Subcellular Biochemistry 83

J. Robin Harris
Jon Marles-Wright *Editors*

Macromolecular Protein Complexes

Structure and Function

Springer

Subcellular Biochemistry

Volume 83

Series editor
J. Robin Harris
University of Mainz, Mainz, Germany

The book series SUBCELLULAR BIOCHEMISTRY is a renowned and well recognized forum for disseminating advances of emerging topics in Cell Biology and related subjects. All volumes are edited by established scientists and the individual chapters are written by experts on the relevant topic. The individual chapters of each volume are fully citable and indexed in Medline/Pubmed to ensure maximum visibility of the work.

More information about this series at http://www.springer.com/series/6515

J. Robin Harris • Jon Marles-Wright
Editors

Macromolecular Protein Complexes

Structure and Function

Springer

Editors
J. Robin Harris
Institute of Zoology
University of Mainz
Mainz, Germany

Jon Marles-Wright
School of Biology
Newcastle University
Newcastle upon Tyne, UK

ISSN 0306-0225
Subcellular Biochemistry
ISBN 978-3-319-46501-2 ISBN 978-3-319-46503-6 (eBook)
DOI 10.1007/978-3-319-46503-6

Library of Congress Control Number: 2017934081

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The last few decades have seen an exceptional expansion of studies on protein molecules and protein complexes by both X-ray crystallography and high-resolution transmission electron microscopy, increasingly incorporating cryo-conditions (cryo-TEM). Advances in cryo-TEM instrumentation and methods now allow the analysis of dynamic structures to resolutions that were once only accessible by X-ray crystallographic analysis. These developments led to cryo-TEM being highlighted as the method of 2015 by Nature Methods[1].

In this book, we have brought together a number of studies of macromolecular complexes from what is now a very large internationally based discipline. These include the bacterial stressosome complex, the Type III CRISPR-Cas prokaryotic intracellular defense system, and the eukaryotic inflammasome complex, to mention but three of the chapters. While numerous topics could have been included in this volume, we have chosen to concentrate on soluble macromolecular complexes in this book. Future volumes in the *Subcellular Biochemistry* series will cover virus protein and nucleoprotein complexes and membrane protein complexes.

The range of material included in the 22 chapters of this book indicates the breadth of knowledge and the recent advancement of structural studies on protein complexes. These studies highlight the importance of closely integrating structural biology with biochemistry and cell biology to understand the functional consequences of the structural arrangements adopted by these complexes. To comment here upon all the interesting chapter topics would duplicate the information given in the Contents list, and certainly it is not in order to select only a few. The strength and diversity of the material reviewed in the chapters here speaks for itself in light of the continual stream of recent and upcoming articles appearing in scientific journals devoted to structural molecular biology.

We would like to thank our chapter authors for their manuscripts and enthusiasm. Electronic publishing and the availability of e-books are making an increasing impact, in particular through the availability of individual chapters and rapid publication of new knowledge. The *Subcellular Biochemistry* series, along with other

[1] Method of the Year 2015. *Nat. Methods* **13,** 1.

established book series, provides a continuing platform for the broad dissemination of knowledge and gives authors an opportunity to reflect on their field and provide a deeper perspective on recent advances in fast-moving fields. We hope that the information presented in this book will be of broad interest and value to molecular structural biologists and the wider scientific community.

Mainz, Germany — J. Robin Harris
Newcastle upon Tyne, UK — Jon Marles-Wright
July, 2016

Contents

Chapter 1
Structure and Function of the Stressosome Signalling Hub

Jan Pané-Farré, Maureen B. Quin, Richard J. Lewis, and Jon Marles-Wright

Abstract The stressosome is a multi-protein signal integration and transduction hub found in a wide range of bacterial species. The role that the stressosome plays in regulating the transcription of genes involved in the general stress response has been studied most extensively in the Gram-positive model organism *Bacillus subtilis*. The stressosome receives and relays the signal(s) that initiate a complex phosphorylation-dependent partner switching cascade, resulting in the activation of the alternative sigma factor σ^B. This sigma factor controls transcription of more than 150 genes involved in the general stress response. X-ray crystal structures of individual components of the stressosome and single-particle cryo-EM reconstructions of stressosome complexes, coupled with biochemical and single cell analyses, have permitted a detailed understanding of the dynamic signalling behaviour that arises from this multi-protein complex. Furthermore, bioinformatics analyses indicate that genetic modules encoding key stressosome proteins are found in a wide range of bacterial species, indicating an evolutionary advantage afforded by stressosome complexes. Interestingly, the genetic modules are associated with a variety of signalling modules encoding secondary messenger regulation systems, as well as classical two-component signal transduction systems, suggesting a diversification in function. In this chapter we review the current research into stressosome systems and discuss the functional implications of the unique structure of these signalling complexes.

J. Pané-Farré
Division of Microbial Physiology and Molecular Biology, University of Greifswald, Greifswald 17487, Germany
e-mail: janpf@uni-greifswald.de

M.B. Quin
Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, St. Paul, MN 55108, USA
e-mail: mbquin@umn.edu

R.J. Lewis
Institute for Cell and Molecular Biosciences, Faculty of Medical Sciences, University of Newcastle, Newcastle upon Tyne NE2 4HH, UK

J. Marles-Wright (✉)
School of Biology, Newcastle University, Devonshire Building, Newcastle upon Tyne NE1 7RU, UK
e-mail: jon.marles-wright1@ncl.ac.uk

J.R. Harris, J. Marles-Wright (eds.), *Macromolecular Protein Complexes*, Subcellular Biochemistry 83, DOI 10.1007/978-3-319-46503-6_1

Keywords Stressosome • Bacillus subtilis • Phosphorylation • Kinase • SigmaB • RsbR • RsbS • YtvA

1.1 Introduction

1.1.1 Environmental Sensing and Signalling in Bacteria

Bacteria have colonized almost every possible environmental niche on earth and are subjected to constant fluctuations in their growth conditions (Aertsen and Michiels 2004). Consequently, they have developed sensing and signalling systems that allow adaptive responses through changes in gene expression and cellular behaviour (Hecker and Völker 2001). These signalling systems can alter some cellular behaviours directly, these include motility (Mitchell and Kogure 2006), the modulation of biochemical pathways by allosteric activation/inhibition (Wang et al. 2008); and affect gene expression through the activation of alternative RNA polymerase sigma factors (Paget 2015), transcription factors, and DNA binding two-component signal transduction systems (Stock et al. 2000; Capra and Laub 2012).

Two-component and hybrid two-component signal transduction systems that regulate gene expression comprise a, usually membrane-embedded but occasionally soluble, sensor histidine kinase that senses environmental signals. The kinase is coupled to a cognate response regulator that mediates a cellular response by affecting the transcription of target genes (Gao and Stock 2009). Upon receipt of the appropriate stimulus, the sensor kinase autophosphorylates on an invariant histidine residue. The phosphoryl group is subsequently transferred to a conserved aspartic acid on the cognate response regulator, triggering a conformational change that activates the effector domain of this protein. Hybrid two-component systems combine the function of both the sensory kinase and the response regulator in a single polypeptide, but function in essentially the same way as the classical two-component systems (Capra and Laub 2012). A number of recent reviews of two component signalling are available, but given the dynamic nature of the field, an exhaustive reference list is impossible because of space limitations (Gao and Stock 2009; Krell et al. 2010; Lowe et al. 2012; Bhate et al. 2015).

RNA polymerase sigma factors are essential for the correct initiation of gene transcription through the recognition of promoters (Paget 2015). In addition to the major sigma factors that are required for the transcription of essential housekeeping genes, many bacteria possess alternative sigma factors that control the expression of distinct subsets of genes. The first identification of an alternative sigma factor was in 1979 by Haldenwang and Losick (Haldenwang and Losick 1979), which was at that time erroneously linked to the sporulation process of *Bacillus subtilis*. Subsequently, it became clear that this particular sigma factor was actually induced by general stress (Moran et al. 1981) and the interest in σ^B waned somewhat. Today, σ^B is known as the sigma factor required for the general stress response in *B. subtilis*

and related bacteria, controlling a large regulon of nearly two hundred genes involved in cellular stress (Price et al. 2001; Petersohn et al. 2001). Owing to the potentially large number of genes that alternative sigma factors can regulate, and the resulting heavy metabolic burden on the cell, the activity of sigma factors are regulated by complex signalling cascades, which include phosphorylation, proteolysis, modulation by the alarmone ppGpp, and partner-switching, to ensure their correct and timely function (Österberg et al. 2011).

