Structural studies on a secretion chaperone from Shigella flexneri and crystallographic explorations with a thermostable aldolase

van Eerde, Jan Hendrik

Chapter 1

The *Shigella* type III secretion system

and

the structural and functional properties of type III secretion chaperones

Bacillary dysentery and its causative agent

Bacteria of *Shigella* spp. and enteroinvasive *Escherichia coli* (EIEC) are responsible for bacillary dysentery, or shigellosis, in humans. Dysentery is a diarrhoeal disease which occurs throughout the world. In particular, bacillary dysentery is endemic in many regions in developing countries and moreover, areas of civil conflict or natural disasters are almost routinely visited by *Shigella*. The latest (conservative) estimates of the global incidence of *Shigella* infections are a number of 165 million episodes of shigellosis throughout the world each year. In over a million cases the disease has a deadly outcome. The great majority (163 million annually) of *Shigella* victims live in developing countries. It is not surprising that children are the major victims of shigellosis. Being weaker than adults, malnutrition and anaemia dangerously diminish their defence systems against Shigellae. Each year, nearly a fifth of the <5 year-old children in developing countries suffer from an episode of shigellosis (Kotloff *et al.*, 1999; WHO, 2005).

The severity of the disease can range from mild diarrhoea to severe dysentery with frequent passage of bloody, mucoid, small-volume stools. Other typical symptoms of bacillary dysentery include fever, abdominal pain and a feeling of general malaise. These symptoms emerge within twelve to fifty hours after infection and can last one to four weeks (US FDA, 2007; WHO, 2005). Long-term infection is known but extremely rare. *Shigella* are host-adapted to the human intestine. However, some primates can also become infected as a result of contact with humans.

Usually the disease is spread by the excreta of infected persons. This can occur either directly, by the faecal-oral route, or by contaminated food or water. The infectious dose of *Shigella* is comparatively low. In studies among volunteers, oral ingestion of as few as ten organisms already caused infection in 10 % of the cases (Janda and Abbott, 1998). High infection rates are typically encountered in overcrowded areas where water supplies are in danger of being contaminated by excreta. This provides a partial explanation for the high incidence of bacillary dysentery outbreaks among populations struck by war or natural disaster. During the 1994 genocide in Rwanda, approximately 20,000 Rwandan refugees who had fled into the North Kivu region of the Democratic Republic of the Congo died in the first month alone from dysentery caused by a strain of *Shigella* that was resistant to all commonly used antibiotics (Goma Epidemiology Group, 1995; WHO, 2007). The general hygienic level in industrialised countries is sufficient to prevent full-fledged *Shigella* epidemics. However, hygienic conditions in some types of institutions, e.g. day-care centres for children and toddlers, can be sub-optimal, which provides opportunities for pathogens like *Shigella* (Kotloff *et al.*, 1999).

Diarrhoeal diseases continue to be one of the leading causes of disease and death in developing countries. Many of the more vigorous forms of diarrhoea can lead to serious dehydration, which can have a deadly outcome. In shigellosis, a strong inflammatory reaction to the infection adds to the severity of the disease. In addition to oral rehydration, shigellosis patients are therefore often treated with antibiotics. This treatment can significantly reduce the duration of infection and prevents further spread by the eradication of the pathogenic organism from the stool. However, the emergence of *Shigella* strains resistant to many of the common anti-microbial drugs is a growing clinical problem, especially in developing countries. At present, few reliable options remain to treat a multi-resistant *Shigella* infection (Kotloff *et al.*, 1999; WHO, 2005).

The pathogenic mechanism of *Shigella*

An important feature of the infection mechanism by *Shigella* spp. is the colonisation of the cell-lining, the epithelium, of the gut (Figure 1). *Shigella* spp. can breach the defensive barrier of the epithelial lining by passing through M cells. M cells normally take up and transport foreign material from the gut to present it to underlying immune effector cells, which then elicit an appropriate immune response (Siebers and Finlay, 1996). *Shigella* spp. readily transmigrate these M cells and kill resident underlying macrophages by inducing their apoptosis. Subsequently, invasion of epithelial cells from the unprotected basolateral side is possible. *Shigella* spp. actively induce their uptake by the normally non-phagocytic epithelial cells by stimulating the formation of membrane ruffles and leaflets around the entering bacterium, which ultimately engulf the bacterium. Entry of *Shigella* into these cells is fast and completed within a matter of minutes (Adam *et al.*, 1995; Adam *et al.*, 1996). Once inside they multiply and spread into adjacent cells (Figure 1). This process of intracellular multiplication and cell-to-cell spread protects *Shigella* from the innate immune response. The immune response caused by the apoptosis of macrophages and the destruction of the epithelial cell lining is relatively strong and leads to a severe inflammation, which is primarily responsible for the clinical symptoms of shigellosis (Fernandez and Sansonetti, 2003; Sansonetti, 2001).

Four species of *Shigella* are distinguished, which are *S. flexneri*, *S. dysenteriae*, *S. sonnei* and *S. boydii*. These species are genetically so closely related to *Escherichia coli* that they could all be regarded as belonging to this species. *S. flexneri* is the most common causative agent of shigellosis and serves as the model organism for the pathogenesis of Shigellae. As compared to other *E. coli* strains, a characteristic genetic feature of *Shigella* spp. and EIEC strains is the presence of a large 213 kb virulence plasmid. This virulence plasmid is necessary for its pathogenis (Sansonetti *et al.*, 1982) and contains ~100 genes (Buchrieser *et al.*, 2000). A 31 kb region on this plasmid is called the entry region, because the genes in this region are necessary and sufficient for *Shigella* to invade epithelial cells (Maurelli *et al.*, 1985; Sasakawa *et al.*, 1988). The genes in this region code for a sophisticated bacterial secretory system belonging to the family of type III secretion systems.

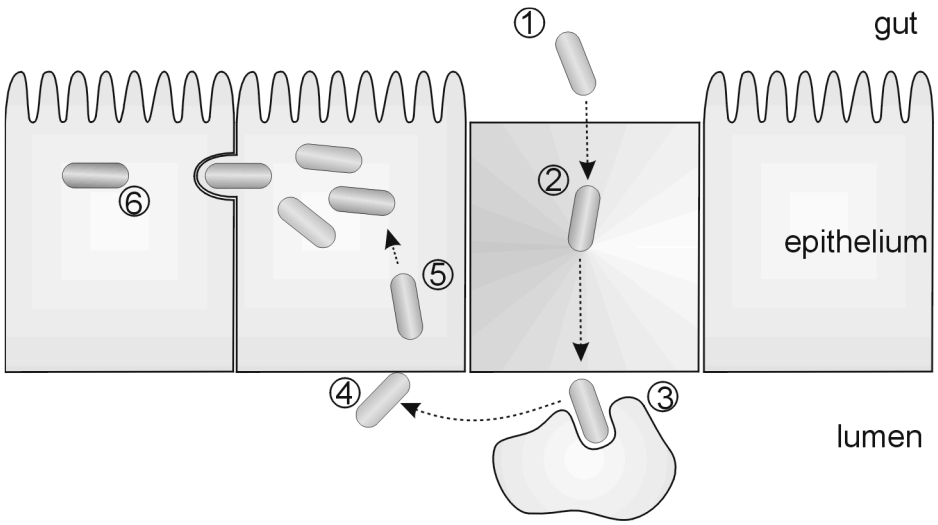

Figure 1: Model of the *Shigella* mechanism of invasion of cells in the colonic epithelium. 1: Arrival of *Shigella* in the gut. 2: Passage through M cells. 3: Killing of underlying macrophages and onset of the immune reaction. 4: Entry of epithelial cells from their unprotected basolateral side. 5: Intracellular multiplication 6: Spread to adjacent epithelial cells.

The *Shigella* type III secretion system

Type III secretion (T3S) systems are used by Gram-negative bacteria to inject virulence proteins in one step into the cytosol of eukaryotic host cells. The system consists of *i*) a T3S apparatus (T3SA) that spans the entire bacterial envelope, *ii*) translocator proteins that transit through the T3SA and insert into the host cell membrane to form a pore (called the translocon), *iii*) effector proteins that transit through the T3SA and the connected translocon and are injected directly into the eukaryotic cell, *iv*) secretion chaperones that bind to translocators and some effectors prior to their transit through the T3SA, and *v*) transcriptional activators.

The ability for direct injection of virulence proteins into host cells confers a distinct advantage on bacteria. T3S systems are therefore often an essential part of the pathogenic mechanism of Gram-negative bacterial pathogens. However, the operation of T3S systems does not necessarily result in disease. Several symbiotic bacteria also utilise the T3S system to maintain communication with their hosts. So far, T3S systems capable of communicating with eukaryotic cells have been encountered in proteobacteria as well as *Chlamydia* spp. (Troisfontaines and Cornelis, 2005). Eukaryotic targets of T3S systems comprise both animal and plant cells, but a T3S system is always adapted to one type of host. The genes coding for the T3S apparatus are usually clustered in a so-called pathogenicity island, e.g. the *Shigella* entry region, and many bacteria are thought to have acquired these clusters through horizontal gene transfer. The genomes of several bacteria even encode more than one T3S system gene cluster. In such a case, each of them probably fulfils a role at a distinct stage of the bacterial pathogenic mechanism (Troisfontaines and Cornelis, 2005).

Among the best characterised T3S systems are, besides the *Shigella* T3S system, the *Yersinia* Ysc system, the *Salmonella* Inv-Spa system and the Entero-pathogenic *E. coli* (EPEC) LEE-encoded Esc system (Troisfontaines and Cornelis, 2005). A similar T3S system is also used in the assembly of the bacterial flagella of both Gram-negative and Gram-positive bacteria. As part of the growing flagellum, it secretes the extracellular building blocks of the hook and the flagellar filament (Macnab, 2003). *Shigella* constitutes a good model system for the study of type III secretion. This pathogen contains only one T3S system, has no flagella, and the involvement of the T3S system in the *Shigella* entry process is relatively well-characterised. Here, an overview of the components and operating principles of a T3S system will be given, with special reference to the *Shigella* T3S system.

