The Role of Flexibility and Protein-Protein Interactions in the Function of the Manganese Stabilizing Protein of Photosystem II

Adele Kim Williamson

A thesis submitted for the degree of Doctor of Philosophy of
The Australian National University
April 2008
This research has been carried out by myself under the supervision of Prof. Tom Wydrzynski and Dr Warwick Hillier. Dr John Liggins also provided helpful discussions and advice in carrying out the calorimetric measurements.

[Signature]
Acknowledgments

I would like to thank the following people:

My supervisors Tom Wydrzynski and Warwick Hillier for giving me the opportunity to do this project, and for their help throughout it. Also John Liggins for his advice and assistance with the calorimetry, without which much of this work would not have been possible. The Photobioenergetics group for generally being a good group of people to share a lab with. Particularly, thanks to my fellow students Brendon Conlan and Iain McConnell for interesting discussion, scientific and otherwise.

The Australian Research Council who granted me an Australian Postgraduate Award, The Australian National University who provided the Slayter scholarship, and the New Zealand Vice-Chancellors’ Committee who granted me the William Georgetti Scholarship.

Finally I must thank my parents Wendy and Bryce Williamson for encouraging my interest in science and nature from a very young age. Mum, I particularly must thank for sharing her enthusiasm about the workings of plants and bacteria, which kindled my interest in these organisms. Dad I must thank for encouraging me to persevere with chemistry as an undergraduate, and for teaching me that thermodynamics is not as hard as it looks. My sister Shannon for always believing in me and encouraging me. My husband Rhys for sticking with me through thick and thin during my studies; first in undergraduate, and then the PhD; even when things weren’t working he always put a smile on my face.

Summary of papers


PS450

iv
Abstract

This research has focused on the manganese stabilizing protein (MSP) of Photosystem II (PSII), an extrinsic subunit which is essential for normal rates of water splitting by this enzyme.

This work extends previous cross-reconstitution studies where the MSP subunit of one species was bound to the PSII complex of a different species. The data show that a non-native fusion protein comprising the MSP subunit from *Thermosynechococcus elongatus* and thioredoxin from *Escherichia coli* (T-MSP-TRX) can rebind spinach PSII and restore higher rates of activity than the native *Spinacia oleracea* (spinach) MSP (S-MSP) or recombinant *T. elongatus* MSP (T-MSP) can. Subsequent characterisation of the thermal unfolding transitions by differential scanning calorimetry (DSC) of isolated S-MSP, T-MSP and T-MSP-TRX fusion protein reveal that T-MSP-TRX has lower thermal stability than T-MSP despite restoring higher rates of activity to the complex. Most interestingly however, these DSC studies reveal that S-MSP completely lacks a cooperative thermal unfolding transition in its isolated (non PSII-bound) state. This is attributed to the previously documented ‘molten globular’ character of S-MSP, and suggest it probably lacks a comparable tertiary structure in isolation. The more compactly folded nature of T-MSP is probably an adaptation to higher temperatures experienced in vivo. The finding that S-MSP lacks a compact structure in isolation led to the hypothesis that tertiary structure increases upon binding to the PSII complex. Isothermal titration calorimetry (ITC) data of S-MSP and T-MSP binding to the PSII complex from spinach support this hypothesis, showing more favourable enthalpy change for binding of S-MSP, and a less unfavourable entropy change for binding of T-MSP. Additionally this work presents the purification and characterisation of the MSP from *Gloeobacter violaceus* (G-MSP). *G. violaceus* is a contemporary cyanobacterium which is thought to represent an intermediate in the evolution of oxygenic photosynthesis. It describes how G-MSP is unable to bind to spinach PSII, and provides in silico models suggesting that this is due to variation in the binding region of PSII and the absence of conserved sequences.

It is argued that the flexibility and non-compact character of the MSP is important for its function during the water splitting reaction, and that this flexibility is altered by interaction of the MSP with other PSII subunits. The possibility that S-MSP forms a more compact structure upon binding the PSII complex, is discussed along with the
importance of this phenomenon when comparing structural data derived from the crystal structure *T. elongatus* PSII with biochemical data from isolated S-MSP. Finally experiments are proposed to investigate the importance of specific residues in MSP function which can be altered by site directed mutagenesis, and future directions for research into the MSP are discussed.
Abbreviations

ANS (1-anilino-naphthalene-8-sulfonate)
ATP (adenosine tri phosphate)
BRC (bacterial reaction centre)
Chl (chlorophyll)
CD (circular dichroism)
CCG (Cluster of carboxylic acid groups)
Cyt b6f (cytochrome b6f complex)
DSC (differential scanning calorimetry)
FD (ferredoxin)
FTIR (fourier transform infrared spectroscopy)
G-MSP (manganese stabilizing protein from Gloeobacter violaceus)
GTP (guanosine tri phosphate)
ITC (isothermal titration calorimetry)
LHCII (light harvesting II antennae complex)
MSP (manganese stabilizing protein)
NMR (nuclear magnetic resonance)
OEC (oxygen evolving complex)
PC (plastocyanin)
PQ (plastiquinone protein)
PSI (photosystem I)
PSII (photosystem II)
QHNDH (quinohemoprotein amine dehydrogenase)
S-MSP (manganese stabilizing protein from spinach)
T-MSP-TRX (fusion protein of T. elongatus MSP with E. coli thioredoxin attached to its N-terminus)
T-MSP (manganese stabilizing protein from, Thermosynechococcus elongatus)
# Table of Contents

**Acknowledgments** ........................................................................................................... iii  
**Summary of papers** .............................................................................................................. iii  
**Abstract** ................................................................................................................................. v  
**Abbreviations** ......................................................................................................................... vii  
**List of Tables** ......................................................................................................................... xiii  
**List of Figures** ......................................................................................................................... xvii  

1. **Introduction** ......................................................................................................................... 1  
1.1 Photosystem II structure and function ................................................................. 1  
  1.1.1 The light reactions of photosynthesis and the role of PSII ............................. 1  
  1.1.2 The PSII subunits ............................................................................................................ 3  
  1.1.3 Electron flow through PSII: charge separation and water oxidation ....... 6  
  1.1.4 The oxygen evolving centre and the S-state cycle ........................................ 7  
  1.1.5 Evolution of PSII ........................................................................................................... 9  
1.2 An overview of the PSII Manganese Stabilising protein ...................................... 12  
  1.2.1 The extrinsic proteins: subunits and differences between species ......... 13  
  1.2.2 Functions of MSP ......................................................................................................... 14  
  1.2.3 Structure of MSP .......................................................................................................... 16  
  1.2.4 Interaction of MSP with the PSII intrinsic proteins ...................................... 23  
  1.2.5 Evolution of MSP ......................................................................................................... 32  
  1.2.6 Role of a possible water channel in MSP ............................................................... 33  
1.3 Research aims ....................................................................................................................... 38  

2. **Methods and Theory** ............................................................................................................. 39  
2.1 Preparation of Spinach MSP .............................................................................................. 39  
  2.1.1 Isolation of PSII-enriched membrane fragments from spinach ................ 39  
  2.1.2 Depletion of the extrinsic proteins from PSII and purification of spinach MSP ................................................................................................................................. 40  
  2.1.3 Reconstitution of extrinsic depleted PSII with MSP ........................................ 40  
2.2 Preparation of recombinant MSP from *T. elongatus* and *G. violaceus* .......... 41  
  2.2.1 Cloning of the psbO gene ............................................................................................. 41  
  2.2.2 Expression and purification of soluble *T. elongatus* MSP ......................... 46  
  2.2.3 Purification of *G. violaceus* MSP by refolding from urea .......................... 48
2.2.4 Purification of G. violaceus MSP by co-expression with chaperonins

2.3 Analytical techniques

2.3.1 Circular dichroism

2.3.2 Fluorescence

2.3.3 Oxygen evolution measurements

2.4 Differential scanning calorimetry

2.4.1 Instrumentation and sample preparation

2.4.2 Thermodynamics of unfolding

2.4.3 Temperature dependence of unfolding

2.5 Isothermal titration calorimetry

2.5.1 Instrumentation and sample preparation

2.5.2 Determination of the thermodynamic parameters

3 Reconstitution of O₂ activity with a recombinant thioredoxin-MSP protein

3.1 Introduction

3.2 Activity of the OEC determined by O₂ evolution

3.3 Discussion

4 Structure and thermal stability of isolated MSP

4.1 Introduction

4.2 Thermal unfolding of spinach MSP

4.2.1 DSC measurements

4.2.2 Effect of calcium

4.2.3 Effect of stabilizing agents

4.2.4 Tryptophan fluorescence

4.2.5 ANS Fluorescence

4.3 Thermal unfolding of T. elongatus MSP

4.3.1 DSC measurements

4.3.2 The effect of disulfide bond reduction

4.3.3 The effect of sarcosine

4.4 Structure and stability of T-MSP-TRX fusion protein

4.4.1 DSC measurements

4.4.2 Circular dichroism

4.4.3 Effect of Calcium and urea

4.5 Discussion

5 MSP binding affinity for PSII in cross-reconstitution experiments
5.1 Introduction .................................................................................................................. 109
5.2 DSC of reconstituted PSII .......................................................................................... 113
5.3 Binding of MSP to PSII measured by ITC ............................................................... 115
5.4 Discussion ................................................................................................................... 118

6 Binding and characterisation of MSP from Gloeobacter violaceus ....... 123

6.1 Introduction ................................................................................................................. 123
6.2 Characterization of recombinant G. violaceus MSP ............................................. 125
6.3 Oxygen evolution from G. violaceus MSP-reconstituted PSII ............................... 126
6.4 Binding of G. violaceus MSP to spinach PSII intrinsic proteins .............................. 127
6.5 Analysis of G. violaceus MSP binding and implications for evolution ................. 127
6.6 Discussion ................................................................................................................... 130

7 General discussion ........................................................................................................ 133

7.1 Comparison of MSP from T. elongatus with MSP from Spinach ........... 133
7.1.1 Differences in Structure ...................................................................................... 133
7.1.2 The basis of protein thermostability .................................................................... 136
7.2 The importance of flexibility for MSP function ...................................................... 144
7.2.1 Summary of flexibility from these experiments ............................................... 144
7.2.2 Examples of other highly flexible proteins and their functions .................... 145
7.2.3 Hypotheses of why MSP is flexible .................................................................... 149
7.3 G. violaceus MSP as a ‘missing link’ .......................................................................... 151
7.3.1 The cyanobacterial periplasm and outer membrane .......................................... 152
7.3.2 The origins of MSP- evolution from a pore? ...................................................... 154
7.4 Conclusions and future directions ........................................................................... 158

References ...................................................................................................................... 161
List of Tables

Table 1-1. The protein subunits of PSII. ................................................................. 5
Table 1-2. Secondary structure of isolated MSP determined by spectroscopy .......... 19
Table 1-3. Summary of site directed mutagenesis of MSP. ..................................... 24
Table 1-4. Cross-reconstitution of MSP-depleted PSII. ........................................... 31
Table 1-5. Summary of cross reconstitution experiments ...................................... 32
Table 2-1. Primers for the amplification of the mature psbO gene .......................... 43
Table 2-2. PCR conditions for amplification of the psbO gene .............................. 44
Table 3-1. O2 yields from cross-reconstituted PSII. ............................................. 65
Table 3-2. Dissociation constants of Calcium for cross-reconstituted PSII .............. 67
Table 4-1. Thermodynamic parameters of the unfolding transition of T. elongatus MSP ........................................................................................................ 79
Table 4-2. Thermodynamic parameters for T. elongatus MSP as a function of of pH 81
Table 4-3. Thermodynamic parameters of the unfolding transition of reduced T. elongatus MSP ........................................................................................................ 83
Table 4-4. ΔG_unf of oxidised and reduced T. elongatus MSP. ............................... 83
Table 4-5. Thermodynamic parameters of T. elongatus MSP unfolding in the presence of sarcosine ........................................................... 85
Table 4-6. Thermodynamic parameters for the unfolding of T-MSP-TRX fusion protein. ........................................................................................................ 89
Table 4-7. Free energies calculated for T. elongatus MSP and T-MSP-TRX .......... 93
Table 4-8. Thermodynamic parameters of unfolding of T. elongatus MSP at varying Ca2+ concentrations ................................................................. 96
Table 4-9. Thermodynamic parameters of unfolding of T. elongatus MSP and T-MSP-TRX fusion protein in the presence of 3M urea ........................................... 99
Table 4-10. Thermodynamics of Ca2+ binding to T. elongatus MSP and T-MSP-TRX ........................................................................................................ 100
Table 5-1. Experimental treatments of PSII carried out in an attempt to assign the ‘A1’ and ‘A2’ thermal unfolding peaks measured by DSC ....................... 111
Table 5-2. Thermodynamics of S-MSP and T-MSP binding to spinach PSII .......... 116
Table 5-3. Thermodynamics of T-MSP-TRX binding to spinach PSII ................. 118
Table 6-1. Conservation of putative water channel residues in PSII .................... 124
Table 6-2. Proportions of secondary structural elements of MSP from *G. violaceus*, *T. elongatus* and Spinach. .................................................................126

Table 6-3. Rates of O2 evolution from spinach PSII reconstituted with MSP from *G. violaceus*, *T. elongatus* and spinach.........................................................127

Table 7-1. Comparison of experiments investigating MSP using either *T. elongatus* or spinach PSII. ..........................................................................................134

Table 7-2. Comparison of the amino acid composition between the MSP from mesophilic cyanobacteria, and the MSP from *T. elongatus*. nce between *T. elongatus* and mesophilic species...........................................140

Table 7-3. Rates of O2 activity restored to extrinsic-depleted PSII when cross-reconstituted with MSP..............................................................142

Table 7-4. Proteins in chemically induced molten globular states..............................................................147

Table 7-5. Examples of molten globular proteins with biological functions........................................149
List of Figures

Figure 1-1. The protein complexes of the light reactions .................................................. 3
Figure 1-2. Electron transport through PSII ........................................................................... 7
Figure 1-3. Kok model of the S-state cycle ............................................................................ 9
Figure 1-4. Crystal structures of the small extrinsic proteins of PSII ................................. 14
Figure 1-5. The crystal structure of MSP from T. elongatus bound to PSII ......................... 17
Figure 1-6. Sequence alignment of MSP from T. elongatus, Spinach and G. violaceus showing conserved regions................................................................. 18
Figure 1-7. Schematic of the depletion/reconstitution of MSP with spinach PSII .......... 24
Figure 1-8. Sequence alignments of MSP from T. elongatus, Spinach and G. violaceus showing residues which are important for interaction with PSII ............................... 28
Figure 1-9. Sequence alignment of T. elongatus MSP, Spinach MSP and G. violaceus MSP showing the putative water channel residues .............................................. 35
Figure 1-10. Putative substrate and product channels in PSII ................................................ 36
Figure 2-1. SDS-PAGE showing purified MSP from spinach ........................................... 40
Figure 2-2. Strategy for cloning the psbO gene from T. elongatus and G. violaceus .......... 43
Figure 2-3. Ligation Independent Cloning strategy ............................................................... 45
Figure 2-4. Map of the cloning and promoter region of pET 32 EK/LIC. ......................... 45
Figure 2-5. SDS-PAGE gel of purified recombinant T. elongatus MSP-TRX fusion protein .................................................................................................................................... 47
Figure 2-6. Over expression and purification of G. violaceus MSP from inclusion bodies .................................................................................................................................. 48
Figure 2-7 Co-expression of G. violaceus MSP-TRX fusion protein with GroEL, and purification of soluble fusion protein ................................................................. 49
Figure 2-8. Schematic representation of a Differential Scanning Calorimetry experiment ................................................................................................................................. 53
Figure 2-9. Schematic representation of an Isothermal Titration Calorimetry experiment ................................................................................................................................. 59
Figure 3-1. Rates of O2 evolution from cross-reconstituted PSII measured at different calcium concentrations .................................................................
Figure 4-1. The thermal unfolding of spinach MSP measured by Differential Scanning Calorimetry. 

Figure 4-2. Differential Scanning Calorimetry scan of S-MSP with the addition of 5mM CaCl₂. 

Figure 4-3. DSC scans of spinach MSP in the presence of sarcosine or reconstitution buffer. 

Figure 4-4. Tryptophan fluorescence from Spinach MSP excited at 278 nm. 

Figure 4-5. Emission spectra of ANS in the presence of Spinach MSP or T. elongatus MSP. 

Figure 4-6. Differential Scanning calorimetry of T. elongatus MSP at pH 6.5. 

Figure 4-7. pH dependence of the Tm of T. elongatus MSP measured by DSC. 

Figure 4-8. Dependence of ΔH_cal on Tm for T. elongatus MSP. 

Figure 4-9. The effect of sarcosine on differential calorimetric scans of T. elongatus MSP. 

Figure 4-10. The effect of sarcosine on the Tm of T. elongatus MSP. 

Figure 4-11. ANS fluorescence in the presence of MSP from T. elongatus with sarcosine. 

Figure 4-12. DSC scan of T-MSP-TRX at pH 6.5. 

Figure 4-13. Thermal unfolding of T-MSP-TRX at different pH, pH2. 

Figure 4-14. pH dependence of Tm for the unfolding of T-MSP-TRX. 

Figure 4-15. Free energies of unfolding for T. elongatus MSP and the MSP portion of the T-MSP-TRX fusion protein. 

Figure 4-16. CD spectra of the T-MSP-TRX fusion protein. 

Figure 4-17. DSC scans of T. elongatus MSP at different Ca²⁺ concentrations. 

Figure 4-18. Dependence of ΔH_cal and Tm of T. elongatus MSP unfolding with increasing Ca²⁺ concentration. 

Figure 4-19. Dependence of ΔH_cal on Tm for T. elongatus MSP when the protein is destabilized by addition of calcium. 

Figure 4-20. DSC scans of T-MSP-TRX fusion protein in the absence of Ca²⁺. 

Figure 4-21. ITC curves of Ca²⁺ binding to T-MSP and Ca²⁺ binding T-MSP-TRX. 

Figure 4-22. Thermogram of unfolding of S-MSP adapted from Kruk et al. (2003). 

Figure 5-1. DSC of PSII adapted from Thompson et al (1986). 

Figure 5-2. DSC scans of MSP-depleted and reconstituted spinach PSII.
Figure 5-3. ITC curves of PSII depleted of MSP, then reconstituted with spinach MSP and *T. elongatus* MSP. .......................................................................................................116

Figure 5-4. ITC curve of PSII depleted of MSP, then reconstituted with T-MSP-TRX. ..................................................................................................................................................118

Figure 6-1. CD spectrum of *G. violaceus* MSP, *T. elongatus* and spinach. .............126

Figure 6-2. The ability of heterologous MSP to bind spinach PSII .........................................................127

Figure 6-3. Sequence alignment of MSP from *T. elongatus*, Spinach, and *G. violaceus*
showing residues thought to be involved in binding to PSII. ....................128

Figure 6-4. 3D structure showing the PSII-binding surface of *T. elongatus* MSP from
and *G. violaceus* MSP ..........................................................................................................................129

Figure 7-1. Sequence alignment of *T. elongatus* MSP with the MSP from eight
mesophilic cyanobacteria ................................................................................................................................141

Figure 7-2. Hypotheses for the basis of thermostability of spinach PSII, and
thermostability of *T. elongatus* PSII. ..........................................................................................................144

Figure 7-3. Structure of the cyanobacterial cell wall .................................................................153

Figure 7-4. Hypothetical scenario for evolution of MSP from a β-barrel outer membrane protein. ........................................................................................................................................156
1 Introduction

Life on earth was changed by the evolution of oxygenic photosynthesis because it is the source of atmospheric O₂, which is used as the terminal electron acceptor in aerobic respiration. Oxygenic photosynthesis has the advantage that water, unlike other reductants, is readily available however, oxidation of water is one of the most thermodynamically demanding reactions in biology, and is achieved only by Photosystem II (PSII). The two main features of PSII that allow it to split water are the P₆₈₀ reaction centre, which is made up a coupled Chl complex, and the ‘oxygen evolving centre’ (OEC) which contains four coupled Mn ions and a Ca ion and templates formation of the O-O bond. Both parts involve bound cofactors acting as electron donors and recipients, but the protein structure surrounding these cofactors is important as well. Here we focus on one of the protein components of the OEC, the manganese stabilising protein, or MSP.

1.1 Photosystem II structure and function

1.1.1 The light reactions of photosynthesis and the role of PSII

Photosynthesis is the process used by organisms to convert light energy from the sun into chemical energy which is then used to drive cellular processes. The photosynthetic reactions can be broadly divided into two phases: the ‘light reactions’, where light energy is absorbed and converted into a chemical potential and reducing equivalents, and the ‘dark reactions’ or the Calvin cycle reactions, where the products of the light reactions are used to assimilate CO₂ into complex organic sugars. Photosynthesis can be classed as either oxygenic or anoxygenic depending on the reductant that ultimately provides electrons in the light reactions. Anoxygenic photosynthesis is carried out by purple bacteria, green sulfur bacteria, green non-sulfur bacteria and heliobacteria. The reductants used in anoxygenic photosynthesis include sulfur containing compounds such as H₂S, reduced organic compounds and H₂. Oxygenic photosynthesis is carried out by cyanobacteria, algae and plants. It has two photosystems joined by an electron transport chain, uses water as the electron donor and releases protons and O₂ as by-products. In addition to fixing carbon which forms the basis of the earth’s food chain,
the O₂ released during oxygenic photosynthesis is used by aerobic organisms for respiration, and maintains the ozone layer which protects the earth’s surface from UV radiation. Water oxidation during oxygenic photosynthesis is carried out by PSII, and is the first step of the light reactions.

The light reactions are carried out by a series of pigment-protein complexes which are embedded in the thylakoid membranes (Figure 1-1). These membrane vesicles partition the outer stromal aqueous phase, which contains the soluble enzymes that carry out the Calvin cycle reactions, and the inner lumenal aqueous phase. Excited electrons are passed through PSII to the membrane-soluble plastiquinone protein (PQ) and are regained by PSII from oxidising water. PQ transfers the electrons to the cytochrome b₆f complex (Cyt b₆f) through a cycle which transfers four protons for every two electrons transferred. From Cyt b₆f the electrons are transferred to Photosystem I (PSI) by the small luminal protein plastocyanin (PC). In PSI light-excited electrons are transferred to ferredoxin (FD) and ultimately to NADP⁺. These electrons are replaced from PC. During electron transport a proton gradient is generated across the thylakoid membrane. This proton motive force drives the synthesis of ATP through the action of the ATP synthase. The details about light reactions and other topics of photosynthesis are covered thoroughly in the book ‘Molecular Mechanisms of Photosynthesis’ by R Blankenship (2002).
1.1.2 The PSII subunits

Several crystal structures of PSII from the thermophilic cyanobacteria *Thermosynechococcus elongatus* and *Thermosynechococcus vulcanus* have been reported, each with increasing and improving resolution (Biesiadka et al. 2004; Ferreira et al. 2004; Fromme et al. 2002; Kamiya and Shen 2003; Loll et al. 2005b; Zouni et al. 2001). The structures at 3.5 Å (Ferreira et al. 2004) 3.2 Å (Biesiadka et al. 2004) and 3.0 Å (Loll et al. 2005b) revealed the detail of most of the amino acid side chains as well as the organisation of the OEC, the site where water oxidation occurs. Only low resolution structures have been reported for eukaryotes (Nield and Barber 2006; Nield et al. 2000a), although further details have been predicted by modelling these on the prokaryotic structure (Nield and Barber 2006).
The PSII complex has a homodimeric structure, with a two-fold axis of symmetry perpendicular to the membrane. Each monomer has one reaction centre, and over 20 individual polypeptide subunits. The subunits can be classed as intrinsic, those embedded in the thylakoid membrane, or extrinsic, proteins attached to the lumenal surface of the intrinsic subunits (Barber and Iwata 2005; Shen and Kamiya 2005; Witt 2005).

The D1 and D2 intrinsic subunits form the centre of the PSII monomer. The subunits D1, D2, CP43, CP47 and Cyt b_{559} form the PSII core, which carries out oxygenic electron transport. D1 and D2 bind all the redox-active cofactors of the acceptor-side electron transport chain. D1 and CP43 bind the \([\text{Mn}_4\text{Ca}]\) metal cluster of the OEC (Nixon 2005). The extrinsic proteins function to stabilise the OEC and provide the correct environment for its optimal function (Bricker and Burnap 2005). All oxygenic photosynthetic organisms have the 33 kDa manganese stabilising protein (MSP) subunit. The subunits of PSII, their functions and cofactors have been extensively reviewed (Barber 2003, 2006a, 2006b; De Las Rivas et al. 2007; Nelson and Yocum 2006; Nield and Barber 2006; Wydrzynski and Satoh 2006), and this information is summarised in Table 1-1.

The distal and proximal antennae proteins associated with PSII facilitate light harvesting and energy transfer to the reaction centre where it is used to drive photochemistry and associated electron transfer reactions. Plants and green algae have the light harvesting II antennae complex (LHCII) which is an integral membrane antennae complex and can form a supercomplex with PSII in the thylakoid membrane. Cyanobacteria and red algae have peripheral antennae proteins called phycobilisomes, which interact with the photosystems at the stromal, and sometimes the lumenal ends (Green and Gantt 2005).
Table 1-1. The protein subunits of PSII. Intrinsic proteins are those which are inserted into the thylakoid membrane, extrinsic proteins are attached to the complex through ionic interactions with other PSII subunits. The core proteins are essential for oxygen production.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Organism</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intrinsic core protein</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1 (PsbA)</td>
<td>All</td>
<td>Reaction centre protein. Binds the P&lt;sub&gt;680&lt;/sub&gt; Chl complex, Chla, pheophytin, and the QA quinone. Also provides most of the ligands to the OEC metal ions, and the tyrosine Y&lt;sub&gt;D&lt;/sub&gt; side chain which transfers electrons from the OEC to P&lt;sub&gt;680&lt;/sub&gt;.</td>
</tr>
<tr>
<td>D2 (PsbD)</td>
<td>All</td>
<td>Reaction centre protein. Binds the redox active quinone Q&lt;sub&gt;B&lt;/sub&gt;. Also has the redox-active tyrosine Y&lt;sub&gt;D&lt;/sub&gt; which is oxidised once by P&lt;sub&gt;680&lt;/sub&gt; after dark adaptation.</td>
</tr>
<tr>
<td>CP43 (PsbC)</td>
<td>All</td>
<td>Core light harvesting protein. Binds Chl and carotenoids and transfers excitation energy to P680 from the light harvesting proteins. Also provides a ligand to the OEC.</td>
</tr>
<tr>
<td>CP47 (PsbB)</td>
<td>All</td>
<td>Core light harvesting protein. Binds Chl and carotenoids and transfers excitation energy to P&lt;sub&gt;680&lt;/sub&gt;.</td>
</tr>
<tr>
<td>Cyt b559 (PsbE, large subunit, PsbF small subunit)</td>
<td>All</td>
<td>A heterodimeric subunit which binds heme. Stabilises the PSII core and may protect against photo-inhibition.</td>
</tr>
<tr>
<td><strong>Core extrinsic proteins</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manganese stabilising protein (MSP, 33 kDa protein, PsbO)</td>
<td>All</td>
<td>Forms part of a luminal ‘cap’ of proteins over the OEC. Stabilises the metal cluster and probably regulates the entry of substrate water and/or the exit of protons.</td>
</tr>
<tr>
<td><strong>Small extrinsic proteins</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PsbP (23 kDa protein)</td>
<td>Higher plants and algae, and some PSII centres in cyanobacteria.</td>
<td>Together with MSP and PsbQ, caps the OEC and optimises Ca&lt;sup&gt;2+&lt;/sup&gt; and Cl&lt;sup&gt;−&lt;/sup&gt; ions needed for water oxidation. Essential for normal rates of water splitting.</td>
</tr>
<tr>
<td>PsbQ (17 kDa protein)</td>
<td>Higher plants and algae, and some PSII centres in cyanobacteria. Absent in G. violaceus.</td>
<td>Together with MSP and PsbP, caps the OEC and optimises Ca&lt;sup&gt;2+&lt;/sup&gt; and Cl&lt;sup&gt;−&lt;/sup&gt; ions.</td>
</tr>
<tr>
<td>PsbR (10 kDa)</td>
<td>Higher plant</td>
<td>Unknown, but appears to be associated with the OEC.</td>
</tr>
<tr>
<td>PsbV (17 kDa protein, cytochrome c 550)</td>
<td>Cyanobacteria</td>
<td>Helps stabilise the OEC, and is necessary if MSP is deleted.</td>
</tr>
<tr>
<td>PsbU (12 kDa protein)</td>
<td>Cyanobacteria</td>
<td>Helps stabilise the OEC, and is necessary if MSP is deleted.</td>
</tr>
</tbody>
</table>
Chapter 1. Introduction

### Small intrinsic proteins

<table>
<thead>
<tr>
<th>Protein</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psb L, M and T</td>
<td>All</td>
</tr>
<tr>
<td>Psb I and X</td>
<td>All</td>
</tr>
<tr>
<td>PsbJ, K, N, Z</td>
<td>All except G. violaceus lacks PsbZ</td>
</tr>
<tr>
<td>PsbY</td>
<td>All except G. violaceus</td>
</tr>
<tr>
<td>Psb W</td>
<td>Higher plants</td>
</tr>
</tbody>
</table>

\(^a\) (Nixon 2005)  
\(^b\) (Bricker and Burnap 2005)  
\(^c\) (Thornton et al. 2005)

#### 1.1.3 Electron flow through PSII: charge separation and water oxidation.

Light energy is absorbed by the antenna complexes, and transferred to the reaction centre through the CP43 and CP47 proteins (Eaton-Rye and Putnam-Evans 2005). From here it excites a group of Chl \(a\) molecules called \(P_{680}\) which are bound by the D1 and D2 proteins. In a few picoseconds, the excited state \(P_{680}^*\) transfers an electron to a pheophytin to generate the radical pair \(P_{680}^{*+}\) Phe\(^-\) (Figure 1-2). In a few hundred picoseconds the electron is passed from Phe\(^-\) to a quinone QA on D2, forming \(P_{680}^{*+}\) Phe QA\(^-\). After a few nanoseconds \(P_{680}^{*+}\) oxidises Tyr 161 (Yz), of the D1 protein, to form Y\(^+\) Phe \(P_{680}\) QA\(^-\), which is probably deprotonated (Diner and Britt 2005). QA\(^-\) in a few milliseconds reduces a quinone on D1, \(Q_B\) to form Y\(^-\) Phe QA \(Q_B^+\) (Renger 1997). At the same time, Yz extracts an electron, and probably a proton from the [Mn4Ca] cluster. After a second photochemical turnover \(Q_B\) is protonated to plastoquinol (PQH\(_2\)) and released to take the electrons to Cyt \(b_{5\alpha}\). After two more turnovers four oxidising equivalents have accumulated on the [Mn4Ca] cluster, which regains four electrons from two molecules of water, releasing H\(^+\) and O\(_2\). For reviews see Barber (2003) and Renger and Holzwarth (2005) and references therein.
Figure 1-2. Electron transport through PSII. The protein structure shown in green is D1, the structure shown in blue is D2. 1. Transfer of an excited electron from P$_{680}$ (red) to Phe (yellow). 2. Reduction of Q$_A$ (light blue) by P$_{680}^{-}$Phe$. 3. Oxidation of Y$_z$ (white) by P$_{680}^{-}$Phe Q$_A$. 4. Extraction of an electron from the [Mn4 Ca] cluster (metal ions shown in purple, putative bicarbonate shown in blue). 5. Transfer of an electron to Q$_B$ (light blue). Image produced from the 1SSL crystal structure using Pymol.

1.1.4 The oxygen evolving centre and the S-state cycle

Four photochemical turnovers are required to oxidise two molecules of water to one molecule of O$_2$. Water oxidation occurs in the OEC which comprises regions of the intrinsic proteins, and extrinsic proteins as well as the [Mn$_4$Ca] cluster (Debus 2000, 2006). The [Mn$_4$Ca] cluster is made up of four Mn ions, one Ca ion, and possibly includes Cl$^-$ and bicarbonate cofactors (van Gorkom and Yocum 2005; van Rensen and Klimov 2005). Its organisation has been determined from recent crystal structures (Ferreira et al. 2004; Loll et al. 2005b) and appears to be a $3+1$ arrangement with three Mn ions and the Ca ion forming a cube-like structure joined by oxo-bridges, and the fourth Mn joined to one of the cube Mn ions via an oxygen. The details of the [Mn$_4$Ca] cluster structures are reviewed in (Barber 2006a, 2006b; Biesiadka et al. 2004; Kern et al. 2007; Yachandra 2005; Yano et al. 2006). The ligands to the metal ions are carboxylic acid and histidine side chains from the luminal loop regions of CP43 and D1.
The extrinsic proteins also bind to the lumenal loops of D1, D2 and CP43. Although they do not provide direct ligands to the [Mn₄Ca] cluster, the extrinsic proteins form a cap over the OEC which provides the necessary environment for water oxidation. The function of the extrinsic proteins in water oxidation is covered in detail in the next section, with a particular focus on the MSP.

During water oxidation the [Mn₄Ca] cluster goes through the ‘S-cycle’ with five intermediate ‘S-states’ (S₀ to S₄). Four of the S-states are generated by sequential photochemical turnovers, while the fifth corresponds to release of O₂ and resetting of the cycle. Figure 1-3 shows the Kok-model (Kok et al. 1970) of the S-cycle, which outlines the general five-step scheme. Many advances have been made recently in understating the S-cycle, using a combination of structural and spectroscopic methods; however some details remain a matter of debate. The S₀ → S₁ and S₁ → S₂ transitions are Mn centred oxidations, but it is not agreed whether the Mn ions, or substrate water are oxidised on the S₂ → S₃. In the latest results (Kulik et al. 2007) the oxidation states of the Mn ions at each step of the S-cycle have been assigned: S₀, Mn₄(III,III,III,IV) S₁, Mn₄(III,III,IV,IV) S₂ Mn₄(III,IV,IV,IV), however which of the Mn atoms in the cluster correspond to which oxidation state has not been determined. At least one water molecule binds to the [Mn₄Ca] cluster in the S₀ state (Hillier and Wydrzynski 2000, 2004), but it is not known which metal ions provide the binding sites for the water molecules. The current theories of the molecular mechanism of water oxidation have been reviewed in the following references:(Ahrling et al. 2005; Barber 2003; Britt et al. 2004; Fromme et al. 2002; Haumann et al. 2005a; Haumann et al. 2005b; Hillier and Wydrzynski 2004; Messinger 2004; Razeghifard et al. 2005). More recently an 8-step ‘I-cycle’ has been proposed as an extension of the S-cycle (Dau and Haumann 2005, 2007). In the I-cycle, protons and electrons are removed alternately leading to 9 intermediate state of the [Mn₄Ca] cluster.
Chapter 1. Introduction

Figure 1-3. Kok model of the S-state cycle. The S-states are intermediate states of the [Mn4 Ca] cluster during water oxidation. The steps $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow S_3$ are all progressed by photochemical turnovers. $S_4 \rightarrow S_0$ is the step where the $O_2$ is released and two water molecules rebind. This figure represents a generalised scheme only. For details see (Kok et al. 1970) and references listed in the text.