1.1.2 B. subtilis *σ^B Partner Switching Cascade*

The σ^B-controlled general stress response of *B. subtilis* is regulated by an intricate, phosphorylation-dependent partner switching cascade comprising over a dozen currently-known components (Fig. 1.1) (Hecker et al. 2007). In unstressed cells, σ^B is prevented from interacting productively with RNA polymerase because σ^B is sequestered in a transcriptionally-inactive complex with its cognate anti-sigma factor, RsbW (where Rsb stands for Regulator of sigma B) (Benson and Haldenwang 1993). Upon the imposition of stress, RsbW releases σ^B, allowing it to bind to RNA polymerase to initiate transcription of the general stress regulon (Alper et al. 1996; Kim et al. 2004a). This effect is mediated by the anti-anti sigma factor RsbV. In its dephosphorylated state, RsbV has a greater affinity for RsbW than RsbW does for σ^B. The switch in binding partner of RsbW from σ^B to RsbV causes the release of σ^B, allowing σ^B to interact with RNA polymerase (Yang et al. 1996). Under non-stressful conditions, RsbV is maintained in a phosphorylated state by the kinase function of RsbW, and when phosphorylated, RsbV has a much lower affinity for RsbW than RsbW has for σ^B and consequently σ^B is preferentially bound by RsbW (Delumeau et al. 2002).

RsbV sits at a branch-point in the signalling cascade, where both energy and environmental stresses are integrated (Voelker et al. 1996). The activator/phosphatase pair, RsbQ/P, responds to the energy status of the cell and controls the energy stress branch of the σ^B pathway. The enzymatic activity of RsbQ is required for the correct response to energy stress (Brody et al. 2001), presumably because the reaction product of the RsbQ α/β hydrolase is needed to stimulate the phosphatase activity of RsbP, towards phosphorylated RsbV (RsbV-P), by binding to its Per-Arndt-Sim (PAS) regulatory domain. When the ATP:ADP ratio falls, the RsbQ/P couple dephosphorylates RsbV, inducing the partner-switching of RsbW away from σ^B and towards RsbV to liberate σ^B (Vijay et al. 2000; Kaneko et al. 2005; Nadezhdin et al. 2011).

In the pathway that responds to environmental stresses, such as high salt and ethanol, the RsbV phosphorylation status is controlled by the phosphatase RsbU, which also dephosphorylates RsbV-P (Yang et al. 1996). In this pathway RsbU is activated by the RsbT kinase. When the cell is unstressed, RsbT is bound by structural proteins, including RsbR and RsbS, which self-assemble into a ~1.8 MDa macromolecular complex known as the stressosome. Upon receipt of an environmental stress

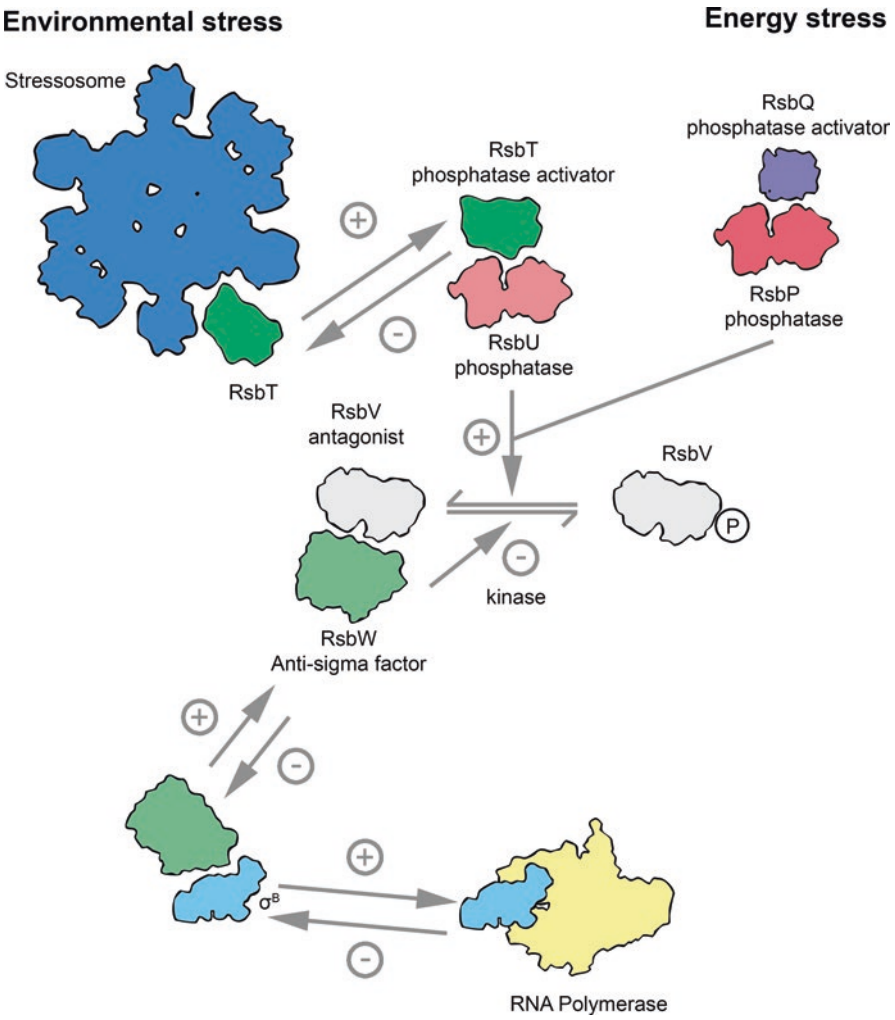

Fig. 1.1 σ^B cascade. The *Bacillus subtilis* σ^B cascade is illustrated to show how partner-switching induces the activity of RNA polymerase and acts to regulate gene expression. Pre-stress, the anti-sigma factor RsbW sequesters σ^B and prevents it from directing RNA polymerase to σ^B-controlled promoters. In this state, RsbV is phosphorylated (RsbV-P) by the kinase activity of RsbW and hence RsbV is inactivated. Under stress conditions, RsbV-P becomes dephosphorylated by one of two phosphatases and attacks the RsbW:σ^B complex and liberates σ^B to direct transcription of its regulon to provide the cell with stress-resistance. RsbV is also the point at which the environmental and energy stress responses converge. Under energetic stress, the phosphatase RsbP is activated by RsbQ and dephosphorylates RsbV-P to allow it to form complexes with RsbW. Environmental stresses are integrated by the stressosome, which sequester the RsbU phosphatase-activator, RsbT, in the absence of stress. Under environmentally stressful conditions, RsbT phosphorylates the STAS domains of the stressosome proteins and disassociates, because of a presumed reduced affinity for the phosphorylated proteins, and RsbT switches its binding partner to the phosphatase RsbU. The RsbT:RsbU complex activates RsbV by its dephosphorylation. The phosphatase RsbX acts to remove phosphoryl groups from the stressosome and to mediate the duration of the stress response by 'resetting' the system. *Ringed plus signs* indicate positive regulatory events affecting σ^B activity, while *ringed minus signs* indicate those that are negative regulatory events

signal, RsbT phosphorylates the structural proteins, initiating its release from the stressosome, and leading to the subsequent activation of RsbU (Kang et al. 1998; Chen et al. 2003; Kim et al. 2004a). The structural proteins that form the stressosome include RsbS; RsbR and the paralogues of RsbR, YkoB, YojH and YqaH (Delumeau et al. 2006), which have been collectively re-named RsbRA, RsbRB, RsbRC and RsbRD, respectively (Kim et al. 2004b); and the blue-light sensor, YtvA (Akbar et al. 2001; Losi et al. 2002; Jurk et al. 2013). The relative stoichiometries of the RsbR paralogues in the stressosome, assuming a single entity is formed in the cell, and the exact mechanism by which environmental stress signals are perceived and transduced by the stressosome are not known at this time. For simplicity, most structural studies have concentrated on the RsbR/RsbS/RsbT triumvirate as a surrogate of the likely more complex stressosome assemblies found in the cell.

By use of the σ^B-dependent promoter of the *ctc* gene (which encodes a component of the large subunit of the ribosome (Truitt et al. 1988; Schmalisch et al. 2002)) fused to a *lacZ* reporter, σ^B activity has been monitored in real time by several groups. From a combination of studies, it is clear that the σ^B signalling cascade is temporally limited. Upon the imposition of environmental stress, σ^B activity rises to a maximum after 20–30 min, after which time σ^B activity begins to decline back to base-line levels. The resetting of the system takes place in two independent steps:

1. RsbW phosphorylates RsbV, resulting in the partner switching of RsbW to resequester and inactivate σ^B, switching off gene transcription
2. The phosphatase RsbX dephosphorylates RsbR-P and RsbS-P (Chen et al. 2004), which resets the system through the re-sequestration of RsbT by the stressosome (Eymann et al. 2011)

The system is also regulated transcriptionally, due to the presence of a σ^B promoter upstream of RsbV, which leads to increased amounts of the RsbW anti-sigma factor and RsbX phosphatase, which are located downstream of RsbV in the *B. subtilis* genome (See 1.7.1) (Dufour et al. 1996).