The T3S apparatus

The *Shigella* T3S apparatus, also called the T3S injectisome, consists of ~ 20 proteins and can be divided into the structural framework of the apparatus and the T3SA machinery.

The structural framework

The structural framework of the T3SA, or needle complex, has been visualised by electron microscopy (EM) (Blocker et al., 1999; Blocker et al., 2001; Sani et al., 2007a; Tamano et al., 2000). The EM images revealed a large cylindrical structure composed of a basal body with a diameter of 26 nm and a height of 32 nm and an extracellular needle with a diameter of 8 nm and a length of 45-50 nm. The basal body consists of two pairs of rings that span the inner and outer membrane (Figure 2A) and consists of the following proteins. The lower pair of rings spans the inner membrane and is formed mainly by MxiG and MxiJ (Allaoui et al., 1995; Allaoui et al., 1992) and presumably has a 24-fold symmetry (Sani et al., 2007a), while the upper pair of rings spans the outer membrane and is formed by MxiD (Allaoui et al., 1993), a member of the secretin family, and the periplasmic secretin pilot MxiM (Lario et al., 2005; Schuch and Maurelli, 1999; Schuch and Maurelli, 2001a) with an approximate 15-fold symmetry (Sani *et al.*, 2007a). MxiJ probably contains a sizeable periplasmic domain and is thought to bridge the two pairs of rings (Schuch and Maurelli, 2001a).

The T3SA structure is hollow, with a continuous channel running from the approximate location of the inner membrane to the needle tip with an internal diameter of ~2-3 nm (Figure 2A). This suggests that exported proteins can pass through this hollow channel in a one-step transport process. The *Shigella* extracellular needle is made up of multiple copies of the 9 kDa MxiH (Blocker *et al.*,

2001; Cordes *et al.*, 2003). The crystal structure of MxiH was recently solved and revealed that MxiH consists of two bent helices, connected by a short turn (Deane *et al.*, 2006). X-ray fibre diffraction studies have shown that these subunits assemble in a helical way into the needle (Cordes *et al.*, 2005) and a detailed model of the needle structure has been made (Deane *et al.*, 2006). The helical assembly and the core feature of the two-helix building block are strikingly similar to flagellin, the subunit of the bacterial flagellar filament. The D0 domain of flagellin, important for the intersubunit interactions also consists of two bent helices and flagellins assemble into a filament with a similar helical pitch (Yonekura *et al.*, 2003).

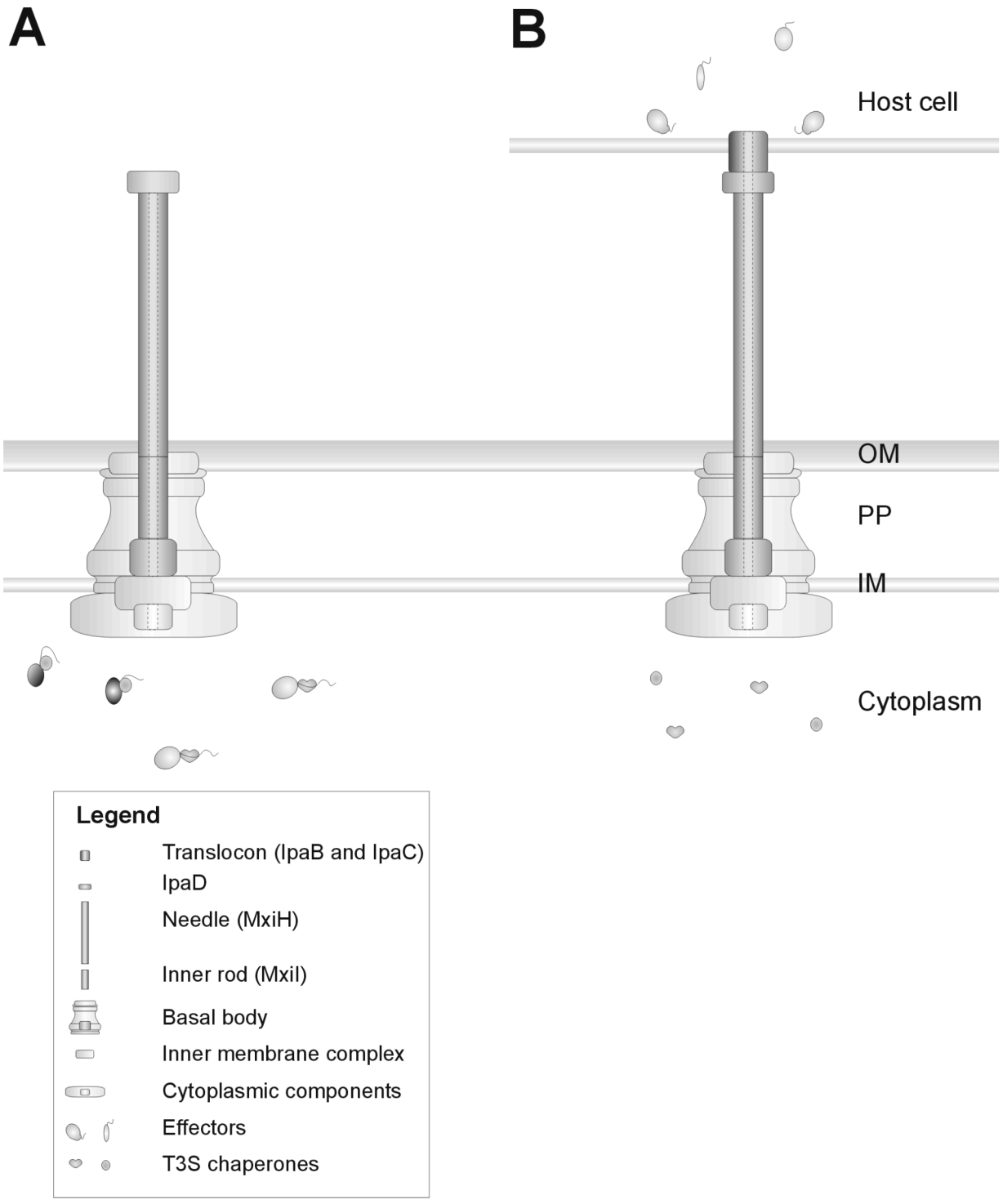

Figure 2: Schematic view of the type III secretion apparatus (T3SA) **(A)** The T3SA under non-secreting conditions. **(B)** The T3SA and the attached translocon structure after contact with a host cell. The structure now forms a conduit from the bacterial cytoplasm to the host cell cytosol.

Further EM analysis with the *Salmonella* Inv-Spa needle complex has revealed a separate inner periplasmic rod structure inside and between the two large ring structures (Figure 2A) (Marlovits *et al.*, 2006; Marlovits *et al.*, 2004). In *Shigella*, MxiI is suggested to be the inner-rod building block (Sani *et al.*, 2007a), in analogy to its homolog PrgH in the *Salmonella* Inv-Spa needle structure (Marlovits *et al.*, 2006). MxiI (11 kDa) shares some features with MxiH, but contains ~14 additional residues at its amino-terminus. Together, these six Mxi proteins form the stable structural framework of the TTS apparatus to which the other components of the T3SA are attached (Figure 2A).

The T3SA machinery

The basal body may provide a scaffold for the assembly of the T3SA machinery, which consists of additional IM and cytoplasmic components. A set of proteins with the characteristics of α-helical integral membrane proteins is assumed to be present within the large IM ring (Figure 2A). In *Shigella* they consist of MxiA, Spa40, Spa9, Spa29 and Spa24 (Ghosh, 2004; Hueck, 1998; Parsot, 2005). These proteins are encoded in a single operon in the entry region on the *Shigella* virulence plasmid. Unlike the proteins that form the structural framework of the T3SA, these inner membrane proteins are highly conserved in sequence among all T3S systems. They are thought to assemble in a large hetero-oligomeric complex and function in the actual secretion process. Each of them is absolutely required for assembly of the T3S structure and secretion (Hueck, 1998).

Secretion is ATP-dependent and is powered by the ATPase Spa47, which is associated with the T3S apparatus at the cytoplasmic side (Jouihri *et al.*, 2003). A homolog of Spa47 from the T3S system of *P. syringae* is HrcN, which can form hexameric rings with a central opening of ~2-3.8 nm (Müller *et al.*, 2006; Pozidis *et al.*, 2003). The T3S ATPases have a conserved domain organisation, consisting of an amino-terminal domain, a central catalytic ATPase domain and a carboxy-terminal helical domain (Zarivach *et al.*, 2007). Other proteins that are associated with the cytoplasmic side of the TTSA include Spa33, Spa32, MxiK, MxiN and, presumably, Spa13 (Jouihri *et al.*, 2003; Magdalena *et al.*, 2002; Morita-Ishihara *et al.*, 2006; Schuch and Maurelli, 2001b; Tamano *et al.*, 2002). These cytoplasmic components are likely to be only peripherally attached (Figure 2A) and a cytoplasmic structure could so far only be visualised by EM with a very mild preparation protocol (Blocker *et al.*, 1999; Morita-Ishihara *et al.*, 2006).

Assembly of the T3SA

Transcription of the genes coding for the T3SA building blocks is activated only under conditions that resemble the environment of the human colon (Parsot, 2005). This allows restriction of the relatively expensive T3SA synthesis to the appropriate conditions. Assembly and operation of the T3S apparatus is a stepwise process. First the T3SA is assembled and translocators and early effectors are synthesised. The fully assembled T3SA remains inactive until contact with a host cell initiates secretion and translocation.