The rates of exchange of water molecules between the PSII active site and the bulk solvent have been measured by mass spectrometry (Hillier et al. 1998; Messinger et al. 1997; Wydrzynski et al. 1996). These experiments show that the two molecules bind with different affinities, with one having a much slower exchange rate than the other. This suggests that the two water molecules have different binding sites (Hillier and Wydrzynski 2000, 2001, 2004). The rates of water exchange are altered by removal of the extrinsic proteins (Hillier et al. 2001). The implications of this observation for the function of the MSP are discussed in Section 1.2.6.

1.1.5 Evolution of PSII

Studies concerning the evolution of photosynthesis, water oxidation and the chloroplast span disciplines from geology to biochemistry to genetics. For general reviews on these topics see Blankenship (1998, 2002), Hashimoto (2003), Rutherford (2003), Olson (2004), Dismukes (2005) and references therein. All modern reaction centres have dimeric membrane spanning core proteins. The reaction centres of purple bacteria, green non-sulfur bacteria, PSI and PSII are heterodimers, whereas those of green sulfur
bacteria and heliobacteria are homodimers. The protein precursor to these reaction centres would have been monomeric but not necessarily photosynthetic. The ancestral monomer probably developed the ability to dimerise, and then gene duplication allowed evolution of the heterodimeric complex (Sadekar et al. 2006). This duplication seems to have occurred independently during evolution of PSI, PSII and the bacterial reaction centre (BRC) of purple bacteria, suggesting a strong selection pressure for a heterodimer (Sadekar et al. 2006).

Depending on their cofactors, reaction centres can be classed as Type I, containing Fe-S electron acceptors and found in green sulfur bacteria, heliobacteria and PSI; or as Type II containing pheophytin-quinone electron acceptors and found in purple bacteria, green non-sulfur bacteria and PSII. It is not clear whether the co-occurrence of the Photosystems I and II in cyanobacteria, plants and algae arose from fusion by lateral gene transfer between two separate organisms, or whether those organisms which only have type I or type II originally had both and then selectively lost the other system. Nor is it known whether water oxidation had already evolved in the PSII precursor before it became linked with PSI through the electron transport chain (Olson and Blankenship 2004).

Structural alignments of the D1 and D2 subunits of PSII with the L and M subunits of B-RC show these have a high degree of structural homology, and the two shared a common ancestor, probably before the gene duplication event which produced a heterodimer (Sadekar et al. 2006). The greatest similarity is in the transmembrane sections and the lowest similarity is in the loop regions connecting them (Sadekar et al. 2006). The primary sequences appear to have diverged considerably and have only 20% homology (Sadekar et al. 2006). D1 and D2 have about 40% sequence homology with each other (Sadekar et al. 2006). CP43 and CP47 are related to each other, and also distantly to the antennae domain of type I reaction centres (Blankenship 1992; Raymond and Blankenship 2004).

To oxidise water the PSII precursor had to develop a more oxidising redox potential, and the ability to store oxidising equivalents. The former was achieved by evolution of chlorophyll pigments that can utilise shorter wavelengths of light than the more primitive bacteriochlorophylls and alteration of the protein environment to increase the redox potential of these Chls at the reaction centre (Rutherford and Faller 2003). The latter step involved the evolution of the [Mn₄Ca] cluster. This allows the four oxidising equivalents that are needed to split water to be accumulated one at a time,
overcoming the problem that they are generated from P$_{680}$ by one electron processes. It is hypothesised that the [Mn$_4$Ca] cluster complex evolved from incorporation of an enzyme with a bi-nuclear Mn catalytic centre. One such enzyme which has been considered is Mn catalase. The binuclear Mn active site of Mn catalase was mapped onto the [Mn$_4$Ca] cluster of PSII, and showed the surrounding protein ligands have some spatial similarity to the PSII ligands (Raymond and Blankenship 2008). However it is not clear whether this similarity is a consequence of the geometry needed to bind Mn, or evidence for a common evolutionary history (Raymond and Blankenship 2008). Intermediates in the evolution of [Mn$_4$Ca]-containing PSII are proposed to have utilised electron donors that are easier to oxidise than water (Blankenship and Hartman 1998; Dismukes et al. 2001). Suggestions include H$_2$O$_2$ and bicarbonate, although bicarbonate is not used as a substrate by modern PSII (Hillier et al. 2006), and early earth probably had very little H$_2$O$_2$ before oxygenic photosynthesis evolved (however see (Liang et al. 2006)).

During the evolution of water oxidation, the PSII precursor also greatly increased the number of subunits from the three or four of the BRC and other reaction centres, to the more than 20 found in the modern PSII. These include the extrinsic proteins, Psb U and V in cyanobacteria, PsbP and Q in plants, and the MSP which is a constituent of all oxygenic photosystems. The extrinsic proteins have no sequence homology with proteins from other non-oxygenic reaction centres suggesting a different evolutionary origin. The extrinsic proteins have no sequence homology with proteins from other non-oxygenic reaction centres suggesting a different evolutionary origin. The subunits PsbP, PsbQ, PsbU and PsbV (Cyt c550) have some sequence homology with non-photosynthetic proteins, but the MSP does not have sequence similarity with any other protein identified yet (De Las Rivas and Roman 2005). This is particularly interesting in the case of the MSP, as this subunit is necessary for functional water splitting, and thus its incorporation into the PSII precursor must have been an important event in the evolution of water oxidation. The sequence and structural homology of MSP with other proteins is discussed in detail in Section 1.2.5.

The light reactions take place in the extensive thylakoid membrane network which in plants and algae is enclosed in the chloroplast. In early oxygenic photosynthetic organisms the PSII, PSI and other light reaction complexes were probably localised in the cytoplasmic membrane along with proteins involved in respiration and solute transport. A living relic of this is the primitive contemporary cyanobacterium *Gloeobacter violaceus* which lacks thylakoid membranes, and is the
subject of study in Chapter 6. The thylakoid membranes are thought to have formed in prokaryotic cyanobacteria from invagination of the plasma membrane to increase surface area, and eventually became internalised. The chloroplast of eukaryotic higher plants and algae is the result of an endosymbiotic event where a cyanobacterial cell was engulfed by a larger unicellular eukaryotic cell (Mcfadden 2001; Raven and Allen 2003). Over time many genes were transferred to the eukaryotic nucleus, although chloroplasts retain some circular bacterial-like DNA as well as prokaryotic type ribosomes and other features. Multiple endosymbiotic events occurred in the evolution of many algae, which have chloroplasts surrounded by multiple membranes (Mcfadden 2001).

1.2 An overview of the PSII Manganese Stabilising protein

The MSP is essential for maintaining maximum rates of oxygen evolution during photosynthesis. Functional aspects have been studied in vivo predominantly in the cyanobacterium Synechocystis, as knockout mutants have been made for the psbO gene, allowing insertion of plasmids which express MSP with point mutations. The function of MSP has been studied in vitro using PSII-enriched membranes of Spinacia oleracea (spinach). The majority of low resolution structural studies of isolated (non-PSII bound) MSP have been done in spinach from both native and recombinant sources, including most of the work on the structural dynamics of MSP. All of the high resolution structures of bound MSP are from the thermophilic cyanobacteria T. elongatus, and T. vulcanus. Much of the structural and functional study of MSP reviewed in these sections, as well as the experiments presented in subsequent chapters was done on the MSP from spinach and T. elongatus and these proteins are referred to as S-MSP and T-MSP respectively from here on. The MSP from G. violaceus is also of interest as it has low sequence conservation. MSP from this species is referred to as G-MSP

A number of reviews about the structure and function of the extrinsic proteins of PSII are available (Bricker and Burnap 2005; Bricker and Frankel 1998; De Las Rivas et al. 2007; Rivas et al. 2004; Roose et al. 2007b; Seidler 1996a), and will be referred to throughout this section.
1.2.1 The extrinsic proteins: subunits and differences between species

The MSP extrinsic protein is the most important for O$_2$ evolution of all the extrinsic subunits of PSII. The MSP common to all oxygenic photoautotrophs and two strains of the cyanobacterium *Prochlorococcus marinus* lack the genes encoding any extrinsic proteins other than MSP (Rivas et al. 2004). Additionally the decrease in oxygen evolution from dissociation of the small extrinsic proteins in isolated PSII can be largely compensated for by addition of Ca$^{2+}$ (Ghanotakis et al. 1984b), but removal of MSP irreversibly decreases the rate of oxygen production (Bricker 1992).

The small extrinsic components of PSII vary depending on the species—higher plants and green algae have PsbP and PsbQ proteins whereas prokaryotic cyanobacteria and eukaryotic red algae have a PsbU and PsbV (Enami et al. 2000). Cross reconstitution experiments indicate differences in the small extrinsic binding sites between higher plants and green algae (Suzuki et al. 2005). Higher plants also have PsbR, and red algae have an additional 20kDa protein, but neither of these appears to be directly involved in water oxidation (De Las Rivas et al. 2007; Enami et al. 1998b). The genomes of most cyanobacteria also include *psbP*- and *psbQ*-like genes (Rivas et al. 2004). Recently it was found that in *Synechocystis* PCC 6803 both the PsbP and PsbQ homologues are associated with the PSII complex in addition to PsbU and PsbV. The PsbP is associated with only some centres, but PsbQ is bound to most PSII, and these have higher activity that those lacking PsbQ (Ishikawa et al. 2005; Roose et al. 2007a; Thornton et al. 2004). Although PsbU and PsbV are functionally replaced by PsbP and PsbQ in plants, there is very little sequence or structural homology between PsbP and PsbQ, and PsbU and PsbV (Nelson and Yocum 2006). PsbP is a mostly β-sheet protein (Figure 1-4), and has some structural similarity to a GTPase from yeast (De Las Rivas and Roman 2005; Ifuku et al. 2004). It is predicted to bind to the head domain of MSP in higher plants, but binds to PSII independently of the other extrinsic proteins in green algae (De Las Rivas et al. 2007; Seidler 1996a). PsbQ is an α-helical protein with a flexible N-terminus. Interestingly, PsbV has a typical class I c-type cytochrome fold with four α-helices, and the N-terminus has some β-sheet structure (Bricker and Burnap 2005). The heme group of PsbV has a low midpoint potential and does not participate in electron transport (Bricker and Burnap 2005; De Las Rivas et al. 2007). PsbU is mostly α-helical.
It is hypothesised that PsbP and PsbQ play a role in oligomerisation of PSII on opposite sides of the thylakoid membrane during granal stacking, and this explains why these proteins replaced PsbV and PsbU (De Las Rivas et al. 2007).

### 1.2.2 Functions of MSP

The importance of MSP for normal rates of $\text{O}_2$ evolution under physiological conditions has been demonstrated *in vivo* by making knockout mutants of the *psbO* gene to produce mutants expressing MSP-less PSII (Al-Khaldi et al. 2000; Mayfield et al. 1987; Philbrick et al. 1991; Yi et al. 2005) and *in vitro* by biochemically releasing the MSP from the PSII complex of PSII-enriched membranes extracted from normal plant leaves (Bricker 1992; Ono and Inoue 1984). Suggested functions of the MSP include stabilizing the active site metal complex (Bricker 1992; Miyao and Murata 1984c) and protection against photoinhibition (Henmi et al. 2004; Henmi et al. 2003b; Yamamoto 2001; Yamamoto et al. 1998).

Mutants of the cyanobacterium *Synechocystis* which lack the *psbO* gene accumulate normal levels of the intrinsic core proteins, and can evolve $\text{O}_2$ at about 70% of the wild-type (Burnap and Sherman 1991; Philbrick et al. 1991). Measurements of
electron transport and measurement of the pattern of flash $O_2$ yield in these cells show that the same numbers of PSII centres are assembled and that $O_2$ evolution occurs by the same basic mechanism. This suggests that, in cyanobacteria at least, the MSP is not essential for PSII function (Burnap et al. 1994). However these $\Delta psbO$ mutants require a supply of $Ca^{2+}$ and $Cl^-$ to grow photoautotrophically, have a greater sensitivity to photoinhibition, have fewer centres which are coupled to $O_2$ evolution and have altered kinetic properties in the S-states of the centres which do oxidise water (Burnap et al. 1994; Burnap and Sherman 1991; Philbrick et al. 1991). Single gene deletion mutants of PsbV can also grow photoautotrophically at reduced rates, but a double deletion mutant which lacks both MSP and PsbV cannot (Al-Khaldi et al. 2000; Morgan et al. 1998). Deletion and mutagenesis of MSP in *Synechocystis* also alters photo-assembly of the [Mn$_4$Ca] cluster (Qian et al. 1997). In contrast to cyanobacteria, deletion of *psbO* from green algae and higher plants abolishes $O_2$ evolution (Mayfield et al. 1987; Yi et al. 2005). Mutants of *A. thaliana* and *C. reinhardtii* that lack the *psbO* gene could not correctly assemble PSII centres and cold not grow photoautotrophically, suggesting that the MSP has a slightly different function in cyanobacteria to plants and algae.

Experiments carried out *in vitro* on PSII-enriched membrane fragments from spinach show an essential role for the MSP in oxygen evolution. Treatment of thylakoid membranes extracted from spinach leaves with detergent separates PSII from PSI, producing PSII-enriched membrane fragments that are capable of high levels of $O_2$ evolution (Berthold et al. 1981; Ford and Evans 1983). The small extrinsic proteins can be specifically dissociated from the PSII complex by washing with NaCl (Ghanotakis et al. 1984a; Ghanotakis et al. 1984c). The MSP can be removed by washing with high concentrations of Tris (Yamamoto et al. 1981), chaotropic agents (Miyao and Murata 1984b) or $CaCl_2$ (Ono and Inoue 1984). In the absence of the MSP the [Mn$_4$Ca] cluster becomes unstable, and two of the four [Mn$_4$Ca] cluster ions of the OEC are released into solution (Miyao and Murata 1984b). Addition of high concentrations of $Cl^-$ during removal of the MSP from PSII stabilises the [Mn$_4$Ca] cluster, producing an MSP-depleted PSII preparation which retains the OEC (Miyao and Murata 1984c). This MSP-depleted PSII complex evolves $O_2$ at $\sim$20% of the original rate of the PSII-enriched fragments if additional $Ca^{2+}$ is present (Bricker 1992; Ghanotakis et al. 1984a; Miura et al. 1997; Miyao and Murata 1984b, 1984c; Murata et al. 1984; Ono and Inoue 1984). Removal of MSP by these biochemical methods stabilises the $S_2$ state, and
inhibits or retards the $S_3-S_4-S_0$ transition (Miyao et al. 1987). The MSP can functionally rebind to this extrinsic depleted PSII, which restores an average of 40% of the original $O_2$ evolving activity (Miyao and Murata 1984c; Ono and Inoue 1984).

1.2.3 Structure of MSP

The structure of MSP has been studied by a number of spectroscopic and biochemical techniques, which provided valuable information in the pre-crystal structure era (reviewed in Bricker and Frankel (1998) De Las Rivas and Heredia (1999) and Seidler (1996a). The results of these experiments are largely in agreement with the structure of PSII-bound MSP in the cyanobacterial crystal structure, and have been used to make predictions about MSP from higher plants (for reviews see Bricker and Burnap (2005), De Las Rivas (2004), Barber (2004) and De Las Rivas and Roman (2005)).

The overall structure of T-MSP bound to PSII has two domains: a $\beta$-barrel core made up of eight antiparallel $\beta$-sheets and an extended head domain inserted between strands $\beta$ 5 and 6 (Figure 1-5). The centre of the $\beta$-barrel is filled with bulky hydrophobic residues, and the head domain is mostly loops and turns with a single $\alpha$-helix. For an in-depth analysis of the structure of T-MSP bound to PSII see (De Las Rivas and Barber 2004).
Chapter 1. Introduction

Head domain
N-terminus

Barrel domain

Cyano loop

Figure 1-5. The crystal structure of MSP from *T. elongatus* bound to PSII. Elements shown in blue are α-helical, elements shown in purple are β-strand, and elements shown in pink are loops. Image produced from PDB,1S5L using the Pymol program.

Primary structure of MSP and conserved regions. MSP is expressed with an N-terminal transit sequence which is cleaved before binding to PSII (Seidler 1996a; Theg and Shi 2005). A multiple alignment of all known MSP sequences shows only 19 residues are completely conserved, and an additional 11 are functionally conserved, having side chains with similar properties. Most conserved residues are found in the highly conserved regions identified in references by Heredia (2003) and De Las Rivas (2004) and shown in Figure 1-6. These include the Cys residues involved in formation of the single disulfide bridge, and sequences involved in binding to PSII. Cyanobacteria have a cyano-loop insertion which protrudes form the barrel in the folded structure (Figure 1-5). Plants have an N-terminal insertion, a glutamate rich region, an Arg-Asp-Gly and a single Trp near the C-terminus. G-MSP has the shortest sequence of any species, and is missing parts of the β1-β2 and N-t loops (Figure 1-5 and Figure 1-6) (De Las Rivas and Barber 2004). All species have a highly conserved DPKGR region which is closest to the [Mn₄Ca] cluster in the *T. elongatus* structure, and is probably essential for its function in stabilising the OEC.
**Figure 1-6.** Sequence alignment of MSP from *T. elongatus* (Therm), Spinach (Spin) and *G. violaceus* (Glo) showing conserved regions. Residues shown in green text are conserved; residues shown in red are not conserved. Residues marked with a * are fully conserved in all species according to the multiple sequence alignment of De Las Rivas (2004). Residues marked with $ are functionally conserved in all species. Residues marked with + are fully conserved in all species except *G. violaceus*. Residues marked with # are functionally conserved in all species except *G. violaceus*. Highly conserved regions identified by De Las Rivas (2004) are boxed in red and numbered 1 through 5. Non conserved regions between plants, cyanobacteria and *G. violaceus* are boxed and labelled in blue. Alignment was produced using CLUSTALW.

**Secondary structure of MSP.** The MSP is a β-barrel protein, with a high content of unstructured loops and turns, and a small amount of α-helix. The structure of the cyanobacterial MSP subunit in its PSII bound form has been solved in the high resolution structures of PSII (Ferreira et al. 2004; Loll et al. 2005b), but no high resolution structural studies of isolated MSP have been reported, and there is no structure available for higher plant PSII, or any other mesophilic organism. Isolated MSP has been characterised by biochemical and biophysical studies, the results of which are summarised in Table 1-2.
Table 1-2. Secondary structure of isolated MSP determined by spectroscopy.

<table>
<thead>
<tr>
<th>Species</th>
<th>Technique</th>
<th>α-helix</th>
<th>β-sheet</th>
<th>Unstructured and turns*</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinach</td>
<td>FTIR¹</td>
<td>27</td>
<td>36</td>
<td>37</td>
<td>(Ahmed et al. 1995)</td>
</tr>
<tr>
<td>Spinach</td>
<td>FTIR</td>
<td>10</td>
<td>48</td>
<td>42</td>
<td>(Heredia and De Las Rivas 2003)</td>
</tr>
<tr>
<td>Spinach</td>
<td>CD²</td>
<td>4</td>
<td>47</td>
<td>49</td>
<td>(Lydakis-Simantiris et al. 1999)</td>
</tr>
<tr>
<td>Spinach</td>
<td>CD</td>
<td>8</td>
<td>33</td>
<td>59</td>
<td>(Shutova et al. 1997)</td>
</tr>
<tr>
<td>Spinach</td>
<td>CD</td>
<td>3</td>
<td>38.5</td>
<td>58.5</td>
<td>(Kruk et al. 2003)</td>
</tr>
<tr>
<td>T. elongatus</td>
<td>CD</td>
<td>19</td>
<td>24</td>
<td>57</td>
<td>(Sonoyama et al. 1996)</td>
</tr>
<tr>
<td>T. elongatus</td>
<td>CD</td>
<td>8</td>
<td>43</td>
<td>48</td>
<td>(Loll et al. 2005a)</td>
</tr>
<tr>
<td>T. elongatus</td>
<td>FTIR</td>
<td>23.6</td>
<td>32.5</td>
<td>43.9</td>
<td>(Loll et al. 2005a)</td>
</tr>
</tbody>
</table>

* Loops, turns and other
¹ Fourier transform infrared spectroscopy (FTIR)
² Circular dichroism (CD)

In most cases FTIR (Fourier transform infrared spectroscopy) predicts a higher proportion of α-helix than CD (Circular dichroism). The average α-helix for S-MSP is 13 %, β-sheet is 30 % sheet and unstructured 49 %. The average for T-MSP is α-helix 17 %, β-sheet is 33 % and unstructured is 50 %. This suggests that S-MSP has slightly more β-sheet than T-MSP, and slightly less α-helix, but both have the same amount of unstructured regions, loops and turns.

Importance of the disulfide bond. The MSP from all species has two conserved cystine residues that form an intramolecular disulfide bond. Reduction of the disulfide in S-MSP by β-mercaptoethanol or dithiothreitol changes the CD spectrum to resemble that of the guanidine denatured protein, and shifts the intrinsic fluorescence maximum from 314 to 350 (Tanaka et al. 1989). Unfolding experiments of isolated S-MSP estimated that reduction of the disulfide bind contributes 4.7 Kcal mol⁻¹ to the free energy change of unfolding (Tanaka et al. 1989; Tanaka and Wada 1988). This chemically reduced MSP had poor PSII binding, and could not restore oxygen evolving activity (Tanaka and Wada 1988). A site directed mutant of Synechocystis which had one of the Cys replaced with a serine could not accumulate MSP, probably because of mis-folding and proteolytic digestion after expression (Burnap et al. 1994). In contrast, site directed mutants of S-MSP which lacked one or both Cys were produced recombinantly in E.
coli and were able to rebind MSP-depleted PSII and restore activity (Betts et al. 1997a). Both the single and double mutants were less stable in their isolated (non-PSII bound) form compared to the wild type, and both could be more easily extracted from PSII with urea (Betts et al. 1997a). Further characterisation of the double mutant showed that it has a perturbed solution structure compared to the wild type, with greater solvent exposure of internal Trp and Tyr residues, and larger apparent molecular weight measured by gel filtration (Wyman and Yocum 2005b). Despite this, the double Cys mutant is able to refold after heating to 90 °C, and the refolded form is still functional (Wyman and Yocum 2005b).

The effect of metal ions and pH on MSP structure. The high resolution crystal structures of cyanobacterial PSII have identified two Ca ions in the vicinity of the OEC (Ferreira et al. 2004; Loll et al. 2005b). One of these is the Ca ion of the [Mn₄Ca] cluster which is essential for water oxidation; the other Ca ion is bound to the MSP near its binding interface with the intrinsic proteins (Murray and Barber 2006). The ligands to this ion are Glu114 which is highly conserved in other species, Glu54 which is less conserved, and His231 which is unique to T. elongatus and is a lysine in most other species (Murray and Barber 2006).

Biochemical and biophysical approaches using spinach PSII have predicted a number of Ca²⁺ binding sites associated with the extrinsic proteins. Using a Ca²⁺ sensitive electrode two high affinity binding sites functional in PSII as well as a number of ‘non-specific’ low affinity sites were observed (Grove and Brudvig 1998). Ca²⁺ can functionally substitute for the loss of the PsbP. The PsbP protein modulates the binding kinetics of Ca²⁺ to PSII (Adelroth et al. 1995). When PsbQ and PsbP are removed by NaCl treatment a Ca²⁺ ion is also released (Miyao and Murata 1984a) and Ca²⁺ must be present during rebinding of the PsbP protein to depleted PSII in order to restore activity (Ghanotakis et al. 1984a). Ca²⁺ may also be able to substitute for MSP, as in cyanobacterial mutants where MSP has been deleted Ca²⁺ was required for photoautotrophic growth (Engels et al. 1994; Philbrick et al. 1991). Ca²⁺ binding motifs have been predicted in MSP primary sequence of higher plants (Wales et al. 1989). Calcium binding to isolated S-MSP induces conformational changes which increase the loop content of the protein and thermally destabilise it (Heredia and De Las Rivas 2003; Kruk et al. 2003), although similar structural changes have not been observed in T. elongatus MSP (Loll et al. 2005a). It has also been proposed that calcium is part of an
activation process for converting the MSP to an ‘open’ conformation which allows controlled access of the substrate water to the active site. At low pHs such as would occur in the lumen during photosynthesis, the MSP core was observed to become more accessible to solvent, (Shutova et al. 1997; Shutova et al. 2003) and able to bind calcium. Binding of calcium causes further structural changes, decreasing the thermal stability of the MSP (Shutova et al. 2005). pH dependent structural changes have been measured in the OEC, and could be related to pH induced structural changes in MSP (Kebekus et al. 1995; Renger et al. 1994). Measurement of the buffering capacity of S-MSP showed that $32 \text{H}^+$ are bound in the pH 3 – 7 region. For about 11 of these the pKa of the amino acid side chains are shifted from the solution values of $\sim 4$ to 5.9 (Shutova et al. 1997). The buffering capacity of S-MSP also depends on the direction of titration, with the proton exchangeability being lower in the acidic to basic direction. Structural rearrangements of S-MSP measured by binding of the fluorescence probe ANS indicate that the hydrophobic core of S-MSP becomes more accessible as the pH is decreased from 7.0 to 3.8. This suggests that when S-MSP is protonated the rigidity of the tertiary structure is loosened allowing the probe to access the core of the S-MSP protein, but the secondary structure is still stable enough for there to be an intact hydrophobic core. Far-UV CD measurements of S-MSP indicate the secondary structure remains unchanged down to pH 3.5. Near UV CD at pH 3.8 indicates a structural change of S-MSP which does not correspond to complete unfolding (Shutova et al. 1997). NMR studies of T-MSP show it remains stable and folded over a pH range of 3.6 to 8.0 (Nowaczyk et al. 2004). Changes in the structure of MSP with decreasing pH have important implications for its biological function as the pH of the lumen decreases during the light reactions.

*The molten globular or natively unfolded structure of MSP.* Some unusual properties of MSP have lead to it being described as either a molten globule (Shutova et al. 2000), or as natively unfolded (Lydakis-Simantiris et al. 1999). Molten globular proteins are characterised by having a significant amount of secondary structure and significant compactness resulting from a hydrophobic core, but with decreased tertiary structure and with differences in packing compared to typical globular proteins (Freire 1995b). Natively unfolded proteins are characterised by a high stability to thermal denaturation, a high percentage of charged amino acids, lots of random coil structure a large Stokes radius, and no hydrophobic core (Weinreb et al. 1996). Both molten globular and natively unfolded proteins differ from the fully denatured protein which lacks both a
hydrophobic core, and secondary structure, with the whole length of the polypeptide chain in a random conformation, solvent exposed. A more detailed discussion of molten globular proteins is given in Chapter 7.

Non-PSII bound S-MSP has unusual mobility properties on an SDS-PAGE gel. It has an extended solution structure as measured by dynamic light scattering, analytical ultracentrifugation and small angle X-ray scattering (Svensson et al. 2004; Zubrzycki et al. 1998). The near UV and far UV spectra of isolated S-MSP also resemble that of classic molten globule proteins (Shutova et al. 2000) and its ability in acidic conditions to bind ANS and increase its fluorescence is diagnostic of the molten globular state (Shutova et al. 2005). In contrast to this, T-MSP appears to be a more typically folded protein. NMR studies of isolated T-MSP show it has a well-folded core, and is probably not a molten globule or natively unfolded (Nowaczyk et al. 2004). PSII-bound T-MSP has a well folded hydrophobic core, and sufficient structure to allow the assignment of all its side chains (De Las Rivas and Barber 2004).

The results presented in chapter 4 show that S-MSP has more molten globular character than T-MSP, which is consistent with the thermophilic ecology of T. elongatus. The implications of the molten globular structure of MSP are discussed in chapter 7.

Thermostability of MSP. Both S-MSP and T-MSP are thermostable, and have unfolding transitions which exhibit low cooperativity which is consistent with their assignment as molten globular proteins (Heredia and De Las Rivas 2003; Loll et al. 2005a; Lydakis-Simantiris et al. 1999; Sonoyama et al. 1996). Both can refold after heating to 90 °C and can restore activity to PSII upon cooling to 25 °C (Lydakis-Simantiris et al. 1999; Sonoyama et al. 1996). The T_m of S-MSP estimated by FTIR is ranges from 56-65 °C (Heredia and De Las Rivas 2003; Lydakis-Simantiris et al. 1999), although intrinsic fluorescence suggests even higher values (Shutova et al. 2005). Differential scanning calorimetric studies of S-MSP measured two unfolding transitions, one at 63 °C, the other between 80 and 100 °C (Kruk et al. 2003). The T_m of T-MSP measured by FTIR and CD is 75-76 °C (Loll et al. 2005a; Sonoyama et al. 1996). The thermal unfolding of S-MSP and T-MSP was measured by differential scanning calorimetry in Chapter 4, and reveal that T-MSP has a typical unfolding transition, but S-MSP does not. This is discussed in terms of structural differences in these proteins, and the temperature optima for oxygen evolution for spinach and T. elongatus.
1.2.4 Interaction of MSP with the PSII intrinsic proteins

*In vivo*, interaction of the extrinsic proteins with the PSII intrinsic proteins is dynamic, and the thylakoid lumen contains a pool of unassembled extrinsic subunits which can bind to PSII (Ettinger and Theg 1991). In eukaryotes all the extrinsic proteins, with the exception of red algal *psbV*, are nuclear encoded and are synthesised in the cytoplasm. The MSP is synthesised as a precursor polypeptide, which has two external domains in its N-terminal sequence- one which targets the protein to the chloroplast in plants, the other which targets it to the lumen and is called the thylakoid transfer domain. MSP transport requires a proton gradient, ATP, and the stromal factor SecA. Translocation becomes saturated, indicating enzyme involvement. Generally the proteins translocated by the SecA pathway have a single lysine adjoining the hydrophobic reigon, rather than the arginine doublet found in the 16 and 32kDa proteins. (Seidler 1996a)

*In vitro* the extrinsic proteins can also be reversibly dissociated from PSII. In PSII-enriched membrane fragments from spinach, the small extrinsic proteins can be specifically dissociated from the complex by washing with NaCl (Ghanotakis et al. 1984a; Ghanotakis et al. 1984c). The MSP can then be dissociated in a way that leaves the OEC intact by washing the membrane fragments with 2.6 M urea and 200 mM CaCl$_2$ (Miyao and Murata 1984b, 1984c; Ono and Inoue 1984). If additional Ca$^{2+}$ is present in the assay buffer, this MSP-depleted PSII complex can still evolve O$_2$ at ~20% of the rate of the untreated PSII (Bricker 1992). The MSP can functionally rebind to the extrinsic depleted PSII, restoring an average of 40% of the original O$_2$ evolving activity (Miyao and Murata 1984c; Ono and Inoue 1984). This strategy has been used extensively to examine the effects of specific amino acid changes in recombinantly expressed MSP on the function of PSII (Figure 1-7). The MSP from one species can also rebind to MSP-depleted PSII from another species and restore varying amounts of activity (Enami et al. 2000; Ono and Inoue 1984).
Figure 1-7. Schematic of the depletion/reconstitution of MSP with spinach PSII. This strategy is used to test the effects of mutations in MSP on the O₂ evolving activity of PSII. 1) Small extrinsic proteins are dissociated by washing with 1M NaCl. 2) MSP is dissociated by washing with urea and NaCl, or with 2M CaCl₂. 3) Recombinant mutated MSP, or MSP from a different species is rebound to the PSII intrinsic proteins, restoring some level of O₂ activity.

Mutagenesis of proposed binding sequences involved in interaction with the intrinsic proteins. A huge number of biochemical experiments have been conducted involving site-directed mutagenesis and chemical modification of isolated MSP from several species (Bricker and Frankel 1998; Eaton-Rye 2005; Han et al. 1994; Miura et al. 1997; Popelkova et al. 2003b; Seidler 1996a; Zhang et al. 2005). These studies have identified regions and individual residues of the MSP involved in binding to the intrinsic proteins of PSII. Much of the biochemical evidence is in agreement with analysis of the MSP binding site from the *T. elongatus* structure which is reviewed in (De Las Rivas and Barber 2004). The results of mutagenesis experiments on MSP are summarised in Table 1-3. The conservation between species of sequences that are involved in MSP binding is shown in the alignment in Figure 1-8.