1.1.3 RsbRST Module Distribution

RsbR, RsbS and RsbT form the core components of the stressosome and the genes encoding these proteins are co-located in the genome at the start of the *rsb* operon. These three proteins form the RsbRST module, which is distributed widely across bacteria and is found in representatives of the Firmicutes, Actinobacteria, proteobacteria, Bacteroides, cyanobacteria and Deinococcus groups. Whilst the core of the module is well conserved, the N-terminal domain of RsbR is less well maintained between species. Furthermore, the downstream components in the *rsb* operon vary considerably, including secondary messenger signalling and two-component regulators, as well as alternative sigma factors, indicating that the stressosome complex may have evolved as a signalling hub controlling a diversity of cellular functions (Pané-Farré et al. 2005).

1.1.4 Chapter Outline

In this chapter we will discuss how the structure of the stressosome relates to its function as a signal integration and transduction hub, and the complex signalling behaviours seen in *B. subtilis*. We will assess the current structural knowledge of the individual components of the stressosome as determined by X-ray crystallography and the structure of the stressosome determined by single particle cryo-EM. We will highlight the modularity of this complex and show that it has been widely adopted among bacterial species as a signalling hub with distinct inputs and outputs. Furthermore, we will discuss the structural basis for its signalling mechanism and the role of different inputs in modulating the output level of the complex. The challenges of structural studies on flexible and heterogeneous complexes will be discussed with reference to the cryo-EM structure. The chapter will close with a discussion of unanswered questions on stressosome signalling.

1.2 Stressosome Components

1.2.1 Stressosome Composition

The simplest stressosome complex comprises the RsbS scaffold protein, which consists of a single STAS (Sulphate Transporter and Anti-Sigma factor antagonist (Sharma et al. 2011)) domain and RsbR, which has a variable N-terminal domain and a conserved C-terminal STAS domain (Chen et al. 2003). The STAS domain scaffold is well-conserved between RsbS and the RsbR paralogues (Fig. 1.2) and the RbsR paralogues can form heterogeneous stressosome complexes both *in vivo* and *in vitro* (Delumeau et al. 2006; Jurk et al. 2013). On the imposition of stress, RsbT phosphorylates conserved serine and threonine residues on the STAS domains of RsbS and the RsbR paralogues (Kim et al. 2004a). In the RsbR/RsbS stressosome surrogate, the rate of phosphorylation of RsbS on Ser59 is quicker *in vitro* than on either RsbR phosphorylation site (Thr171, Thr205) (Chen et al. 2003), strongly suggesting that RsbT interacts predominantly with RsbS in the stressosome. How the kinase is held in an inactive state in the absence of stress signals is unknown. Moreover, the effect and purpose of phosphorylation in the stressosome is also unclear (Kim et al. 2004a; Chen et al. 2004; Liebal et al. 2013; Gaidenko and Price 2014). It is generally agreed, however, that RsbT must dissociate from the stressosome to activate RsbU and the downstream components of the σ^B cascade in order to respond to the imposition of stress, (Kang et al. 1998). In the absence of RsbT, the RsbX phosphatase is able to dephosphorylate RsbS-P and RsbR-P (Chen et al. 2004). RsbX has a differential activity against the different phosphorylation sites on these proteins, which has been suggested to relate to an adaptive response of the stressosome to sustained, or repeated, stresses (Eymann et al. 2011).

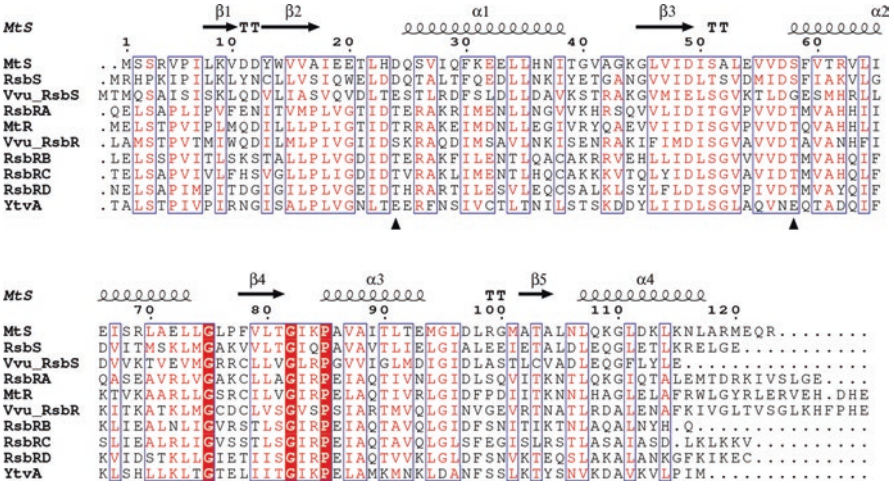

Fig. 1.2 Alignment of STAS domain sequences. Multiple sequence alignment generated for the STAS domains of stressosome proteins from *Bacillus subtilis* (RsbS, RsbRA-D, YtvA), *Moorella thermoacetica* (MtS, MtR), and *Vibrio vulnificus* (Vvu_RsbS, Vvu_RsbR) to highlight sequence conservation in these proteins. Regions outlined in *blue* show areas of sequence conservation, residues *coloured red* indicate partial conservation and those outlined in *red* are completely conserved across the aligned sequences. The secondary structure elements from the X-ray crystal structure of *Mt*S (PDBID: 2VY9) are shown above the alignment with residues numbered for the *Mt*S sequence. The positions of conserved serine and threonine residues phosphorylated by RsbT kinase are shown with *black triangles* below the alignments. Note that RsbS and its homologues are only phosphorylated on the conserved serine at position 58 in *Mt*S and 59 in RsbS. VVu_RsbS has a glycine at this position, although there is a serine at position 62 in this protein; there is currently no published experimental evidence that the Vvu_RsbS protein is phosphorylated. Figure 1.2 was prepared using EsPript (Gouet et al. 2003)

1.2.2 Structure of RsbS

RsbS is a single STAS domain protein, and STAS domains are also found in anion transporters, the SpoIIAA anti-anti-sigma factor antagonist in *B. subtilis* (Kovacs et al. 1998), and as essential components of the stressosome. STAS domains appear to function primarily as a scaffold for the recruitment of other proteins, particularly in partner-switching networks (Aravind and Koonin 2000). The *B. subtilis* RsbS protein forms stable stressosome complexes with the RsbR paralogues both *in vivo* and *in vitro* (Chen et al. 2003; Kim et al. 2004b; Delumeau et al. 2006; Reeves et al. 2010). The RsbS STAS domain has a five-stranded beta sheet core with three alpha helices on one face and a single C-terminal helix on the other face (Fig. 1.3). Phosphorylation takes place on a conserved serine/threonine at the N-terminus of helix 2, the central of the three helices of the STAS domain. Structures of the STAS domains of SpoIIAA (Seavers et al. 2001) and RsbS (Marles-Wright et al. 2008; Quin et al. 2012) in both phosphorylated and non-phosphorylated forms do not reveal any significant conformational changes in protein structure, suggesting that

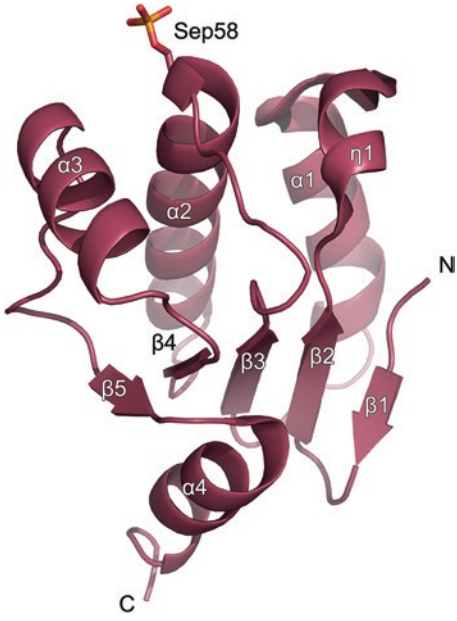

Fig. 1.3 X-ray crystal structure of MtS. The X-ray crystal structure of the phosphorylated form of MtS (PDBID: 3TZB) is shown as a cartoon with secondary structure elements labelled from the N- to C-terminus. Serine 58 is phosphorylated in this structure and the phosphoryl group is depicted as *orange and red sticks*

release of RsbT from the stressosome as a function of phosphorylation on RsbS-Ser59, and/or its equivalent residue in RsbR, Thr205 (and/or the non-equivalent Thr171), is likely to occur solely by electrostatic and physical repulsion. Similarly, the anti-anti-sigma factor SpoIIAA is released from its complex with cognate dual-function kinase and anti-sigma factor SpoIIAB by the same mechanism (Seavers et al. 2001; Masuda et al. 2004). However, in the absence of a high resolution structure of a phosphorylated stressosome complex, rather than isolated RsbS and RsbS-P structures, significant conformational changes in STAS domain architecture during stressosome activation cannot be excluded (Kumar et al. 2010).

