids. A search based on the IpaB sequence bound to the pockets in the groove revealed similar sequence in homologous substrates (Fig. 6.10). The effect of individual residues interacting with the pockets was further assessed by ITC.

Unfavorable mutation in any pocket binding residue i.e., P65R, L67N and P70D resulted in strong reduction in binding to IpgC compared to wild type sequence. Whereas a salt bridge disrupting mutation K68A showed reduced affinity.

	P1		P2			P3
IpaB (*S.flexneri*)	P	E	L	K	A	P
IpaB (*E.coli*)	P	E	L	K	A	P
BipB (*B.mallei*)	P	A	L	R	A	P
YopB (*Y.enterocolitica*)	V	Q	L	P	A	P
YopD (*Y.enterocolitica*)	P	E	L	I	K	P
SopB (*S.typhimurium*)	V	G	L	K	P	P

Fig. 6.10: Substrate motif for class II chaperones. Local sequence alignment of the predicted chaperone binding domain from six different effector proteins. Binding pockets are labelled and bridging residues are highlighted in white.

To validate the identification of chaperone binding motif, binding of IpgC homologue to its substrates was assessed. Binding of SycD to the predicted chaperone binding motifs in substrates- YopB and YopD was checked and confirmed. SycD bound to the peptides VQLPAP (YopB) and PELIKP (YopD) with an affinity of 445 µM and 885 µM respectively which is comparable to an affinity of 625 µM shown for binding of IpgC to IpaB peptide (PELKAP).

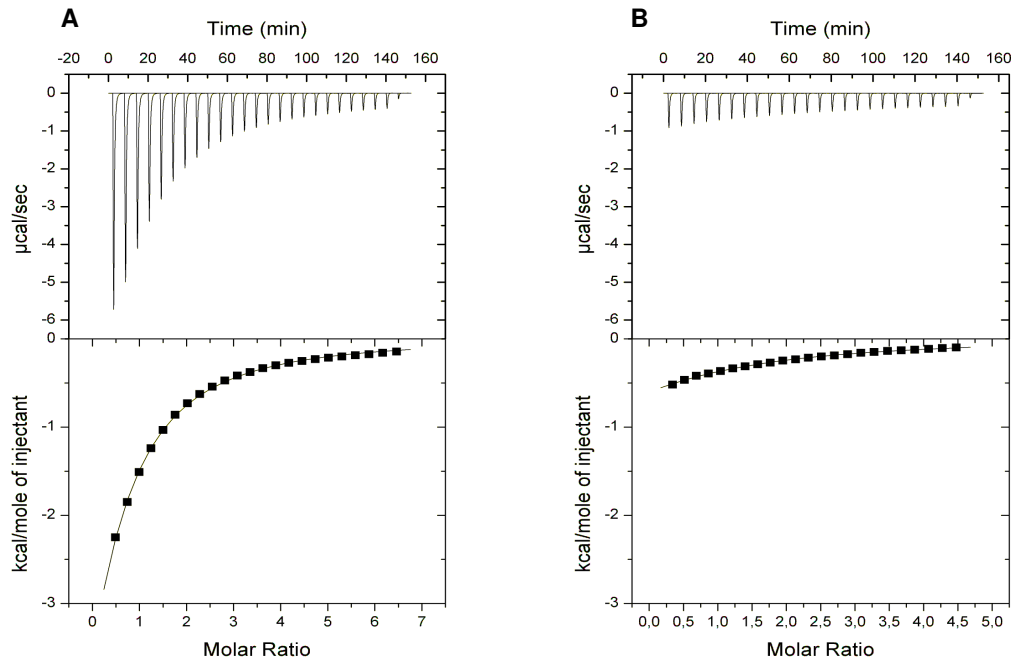


Fig. 6.9: **A**, Binding of YopB peptide (VQLPAP) showed an affinity of 445 ± 19 µM to SycD while **B**, YopD (PELKIP) showed an affinity of 885 ± 78 µM to SycD.

In summary, based on the IpaB-IpgC structure and considering the sequences of homologous substrates, a chaperone binding motif in TTSS effector proteins is identified. This motif (Fig. 6.8), in agreement with other observations [192], is located within the 100 N-terminal residues of TPR-chaperone substrates. The sequence of the motif implements the following criteria; a conserved Pro or Val occupies P1, while the conserved residues Leu and Pro occupy P2 and P3 respectively. An amino acid, preferentially with a negative charge, is sandwiched between P1 and P2. While a basic residue and a non-polar amino acid make a bridge between P2 and P3 in the substrate.