Components of the basal body and inner membrane proteins, such as MxiG, MxiJ, MxiM, MxiD, MxiA, Spa40, Spa28, Spa24 and Spa9 all have a typical amino-terminal signal peptide that ensures their recognition by the general Sec protein export pathway (Hueck, 1998). The Sec machinery is therefore probably involved in their insertion in the IM or transport to the periplasm (in the case of the secretin MxiD and the secretin pilot MxiM). Among the first proteins to be secreted by the nascent TTS apparatus are MxiI and MxiH, which subsequently assemble into the inner rod and the needle. Secretion of MxiI and MxiH requires the cytoplasmic T3SA components MxiK and MxiN, which interact with Spa47 and Spa33 (Jouihri *et al.*, 2003).

The assembly of the needle and the determination of its length is further controlled by Spa32, which is transiently associated with the T3S system (Magdalena *et al.*, 2002; Tamano *et al.*, 2002). Spa32 interacts with the IM component Spa40 (Magdalena *et al.*, 2002). A *spa32* mutant is not able to secrete proteins other than MxiH and MxiI, suggesting a substrate specificity switch function for Spa32. Comparative analysis with the homologous *Salmonella* Inv-Spa T3S system (Marlovits *et al.*, 2006) suggests that Spa32 senses the completion of the inner rod structure, which triggers the release of Spa32 from the T3SA through its secretion. The T3SA then no longer secretes MxiH and MxiI, but is ready to accept other substrates like translocators and effectors. The relevance of this model for the

Shigella T3S system still has to be confirmed experimentally.

The translocators

Three translocator proteins are necessary to form a pore in the plasma membrane of the host cell and maintain a link with the T3S needle (Ménard *et al.*, 1993). In *Shigella*, IpaB and IpaC form the actual translocation pore and IpaD probably links this pore to the T3SA (Figure 2B). EM images of the needle tip revealed a distinct tip structure, which contains IpaD (Sani *et al.*, 2007b) and which assembles under non-secreting conditions. The IpaD crystal structure has been solved and reveals a protein with two domains, sharing two long helices (Johnson *et al.*, 2007). Details of the crystal packing led to the proposal that IpaD forms a five-subunit helical ring, with the same helical pitch as the needle, allowing it to form a crown on top of the T3SA (Figure 2A). The carboxy-terminal domain and the two long helices are involved in forming this putative pentameric structure, which may close off the needle under non-secreting conditions (Johnson *et al.*, 2007).

IpaB interacts with IpaD to block T3S secretion under non-secreting conditions (Ménard *et al.*, 1994a) and this IpaD-IpaB complex is probably present at the T3SA needle tip (Olive *et al.*, 2007; Veenendaal *et al.*, 2007). Contact with a host cell induces activation of the T3SA (Ménard *et al.*, 1994a; Watarai *et al.*, 1995). The presence of IpaD and IpaB at the tip of the needle (Espina *et al.*, 2006; Sani *et al.*, 2007b) may suggest a mechanism in which IpaD serves as a platform for insertion of IpaB and IpaC into the membrane (Espina *et al.*, 2006; Picking *et al.*, 2005; Watarai *et al.*, 1996). Sequence analysis indicates that IpaB and IpaC have hydrophobic regions which may facilitate this membrane insertion. It has been established *in vitro* that they interact with each other and that they may permeate membranes, further supporting their function as pore components (De Geyter *et al.*, 1997; De Geyter *et al.*, 2000; Ménard *et al.*, 1994b). Cholesterol facilitates the insertion of the translocation pore in the host cell membrane and it is not impossible that cholesterol itself also triggers the secretion of these translocators (Hayward *et al.*, 2005; van der Goot *et al.*, 2004). The IpaB-IpaC pore has been estimated to have an inner diameter of ~2-3 nm, similar to the inner diameter of the needle of the T3S apparatus (Blocker *et al.*, 1999) (Figure 2B).

The Effectors

The *Shigella* T3SA injects ~25 different effectors in the cell cytosol. The effectors can be divided into two groups, distinguished by the way their synthesis and secretion are controlled (Le Gall *et al.*, 2005). The first group are effectors with, presumably, a role in the entry mechanism. They are synthesised along with the T3SA components, and in fact, several of these effectors are encoded in the entry region. These effectors are stored in the bacterial cytoplasm under non-secreting conditions, many of them in association with specific T3S chaperones. Their storage allows the rapid injection of these early effectors after contact with a host cell has activated the T3SA. The second group of effectors is not produced until the T3SA is activated (Demers *et al.*, 1998) and is therefore thought to have functions later in the invasion cycle of *Shigella*.

Interestingly the translocator proteins IpaB and IpaC are multifunctional proteins with additional roles in the *Shigella* invasion mechanism. Specifically, IpaC elicits actin polymerisation in the host cell. The growing actin filaments induce membrane ruffles, which ultimately engulf the bacterium (Tran Van Nhieu *et al.*, 1999). Several early effectors work synergistically to modulate *Shigella* entry by targeting host signalling and cytoskeletal reorganisation pathways (Parsot, 2005). IpgB1 localises to the plasma membrane, where it mimics the role of RhoG and triggers a signalling cascade that leads to membrane ruffling (Alto *et al.*, 2006; Handa *et al.*, 2007; Ohya *et al.*, 2005). VirA induces microtubule destabilisation, promoting Rac1 activity and membrane ruffling (Yoshida *et al.*, 2002). IpgD converts phosphatidyl-inositol 4,5-bisphosphate into phosphatidyl-inositol 5-phosphate at the site of entry, which uncouples the plasma membrane from the cellular cytoskeleton, thereby facilitating local plasma membrane ruffling (Niebuhr *et al.*, 2002), but also activates a signalling pathway which prolongs host cell survival during *Shigella* infection (Pendaries *et al.*, 2006). IpaA localises to the plasma membrane close to the site of entry and specifically binds and activates

vinculin, which subsequently caps actin filaments, aiding in the shaping of membrane ruffles into an entry structure that leads to uptake of *Shigella* (Bourdet-Sicard *et al.*, 2000; DeMali *et al.*, 2006; Hamiaux *et al.*, 2006; Ramarao *et al.*, 2007).

Some of the other effectors are thought to have roles in post-invasion stage of infection. IcsB contributes to intracellular survival and cell-to-cell spreading by inhibiting a host autophagy defence mechanism (Ogawa *et al.*, 2003; Ogawa *et al.*, 2005). OspE2 is necessary to maintain an intact host cell morphology, which aids cell-to-cell spreading (Miura *et al.*, 2006). Also, several effectors interfere with immune response mechanisms. OspF is a phosphothreonine lyase (Li *et al.*, 2007), which inactivates MAP kinases (Kramer *et al.*, 2007; Pendaries *et al.*, 2006), preventing histone H3 phosphorylation. This blocks activation of a subset of NF-κB-responsive genes, leading to compromised recruitment of polymorphonuclear leukocytes to infected tissues (Arbibe *et al.*, 2007). OspG is a protein kinase which interferes with innate immune responses by targeting ubiquitin-conjugating enzymes (Kim *et al.*, 2005). IpaH9.8 binds to a splicing factor U2AF(35), causing a reduction in production of certain inflammatory factors (Okuda *et al.*, 2005).

Regulatory processes

Regulation of the T3SA activity is one of the key features of the *Shigella* T3S system. Examples of regulation include the restriction of T3SA assembly to favourable conditions, the switch of secretion specificity to translocators and effectors after completion of the needle, and activation of the T3SA upon contact with a host cell. In addition, transcription and secretion of a set of ~ twelve effectors is only initiated after the T3SA has been activated (Kane *et al.*, 2002; Le Gall *et al.*, 2005; Parsot *et al.*, 2005). Activity of the T3SA can also be switched off again about an hour after cell invasion and reactivated at later stages (Demers *et al.*, 1998).

Regulatory mechanisms are highly specific for a certain T3S system. The ability to regulate when and which proteins are secreted or translocated is one of the main factors that explain the successful adaptation of T3S systems by pathogenic and symbiotic bacteria to their specific infection strategy.

The process of secretion in the T3S system

The understanding of events that take place during secretion of a protein by the T3SA is very limited. The narrow size of the central channel through the needle complex requires proteins to be unfolded during export. Arguably, secretion substrates should be recognised and interact with components of the T3S machinery on the cytoplasmic side of the T3SA. Spa33 has e.g. been reported to interact with several secretion substrates (IpaB, IpaC, IpgB1, VirA) (Morita-Ishihara *et al.*, 2006), and MxiK may interact with IpaB (Page *et al.*, 2001). However, the significance of these interactions is at present not clear. To complicate matters further, the T3SA ability to switch substrate specificity from needle components to translocators and effectors, indicates that substrate recognition can be modulated. The associated T3S ATPase, Spa47 in *Shigella*, is however required for the secretion of all T3S substrates. In analogy to what is proposed for other T3S ATPases (Akeda and Galán, 2005; Sorg *et al.*, 2006; Thomas *et al.*, 2004), Spa47 can be anticipated to unfold secretion substrates and initiate them into the T3SA.