1.2.3 Structure of RsbR

The *B. subtilis* RsbR protein has a C-terminal STAS domain and an N-terminal non-heme globin domain (Murray et al. 2005). While there is no crystal structure of the RsbR STAS domain, it has 30% sequence identity to that of RsbS and will presumably adopt the same fold. Unlike RsbS and other characterised single domain STAS proteins, RsbR is phosphorylated on two conserved threonine residues, Thr171 and Thr205 (Gaidenko et al. 1999; Eymann et al. 2011). Thr205 is equivalent to Ser59 of RsbS and is likely to be located at the N-terminus of helix 2, whereas the unique phosphorylation site of RsbR, Thr171, is positioned at the N-terminus of helix 1. Thr171 tends to be occupied by serine or threonine in the two-domain RsbR paralogues, but glutamate or aspartate in the single domain RsbS paralogues and YtvA,

suggesting that the Asp/Glu residues mimic the effect of phosphorylation on RsbR in the latter group of proteins, which apparently increases the activity of RsbT towards RsbS (Chen et al. 2004). Whilst phospho-ablative mutations have been made in RsbR and the effects on phosphorylation patterns, rates and signalling have been measured (Chen et al. 2004; Gaidenko and Price 2014), similar experiments have not been performed to the best of our knowledge on Asp25Thr (or alanine) mutations in RsbS in order to decipher the role of phosphorylation on Thr171 in RsbR.

The structure of the N-terminal domain of RsbR (N-RsbR) displays an all alpha-helical non-heme globin fold (Murray et al. 2005) (Fig. 1.4a, b). A comparison of N-RsbR to other globin-like proteins, such as the HemAT aerotaxis sensor (Hou et al. 2000; Zhang and Phillips 2003) shows that RsbR does not possess a heme binding pocket and also lacks the heme coordinating residues found in the hemoglobins (Murray et al. 2005). In the absence of a heme binding pocket, it is thought that RsbR may interact with an as yet unidentified small molecule ligand in an analogous way to the amino acid-binding globin sensor proteins (Kitanishi et al. 2011), or to the ligand-binding sites in the non-heme globin sensors that regulate the entry of *Bacillus anthracis* into sporulation (Stranzl et al. 2011).

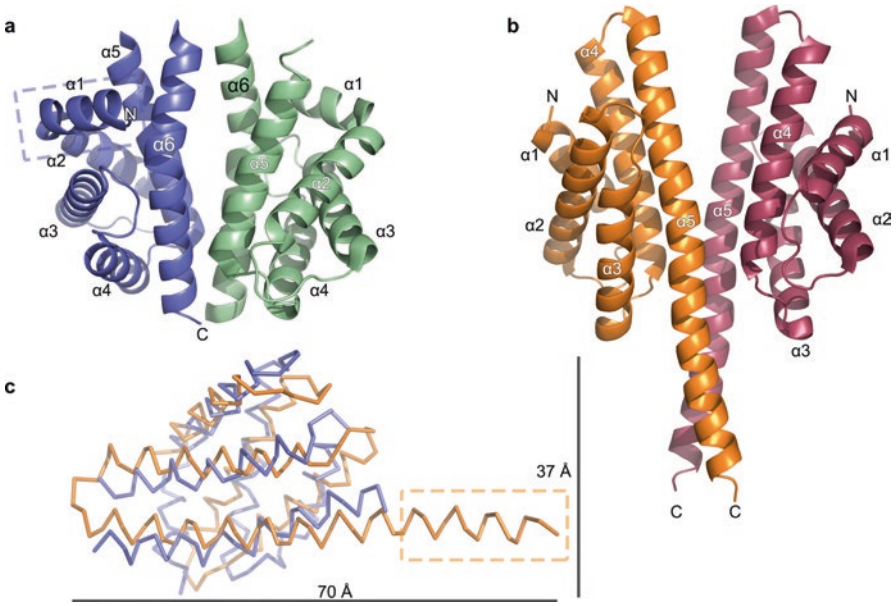

Fig. 1.4 X-ray crystal structures of RsbR and MtR. The X-ray crystal structures of the *B. subtilis* RsbR (**a**) (PDBID: 2BNL) and *Moorella thermoacetica Mt*R (**b**) (PDBID: 3TZA) are shown as cartoons with secondary structure elements labelled. Both proteins are depicted as the physiologically relevant dimer forms. (**c**) Secondary structure comparison of RsbR and *Mt*R monomers; RsbR is shown in *blue*, *Mt*R in *orange*. While the two structures have the same non-heme globin fold, RsbR has an additional N-terminal α-helix (*blue-dashed box*) and in the X-ray crystal structure lacks the C-terminal portion of the α-helix that links to the STAS domain (*orange-dashed box*)

The N-terminal domain of the *Moorella thermoacetica* RsbR orthologue, *Mt*R (N-*Mt*R), has only 12% sequence identity to that of *B. subtilis* RsbR, yet an almost identical fold (Fig. 1.4c). The N-*Mt*R structure has an extended alpha helix at the C-terminus of the globin domain that links this with the C-terminal STAS domain (Quin et al. 2012). This extended, terminal helix is equivalent to the 'J' helix of the LOV (light-oxygen-voltage) domain of the RsbR paralogue YtvA, which has been suggested to convey blue light-dependent signalling (Jurk et al. 2010).

The N-terminal domains of the RsbR paralogues encoded in the *B. subtilis* genome share limited sequence identity, although they are each found in stressosomes *in vivo* and can associate with RsbS to form stressosome-like structures using purified recombinant proteins *in vitro* (Delumeau et al. 2006). While the structures of their N-terminal domains are yet to be determined, they are likely to form sensory domains given the fact that a number of them belong to the PAS (Per-Arnt-Sim) domain superfamily, which are usually found in signalling systems as sensors (Ponting and Aravind 1997).

1.2.4 Structure of YtvA

While the exact role of the RsbR paralogues is not known, it appears that the detection of blue light by YtvA is modulated by the other RsbR paralogues (van der Steen et al. 2012). Indeed, YtvA is the only RsbR related stressosome component in *B. subtilis*, and potentially *Listeria monocytogenes*, for which the activating signal has been confirmed (Möglich and Moffat 2007; Ondrusch and Kreft 2011). This protein has a C-terminal STAS domain and an N-terminal LOV domain that binds a flavin mononucleotide cofactor and undergoes a light-dependent conformational change in the nucleotide binding site (Losi et al. 2002; Möglich and Moffat 2007; Herrou and Crosson 2011) (Fig. 1.5a). This structural rearrangement is transmitted to the 'J' helix, which links the sensory domain to the STAS domain, resulting in the movement of this helix away from the axis of the dimer interface (Möglich and Moffat 2007). These structural changes hint at the signal transduction mechanism from the sensory domains of RsbR proteins to the STAS domain, although solution small-angle X-ray scattering experiments on the full length YtvA failed to show any significant structural rearrangements upon illumination with blue light (Jurk et al. 2010).