**Table 1-3.** Summary of site directed mutagenesis of MSP. These experiments measured the effect of the mutation on the ability of the MSP to restore O₂ evolving capability to extrinsic-depleted PSII.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Nature of change</th>
<th>Effect</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>None</td>
<td>Restored oxygen evolving activity at ~40% of the rate of the</td>
<td>Spinach</td>
<td>(Miyao and Murata 1984c; Ono and Inoue 1984)</td>
</tr>
<tr>
<td>protein</td>
<td></td>
<td>untreated PSII</td>
<td></td>
<td>(Burnap et al. 1992; Burnap and Sherman 1991; Philbrick et al. 1991)</td>
</tr>
<tr>
<td>ΔPsbO*</td>
<td>No MSP</td>
<td>Decreased O₂ evolution.</td>
<td><em>Synechocystis</em></td>
<td>(Popelkova et al. 2002a)</td>
</tr>
<tr>
<td>Pre-MSP</td>
<td>Has the 84 amino</td>
<td>Decreased binding affinity, and non-specific binding, but</td>
<td>Spinach</td>
<td></td>
</tr>
<tr>
<td></td>
<td>acid precursor</td>
<td>still restored O₂</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>sequence.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 1. Introduction

<table>
<thead>
<tr>
<th>ΔG3M</th>
<th>Missing the first three amino acids</th>
<th>Non-specific binding but still evolves O₂.</th>
<th>Spinach (Popelkova et al. 2002a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔR5M</td>
<td>Missing the first 5 amino acids</td>
<td>Non-specific binding but still evolves O₂.</td>
<td>Spinach (Popelkova et al. 2003a)</td>
</tr>
<tr>
<td>ΔL6M</td>
<td>Missing the first 6 amino acids</td>
<td>Non-specific binding but still evolves O₂.</td>
<td>Spinach (Popelkova et al. 2003a)</td>
</tr>
<tr>
<td>ΔT7M</td>
<td>Missing the first 7 amino acids</td>
<td>Lower affinity and only ever bound one copy.</td>
<td>Spinach (Popelkova et al. 2003a)</td>
</tr>
<tr>
<td>D9K*</td>
<td>Negatively charged to positively charged.</td>
<td>No effect</td>
<td>Synechocystis (Bumap et al. 1994)</td>
</tr>
<tr>
<td>ΔK14M</td>
<td>Missing the first 14 amino acids</td>
<td>Decreased O₂ evolution.</td>
<td>Spinach (Popelkova et al. 2002b)</td>
</tr>
<tr>
<td>ΔT15M</td>
<td>Missing the first 15 amino acids</td>
<td>Non-specific binding and low O₂ evolution.</td>
<td>Spinach (Popelkova et al. 2003a)</td>
</tr>
<tr>
<td>ΔE18M</td>
<td>Missing the first 18 amino acids</td>
<td>Decreased binding affinity, and low O₂ evolution.</td>
<td>Spinach (Popelkova et al. 2002b)</td>
</tr>
<tr>
<td>C20S*</td>
<td>Removed disulfide bridge.</td>
<td>Normal levels of mRNA expression, but no protein expression.</td>
<td>Synechocystis sp.PCC6803 (Bumap et al. 1994)</td>
</tr>
<tr>
<td>C28AC</td>
<td>Removed disulfide bridge.</td>
<td>Altered solubility. Is able to restore activity.</td>
<td>Spinach (Betts et al. 1996a)</td>
</tr>
<tr>
<td>51A,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C28A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C51A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K59Q</td>
<td>Positive charge to a negative charge.</td>
<td>No change.</td>
<td>T. elongatus (Motoki et al. 2002)</td>
</tr>
<tr>
<td>R73Q</td>
<td>Positive charge to a negative charge.</td>
<td>No change.</td>
<td>T. elongatus (Motoki et al. 2002)</td>
</tr>
<tr>
<td>D109K</td>
<td>Negative charge to positive charge.</td>
<td>Very low expression.</td>
<td>Spinach (Seidler 1996b)</td>
</tr>
<tr>
<td>K123Q</td>
<td>Positive charge to a negative charge.</td>
<td>No change</td>
<td>T. elongatus (Motoki et al. 2002)</td>
</tr>
<tr>
<td>V148T</td>
<td>Non-polar to polar.</td>
<td>No change</td>
<td>T. elongatus (Motoki et al. 2002)</td>
</tr>
<tr>
<td>P149A</td>
<td>Special backbone to normal.</td>
<td>No change</td>
<td>T. elongatus (Motoki et al. 2002)</td>
</tr>
<tr>
<td>Y151F</td>
<td>Aromatic to another aromatic.</td>
<td>No change</td>
<td>T. elongatus (Motoki et al. 2002)</td>
</tr>
<tr>
<td>R152K</td>
<td>Decreased the length of the positively charged side chain.</td>
<td>Decreased O₂ evolution.</td>
<td>T. elongatus (Motoki et al. 2002)</td>
</tr>
<tr>
<td>Sequences T153 – N155</td>
<td>Changed single residues, or all three in this sequence.</td>
<td>No change</td>
<td>T. elongatus (Motoki et al. 2002)</td>
</tr>
<tr>
<td>F156L</td>
<td>Non-polar to positively</td>
<td>Loss of reactivation</td>
<td>T. elongatus (Motoki et al. 2002)</td>
</tr>
</tbody>
</table>
### Chapter 1. Introduction

<table>
<thead>
<tr>
<th>Charged/Non-charged Changes</th>
<th>Description</th>
<th>Organism</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>F156Y</strong> (charged)</td>
<td>Changed one non-polar aromatic residue to another one.</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>156+M+157</strong> (long non-polar residue)</td>
<td>Insertion of a short-chain non-polar residue between aromatic and long chain non-polar residues</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>156+G+157</strong> (long non-polar residue)</td>
<td>Insertion of a short-chain non-polar residue between aromatic and long chain non-polar residues</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>156+A+157</strong> (long non-polar residue)</td>
<td>Insertion of a short-chain non-polar residue between aromatic and long chain non-polar residues</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>156+V+157</strong> (long non-polar residue)</td>
<td>Insertion of a short-chain non-polar residue between aromatic and long chain non-polar residues</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>Deletion L157</strong></td>
<td>Removed a long-chain non-polar residue.</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>L157M</strong> (long non-polar residue)</td>
<td>Non-polar residue to another of similar chain length.</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>R162Q</strong> (charged)</td>
<td>Positively charged changed to uncharged.</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>R162K</strong> (long non-polar residue)</td>
<td>Decreased the length of the positively charged side chain.</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>G163Q</strong> (charged)</td>
<td>Changed a non-polar residue to a polar long-chain one.</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>Sequenc e L164-S166</strong></td>
<td>Changed single residues, or all three in this sequence.</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>G167Q</strong> (charged)</td>
<td>Non-polar to polar.</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>Y168F</strong> (charged)</td>
<td>Aromatic to another aromatic</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>V186S</strong> (charged)</td>
<td>Non-polar to neutral polar</td>
<td>Arabidopsis</td>
<td>(Murakami et al. 2002)</td>
</tr>
<tr>
<td><strong>K188Q</strong> (charged)</td>
<td>Positive to negative.</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td><strong>E212N</strong> (charged)</td>
<td>Polar to polar</td>
<td>Spinach</td>
<td>#20 (Seidler 1996b)</td>
</tr>
</tbody>
</table>

**Note:** Motoki et al. 2002: This research indicates a significant change in reactivation capability, with strong non-specific binding. The loss of reactivation capability, along with decreased O2 evolution, is observed in T. elongatus. Significant changes in protein expression are noted, with decreased O2 evolution in Arabidopsis and Spinach. Very low expression in Spinach is associated with a functional O2 evolution.
<table>
<thead>
<tr>
<th>Mutant</th>
<th>Description</th>
<th>Effect</th>
<th>Organism</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>K234Q</td>
<td>Positive to negative.</td>
<td>No effect on binding.</td>
<td>T. elongatus</td>
<td>(Motoki et al. 2002)</td>
</tr>
<tr>
<td>V235A</td>
<td>Removed two methyl groups from a non-polar amino acid.</td>
<td>Decreased re-binding at low temperatures.</td>
<td>Spinach</td>
<td>(Betts et al. 1998)</td>
</tr>
<tr>
<td>Q244@</td>
<td>4 amino acid C-terminal truncation</td>
<td>Decreased binding, no reactivation.</td>
<td>Spinach</td>
<td>(Betts et al. 1998)</td>
</tr>
<tr>
<td>Q245@</td>
<td>2 amino acid C-terminal truncation</td>
<td>No binding.</td>
<td>Arabadopsis</td>
<td>(Betts et al. 1998)</td>
</tr>
<tr>
<td>L245E</td>
<td>Non-polar to negative.</td>
<td>Decreased O₂ evolution.</td>
<td>Spinach</td>
<td>(Betts et al. 1998)</td>
</tr>
<tr>
<td>L246I</td>
<td>Isomer of this non-polar amino acid</td>
<td>Decreased O₂ evolution.</td>
<td>Arabadopsis</td>
<td>(Murakami et al. 2002)</td>
</tr>
<tr>
<td>E246K</td>
<td>Negative charged to positive.</td>
<td>No change.</td>
<td>Spinach</td>
<td>(Betts et al. 1998)</td>
</tr>
<tr>
<td>C-termina</td>
<td>Added 6 His residues to the C-terminus.</td>
<td>Weakened binding.</td>
<td>Spinach</td>
<td>(Betts et al. 1998)</td>
</tr>
<tr>
<td>1 his-tag</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Mutants were made in ΔpsbO strains, and not reconstituted systems.
Chapter 1. Introduction

Figure 1-8. Sequence alignments of MSP from *T. elongatus* (Therm), Spinach (Spin) and *G. violaceus* (Glo), showing residues which are important for interaction with PSII. The top sequence shows residues identified by mutagenesis from Table 1-3. **Yellow** indicates a residue where charge is important for binding or function, **pink** indicates a residue which causes structural perturbation when mutated, and **grey** indicates residues which are highly conserved, but did not show any effects from mutation. **Red** indicates the residue is not conserved in that species. Bottom sequence shows residues indicated to be involved in binding based on the crystal structures as analysed in (De Las Rivas and Barber 2004). **Green** indicates ‘regions’ as defined in (De Las Rivas and Barber 2004), **blue** indicates residues which form charge pairs. **Yellow** and **red** highlighted on S-MSP or G-MSP indicate lack of conservation of a residue which is involved in binding in T-MSP.

Both biochemical and structural analysis show that MSP binding to PSII occurs though electrostatic interactions, salt bridges and hydrogen bonds, and that non-interacting regions of the proteins are essential for providing the correct conformation for binding. The stoichiometry of binding is a matter of debate. There is some
biochemical evidence for binding of two copies per reaction centre in plants, and only
one in cyanobacteria (Popelkova et al. 2002a; Popelkova et al. 2002b, 2003a). However
high resolution structures of the cyanobacterial PSII show only one copy is bound
(Ferreira et al. 2004; Loll et al. 2005b), and low resolution structures of higher plants
also indicate only one copy (Nield et al. 2002; Nield et al. 2000b). Models built of the
spinach PSII from the low resolution structure, and predictions from this cyanobacterial
structure show that placement of MSP is the same in both species, and both have only
one MSP (Nield and Barber 2006). The binding affinity of MSP for PSII was estimated
to be between $8.0 \times 10^7$ to $6.0 \times 10^8$ M$^{-1}$ (Leuschner and Bricker 1996; Miyao and
Murata 1989). The binding of S-MSP to the intrinsic proteins of spinach PSII was
measured by isothermal titration calorimetry and is presented in Chapter 5. This method
gives a reliable measure of the stoichiometry and the binding affinity of this interaction.

The MSP binding sites on the intrinsic proteins of PSII are the loop regions of
CP43 and CP47, and D1 and D2. The T. elongatus structure shows that the MSP from
one monomer also makes contact with the CP47 protein of the other monomer,
stabilising the dimer (De Las Rivas and Barber 2004). It also shows that PsbU is bound
to MSP through the extended head domain but PsbV is not in contact with MSP (De Las
Rivas and Barber 2004). The presence of PsbV is necessary for PsbU to bind to PSII,
but PsbU can still bind in the absence of MSP (Eaton-Rye 2005). In plants, PsbP is
bound to MSP, probably through conserved positive regions which dock with negative
areas on the cylinder domain of MSP in higher plants (De Las Rivas et al. 2007; De Las
Rivas and Roman 2005).

Structural changes upon interacting with PSII. Isotope editing FTIR studies have been
used to estimate the extent of structural rearrangement occurring in the MSP protein
upon binding to the PSII intrinsic proteins. This technique involved labelling MSP with
$^{13}C$ by recombinant expression in E. coli grown on $^{13}C$ supplemented minimal media.
This allowed distinction of the labelled MSP from the unlabeled background of the PSII
protein. The spectra of PSII which had been reconstituted with unlabeled MSP were
subtracted from the sample reconstituted with $^{13}C$ labelled protein. The resulting double
difference spectra included only structural contributions from the labelled MSP, and
eliminated those from the rest of the PSII complex. A direct comparison of these with
spectra of solution MSP allowed evaluation of structural differences upon binding. This
suggests that 30 – 40% of the MSP backbone undergoes structural rearrangement upon
binding, with a conversion of unstructured regions to β-sheet (Hutchison et al. 1998a, 1998b). This structural change persists for some time in S-MSP after extraction from PSII, but is gradually lost, implying that PSII provides stability to this secondary structure (Svensson et al. 2004). Isothermal titration experiments presented in Chapter 5 indicate that a greater extent of structural rearrangement occurs when S-MSP binds to PSII than when T-MSP binds to PSII.

**GTP – mediated release of MSP from PSII.**

During the normal function of PSII in the light reactions, the D1 core intrinsic protein becomes damaged by the oxidants produced in the OEC. This damage is accelerated under high light conditions. The PSII complex is repaired by de novo synthesis and replacement of the D1 subunit. During this repair process the extrinsic proteins are released from PSII and D1 is degraded and replaced (Chow and Aro 2005; Henmi et al. 2003a; Henmi et al. 2004; Henmi and Yamamoto 2002). In the absence of MSP, D1 is not degraded, and forms aggregates with CP43 and CP47 (Yamamoto 2001; Yamamoto et al. 1998). There is strong evidence that in spinach, release of MSP from PSII is regulated by GTP binding. GTP is synthesised in the thylakoid lumen, and is required for degradation of D1 under photoinhibitory conditions (Spetea et al. 1999). GTP binds to proteins in the lumen including MSP (Spetea et al. 2004), and recent work has shown that the MSP from higher plants has GTPase activity (Lundin et al. 2007). This is imparted by portions of S-MSP structure which are not present in MSP from cyanobacteria or green algae (Lundin et al. 2007). The GTPase activity of S-MSP is greatest in high light and when associated with PSII, and is thought to induce conformational changes which cause S-MSP to release PSII so that photodamaged D1 can be replaced (Lundin et al. 2007).

**Cross reconstitution of MSPs from different species.** The conservation and function of the MSP in different species was studied by cross-reconstitution experiments where the MSP of one species was bound to MSP-depleted PSII from another species. The results of these experiments are presented in Table 1-4, and summarised in Table 1-5. They show that MSPs can bind to PSII from any other species and in most cases restore high levels of O₂ production activity. This suggests the binding sites on both the PSII and
MSP proteins are well conserved, and that MSP in all species carries out the same function in the same manner.

In all cases the native protein restored better activity than the cross reconstituted one, although for *A. thaliana* the difference is negligible. The lowest rate was for reconstitution of *C. reinhardtii* PSII with S-MSP; however the control rate of *C. reinhardtii* PSII reconstitution with its native MSP was equally low, suggesting this is specific to this organism’s PSII rather than an inability of S-MSP to function in the cross-reconstituted complex.

**Table 1-4.** Cross-reconstitution of MSP-depleted PSII. The MSP subunit of one species was rebound to the PSII of a different species, and the O₂ evolving rate restored was measured. All Values are the percentage of the untreated PSII O₂ evolution rate at 25 °C in the presence of 5mM CaCl₂.

<table>
<thead>
<tr>
<th>PSII species</th>
<th>MSP species</th>
<th>PSII/ MSP</th>
<th>O₂ (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinach</td>
<td>Spinach</td>
<td>Control</td>
<td>53</td>
<td>(Enami et al. 2000)</td>
</tr>
<tr>
<td>Spinach</td>
<td><em>Arabidopsis thaliana</em></td>
<td>HP/HP</td>
<td>52</td>
<td>(Bettset al. 1994)</td>
</tr>
<tr>
<td>Spinach</td>
<td><em>Chlamydomonas reinhardtii</em></td>
<td>HP/GA</td>
<td>54</td>
<td>(Suzuki et al. 2005)</td>
</tr>
<tr>
<td>Spinach</td>
<td><em>Cyanidium. caldarium</em></td>
<td>HP/TRA</td>
<td>51</td>
<td>(Enami et al. 2000)</td>
</tr>
<tr>
<td>Spinach</td>
<td><em>Synechococcus vulcanus</em></td>
<td>HP/TCB</td>
<td>39</td>
<td>(Enami et al. 2000)</td>
</tr>
<tr>
<td>Spinach</td>
<td><em>Phormidium laminosum</em></td>
<td>HP/TCB</td>
<td>30*</td>
<td>(Pueyo et al. 2002)</td>
</tr>
<tr>
<td><em>Chlamydomonas reinhardtii</em></td>
<td><em>Chlamydomonas reinhardtii</em></td>
<td>Control</td>
<td>10</td>
<td>(Suzuki et al. 2005)</td>
</tr>
<tr>
<td><em>Chlamydomonas reinhardtii</em></td>
<td>Spinach</td>
<td>GA/ HP</td>
<td>6</td>
<td>(Suzuki et al. 2005)</td>
</tr>
<tr>
<td><em>Cyanidium. caldarium</em></td>
<td><em>Cyanidium. caldarium</em></td>
<td>Control</td>
<td>41</td>
<td>(Enami et al. 2000)</td>
</tr>
<tr>
<td><em>Cyanidium. caldarium</em></td>
<td>Spinach</td>
<td>TRA/ HP</td>
<td>39</td>
<td>(Enami et al. 2000)</td>
</tr>
<tr>
<td><em>Cyanidium. caldarium</em></td>
<td><em>Synechococcus vulcanus</em></td>
<td>TRA/ TCB</td>
<td>37</td>
<td>(Enami et al. 2000)</td>
</tr>
<tr>
<td><em>Synechococcus vulcanus</em></td>
<td><em>Synechococcus vulcanus</em></td>
<td>Control</td>
<td>69-75</td>
<td>(Enami et al. 2000; Koike and Inoue 1985)</td>
</tr>
<tr>
<td><em>Synechococcus vulcanus</em></td>
<td>Spinach</td>
<td>TCB/ HP</td>
<td>40</td>
<td>(Enami et al. 2000)</td>
</tr>
<tr>
<td><em>Synechococcus vulcanus</em></td>
<td>Spinach</td>
<td>TCB/ HP</td>
<td>60</td>
<td>(Koike and Inoue 1985)</td>
</tr>
<tr>
<td><em>Synechococcus vulcanus</em></td>
<td><em>Cyanidium. caldarium</em></td>
<td>TCB/ TRA</td>
<td>23</td>
<td>(Enami et al. 2000)</td>
</tr>
</tbody>
</table>

1 Control = reconstituted with MSP from the same species  
2 HP = higher plant  
3 TRA = Thermophilic red algae  
4 GA = green algae  
5 TCB = thermophilic cyanobacterium.  
* The spinach PSII reconstituted with S-MSP was also 30% in this study, so this restores the same activity as the control.
Table 1-5. Summary of cross reconstitution experiments from Table 1-4. Values are the percentage of the untreated PSII O₂ evolution rate in the presence of 5mM CaCl₂. Abbreviations: HP = higher plant, RA = red algae, GA = green algae, TCB = thermophilic cyanobacterium

<table>
<thead>
<tr>
<th>Species of PSII</th>
<th>Species of MSP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HP¹</td>
</tr>
<tr>
<td>HP</td>
<td>52-53</td>
</tr>
<tr>
<td>GA</td>
<td>6</td>
</tr>
<tr>
<td>TRA</td>
<td>39</td>
</tr>
<tr>
<td>TCB</td>
<td>40</td>
</tr>
</tbody>
</table>

¹ HP = higher plant
² GA = green algae
³ TRA = Thermophilic red algae
⁴ TCB = thermophilic cyanobacterium.

The work presented here extends these cross-reconstitution experiments. Chapter 3 shows that a fusion protein between T-MSP and E. coli thioredoxin restore higher rates of O₂ evolving activity to spinach PSII than S-MSP does. Chapter 6 shows that G-MSP, which is poorly conserved, compared to the MSP of other species cannot bind to spinach PSII.

1.2.5 Evolution of MSP

The sequential order in which the extrinsic proteins evolved is thought to begin with MSP, as it is common to all oxygenic photosystems. This was followed by PsbU and PsbV of the cyanobacterial PSII as both of these are found in G. violaceus. PsbP probably came next as homologues of this gene are found in most cyanobacteria, and last, PsbQ as it is not found in G. violaceus or Prochlorococcus species. PsbU and PsbV were probably lost from the plant and green algal genome when they were functionally replaced by PsbP and PsbQ (De Las Rivas et al. 2007; Rivas et al. 2004). The psbO gene is a good candidate for studying the evolution of oxygenic photosynthesis because it is a functionally conserved part of all oxygenic photosystems, it is nucleus encoded in eukaryotes and so will not be affected by the difference in rates of evolution between chloroplast and nuclear genes, and it has a conserved function (De Las Rivas and Roman 2005; Rivas et al. 2004).

Sequence conservation of the MSP between different species is high, ~ 80% between higher plants, ~60% between higher plants and algae, and about ~40% between cyanobacteria and plants. The exception is G. violaceus, which is only 29% conserved
with plants (De Las Rivas and Roman 2005; Rivas et al. 2004). Homology between PsbU and PsbV is about 40% between all cyanobacteria, including *G. violaceus* (Rivas et al. 2004).

The evolutionary origins of the MSP are not known and no proteins that share a common ancestor have been identified. A homology search could not find the MSP or any homologous protein in non-oxygenic photosynthetic bacteria, including those with Type II photosystems (Rivas et al. 2004). By contrast, the PsbP and PsbQ proteins have homology with the plant proteins TL26 and TL40, the function of which are not known, and PsbV has sequence and structural homology with cytochrome c-6 (Rivas et al. 2004). The results presented in Chapter 6 show that, unlike other species, G-MSP cannot be cross-reconstituted with spinach PSII. This is interpreted as a consequence of the primitive nature of G-MSP, and suggests it may represent an intermediate in the evolution of other MSPs. The possible origins of MSP are discussed in Chapter 7.

### 1.2.6 Role of a possible water channel in MSP

In addition to its role in maintaining the integrity of the WOC, the MSP may also be involved in forming a channel that regulates the access of small molecules to the metal-containing catalytic site. Such a channel could provide a pathway for the release of H⁺ and O₂, and allow the controlled entry of substrate water in order to prevent unwanted side reactions by exposure of the OEC to the bulk solvent (De Las Rivas and Barber 2004; Hillier et al. 2001; Ishikita et al. 2006a; Wydrzynski et al. 1996). The role of MSP as a water channel might serve to restrict the access of substrate water to the [Mn₄Ca] cluster, an idea first proposed by Wydrzynski *et al* over ten years ago (Wydrzynski et al. 1996). Support for a water channel involving the MSP comes from a number of sources. The accessibility of water analogues which inhibit O₂ evolution is increased by depletion of the small extrinsic proteins (Mei and Yocum 1993; Radmer et al. 1986; Radmer and Ollinger 1982, 1983; Tamura and Cheniae 1985). The rate of water exchange in spinach PSII measured by mass spectrometry was actually slowed by removal of the small extrinsic proteins, and was slowed even further by depletion of the MSP (Hillier et al. 2001). The reason for this is likely the disruption of the hydrophilic pathway or water chain delivering the substrate waters to the active site. The presence of a controlled substrate delivery and product removal system is observed in many enzymes, and is expected to be vital in PSII. During water splitting
reactive oxygen intermediates may be produced, which could undergo side reactions with non-substrate waters to produce peroxide which damages proteins (Wydrzynski et al. 1996).

The structure of bound T-MSP shows that the β-barrel cannot provide such a channel as it is filled with the bulky hydrophobic side-chains that define the hydrophobic core (De Las Rivas and Barber 2004). However a channel involving residues from MSP, as well as regions of D1 and D2, and the MSP-bound Ca\(^{2+}\) ion was proposed instead (Figure 1-9, dark blue). The hydrophilic channel runs from the [Mn\(_4\)Ca] cluster to the lumen, along the binding interface of MSP and the intrinsic proteins (De Las Rivas and Barber 2004; Ferreira et al. 2004; Murray and Barber 2006) (Figure 1-10 A). The putative channel residues show an increase in calculated pKa from the OEC to the lumenal exit, which is predicted to increase in the S\(_4\) state. This could act to remove protons released from water splitting reaction and expel them into the lumen (Ishikita et al. 2006b). Further analysis of the 3.5 Å crystal structure (Ferreira et al. 2004) was carried out by Murray and Barber (2007) using a programme which predicts cavities in proteins, and indicated two additional channels leading from the [Mn\(_4\)Ca] cluster (Figure 1-10 B). Channel (i) is thought to act as an O\(_2\) exit channel, and channel (ii) as a water entry/proton exit channel. Channel (iii) which involves MSP is suggested to act as a proton exit channel. A similar analysis of the 3.0 Å (Loll et al. 2005b) and 3.5 Å (Ferreira et al. 2004) structures was used by Ho and Styring (2008). This study also identified three channels, although these differed somewhat to those proposed by Murray and Barber. One of the Ho channels also includes residues from MSP (Figure 1-9 purple), but these do not correspond to the sequences proposed by Murray and Barber (2007) (Figure 1-9 dark blue). Shutova et al. (2007) have identified a cluster of carboxylic acid groups (CCG) which they suggest form a pool of bound substrate water at the OEC (Figure 1-9 light blue) (Shutova et al. 2007). This would provide the hydrophobic environment necessary for the functioning of the [Mn\(_4\)Ca] cluster while still supplying water for oxidation by PSII. Shutova et al. (2007) also suggest the CCG motif acts as a ‘proton antennae’ removing them from the OEC, and delocalising them over a number of acceptors.
Figure 1-9. Sequence alignment of *T. elongatus* MSP (Therm), Spinach MSP (Spin) and *G. violaceus* MSP (Glo) showing the putative water channel residues. Residues shown in green text are conserved; residues shown in red are not conserved. Residues marked with a * are fully conserved in all species according to the multiple sequence alignment of (De Las Rivas and Barber 2004). Residues marked with $ are functionally conserved in all species. Residues marked with a + are fully conserved in all species except *G. violaceus*. Residues marked with # are functionally conserved in all species except *G. violaceus*. Residues highlighted in dark blue are proposed to be involved in a water/proton channel by (Murray and Barber 2007). Residues highlighted in purple are proposed to be involved in a channel by (Ho and Styring 2008). Residues highlighted in light blue are proposed to form a cluster of carboxylic acids involved in proton transfer (Shutova et al. 2007). Alignment produced in CLUSTALW 35.
Figure 1-10. Putative substrate and product channels in PSII. These figures are adapted from references (Murray and Barber 2006) and (Murray and Barber 2007).  

A) Proton channel from (Murray and Barber 2006) with residues provided by D1 in yellow, D2 in orange, and D3 in blue.  

B) Channel system predicted for the PSII OEC using the CAVER programme (see (Murray and Barber 2007) for details). D1 shown in yellow, D2 in orange, CP43 in green, PsbV in blue and PsbU in purple. Channel iii corresponds to the proton channel shown in A, although the MSP residues are not shown in this view.

When comparing these putative channels and cavities predicted in PSII from x-ray crystal data, it is important to first consider how conserved they are. The sequence alignment in Figure 1-9 shows that the Murray channel and Shutova CCG motif share many residues in common (dark blue and light blue), and many are highly conserved across species (De Las Rivas and Barber 2004). This is significant since such an important function of the MSP would probably be common to all PSII systems. The exception of this conservation is G-MSP which only conserves one of the Asp residues.
common to both the Murray and Shutova models. The alignment presented here differs slightly from that given in (De Las Rivas and Barber 2004), as only the sequences of S-MSP, T-MSP and G-MSP were used. In the alignment of all species the final Glu of these channels is not conserved in G-MSP either. The lack of conservation of these residues in G-MSP makes this species of particular interest, and is discussed in detail in Chapter 6. In contrast to the Murray and Shutova models, only one of the residues proposed by Ho to form part of a channel is partly conserved across all species other than G. violaceus. A second consideration is how these predictions correlate with other experimental evidence. All but one of the residues of the Ho channel have been examined by site directed mutagenesis in T. elongatus which had no effect on steady-state O₂ evolving capacity (Motoki et al. 2002). Shutova’s CCG cavity was consistent with other work by that group which showed a hysteresis in pH titration of S-MSP (Shutova et al. 1997), and with the work by Hillier et al. (2001) suggesting that MSP functions to control supply of water to the active site. Finally, it is significant that in all three models the key residues are found in the loop regions of the MSP, and all three models are based on predictions from the crystal structure. Although solving the high resolution structure of T. elongatus PSII ushered in a new era of photosynthesis research, this structure depicts a single conformation in one of the five S-state intermediates, in a single organism. Because of the flexibility of the loop regions of MSP, the positions of the side chains could change during the reaction cycle, and might open alternate substrate pathways and close others. The lack of conservation in the Ho channel would not matter if the key residues are found in different loops in other organisms, and could still interact in the same manner. Additionally the plasticity of the loop regions could allow MSP to tolerate mutations to functional residues better than would be expected for most proteins. In the Murray and Ho models, where MSP is thought to contribute to a proton rather than a water channel, MSP might also influence the substrate water supply by causing the intrinsic proteins to adopt the correct conformation for the substrate channel.

The importance of MSP’s flexibility in its function is one of the overall themes of this work, and as discussed in the following chapters, the degree of flexibility probably varies between species and between the bound and un-bound conformations. The importance of flexibility in MSP’s contribution to a product/substrate channel is considered in the context of these results in Chapter 7.
1.3 Research aims

The aims of this research are first, to extend the cross reconstitution experiments described in Section 1.2.4 of Chapter 3 and include reconstitution of MSP-depleted PSII with a recombinant fusion protein of *T. elongatus* MSP with *E. coli* thioredoxin attached to its N-terminus (T-MSP-TRX). The activity reconstituted by the T-MSP-TRX fusion protein is compared to the activity reconstituted by rebinding T-MSP and S-MSP to MSP-depleted PSII. The effect of increasing concentrations of CaCl$_2$ on these reconstituted systems is measured. Second, to determine the difference in thermal stability between MSP from *T. elongatus* and MSP from spinach. The *T. elongatus* has evolved to live at 57 °C, whereas spinach normally experiences temperatures between 15 and 25 °C. The effect of a variety of solvent conditions on the thermal stability of these proteins is also compared. Third, to characterise the thermodynamics of binding of MSP to PSII using isothermal titration calorimetry, and compare differences between *T. elongatus* and spinach MSPs. Fourth, to produce and characterise MSP from *G. violaceus*, and investigate its ability to bind and restore activity to spinach PSII. *G. violaceus* is thought to preserve features of ancestral oxygenic photosynthetic organisms, and the sequence from *G. violaceus* indicates that it diverged earliest during the evolutionary radiation of cyanobacteria.
2 Methods and Theory

2.1 Preparation of Spinach MSP

Photosystem II-enriched membrane fragments with high rates of O\textsubscript{2} producing activity were isolated from spinach leaves. These membranes were used in activity assays by removing their native extrinsic proteins and in their place binding recombinant proteins from other species. This native spinach PSII is also the source of isolated S-MSP used in other experiments.

2.1.1 Isolation of PSII-enriched membrane fragments from spinach

Spinach PSII enriched membranes were prepared from market spinach based on earlier methods (Berthold et al. 1981). The stalks and mid veins of the spinach plants were removed, and the leaves were washed with distilled water. The leaves were dark adapted for 10 minutes, the homogenised in batches with blend buffer (100 mM sucrose, 50 mM KH\textsubscript{2}PO\textsubscript{4}, 50 mM NaH\textsubscript{2}PO\textsubscript{4}, 200 mM NaCl, pH 7.5) using a kitchen blender. The homogenised leaves were filtered through two layers of cheese cloth, and two layers of nylon mesh 40 μm to remove fibrous material. The broken chloroplasts were collected by centrifugation at 10 000 xg for 10 minutes. The pelleted chloroplasts were washed by resuspending in blend buffer using a glass homogeniser, and collecting again by centrifugation. The washed chloroplasts were resuspended in 400 mM sucrose, 20 mM MES pH 6.5, 15 mM NaCl (SMN buffer) with 5 mM MgCl\textsubscript{2} to a Chl concentration of 2.5 mg ml\textsuperscript{-1} and incubated at 4 °C in the dark for 90 minutes to allow membrane stacking (Ford and Evans 1983). Triton X100 was added drop wise to the stacked membranes with gentle stirring to a final concentration of 5 % per volume. The solubilised membranes were collected by centrifugation at 35 000 g for 30 minutes, and washed in SMN buffer. The starch was removed from the membranes as a pellet by centrifuging at 3 000 rpm for 5 minutes (Campbell et al. 1998). Finally the membranes were collected by centrifugation and resuspended in SMN buffer with 5 mM CaCl\textsubscript{2} for storage at -80 °C. All steps were carried out at 4 °C under safe green light. The Chl concentration of all photosynthetic samples was determined by the method of Porra et al. (1989).
2.1.2 Depletion of the extrinsic proteins from PSII and purification of spinach MSP

PSII membranes depleted of the small extrinsic proteins were prepared by resuspending PSII-enriched membranes in SMN buffer with 1 M NaCl and incubating for 30 minutes in the dark on ice (Kuwabara et al. 1985). Small extrinsic-depleted PSII membranes were collected by centrifugation, washed once and resuspended in SMN with 5 mM CaCl₂.

The native S-MSP was removed by resuspending the small extrinsic depleted PSII in 2.6 M urea, 200 mM NaCl and incubating for 30 minutes as described by (Bricker 1992), collecting the membranes by centrifugation, washing, and resuspending in SMN with 5 mM CaCl₂. All PSII samples were stored at -80 °C.

The supernatant from the urea wash containing the purified S-MSP was dialysed against 20 mM Mes, pH 6.5 overnight, then centrifuged at 35 000 g to remove any precipitated proteins. The purity of the MSP was checked on an SDS-PAGE gel shown in Figure 2-1. The concentration of S-MSP was estimated spectrophotometrically using an extinction coefficient of 16 mM⁻¹ cm⁻¹ (Kuwabara et al. 1985).

![SDS-PAGE showing purified MSP from spinach.](image)

2.1.3 Reconstitution of extrinsic depleted PSII with MSP

Urea washed PSII was reconstituted with MSP, MSP-TRX fusion protein or TRX using an earlier method (Ono and Inoue 1984). MSP in reconstitution buffer (300 mM sorbitol, 10 mM NaCl, 10 mM CaCl₂, 20 mM MES) was added to extrinsic depleted
PSII membranes at a mol to mol ratio of 5 to 1. The volume was adjusted so that the Chl was 0.25 mg/ml and the reaction was incubated on ice in the dark for 15 minutes.

2.2 Preparation of recombinant MSP from *T. elongatus* and *G. violaceus*

The MSP proteins from the cyanobacteria *T. elongatus* (T-MSP) and *G. violaceus* (G-MSP) were expressed as recombinant proteins in *E. coli* using the pET32 plasmid system (Novagen). Protein expression from the pET32 produces the target protein with *E. coli* thioredoxin (TRX) and a poly-histidine tag (6-His) fused to the N-terminus. The T-MSP-TRX fusion protein is expressed in *E. coli* strain BL21(DE3)PlysS (Novagen) as soluble protein, whereas G-MSP-TRX fusion protein is expressed as insoluble inclusion bodies. Soluble G-MSP-TRX was produced using two strategies: solubilisation and refolding of the inclusion bodies from urea, or by co-expression with chaperone proteins. The TRX and 6-His tags were removed from the MSPs by digestion with enterokinase (Novagen).