All secretion substrates of the system are recognised by a signal which resides within their amino-terminal 10-30 residues (Ghosh, 2004). This notion is supported by deletion experiments, which have shown that proteins required their native amino-terminus to be recognised as a secretion substrate. Other proteins were secreted after fusion to the amino-terminal residues of a native T3SA substrate (Schesser *et al.*, 1996; Sory *et al.*, 1995). The nature of this secretion signal is uncertain. Unlike the signal peptide in the Sec pathway, which is a clearly recognisable hydrophobic sequence at the amino-terminus of the protein, no special requirements have been identified for the residues at the amino-terminus of the substrate. Analysis of amino-terminal sequences of effectors from different bacterial T3S systems has suggested enrichment in Ser, Thr, Ile and Asn and a paucity of Asp (Guttman *et al.*, 2002; Lloyd *et al.*, 2002). Alternating Ser and Ile residues at the amino-terminus are sufficient to enable secretion of YopE by the T3S apparatus (Lloyd *et al.*, 2002).

The lack of any specific motifs has spurred the suggestion that presentation of the T3S signal in

the form of flexible or disordered amino-terminal residues is a major determinant in the recognition process by the T3SA. Indeed, many structure determinations of secreted proteins in recent years have revealed that amino-terminal residues are not part of a well-ordered domain. The amino-terminal residues of the translocators LcrV (residues 1-27) (Derewenda *et al.*, 2004), BipD (residues 1-34) (Erskine *et al.*, 2006; Johnson *et al.*, 2007) and IpaD (residues 1-39) (Johnson *et al.*, 2007), the needle subunits MxiH (residues 1-14) (Deane *et al.*, 2006) and BsaL (residues 1-7, with residues 8-27 partly ordered) (Zhang *et al.*, 2006), and the effectors YopM (residues 1-33) (Evdokimov *et al.*, 2001a), HopF1 (residues 1-28) (Singer *et al.*, 2004), AvrB (residues 1-27) (Lee *et al.*, 2004) and AvrPto (residues 1-28) (Wulf *et al.*, 2004) were all found to be disordered. In addition, the terminal regions of flagellar T3S substrates, the flagellar hook and filament building blocks, are disordered (Daughdrill *et al.*, 1997; Vonderviszt *et al.*, 1992). An apparent exception to these observations is however formed by the amino-terminal residues of YopH, which form a well-ordered α-helix in two independently determined crystal structures as well as in its solution structure, and is an integral part of the well-structured YopH amino-terminal domain (Evdokimov *et al.*, 2001b; Khandelwal *et al.*, 2002; Smith *et al.*, 2001). However, this protein can only be secreted in the presence of its chaperone SycH, and its amino-terminus may well be disordered and accessible in association with SycH (see below). An accessible, poorly ordered amino-terminus with a bias towards certain amino acids may therefore be sufficient to direct a protein to the T3SA.

An alternative hypothesis has been put forward that suggests that the secretion signal is contained in the mRNA (Anderson and Schneewind, 1997) and directs secretion co-translationally. This model is controversial (see Cornelis, 2003 and Ramamurthi and Schneewind, 2003) and moreover has little relevance for the following discussion of the essentially post-translational function of T3S chaperones. Therefore it will not be treated in this text.

T3S Chaperones

Due to the regulation of T3SA activity by external signals, several secretion substrates are stored in the bacterial cytoplasm under non-secreting conditions. Several of them depend on dedicated T3S chaperones for their proper secretion. These chaperones bind to their protein before their secretion. Upon secretion, the protein complex dissociates; the secretion substrate is exported and the chaperone remains behind in the bacterial cytoplasm. The T3S chaperones may be divided into three different classes on the basis of the type of protein that they assist (Cornelis, 2006; Parsot *et al.*, 2003); effector proteins (class I), pore-forming translocator proteins (class II), and T3S-exported parts of the T3S apparatus (class III).

Class III T3S chaperones

The extracellular helical components of structures like the needle components of the T3S apparatus or the flagellar hook and filament subunits must be prevented from premature polymerisation in the bacterial cytoplasm. The composite group of class III T3S chaperones bind these components and mask the regions that drive polymerisation. In the flagellar T3S system FlgN is the chaperone for the two homologous hook-associated proteins FlgK and FlgL (Fraser *et al.*, 1999), FliS is the chaperone for the flagellin FliC (Auvray *et al.*, 2001) and FliT is the chaperone for the flagellar cap protein FliD (Fraser *et al.*, 1999). These chaperones bind to the carboxy-terminus of their cognate flagellar components in a 1:1 stoichiometry and stabilise them, presumably by preventing their polymerisation, which is dependent on the carboxy-terminus of these proteins (Auvray *et al.*, 2001; Minamino and Namba, 2004). Flagellar chaperones are four-helix bundle proteins (Badger *et al.*, 2005; Evdokimov *et al.*, 2003) of ~120-140 residues. The crystal structure of the complex between *Aquifex aeolicus* FliS and the C-terminal residues 464-518 of FliC has been determined and shows that FliS binds the carboxy-terminus of FliC in an extended conformation (Figure 3A). The FliC polypeptide binds in grooves on the surface of the FliS chaperone, which provides specificity and affinity to the interaction (Evdokimov *et al.*, 2003). Other flagellar chaperones are predicted to have the same fold as FliS and associate with their cognate flagellar components in a similar way.

Needle subunits of a subset of the pathogenic injectisomes also associate with small cytoplasmic proteins. In the cytoplasm of *Pseudomonas aeruginosa* the needle component PscF is sequestered in a 1:1:1 complex with PscE and PscG (Quinaud *et al.*, 2005). A similar process occurs in *Yersinia* spp. where the needle subunit YscF is bound by YscE and YscG (Cornelis, 2006; Wilharm *et al.*, 2007). In *Shigella*, MxiH does not seem to require such proteins for its assembly (Jouihri *et al.*, 2003).

Finally, the needle of the EPEC T3SA contains a long sheath-like extension formed by EspA, which connects the needle and the pore-forming translocators (Knutton *et al.*, 1998). In the cytoplasm, EspA is chaperoned by the 12 kDa chaperone CesA, which prevents the polymerisation of EspA (Creasey *et al.*, 2003b). The structure of the CesA:EspA complex has been solved (Yip *et al.*, 2005), revealing that EspA associates with CesA in a 1:1 complex (Figure 3B). The structure of CesA consists of three α-helices, which clamp two α-helices located at the amino- and carboxy-terminus of EspA. These EspA regions are probably involved in intersubunit contacts in the EspA filament, therefore the structure explains how CesA prevents EspA polymerisation (Yip *et al.*, 2005). CesA is also remarkable for its basic pI, which is unlike most other T3S chaperones.

All known examples of class III T3S chaperones seem to share their function as polymerisation-preventers. In each case they cover the region of their partner that promotes polymerisation or oligomerisation. In this way, class III T3S chaperones maintain their substrates in a secretion-competent state. Besides their overall α-helical folds and small size, little seems to be conserved between the different types of class III T3S chaperones.

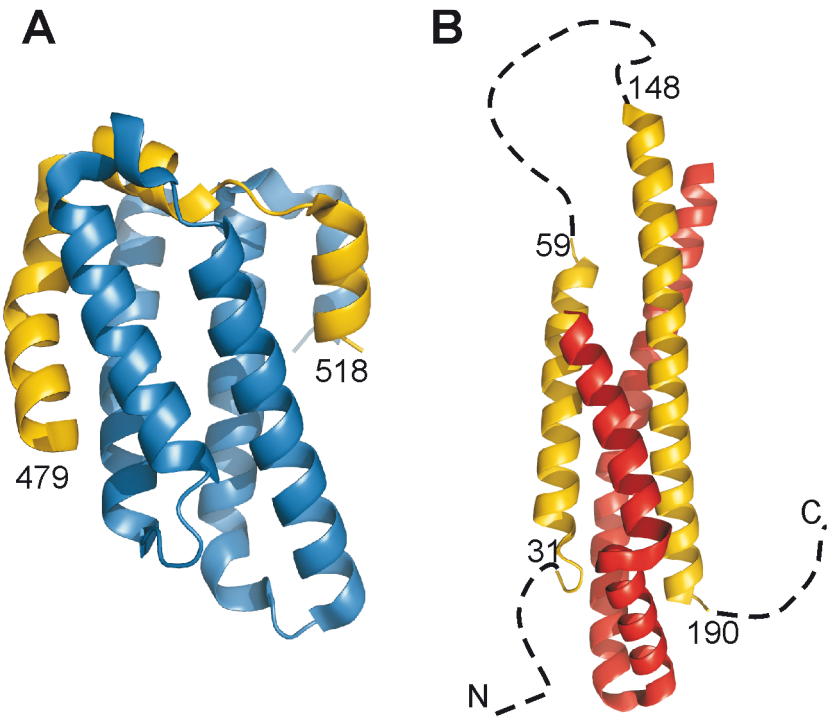

Figure 3: Structures of class III T3S chaperones in complex with their binding partner. **(A)** Crystal structure of FliS (blue) in complex with FliC(464-518) (gold) (pdb code 1ORY). **(B)** Crystal structure of EPEC CesA (red) in complex with EspA (gold) (pdb code 1XOU). Disordered parts of EspA are indicated with dashed lines. Figures 3, 4 and 5 were created with Pymol (http://pymol.sourceforge.net).

Class II T3S chaperones

Class II T3S chaperones assist the two pore-forming translocators (Parsot *et al.*, 2003). They are all small (~18 kDa) proteins with an acidic pI and are predicted to have an α-helical fold. In particular, their central part contains three so-called tetratricopeptide repeats (TPRs) (Edqvist *et al.*, 2006; Pallen *et al.*, 2003), where each TPR constitutes a helix-turn-helix structural motif. Class II T3S chaperones have a partitioning function, as e.g. in the absence of their class II T3S chaperone IpgC, the translocators IpaB and IpaC prematurely associate in the *Shigella* cytoplasm (Ménard *et al.*, 1994b). In the IpaB:IpgC and IpaC:IpgC complexes, the translocators are stable and protected during storage in the cytoplasm (Ménard *et al.*, 1994b; Page *et al.*, 1999). Yeast two-hybrid (Y2H) screening and deletion experiments have been used to identify a chaperone-interacting region (Harrington *et al.*, 2003; Page *et al.*, 1999). For both proteins this has been located relatively close to the amino-terminus of the translocator; residues 58-72 for IpaB (Page *et al.*, 2001) and residues 73-122 (Page *et al.*, 2001) or residues 50-80 (Harrington *et al.*, 2003) for IpaC. Association of IpgC to these approximate regions is probably sufficient to prevent any premature associations involving these translocators.