The solution structure of the full length *B. subtilis* YtvA protein was published recently in the PDB (PDBID: 5MWG) (Fig. 1.5b). While the dimer interface in the crystal structure of the YtvA LOV domain is formed primarily by interactions between the beta-sheets of the two monomers, the solution structure has a distinct dimer arrangement with an additional N-terminal α-helix modelled to form the primary point of contact between the monomers. This particular arrangement is consistent with the dimer arrangement seen in the crystal structures of both the *B. subtilis* and *M. thermoacetica* RsbR proteins (Fig. 1.4). These intriguing structural differences may be biologically significant, or represent artefacts of the different

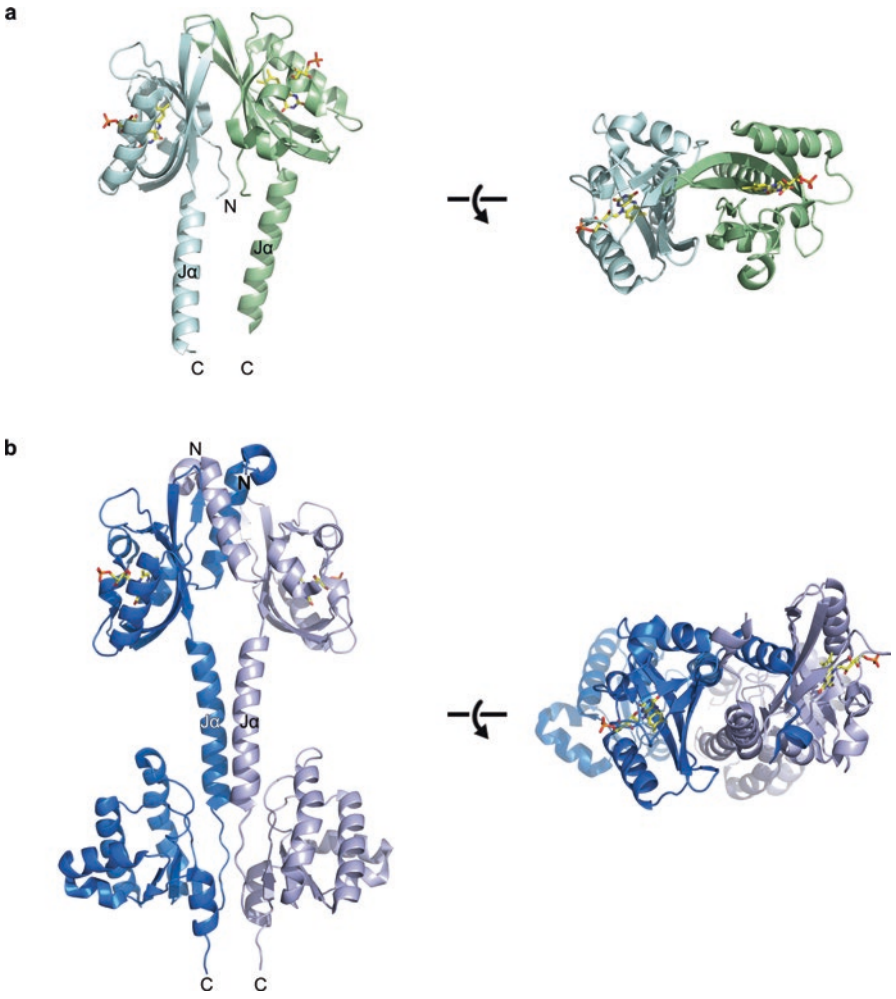

Fig. 1.5 Structures of YtvA. (**a**) Crystallographic model of the LOV domain of *B. subtilis* YtvA in the dark state (PDBID: 2PR5). The model comprises a dimer in the asymmetric unit, which are shown as cartoons with the bound FMN cofactor as sticks coloured by atom. In the light state structure (PDBID: 2PR6) the Jα helices are displaced by up to 2 Å away from the dimer axis. (**b**) Solution structure of the full length *B. subtilis* YtbA protein (PDBID: 2MWG). The LOV domain is shown at the *top* and STAS domain to the *bottom* in the *left* panel. The full length structure displays a distinct dimer organisation to the isolated LOV domain with an N-terminal helix, which is not modelled in the crystal structure, forming substantial contacts in the dimer interface

constructs and methods used to produce the two structures. The orientation of the LOV domain in relation to the STAS domains in the full length dimer structure hints at a mechanism of signal transduction through the Jα helix to alter the orientation of the LOV domain relative to the STAS domain, either through a movement of the LOV domain itself, or of the STAS domains in the stressosome core.

YtvA is the only RsbR paralogue that is not phosphorylatable, encoding glutamate at both positions equivalent to Thr171 and Thr205 in RsbR. The ready light-to-dark reversibility of YtvA is indicated by a half-life of the photo-excited state of some 40 min (Losi et al. 2003). The effect of mutation of Glu142 and Glu202 in YtvA to serine/threonine, resulting in the potential for phosphorylation of these residues by RsbT, and the potential for phosphoregulation of YtvA and its impact on σ^B activity, are as-of-yet unconducted experiments that might explain the regulatory role of YtvA in stressosome signalling.

1.2.5 Structure of RsbT

To date there is no X-ray crystal structure of the RsbT protein kinase. However, there is a structure of a *Bacillus stearothermophilus* homologue, the SpoIIAB anti-sigma factor, with which it shares 28% sequence identity (Campbell et al. 2002). This protein binds to the sporulation sigma factor, σ^F, in an ATP dependent manner to suppress the activity of the sigma factor. SpoIIAB also binds in an ADP-dependent fashion to the anti-anti-sigma factor protein SpoIIAA, which is a STAS domain protein that disrupts the binding between SpoIIAB and σ^F to release the sigma factor (Alper et al. 1994). The nucleotide-dependencies of these interactions are probably a reflection of the rather poor stability of recombinant SpoIIAB in the absence of ADP or ATP (Lord et al. 1996). The homologous RsbW can bind to both its cognate sigma factor (σ^B) and anti-anti-sigma factor *in vitro* in the absence of additional nucleotide (Delumeau et al. 2002). Both SpoIIAB and RsbT are members of the GHKL superfamily of protein kinases and the structure of SpoIIAB displays an α/β sandwich with an antiparallel β sheet flanked by alpha helices (Dutta and Inouye 2000). The ATP binding pocket is found in a deep crevice between the helices and the β sheet. The pocket has a highly flexible lid that is thought to communicate the nucleotide binding status of the protein through changes in its structure. Unlike SpoIIAB, which acts as a dimer, RsbT appears to act as a monomer in solution, because it lacks a C-terminal helix found in SpoIIAB that mediates dimerisation (Campbell et al. 2002; Masuda et al. 2004; Delumeau et al. 2006).

1.2.6 Structure of RsbX

RsbX is the protein phosphatase responsible for resetting the stressosome to the resting state after the imposition of stress and its phosphorylation by RsbT. RsbX acts on both RsbS-P and RsbR-P, although it displays differential activity against the different proteins *in vitro* and *in vivo* (Chen et al. 2003; Eymann et al. 2011). RsbX is a member of the PP2C family of protein phosphatases (Das et al. 1996) and the structures of both the *B. subtilis* and *M. thermoacetica* homologues (*Mt*X) have been determined by X-ray crystallography (Quin et al. 2012; Teh et al. 2015). These

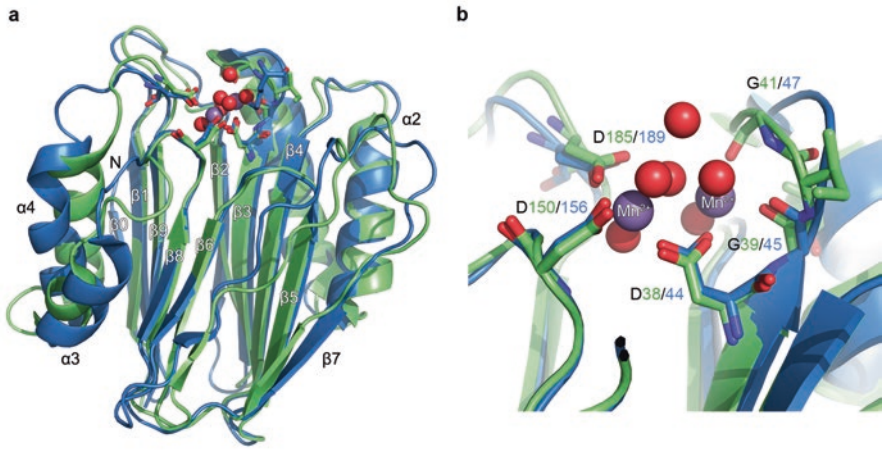

Fig. 1.6 X-ray crystal structure of RsbX and *Mt* X. (**a**) Crystallographic models of both the *B. subtilis* RsbX (PDBID: 3ZT9) (*green*) and *Moorella thermoacetica Mt*X (PDBID: 3W40) (*blue*) protein phosphatases have been published. The models are depicted as cartoons with secondary structure features labelled from the N- to the C-terminus. Manganese ions required for the catalytic activity of the protein bound in the active site of the protein are depicted as *purple spheres*, with water molecules shown as *red spheres*. (**b**) Metal binding site of RbsX and *Mt*X showing residues coordinating the bound manganese ions. The two manganese ions (*purple spheres*) are coordinated by the carboxylic acid groups of aspartic acid residues and the backbone carbonyl groups of two glycine residues (shown as stick representations). A number of ordered solvent residues are also present in the coordination shell of the manganese ions (*red spheres*)

proteins share only 25% sequence identity yet superimpose with an RMSD Cα of 1.9 Å over 182 Cα. The two proteins display the same phosphatase fold with an αββα architecture, with the central beta-sandwich flanked by two pairs of alpha helices on the outer faces of the protein (Fig. 1.6). The catalytic centre of the protein is located at one edge of the beta-sandwich and centred around a cluster of acidic residues, which are highly conserved in the PPM phosphatase family (Quin et al. 2012). These residues coordinate two divalent metal ions that are required for the function of the protein. RsbX and *Mt*X show a strong preference *in vitro* for the presence of Mn^{2+} in this metal binding site for their activity against both their native and synthetic phosphatase substrates (Quin et al. 2012; Teh et al. 2015). A depression on the surface of these proteins above the active site provides a potentially ideal site for interaction with their targets (Quin et al. 2012; Teh et al. 2015). The dephosphorylation reaction is likely to proceed in a manner analogous to other members of the PP2C family, where a metal-ion bridged water molecule acts as a nucleophile against the phosphate group on the target protein and a second water molecule protonates the dephosphorylated serine/threonine residue (Das et al. 1996; Barford et al. 1998).