2.2.1 Cloning of the psbO gene

*Extraction of genomic DNA from *T. elongatus* and *G. violaceus.** The genomic DNA was extracted from a stationary culture of *T. elongatus* which had been grown in BG11 media at 55°C under constant light from white fluorescent lamps at an irradiance of 50 μmol m² s⁻¹, or *G. violaceus* which had been grown in BG11 media at a irradiance of 10 μmol photons m⁻² s⁻¹ and 25 °C to stationary phase. The extraction protocol was adapted from a published method for extracting genomic DNA from filamentous cyanobacteria (Wu et al. 1999). The cells were pelleted by centrifugation at 6 000 g 20 °C for 20 minutes then resuspended in 100 mM Tris, 50 mM EDTA, 100 mM NaCl pH 8.0. Sarkosyl was added to a final concentration of 0.1 % to remove the gelatinous sheath that often surrounds cyanobacterial cells, and the sample was incubated at room temperature for 30 minutes. The cells were collected by centrifugation at 8 000 g, 20 °C for 10 minutes, resuspended in saturated sodium iodide solution and incubated at 37 °C for 20 minutes. The sample was pelleted by centrifugation at 8 000 g, 20 °C for 5 minutes, and then resuspended in 50 mM Tris, 5 mM EDTA, 50 mM NaCl, pH 8.0. Lysozyme was added to a final concentration of 0.5 mg ml⁻¹, and the solution was incubated at 37 °C for 30 minutes. Proteinase K was added, and the solution incubated
at 37 °C for another hour. The cells were checked under a light microscope to confirm they had been broken, then SDS was added to a final concentration of 1 %, and the solution was incubated at 37 °C with stirring.

**Design of primers and amplification of the psbO gene.** Cloning of the section of the psbO gene encoding the mature MSP was done in two steps: (1) amplification of the gene from the genomic DNA, and (2) addition of the LIC extensions for insertion into the pET 32 EK LIC vector. (Figure 2-2). The primers shown in Table 2-1 were designed from the published psbO sequences for *T. elongatus* psbO and the *G. violaceus* (Miura et al. 1993; Nakamura et al. 2002), and were purchased from Novagen USA. For *G. violaceus*, the cleavage site of the transit peptide from the mature protein was predicted from the published gene sequence (Nakamura et al. 2003) using the SignalP 3.0 server ([http://www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/)), and was in agreement with other published mature sequences (De Las Rivas and Barber 2004). The primers for the first PCR reaction for *T. elongatus* did not include the end nucleotides of the gene. This was done to maximise the GC content and ensure greater specificity during cloning. The second PCR reaction added the LIC extension for insertion into the pET32 plasmid (underlined) as well as the missing nucleotides from the *T. elongatus* psbO gene. The PCR conditions are given in Table 2-2. The product from each reaction was electrophoresed on a 1% agarose gel and the bands corresponding to the expected nucleotide length were cut out and purified using the QIAquick Spin Gel Extraction kit (Qiagen, Australia).
Chapter 2. Methods and Theory

(1) Cloning the majority of the gene out of the genomic DNA

(2) Adding the end nucleotides of the gene, and the LIC ends

Figure 2-2. Strategy for cloning the psbO gene from T. elongatus and G. violaceus. (1) Most of the psbO gene was cloned from the genomic DNA using a set of primers designed to have high specificity to the ends of the psbO sequences. (2) The LIC ends, and a few missing terminal nucleotides of the psbO gene were added by another round of PCR, using the first PCR products as a template, and having a new set of primers with the end sequences on them.

Table 2-1. Primers for the amplification of the mature psbO gene from T. elongatus and G. violaceus. Primers for reaction (1) were used to amplify the gene from the genomic DNA. Primers for reaction (2) were used to add the LIC extensions for insertion into the pET 32 EKLIC plasmid

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>T. elongatus forward</td>
<td>5’GCAAAACAGACCTTTAACC3’</td>
</tr>
<tr>
<td>(1)</td>
<td>T. elongatus back</td>
<td>5’CTAGGCCGGTTCTACGCGCCG 3’</td>
</tr>
<tr>
<td>(1)</td>
<td>G. violaceus forward</td>
<td>5’GCTCCGATCGTGCACCC3’</td>
</tr>
<tr>
<td>(1)</td>
<td>G. violaceus back</td>
<td>5’TTAGCTCGCTCTC 3’</td>
</tr>
<tr>
<td>(2)</td>
<td>T. elongatus forward</td>
<td>5’GACGACGACAAGATCGCAGAAACAGACTTTAACC3’</td>
</tr>
<tr>
<td>(2)</td>
<td>T. elongatus back</td>
<td>5’GAGGGAAGGAGCAGGTAGGCTCTGCTCGAATGC3</td>
</tr>
<tr>
<td>(2)</td>
<td>G. violaceus forward</td>
<td>5’GACGACGACAAGATCGCTCCGATCGTGCACCC3’</td>
</tr>
<tr>
<td>(2)</td>
<td>G. violaceus back</td>
<td>5’GAGGGAAGGAGGCTCTCTCTC3’</td>
</tr>
</tbody>
</table>
Chapter 2. Methods and Theory

Table 2-2. PCR conditions for amplification of the psbO gene from *T. elongatus* and *G. violaceus*. This cycle is repeated 34 times.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
<th>Cycle</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>95 °C</td>
<td>5min</td>
<td>No</td>
<td>Denatures the long strands of genomic DNA</td>
</tr>
<tr>
<td>95 °C</td>
<td>1min</td>
<td>yes</td>
<td>Denatures the short amplified product</td>
</tr>
<tr>
<td>55 °C</td>
<td>1min</td>
<td>yes</td>
<td>The primers are able to anneal to the ends of the gene</td>
</tr>
<tr>
<td>72 °C</td>
<td>1min</td>
<td>yes</td>
<td>The DNA polymerase elongates the primers by adding nucleotides</td>
</tr>
<tr>
<td>72 °C</td>
<td>10min</td>
<td>no</td>
<td>All the remaining primers and genes are used to make lots of copies of the gene, and any partially finished elongation reactions are completed. The gene cannot be unwound during this time.</td>
</tr>
<tr>
<td>4 °C</td>
<td>Indefinitely</td>
<td>no</td>
<td>Prevents DNA being broken down</td>
</tr>
</tbody>
</table>

*Insertion of the psbO gene into pET 32 EK LIC by Ligation Independent Cloning*

The amplified *psbO* products with LIC ends were cloned into pET32 EK/LIC plasmids (Novagen, USA) using the LIC method (Aslanidis and Dejong 1990). In addition to the LIC site shown in Figure 2-3, the pET32 EK/LIC cloning region has a number restriction sites and an ampicillin resistance site for selection on antibiotic media (Figure 2-4). The annealing reaction was carried out as described in the Novagen product manual, and the resulting constructs are referred to as pET32 T-psbO for the *T. elongatus* insert, and pET32 G-psbO for the *G. violaceus* insert.
Chapter 2. Methods and Theory

Figure 2-3. Ligation Independent Cloning (LIC) strategy. (1) LIC sequences added onto the gene during the second PCR reaction. (2) Overhangs created from the LIC sequences by treatment with T4 DNA polymerase. (3) LIC sequences in the linearised pET32EK/LIC plasmid. The overhangs on the gene and the plasmid are complementary allowing the gene to anneal and create a circularised plasmid (figure adapted from Novagen).

Figure 2-4. Map of the cloning and promoter region of pET 32 EK/LIC. The regions important for this work are the T7 promoter that initiates transcription upon induction of protein expression, the TRX•Tag which generates an N-terminal fusion with the target protein, the His•Tag which encodes six sequential histidines in the linker region between the TRX and target protein, and allows purification by metal affinity chromatography, the enterokinase site which encodes the recognition sequence for the protease allowing removal of the N-terminal tags from the target protein, and the LIC site which allows cloning of the gene into the plasmid (figure adapted from Novagen).

Selection for Positive clones, plasmid amplification and transformation into the expression strain. The constructs were transformed into the E-coli strain NovaBlue (Novagen, USA) by heat shock, and the transformed cells were plated onto 2YT media.
with 50 µg ml\(^{-1}\) ampicillin for selection of positive clones, and grown overnight at 37 °C. Single colonies from the plates were grown in LB liquid media with 50 µg ml\(^{-1}\) ampicillin, and the plasmids were purified using QIAprep Miniprep system (Qiagen), and sequenced to check for point mutations (Australian Genome Research Facility, Australia).

The purified plasmids were used to transform expression strains of E-coli. For expression of soluble T-MSP, or G-MSP as inclusion bodies Bl21(DE3)PlysS (Novagen, USA) was used. For co-expression of G-MSP with chaperone proteins the Bl21(DE3) strain was used, and the details are given in section 2.2.4.

### 2.2.2 Expression and purification of soluble T. elongatus MSP

**Expression and extraction of T-MSP-TRX fusion protein.** The T-MSP-TRX fusion protein was expressed as described previously (Tohri et al. 2002), with some modifications. A liquid culture of BL21 was grown to an OD\(_{600}\) of 0.5 AU then induced by addition of 1mM IPTG. The cells were grown for another four hours with vigorous aeration, collected by centrifugation at 4225 g 25°C for 5 minutes and then frozen at -20 °C overnight. 1L of culture yielded approximately 5 g of cells. The pelleted cells were thawed and resuspended in at 125 mg ml\(^{-1}\) in 50 mM phosphate pH 8 at. To lyse the cells, lysozyme was added at 125 µg ml\(^{-1}\), DNase A was added at 50 µg ml\(^{-1}\), and the solution was incubated at 37 °C for 20 minutes with vigorous shaking. To ensure complete cell breakage the mixture was sonicated on full power for two minutes on ice. The lysed cell mixture was centrifuged at 35,000 g for 30 min to separate soluble from insoluble material. T-MSP-TRX fusion protein was soluble and so could be purified from the cell extract. *G. violaceus* MSP-TRX was insoluble, and needed to be solubilised in 7 M urea prior to purification and refolding.

The MSP-TRX fusion protein was partially purified by heating the solution to 70 °C for 20 minutes before loading onto a 5 ml Ni-sepharose high performance column (Amersham). The T-MSP-TRX fusion protein was eluted on a linear gradient from 0 to 250 mM imidazole, with the T-MSP-TRX eluting at 70 mM imidazole. Fractions containing T-MSP-TRX were pooled, and dialysed into either 20 mM MES (pH 6.5) for reconstitution with PSII, 20 mM phosphate (pH 6.5) for DSC and CD measurements or 20 mM Tris (pH 6.5) to remove imidazole before cleavage into T-MSP and TRX.
attached to the linker. Purified T-MSP-TRX is shown on the SDS-PAGE gel in Figure 2-5.

![SDS-PAGE gel of purified recombinant T. elongatus MSP-TRX fusion protein (lane A) and T-MSP (lane B)](image)

**Figure 2-5.** SDS-PAGE gel of purified recombinant *T. elongatus* MSP-TRX fusion protein (lane A) and T-MSP (lane B)

*Removal of the TRX tag.* The fusion protein was cleaved in 20 mM Tris (pH8), 50 mM NaCl at 2.5 mg ml\(^{-1}\) with 5 units of recombinant enterokinase (Novagen) per mg of T-MSP-TRX. The reaction was stopped after 2 hours at room temperature by heating for 20 minutes at 70 °C. The mixture of T-MSP and TRX was separated by applying to another Ni-sepharose column cluster and equilibrating with 20 mM Tris (pH 8). The flow through contained the T-MSP. The TRX with the linker and His tag attached was eluted in a single step with 100 mM imidazole. The T-MSP and TRX were dialysed into 20 mM MES (pH 6.5) for reconstitution with PSII and calorimetry or into 20 mM phosphate (pH 6.5) for CD measurements. Purified T-MSP is shown in the SDS-PAGE gel in Figure 2-5 lane B. The MSP concentrations were estimated spectrophotometrically at 280 nm using extinction coefficients based on the number of tryptophan, tyrosines and disulfide bonds; 21.8 mM\(^{-1}\) cm\(^{-1}\) for T-MSP-TRX and 7.8 mM\(^{-1}\) cm\(^{-1}\) for T-MSP (Gill and Vonhippel 1989).
2.2.3 Purification of *G. violaceus* MSP by refolding from urea

The G-MSP-TRX fusion protein was expressed as described for T-MSP-TRX in section 2.2.2, and lysed cells were centrifuged to collect the TRX-MSP fusion protein which was expressed as inclusion bodies (Figure 2-6A). Almost no G-TRX-MSP was expressed in the soluble fraction (Figure 2-6B), and this was not increased by expression at 15 °C or induction by lower concentrations of IPTG. The inclusion bodies were solubilised in buffer A (1mM glutathione reduced and 20 mM phosphate pH 8, 0.1 % triton X100) with 7 M urea. The urea solubilised protein was purified on a 5 ml Ni-sepharose high performance column cluster (Amersham), eluting at about 100 mM imidazole. The purified G-MSP-TRX was diluted into 10X the volume of refolding buffer (100 mM NaCl, 0.1 mM glutathione oxidized, 5% glycerol, 20 mM phosphate buffer pH 8) and incubated at 4 °C for 12 hours. The diluted protein was then concentrated by binding to Ni-sepharose and eluting with buffer A, 0.7 M urea, 500 mM imidazole. The G-MSP-TRX was diluted and concentrated again, and dialysed into 50 mM Tris pH 8 (Figure 2-6C). The TRX-linker was removed from the refolded G-MSP as described for T-MSP in section 2.2.2, and the G-MSP was dialysed into 20mM MES pH 6.5 (Figure 2-6C). The G-MSP concentration was estimated spectrophotometrically at 280 nm using an extinction coefficient of 17.45 mM⁻¹ cm⁻¹ based on the number of tryptophan, tyrosines and disulfide bonds (Gill and Vonhippel 1989).

![Figure 2-6. Over expression and purification of *G. violaceus* MSP from inclusion bodies. A) Crude cell extract. B) Soluble fraction of cell extract. C) Purified and refolded G-MSP-TRX fusion protein. D) G-MSP with TRX tag removed.](image-url)
2.2.4 Purification of G. violaceus MSP by co-expression with chaperonins

G-MSP-TRX was co-expressed in *E. coli*, with chaperone proteins to make it soluble. The pGro7 plasmid (Takara Bio inc. Japan) which encodes the chaperone proteins GroEL and GroES was used to transform the *E. coli* expression strain BL21(DE3) according to the manufacturer’s directions (Nishihara et al. 1998). Competent cells were prepared from positive transformants and were subsequently transformed with the pET32 G-psbO construct. Transformants retaining both the pGro7 and pET 32 EK/LIC plasmids were selected for using media containing 50 µg ml⁻¹ ampicillin, and 20 µg ml⁻¹ chloromphenicol. Co-expression of G-MSP with GroEL and GroES was induced by adding 0.5 mg ml⁻¹ L-arabinose and 0.1 mg ml⁻¹ IPTG to a log phase culture at 37 °C, and growing for four hours.

Both G-MSP-TRX and GroEL were over expressed under these conditions (Figure 2-7 A). Although most of the G-MSP was expressed in the insoluble fraction (Figure 2-7 B) a significant amount was also found in the soluble fraction along with a large amount of GroEL (Figure 2-7 C) A large portion of the GroEL was also found in the insoluble fraction (Figure 2-7 B). The soluble fraction of G-MSP-TRX was purified by Ni-sepharose affinity chromatography, and the TRX-linker was removed from the refolded G-MSP as described for T-MSP in section 2.2.2, and the G-MSP was dialysed into 20 mM MES pH 6.5 (Figure 2-6 C).

![Figure 2-7](image)

**Figure 2-7** Co-expression of *G. violaceus* MSP-TRX fusion protein with GroEL, and purification of soluble fusion protein. A) Crude cell extract. B) Insoluble fraction of cell extract. C) Soluble fraction of the cell extract D) Purified soluble G-MSP-TRX.

2.3 Analytical techniques

2.3.1 Circular dichroism

Circular dichroism spectropolarimetry (CD) is used to monitor changes in the secondary and tertiary structure of a protein. CD measures the rotation of plane polarized light, or ellipticity (Pain 2004). This rotation occurs when plane polarized light from the
Networks are passed through a sample solution that has chromophores which are in an optically a-symmetric environment. The chromophores absorb the left- and right-handed polarized light differently, so when the two components are recombined the resulting vector is rotated compared to the original plane. In proteins, optically active chromophores include Trp, Tyr and Phe in the near-UV region (240 – 320 nm) and the peptide bond in the far- UV region. Disulfide bonds and His also contribute generally to the CD in both regions. Far-UV CD is sensitive to secondary structure as the regular α-helix and β-sheet structures are conformationally asymmetric. The far-UV CD spectra of proteins can be deconvoluted to give an estimate of their secondary structural composition by comparison with a reference set of spectra from proteins with known structure.

CD Spectra were recorded on a Jobin-Yvon Type III Plus spectrometer from 182 to 240 nm at room temperature using a quartz cuvette with a path length of 0.05 cm., The protein samples were extensively dialysed into 20 mM phosphate buffer pH 6.5, filtered through a pore size of 0.22 μm to remove any precipitate and degassed by vacuum before measurements (Matsuo et al. 2005).

2.3.2 Fluorescence

Intrinsic fluorescence of a protein arises from Tyr and Trp emitters. The intensity of fluorescence emission gives information about the hydrophobicity of the environment these residues are located in, as they emission increases in a non-polar environment. In S-MSP fluorescence emission arises from a single Trp located in the hydrophobic core of the protein, and a number of Tyr residues found throughout the protein (Shutova et al. 2001). Folded S-MSP with the disulfide bond has a characteristic fluorescence peak at 310 nm. When S-MSP is unfolded or the disulfide bond reduced the maximum shifts to ~365 and is significantly broadened. Intrinsic fluorescence of S-MSP thus gives a rapid and accurate method to determine whether the protein is in its correctly folded state (Shutova et al. 1997).

Fluorescence emission from external emitters such the hydrophobic dye 1-anilino-naphthalene-8-sulfonate (1, 8 ANS, here called ANS) can be used to examine the tertiary structure of a protein (Poklar et al. 1997; Semisotnov et al. 1991). Transfer of ANS to a hydrophobic environment increases the intensity of its fluorescence emission. In a compactly folded protein the hydrophobic core is sequestered and
inaccessible to the probe; however in molten globular proteins with loosened tertiary structure the probe can enter and bind (Semisotnov et al. 1991).

All fluorescence measurements were made using a Fluoromax-3 (Jobin Yvon HORBIA) spectrofluorimeter. The intrinsic fluorescence of S-MSP was measured with a slit width of 5 nm and a concentration of approximately 5 μM. Samples were excited at 278 nm and recorded from 390 to 500 nm. For ANS fluorescence measurements a protein to ANS ratio of 1:10 was used, with protein concentrations of 2.5 μM and an ANS concentration of 25 μM. These concentrations were in the linear range of the detector when the slit width is set at 2 nm. The ANS was excited at 360 nm and the emission spectrum was recorded between 390 and 620 nm. The background fluorescence of the probe in buffer and subtracted it for each treatment.

2.3.3 Oxygen evolution measurements.

O₂ evolution was measured using a Clarke-type electrode (Hansatech) at 25 °C using either continuous saturating white light, or repetitive 15 Hz flashes of saturating white light to excite the sample. Assay conditions were 20 mM MES (pH 6.5), 300 mM sucrose, 10 mM NaCl, with para-phenyl-benzoquinone as a PSII electron acceptor plus varying amounts of CaCl₂ at a sample concentration of 10 μg of Chl for continuous light, or 30 μg of Chl for flashing light.

2.4 Differential scanning calorimetry

Differential scanning calorimetry (DSC) measures protein thermostability by monitoring heat absorbed by protein in solution as it unfolds. This allows direct measurement of the change in enthalpy during protein unfolding and the melting temperature. DSC is considered the technique of choice for characterising the thermodynamics of protein unfolding. As there is a vast body of literature describing the technique, its theory and applications, this section includes only the information necessary to interpret the data presented in the following chapters. For a more thorough treatment of the topic the reader is directed to the many comprehensive reviews listed in the references section (Edgcomb and Murphy 2000; Freire 1995a; Jelesarov and Bosshard 1999; Privalov 1997; Sturtevant 1987). The equations listed in this chapter are commonly used thermodynamic equations, and their derivations can be found in the reviews by Freire (1995a) Privalov (1997) and Sturtevant (1987).
2.4.1 Instrumentation and sample preparation

Instrumentation. The differential scanning calorimeter has two cells of equal volume-a sample cell which is filled with the protein dissolved in buffer and a reference cell containing only buffer (Figure 2-8). The cells are joined by a thermopile which measures the temperature difference and feeds back to a heater that maintains them at identical temperatures. Both cells are heated or cooled (scanned) at the same rate and the heat which is absorbed by the protein as it unfolds is recorded as the voltage from the thermopile. This method measures only the temperature change associated with protein unfolding; i.e. the endothermic unfolding transition and the change in heat capacity between the folded and unfolded forms of the protein. Because the reference cell contains only buffer, the contribution of the buffer’s heat capacity in both cells cancels out, leaving only the change in heat capacity due to protein unfolding, and making this technique extremely sensitive. The DSC experiments described here were carried out using a VP-DSC calorimeter (MicroCal, USA) which has disc-shaped cells made of tantalum with a volume of about 500μl. The temperature range of the calorimeter is -10 °C to +130 °C with a maximum scan rate of 90 °C per hour and a sensitivity of 0.5 μcal (°C)^{-1}.
Both cells heated linearly

Reference (buffer)

Sample (buffer + protein)

Compensation heating to maintain identical temperature

ΔT

Output

Folded

Un-folded

Temperature (°C)

Calories/mole °C

Figure 2-8. Schematic representation of a Differential Scanning Calorimetry experiment. Both the reference and sample cells have power supplied to raise the temperature in a linear manner. The sample cell contains the macromolecule, and the reference cell contains buffer. As the temperature is raised the protein begins to unfold, and increases its heat capacity. The temperature difference between the reference and sample cell is registered by the thermopile joining them (ΔT), which feeds back the sample cell and adjusts the temperature accordingly. The power supplied to the sample cell is recorded in the output as the change in heat capacity with temperature.

Sample preparation and experimental procedures. Protein solutions with concentrations between 5 and 20 μM (0.1-0.4 mg ml⁻¹) were extensively dialysed against several changes of buffer to ensure complete equilibration. Unless otherwise specified, the protein was dialysed against 20 mM MES, 1 mM EDTA pH 6.5, and then into several changes of 20 mM MES pH 6.5. The buffer from the final dialysis was saved and used in the reference cell. Both sample and reference solutions were degassed at 10 °C by vacuum and filtered through a 0.25 μm pore size before experiments. Calorimetric cells were kept at an excess pressure of about 25 psi to prevent degassing of the solution during the scan. Several buffer baselines where both cells are filled with buffer were obtained before each run of protein to equilibrate the calorimeter, and for each protein the temperature range was scanned four times at the maximum scan rate.
Data analysis. All data analysis was performed using the Micorcal Origin VP DSC data analysis software Version 7.0 programme provided with the instrument. A typical unfolding transition measured by DSC has a bell-shaped heat capacity curve, flanked on either side by relatively flat pre- and post-transitional baselines. The units of the raw data are cal (°C)⁻¹. For each sample scanned the corresponding buffer baseline was subtracted, and the data was normalised to the molar protein concentration, giving units of Kcal mol⁻¹ (°C)⁻¹.

A merged baseline was fitted to the pre- and post-transitional baselines, and subtracted from the normalised data. A theoretical model was then fitted to the data using the routines provided in the programme and the thermodynamic parameters were determined. The enthalpies of unfolding listed in the results section were calculated from the second scan, and then corrected for any irreversibility, unless otherwise stated. The reversibility of scans of the same sample was calculated by subtracting consecutive scans and calculating the enthalpy from the area of the difference.

2.4.2 Thermodynamics of unfolding

The equilibrium constant (K_{unf}) of the unfolding transition is given by:

\[ K_{unf} = \frac{[U]}{[N]} \]  

(1)

where the symbols [U] and [N] denote concentrations the of the unfolded and folded of the states protein. In the DSC experiment the fraction of folded to unfolded protein is determined by the heat absorbed by the protein as a function of temperature. The total area under the unfolding curve is proportional to the energy needed to unfold all of the proteins molecules in the sample. Thus, the area under the transition curve up to a particular temperature is the energy needed to unfold that fraction of the protein molecules. The unfolding transition is characterised by the melting temperature (T_m) where half of the molecules are in the native state and half are in the unfolded state. At T_m, K, is 1 and the Gibbs free energy of unfolding (ΔG_{unf}) is 0. The higher the T_m is, the more stable the protein is considered to be.

ΔG_{unf} also gives a measure of the stability of the protein. Processes will spontaneously proceed in the forward direction only when the ΔG is negative. So the more positive the ΔG_{unf} of a protein is at a certain temperature, the more stable it is. The ΔG_{unf} of a protein is determined as:
\[ \Delta G_{\text{unf}} = R \, T \ln K_{\text{unf}} \] (2)

where \( R \) is the gas constant (\( R = 1.98 \text{ cal mol}^{-1} \text{ K}^{-1} \)), \( T \) is the temperature in °K and \( K_{\text{unf}} \) is the equilibrium constant of unfolding.

The change in free energy of unfolding at a certain temperature (\( T_{\text{ref}} \)) has enthalpic, (\( \Delta H_{\text{unf}} \)), and entropic, (\( \Delta S_{\text{unf}} \)), contributions:

\[ \Delta G_{\text{unf}} = \Delta H_{\text{unf}} - T_{\text{ref}} \, \Delta S_{\text{unf}} \] (3)

\( \Delta H_{\text{unf}} \) at \( T_m \) can be measured directly during the DSC experiment, and is proportional to the area under the transition curve. When measured in this way, \( \Delta H_{\text{unf}} \) is referred to as the calorimetric enthalpy (\( \Delta H_{\text{cal}} \)). \( \Delta H_{\text{cal}} \) is approximately proportional to the number of non-covalent bonds in the protein (Edgcomb and Murphy 2000). \( \Delta H_{\text{unf}} \) can also be determined by the temperature dependence of the equilibrium constant as shown in:

\[ \Delta H^\text{vH} = RT^2 \left[ d \left( \ln K_{\text{unf}} \right) / d T \right] \] (4)

where \( R \) is the gas constant, \( T \) is the absolute temperature in °K, and \( K_{\text{unf}} \) is the equilibrium constant. When \( \Delta H_{\text{unf}} \) is calculated using this method, is referred to as the van’t Hoff enthalpy (\( \Delta H^\text{vH} \)).

The ratio of \( \Delta H^\text{vH} \) to \( \Delta H_{\text{cal}} \) is related to whether the transition is a two-state or a non two-state phenomenon. If the transition is strictly two-state the ratio is 1 and protein unfolding is ‘all-or none’ with the only states present being the completely native and the completely unfolded. Ratios less than 1 indicate that unfolding proceeds via intermediate states. Ratios greater than one indicate a degree of intermolecular cooperativity such as dimerisation during unfolding. The theoretical treatment and curve fitting of non-two state unfolding transitions are thoroughly described in a review by Sturtevant (1987).

The overall change in entropy, (\( \Delta S_{\text{unf}} \)) upon unfolding has contributions from both the protein and the solvent interactions. The gain in conformational freedom of the linear polypeptide when it unfolds increases the entropy, and favours the unfolded state. However the unfolded polypeptide has a greater surface for interaction with the solvent molecules, causing them to become more ordered. This ordering decreases the entropy of the system, and favours the folded state of the protein (Edgcomb and Murphy 2000).

The \( \Delta S_{\text{unf}} \) at \( T_m \) is determined by:

\[ \Delta S_{m\text{ un}} = \Delta H_{m\text{ un}} / T_m \] (5)
Heat capacity \((C_p)\) is how much heat it takes to raise the temperature of 1 mol of a substance by 1 °C. The heat capacity is different in the native and denatured states of a protein. The change in heat capacity upon unfolding of the protein \((\Delta C_p)\) is given by:

\[
\Delta C_p = C_{p,N} - C_{p,U}
\]  

\((6)\)

\(\Delta C_p\) can be calculated directly from the thermogram by the difference between the pre-transition and post-transitional baselines. Alternatively, the \(\Delta C_p\) can be derived from the slope of a \(\Delta H_{cal} vs T_m\) plot if it is assumed that \(\Delta C_p\) is constant in the investigated temperature region. The easiest way to do this is to scan the protein under investigation at a series of different pHs. At moderately acidic or alkali pHs globular proteins are often destabilised and unfold at lower temperatures, giving a lower \(T_m\) for the transition. A plot of the lower \(T_m\) \(\Delta H_{cal}\) pairs should have a linear relationship, the slope of which is proportional to the \(\Delta C_p\). This relationship is expressed as:

\[
\Delta H_{cal1} - \Delta H_{cal2} = \Delta C_p (T_{m1} - T_{m2})
\]  

\((7)\)

where the subscripts 1 and 2 denote the different melting temperatures. The difference in heat capacity between the folded and unfolded states is generally thought to be directly proportional to the change in solvent accessible polar and apolar surface area in the folded and unfolded states, and is a measure of the hydrophobic stabilisation of a protein. This relationship is given in:

\[
\Delta C_p = \Delta C_{p,ap} \cdot \Delta A_{ap} + \Delta C_{p,pol} \cdot \Delta A_{pol}
\]  

\((8)\)

where \(\Delta C_{p,ap}\) and \(\Delta C_{p,pol}\) are the apolar and polar contributions to the change in heat respectively.

### 2.4.3 Temperature dependence of unfolding

If a number of proteins with different unfolding temperatures are being compared, it is useful to compare them at a common reference temperature. To do this the thermodynamic parameters for these proteins are calculated at different temperatures. At temperatures outside of the unfolding transition, all the protein molecules will be in the folded state (below the transition) or the unfolded state (above the transition). The \(\Delta C_p\) value derived from the \(T_m vs \Delta H_{cal}\) plot can be used to calculate the enthalpy of
unfolding at a different temperature. From this the entropy and free energy can be calculated.

The enthalpy change of unfolding at any temperature $\Delta H_{\text{unf}}(T_{\text{ref}})$ can be calculated using:

$$\Delta H_{\text{unf}}(T_{\text{ref}}) = \Delta H_{\text{cal}} + \Delta C_p (T_{\text{ref}} - T_m) \quad (9)$$

where $T_m$ is the melting temperature, $\Delta H_{\text{cal}}$ is the enthalpy of unfolding at the $T_m$, $\Delta C_p$ is the change in heat capacity determined from the $\Delta H_{\text{cal}}$ vs $T_m$ plot, and $T_{\text{ref}}$ is the reference temperature the parameters are being calculated for. The entropy change of unfolding at a non-$T_m$ temperature ($\Delta S_{\text{unf}}(T_{\text{ref}})$) is given in:

$$\Delta S_{\text{unf}}(T_{\text{ref}}) = \Delta S_{\text{unf}} + \Delta C_p * \ln(T_{\text{ref}} / T_m) \quad (10)$$

where $\Delta S_{\text{unf}}$ is the entropy of unfolding at $T_m$, and all other parameters are as defined in equation (9).

The free energy of unfolding at a non-$T_m$ temperature ($\Delta G_{\text{unf}}(T_{\text{ref}})$) is given by:

$$\Delta G_{\text{unf}}(T_{\text{ref}}) = \Delta H_{\text{cal}} - T_{\text{ref}} \Delta S_{\text{unf}} + \Delta C_p (T_{\text{ref}} - T_m - T_{\text{ref}} \ln(T_{\text{ref}} / T_m)) \quad (11)$$

where all the parameters are as defined in equations (9) and (10). The free energy of unfolding as a function of temperature has a maximum where the protein is maximally stable.

2.5 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) monitors the binding of a ligand to a macromolecule by measuring the heat released or absorbed when the two species interact. The enthalpy of binding is directly measured as the area under the injection peak. By performing a series of injections until the macromolecule is saturated, the binding constant and entropy of the interaction can be determined. ITC is the only technique that can resolve the enthalpic and entropic contributions of binding, and a number of excellent reviews on the general technique are available (Doyle 1997; Ladbury 2004; Leavitt and Freire 2001; Pierce et al. 1999; Velazquez Campoy and Freire 2005; Wadso 1997). ITC has been used extensively to characterise the thermodynamics of binding of drugs to drug targets (Gaisford and Buckton 2001; Todd et al. 2000), as well as protein-protein interactions (Pierce et al. 1999), nucleic acid-ligand interactions (Feig 2007; Lane and Jenkins 2000), lipid-peptide interactions
(Wieprecht and Seelig 2002) and protein-carbohydrate interactions (Dam and Brewer 2004).

2.5.1 Instrumentation and sample preparation

Instrumentation

An ITC calorimeter has two cells of identical volume surrounded by an adiabatic jacket. A thermopile circuit detects differences in the two cells, and a heater on each cell maintains equal temperatures. One of the binding partners, usually the macromolecule, is in the stirred sample cell, and the other, usually the ligand, is in the syringe which feeds into the sample cell (Figure 2-9). The reference cell contains water. Before injection of the titrant, a constant power is applied to the reference cell, which forms a baseline and keeps the cells at a constant temperature slightly above the ambient temperature of the jacket. The observed signal is the time-dependent input of power needed to maintain equal temperatures in the sample and reference cells.
Both cells kept at constant temperature

Reference (Water)

Sample (Macromolecule)

Syringe (Ligand)

Compensation heating to maintain identical temperature

ΔT

Output

Time (min)

μcal/sec

Binding

Figure 2-9. Schematic representation of an Isothermal Titration Calorimetry experiment. Both the reference and sample cells have power supplied to keep them at a constant temperature. The sample cell contains the macromolecule and is stirred by the needle from the syringe. Aliquots of the ligand in the syringe are injected and bind to the macromolecule, absorbing or (in this case) releasing heat. The temperature difference between the reference and sample cell is registered by the thermopile joining them (ΔT), which feeds back the sample cell and adjusts the temperature accordingly. The power supplied to the sample cell is recorded in the output as a series of injection peaks with time.