IpgC also acts as a transcriptional co-activator, as it is involved in the transcriptional activation of late effector genes (Mavris *et al.*, 2002a; Mavris *et al.*, 2002b; Parsot *et al.*, 2005). Under non-secreting conditions, IpgC is sequestered in complexes with IpaB and IpaC. After activation of the T3SA, IpaB and IpaC are exported and IpgC remains behind in the Shigella cytoplasm. IpgC is then available to interact with the transcriptional activator MxiE, which is a member of the AraC family of transcriptional regulators. MxiE binds to the promotor regions of a set of twelve effector genes, activating their transcription during entry (Kane *et al.*, 2002; Le Gall *et al.*, 2005; Parsot *et al.*, 2005). Through this mechanism the activity of the T3SA is coupled to transcription of effector genes.

Class I T3S chaperones

Class I T3S chaperones assist effectors and are small (12-18 kDa), dimeric proteins with an acidic pI. *Yersinia* YopE (*Yersinia* outer protein E) was the first effector for which a dedicated chaperone, SycE (Specific Yop chaperone E), was described. Deletion of the gene coding for SycE resulted in the same phenotype as deletion of the effector gene, indicating the importance of the chaperone. Absence of SycE resulted in a modest decrease of the intrabacterial levels of stored YopE under non-secreting conditions (Forsberg and Wolf-Watz, 1990; Lloyd *et al.*, 2001; Wattiau and Cornelis, 1993) and impairment in secretion of stored YopE after activation of the T3SA (Lloyd *et al.*, 2001). YopE which was synthesised under conditions in which the T3SA was active, was however secreted normally. At the same time, stability and secretion of other *Yersinia* effectors were unaffected (Forsberg and Wolf-Watz, 1990; Wattiau and Cornelis, 1993). SycE tightly bound to YopE, which apparently facilitated its storage and subsequent secretion (Wattiau and Cornelis, 1993).

Several class I T3S chaperones were subsequently identified. In each case they fulfilled similar roles in maintaining cytoplasmic stability and secretion of their effectors (i.e., in the absence of the T3S chaperone, the secretion of stored effectors was diminished), although not in each case to the same extent. Class I T3S chaperones can be found in many T3S systems, encompassing both plant and animal pathogens and symbionts and the T3S systems in proteobacteria and *Chlamidia* spp., but are, not surprisingly, absent in flagellar T3S systems. The number of identified class I T3S chaperones varies between the T3S systems of different pathogens, with three in the *Shigella* Mxi-Spa T3S system (Parsot, 2005), five in the *Yersinia* Ysc system (Letzelter *et al.*, 2006), three in the *Salmonella* Inv-Spa system (Bronstein *et al.*, 2000; Darwin *et al.*, 2001; Fu and Galán, 1998), one (putative) in the *Salmonella* Ssa system (Yu *et al.*, 2004) and three in the EPEC LEE (locus of enterocyte effacement) encoded system (O'Connell *et al.*, 2004; Pallen *et al.*, 2005). Among different pathovars of the plant pathogen *Pseudomonas syringae* seven class I T3S chaperones can be identified (Lindeberg *et al.*, 2005). The number of known or putative class I T3S chaperones is invariably smaller than the number of known or putative effectors. Their wide distribution and general importance indicate that class I T3S chaperones form an integral part of the operation of T3S systems that communicate with eukaryotic cells.

Class I T3S chaperones have been further subdivided into class IA T3S chaperones, which are specific for one effector (but see below) and class IB, which function as T3S chaperones for multiple effectors. The genes encoding class IA T3S chaperones are often, but not always, adjacent to the gene encoding their effector. In *Shigella*, the class IA chaperones IpgA and IpgE are encoded next to the genes for their respective effectors IcsB and IpgD (Buchrieser *et al.*, 2000). In contrast, class IB T3S chaperones, like Spa15 in *Shigella*, are encoded in operons also coding for structural components of the T3S apparatus (Parsot *et al.*, 2003). In addition a subclass IC will be distinguished here, consisting of heterodimeric class I T3S chaperones. The *Shigella* T3S system contains no representative of this last subclass.

Structural and functional properties of class I T3S chaperones

The involvement of secretion chaperones in the T3S system represents an additional level of complexity in an already highly sophisticated bacterial export system. The function of this added complexity is not entirely clear, especially since several effectors apparently do not require dedicated T3S chaperones. Study of the function of class I T3S chaperones in the context of the T3S system is severely complicated by the potential pleiotropic effects of mutations or alterations in T3S components. It has been proposed that class I T3S chaperones are required to maintain stored effectors in a secretion-competent state. Alternatively, it has also been proposed that class I T3S chaperones actively target their effectors to the T3SA and aid in their secretion.

Expansion of the amount of structural information about class I T3S chaperones and about their association with effectors would be able to shed more light on the functioning of these proteins. Furthermore, sequence similarities between class I T3S chaperones, in particular between the different subclasses, are very low. This means that structural information of each subclass is necessary to determine whether structural aspects of chaperoning are the same in each subclass.

In recent years, many structural and functional studies have been carried out on class I T3S chaperones. Even though they have far from resolved all the questions regarding their function, they have put the knowledge of these chaperones in a structural perspective. This has led to a better understanding of the functions and modes of action of the class I T3S chaperones, and the similarities and differences between the subclasses. These studies are reviewed here, including the results of the studies performed with Spa15 which are presented in detail in the following chapters of this thesis.

Class IA T3S chaperones have a conserved structure

Class I T3S chaperones all have similar characteristics, but generally have a very low sequence similarity among each other. In recent years, the crystal structures of several class IA T3S chaperones have been determined, which established that, despite the low sequence similarity, all class IA T3S chaperones have the same novel fold (Birtalan and Ghosh, 2001; Birtalan *et al.*, 2002; Büttner *et al.*, 2005; Evdokimov *et al.*, 2002; Locher *et al.*, 2005; Luo *et al.*, 2001; Singer *et al.*, 2004; Stebbins and Galán, 2001; Trame and McKay, 2003). The chaperones are dimeric, and each subunit has a compact fold, consisting of a five-stranded β-sheet surrounded by three α-helices. The central β-sheet and the carboxy-terminal α-helix α3 are structurally well-conserved (Figure 4A). This fold is distinctly different from class III T3S chaperones of known structure and does not contain TPRs as predicted for class II T3S chaperones.

The loop region connecting α1 and strand β1 forms an exception to the generally compact and rigid structure. This loop is partly disordered in the SigE structure (Luo *et al.*, 2001), and follows a different course in each of the other chaperone structures, sometimes possibly due to crystal contacts. Furthermore, while some chaperones end immediately after last α-helix α3, others have an additional 5-30 residues, which are probably disordered (Birtalan *et al.*, 2002; Rumpel *et al.*, 2005; Trame and McKay, 2003). Residues from this C-terminal tag of CesT have been implicated in T3SA targeting (Thomas *et al.*, 2005). Overall, however, class IA chaperones are remarkably conserved (Figure 4A), which points towards a conserved function of these chaperones.

Class IA chaperones bind to a defined region of their effector

Class IA T3S chaperones generally associate with an amino-terminal region of their effector located next to the amino-terminal secretion signal. The amino-terminal secretion signal itself is dispensable for binding. Deletion experiments identified a minimal SycE-binding region between residues 15-50 of YopE, with additional interactions of YopE residues 51-77 (Boyd *et al.*, 2000; Woestyn *et al.*, 1996). Interestingly, deletion of the chaperone-binding region (CBR) also removed the requirement for SycE (Boyd *et al.*, 2000). Similar results were obtained with SycH:YopH (residues 20-70) (Woestyn *et al.*, 1996), SycO:YopO (20-78) (Letzelter *et al.*, 2006), EPEC Cest:Tir (~ 1-100) (Abe *et al.*, 1999), CesT:Map (~ 1-101) (Creasey *et al.*, 2003a) and *Salmonella* SicP:SptP (~ 15-100) (Fu and Galán, 1998).