1.3 Stressosome Complexes

The RsbR paralogues and RsbS form a stable 1.5 MDa stressosome complex in *B. subtilis* that sequesters RsbT in the absence of stress. These complexes can be isolated directly from *B. subtilis* cells and are present throughout the life of individual bacterial cells (Kim et al. 2004b; Gaidenko and Price 2014). Minimal stressosome complexes can be reconstituted *in vitro* by mixing RsbS with the RsbR paralogues to simplify the structural biology, although RsbR has a tendency to self-associate into stressosome-like particles in the absence of RsbS. In the absence of RsbS, stressosome complexes are unable to sequester RsbT, and deletion of the *rsbS* gene in *B. subtilis* leads to constitutive activation of RsbT and a resulting small-colony phenotype. This is presumably caused by deleterious effects of the products of the σ^B regulon, or a negative effect of competition for cellular resources with σ^A regulated housekeeping genes (Kang et al. 1996).

The minimal stressosome complexes of individual RsbR paralogues and RsbS are competent to bind RsbT (Delumeau et al. 2006) and can be phosphorylated by the kinase on conserved residues in their C-terminal STAS domains (Reeves et al. 2010). RsbR paralogues in stressosomes can also be exchanged *in vitro*, implying that these complexes have some dynamic properties in solution (Delumeau et al. 2006). However, immunofluorescence experiments using antibodies specific for the N-terminal domain of RsbR show that RsbR – and hence the stressosome - forms punctate foci that persist in cells throughout the stress response and its recovery (Marles-Wright et al. 2008). These results suggest that in wildtype *B. subtilis* the stressosome always contains RsbR and that the RsbR paralogues may exchange into stressosomes through the life of the cell. It is not known if distinct populations of stressosomes with different RsbR paralogues exist at the same time within the cell. *B. subtilis* stressosomes purified by anti-N-RsbR affinity purification contain at least three paralogues of RsbR (Delumeau et al. 2006); thus, single, double, and triple knockouts are effectively complemented by the remaining paralogues in normally growing *B. subtilis* (Kim et al. 2004b). A quadruple knockout of the RsbR paralogues leads to constitutive σ^B activation in much the same manner as the RsbS knockout, as no competent stressosome complexes are able to form in the absence of the RsbR paralogues, leading to the presence of free RsbT in the cell, up-regulated RsbU activity and σ^B liberated to interact with RNA polymerase.

1.4 Production of Recombinant *B. subtilis* Stressosomes for Structural Analysis

1.4.1 RsbR/RsbS Binary Complex

Recombinant *B. subtilis* stressosome complexes for structural analysis were produced by plasmid-based co-expression of RsbR and RsbS from a bi-cistronic operon in *Escherichia coli* and purified to homogeneity in a multi-step protocol

(Marles-Wright et al. 2008). As with any high molecular weight protein complex destined for structural analysis, the purification protocol was extensively refined to ensure the final sample was as homogeneous as possible. The optimised protocol included anion-exchange, size-exclusion gel-filtration, hydrophobic interaction chromatography, and a second gel-filtration step. Using this protocol, stressosome complexes were separated from ribosomes, and other high-molecular weight proteins and macromolecular complexes from the *E. coli* host strain. Despite the relatively high levels of protein expression and the multi-step purification protocol, it was still possible to identify contamination by the cubic core of the pyruvate dehydrogenase complex in some micrographs (Marles-Wright and Lewis, unpublished observations).

1.4.2 RsbR/S/T Ternary Complex

Production of the RsbR/S/T ternary complex was achieved by expression of RsbT as an N-terminal GST-fusion to enhance protein production levels and solubility. The RsbT protein was purified by glutathione affinity chromatography in the presence of 1 mM ADP to ensure the nucleotide-binding site of the kinase was occupied, and subsequent on-column cleavage of the GST-tag was achieved using the HRV-3C protease. Purified ADP-loaded RsbT was mixed in excess with purified RsbR/S minimal stressosomes and subjected to an additional round of size-exclusion gel-filtration chromatography to remove unbound RsbT (Marles-Wright et al. 2008).

1.5 Cryo-EM Structure of the *Bacillus subtilis* Stressosome

Initial negative stain electron microscopy studies of recombinant *B. subtilis* stressosomes identified a 20 nm ring-shaped structure (Fig. 1.7) (Chen et al. 2003). Due to the absence of distinct views other than the characteristic ring, the complex was initially thought to form a doughnut-like oligomer that exhibited a preferential orientation on the carbon-film of the EM grid (Delumeau et al. 2006). Collection of data on unstained samples by cryo-EM and single-particle analysis allowed the calculation of 3D–reconstructions for a stressosome core structure comprising the full length RsbS protein and an N-terminally truncated RsbR; the full length RsbR/RsbS stressome; and a ternary complex between RsbR/RsbS and the kinase RsbT. We will discuss each of these in turn below.

1.5.1 RsbR/RsbS Core Structure

The single-particle cryo-EM reconstruction of the $RsbR_{(146–274)}$:RsbS core STAS-domain complex was determined to a resolution of 6.5 Å by imposing icosahedral symmetry restraints on the calculation of the molecular envelope as indicated by

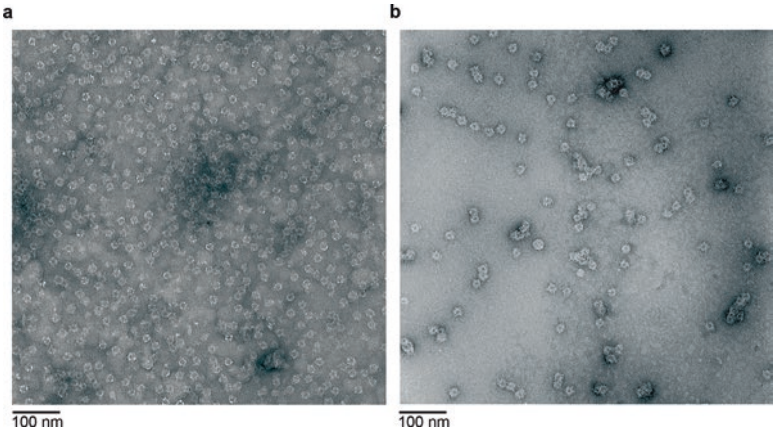

Fig. 1.7 Negatively-stained transmission electron micrograph of the *B. subtilis* stressosome. Purified recombinant *B. subtilis* stressosome complexes were stained with 1% uranyl acetate and imaged by transmission electron microscopy. (**a**) RsbR/RsbS stressosome complexes have a ring-shaped appearance in projection, with rough edges. They are approximately 25–30 nm in size when imaged by TEM. (**b**) RsbR/RsbS/RsbT ternary complexes have a similar appearance to the RsbR/RsbS complexes with more pronounced features at their edges. The ternary complex has a similar diameter to the binary complex. These images were taken with the assistance of Professor J. Robin Harris

analysis of the initial Eigenimages from the first round of reference-free class-averaging. The reconstructed density was a hollow shell with an outer radius of 90 Å and an inner radius of 45 Å. The final reconstruction displayed clear tubes and flat sheets of electron density consistent with the alpha helices and beta sheet seen in the crystal structure of *Mt*S (Fig. 1.8). The position of the STAS domains in the EM-derived molecular envelope places the N-terminus of the STAS domain at the external surface. This distribution is consistent with their covalent attachment to the C-terminus of the N-terminal domains of RsbR, to place these domains on the outside of the stressosome complex so as to most easily interact with other proteins, ligands and stimuli. Indeed, it has already been suggested that the groove in the dimer interface of N-RsbR could be utilised for binding upstream signalling partners, based on its structural similarity to HemAT and the location of the unique 'Z' helix of HemAT in the dimer interface groove; the binding of KaiC peptides in the circadian clock complex KaiA/KaiC; and the interaction of *Vitreoscilla* haemoglobin with a partner dioxygenase (Murray et al. 2005).