The instrument used for these experiments is the Microcal VP-ITC (MicroCal Northampton USA). The sample and reference cells are made of Hastelloy® Alloy C-276, and have a volume of approximately 1.5 mls. The syringe has a volume of 500 µl, and the syringe needle doubles as a stirrer for the sample cell.

Sample preparation and experimental procedure. For the MSP-Ca\(^{2+}\) binding experiments in Chapter 4, MSP was dialysed extensively, and buffer from the last dialysis was used when making up the CaCl\(_2\) solution. The MSP concentration was approximately 20 µM (0.4 mg ml\(^{-1}\)) and the CaCl\(_2\) concentration was between 50 and 200 µM. All experiments were run at 25 °C with a stirring rate of 310 rpm. Each experiment consisted of 30 injections of 10 µl aliquots, spaced 2 minutes apart. For the
PSII-MSP binding experiments in Chapter 5, both species were resuspended in reconstitution buffer and any buffer mismatch was accounted for when the baseline was subtracted. PSII concentrations were 0.5 – 1 μM PSII in the sample cell, based on a stoichiometry of 2.5 μM PSII per (mg)$^{-1}$ of Chl (ml)$^{-1}$. MSP concentrations were 5 – 15 μM (0.1- 0.3 mg ml$^{-1}$) MSP in the syringe. All experiments were done at 10 °C, at a stirring rate of 270 rpm. Each experiment had 30 injections of 10 μl, spaced 2 minutes apart.

2.5.2 Determination of the thermodynamic parameters

Data analysis. The direct output from the calorimeter is a series of injection peaks. These are proportional to the energy input needed to compensate for the heat absorbed or evolved upon the ligand interactions and maintain the cells at the same temperature. The area under each peak is related to the enthalpy of binding of the ligand to macromolecule, and is proportional to amount of macromolecule/ligand complex formed. The amount of bound and unbound ligand in the cell at each injection is determined by the association constant $K_a$. As the concentration of the ligand in the cell increases with progressive injections, the system goes through different equilibrium states and the magnitude of the peaks decreases with each injection. The amount of heat evolved or absorbed (Q) upon binding is given in:

$$Q = V_0 \Delta H_{app} [M]_\text{tot} K_a[L] / (1 + K_a[L])$$

(12)

where $V_0$ is the cell volume, $\Delta H_{app}$ is the enthalpy of binding per mol of ligand, $[M]_\text{tot}$ is the total macromolecule concentration including both the bound and free states, $K_a$ is the association constant, and $[L]$ is the concentration of free ligand.

In addition to the heat change from the enthalpy of binding, the calorimeter also measures the enthalpies of dilution of the ligand and macromolecule which are visible as smaller changes after saturation. Normally these changes are small in comparison to the enthalpy of binding and can be corrected by subtracting a baseline either from the enthalpy change after saturation or by determining ligand and macromolecule dilution baselines from separate experiments. The data presented here were analysed after subtraction of a baseline using the points after saturation. Binding models which estimate the $K_a$ and ligand stoichiometry (n) were fitted to the corrected data using the Origin software provided with the VP-ITC instrument (version 7.0). The available
models for fitting the ligand-macromolecule are: (i) a single set of identical sites, (ii) two sets of independent sites, or (iii) multiple sequential sites. The model is fitted to the experimental data by an iterative process the details of which are given in the user manual. The association constant ($K_a$) is the strength of binding between the two molecules. $K_a$ is the inverse of the dissociation constant $K_d$, which is the concentration of ligand at which half the macromolecule sites are filled. For a reversible interaction the $K_a$ is given by:

$$K_a = \frac{[ML]}{[M][L]}$$  \hspace{1cm} (13)

where $[M]$, $[L]$ and $[ML]$ are the concentrations of the macromolecule, the ligand and the complex, respectively. The $K_a$ is determined directly from the ITC data by applying a linear regression model to the plot of $\Delta H_{bind}$ vs molar ratio. When a ligand and macromolecule go from free molecules to a complexed state, solvent-ligand and solvent-macromolecule interactions are lost and ligand-macromolecule interactions are gained. The overall free energy of binding ($\Delta G_{bind}$) is the net thermodynamic effect of these changes. $\Delta G_{bind}$ is determined by the $K_a$:

$$\Delta G_{bind} = -RT \ln K_a$$  \hspace{1cm} (14)

$\Delta G_{bind}$ has contributions from the binding enthalpy ($\Delta H_{bind}$) and the binding entropy ($\Delta S_{bind}$):

$$\Delta G = \Delta H - T \Delta S$$  \hspace{1cm} (15)

Different enthalpic and entropic contributions of binding can give the same $\Delta G_{bind}$. The more negative the $\Delta G_{bind}$, the more favourable binding is. This means that for the binding interaction to occur, the $\Delta H_{bind}$ of the reaction must be large and negative, or the $\Delta S_{bind}$ must be large and positive so that the $-T \Delta S_{bind}$ term gives an overall negative $\Delta G_{bind}$. The $\Delta H_{bind}$ reflects the strength of non-covalent interactions such as van der Waals interactions and hydrogen bonds between the interacting molecules, relative to their interaction with the solvent. The loss of macromolecule-solvent and ligand-solvent interactions is unfavourable, but the gain of new macromolecule-ligand bonds is favourable. A large gain of interactions between the molecules and a relatively small loss of interactions with the solvent has a net favourable enthalpy term. It is important
to note that the $\Delta H_{\text{bind}}$ for a reaction is the same throughout the titration regardless of the ligand concentration. The area under the injection peaks, which is a measure of the enthalpy of the reaction, decreases as the titration progresses because each injection represents a different equilibrium state with a different concentration of ligand.

The $\Delta S_{\text{bind}}$ has a favourable contribution to the interaction when solvent molecules, which were restricted by interaction with the macromolecule and ligand, are released. In aqueous solvent this destroys the hydrogen bond networks, which reduced their degrees of freedom. The increase in entropy from freedom of the solvent molecules is offset by the loss of conformational freedom of the ligand and macromolecule. The $\Delta S_{\text{bind}}$ at a certain temperature is determined from equation (15).

The change in heat capacity of binding ($\Delta C_p \text{ bind}$) depends partly on the burial of hydrophobic surface upon binding. $\Delta C_p \text{ bind}$ is determined by measuring the $\Delta H_{\text{bind}}$ of an interaction at a number of temperatures:

$$\Delta C_p = \frac{d (\Delta H_{\text{bind}})}{dT} \quad (16)$$
3 Reconstitution of O₂ activity with a recombinant thioredoxin- MSP protein

3.1 Introduction

The MSP protein has been produced for a number of species using recombinant methods, including expression as a fusion protein with *E. coli* thioredoxin (TRX) (Tohri et al. 2002). Fusion of target proteins with an *E. coli* protein such as TRX can increase their solubility and level of expression in *E. coli* cells (Lavallie and McCoy 1995). Often an affinity tag such as poly-histidine sequence is added to simplify purification as well. Both these strategies were employed in the purification of large amounts of soluble T-MSP as described in Section 2.2 of the previous chapter.

The current chapter shows that the fusion protein of the MSP from the thermophilic cyanobacterium *T. elongatus* and TRX from *E. coli* (T-MSP-TRX), can rebind to MSP-depleted spinach PSII. Most interestingly, it restores higher rates of O₂ evolving activity to MSP-depleted PSII than either the wild-type spinach MSP (S-MSP) or the isolated *T. elongatus* MSP (T-MSP).

3.2 Activity of the OEC determined by O₂ evolution

MSP can be reversibly dissociated from PSII in such a way that leaves the metal cluster intact so that PSII can still evolve O₂ at a very low rate. The depleted PSII can then be reconstituted with MSP from different sources and restore O₂ activity. Figure 3-1 and Table 3-1 show the O₂ evolution activity as a function of added calcium in extrinsic protein-depleted and reconstituted PSII samples from spinach. The untreated control (Figure 3-1 (●), Table 3-1) shows a Ca²⁺ effect under steady-state light in which the O₂ evolving activity increases from 460 (±15) to 663 (± 14) μmol O₂ (mg Chl)⁻¹ h⁻¹ or about 30 % of the total activity in the presence of 1 mM Ca²⁺. The percentage increase under repetitive flashing light shown in Table 3-1 is not as high, which indicates that in samples with no added calcium some centres have become reversibly inactivated while other centres are being rate limited. The inactivated centres may have been damaged by the loss of the MSP and the [Mn₄Ca] cluster during sample preparation (Miyao and Murata 1984a).
Figure 3-1. Rates of O₂ evolution from cross-reconstituted PSII measured at different calcium concentrations. Rates are expressed as a percentage of the maximum rate of the control PSII (■) measured at 3 mM calcium (100% of total activity was 663 (± 14) μmol O₂⁻¹ mg Chl⁻¹ hour); for PSII which has had the small extrinsic proteins removed (●); PSII with all the extrinsic proteins removed (▲); extrinsic depleted PSII reconstituted with MSP from spinach (◇); extrinsic depleted PSII reconstituted with MSP from T. elongatus (♦), or MSP-TRX fusion protein (○); PSII incubated with TRX(linker) (◁).
Table 3-1. O$_2$ yields from cross-reconstituted PSII under repetitive excitation with single turnover flashes, in the presence and absence of calcium.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Ca$^{2+}$ (mM)</th>
<th>Flash Illumination $^a$</th>
<th>Continuous Illumination $^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>O$_2$ activity (%)</td>
<td>Relative increase (%)</td>
<td>O$_2$ activity (%)</td>
</tr>
<tr>
<td>Untreated</td>
<td>0</td>
<td>87 (± 6)</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>100</td>
<td>13 (± 5)</td>
</tr>
<tr>
<td>NaCl washed</td>
<td>0</td>
<td>51 (± 7)</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>68 (± 7)</td>
<td>17 (± 7)</td>
</tr>
<tr>
<td>Urea washed</td>
<td>0</td>
<td>34 (± 4)</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>36 (± 6)</td>
<td>0 (± 6)</td>
</tr>
<tr>
<td>Reconstituted with S-MSP</td>
<td>0</td>
<td>35 (± 3)</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>45 (± 2)</td>
<td>9 (± 2)</td>
</tr>
<tr>
<td>Reconstituted with T-MSP</td>
<td>0</td>
<td>43 (± 3)</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>43 (± 6)</td>
<td>0 (± 6)</td>
</tr>
<tr>
<td>Reconstituted with T-MSP-TRX fusion protein</td>
<td>0</td>
<td>45 (± 2)</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>48 (± 1)</td>
<td>3 (± 11)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>72 (± 11)</td>
<td>27 (± 11)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>37 (± 3)</td>
<td>-7 (± 11)</td>
</tr>
</tbody>
</table>

$^a$ Rates are given as a percentage of the untreated sample with 5mM calcium 663 ± 14 μmol O$_2$ (mg Chl)$^{-1}$ hour$^{-1}$.

$^b$ Rates are given as a percentage of the untreated sample with 5mM calcium which has a photosynthetic unit size of 430 Chls per PSII.

Upon depletion of the two small extrinsic proteins by NaCl washing, the rate of O$_2$ evolution with continuous light drops to ~40% of the maximum rate for the untreated PSII. With both continuous and flashing light, the addition of 10 mM calcium to the NaCl washed sample raised the maximum rate to ~70% of the control (Figure 3-1 (●), Table 3-1) indicating that suboptimal calcium concentrations caused a about 30% rate limitation superimposed by 30% irreversible inactivation of the centres.

When the MSP as well as the small extrinsic proteins have been removed by urea washing, the rate drops to ~8% of the control and increases to only ~12% with 10 mM calcium, while in flashing light the number of centres are about a third of the control and there is no effect of calcium (Figure 3-1 (●), Table 3-1); however, there is a further small increase in the rate at higher calcium concentration where the maximum...
rate of O$_2$ production is achieved at 20 mM calcium to about 20 % of the control (Figure 3-1 (•)).

When extrinsic-protein depleted PSII (urea washed) is reconstituted with either S-MSP or T-MSP (Figure 3-1, (•) and (*) respectively and Table 3-1) O$_2$ evolution recovers to ~45 % of the original activity. Both samples still required the addition of 20 mM calcium to reach their maximum rates under continuous light; however under repetitive flashing light T-MSP reconstituted PSII showed no additional activation with the addition of calcium. This suggests that while calcium is reactivating centres in spinach reconstituted PSII, suboptimal calcium is rate limiting in T-MSP reconstituted PSII.

Most surprisingly however, when extrinsic-protein depleted PSII is reconstituted with the T-MSP-TRX fusion protein, over 60 % of the original activity is restored under optimal calcium concentrations (Figure 3-1 (▲)), 15 % more activity than either the native S-MSP, or T- MSP restored. The ~30 % difference between the intact control PSII and the NaCl-washed PSII is due to removal of the small extrinsic proteins. The loss of this activity is reversible, and can be restored by rebinding the small extrinsic proteins (Ghanotakis et al. 1984c). The ~50 % loss of activity between the NaCl-washed (small extrinsic depleted) PSII and the urea-washed PSII is due to removal of the MSP. Upon rebinding of the MSP to urea-washed (all extrinsic depleted) PSII the activity increases to about 30 % (Figure 3-1 (▲), Table 3-1). However there is still a 20 % difference between the MSP reconstituted PSII, and the NaCl-washed (small extrinsic depleted) PSII, both of which are comprised of the same complement of proteins, i.e. the core PSII membrane spanning proteins and the MSP. This discrepancy is most likely due to damage to the metal centres when the PSII has the MSP removed by the urea washing. This damage is irreversible such that the maximum activity that can be recovered by rebinding the MSP to PSII will always fall short of the activity of PSII which has had its extrinsic proteins removed by NaCl washing.

When PSII is reconstituted with T-MSP-TRX fusion protein, the maximum rate of O$_2$ evolution increases to ~65 %. If the 20 % loss in activity caused by removing and replacing MSP is taken into account for urea-washed PSII that is reconstituted with MSP-TRX, then the maximum activity would be very close to that of the control PSII.
which has the small extrinsic proteins as well as the MSP intact. This suggests that the fusion protein is reconstituting the activity by compensating for all of the extrinsic proteins, not just the MSP. The higher rate of \( \text{O}_2 \) evolution when the T-MSP-TRX fusion protein is rebound is not due to electrostatic interaction of the TRX-linker portion with PSII. Adding TRX-linker to the reconstitution reaction in place of MSP does not greatly increase the rate of \( \text{O}_2 \) evolution above that of urea-washed PSII (Figure 3-1, (a)). Additionally, reconstituting urea-washed PSII with a 1:1 mixture of T-MSP and TRX-linker did not produce a rate increase above that of T-MSP reconstitution alone suggesting that the TRX needs to be covalently linked to the MSP (data not shown).

The T-MSP-TRX fusion protein also appears to require a lower concentration of \( \text{CaCl}_2 \) to reach its maximum activity than either native spinach MSP or recombinant \( T. \ elongatus \) MSP. In Table 3-2, analysis of the affinity of PSII for calcium indicates that PSII reconstituted with MSP-TRX fusion protein binds calcium better than PSII reconstituted with MSP alone, or PSII depleted of the small extrinsic proteins. It also shows that spinach MSP has a lower affinity for calcium than \( T. \ elongatus \) MSP.

Above 5 mM \( \text{CaCl}_2 \), the rate of \( \text{O}_2 \) evolution restored by rebinding of the MSP-TRX fusion protein decreases back down to the level of the \( T. \ elongatus \) and spinach MSPs. A decrease in \( \text{O}_2 \) production activity is also observed for the untreated PSII above 20 mM \( \text{CaCl}_2 \). In both cases, the calcium may be dissociating the MSP, and for the intact PSII the small extrinsics, from the PSII core. The decrease in activity of the MSP-TRX reconstituted PSII at lower calcium concentrations could be because the decreased stability of the MSP-TRX fusion protein makes it easier to remove (see below).

### Table 3-2. Dissociation constants of Calcium for cross-reconstituted PSII.

<table>
<thead>
<tr>
<th>Sample</th>
<th>( \text{O}_2 ) evolving activity (%)</th>
<th>( \text{Ca}^{2+} \text{K}_d ) (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untreated</td>
<td>100</td>
<td>0.07</td>
</tr>
<tr>
<td>NaCl washed</td>
<td>70</td>
<td>1.22</td>
</tr>
<tr>
<td>Urea washed</td>
<td>18</td>
<td>10.7</td>
</tr>
<tr>
<td>Reconstituted with S- MSP</td>
<td>47</td>
<td>7.25</td>
</tr>
<tr>
<td>Reconstituted with T-MSP</td>
<td>52</td>
<td>3.85</td>
</tr>
<tr>
<td>Reconstituted with T-MSP-TRX</td>
<td>64</td>
<td>0.37</td>
</tr>
</tbody>
</table>

\( ^a \) Rates are given as a percentage of the untreated sample with 5mM calcium \( 663 \pm 14 \mu \text{mol} \text{ O}_2 \) (mg Chl)\(^{-1} \) hour\(^{-1} \).
3.3 Discussion

The results reported here show that a chimeric fusion protein of T-MSP and E. coli TRX restores higher rates of O₂ evolving activity when rebound to PSII than either the native S-MSP, or a recombinant T-MSP (Figure 3-1, Table 3-1). There are two possibilities for why fusion of TRX to T-MSP gives higher rates of O₂ evolution in reconstituted samples than S-MSP or T-MSP. Firstly, T-MSP-TRX might bind more readily to the PSII intrinsic proteins, meaning that a greater number of the PSII centres are actually reconstituted. The TRX (linker) is attached at the N-terminus of the MSP, a region which is involved in binding. Higher plant MSP has an insert in the N-terminus, which upon mutation alters the binding to PSII (Popelkova et al. 2002a; Popelkova et al. 2002b, 2003a). It is possible that attachment of TRX to T-MSP alters the structure in this region and allows PSII contact with regions on the MSP that are less accessible in the native MSP. The second possibility is that the MSP portion of the fusion protein reconstituted the same as T-MSP or S-MSP, but somehow has a superior capacity to restore activity once bound.

In addition to an increased in overall O₂ evolving activity, the MSP-TRX fusion protein appears to require a lower concentration of calcium to restore the maximum rate of O₂ evolution than either of the wild type MSPs (Table 3-2). However as discussed in the next chapter, the binding affinity of MSP for Ca^{2+} is much higher than the constants measured here, so it is likely that this effect is a consequence of the biphasic nature of the Ca^{2+} dependence of T-MSP-TRX reconstituted PSII.

From this data it is interesting to speculate that the TRX-linker portion is substituting for the function of one of the small extrinsic proteins. As discussed in Section 1.2.1 the extrinsic components of PSII vary between species, with the PsbP and PsbQ of higher plants being functionally replaced by PsbU and PsbV in most PSII centres of cyanobacteria. These proteins do not have a high sequence homology or a similar structure, but play a similar role in activating O₂ evolution.

The current chapter shows that interaction of the MSP with an unrelated protein improves its ability to restore activity to extrinsic-depleted PSII. In the following chapters calorimetric and structural data show that this increased activity is accompanied by structural changes which decrease the stability of the MSP, and dramatically increase its affinity for Ca^{2+}. ITC experiments indicate that T-MSP-TRX does not bind preferentially to PSII when compared to T-MSP or S-MSP, which suggest
that some property related to MSP function has been improved by fusion to TRX. The importance of protein-protein interactions and the flexibility of MSP for its role in $O_2$ evolution are discussed in detail in the final chapter.
4 Structure and thermal stability of isolated MSP

4.1 Introduction

In Chapter 3 it was shown that the PSII intrinsic subunits can be functionally reconstituted with S-MSP and T-MSP, and that these give similar rates of O₂ evolution. It was also shown that reconstitution of PSII with a T-MSP-TRX fusion protein gives higher rates of activity than either S-MSP or T-MSP. The purpose of the current chapter is to investigate the thermal stability of isolated, as opposed to PSII-bound, MSP. One of the most important questions addressed here is how the thermal stabilities of T-MSP from a thermophilic organism and S-MSP from a mesophilic organism compare. The following results show that while T-MSP has a cooperative unfolding transition measured by DSC, the S-MSP does not. Although this property of S-MSP precludes analysis of its thermal by conventional methods, it reveals that in the isolated state the unfolding process of S-MSP is very different to that of T-MSP, and does not conform to the expectations of a ‘well behaved’ globular protein. The implications of this are briefly summarised at the conclusion of this chapter, and are discussed in detail in Chapter 7.

The subsequent sections of this chapter deal only with experiments carried out on T-MSP and T-MSP-TRX since the thermal unfolding transitions of these proteins can be observed by DSC. The stability of these proteins was compared to investigate why the T-MSP-TRX protein restored activity to PSII better than T-MSP. The effects of pH and addition of Ca²⁺ and the osmolyte sarcosine were also investigated.

4.2 Thermal unfolding of spinach MSP

4.2.1 DSC measurements

DSC scans of S-MSP reveal that it lacks the unfolding peak typical of a cooperative unfolding transition (Figure 4-1). Instead there is a large upward slope in the heat capacity throughout the temperature range studied. This slope is fully reversible as seen by the repeat scan shown in red in Figure 4-1. The lack of an unfolding peak and the presence of the upward slope were seen in multiple independently prepared samples of native S-MSP, and samples could reconstitute PSII activity as described in Chapter 3.
The upward slope was not seen in scans of the buffer left over from the dialysis of the S-MSP suggesting this is a contribution from the protein.

\[ \text{Figure 4-1. The thermal unfolding of spinach MSP measured by Differential Scanning Calorimetry. This shows the lack of a typical unfolding peak. After subtraction of the instrumental baseline the thermogram was normalized to the molar protein concentration, and the start point of the scans set at zero. The upwardly sloping feature is reversible as shown by the repeatability of the first (black) and second (red) scans.} \]

\subsection*{4.2.2 Effect of calcium}

In Chapter 3 it was shown that the addition of Ca$^{2+}$ significantly increases the activity of PSII which has been reconstituted with S-MSP. Ca$^{2+}$ has also been shown to induce structural changes in S-MSP (Heredia and De Las Rivas 2003; Kruk et al. 2003). This raised the possibility that S-MSP is stabilised by Ca$^{2+}$ and may require it as a cofactor to assume its fully folded conformation. Addition of CaCl$_2$ however did not induce a cooperative unfolding transition (Figure 4-2). The only effect was to decrease the magnitude of the slope slightly.
In further attempts to induce a cooperative unfolding transition in S-MSP, its thermal unfolding was measured in the presence of sarcosine and reconstitution buffer. Sarcosine is a so-called protecting osmolyte that stabilises proteins, in some cases increasing their denaturation temperatures by up to 20 °C when added in molar concentrations (Santoro et al. 1992). Stabilisation is due to preferential hydration of the protein as sarcosine is excluded from the surface (Arakawa and Timasheff 1985). The interaction of sarcosine with the native protein is unfavourable compared to the interaction of water with the native protein; however, interaction of sarcosine or water with the unfolded protein is even more unfavourable. For this reason the protein is maintained in its native state, and the water concentration at the protein surface is increased. Most of the effect involves unfavourable interaction of the polypeptide backbone with the solvent, which overrides favourable interactions of the side chains with sarcosine (Liu and Bolen 1995). This increase in hydration favours a decreased surface area, causing the protein to adopt more compact structure. Addition of sarcosine
to random coil proteins also causes them to contract, which in turn is expected to promote folding by decreasing the entropy of the denatured state, and increasing the density of hydrophobic side chains in the denatured ensemble (Qu et al. 1998). For these reasons sarcosine was considered a potential agent for inducing a compactly folded state of S-MSP.

Reconstitution buffer is the buffer S-MSP is in when it functionally rebinds PSII as described in Section 2.1.3 of Chapter 2, and the results in Chapter 3. It contains 300 mM sorbitol, another osmolyte which has been shown to increase protein stability and induce folding in unstructured proteins (Kamiyama et al. 1999; Petersen et al. 2004; Wu and Bolen 2006). It also contains 10 mM CaCl$_2$, which was shown to induce conformational changes in S-MSP (Heredia and De Las Rivas 2003). Unfolding of S-MSP in reconstitution buffer was carried out to see whether the combination of these additives was able to induce a typical unfolding transition.

Surprisingly however, neither of these stabilising agents, sarcosine or sorbitol were able to induce a cooperative endothermic transition typical of the unfolding of a compactly folded globular protein (Figure 4-3). Sarcosine added at 1 M and 4 M increased the magnitude of the slope by the same amount, and reconstitution buffer increased it to a lesser extent. It is possible that this increase in the slope is due to stabilisation of the S-MSP protein by these chemicals. This slope feature is completely reversible for both buffers.
4.2.4 Tryptophan fluorescence

To confirm that the S-MSP protein used in the DSC experiments was correctly folded, its intrinsic fluorescence was monitored after excitation at 278 nm. This fluorescence arises from the single Trp residue and a number of Tyr (Shutova et al. 2001). In the correctly folded S-MSP with the disulfide bond oxidised, these emitters are in the hydrophobic interior of the protein and the emission has a maximum at \( \sim 310 \) nm. In unfolded reduced S-MSP it has a broad peak at 350nm (Shutova et al. 1997). As seen in Figure 4-4 the results confirm that S-MSP has retained its native structure.

Figure 4-3. DSC scans of spinach MSP in the presence of sarcosine or reconstitution buffer. 1M sarcosine (Blue) 4M sarcosine (green), reconstitution buffer (300mM sorbitol, 10mM NaCl, 10mM CaCl\(_2\)) (pink) or with no additions (black). All scans were completely reversible.
4.2.5 ANS Fluorescence

In contrast to S-MSP, T-MSP does exhibit a cooperative endothermic unfolding transition measured by DSC, which is discussed in detail in Section 4.3 on. To investigate whether the lack of a typical DSC transition in S-MSP can be correlated with a higher degree of molten globular character compared to T-MSP, the binding of both proteins to ANS was investigated.

The enhancement of fluorescence from the hydrophobic ANS is considered diagnostic of the molten globular state of proteins, and has been used extensively to identify acid denatured states and folding intermediates (Poklar et al. 1997; Semisotnov et al. 1991). Transfer of ANS to a hydrophobic environment increases the intensity of its fluorescence emission. In a compactly folded protein the hydrophobic core is sequestered and inaccessible to the probe; however in molten globular proteins with loosened tertiary structure the probe can enter and bind (Semisotnov et al. 1991). ANS binding to S-MSP has been studied previously as a function of pH. S-MSP was shown...
to cause no enhancement of fluorescence at neutral pH, but produced a considerable increase at pH 3.8 (Shutova et al. 1997).

The results presented here are in agreement with previously published data indicating a large enhancement of ANS fluorescence in the presence of S-MSP at low pH (Figure 4-5). Most interestingly the increase produced by T-MSP at any pH is lower than that of S-MSP at the same pH, indicating that the hydrophobic core of S-MSP is more accessible than that of T-MSP.
Figure 4-5. A) Emission spectra of ANS in the presence of Spinach MSP or *T. elongatus* MSP. Spinach MSP at pH 3 (dark green), pH 5 (red) and pH 8 (dark blue). *T. elongatus* MSP at pH 3 (light green) pH 5 (pink) or pH 8 (light blue). Each sample had the buffer emission at that pH subtracted, and was normalised to the protein concentration. Excitation = 346 nm. B) Fluorescence emission intensity of ANS at 475 nm in the presence of T-MSP (black) and S-MSP (red) as a function of pH.
4.3 Thermal unfolding of *T. elongatus* MSP

4.3.1 DSC measurements

Repeated scanning of T-MSP at pH 6.5 shows unfolding is highly reversible (> 90 %), indicating that this technique provides a reliable thermodynamic description of T-MSP unfolding. The unfolding transition fits to a non-two state model which indicates a high degree of intermolecular cooperativity (Figure 4-6, Table 4-1). The $\Delta H_{\text{cal}}$ was calculated directly as the area under the transition peak fitted by the model, and the $T_m$ is the midpoint of the transition. The $\Delta H^\text{vH}$ and $\Delta S_{\text{unf}}$ were calculated using equations (4) and (5) respectively.

**Table 4-1.** Thermodynamic parameters of the unfolding transition of *T. elongatus* MSP at pH 6.5. The sample was scanned from 20°C to 85°C. Values are from the second scan, corrected for irreversible denaturation.

<table>
<thead>
<tr>
<th>$\Delta H_{\text{cal}}$ (Kcal mol$^{-1}$)</th>
<th>$T_m$ (°C)</th>
<th>$\Delta H^\text{vH}$ (Kcal mol$^{-1}$)</th>
<th>Ratio of $\Delta H^\text{vH}/\Delta H_{\text{cal}}$</th>
<th>$\Delta S_{\text{unf}}$ (Kcal mol$^{-1}$ K)</th>
<th>Reversibility (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>57.7</td>
<td>76.7</td>
<td>145.8</td>
<td>2.6</td>
<td>0.4</td>
<td>91.1</td>
</tr>
<tr>
<td>(± 0.4)</td>
<td>(± 0.02)</td>
<td>(±1.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4-6. Differential Scanning calorimetry of *T. elongatus* MSP at pH 6.5 (black). After subtraction of the instrumental baseline the thermogram was normalized to the molar protein concentration and a merged pre- and post-transitional baseline were subtracted using the software supplied with the instrument. The theoretical fit of the unfolding transition to a non-two state model is shown in red.

**Effect of pH and calculation of the ΔC_P of unfolding.** The change in heat capacity between the folded and unfolded states of the protein (ΔC_P unf) was determined from the slope of a plot of ΔH_cal with T_m at pHs where the T_m is strongly pH dependent. T-MSP was scanned over a pH range of 2-8. There was no transition evident at pH 2, but in the range of pH 3 to pH 8 the T_m spanned a range of 17 °C with the greatest stability between pH 4 and 5 (Table 4-2, Figure 4-7). The ΔC_P unf determined by the dependence of ΔH_cal on T_m is 0.52 Kcal mol⁻¹ (°C)⁻¹ (Figure 4-8).
Table 4-2. Thermodynamic parameters for *T. elongatus* MSP as a function of pH. T-MSP was extensively dialysed into 50mM phosphate buffer at each pH.

<table>
<thead>
<tr>
<th>pH</th>
<th>ΔH&lt;sub&gt;cal&lt;/sub&gt; (Kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>T&lt;sub&gt;m&lt;/sub&gt; (°C)</th>
<th>ΔH&lt;sup&gt;°H&lt;/sup&gt; (Kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>ΔS&lt;sub&gt;unf&lt;/sub&gt; (Kcal mol&lt;sup&gt;-1&lt;/sup&gt; K)</th>
<th>Reversibility (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>28.2 (±0.6)</td>
<td>72.9 (±0.1)</td>
<td>146.2 (±3.9)</td>
<td>0.1</td>
<td>60.8</td>
</tr>
<tr>
<td>4</td>
<td>33.1 (±0.5)</td>
<td>80.6 (±0.1)</td>
<td>180.2 (±3.6)</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>5</td>
<td>36.8 (±0.4)</td>
<td>80.3 (±0.1)</td>
<td>167.5 (±2.4)</td>
<td>0.1</td>
<td>18.8</td>
</tr>
<tr>
<td>6</td>
<td>30.8 (±0.5)</td>
<td>77.5 (±0.1)</td>
<td>162.3 (±3.6)</td>
<td>0.1</td>
<td>91.2</td>
</tr>
<tr>
<td>7</td>
<td>34.6 (±0.5)</td>
<td>72.3 (±0.1)</td>
<td>155.3 (±3.3)</td>
<td>0.1</td>
<td>85.4</td>
</tr>
<tr>
<td>8</td>
<td>28.6 (±0.4)</td>
<td>68.5 (±0.1)</td>
<td>134.6 (±2.9)</td>
<td>0.1</td>
<td>64.3</td>
</tr>
</tbody>
</table>

Figure 4-7. pH dependence of the *T<sub>m</sub>* of *T. elongatus* MSP measured by DSC.
The thermal unfolding transition has a component contributed by the enthalpy of ionisation of the buffer as protons are released or taken up during protein unfolding (Fukada and Takahashi 1998). MES buffer has a large enthalpy of ionisation compared to phosphate buffer, so the effect of exposing ionisable groups is greater. The difference between the $\Delta H_{\text{cal}}$ measured in phosphate and the $\Delta H_{\text{cal}}$ measured in MES buffer at pH 6.5 is 23.9 Kcal mol$^{-1}$. The number ionisable side chains which are exposed during unfolding were estimated to be 6.4 groups per MSP. This was calculated by dividing the difference in the enthalpy of unfolding in the two different buffers by the enthalpy of ionization of MES, 3.7 Kcal mol$^{-1}$ (Fukada and Takahashi 1998).

4.3.2 The effect of disulfide bond reduction

The single disulfide bond of T-MSP was reduced by addition of β-mercaptoethanol and purging with nitrogen gas under vacuum for one hour. DSC scans were performed under anaerobic conditions to prevent re-oxidation. Reduction of the disulfide bond decreased the $T_m$ by more than 10 °C. The decrease in the $\Delta H_{\text{cal}}$ is within the range expected for such a decrease in $T_m$ from the plot in Figure 4-8. Reduction of the
disulfide bond also decreased the reversibility of the transition, probably due to aggregation.

Table 4-3. Thermodynamic parameters of the unfolding transition of reduced T. elongatus MSP at pH 6.5. Values are from the second scan, corrected for irreversible denaturation.

<table>
<thead>
<tr>
<th>Disulfide bond</th>
<th>( \Delta H_{\text{cal}} ) (Kcal mol(^{-1}))</th>
<th>( T_m ) (°C)</th>
<th>( \Delta H^{\text{vH}} ) (Kcal mol(^{-1}))</th>
<th>Ratio of ( \Delta H^{\text{vH}}/\Delta H_{\text{cal}} )</th>
<th>( \Delta S_{\text{unf}} ) (Kcal mol(^{-1}) K)</th>
<th>Reversibility (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxidised</td>
<td>57.7 (± 0.4)</td>
<td>76.7 (± 0.1)</td>
<td>145.8 (± 1.5)</td>
<td>2.6</td>
<td>0.4</td>
<td>91.1</td>
</tr>
<tr>
<td>Reduced</td>
<td>29.7 (±5.6)</td>
<td>65.3 (± 1.2)</td>
<td>98.2 (±36.7)</td>
<td>3.31</td>
<td>0.3</td>
<td>62.8</td>
</tr>
</tbody>
</table>

The \( \Delta G_{\text{unf}} \) for reduced and oxidised T-MSP was calculated from equation (3), and is given in Table 4-4. At 65 °C, the \( T_m \) of reduced T-MSP, the oxidised disulfide bond contributes 1.8 Kcal mol\(^{-1}\) to the protein stability. At 60 °C, the upper temperature limit at which T. elongatus cells can photosynthesise, the oxidised form is 2 Kcal mol\(^{-1}\) more stable than the reduced form.