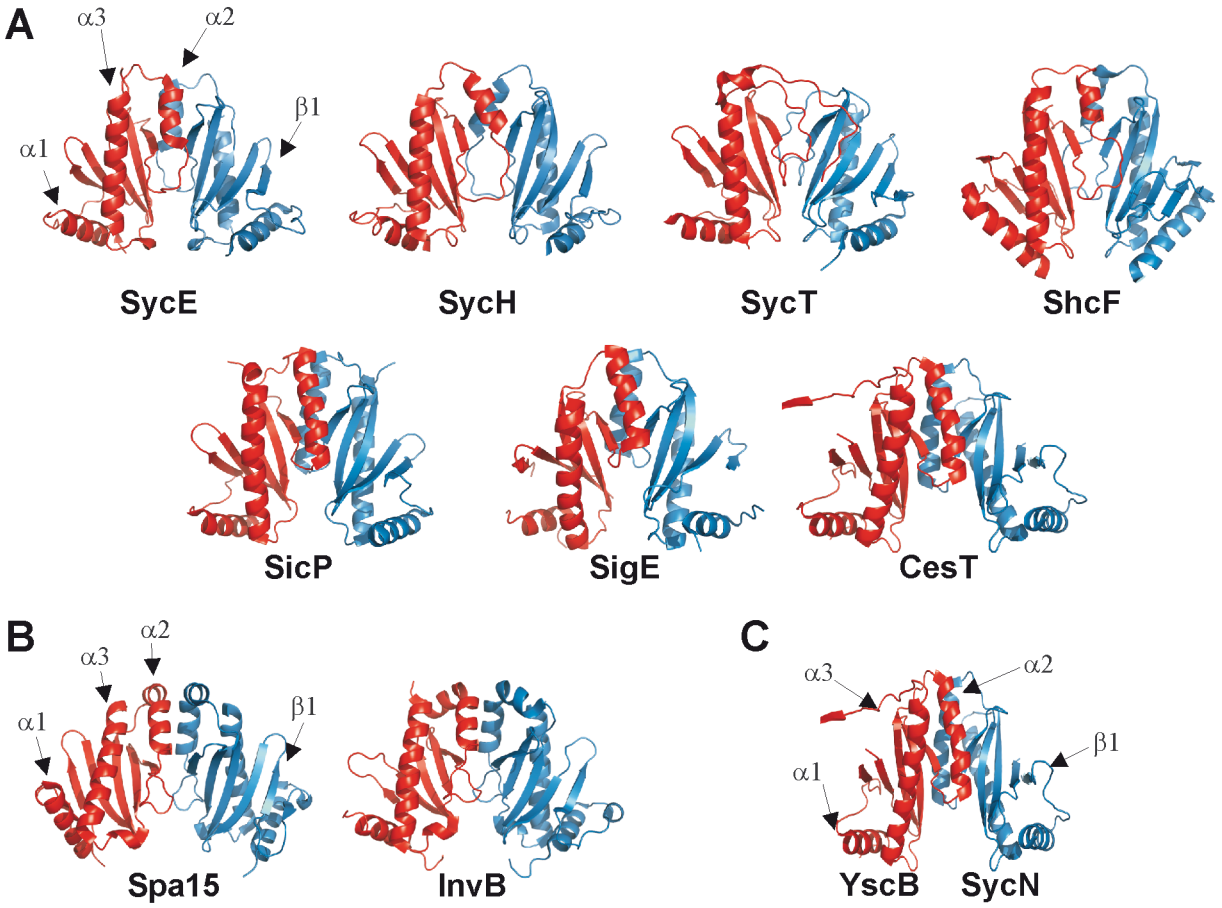

Figure 4: Structures of class I T3S chaperones. **(A)** Structures of the class IA T3S chaperones SycE, SycH, SycT, ShcF, SicP, SigE and CesT, respectively. The CesT structure was corrected for domain swapping as suggested by Luo *et al.* (2001) **(B)** Structures of the class IB T3S chaperones Spa15 and InvB and **(C)** Structure of the class IC heterodimeric T3S chaperone formed by SycN (blue) and YscB (red). Helices α1, α2 and α3 and strands β1 are indicated.

Class IA T3S chaperones do not affect the folding of the region carboxy-terminal of the CBR. Enzymatic activity could be detected for the carboxy-terminal effector domains of SopB (Luo *et al.*, 2001), YopE (Birtalan *et al.*, 2002), SptP (Akeda and Galán, 2005), YopH (Neumayer *et al.*, 2004) and YopO (Letzelter *et al.*, 2006) while they were associated with their cognate chaperones, and the EPEC effector Tir was still able to interact with its binding partners α-actinin and intimin in the presence of CesT (Goosney *et al.*, 2000; Luo *et al.*, 2001). This also suggests that the CBR forms a functionally and structurally distinct part of the effector.

Detailed information about the binding of class IA T3S chaperones with the CBR of their effector is available for two chaperone:CBR complexes. The crystal structures of SycE in complex with YopE(23-78) (Birtalan *et al.*, 2002), and of SicP in complex with SptP(36-139) (Stebbins and Galán, 2001) have been solved (Figure 5A). In both cases the CBR adopts an extended conformation and wraps around the chaperone dimer. Specifically, residues located at the amino-terminal end of the CBR adopt a β-strand conformation and complement the β-sheet of the chaperone, adjacent to strand β1. A hydrophobic side chain at one end of this extra β-strand occupies a deep hydrophobic pocket on the chaperone surface. Furthermore, a downstream helical part of the effector polypeptide is bound in a groove formed by the exposed side of the chaperone β-sheet. The effector chain then continues around the surface chaperone dimer and the remainder of the effector chain occupies the corresponding binding sites on the other chaperone subunit in a similar fashion, even though the primary sequence of the effector in this region is decidedly different (Birtalan *et al.*, 2002; Stebbins

and Galán, 2001). The cross-over region of SptP folds into additional helices (Stebbins and Galán, 2001), which make the CBR of SptP almost twice as long as the CBR of YopE.

Overall, a considerable amount of surface is buried in the chaperone:CBR interface, leading to a high affinity between chaperone and effector; the SycE:YopE complex has e.g. a K_D of 0.3 nM (Cheng and Schneewind, 1999). The asymmetric binding, in which regions of the CBR which differ in sequence occupy identical sites on the chaperone surface, does however indicate that the individual effector-binding sites are not highly specific. Therefore it seems that specificity is in part achieved by the linear combination of binding elements leading to the large buried surface area. A further point is that the distribution of polar and hydrophobic residues across the surface of the chaperones is not conserved. This probably also contributes to the specificity of the chaperone:CBR interaction. The similar structures of these two complexes suggested that the general mode of binding between class IA T3S chaperones and their CBRs was conserved.

A general conservation of class IA T3S chaperone:CBR interactions was further supported by examination of the structures of other chaperones (Büttner *et al.*, 2005; Locher *et al.*, 2005; Luo *et al.*, 2001; Phan *et al.*, 2004; Singer *et al.*, 2004). Despite the limited sequence conservation, the surface of each of them displayed similar grooves as SicP and YopE, which would allow binding of similar structural elements of the CBR. A further crystal structure of a complex between SycH and YscM2 also seemed to confirm the importance of the β-strand binding groove and the curved β-sheet in effector binding. However, the experimental data on this particular complex were not of sufficient quality to allow further conclusions to be drawn (Phan *et al.*, 2004). Furthermore, several mutational studies indicated the general relevance of hydrophobic as well as polar interactions for the stability of the chaperone:CBR interaction (Knodler *et al.*, 2006; Lilić *et al.*, 2006; Schesser *et al.*, 1996).

Class IB T3S chaperones have the same fold as class IA T3S chaperones

Class IB T3S chaperones have so far been characterised in the *Shigella* T3S system and in the *Salmonella* Inv-Spa T3S system. The interaction of the *Shigella* class IB T3S chaperones Spa15 with multiple effectors, including IpaA, IpgB1 and OspC3, was initially suggested by Y2H screening (Page *et al.*, 2002). Whereas IpaA and IpgB1 are both encoded in the entry region of the *Shigella* virulence plasmid, OspC3 is encoded in a different locus on the virulence plasmid (Buchrieser *et al.*, 2000). Spa15 did not affect the stability of IpaA, but was required for the secretion of stored IpaA upon activation of the T3SA. In contrast, production or secretion of IpgB1 could not be detected in the absence of Spa15 (Page *et al.*, 2002). Recent studies have established that Spa15 is also the chaperone of the secreted anti-activator OspD1. Therefore Spa15 is involved in the regulatory network that controls transcription of late effector genes in *Shigella* (Parsot *et al.*, 2005).

Spa15 was the first class I T3S chaperone for which it was known that it interacted with multiple effectors. This, together with its exceptional genetic location provided the original basis of its distinction as a separate subclass IB T3S chaperone (Page *et al.*, 2002; Parsot *et al.*, 2003). Subsequently InvB (33 % sequence identity to Spa15) was found to be the class IB T3S chaperone of multiple *Salmonella* effectors, including at least SipA (Bronstein *et al.*, 2000), SopA (Ehrbar *et al.*, 2004), SopE (Lee and Galán, 2003) and SopE2 (69 % seq. identity to SopE) (Ehrbar *et al.*, 2003). T3S gene clusters of five other bacteria encode clear Spa15 homologues, which are InvB from *Chromobacterium violaceum* (Betts *et al.*, 2004) (30 % sequence identity), BsaR (InvB) from *Burkholderia* spp. (Attree and Attree, 2001) (25 % sequence identity), YsaK from the Ysa T3S system of *Y. enterocolitica* (Haller *et al.*, 2000) (24 % sequence identity) and InvB from *Sodalis glossinidius* (Dale *et al.*, 2001) (25 % sequence identity). In addition, the fish pathogen *Edwardsiella ictaluri* contains a Spa15 homolog (15 % sequence identity) on a small plasmid, i.e. outside a T3S gene cluster (Fernandez *et al.*, 2001). This level of similarity is remarkable, considering the overall low similarity among class IA T3S chaperones.