The arrangement of the STAS domains across the icosahedral dimer interface of the core stressosome reconstruction was distinct to the dimeric arrangement seen within the *Mt*S crystal structure; the primary interactions in the stressosome core are between the first β-strand and the C-terminal helix of the STAS domain, whereas the crystallographic dimer of *Mt*S is formed between α3 and β5 across the two chains in the dimer, and probably does not represent a stable assembly as determined by the PISA server (Krissinel and Henrick 2007) .

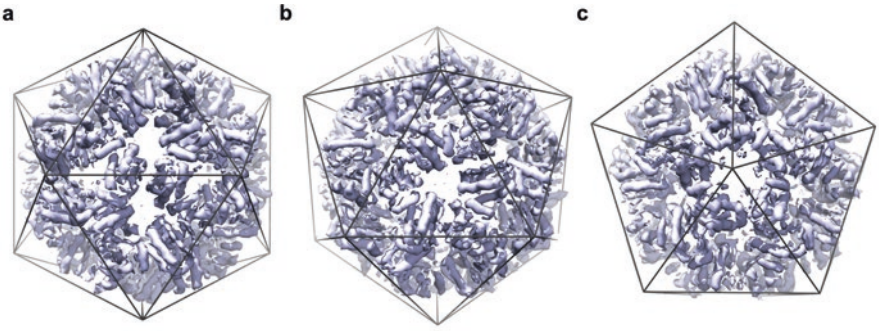

Fig. 1.8 Single particle reconstruction of the RsbR$_{(146–274)}$/RsbS stressosome core complex. The final experimental EM-derived icosahedral reconstruction of the RsbR$_{146–274}$/RsbS stressosome core complex is shown as a *blue* surface contoured at 3σ. Views down the icosahedral two (**a**), three (**b**) and fivefold (**c**) axes are shown with an icosahedral net for reference. The stressosome core has radius of 90 Å

1.5.2 RsbR/RsbS Structure

The cryo-EM reconstruction of the full-length RsbR/RsbS minimal stressosome was determined to a nominal resolution of 8.0 Å. Inspection of experimental Eigenimages generated from the initial particle set showed that the complex possessed a mixed-symmetry, with an icosahedral core and projections from this core that obeyed a lower, D_2 symmetry (Fig. 1.9). While unusual, this type of symmetry mismatch has been found in in the capsids of bacteriophage, where nucleic-acid portal translocases inhabit unique vertices (Dube et al. 1993), and also in multicomponent proteasome complexes (Beuron et al. 1998). Analysis of this symmetry mismatch and manual model building with a child's magnetic toy led to a single plausible model for the RsbR/RsbS stressosome containing 40 copies of RsbR and 20 copies of RsbS (Fig. 1.10). In this model, the core of the stressosome is comprised solely of 60 STAS domains, but with an additional 20 projections that obey D_2 symmetry to yield 20 peripheral 'turrets' made up of dimers of N-RsbR (Fig. 1.11). This model was validated experimentally by the production of Eigenimages from back-projections of the final stressosome model with D_2 symmetry imposed (Fig. 1.9c). The final reconstruction had a core with a radius of 90 Å and total radius with the projections of 150 Å. The density for the peripheral projections was less well defined than the core density, implying a level of heterogeneity in the position of the N-termini of RsbR in the complex relative to the core.

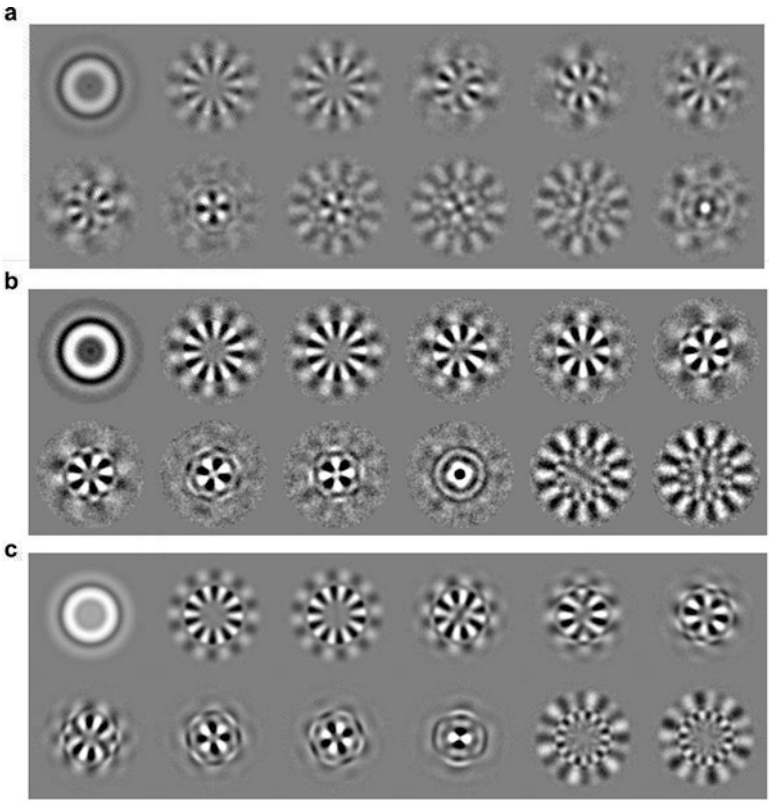

Fig. 1.9 Experimental Eigenimages of the RsbR/RsbS stressosome complex. (**a**) Eigenimages showing symmetry elements associated RsbR/RsbS stressosome showing the mixed symmetry that appears due to the N-terminal projections of RsbR protein. (**b**) Eigenimages from the full length RsbR/RsbS reconstruction, which appear to show a clear tenfold symmetry (**c**) Eigenimages from re-projections of a model RsbR/RsbS with imposed D_2 symmetry, which also exhibit a clear tenfold symmetry, suggesting the tenfold feature is a consequence of the centring and not a structural feature

1.5.3 RsbR/RsbS/RsbT Ternary Complex

The ternary complex of RsbR/RsbS and RsbT was determined to 8.3 Å resolution and displayed the same symmetry mismatch as the binary RsbR/RsbS complex. This reconstruction had identical core and total dimensions to the RsbR/RsbS complex, while the presence of additional density features was apparent above the regions of the core not occupied by the N-RsbR turrets (Fig. 1.12). These additional density features were attributed to the presence of the RsbT protein. In this reconstruction the turrets appeared less well distinguished than in the RsbR/RsbS complex implying a greater degree of positional heterogeneity than in the binary complex.

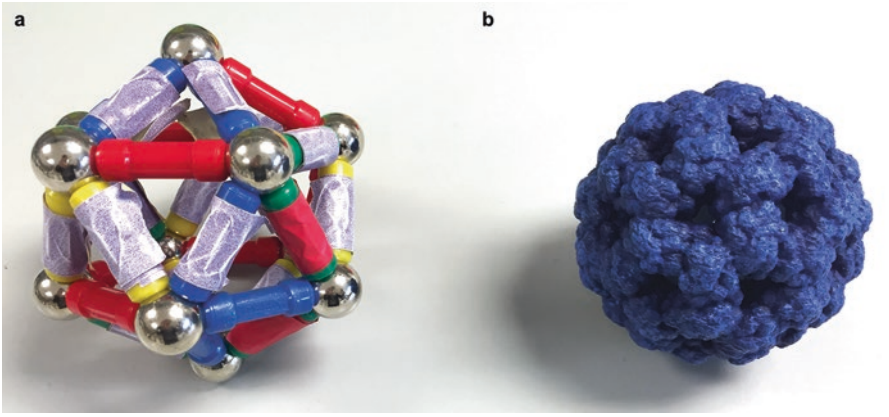

Fig. 1.10 Stressosome model building. To aid with the determination of the subunit arrangement within the stressosome reconstructions a model of the stressosome core was built using a magnetic modelling set (**a**). In this model the magnetic bars represent dimers of RsbR (*blue*) and RsbS (*red*), and the steel balls serve as anchor points between the bars, but represent holes at the fivefold axes of the stressosome. The model was built using a set of rules for protein:protein interactions determined from solution studies of the individual components, namely: (1) RsbS is a dimer in solution, but doesn't interact to form higher-order structures. (2) RsbR is a dimer in solution and can interact with itself to form higher-order structures; however, higher-order interactions with RsbS are stronger than self-interactions. (3) The final model would take the form of an icosahedron. These rules lead the model that was ultimately shown to represent the cryo-EM data. A 3D–printed model of the $RsbR_{(146–274)}$:RsbS stressosome core is shown in (**b**), where all STAS domains are coloured *blue*