Table 4-4. \( \Delta G_{\text{unf}} \) of oxidised and reduced T. elongatus MSP.

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>( \Delta G_{\text{unf}} ) (Kcal mol(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxidised</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>65</td>
</tr>
<tr>
<td>Reduced</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>76</td>
</tr>
</tbody>
</table>

4.3.3 The effect of sarcosine

The effect of sarcosine on T-MSP stability. As shown in Figure 4-9 and Figure 4-10 (A), addition of the osmolyte sarcosine to T-MSP increases the denaturation temperature of T-MSP at pH 6.5 in a linear manner up to 2M sarcosine. The effect tails off at higher concentrations where sarcosine becomes destabilising as has been seen in other studies (Santoro et al. 1992). Surprisingly, however, the \( \Delta H_{\text{cal}} \) of T-MSP decreases with increasing sarcosine at all concentrations above 0.5 M, giving a negative dependence of \( \Delta H \) on \( T_m \) (Table 4-5, Figure 4-10 B). Additionally, unfolding at concentrations of sarcosine above 1 M was irreversible, even when samples were scanned only to the end of the transition (data not shown). There was no evidence of aggregation during
unfolding which would be a large drop in heat capacity, and no aggregates were visible in the sample after heating.

Figure 4-9. The effect of sarcosine on differential calorimetric scans of *T. elongatus* MSP at pH 6.5. **Main figure:** Transitions after subtraction of a merged baseline clearly showing the increase in *T*^m^r. Scans are offset by ~ 5 kcal mol⁻¹ (°C)⁻¹ for clarity. Sarcosine concentrations: 0 (black), 0.5 M (red), 1 M (dark blue), 2 M (light blue), 3 M (yellow), 4 M (green). **Inset:** Scans before subtraction of baselines, showing the large positive change in heat capacity. For clarity the beginning of each transition was adjusted to zero.
### Table 4-5. Thermodynamic parameters of *T. elongatus* MSP unfolding in the presence of sarcosine.

<table>
<thead>
<tr>
<th>Sarcosine(^\ast) (mM)</th>
<th>(\Delta H_{\text{cal}}) (Kcal mol(^{-1}))</th>
<th>(T_m) (°C)</th>
<th>(\Delta H^{\text{VH}}) (Kcal mol(^{-1}))</th>
<th>Ratio of (\Delta H^{\text{VH}}/\Delta H_{\text{cal}})</th>
<th>Reversibility (%)</th>
<th>(\Delta S_{\text{unf}}) (Kcal mol(^{-1}) K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>57.7 ± (0.3)</td>
<td>76.7</td>
<td>145.0 ± (1.0)</td>
<td>2.5</td>
<td>91</td>
<td>0.2</td>
</tr>
<tr>
<td>0.5</td>
<td>59.2 ± (5.3)</td>
<td>80.2</td>
<td>140.0 ± (15.6)</td>
<td>2.4</td>
<td>97</td>
<td>0.2</td>
</tr>
<tr>
<td>1</td>
<td>55.9 ± (2.3)</td>
<td>84.2</td>
<td>172.0</td>
<td>3.1</td>
<td>62</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>54.1 ± (2.1)</td>
<td>89.4</td>
<td>195.0</td>
<td>3.6</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>50.1 ± (6.1)</td>
<td>93.6</td>
<td>229.0</td>
<td>4.6</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>47.8 ± (3.6)</td>
<td>94.4</td>
<td>318.0</td>
<td>6.7</td>
<td>0</td>
<td>0.1</td>
</tr>
</tbody>
</table>

* Solid sarcosine was added directly to aliquots of T-MSP in 20mM MES pH 6.5, which had a known concentration. The T-MSP concentration used in calculations was adjusted for the volume displaced by sarcosine.
Chapter 4. Structure and thermal stability of isolated MSP

Sarcosine concentration (M)

Figure 4-10. (A) Main figure: The effect of sarcosine on the $T_m$ of *T. elongatus* MSP. Inset: The linear fit the $T_m$ with 0, 0.5 and 1M concentrations of sarcosine. Slope = 7.54, and $R^2 = 0.99$. (B) Main figure: Plot of variation of $\Delta H_{cal}$ with $T_m$ when the $T_m$ is raised by addition of sarcosine. The slope of the fit is -0.57, with an $R^2 = 0.87$. Inset: Decrease of $\Delta H_{cal}$ with increasing sarcosine concentration Slope = -2.79 and $R^2 = 0.95$. 
The raw scans of T-MSP with sarcosine reveal a feature which could be significant- a large slope in the baseline which is more prominent at higher concentrations of sarcosine (Figure 4-9 inset). This is very similar to the feature seen in DSC scans of S-MSP (Figure 4-1). This slope is seen in repeat scans, even at concentrations where the unfolding peak was not repeatable. Control experiments of unfolding of RNAse with 4M sarcosine gave similar results to previously published data (Delpino and Sanchezruiz 1995). The sloping feature was not seen in scans for RNase at 71 μM, (0.9 mg ml⁻¹) indicating it is specific to T-MSP, and not caused by displacement of sarcosine by the protein. If the sloping feature in the S-MSP scans in Figure 4-1 is attributed to the unfolding of the molten globular state (see section 4.2 and the discussion below for the details of this hypothesis), then the similar feature in T-MSP with sarcosine scans suggests an explanation for the decrease in ΔH_cal with increasing T_m. If sarcosine is causing some regions of T-MSP to become more disordered, then these could give rise to the sloping feature. Other parts of T-MSP that remain well folded and still undergo a near two-state transition would be stabilised by sarcosine, increasing their T_m, but as less of the protein participates in the transition the apparent ΔH_cal is lower.

The effect of sarcosine on ANS fluorescence. To test whether addition of sarcosine increases the molten globular character of T-MSP, the effect of sarcosine on ANS binding by T-MSP was measured. ANS fluorescence was greatly increased by the addition of sarcosine to T-MSP, indicating that the hydrophobic core of T-MSP is more accessible (Figure 4-11). The addition of sarcosine alone to ANS greatly increases its fluorescence (Figure 4-11 inset), so the data presented here have had the buffer effect subtracted.
Chapter 4. Structure and thermal stability of isolated MSP

4.4 Structure and stability of T-MSP-TRX fusion protein

4.4.1 DSC measurements

The thermal unfolding of the T-MSP-TRX fusion protein has two distinct transitions which correspond to the unfolding of the MSP component (T_m = 74 °C) and the TRX component (T_m = 85°C) (Figure 4-12). Although the T_m of the MSP component is 2 °C lower than the T_m of the isolated T-MSP, the ΔH_cal and ΔH_vH of both proteins are very similar. The reversibility of the MSP peak in the T-MSP-TRX protein is probably lower because this protein was scanned to a higher temperature than T-MSP (95 °C rather than 85 °C). When only scanned to 85 °C the MSP peak of T-MSP-TRX is 89% reversible (Table 4-6).

Figure 4-11. ANS fluorescence in the presence of MSP from *T. elongatus* with sarcosine. **Main figure:** Emission spectra of ANS when added to *T. elongatus* MSP at pH 6.5 with the addition of 4M sarcosine (green), 1M sarcosine (red) no sarcosine (black). Each sample had the buffer emission at that sarcosine concentration subtracted, and was normalised to the protein concentration. **Inset:** Raw emission spectra before subtraction of buffer; ANS with *T. elongatus* MSP and 1M sarcosine (red), ANS with *T. elongatus* MSP and 4M sarcosine (green). Emission spectra with buffer; ANS with 4M sarcosine (blue) ANS with 1M sarcosine (pink).
Table 4-6. Thermodynamic parameters for the unfolding of T-MSP-TRX fusion protein at pH 6.5. Samples were scanned from 20 °C to 95 °C.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$\Delta H_{cal}^*$</th>
<th>Tm $^*$</th>
<th>$\Delta H^{\text{VH}}$</th>
<th>Ratio of $\Delta H^{\text{VH}}/\Delta H_{cal}$</th>
<th>$\Delta S_{unf}$</th>
<th>Reversibility $^\dagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Kcal mol$^{-1}$)</td>
<td>(°C)</td>
<td>(Kcal mol$^{-1}$)</td>
<td></td>
<td>(Kcal mol$^{-1}$ K)</td>
<td>(%)</td>
</tr>
<tr>
<td>T-MSP-TRX (MSP peak)</td>
<td>58.5</td>
<td>74.5</td>
<td>133.7</td>
<td>2.3</td>
<td>0.2</td>
<td>86 (89)</td>
</tr>
<tr>
<td>T-MSP-TRX (TRX peak)</td>
<td>25.3</td>
<td>85.2</td>
<td>123.6</td>
<td>2.7</td>
<td>0.1</td>
<td>94</td>
</tr>
<tr>
<td>TRX-linker</td>
<td>49.3</td>
<td>85.6</td>
<td>127.1</td>
<td>2.6</td>
<td>0.1</td>
<td>90</td>
</tr>
</tbody>
</table>

* Values are from the second scan, corrected for irreversible denaturation.

$^\dagger$ The value given in brackets is the reversibility of this transition when the sample is only scanned to 85 °C.
Figure 4-12. (A) DSC scan of T-MSP-TRX at pH 6.5 (black). The theoretical fit of the unfolding transition to a non-two state model with two transitions is shown in red. The best fit of the two transitions predicted from the model is shown in blue. (B) Superimposed thermograms from T-MSP (black) and TRX-linker (blue).
Effect of pH on T-MSP-TRX stability. Scans were measured of T-MSP-TRX at pHs in the range of 2- 6.5. Single transitions peaks are seen at over the pH range of 3, 4 and 5, whereas pH 6.5 which two transitions corresponding to the MSP and the TRX portions of the protein (Figure 4-13). The Tm of the single transitions at pHs 2 to 5 span a range of 20 °C (Figure 4-14). This single transition presumably arises from an overlap of the unfolding peaks of the MSP and the TRX components. A plot of the pH dependence of the Tm of T-MSP-TRX shows it has maximum stability at pH 5. Because the two unfolding peaks cannot be resolved, the ΔCp_unf of T-MSP-TRX cannot be determined.

Figure 4-13. Thermal unfolding of T-MSP-TRX at different pH. pH2 (black), pH3 (red), pH4 (green), pH 5 (dark blue), pH 6.5 (light blue). Scans are offset by ~10 kcal mol⁻¹ °C⁻¹ for clarity.
Figure 4-14. pH dependence of $T_m$ for the unfolding of T-MSP-TRX. $T_m$s at pHs 2-5 were for the single peak which is assumed to be a composite of the MSP and TRX transitions that are resolved at pH 6.5. The $T_m$ of the 6.5 point is the MSP transition only.

The temperature dependence of the $\Delta G_{\text{unf}}$ of T-MSP and T-MSP-TRX was calculated from equation (3). A plot of these values shows that T-MSP and T-MSP-TRX have very similar stabilities between 0 and 90 °C.
4. Structure and thermal stability of isolated MSP

Figure 4-15. Free energies of unfolding for \textit{T. elongatus} MSP (black) and the MSP portion of the T-MSP-TRX fusion protein (red) at pH 6.5. The values for \( \Delta H \) and \( T_m \) are as given in Table 4-2 and Table 4-6. The \( \Delta C_p \) value used in both calculations for the T-MSP protein are, given in Figure 4-8.

Table 4-7. Free energies calculated for \textit{T. elongatus} MSP and T-MSP-TRX at pH 6.5.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Temperature (°C)</th>
<th>( \Delta G ) (Kcal mol(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-MSP</td>
<td>25</td>
<td>6.43</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>2.54</td>
</tr>
<tr>
<td>T-MSP-TRX</td>
<td>25</td>
<td>8.32</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>2.28</td>
</tr>
</tbody>
</table>

4.4.2 Circular dichroism

In order to see whether covalent linkage of TRX to MSP alters the MSP secondary structure the CD spectra were measured. The results from the T-MSP-TRX fusion protein and a 1:1 mixture of the MSP plusTRX-linker are shown in Figure 4-16, as well as the result for T-MSP alone. There is a qualitative difference in the spectra, with the fusion protein having a higher ellipticity below about 205nm and then a more negative ellipticity above 205. A similar result has been reported that when the MSP is heated up
to 95 °C there is a significant increase in the negative ellipticity between 200 and 210 nm, which does not correspond to a large change in secondary structure (Loll et al. 2005a). The authors interpreted this result as a difference in folding or packing of the fusion protein, without a change in the secondary structure.

Figure 4-16. CD spectra of the T-MSP-TRX fusion protein (Black), a 1:1 mixture of T-MSP and TRX-linker (red), and T-MSP (green). The spectra have been normalized to the same concentration.

4.4.3 Effect of Calcium and urea

The rate of O$_2$ evolution in PSII samples reconstituted with T-MSP and T-MSP-TRX was strongly dependent on the concentration of calcium added to the assay buffer (see Chapter 3). The requirement of T-MSP reconstituted PSII for calcium was higher (10 mM for maximum activity) compared to T-MSP-TRX (3 mM for maximum activity). The enhancement of the rate by calcium in the T-MSP-TRX sample was biphasic, decreasing at concentrations above 3 mM (Figure 3-1). To investigate the effect of calcium on the thermal stability of isolated T-MSP and T-MSP-TRX, DSC scans were done at MSP dialysed into various concentrations of calcium.
The effect of calcium on the thermal stability of T-MSP. The addition of Ca\(^{2+}\) to T-MSP decreases $T_m$ and $\Delta H$ of unfolding, and the reversibility of refolding (Table 4-8 and Figure 4-18). By contrast the apparent van’t Hoff’s enthalpy is increased, which increases the $\Delta H^\text{vH}/\Delta H_{\text{cal}}$ ratio with increasing concentration, suggesting a large degree of oligomerisation at high Ca\(^{2+}\) concentrations (Table 4-8).

**Figure 4-17.** DSC scans of *T. elongatus* MSP at different Ca\(^{2+}\) concentrations. **Black** 0mM, **red** 0.25mM, **green** 1mM, **dark blue** 2mM, **light blue** 3mM, **pink** 5mM.
Table 4-8. Thermodynamic parameters of unfolding of *T. elongatus* MSP at varying Ca\(^{2+}\) concentrations.

<table>
<thead>
<tr>
<th>Ca(^{2+}) (mM)</th>
<th>(\Delta H_{\text{cal}}) *</th>
<th>Tm * (°C)</th>
<th>(\Delta H^{\text{VH}}) * (Kcal mol(^{-1}))</th>
<th>Ratio of (\Delta H^{\text{VH}}/\Delta H_{\text{cal}})</th>
<th>(\Delta S_{m}) (Kcal mol(^{-1}) K)</th>
<th>Reversibility# (%)</th>
<th>Activity † (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (±0.4)</td>
<td>57.7</td>
<td>76.7</td>
<td>145.8</td>
<td>2.5</td>
<td>0.2</td>
<td>91.1</td>
<td>30.8</td>
</tr>
<tr>
<td>0.25 (±1.3)</td>
<td>54.5</td>
<td>77.1</td>
<td>180.5</td>
<td>3.3</td>
<td>0.2</td>
<td>92.6</td>
<td>33.6</td>
</tr>
<tr>
<td>1 (±1.0)</td>
<td>47.6</td>
<td>76.3</td>
<td>175.6</td>
<td>3.7</td>
<td>0.2</td>
<td>88.8</td>
<td>36.8</td>
</tr>
<tr>
<td>2 (±4.5)</td>
<td>45.9</td>
<td>75.8</td>
<td>176.6</td>
<td>3.9</td>
<td>0.2</td>
<td>86.0</td>
<td>39.0</td>
</tr>
<tr>
<td>3 (±0.7)</td>
<td>49.1</td>
<td>75.4</td>
<td>179.8</td>
<td>3.7</td>
<td>0.1</td>
<td>85.3</td>
<td>40.0</td>
</tr>
<tr>
<td>5 (±2.7)</td>
<td>31.4</td>
<td>74.2</td>
<td>235.9</td>
<td>7.5</td>
<td>0.1</td>
<td>80.0</td>
<td>46.7</td>
</tr>
</tbody>
</table>

* Values \(\Delta H_{\text{cal}}, \text{Tm}\) and \(\Delta H^{\text{VH}}\) are from the second scan, corrected for irreversible denaturation.
# Reversibility is calculated from the area lost due to irreversible denaturation in consecutive scans.
† Activity is the percentage of O\(_2\) evolution rate as given in Table 3-1.

Figure 4-18. Dependence of \(\Delta H_{\text{cal}}\) and Tm of *T. elongatus* MSP unfolding with increasing Ca\(^{2+}\) concentration.
The ΔCₚ determined by a plot of the dependence of ΔH_cal on Tₘ gives a value of 8.2 Kcal mol⁻¹ °C⁻¹ (R² = 0.84) (Figure 4-19) which is much larger than the ΔCₚ measured by changing the pH. This suggests that the destabilising effect of CaCl₂ changes the mechanism of unfolding of these proteins. Addition of 1 mM NaCl did not change the unfolding of T-MSP, indicating that this is a Ca²⁺ effect.

![Graph](image)

**Figure 4-19.** Dependence of ΔH_cal on Tₘ for *T. elongatus* MSP when the protein is destabilized by addition of calcium.

*The effect of calcium on the stability of T-MSP-TRX.* Addition of calcium above 0.25 mM to T-MSP-TRX made the thermal transition completely irreversible due to aggregation. This is indicated by the sharp drop in Cᵥ of the scan above the Tₘ (Figure 4-20) and the complete lack of a transition upon rescanning (data not shown). The irreversible nature of the transition prevents equilibrium thermodynamic analysis of the data, and makes it is difficult to ascribe the apparent drop in Tₘ to a real decrease in stability as occurrence of aggregation upon unfolding has distorted the thermogram. Attempts to prevent aggregation by addition of glycerol were unsuccessful. Samples scanned to 85°C, the endpoint of the MSP-portion transitions, were also completely irreversible at all measured Ca²⁺ concentrations (0.25, 1, 2, 3 and 5 mM).
The effect of urea and calcium on T-MSP and T-MSP-TRX stability. The effect of CaCl₂ on the stability of T-MSP and T-MSP-TRX described above is somewhat counter-intuitive. For Ca²⁺ to bind to MSP the thermodynamics must be favourable, and this would stabilise the folded form. But the DSC measurements show that additional CaCl₂ in fact destabilises the MSP protein. A possible explanation is that Ca²⁺ binds to and stabilises an intermediate of MSP that has lower thermal stability. For this reason the thermal unfolding of T-MSP and T-MSP-TRX was examined in the presence of 3 M urea with and without CaCl₂. At this urea concentration the MSP subunit is released from PSII and may be in an intermediate non-functional state which lacks capacity to bind PSII.

The unfolding of both proteins was completely irreversible; however there was no evidence of aggregation in the thermogram (unlike Figure 4-20) or by visual inspection of the sample after heating. For this reason it is assumed that equilibrium thermodynamics could be applied, and the data from the first scan have been analysed by fitting to a non-two-state model as described above. The addition of 3 M urea...
increased the $T_m$ of T-MSP and T-MSP-TRX by over 14 and 13°C respectively (Table 4-9).

Interestingly, in the presence of 3 M urea, the addition of 5 mM Ca$^{2+}$ increased the $T_m$ of both proteins by almost 2°C, which does support the idea that Ca$^{2+}$ can bind to and stabilise partially unfolded intermediate states of MSP. The model proposed here is that Ca$^{2+}$ binds to an intermediate state of T-MSP, so addition of Ca$^{2+}$ to native T-MSP pulls the equilibrium towards this intermediate. Addition of urea to T-MSP also forces it into an intermediate state. In both cases the intermediate state has a lower $T_m$ than the native state. The intermediate state is thermally stabilised by Ca$^{2+}$, therefore addition of Ca$^{2+}$ to T-MSP in 3M urea causes an increase in $T_m$. The intermediate has a lower $T_m$ than the native state, so addition of Ca$^{2+}$ to native T-MSP appears to decrease the $T_m$.

### Table 4-9. Thermodynamic parameters of unfolding of *T. elongatus* MSP and T-MSP-TRX fusion protein in the presence of 3M urea.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Ca$^{2+}$</th>
<th>$T_m$ * (°C)</th>
<th>$\Delta H_{cal}$ * (Kcal mol$^{-1}$)</th>
<th>$\Delta H^{-1H}$ (Kcal mol$^{-1}$)</th>
<th>Ratio of $\Delta H^{+H} / \Delta H_{cal}$</th>
<th>$\Delta S_m$ (Kcal mol$^{-1}$ K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-MSP</td>
<td>+</td>
<td>63.1 (± 0.6)</td>
<td>22.9 (± 0.1)</td>
<td>10.8 (± 0.1)</td>
<td>0.5</td>
<td>0.1</td>
</tr>
<tr>
<td>T-MSP</td>
<td>-</td>
<td>61.4 (± 0.3)</td>
<td>31.6 (± 0.1)</td>
<td>13.6 (± 0.1)</td>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td>T-MSP-TRX</td>
<td>+</td>
<td>62.6 (± 0.5)</td>
<td>38.7 (± 0.1)</td>
<td>72.6 (± 0.1)</td>
<td>1.9</td>
<td>0.1</td>
</tr>
<tr>
<td>T-MSP-TRX</td>
<td>-</td>
<td>60.5 (± 0.6)</td>
<td>37.1 (± 0.1)</td>
<td>68.7 (± 0.1)</td>
<td>1.9</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Symbols: (-) = no Ca$^{2+}$ addition (+) = 5 mM Ca$^{2+}$ addition.

*All values are given from the first scan as all transitions were irreversible.

### ITC of calcium binding. In Chapter 3, binding of Ca$^{2+}$ to PSII reconstituted with T-MSP and T-MSP-PSII was determined by measuring the stimulation of O$_2$ evolving rates with increasing Ca$^{2+}$ concentrations. These experiments showed that the affinity of T-MSP-TRX reconstituted PSII was about 10 fold higher than the affinity of T-MSP reconstituted PSII.

ITC experiments measuring Ca$^{2+}$ binding by isolated T-MSP and T-MSP-TRX show they bind about 1 Ca$^{2+}$ ion each (Figure 4-21 and Table 4-10). Binding by T-MSP-TRX is three orders of magnitude higher than T-MSP and in both cases is a lot higher than for PSII reconstituted with T-MSP and T-MSP-TRX (Table 4-10). Ca$^{2+}$ binding by both proteins has a favourable $\Delta H_{bind}$, which is much larger for T-MSP-TRX than T-
MSP. The $TAS_{\text{bind}}$ term is very small for both proteins, but is slightly favourable for T-MSP and slightly favourable for T-MSP-TRX (Table 4-10).

Figure 4-21. ITC curves of (A) $Ca^{2+}$ binding to T-MSP and (B) $Ca^{2+}$ binding T-MSP-TRX. T-MSP and T-MSP-TRX had been dialysed against 1mM EDTA and then against several changes of MES pH 6.5. The raw titration data (top panel) shows that both interactions have a negative enthalpy of binding. The integrated area under the titration peaks (bottom panel, black) was fitted to a single set of sites model after subtraction of the baseline.

Table 4-10. Thermodynamics of $Ca^{2+}$ binding to *T. elongatus* MSP and T-MSP-TRX. Values calculated from fits of ITC data in Figure 4-21 to a single set of sites model.

<table>
<thead>
<tr>
<th></th>
<th>$N$ (number of sites)</th>
<th>$K_a (M^{-1})$</th>
<th>$\Delta H_{\text{bind}}$ (Kcal mol$^{-1}$)</th>
<th>$TAS$</th>
<th>$\Delta G$ (Kcal mol$^{-1}$)</th>
<th>$K_a (M^{-1})$ of reconstituted PSH*</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-MSP</td>
<td>0.81</td>
<td>$4.38 \times 10^4$</td>
<td>-6.23</td>
<td>0.10</td>
<td>-6.33</td>
<td>$2.60 \times 10^3$</td>
</tr>
<tr>
<td></td>
<td>(±0.219)</td>
<td>($\pm 1.53 \times 10^4$)</td>
<td>($\pm 1.897$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T-MSP-TRX</td>
<td>0.71</td>
<td>$4.29 \times 10^7$</td>
<td>-11.40</td>
<td>-0.99</td>
<td>-10.41</td>
<td>$2.70 \times 10^3$</td>
</tr>
<tr>
<td></td>
<td>(±0.008)</td>
<td>($\pm 2.40 \times 10^7$)</td>
<td>($\pm 0.113$)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* From Table 3-2 Chapter 3
4.5 Discussion

Comparison of the stability of S-MSP and T-MSP. The finding that S-MSP completely lacks a cooperative two-state transition measured by DSC, shown in Figure 4-1 is extremely surprising. However this result has some precedent and is probably due to its 'molten globular' or 'natively unfolded' structure (Lydakis-Simantiris et al. 1999; Shutova et al. 2000). A molten globular protein is defined as having a high proportion of secondary structure and a hydrophobic core, while lacking any significant tertiary structure (Uversky 2002b). It has been shown previously that some molten globular proteins lack a DSC unfolding transition. The apo form of α-lactalbumin has no transition measured by DSC, despite having normal secondary structural content measured by CD (Yutani et al. 1992b). Retinol A in its apo form at pH 2 has an extremely broad, low DSC transition, while the holo and neutral pH forms have normal transitions (Bychkova et al. 1992). α-fetoprotein at pH 3 in water completely lacks a cooperative unfolding transition measured by DSC but has a native Trp fluorescence and far UV CD (Uversky et al. 1995). The oligomeric multi-domain mustard globulin lacks a transition at pH 3, but dialysis into pH 7 buffer restores the typical transition (Marcone et al. 1997). These proteins all have considerable secondary structure measured by far UV CD, all have a hydrophobic core measured by near UV CD or intrinsic fluorescence, and all are considered to be molten globular judged by increased ANS emission or a larger apparent molecular weight by gel chromatography. The lack of a two-state transition measured by non-cooperative denaturant unfolding, or lack of a DSC transition is considered in some reviews to be another diagnostic feature of the molten globular state (Ptitsyn et al. 1995; Ptitsyn and Uversky 1994; Uversky 2002a). Many other molten globular proteins do have unfolding transitions with the typical bell-shaped heat capacity change measured by DSC, but in most cases the transition is shifted to lower temperatures, broadened compared to the transition of the fully native state and has low cooperativity with small $\Delta H^{\text{non-coop}} / \Delta H_{\text{cal}}$ ratios (Bychkova et al. 1996; Freire 1995b; Poklar et al. 1997). The lack of an endothermic unfolding peak in molten globular proteins is due to a lack of cooperativity. The thermodynamics of molten globular unfolding are reviewed in detail in (Freire 1995b). The large gradual increase in heat capacity during the DSC scan must be the result of some unfolding transition in S-MSP over this temperature range, albeit an atypical one. A study of natively disordered fragments of E. coli thioredoxin indicated an increase in $\Delta C_p$ values, even
without a cooperative DSC transition, is indicative of burial of hydrophobic area (Mendoza et al. 2003). The fragments used in these experiments are much smaller than S-MSP, between 30 and 70 residues, and are ‘intrinsic coils’ with little structure. The exposure of hydrophobic surface area during unfolding of the larger more structured S-MSP might account for the large change in apparent heat capacity.

In contrast to S-MSP, T-MSP does have a typical unfolding transition with a T_m of 78 °C (Figure 4-6). This would suggest that T-MSP has considerably more tertiary structure than S-MSP, and less molten globular character. This is supported by the ANS experiments which show that ANS has much higher emissions in the presence of S-MSP than T-MSP; especially at low pH (Figure 4-5). The findings presented in this chapter resolve some discrepancies in the literature which report that i) MSP is molten globular or nativly unfolded (Lydakis-Simantiris et al. 1999; Shutova et al. 2000; Zubrzycki et al. 1998) or ii) MSP has a compactly folded core with some flexible loop regions (Nowaczyk et al. 2004), as the former conclusion was based on experiments using S-MSP, while the latter used T-MSP.

The question now arises, why does T-MSP appear to have more tertiary structure than S-MSP? At first the answer might seem obvious- S-MSP has evolved to tolerate temperatures in the 20 to 30 °C range, while T-MSP evolved to function at 57 °C (Yamaoka et al. 1978). However this still means that in its normal functioning temperature range isolated S-MSP lacks a compact tertiary structure while T-MSP retains considerable tertiary structure until well above its functioning temperature. It also means that while T-MSP loses its secondary and tertiary structure simultaneously, S-MSP does not. One possibility is that S-MSP tertiary structure forms upon binding with the PSII intrinsic proteins, and is lost upon dissociation from PSII. This idea is described in more detail in Chapter 5, where measurements of the thermodynamics of the binding between MSP and PSII are presented.

Estimates of the S-MSP T_m have been reported based on unfolding of its secondary structure. A study using FTIR estimated the T_m to be 56 °C, although a significant portion of β-sheet structure remained between 60 and 70 °C (Heredia and De Las Rivas 2003). The T_m measured by far UV CD was estimated to be slightly larger at 61 °C, and is noted to occur over a broad range and show no extensive cooperativity (Lydakis-Simantiris et al. 1999). Near UV CD revealed that the hydrophobic core of S-MSP also re-forms after heating when the protein is cooled (Lydakis-Simantiris et al.
Some research reports aggregation of S-MSP above 70 °C (Heredia and De Las Rivas 2003), while others show that refolding is reversible up to 90 °C (Lydakis-Simantiris et al. 1999).

A previous DSC study of S-MSP by Kruk et al. (2003) reported an endothermic peak at 63 °C, and a broad transition spanning the 80 to 100 °C range. This thermogram is reproduced in Figure 4-22 (A), and quite obviously is not typical of a normal unfolding transition. Unlike the thermogram for S-MSP reported here, it has two separate transition peaks and the excess heat capacity drops after unfolding. However the Kruk et al. thermogram still does not conform to a typical protein unfolding transition as there are no distinct pre- and post-transitional baselines recorded; rather the transitions resembled two continuous humps. It is possible the difference between the thermograms from the Kruk et al. study and the present study is due to differences in sample preparation. Kruk et al. used Tris treatment to extract the native S-MSP and purified it by precipitation with ammonium sulfate, which could have damaged the protein in some way. The native S-MSP in the current study was prepared by sequential extraction of S-MSP with 2.6 M urea, 200 mM NaCl, and dialysis against MES buffer (see Section 2.1. for details) and was shown to be active in restoring O₂ evolution, and to have the intrinsic fluorescence maximum expected for the natively folded protein. An alternate possible explanation for the discrepancy on the results between the Kruk et al. study and the present study is their approach to the data analysis. The results presented in the current study closely resemble those of Kruk et al. if a merged baseline is fitted using the lowest and highest temperature points of the thermogram as the pre- and post-transitional baselines, and subtracted from the normalised data Figure 4-22 (B). If the data in the Kruk et al. study were analysed in this way is would explain the differences between these results- however no such analysis was mentioned in the paper. This approach to the data analysis was not considered an appropriate way to process the S-MSP unfolding data from the present study, as clear pre- and post-transitional baseline can not be distinguished from the transition.
Figure 4-22. (A) Thermogram of unfolding of S-MSP adapted from Kruk et al. (2003). (B) Data from Figure 4-1, after subtraction of a baseline which used the lowest and highest points of the transition as the pre- and post-transitional heat capacities. This treatment of the data is incorrect, and is only shown as an example.
Thermodynamics of T-MSP unfolding. The T_m of T-MSP measured here by DSC (Table 4-1) is in good agreement with previous estimates discussed in Section 1.2.3 (Loll et al. 2005a; Sonoyama et al. 1996). The stability of T-MSP at mid to low pH shown in Table 4-2 and Figure 4-7 is consistent with studies on S-MSP which show no change in secondary structure between pH 7 and 3.8 (Shutova et al. 1997). The difference in ΔH_cal for unfolding in phosphate and MES at pH 6.5 indicates about 6 ionisable groups are exposed with unfolding. At pH 6.5 normally only His with a solution pKa of 6 is titrateable. But measurement of the buffer capacity of S-MSP showed that the pKa of several Asp and Glu residues is increased to about 6 (Shutova et al. 2005). T-MSP has only two His residues, but based on the behaviour of S-MSP, some of the 25 Asp and Glu residues are probably ionisable as well.

As discussed previously in Section 1.2.3, MSP has a single disulfide bond which is necessary for its correct solution structure. It was previously estimated from chemical denaturation experiments that the disulfide contributes about 5 Kcal mol⁻¹ to the stability of S-MSP (Tanaka et al. 1989). This is in excellent agreement with the value reported here for T-MSP (Table 4-3 and Table 4-4). A site directed mutant of S-MSP which lacked the Cys residues that form the disulfide bond was reported to be thermostable, judged by its ability to reconstitute activity after heating to 90 °C (Wyman and Yocum 2005b). That study reported about 50 % of the sample aggregated, which is in agreement with the data reported here that the scans were only 60 % reversible. The S-MSP study did not examine the effect of temperatures between 25 °C and 90 °C, so no comparison in the change of T_m between reduced S-MSP and T-MSP can be made.

The addition of sarcosine to T-MSP produced a very interesting result as there is a negative dependence of ΔH_cal on T_m shown in Figure 4-9 Figure 4-10, and Table 4-5. The raw data also had a large sloping feature in the transition seen in the inset of Figure 4-9, similar to that seen in scans of S-MSP in Figure 4-1. A possible explanation is that while sarcosine stabilises some regions of T-MSP, increasing their T_m, other portions become disordered which produces the baseline slope, and accounts for the decreased ΔH_cal. This suggestion is supported by ANS fluorescence data which show an increase in the molten globular character of T-MSP with addition of sarcosine (Figure 4-11). This result is unexpected as sarcosine normally acts to stabilise protein structures and cause them to adopt more compact conformations, increasing the ΔH_cal along with the T_m (Ibarra-Molero et al. 2000; Kamiyama et al. 1999; Qu et al. 1998; Santoro et al.
A negative dependence of $\Delta H_{\text{cal}}$ on $T_m$ has not been reported, although the $T_m$ of RNAse A with 8 M sarcosine stays constant despite an increase in $T_m$ of over 20 °C (Santoro et al. 1992). This explanation for the unusual behaviour of T-MSP in sarcosine is a hypothesis only. A considerable amount of experimental work must be done before conclusions can be drawn about the molecular basis of this result, and about the implications of this for the structure of T-MSP. Measurement of a range of pHs at each sarcosine concentration would determine the $\Delta C_P$ and allow calculation of the temperature dependence of $\Delta G$ for each sarcosine concentration.