Spa15 and other class IB T3S chaperones share some characteristics with class IA T3S chaperones, but no significant primary sequence similarity could be detected. Therefore it was not known whether class IB T3S chaperones belonged to the same structural class as class IA T3S chaperones. The crystal structure of Spa15 (chapter 2) revealed that Spa15 has a similar fold as class

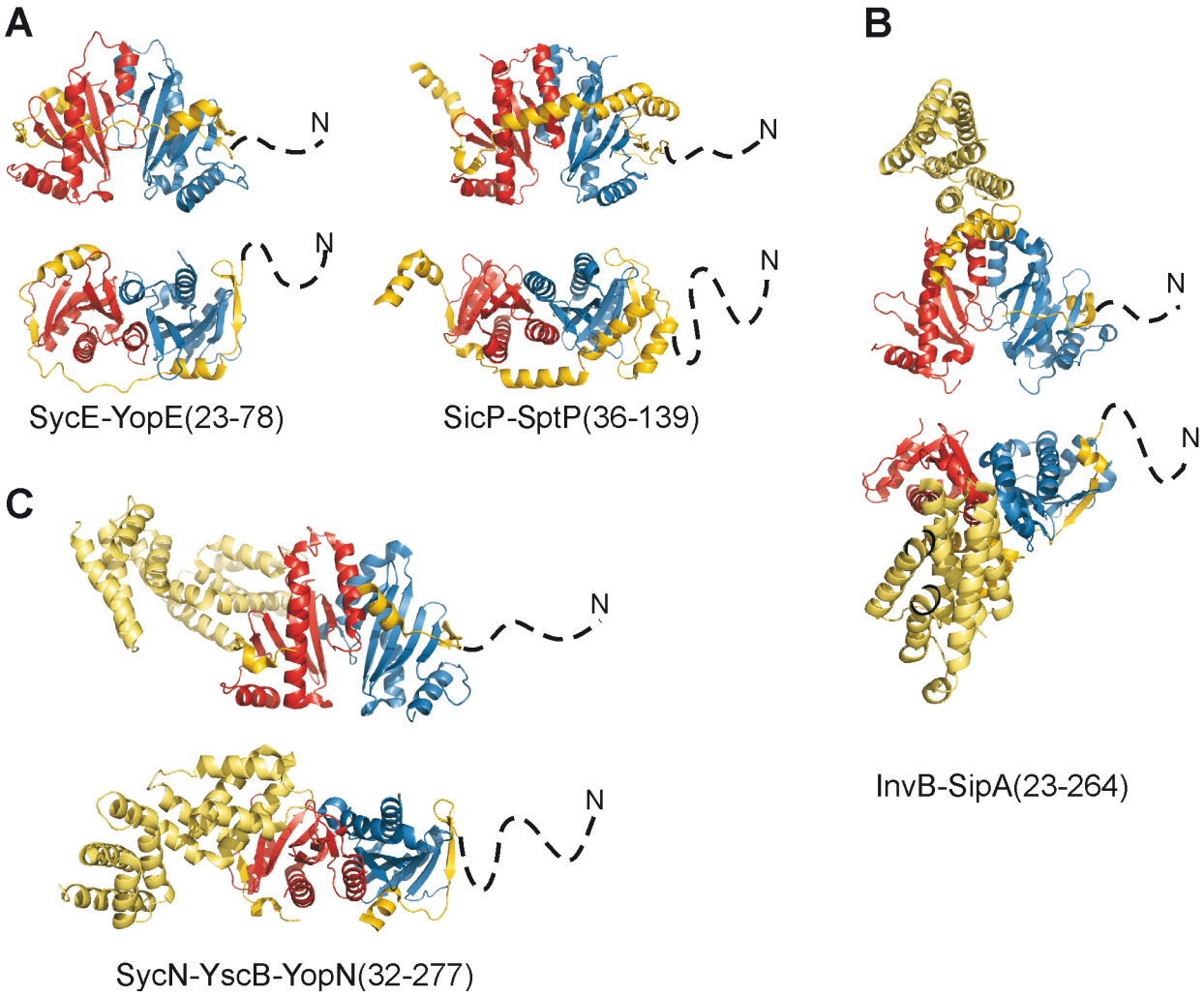

Figure 5: Orthogonal views of structures of class I T3S chaperones in complex with a part of their effector. **(A)** SycE:YopE(23-78) and SicP:SptP(36-139). The carboxy-terminal helix of SptP(36-139) is involved in a domain swap in the crystal, but also makes contact with a groove formed by the exposed side of the β-sheet of SicP. **(B)** InvB:SipA(23-264) **(C)** SycN:YscB:YopN(32-277). Chaperone subunits are in red and blue. Effector CBRs are in dark yellow and other effector domains in light yellow. For illustrative purposes, the approximate location of the amino-terminal secretion signals, which are missing from the crystal structures, are indicated schematically with a dashed line.

IA T3S chaperones of known structure (Figure 4B). This conclusively established the relationship between class IA and IB T3S chaperones. The structure of Spa15 however also revealed some remarkable differences with the other known class I T3S chaperone structures. The subunit orientation of Spa15 was decidedly different from class IA T3S chaperones like SycE and SicP and is characterised by a different orientation of ~ 30°. These differences provided a structural basis for the classification of Spa15 and its homologues as a separate class IB (chapter 2). The structure of *Salmonella* InvB (Figure 4B) confirmed that this divergent subunit orientation is a distinct feature of the class IB T3S chaperones (Lilić *et al.*, 2006).

Class IB T3S chaperones bind an amino-terminal CBR in their effector

How do class IB T3S chaperones bind their effectors? Deletion experiments indicated that residues 1-170 of IpaA and residues 22-190 of IpgB1 contain a Spa15 binding region (chapter 3). Furthermore, Y2H screening indicated that a common region in OspC3 from residues 23-141 was identified by Spa15 preys (Page *et al.*, 2002). Likewise, deletion experiments suggested that residues 27-264 of SipA (Lilić *et al.*, 2006) contain an InvB binding region, and fusion of residues 1-45 of SopA to a

reporter protein was sufficient for InvB binding and secretion of the reporter (Higashide and Zhou, 2006). This suggested that, like many class IA T3S chaperones, class IB T3S chaperones associate with an amino-terminal region in their effectors.

The structure of Spa15 gave indications that binding patches are conserved in class IA as well as IB T3S chaperones, suggesting similarities in binding mode between class IA and class IB T3S chaperones (chapter 2). The crystal structure of the InvB:SipA(22-268) complex (Lilić *et al.*, 2006) gives structural data for one class IB T3S chaperone:effector complex. SipA residues 23-43 adopt an extended conformation and engage in interactions with one of the chaperone subunits. The binding of this SipA fragment to InvB consists of the same two elements as class IA T3S chaperone:effector complexes. SipA donates a β-strand, and interacts with the β-sheet groove (Lilić *et al.*, 2006). However, unlike class IA T3S chaperone:effector complexes, the effector occupies these sites on only one of the two chaperone subunits. Indeed, the remainder of SipA (residues 52-268) forms a globular domain, which is necessary for the formation of a stable complex (Figure 5B). Two protruding helices that are part of this domain are involved in additional interactions with the chaperone (Lilić *et al.*, 2006) and can be considered as part of the CBR of SipA.

The heterodimeric SycN-YscB chaperone is the archetype of class IC T3S chaperones

Finally, the *Yersinia* heterodimeric chaperone SycN:YscB (Day and Plano, 1998) forms the archetype of another subgroup of class I T3S chaperones, which will be designated here as class IC. *Yersinia* SycN:YscB and its (putative) homologues (Pallen *et al.*, 2005; Sundin *et al.*, 2004; Yang *et al.*, 2007; Yu *et al.*, 2004) are the only known heterodimeric T3S chaperones. The YscB subunit forms an exception to the observation that T3S chaperones generally have a low pI with its predicted pI of 9.3 (Jackson *et al.*, 1998), while the SycN subunit has a predicted pI of 5.2 (Day and Plano, 1998). Similar pI distributions can be found in other (putative) class IC T3S chaperones. In addition, they are not the specific chaperone of an effector, but of a regulatory protein (YopN in *Yersinia* spp.), which blocks secretion under non-secreting conditions and is secreted upon activation of the T3S apparatus (Cheng *et al.*, 2001). The *Shigella* system contains a putative homolog of YopN (which is MxiC), but apparently it does not require a chaperone for its function.

The crystal structure of the heterotrimeric complex between SycN:YscB and YopN(32-277) has been solved (Schubot *et al.*, 2005). SycN and YscB form a dimer, which is similar to the structures of class IA and IB chaperones, but has its own variations (Figure 4C). The largest variations are at the dimer interface in the orientations of the α2 helices of both subunits. Unlike Spa15 and InvB, these structural variations do not affect the relative orientation of the chaperone subunits and the entire heterodimer superimposes well on class IA homodimeric T3S chaperones. YopN residues 32-76 have an extended conformation and wrap around the chaperone in a similar way as observed for class IA T3S chaperone-effector complexes, with YopN residues 32-50 contacting the SycN subunit whereas residues 51-57 and 65-85 are mainly involved in interactions with YscB (Figure 5C). The carboxy-terminal folded domain of YopN makes additional interactions with the exposed side of the YscN β-sheet (Schubot *et al.*, 2005).

Redefinition of class IA T3S chaperones

The original description of class IA T3S chaperones defined them as mono-specific chaperones. Recent data suggest that some class I T3S chaperones, which can be expected on sequence and profile similarity considerations to belong structurally to class IA, function as T3S chaperone for multiple effectors. CesT from EPEC specifically interacts with the unrelated effectors Tir and Map, and in the absence of CesT, secretion of other unrelated effectors (NleA, NleF, NleG/I, NleH, EspH, SepZ) is affected (Abe *et al.*, 1999; Creasey *et al.*, 2003a; Elliott *et al.*, 1999; Li *et al.*, 2006; Thomas *et al.*, 2005). Another example is HpaB from the T3S system of the plant pathogen *Xanthomonas campestris*, which is the chaperone of at least AvrBs1 and AvrBs3 (Büttner *et al.*, 2004), while in the absence of HpaB secretion of additional effectors is affected (Büttner *et al.*, 2006). The *P. syringae* ShcO1, ShcS1, and ShcS2 chaperones can all bind the effectors HopO1-1, HopS1, and HopS2 (Guo *et al.*, 2005; Kabisch *et al.*, 2005) and facilitate their secretion. Likewise, *Yersinia* SycE is able to interact not

only with YopE but also with YscM1 and YscM2 (Swietnicki *et al.*, 2004). The biological relevance of these last interactions is however not clear. In many of these cases, the requirement of the effector on the chaperone has not been precisely investigated. It is therefore not known whether these chaperones interact with these effectors through a defined amino-terminal CBR. In light of these examples, the class IA T3S chaperones should be redefined as chaperones able to serve one or more effectors. The comparatively high degree of sequence similarity, conserved gene location and divergent 3D structures of class IB chaperones still sets them apart as a separate class.