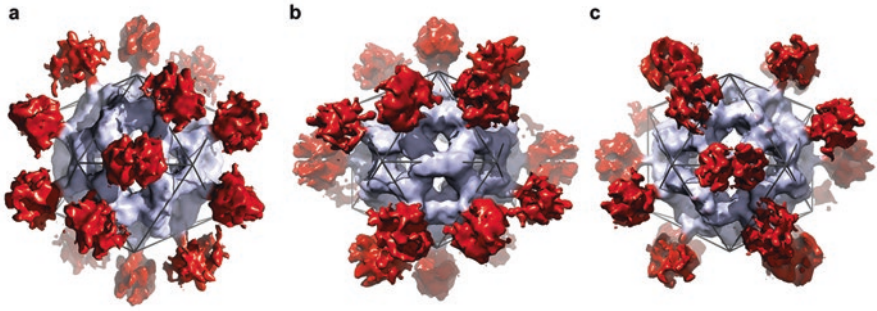

Fig. 1.11 Single particle reconstruction of the RsbR/RsbS stressosome complex. The final experimental EM-derived D_2 symmetry reconstruction of the RsbR/RsbS stressosome complex is shown as a surface contoured at 2σ. The RsbR and RsbS core STAS domains are coloured *blue*, while the N-terminal RsbR signalling domains are coloured *red*. Views down the three unique twofold symmetry axes are shown in panels **a**–**c**, with an icosahedral net for reference. The total radius of the stressosome is 150 Å

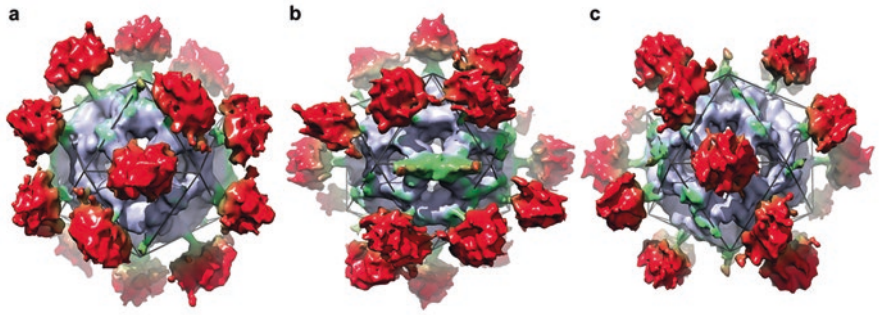

Fig. 1.12 Single particle reconstruction of the RsbR/RsbS/RsbT ternary stressosome complex. The final experimental EM-derived D_2 symmetry reconstruction of the RsbR/RsbS: RsbT ternary stressosome complex is shown as a surface contoured at 1.5σ. The RsbR and RsbS core STAS domains are coloured *blue*, the N-terminal RsbR signalling domains are coloured *red*, the additional density present when this reconstruction is compared to the RsbR/RsbS reconstruction is attributed to the presence of RsbT and is shown in *green*. Views down the three unique twofold symmetry axes are shown in panels **a**–**c**, with an icosahedral net for reference. The total radius of the stressosome is 150 Å as before

1.5.4 Pseudo-atomic Model of the B. subtilis *Stressosome*

With cryo-EM reconstructions available for the core stressosome, RsbR/RsbS binary complex, and RsbR/RsbS/RsbT ternary complex it was possible to dock the crystal structures of the individual stressosome components into the reconstructed density maps to produce a pseudo-atomic model of the stressosome complexes.

Because the turret-like peripheral projections from the RsbR/RsbS and RsbR/RsbS/RsbT reconstructions were absent from the $RsbR_{(146–274)}$/RsbS complex, these regions were attributed to the N-terminal domains of RsbR. The initial model was based on the structure of the *B. subtilis* N-RsbR protein (Murray et al. 2005; Marles-Wright et al. 2008). The crystallographic dimer fitted the density well; however, the model lacked the linker between the N-terminal globin domain and the C-terminal STAS domain; therefore, the density visible between the turrets and core stressosome was not occupied by this model. The subsequent structure of the *M. thermoacetica* RsbR homologue, *Mt*R, had fifteen additional C-terminal residues that formed four turns of an α-helix that extended from the bottom of the globin domain to the N-terminus of the STAS domain (Quin et al. 2012). Docking the *Mt*R model into the stressosome reconstructions accounted completely for the turret density and the neck region between the peripheral turrets and the inner core (Fig. 1.13).

A monomer of the *Mt*S structure was used as a model for the STAS domain and was fitted to the clear secondary structure elements seen in the stressosome core reconstruction (Fig. 1.14). With the position of RsbR determined by the position of the turrets in the RsbR/RsbS reconstructions, it was possible to assign the positions of the RsbR and RsbS STAS domains with reference to the turrets. The three recon-

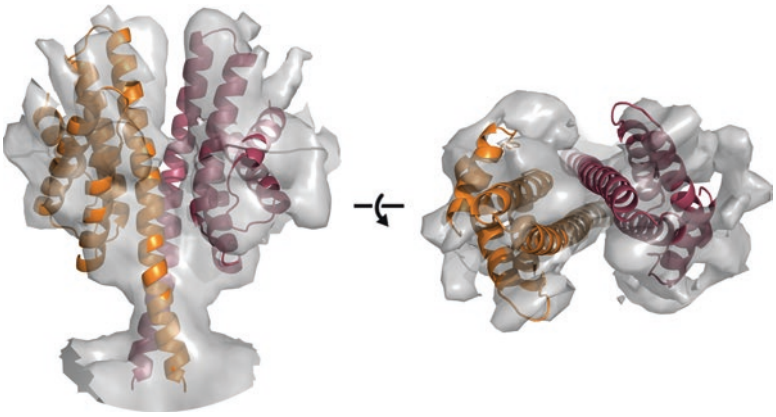

Fig. 1.13 Docking of MtR to the stressosome reconstruction envelope. Orthogonal views of the secondary structure fit of the X-ray crystal structure of MtR in the RsbR/RsbS stressosome reconstruction; the density map is contoured at 2σ. The model of MtR is shown as a secondary structure cartoon coloured *orange* and *red*

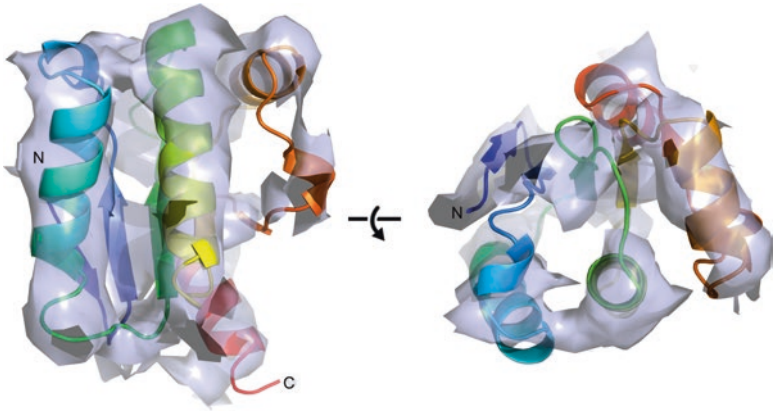

Fig. 1.14 Docking of MtS model to the stressosome core reconstruction envelope. Orthogonal views of the secondary structure fit of the X-ray crystal structure of MtS (PDBID: 3TZA) in the icosahedral stressosome core ($RsbR_{146–274}$/RsbS) reconstruction. The density map is contoured at 2.5σ to emphasise α-helices. The model of MtS is shown as a secondary structure cartoon and colour ramped from *blue* at the N-terminus, to *red* at the C-terminus

structions were aligned to the same coordinate frame and STAS domains below the turrets were assigned as RsbR and those with no turret as RsbS.

The RsbR/RsbS/RsbT ternary complex reconstruction showed clear additional density above the positions assigned as RsbS in the RsbR/RsbS reconstruction; this density was therefore attributed to the presence of the RsbT protein. The co-crystal structure of the SpoIIAA/SpoIIAB complex was used to guide fitting of a SpoIIAB-based homology model of the RsbT kinase in the RsbR/RsbS/RsbT reconstruction