**Thermodynamics of T-MSP-TRX unfolding.** The DSC data presented in Figure 4-12 and Table 4-6 show that the MSP portion of the fusion protein has a slightly lower thermal stability than the MSP alone. The CD data in Figure 4-16 indicate that the T-MSP-TRX fusion protein has an altered structure compared to a mixture of the two in solution. These results taken together with the activity measurements in Chapter 3 suggest that covalently linking TRX to the N-terminus of MSP alters its structure in a way that increases its flexibility or causes it to be more loosely packed, and this improves its binding capacity or activity. The next chapter shows that T-MSP-TRX binds with much the same affinity as T-MSP and S-MSP.

**The thermodynamics of $Ca^{2+}$ binding to MSP.** The binding affinity of $Ca^{2+}$ measured for reconstituted PSII in Chapter 3 is very different from that measured for isolated MSP in the current chapter, presented in Figure 4-21 and Table 4-10. This is because the reconstituted PSII samples were not depleted of $Ca^{2+}$, and what was being measured was the effect of added calcium. The isolated MSP samples had been dialysed against EDTA to remove any tightly bound $Ca^{2+}$ ions. Based on the $K_a$ values measured by ITC, the high affinity $Ca^{2+}$ binding sites on MSP were already filled in the PSII experiments. This would mean that the stimulation of activity by increased $Ca^{2+}$ was due to other effects at other sites on the PSII protein complex.

This suggests that despite isolated T-MSP-TRX having a higher $Ca^{2+}$ affinity than T-MSP, the apparent high affinity of the T-MSP-TRX reconstituted PSII complex is not due to $Ca^{2+}$ binding to MSP. The apparent high affinity of PSII reconstituted with T-MSP-TRX is probably a consequence of the biphasic nature of the Ca dependence of this chimeric complex. If the inhibitory effect of high CaCl$_2$ levels were not observed
then the Ca\(^{2+}\) dependence might more closely resemble that of PSII reconstituted with S-MSP and T-MSP, and the O\(_2\) evolving rates would be further increased at higher CaCl\(_2\) concentrations. A clue to the origin of the decrease in activity of PSII reconstituted with T-MSP-TRX at high Ca concentrations may be found in Figure 4-20. In the presence of increasing Ca concentrations T-MSP-TRX aggregates upon unfolding, and it is possible that a degree of aggregation is occurring in the conditions of the oxygen electrode chamber which interferes with the function of the T-MSP-TRX. Binding of Ca\(^{2+}\) slightly lowers the thermal stability of T-MSP and T-MSP-TRX, shown in Table 4-8 and Table 4-9, and in Figure 4-17, Figure 4-18, Figure 4-19 and Figure 4-20. This is consistent with previous findings that show calcium binding to isolated S-MSP induces conformational changes which increase the loop content of the protein and thermally destabilise it (Heredia and De Las Rivas 2003; Kruk et al. 2003). Although similar secondary structural changes were not observed in T-MSP (Loll et al. 2005a), a bound calcium ion was identified in the crystal structure near the exit of a proposed water channel (Murray and Barber 2006). The binding constant for Ca\(^{2+}\) binding to T-MSP measured here by ITC (4.38 x 10\(^{-4}\) M\(^{-1}\)), is lower than that reported previously for S-MSP (1 x 10\(^{6}\) M\(^{-1}\)), but is similar to the value for Mn binding to S-MSP (1 x 10\(^{4}\) M\(^{-1}\)) (Shutova et al. 2005).

It has been suggested that Ca\(^{2+}\) binding is part of an activation process for converting the MSP to an ‘open’ conformation which allows controlled access of the substrate water to the active site. At low pHs such as would occur in the lumen during photosynthesis, the core of S-MSP was observed to become more accessible to solvent (Shutova et al. 1997; Shutova et al. 2003), and more able to bind calcium. Binding of calcium causes further structural changes and decreasing the thermal stability of the MSP (Shutova et al. 2005). It is possible the reason T-MSP-TRX has a higher binding affinity for Ca\(^{2+}\) is that the presence of TRX in the fusion protein alters the structure of the MSP portion so it is pre-organized for calcium binding.
5 MSP binding affinity for PSII in cross-reconstitution experiments

5.1 Introduction

The intention of this chapter is to examine the thermodynamics of binding of S-MSP T-MSP and T-MSP-TRX to the PSII intrinsic proteins from spinach. One of the main questions this chapter intends to answer is: does S-MSP assume a more compact tertiary structure upon binding to PSII? As discussed in Chapter 4, DSC shows that isolated T-MSP has both secondary and tertiary structure in solution at 55 °C, its functioning temperature, and that these structures are lost simultaneously upon heating. However, S-MSP lacks a cooperative unfolding transition measured by DSC, even though it has been shown previously to retain its secondary structure at temperatures well above 25 °C – the temperature it functions at (Heredia and De Las Rivas 2003; Lydakis-Simantiris et al. 1999). The lack of a cooperative DSC transition is interpreted as a lack of tertiary structure which is consistent with the assignment of MSP as a molten globule. Together the DSC and ANS fluorescence data indicate that isolated S-MSP has far more molten globular character than isolated T-MSP. This leads to the hypothesis that unbound S-MSP is molten globular and may adopt a compact tertiary structure upon binding to the PSII intrinsic subunits.

The importance of MSP thermostability for PSII function was discussed by Pueyo et al. (2002) in their paper investigating the increased heat tolerance of O2 evolution from spinach PSII reconstituted with MSP from the thermophilic cyanobacterium *P. laminosum* (Pueyo et al. 2002). They suggest that loss of O2 evolution from spinach PSII with increasing temperature is due to dissociation of the native MSP from the PSII complex, and that the thermal tolerance in the *P. laminosum* MSP-PSII complex is because this MSP remains bound at higher temperatures. They suggest that dissociation of the native S-MSP arises from conformational changes in the 40 to 50 °C range, below the temperature where secondary structure is lost. Pueyo et al. (2002) also suggested that these changes correspond to the thermal unfolding transitions measured by Cramer et al. (1981) and Thompson et al (1986, 1989) in DSC experiments of whole PSII complexes. Cramer et al. (1981) carried out DSC studies of thylakoids. This revealed six endothermic transitions at 42, 54, 65, 72, 79, and 84 °C, denoted A, B,
C, D, E and F (Cramer et al. 1981). The first three of these transitions were further resolved by Thompson et al. (1986) who repeated the DSC experiments on PSII enriched fragments. In these experiments the ‘A’ peak was deconvoluted into two transitions denoted ‘A1’ and ‘A2’ shown in Figure 5-1 (Thompson et al. 1986). Thermal gel analysis, thermal inactivation experiments and EPR measurements allowed assignment of peaks B C to the PSII core proteins. Peak D includes the components of the LHCP, and a 29 kDa protein (Thompson et al. 1989). The ‘A1’ and ‘A2’ peaks were generally assigned to ‘denaturation of the OEC’ in all three studies. The ‘A2’ peak generated the most interest as it was the sharpest and most variable. A number of treatments of the PSII were used to manipulate the ‘A2’ peak. These experiments are summarised in Table 5-1.

Figure 5-1. DSC of PSII adapted from Thompson et al (1986). Red, raw experimental data. Black, endothermic component curves fitted by deconvolution of the experimental data, Blue, sum of the component curves.
Table 5-1. Experimental treatments of PSII carried out in an attempt to assign the ‘A1’ and ‘A2’ thermal unfolding peaks measured by DSC.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Effect</th>
<th>Conclusion</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heat treatment to 49 °C for 5 minutes or treatment with trypsin prior to scanning</td>
<td>Abolishes ‘A’ transition</td>
<td>‘A’ corresponds to the denaturation of some donor side component of PSII.</td>
<td>(Cramer et al. 1981)</td>
</tr>
<tr>
<td>Measured the rate of O₂ evolution at the ‘A’ peak Tₘ</td>
<td>Abolition of O₂ evolution coincides with ‘A’ transition</td>
<td>‘A’ represents denaturation of some OEC component</td>
<td>(Cramer et al. 1981)</td>
</tr>
<tr>
<td>Measurement of Cyt b559 potential at the same temperature as the ‘A’ peak</td>
<td>Loss of the high potential form of Cyt b559 coincides with the ‘A’ peak</td>
<td>‘A’ represents oxidation of Cyt b559</td>
<td>(Cramer et al. 1981)</td>
</tr>
<tr>
<td>Measurement of released [Mn₄Ca] cluster at the same temperature as the ‘A’ peak</td>
<td>Release of Mn²⁺ from the OEC coincides with the ‘A’ peak</td>
<td>Denaturation at ‘A’ leads to breakdown of the [Mn₄Ca] cluster</td>
<td>(Cramer et al. 1981)</td>
</tr>
<tr>
<td>NaCl depleted PSII before scanning samples</td>
<td>Abolishes the ‘A2’ transition</td>
<td>Removal of the small extrinsic lowers the potential of Cyt b559, so the loss of ‘A2’ is actually due to the loss of high-potential Cyt b559</td>
<td>(Thompson et al. 1986)</td>
</tr>
<tr>
<td>Oxidation of Cyt b559</td>
<td>Abolishes the ‘A2’ transition, leaving a broad peak with a lower Tₘ</td>
<td>Oxidation causes loss or broadening of the ‘A1’ or ‘A2’ peak, possibly by reducing its cooperativity.</td>
<td>(Thompson et al. 1986)</td>
</tr>
<tr>
<td>Tris washing to remove all three extrinsic proteins</td>
<td>Abolishes ‘A2’</td>
<td>Removal of the all three extrinsic proteins lowers the potential of Cyt b559, which in turn reduces the peak cooperativity.</td>
<td>(Thompson et al. 1986)</td>
</tr>
<tr>
<td>Thermal gel analysis of PSII</td>
<td>No protein denaturation corresponding to ‘A2’</td>
<td>The A2 transition is actually due to loss of the [Mn₄Ca] cluster from PSII</td>
<td>(Thompson et al. 1989)</td>
</tr>
<tr>
<td>Cl⁻ depletion</td>
<td>Broadens ‘A2’ and lowers its Tₘ</td>
<td>Consistent with it being part of the OEC</td>
<td>(Thompson et al. 1989)</td>
</tr>
<tr>
<td>Higher pH (7.5)</td>
<td>Lowers Tₘ</td>
<td>OH⁻ is competing for the Cl⁻ binding site.</td>
<td>(Thompson et al. 1989)</td>
</tr>
</tbody>
</table>

An interesting possibility is that ‘A2’ corresponds to a structural transition occurring in the MSP that leads to its release from the PSII complex. This possibility is consistent with the findings from previous research, presented in
Table 5-1: First, the concurrent of loss of $O_2$ evolving activity and dissociation of the Mn ions of the $[\text{Mn}_4\text{Ca}]$ cluster occurring at the same temperature as the ‘A2’ transition is consistent with dissociation of the MSP. (Miyao and Murata 1984c). Second, heat treatment, trypsin digestion and Tris washing which were all observed to abolish the ‘A2’ transition also remove MSP (Enami et al. 1994; Yamamoto et al. 1981). Third, the removal of the two small extrinsic proteins also abolishes ‘A2’. The small-extrinsic proteins are bound in the vicinity of the MSP, and PsbP interacts directly with MSP. Such interactions could influence the structure and stability of the MSP in the PSII complex, so in the absence of the small extrinsic proteins the MSP’s transition would be different. The data presented in Chapter 4 also support the assignment of these transitions to MSP dissociation. In particular thermal gel analysis showed no protein denaturation at the transition, which is consistent with the hypothesis presented here that the transition is due to loss of tertiary structure rather than complete denaturation. Additionally, the pH sensitivity of the ‘A2’ transition is consistent the calorimetric data for T-MSP showing that it is less stable at higher pHs.

To investigate the hypothesis that the ‘A2’ transition corresponds to the release of MSP from the PSII complex, DSC scans were performed on spinach PSII which had first been depleted of its native MSP and then been reconstituted with T-MSP and S-MSP. Although it was reported that the ‘A2’ peak was abolished by removal of the small extrinsic proteins it was hoped that using a double difference approach where scans of MSP-depleted PSII are subtracted from scans of MSP-reconstituted PSII the contribution of the MSP to the overall thermal unfolding profile could be derived.

A second approach was also undertaken to measure the thermodynamics of MSP binding to PSII. As discussed in Chapter 2 Section 2.4.2, the enthalpy change upon interaction of a ligand and a macromolecule involves contributions from the binding partners (usually favourable) and the solvent (usually unfavourable due to loss of interactions with the solvent molecules). The enthalpy change contributed by the binding partners arises from non-covalent interactions at the binding face as well as the enthalpy change from any conformational change in the macromolecule and ligand upon binding. This approach has never been used to measure the interaction between the
MSP and PSII. If as hypothesised, S-MSP undergoes a significant conformational change upon binding to PSII but T-MSP does not, it is expected that this will be reflected in its enthalpy of binding. In this scenario, S-MSP is expected to have a larger $\Delta H_{\text{bind}}$ compared to T-MSP. A large conformational change in either the ligand or the macromolecule upon binding will also be reflected in the entropy change upon binding. If S-MSP does increase its tertiary structure when it binds to PSII the entropy is likely to be more unfavourable as S-MSP will experience conformational restrictions.

ITC experiments also provide a more accurate measure of the binding constant and stoichiometry of MSP to PSII, and it is hoped that these experiments will settle the debate that has existed for a number of years over the number of copies of MSP bound to higher plant PSII. Previous determinations of the MSP per PSII stoichiometry relied on correlation of O$_2$ evolving activity with the intensity of MSP bands on an SDS-PAGE gel, and have lead to disagreement about the number of copies of S-MSP bound to spinach PSII (Murata et al. 1984; Xu and Bricker 1992). A low resolution three dimensional structure of spinach PSII indicated there is only one copy of bound S-MSP (Nield et al. 2002; Nield et al. 2000b). But mutagenesis of recombinant S-MSP has identified a sequence found only in higher plants which was hypothesised to be responsible for rebinding a second copy (Popelkova et al. 2002a; Popelkova et al. 2002b, 2003a). In contrast to any assertions about the stoichiometry of MSP in plants, the high resolution crystal structures of *T. elongatus* show only one copy is bound in cyanobacterial PSII (Ferreira et al. 2004; Loll et al. 2005b).

This chapter also uses ITC to examine the binding of the T-MSP-TRX fusion protein to spinach PSII. As discussed in Chapter 3 in Section 3.3, the ability of T-MSP-TRX to reconstitute higher rates of activity than T-MSP or S-MSP could be due an enhanced ability to bind to spinach PSII, or to a superior function compared to S-MSP and T-MSP once it is bound. ITC provides an accurate method for measuring the binding constant of these MSPs to spinach PSII.

### 5.2 DSC of reconstituted PSII

The results from DSC scans of PSII-enriched membranes are very similar to those of the earlier research (Thompson et al. 1989; Thompson et al. 1986). Thompson et al. (1986, 1989) observed that the ‘A2’ transition is not observed in PSII preparation lacking the small extrinsic proteins, and the ‘A2’ is not visible in any of the scans in
Figure 5-2. The difference graph where the extrinsic depleted scan was subtracted from the reconstituted scans showed that both reconstituted samples had a broad transition (Figure 5-2 B) between 40 and 50 °C. These transitions were not present in scans of the MSP-depleted sample. Since both T-MSP and S-MSP-reconstituted samples have similar transitions this is not the release of S-MSP. Attempts to fit a model to this peak were unsuccessful because the upper and lower baselines can not easily be distinguished from the rest of the scan. It is possible that this feature corresponds to the ‘A1’ peak or to a broadened ‘A2’ peak and may be related to unfolding of the intrinsic proteins. There were also peaks in the 65-75 °C regions in the subtracted scans which are probably the consequence of comparatively small differences in the large heat capacity changes being subtracted from each other. The unfolding of T-MSP is expected to occur in this temperature range, but is difficult to distinguish from these large peaks. The absence of the slope attributed to S-MSP is harder to explain. It could be attributed to a potential ‘swamping’ of the S-MSP transition in the much larger signal from the whole PSII. An alternate possibility is that a transition is now occurring between 60-70 °C, but this is indistinguishable from the other peaks in this region. There is a shoulder on the difference thermogram from S-MSP reconstituted PSII which is absent in T-MSP reconstituted sample. However this could also be due to experimental error it was noted in the previous calorimetric investigations that scans showed a large degree of variability, even of the same sample.
Figure 5-2. A) DSC scans of MSP-depleted and reconstituted spinach PSII. Green, PSII depleted of MSP Black, PSII depleted and then reconstituted with spinach MSP Red, PSII depleted and then reconstituted with *T. elongatus* MSP B) Difference between reconstituted scans and depleted scans. Black, spinach MSP-reconstituted PSII minus depleted PSII. Red, *T. elongatus* MSP-reconstituted PSII minus depleted PSII.

5.3 Binding of MSP to PSII measured by ITC

ITC experiments show that S-MSP and T-MSP both bind to spinach PSII with a stoichiometry of approximately one (Figure 5-3 Table 5-2). The binding affinities of
both MSPs are similar, with T-MSP being slightly higher. The binding of both MSPs is
driven by a negative enthalpy change, which is 15 Kcal more favourable for S-MSP
than for T-MSP. In both cases the entropy change is unfavourable but this is to a lesser
extent for T-MSP than for S-MSP.

Figure 5-3. ITC curves of PSII depleted of MSP, then reconstituted with (A) spinach MSP and (B) T.
elongatus MSP. The raw titration data is given in the upper panels. The integrated area under the titration
peaks is shown in the bottom panels (black) and was fitted to a single set of sites model after subtraction
of the baseline (red).

Table 5-2. Thermodynamics of S-MSP and T-MSP binding to spinach PSII calculated from fits ITC data
from Figure 5-3 to a single set of sites model.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>N</th>
<th>$K_a$ (M$^{-1}$)</th>
<th>$\Delta H_{bind}$ (Kcal mol$^{-1}$)</th>
<th>$T \Delta S_{bind}$ (Kcal mol$^{-1}$)</th>
<th>$\Delta G_{bind}$ (Kcal mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-MSP</td>
<td>1.13(±0.1)</td>
<td>1.7 x 10$^{-7}$ (±5.5 x 10$^{-6}$)</td>
<td>-89.8 (±4.7)</td>
<td>-80.0</td>
<td>-9.9</td>
</tr>
<tr>
<td>S-MSP</td>
<td>1.27 (±0.1)</td>
<td>6.07 x 10$^{-6}$ (±1.5 x 10$^{-6}$)</td>
<td>-105.5 (±10.6)</td>
<td>-96.3</td>
<td>-9.3</td>
</tr>
</tbody>
</table>

The binding of the T-MSP-TRX fusion protein to PSII was also examined using
ITC. T-MSP-TRX binds to PSII with a stoichiometry greater than 1, and has a best fit to
a model with two consecutive binding sites (Figure 5-4). The binding affinity of the first site is almost an order of magnitude higher than the second site, and is very similar to the affinity measured for S-MSP and T-MSP binding to PSII. The enthalpy for both sites is favourable, but less so than for either S-MSP or T-MSP, and the entropy is unfavourable. The first binding site is probably the same as for S-MSP and T-MSP binding to PSII, and probably represents functional binding. The second site may be non-specific as has been observed in many mutants with perturbed binding properties (Betts et al. 1998; Betts et al. 1996b, 1997b; Frankel et al. 1999; Motoki et al. 1998; Popelkova et al. 2002a; Popelkova et al. 2002b). This seems particularly likely given the propensity of T-MSP-TRX to self-associate with low concentrations of CaCl₂ shown in DSC experiments in Chapter 4, Section 4.4.3. Assuming the first site is functional then the higher activity restored to PSII by rebinding T-MSP-TRX is not due to a superior ability of T-MSP-TRX to rebind to PSII. Instead, T-MSP-TRX must work better once it is bound.
Figure 5-4. ITC curve of PSII depleted of MSP, then reconstituted with T-MSP-TRX. The raw titration data is given in the upper panels. The integrated area under the titration peaks is shown in the bottom panels (black) and was fitted to a sequential sites model with two binding sites after subtraction of the baseline.

Table 5-3. Thermodynamics of T-MSP-TRX binding to spinach PSII calculated from the fit of ITC data from Figure 5-4 to sequential binding with two sites model.

<table>
<thead>
<tr>
<th>Site number</th>
<th>$K_a$ (M$^{-1}$)</th>
<th>$\Delta H_{\text{bind}}$ (Kcal mol$^{-1}$)</th>
<th>$-T\Delta S_{\text{bind}}$ (Kcal mol$^{-1}$)</th>
<th>$\Delta G_{\text{bind}}$ (Kcal mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$7.7 \times 10^6$ ($\pm 4.6 \times 10^6$)</td>
<td>$-39.8$ ($\pm 3.4$)</td>
<td>$-30.4$</td>
<td>$-9.3$</td>
</tr>
<tr>
<td>2</td>
<td>$9.5 \times 10^7$ ($\pm 2.8 \times 10^7$)</td>
<td>$-35.1$ ($\pm 5.4$)</td>
<td>$-26.9$</td>
<td>$-8.1$</td>
</tr>
</tbody>
</table>

5.4 Discussion

DSC of PSII-enriched membranes. The DSC scans of PSII-enriched membranes (Figure 5-2) did not provide any compelling evidence for S-MSP undergoing a low-temperature transition upon release from PSII. The only transition observed in the 40 – 50 °C range was seen in both T-MSP and S-MSP reconstituted PSII. It is still possible that this
transition may be related to MSP release; however it would come from a structural change in the intrinsic proteins. Since isolated T-MSP does not have any transition in this temperature range, it must be assumed in the T-MSP reconstituted sample that this feature arises from unfolding of the PSII intrinsic proteins. Since the transition is about the same magnitude in the S-MSP reconstituted sample as the T-MSP reconstituted sample there is no addition from simultaneous unfolding of S-MSP. Presumably both T-MSP and S-MSP underwent unfolding transitions after release from PSII above 50 °C. The most likely reason the T-MSP transition is not visible in the difference thermogram is that it would coincide with the large peaks of the intrinsic proteins denaturing. The slope of S-MSP is probably not seen as it is broad and not easily distinguished from the rest of the baseline.

In the studies by Thompson et al. (1986) denaturation of the core proteins of PSII were assigned to peak ‘B’, between 50 and 60 °C. The 40 – 50 °C peak observed here might be from changes in only the regions where the extrinsic proteins and the OEC metal cluster bind, rather than denaturation of the whole protein which happens at a higher temperature. As mentioned by Thompson et al. (1986), the ‘A2’ transition is highly cooperative, indicated by its sharpness, and is not seen in samples depleted of the small extrinsic proteins. The cooperativity of ‘A2’ could be due to enthalpic contributions from a number of PSII subunits and cofactors undergoing concerted rearrangements upon dissociation of the MSP. These would include the small extrinsic proteins which would be released concomitantly with the MSP, and the manganese cluster.

ITC of MSP rebinding to PSII. ITC experiments measuring the thermodynamics of MSP binding to the PSII intrinsic proteins provide unique thermodynamic insights into the binding of MSP to PSII. The results support the hypothesis that S-MSP undergoes a conformation change upon binding. Binding of both S-MSP and T-MSP to PSII is driven by a favourable enthalpy change on the order of 100 Kcal mol⁻¹, and has an unfavourable entropy change of ~90 Kcal mol⁻¹. As discussed in Chapter 2 Section 2.5.2, a negative ΔHₘₚₑₙ means that the non-covalent interactions between the binding partners are greater than the interaction between the molecules which are lost upon binding. This is consistent with the numerous studies reviewed in Table 1-3 which show that MSP and PSII electrostatic interactions drive the binding of MSP to PSII. The
\( \Delta H_{\text{bind}} \) for S-MSP is more negative than for T-MSP, which would be consistent with structural rearrangements occurring when it binds, if it is forming a more compact structure. If as hypothesised the isolated T-MSP already has the ‘bound’ tertiary structure there will be a smaller contribution from conformational changes. The unfavourable entropy (Table 5-2) indicates that the penalty of restricted conformation of the MSP and PSII is not offset by freedom gained from the solvent molecules. The slightly more favourable \(-T\Delta S_{\text{bind}}\) of T-MSP (-80 for T-MSP and -97 for S-MSP) could be from it having a more compact tertiary structure, and already being more pre-organised, so the entropy penalty is lower when it binds. This could be an adaptation to life at high temperatures. We already know that O\(_2\) evolution is temperature sensitive (Enami et al. 1994; Pueyo et al. 2002), and if the MSP component does not have to form a compact structure when it binds, it will not suffer the associated entropy penalty, which has an even greater effect at higher temperatures. This is strategy is used in drug design to improve binding affinity; pre-constrained conformations use entropy rather than enthalpy to drive binding, and give a large increase in affinity without the need to engineer specific reactions. (Todd et al. 2000). Previous studies indicate that 30- 40 \% of the backbone of S-MSP undergoes structural changes upon binding to PSII, which is consistent with this work (Hutchison et al. 1998a, 1998b). Studies of T-MSP using NMR also indicate it has less structure in isolation than when bound to PSII (Nowaczyk et al. 2004) so although the conformational changes may not be as great as for S-MSP, there is probably some stabilisation of disordered regions.

Although these ITC of experiments support the hypothesis that S-MSP has a greater conformational change upon binding to PSII than T-MSP, more work is needed to confirm this. A series of ITC MSP-PSII binding experiments run at multiple temperatures will allow a determination of the heat capacity of the binding interaction. Multiple temperature experiments have been used to measure the enthalpy contribution of the molten globular-to-native transition which occurs when \(\alpha\)-lactalbumin binds Ca\(^{2+}\) ions (Vanhooren et al. 2002). A similar study of the MSP-PSII interaction would be highly desirable. Additional experiments to determine the release or uptake of protons with binding will also be necessary, as differences in protonation could also account for different \( \Delta H_{\text{bind}} \) values. An ITC study of T-MSP binding to PSII from \(T.\ elongatus\) is also necessary. A weakness of the system used here is that because T-MSP is binding to the PSII from a different species, binding will not be as optimal as for
binding to PSII from *T. elongatus* and the smaller enthalpy change of binding could result from fewer interactions at the binding site compared to binding of S-MSP to its native partner, spinach PSII. Another interesting application of ITC to PSII would be to measure the interaction of spinach PSII and S-MSP with binding site mutations like those described in Chapter 1, Section 1.2.4 to determine the enthalpic contribution of these specific sequences to the $\Delta H_{\text{bind}}$.

ITC provides a powerful tool for characterising the thermodynamics of binding between proteins and ligands, and no doubt will prove a useful approach in further determination of the PSII-MSP interaction. However, the reader must appreciate here the difficulty involved in running ITC experiments like the ones described in this chapter. Most ITC experiments involve measuring binding of relatively small ligands like metal ions, oligonucleotides and drugs to proteins. Of the protein-protein interactions studied by ITC, nearly all have been of one monomeric stable protein binding to another. In this study the macromolecule is a large multi-subunit protein complex in a membrane which is sensitive to both light and temperature. To ensure that the PSII remained intact, all sample preparation and degassing was done in the dark. The experiments were run at 10 °C, which introduced very long equilibration times, as the calorimeter had to cool to this temperature. Optimising the sample concentrations was also difficult. Running the experiment with dilute samples did not give a clear saturation of binding, but higher sample concentrations were limited by the concentration of MSP which had to be 10 to 15 times that of PSII, and by aggregation of the PSII sample in the cell at high concentrations. The data presented are the result of extensive optimisation of the reaction conditions to get a reproducible result. Running these experiments under different conditions, particularly different temperatures will no doubt present similar challenges. Despite this, more ITC experiments with the binding of MSP to PSII are planned to further characterise the thermodynamics of this interaction.
6 Binding and characterisation of MSP from *Gloebacter violaceus*

6.1 Introduction

In this chapter the structure of the recombinant G-MSP from the primitive contemporary cyanobacterium *Gloebacter violaceus* is characterised and its surprising inability to rebind to spinach PSII is investigated.

*G. violaceus* is believed to have diverged the earliest during the evolutionary radiation of cyanobacteria and has a number of interesting features as a consequence of its position in the evolutionary tree. One of these is the absence of thylakoids (Gugliemi 1981; Turner et al. 1999). Instead the Photosystem I (PSI) and PSII complexes are inserted into the plasma membrane with the ‘lumenal’ portions oriented toward the periplasmic space between the plasma membrane and the cell wall (Gugliemi 1981; Koenig and Schmidt 1995). The MSP from *G. violaceus* is an interesting candidate for structural and functional study for three reasons. First, it is the most divergent MSP of any species thus far identified, and may therefore retain primitive features of a cyanobacterial ancestor which could give clues to the evolution of photosynthetic water oxidation. The MSPs of all species have a low primary sequence homology (~10 % maximum) with other known proteins (De Las Rivas and Barber 2004), and G-MSP may provide a structural ‘missing link’ to the origins of the MSP. Second, the binding of G-MSP to PSII occurs in the periplasmic space rather than the lumen of a thylakoid and therefore it functions in a different environment to other MSPs. Third, G-MSP does not contain certain residues which appear conserved in all other species and have been implicated in substrate or product channels (De Las Rivas and Barber 2004; Ferreira et al. 2004; Murray and Barber 2007). The conservation of these amino acids is summarised in Table 6-1. Of the proposed channel residues contributed by the MSP R152 and D159 are conserved in all species, and are part of the DPKGR conserved region which is thought to stabilize the catalytic manganese cluster (De Las Rivas and Barber 2004).
Table 6-1. Conservation of residues contributed by MSP that are proposed to form part of a water channel in PSII.

<table>
<thead>
<tr>
<th>Residue in T. elongatus</th>
<th>Residue in most other species</th>
<th>Residue in G. violaceus</th>
<th>Conservation in G. violaceus</th>
</tr>
</thead>
<tbody>
<tr>
<td>R152</td>
<td>R</td>
<td>R</td>
<td>Fully conserved</td>
</tr>
<tr>
<td>D159</td>
<td>D</td>
<td>D</td>
<td>Fully conserved</td>
</tr>
<tr>
<td>D222</td>
<td>D</td>
<td>G</td>
<td>Acidic to hydrophobic</td>
</tr>
<tr>
<td>D223</td>
<td>T</td>
<td>L</td>
<td>Acidic to basic</td>
</tr>
<tr>
<td>D224</td>
<td>D</td>
<td>A</td>
<td>Acidic to hydrophobic</td>
</tr>
<tr>
<td>H228</td>
<td>K</td>
<td>E</td>
<td>Basic to acidic</td>
</tr>
<tr>
<td>E229</td>
<td>A, V, P (hydrophobic), T or Q (hydrophilic)</td>
<td>V</td>
<td>Acidic to hydrophobic</td>
</tr>
</tbody>
</table>

As discussed in Section 1.2.1, a very interesting feature of the MSP-PSII interaction is that PSII intrinsic subunits can be functionally reconstituted with the MSP from other species that comes from either native or recombinant sources. The spinach PSII intrinsic proteins, for example, have been reconstituted with MSP from another higher plant *A. thaliana*, a red alga *C. caldarium*, a green alga *C. reinhardtii*, and the thermophilic cyanobacteria *T. vulcanus*, *T. elongatus* and *P. laminosum* (Betts et al. 1994; Enami et al. 2000; Koike and Inoue 1985; Pueyo et al. 2002). Rebinding of these seemingly quite heterologous MSPs generally restores similar rates of O$_2$ evolving activity to PSII as rebinding the native spinach MSP. The experiments discussed in previous chapters have extended these findings; Chapter 3 describes how reconstitution of spinach PSII with the T-MSP-TRX fusion protein restores higher rates of O$_2$ evolving activity than reconstitution with either native S-MSP, or T-MSP alone, and Chapter 5 shows that the thermodynamics of T-MSP rebinding to spinach PSII are very similar to those of the rebounding of native S-MSP. Both results confirm the remarkable exchangeability of the MSP-PSII interaction and make the inability of G-MSP to rebind spinach PSII all the more surprising.

To investigate the why G-MSP is unable to bind to spinach PSII, the three dimensional structure of G-MSP was simulated using the Swiss-Model comparative modelling server (http://swissmodel.expasy.org/SWISS-MODEL.html). This server generates a model based on alignment with known structures with high sequence
homology; in this case the *T. elongatus* MSP from the 1S5L crystal structure (Ferreira et al. 2004). The initial model is then energy minimized using the Gromos96 program. As Swiss-Model did not predict the N- and C-termini of the *G. violaceus* MSP because of low sequence homology with *T. elongatus* MSP in these areas, these were built manually and energy minimized using the HyperChem program (Hypercube USA). Sequence alignment was done using CLUSTALW. Residues involved in the interaction with the PSII intrinsic proteins were identified from De Las Rivas and Barber’s analysis of the 1S5L structure (De Las Rivas and Barber 2004). The results of the modelling study indicate that the residues at the binding interface with PSII differ greatly in G-MSP compared to T-MSP and S-MSP, which is probably why it cannot rebind to spinach PSII. The implications of this for the evolution of MSP are discussed at the end of this chapter.

6.2 Characterization of recombinant *G. violaceus* MSP

*G. violaceus* has proven challenging to culture in large quantities. It is very sensitive to light and temperature and has a long generation time. It also has low levels of O2 production and a low PSII content in its membranes. As these features make extraction and characterization of native PSII proteins from *G. violaceus* very difficult, a protocol for expression and purification of recombinant G-MSP has been developed. Recombinant G-MSP is expressed as inclusion bodies in *E. coli* under normal expression conditions, and two different strategies were used to obtain soluble protein; refolding of purified inclusion bodies by denaturation and dilution from urea, and co-expression with chaperone proteins.

SDS-PAGE gels of the purified refolded G-MSP, and purified soluble G-MSP from co-expression with chaperones in Sections 2.2.3 and 2.2.4 shows single bands at the expected apparent molecular weight of ~30 kDa. The circular dichroism (CD) spectrum (Figure 6-1) of G-MSP was almost identical to that of native spinach MSP and recombinant *T. elongatus* MSP, indicating the protein is correctly folded. Analysis of the CD spectrum by the K2d program (Merelo et al. 1991) shows that G-MSP contains similar proportions of secondary structural elements to S-MSP and T-MSP (Table 6-2). The CD spectra for G-MSP prepared by the two different methods were identical indicating that the method of refolding did not affect the secondary structure (data not shown).
Figure 6-1. CD spectrum of *G. violaceus* MSP (green), *T. elongatus* (red) and spinach (black).