Structural and functional properties of the CBRs of effectors

The characterisation of chaperone:effector complexes has established that effector CBRs are defined amino-terminal effector regions, which function to bind a class I T3S chaperone in the bacterial cytoplasm. Every so-far determined chaperone:CBR structure has unique features, making it difficult to infer information about chaperone:effector interactions for which no structural information is available. Accordingly, it has thus far not been possible to identify putative CBRs in effectors based on their amino acid sequence or secondary structure profiles. A motif has been assembled on the basis of the structurally most conserved part of the CBRs, the β-strand (Lilić *et al.*, 2006), but this is not very specific and has little predictive power. Moreover, even though the 3D structures have thus far shown that CBRs wrap around both class IA and IC chaperone subunits and around only one of the subunits of a class IB chaperone, it cannot be excluded that this situation is different in individual cases. Although the structural differences between class IA, IC and IB chaperones are clear, these differences do not dictate structural differences between the CBRs. Structural characterisation of additional chaperone-effector complexes is therefore necessary to fully explore the conformational range of chaperone-bound CBRs.

CBRs lack any obvious sequence similarity among each other, are variable in size and (consequently) are generally not well-defined in those cases where no structural information is available. This variability makes it unlikely that the CBRs have a second conserved function in addition to the binding of their T3S chaperones. In individual cases an additional function could however be identified. Deletion experiments have shown that the region that functions as CBR in the bacterial cytoplasm may function as a targeting region after injection in the host cell cytosol. In YopE, the CBR directs the effector to the membrane, its proper location in the host cell (Krall *et al.*, 2004). Similarly, the CBR of YopO, YopT and the putative CBR of ExoS function in this way. SopE and SptP do not target the membrane directly, but associate with membrane-associated Cdc42 through their CBR (Cain *et al.*, 2004). Furthermore, the CBR of IpgB1 overlaps with a region that ensures colocalisation of IpgB1 with the plasma membrane (chapter3; Handa *et al.*, 2007). Other effectors are also targeted to distinct locations in the host cell, but the relation between this targeting and the CBR is not known in these cases.

The fact that several CBRs may represent targeting domains apparently does not imply a similar structure. The residues that comprise the minimal membrane-targeting region in YopE (54-75) correspond roughly to the last α-helix of the CBR (Krall *et al.*, 2004). Comparison of the available structures shows that this part of the CBR is the most widely divergent part of the CBRs thus far characterised structurally. The only readily detectable similarity among thus-far characterised CBRs appears to be in some physicochemical properties. The CBR of most effectors has more positively charged than negatively charged residues. This probably enhances affinity for the negatively charged chaperones. In the case where the CBR functions as MLD, such as for YopE, YopO and YopT (Letzelter *et al.*, 2006), as well as the putative CBR of the *P. aeruginosa* effector ExoS (Krall *et al.*, 2004), this may also increase membrane affinity.

The extended structure of the chaperone-bound form of the CBRs of SptP, YopE, YopN and SipA raises the question whether a CBR adopts a folded structure without its chaperone. A direct answer to this question is unfortunately not possible, as no structures are available for a single effector CBR in both its chaperone-bound and uncomplexed state. Structural information of the uncomplexed form of the CBR is available for two effectors, YopH (Evdokimov *et al.*, 2001b; Khandelwal *et al.*, 2002; Smith *et al.*, 2001) and HopF1 (Singer *et al.*, 2004), but in the case of HopF1 the CBR has not been

precisely defined. Comparison of these structures establishes that even though both proteins possess a well-structured amino-terminal α/β domain, they do not have the same fold. Comparison of the CBR of YopH identified by deletion experiments (18-71) (Sory *et al.*, 1995; Wattiau *et al.*, 1994) or limited proteolysis (31-92) (Phan *et al.*, 2004) with the structure of YopH(1-130) shows that these residues form an integral part of the amino-terminal domain. This suggests that the amino-terminal domain of YopH should be unfolded to associate with SycH. A complex can be formed by mixing purified SycH with the separately purified SycH partners YopH (residues 1-130 or full-length), or YscM1, or YscM2 (Neumayer *et al.*, 2004; Phan *et al.*, 2004; Wilharm *et al.*, 2003), suggesting that SycH can actively induce the local unfolding of the CBR of its three separate binding partners. More study is needed to answer questions about the structure of CBRs in general, but it is clear that unlike their dedicated chaperones, effector CBRs are structurally highly diverse.

The function of class I T3S chaperones

The T3S chaperones of classes IA, IB and IC all have the same fold, share a similar size, have a similar heart-shaped form, and all use the same characteristic features of their fold to bind effectors (Figure 4 & 5). Each subclass harbours chaperones with distinct structural properties, which are quite well-defined by the currently available structural information. The overall similarities between these secretion chaperones also suggest an overall similarity in function.

Class I T3S chaperones are necessary to secrete effectors at the required levels at the required moment. This can be understood by considering two basic properties of the chaperone:effector complexes, revealed by structural characterisation of several chaperone:effector complexes. First, there is good evidence that the amino-terminal residues which constitute the secretion signal are kept in an accessible state in the chaperone:effector complexes, thereby maintaining a secretion competent state. Second, the binding to aggregation-prone or relatively unstable CBRs is probably necessary to maintain significant concentrations of these effectors under non-secreting conditions, which explains the beneficial effect of class I T3S chaperones on stability of the effectors.

For each chaperone:effector pair the roles in effector stability or secretion seem to be relevant to different degrees. In some cases effector stability is predominant, like Spa15:IpgB1, InvB:SopE2 and SicP:SptP, in other cases the role of chaperones in secretion of stored effectors is more apparent, like Spa15:IpaA, InvB:SopE, SycH:YopH and SycE:YopE. Other effectors might not require class I T3S chaperones because they do not contain aggregation-prone targeting domains or are never produced in sufficient quantities to result in aggregation, are not stored in the bacterial cytoplasm before initiation of T3SA activity or contain a readily accessible secretion signal.

In the previous model class I T3S chaperones have a passive role during storage of effectors in the bacterial cytoplasm. It has also been suggested that class I T3S chaperones have a more important role in the actual secretion process and actively target their effectors to the T3SA. This model has gained popularity since it was found that CesT can co-purify with the EPEC T3S ATPase EscN from an EPEC lysate (Gauthier and Finlay, 2003), and that SicP can interact with the *Salmonella* T3S ATPase InvC, which is able to unfold SptP and disrupt the SicP-SptP interaction (Akeda and Galán, 2005). This last study suggested parallels between T3S ATPases and the function of AAA+ (ATPases associated with diverse cellular activities) machines (Ogura and Wilkinson, 2001; Sauer *et al.*, 2004). The ability of the T3S ATPases to organize in hexameric rings (Pozidis *et al.*, 2003), to disassemble protein complexes (Akeda and Galán, 2005) and to recognise and unfold protein substrates closely resembles activities of AAA+ disassembly machines (Dougan *et al.*, 2002; Sauer *et al.*, 2004). The similarities seem to extend to essential features of the substrate recognition process, because T3S systems and AAA+ machines recognise a poorly conserved set of unstructured short signals located at the amino terminus of the target proteins, which sometimes experience enhanced targeting due to a special adaptor protein. Class I T3S chaperones might function as an adaptor in the secretion of their effectors. The type III secretion process and the translocation of substrates into AAA+ machines might therefore share some mechanistic features (Akeda and Galán, 2005; Galán and Wolf-Watz, 2006; Wilharm *et al.*, 2007). A careful comparison with other systems in which such ATPases function may lead to a better understanding of these processes.

So far, it is not clear which selective advantage could be offered by the additional complexity of such an extra chaperone-mediated targeting mechanism, but it as been suggested that enhanced targeting may speed up T3S-mediated translocation or that selective targeting by T3S chaperones mediates an export hierarchy (Birtalan *et al.*, 2002; Feldman and Cornelis, 2003; Ghosh, 2004; Lilić *et al.*, 2006). This has however so far not been observed in native T3S systems. Furthermore, it is not clear what the targeting motif or additional secretion signal might be. The structural similarities between different class I T3S chaperone-effectors are very limited and mainly confined to the first β-strand motif which is immediately adjacent to the N-terminal secretion signal (Lilić *et al.*, 2006). However, no obvious conservation is present in the accessible residues of this part of the four known chaperone:effector structures. The β-motif may also be required to facilitate chaperone:effector dissociation. Indeed, the chaperone:effector interactions may be strong, but should be compatible with its final dissociation. Lateral force applied to the β-motif may lead to a relatively straightforward dissociation of the chaperone:effector complex.

Certain restraints on the size and shape of the chaperone-effector complexes may be in force, since the complexes have to access the component(s) of the T3SA, which facilitate dissociation and effector secretion, in the context of the fully assembled T3SA. The observed similarity in shape and the conserved features of effector binding could therefore also be consequences of such size and shape restraints.

The further development of a model for substrate recognition and secretion by T3S systems, in particular the spatial organisation of cytoplasmic components and the mechanistic aspects of substrate unfolding and disassembly, may allow a definitive answer to the questions regarding the function(s) of class I T3S chaperones.

Acknowledgements

C. Hamiaux and C. Parsot are thanked for useful discussions. Help of M. Bokhove in constructing Figure 1 is gratefully acknowledged.