Table 6-2. Proportions of secondary structural elements of MSP from *G. violaceus*, *T. elongatus* and Spinach.

<table>
<thead>
<tr>
<th>Sample</th>
<th>% α-helix</th>
<th>% β-sheet</th>
<th>% random coil</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>G-MSP</td>
<td>4</td>
<td>48</td>
<td>48</td>
<td>CD data analysed using K2d</td>
</tr>
<tr>
<td>T-MSP</td>
<td>23</td>
<td>33</td>
<td>44</td>
<td>FTIR (Loll et al. 2005a)</td>
</tr>
<tr>
<td>S-MSP</td>
<td>8</td>
<td>33</td>
<td>59</td>
<td>CD (Shutova et al. 1997)</td>
</tr>
</tbody>
</table>

6.3 Oxygen evolution from *G. violaceus* MSP-reconstituted PSII

Reconstitution of spinach PSII with G-MSP produced by both refolding and chaperone co-expression slightly increased the rate of O$_2$ evolution above the control (MSP-depleted PSII) rate (Table 6-3). However this is similar to the increase seen upon addition of TRX (Chapter 3) and is probably due to non-specific interaction with the PSII intrinsic proteins (Table 6-3). Control O$_2$ production rates of PSII reconstituted with S-MSP and T-MSP (Table 6-3) are in agreement with those shown in Chapter 3.
Chapter 6. Binding and characterisation of MSP from *G. violaceus*

Table 6-3. Rates of O₂ evolution from spinach PSII reconstituted with MSP from *G. violaceus*, *T. elongatus* and spinach.

<table>
<thead>
<tr>
<th>Treatment of spinach PSII</th>
<th>Protein reconstituted with spinach PSII</th>
<th>O₂ evolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untreated</td>
<td></td>
<td>663 ± 100 mL O₂ (mg Chl)⁻¹ hour⁻¹</td>
</tr>
<tr>
<td>MSP depleted</td>
<td>S-MSP</td>
<td>285 ± 15 mL O₂ (mg Chl)⁻¹ hour⁻¹</td>
</tr>
<tr>
<td>MSP depleted</td>
<td>T-MSP</td>
<td>332 ± 50 mL O₂ (mg Chl)⁻¹ hour⁻¹</td>
</tr>
<tr>
<td>MSP depleted</td>
<td>G-MSP</td>
<td>179 ± 27 mL O₂ (mg Chl)⁻¹ hour⁻¹</td>
</tr>
<tr>
<td>MSP depleted</td>
<td>TRX</td>
<td>152 ± 23 mL O₂ (mg Chl)⁻¹ hour⁻¹</td>
</tr>
</tbody>
</table>

* percentage of the untreated sample with 5 mM CaCl₂.

6.4 Binding of *G. violaceus* MSP to spinach PSII intrinsic proteins

Analysis of the MSP binding to the PSII intrinsic proteins indicates that G-MSP is unable to bind strongly to spinach PSII. Figure 6-2, lane B shows that the urea extract from the G-MSP-reconstituted spinach PSII has no band corresponding to the MSP molecular weight. As a control, the urea extract from T-MSP-reconstituted PSII was also run, and shows the MSP band on the gel (Figure 6-2, A).

![Figure 6-2](image-url)

Figure 6-2. The ability of heterologous MSP to bind spinach PSII analysed by urea extraction of the reconstituted complex. The presence of a band indicates binding. A) MSP from *T. elongatus*. B) MSP from *G. violaceus*.

6.5 Analysis of *G. violaceus* MSP binding and implications for evolution

The G-MSP is probably unable to bind specifically to spinach PSII because the residues involved in the binding interface are poorly conserved. 15 residues have been identified in the *T. elongatus* crystal structure as being involved in charge pairs and hydrogen bonds with PSII (De Las Rivas and Barber 2004). Of these, 11 (73 %) are conserved in S-MSP, where as only seven (37 %) are in G-MSP. In addition to residues forming bonds there are residues which are in contact or close proximity with the PSII
intrinsic proteins (defined as ‘regions’ by De Las Rivas and Barber, and in Figure 6-3 and Figure 6-4) which surround the binding residues and will influence their conformation. Of these, 86% are conserved in the spinach MSP, but only 34% in the G-MSP. The easiest way to visualize these differences is in Figure 6-4, where a model has been simulated of the G-MSP, and compared to the crystal structure of *T. elongatus*. The view is looking down onto the face where the intrinsic PSII proteins would bind. In the T-MSP structure the residues involved in binding are coloured in blue. Where these are conserved in G-MSP, they are also blue, but where they are not they are coloured red. Residues in green in both structures are the ‘regions’ involved surrounding the binding amino acids, but are coloured orange in the G-MSP structure where they are not conserved.

![Figure 6-3](image)

**Figure 6-3.** Sequence alignment of MSP from *T. elongatus* (Therm), Spinach (Spin), and *G. violaceus* (Glo) showing residues thought to be involved in binding to PSII. Blue, residues that are involved in binding interactions such as hydrogen bonds or charge pairs with the PSII intrinsic proteins, and are conserved compared to the *T. elongatus* sequence. Red, residues that would be involved in binding interactions, but are not conserved with respect to the *T. elongatus* sequence. Green, and orange are in regions which are in contact or close proximity with the intrinsic proteins of PSII, but do not form bonds. Green indicates conservation with the *T. elongatus* sequence. Orange indicates non conserved residues.
Figure 6-4. 3D structure showing the PSII-binding surface of A) *T. elongatus* MSP from the 1SS5L crystal structure and B) *G. violaceus* MSP predicted by Swiss-Model. **Red**, residues that would be involved in binding interactions, but are not conserved with respect to the *T. elongatus* sequence. **Green**, and **orange** are in regions which are in contact or close proximity with the intrinsic proteins of PSII, but do not form bonds. **Green** indicates conservation with the *T. elongatus* sequence. **Orange** indicates non conserved residues. **Purple**, Portions of the protein which are not involved in binding.

In particular, G-MSP has very low sequence conservation at its N-terminus, and lacks both of the ‘TY’ sequences which mutagenesis experiments have shown to be essential for the MSP-PSII interaction in spinach (Eatonrye and Murata 1989; Popelkova et al. 2002a; Popelkova et al. 2002b). One ‘TY’ sequence is conserved in all other MSPs which have been sequenced, the other is only conserved in higher plants and has been proposed to be responsible for the binding of a second copy of MSP to PSII (Popelkova et al. 2002b).
Chapter 6. Binding and characterisation of MSP from *G. violaceus*

6.6 Discussion

G-MSP is an unusual MSP from an unusual cyanobacterium. *G. violaceus* lacks the internal thylakoid membranes of other oxygenic photoautotrophs, and instead has its photosynthetic machinery located in the cytoplasmic membrane (Gugliemi 1981; Koenig and Schmidt 1995). G-MSP has the lowest sequence conservation compared to the MSP of any other species, and has the shortest primary sequence of any species. The inability of G-MSP to bind specifically to spinach PSII intrinsic proteins is probably a biochemical indicator of the divergent nature of its photosynthetic apparatus, and in particular of its PSII proteins. In addition to the morphological and molecular evidence indicating that *G. violaceus* is evolutionarily isolated, the data presented here shows that the G-MSP is biochemically different and predict that the sequence variations do indeed translate into structural variations of the protein.

The first and most obvious indications of the divergent nature of *G. violaceus* were that initial morphological examination of the cells show it has no internal membrane structure or invaginations and probably retains the cellular structure of photosynthetic prokaryotes before evolution of thylakoids. Subsequent investigation revealed that the phycobilisomes of *G. violaceus* are different to those of other cyanobacteria (Bryant et al. 1981; Gugliemi 1981), and it lacks the fatty acid sulfoquinovosyl diacylglycerol which is thought to stabilize the Photosystems (Selstam and Campbell 1996). Molecular studies have added to the evidence that *G. violaceus* retains features of an ancient ancestor. Analysis of the GC content of the *G. violaceus* genome revealed a different composition to that of other cyanobacteria (Rippka 1974), and phylogenetic relationships based on small subunit RNA sequences placed *G. violaceus* at the deepest branching lineage of cyanobacteria (Turner et al. 1999). More recently the complete genome structure of *G. violaceus* has been determined and shows a lack of conservation in photosynthetic protein sequences, including the absence of the PsaI, PsaJ, PsaK, and PsaX subunits for Photosystem I (PSI) and PsbY, PsbZ and Psb27 subunits for Photosystem II (Nakamura et al. 2003). The sequence conservation in the genes encoding the PSII extrinsic subunits, *psbU psbV* and *psbO* (which encodes the MSP) are low compared to other cyanobacteria. The PSI complex of *G. violaceus* has been extensively studied due to the absence of the long-wavelength emission in the 77 K fluorescence (Koenig and Schmidt 1995; Mimuro et al. 2002). Functional membrane-bound PSI from *G. violaceus* has been isolated from the plasma membrane and analysed.
Chapter 6. Binding and characterisation of MSP from *G. violaceus*

by SDS-PAGE, immunoblotting, gel filtration and electron microscopy (Mangels et al. 2002). No such studies have been reported on the PSII of *G. violaceus*, and the data presented here represent the first experimental findings on any isolated component of the PSII from this organism.

The recombinant expression of G-MSP in *E. coli* will provide sufficient material to undertake further characterization of the structure and function of this protein, and address some of its other unusual features. In particular G-MSP lacks a number of residues which are proposed to form part of the proposed product/substrate channel in *T. elongatus* PSII. Most of these residues are conserved in the MSP from all other species. Because these residues are not conserved the MSP of *G. violaceus* is not expected to participate in the potential channel identified in *T. elongatus*, although it may form an alternate channel which has not yet been identified.
7  General discussion

7.1  Comparison of MSP from *T. elongatus* with MSP from Spinach

7.1.1  Differences in Structure

Low resolution 3-D structures of PSII from spinach and *C. reinhardtii* have been obtained by electron microscopy, and show that higher plants and algae have very similar PSII structures (Nield et al. 2000a). These studies also show that the intrinsic proteins of the plant and algal PSII are similar to *T. elongatus* PSII, but differences are seen in the placement of the extrinsic proteins (Nield et al. 2000a). The higher resolution structure of spinach PSII has been predicted based on the 3.5 Å crystal structure of *T. elongatus*, and this shows that the MSP is located in the same place but the small extrinsic proteins are positioned differently (Nield and Barber 2006). All species have MSP as a constituent of their PSII, and in most cases the MSP of one species can functionally bind to the PSII of another species, which indicates a high degree of similarity (Betts et al. 1994; Enami et al. 2000; Suzuki et al. 2005).

However a number of other experiments have shown differences in the function and structure between MSPs from different species. First, the deletion of the *psbO* gene from green algae and higher plants shows that the MSP is required to form a stable OEC (Mayfield et al. 1987; Yi et al. 2005). In contrast, when the *psbO* gene is deleted from cyanobacteria they can still evolve O₂ (Burnap et al. 1992; Burnap and Sherman 1991; Mayes et al. 1991; Philbrick et al. 1991). Second, depletion of MSP from PSII in thylakoid membranes slows the water exchange rate measured in spinach, but not the rate measured in *Synechocystis* (Hillier et al. 2001). Third, isolated MSPs from red algae, cyanobacteria, green algae and higher plants appear to have different solution structures based on their differing sensitivities to protease digestion (Tohri et al. 2002). Finally, earlier research which suggests that the S-MSP has a molten globular structure (Lydakis-Simantiris et al. 1999; Shutova et al. 2000; Zubrzycki et al. 1998) whereas T-MSP has a compactly folded core with some flexible loop regions (Nowaczyk et al. 2004) is in agreement with the findings of this thesis. Many of the experiments that define our understanding of the structure and function of MSP were carried out using either T-MSP or S-MSP and a summary of these is given in Table 7-1.
### Table 7-1. Comparison of experiments investigating MSP using either *T. elongatus* or spinach PSII.

<table>
<thead>
<tr>
<th>Experimental observation</th>
<th>Species</th>
<th>Intactness of the system</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Secondary structure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ca(^{2+}) binding induces structural changes</td>
<td>Spinach</td>
<td>Isolated (^1)</td>
<td>(Heredia and De Las Rivas 2003; Kruk et al. 2003)</td>
</tr>
<tr>
<td>Ca(^{2+}) has no affect on the structure</td>
<td><em>T. elongatus</em></td>
<td>Isolated</td>
<td>(Loll et al. 2005a)</td>
</tr>
<tr>
<td>Resolved a bound Ca(^{2+}) ion</td>
<td><em>T. elongatus</em></td>
<td>Crystal structure (^2)</td>
<td>(Murray and Barber 2006)</td>
</tr>
<tr>
<td>The interaction with Ca(^{2+}) is pH dependent</td>
<td>Spinach</td>
<td>Isolated</td>
<td>(Shutova et al. 2005)</td>
</tr>
<tr>
<td>Stable at low pH and has a hysteresis in the buffering capacity</td>
<td>Spinach</td>
<td>Isolated</td>
<td>(Shutova et al. 1997)</td>
</tr>
<tr>
<td>Predicted increase in amino acid pKa away from the OEC</td>
<td><em>T. elongatus</em></td>
<td>Crystal structure</td>
<td>(Ishikita et al. 2006a)</td>
</tr>
<tr>
<td>Thermostable and recovers secondary structure after heating</td>
<td>Spinach</td>
<td>Isolated</td>
<td>(Lydakis-Simantiris et al. 1999; Shutova et al. 2001; Shutova et al. 2003)</td>
</tr>
<tr>
<td>Thermostable and recovers secondary structure after heating</td>
<td><em>T. elongatus</em></td>
<td>Isolated</td>
<td>(Sonoyama et al. 1996)</td>
</tr>
<tr>
<td>The disulfide bond has no essential for function</td>
<td>Spinach</td>
<td>Membrane fragments (^3)</td>
<td>(Betts et al. 1996a)</td>
</tr>
<tr>
<td>Secondary structural changes upon binding PSII</td>
<td>Spinach</td>
<td>Membrane fragments</td>
<td>(Hutchison et al. 1998b)</td>
</tr>
<tr>
<td><strong>Tertiary structure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Molten globular structure</td>
<td>Spinach</td>
<td>Isolated</td>
<td></td>
</tr>
<tr>
<td>Not a molten globule</td>
<td><em>T. elongatus</em></td>
<td>Isolated</td>
<td>(Lydakis-Simantiris et al. 1999; Shutova et al. 2000; Zubrzycki et al. 1998)</td>
</tr>
<tr>
<td><strong>PSII binding dynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-linked MSP can rebind PSII</td>
<td>Spinach</td>
<td>Membrane fragments</td>
<td>(Enami et al. 1998a)</td>
</tr>
<tr>
<td>Extensive mutagenesis of amino acids involved in binding to PSII</td>
<td>Spinach</td>
<td>Membrane fragments</td>
<td>See Table 1-3 for details and references</td>
</tr>
<tr>
<td>Extensive mutagenesis of amino acids involved in binding to PSII</td>
<td><em>T. elongatus</em></td>
<td>Membrane fragments</td>
<td>See Table 1-3 for details and references</td>
</tr>
<tr>
<td>Characterisation of the MSP-PSII binding interface</td>
<td><em>T. elongatus</em></td>
<td>Crystal structure</td>
<td>(De Las Rivas and Barber 2004)</td>
</tr>
<tr>
<td>Release from PSII is stimulated by hydrolysis of GTP</td>
<td>Spinach</td>
<td>Membrane fragments</td>
<td>(Lundin et al. 2007)</td>
</tr>
<tr>
<td><strong>MSP function</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-linked MSP cannot restore O(_2) evolving activity</td>
<td>Spinach</td>
<td>Membrane fragments</td>
<td>(Enami et al. 1998a)</td>
</tr>
</tbody>
</table>
Predicted a product or substrate channel or pool in PSII involving MSP

<table>
<thead>
<tr>
<th>Predicted a product or substrate channel or pool in PSII involving MSP</th>
<th>T. elongatus</th>
<th>Crystal structure</th>
<th>See Figure 1-10 for details and references</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP plays a role in PSII turnover</td>
<td>Spinach</td>
<td>Membrane fragments</td>
<td>(Yamamoto et al. 1998)</td>
</tr>
<tr>
<td>The absence of MSP alters water accessibility</td>
<td>Spinach</td>
<td>Thylakoids</td>
<td>(Hillier et al. 2001)</td>
</tr>
</tbody>
</table>

1 Isolated= MSP not bound to PSII
2 Crystal structure= PSII-bound MSP in either the 3.5 Å or 3.0 Å crystal structures
3 Membrane fragments= PSII-bound MSP in PSII-enriched membrane fragments
4 Thylakoids= PSII-bound MSP in whole thylakoid membranes

The DSC experiments presented in this work reveal that T-MSP unfolds cooperatively with a typical two-state like transition (Figure 4-6), whereas S-MSP does not (Figure 4-1). The increase in ANS fluorescence measured in the presence of T-MSP and S-MSP (Figure 4-5) indicates that S-MSP has a more accessible hydrophobic core than T-MSP; particularly at low pH. These results indicate that S-MSP has structure with more molten globular character than T-MSP when both are in solution. The ITC experiments measuring the binding of MSP to PSII also indicate that S-MSP undergoes a larger degree of structural rearrangement than T-MSP during the binding reaction (Table 5-2). This suggests that S-MSP adopts a more compact tertiary structure when it binds to PSII. Taken together these experiments show that isolated S-MSP and T-MSP do not have the same structure, and that isolated S-MSP and PSII-bound S-MSP probably do not either. These findings are important to consider because all structural information is based on T-MSP, and no high-resolution structural information of plant MSP exists. The DSC and ITC studies suggest that these structures can not be assumed to be the same.

It is expected that proteins from spinach and T. elongatus will have different thermal stabilities, as these two organisms inhabit niches with quite different environmental conditions. Spinach is a multicellular, terrestrial, mesophilic eukaryote that has evolved to tolerate temperatures in the range of 10 – 30 °C. T. elongatus is a thermophilic single celled cyanobacterium which was isolated from a hot spring with a temperature of 57 °C. The optimum temperature for growth rate and whole-cell photosynthesis coincides with the environmental temperature, and has an upper limit of 65 °C (Yamaoka et al. 1978). The photosynthetic activity of isolated T. elongatus thylakoid membranes is optimal at 50 °C with an upper limit of 60 °C. Below 40 °C the photosynthetic rate is less than half that of the optimal rate which shows that the
photosynthetic apparatus itself is adapted to these higher temperatures (Yamaoka et al. 1978).

### 7.1.2 The basis of protein thermostability

*Adaptations to thermophily in other proteins.* Extremophilic microorganisms can survive temperatures as low as -40 °C and as high as 115 °C, as well as high concentrations of salt, extreme pHs and high pressure. Psychrophiles are cold adapted organisms which have optimal growth temperatures between 0 and 20 °C, thermophiles have optima between 50 and 80 °C, and hyperthermophiles have optima above 80 °C. For organisms to grow at these temperatures their cellular proteins must be adapted to function correctly under these extreme conditions. Reviews of proteins from extremophiles include (Jaenicke and Bohm 1998; Luke et al. 2007; Vieille and Zeikus 2001).

Protein conformational flexibility and dynamics are important for enzyme function, and the extent of this flexibility depends on the temperature the enzyme is operating at. For this reason the structural stability and dynamics of proteins are usually optimised to allow them to function in the temperature range the organism lives at (Fields (2001) and references therein). If the temperature is too high the protein structure will be too loose to bind the substrates and carry out the reaction, however if the temperature is too low catalysis will occur too slowly to be useful (Collins et al. 2003). At extremely high temperatures the protein will completely denature. A consequence of the need to balance structural integrity with conformational flexibility is that the rate of catalysis thermostable enzymes is generally decreased at lower temperatures because they lack the necessary flexibility to function properly. Plots of the temperature dependence of enzyme catalysed reactions generally show an increase in reaction rate up to an optimum, followed by a rapid decrease when the protein structure looses integrity. The affinity of an enzyme for its substrate generally decreases with increasing temperature as the areas of the protein structure needed to bind and orient the substrates can no longer maintain the correct geometry (Fields 2001).

Insights into temperature dependent protein stability have been gained by comparing the structures of enzymes from mesophilic, thermophilic and psychrophilic organisms which carry out the same function and are derived from the same ancestral
form. Usually the overall structures of the proteins are conserved and the active site residues remain identical (Fields 2001). Cold stable enzymes generally have increased flexibility allowing greater thermal motion to carry out their reactions (Collins et al. 2003). Heat stable enzymes usually are more rigid, and have a decreased surface area and increased compactness (Bell et al. 2002).

There are many different mechanisms which increase the stability of proteins from thermophiles. Thermostable proteins often have an increased number of intramolecular ion pair bonds compared to mesostable proteins. These are particularly important when they occur between distantly placed residues in the linear polypeptide chain (Bae and Phillips (2004, 2005), Karshikoff and Ladenstein (2001) and references there in). Many thermostable proteins also have charged residues at their surfaces which are involved in inter-subunit interactions or increase the proteins’ solubility (Martin et al. 2001). An increased number of disulfide bonds and Pro residues relative to mesostable enzymes both increase thermostability by constraining the conformation of the protein and making its unfolded form less favourable (Boutz et al. 2007) (Vieille and Zeikus 2001) and references there in). Conversely, Gly residues have the most degrees of conformational freedom of any amino acid in the polypeptide back-bone and are expected to encourage flexibility more than other residues. Replacement of Gly with other amino acids has been observed in artificially evolved thermostable proteins (Brouns et al. 2005), but is not always seen in natural thermostable proteins, and some thermostable proteins have higher numbers of glycine than their mesostable counterparts (Ibrahim and Pattabhi 2004). An overall increase in weak interactions has also been observed, particularly an increase in hydrogen bonds (Ibrahim and Pattabhi 2004).

Secondary structural adaptations include a higher propensity of β-sheets in thermostable proteins compared to other types of secondary structure (Fields 2001). α-helical secondary structure is stabilised in several ways; (i) by helix capping where the side chain of the first residue forms a hydrogen bond with the main-chain of the third residue, (ii) by increasing the dipole of the helix, (iii) by having a proline in the first position on the helix which fits the first turn of the helix (iv) by having additional intra-helix ion pairs, and having ion pairs between helices (Fields 2001; Hashimoto et al. 1999). Loop regions are often reduced in size, and are sometimes stabilised by ion pairs (Arnott et al. 2000).
The protein tertiary structure plays a large role in determining thermostability, and changes include increased hydrophobic interactions between domains, an increase in packing density, and a reduction in the size and number of internal cavities ( Arnott et al. 2000; Bell et al. 2002; Chen and Stites 2004; Martin et al. 2001). Intra-subunit ionic networks are also common ( Arnott et al. 2000; Bogin et al. 2002). Many thermostable proteins are also multimeric, and have increased hydrophobic and ionic interactions between the subunits ( Arnott et al. 2000; Sambongi et al. 2002). The consequence of this for the amino acid composition is that thermostable proteins often have fewer uncharged polar residues compared to mesostable proteins. These are replaced by non polar and charged amino acids which are involved in inter and intra subunit interactions. There is an increase in Pro content and increased residue volume which may increase packing in the core ( Fields et al. 2001).

Because of the variety of strategies that have evolved to adapt proteins to function at high temperatures, it is not possible to predict how the substitution of one amino acid for another in a linear primary sequence will affect the overall thermal stability of the protein ( Fields (2001) and references there in). Analysis of proteomes and sequences can identify trends in the amino acid composition. But stabilisation comes from the interaction of a large number of relatively small individual changes which are difficult to distinguish from mutations from genetic drift between species. Elucidation of the stabilising mechanisms in native proteins has mostly come from comparison of 3D crystal structures of proteins adapted to different temperatures in combination with other biophysical techniques ( Bae and Phillips 2004; Bell et al. 2002; Boutz et al. 2007; Collins et al. 2003; Sambongi et al. 2002). Site directed mutagenesis and attempts to engineer stabilising interactions have provided some insight into the role of individual residues in stabilisation, but this method can prove difficult as altering single amino acids can have unintentional deleterious effects on the overall protein structure ( Bae and Phillips 2005; Bogin et al. 2002). A more successful strategy has involved engineering chimeric proteins by ‘transplanting’ whole portions of secondary structure or domains from a thermostable protein to a mesostable protein, or vice versa and evaluating the stability of the chimera ( Arnott et al. 2000; Chen and Stites 2004). Directed evolution of thermostable proteins from mesostable precursors has been used primarily to develop thermostable proteins for biotechnological applications, but also provides useful insights into forces which stabilise naturally occurring thermostable
What does this mean for the MSP? The amino acid composition of T-MSP compared to other mesophilic cyanobacteria is remarkably similar (Table 7-2). The most noticeable changes are fewer Gly (2.4% fewer), and more Thr (1.4% more). There are also more Ala and Gln. Of these differences, the lower glycine content is most likely to be significant as glycine has the smallest side chain of any amino acid and therefore the greatest conformational flexibility. These residues are highlighted in the sequence alignment in Figure 7-1. Residues marked with * are unique to T-MSP and not functionally conserved. Residues marked with a + are unique to T-MSP, but are functionally conserved with other species. Residues marked with – are where a conserved or partially conserved glycine (occurring in two or more other species) is replaced with something else. The 14 non-conserved glycines are distributed evenly between structure and unstructured regions. These glycine are replaced with Ala (4), Thr (4) and Ser (3) which are all still small, one Asn and two charged residues Asp and Gln. Replacement of Gly with Ser and Val conferred stability to mesostable proteins in directed evolution experiments (Brouns et al. 2005), so it is likely that at least some of these differences relate to the thermostability of T-MSP. Of the unique residues which are not functionally conserved, nine are found in secondary structural regions compared to five in loops. The secondary structural regions of T-MSP have four unique hydrophobic residues, two Ile, a Leu, and a Val. There are also three unique charged residues; Glu, Lys and His as well as Gln, Arg and Gly. A previous sequence alignment of T-MSP with MSP from cyanobacteria and plants indicated several unique amino acids in T-MSP that were suggested to be responsible for its thermostability (Miura et al. 1993). Several of these match the analysis here, and those which are different are due to exclusion of higher plant sequences in the current analysis, and inclusion of more cyanobacterial sequences. The residues from the previous analysis which are not unique in this alignment are: Gln46 and Leu118 which are found in several of the species here, and Met110, Ala111, Ile116 and Leu220 that are in this alignment. Other thermophilic cyanobacteria include species closely related to T. elongatus, including some that can survive temperatures up to 74 °C (Steunou et al. 2006). Several filamentous thermophilic species have also been characterised recently, and appear to represent a distinct lineage (Finsinger et al. 2008); however, the sequences for the MSP
from these are not available for comparison. Although it is interesting to speculate on the structural basis of the thermostability of T-MSP, a direct comparison with a high resolution structure from a mesophilic organism is needed to draw firmer conclusions.

**Table 7-2**: Comparison of the amino acid composition between the MSP from mesophilic cyanobacteria, and the MSP from *T. elongatus*. Values given are for the average percentage composition of that amino acid. Residues highlighted in bold have a 1% or greater difference between *T. elongatus* and mesophilic species.

<table>
<thead>
<tr>
<th></th>
<th>Mesophilic MSP *</th>
<th>T-MSP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala (A)</td>
<td>6.9</td>
<td>8.1</td>
</tr>
<tr>
<td>Arg (R)</td>
<td>3.9</td>
<td>4.5</td>
</tr>
<tr>
<td>Asn (N)</td>
<td>3.6</td>
<td>3.7</td>
</tr>
<tr>
<td>Asp (D)</td>
<td>7.3</td>
<td>6.5</td>
</tr>
<tr>
<td>Cys (C)</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>Gln (Q)</td>
<td>4.1</td>
<td>5.3</td>
</tr>
<tr>
<td>Glu (E)</td>
<td>7.7</td>
<td>7.7</td>
</tr>
<tr>
<td>Gly (G)</td>
<td>9.7</td>
<td>7.3</td>
</tr>
<tr>
<td>His (H)</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>Ile (I)</td>
<td>5</td>
<td>4.9</td>
</tr>
<tr>
<td>Leu (L)</td>
<td>7.2</td>
<td>8.1</td>
</tr>
<tr>
<td>Lys (K)</td>
<td>5.9</td>
<td>6.1</td>
</tr>
<tr>
<td>Met (M)</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>Phe (F)</td>
<td>5.2</td>
<td>4.5</td>
</tr>
<tr>
<td>Pro (P)</td>
<td>5.5</td>
<td>5.3</td>
</tr>
<tr>
<td>Ser (S)</td>
<td>7.4</td>
<td>6.5</td>
</tr>
<tr>
<td>Thr (T)</td>
<td>8.5</td>
<td>9.8</td>
</tr>
<tr>
<td>Trp (W)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tyr (Y)</td>
<td>3.1</td>
<td>2.4</td>
</tr>
<tr>
<td>Val (V)</td>
<td>7.4</td>
<td>6.9</td>
</tr>
</tbody>
</table>

Figure 7-1. Sequence alignment of *T. elongatus* MSP (highlighted in yellow) with the MSP from eight mesophilic cyanobacteria. Ala are highlighted in light blue, Gln in red, Gly in pink, and Thr in green. Residues marked with * are unique to T-MSP and not functionally conserved. Residues marked with a + are unique to T-MSP, but are functionally conserved in other species. Residues marked with – are where a conserved or partially conserved glycine (occurring in two or more other species) is replaced with something else. The secondary structural elements of T-MSP from (De Las Rivas and Barber 2004) are shown above the sequence, α- for the α-helices, β- for the β-strands. The mesophilic cyanobacteria are: Nost (*Nostoc* sp. PCC 7120) Cyano (*Cyanothece*, strain ATCC 51142), Croc (*Crocospheara watsonii*) Micro (*Microcystis aeruginosa* PCC 7806), Tric (*Trichodesmium erythraeum*, strain IMS101), Acac (*Acaryochloris marina* MBIC11017), Proc (*Prochlorococcus marinus* subsp. marinus str. CCMP1375).
It is important to note that MSP from thermophiles performs as well as MSP from mesophiles in the reconstitution reactions carried out at 25 °C in Chapter 3 when considering the similarities and differences between them. Cross-reconstitution experiments reviewed in Table 1-4 of Chapter 1 included reconstitution of PSII with MSP from both thermophilic and mesophilic organisms. These results are summarised in Table 7-3, where the PSII intrinsic proteins and the MSP are classified based on whether they are thermophilic or mesophilic and shows that thermophilic MSP performs as well in reconstituting O₂ evolving activity at 25 °C as mesophilic MSP. This suggests that there is no negative effect from the thermophilic MSP being more rigid than the mesophilic one on the water splitting reaction. By contrast the O₂ evolving activity of photosynthetic membranes from *T. elongatus* is extremely temperature dependent and very low at 25 °C (Yamaoka et al. 1978).

Table 7-3. Rates of O₂ activity restored to extrinsic-depleted PSII when cross-reconstituted with MSP.

<table>
<thead>
<tr>
<th></th>
<th>Thermophilic MSP *</th>
<th>Mesophilic MSP †</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermophilic PSII</td>
<td>45 %</td>
<td>46 %</td>
</tr>
<tr>
<td>Mesophilic PSII</td>
<td>47 %</td>
<td>53 %</td>
</tr>
</tbody>
</table>

* S. *vulcanus*, S. *elongatus*, C. *Calderarium* and *P. laminosum*
† *Spinach*, *A. thaliana*, and *C. reinhardtii*

The observation that *T. elongatus* PSII has low O₂ evolving activity at 25 °C, but that T-MSP can restore normal rates of O₂ evolving activity in reconstituted systems at 25 °C suggests that the high thermal stability of T-MSP is just one of several stabilising adaptations of *T. elongatus* PSII. It is likely that the *T. elongatus* intrinsic proteins are optimised for high temperatures and exhibit decreased flexibility which impairs their function at lower temperatures. By contrast the function the individual T-MSP subunit does not appear to be adversely affected by a decrease in temperature. When bound to spinach PSII at 25 °C T-MSP appears to have similar enough structure and sufficient flexibility to function as well as S-MSP in restoring O₂ evolution. This could be because both S-MSP and T-MSP are in the same conformation when bound to PSII, which is supported by the experiments here and previously published data (Pueyo et al. 2002).

The major difference between S-MSP and T-MSP found in this work is that T-MSP has a pre-formed tertiary structure while S-MSP does not. However it seems likely that S-MSP adopts the same structure as T-MSP when it binds to PSII. This leads
to the following hypothesis, summarised in Figure 7-2: S-MSP lacks a tertiary structure in solution, and does not exhibit a cooperative unfolding transition measured by DSC, as shown in Figure 4-1. At 25 °C, S-MSP binds to the spinach PSII-intrinsic proteins and S-MSP undergoes a structural transition to form a compact tertiary structure (Figure 7-2 A). This transition contributes to the favourable enthalpy change and the unfavourable entropy change measured by ITC during binding (Table 5-2). At 60 °C, S-MSP cannot assume the compact tertiary conformation because the entropy is too unfavourable, and so S-MSP will not bind to spinach PSII and the complex cannot evolve oxygen (Figure 7-2 B). T-MSP has a compact tertiary structure that unfolds cooperatively during DSC experiments with a T_m of 75 °C (Figure 4-6). At 25 °C T-MSP can bind to T. elongatus PSII- however this complex cannot evolve O_2 at high rates as the T. elongatus PSII-intrinsic proteins do not have sufficient flexibility to allow catalysis to proceed (Figure 7-2 C). At 60 °C, T-MSP can bind to T. elongatus PSII, and the complex can evolve oxygen (Figure 7-2 D). The reason that S-MSP cannot bind to PSII at 60 °C, whereas T-MSP can, is that S-MSP must undergo a structural transition to bind and this is unfavourable at high temperatures. T-MSP however has much the same structure in its bound and free states, and therefore a change in structure is not a prerequisite for its functional binding.
Figure 7-2. Hypotheses for the basis of thermolability of spinach PSII (green cartoon top two panels), and thermostability of T. elongatus PSII (blue cartoon, bottom two panels). In all cases MSP bound to PSII is in equilibrium with free MSP. For T-MSP the bound and free forms are the same, but for S-MSP the free form is molten globular. At 25 °C the equilibrium favours the bound MSP-PSII complex, and spinach PSII is active. But because of the rigidity of the T. elongatus intrinsic PSII proteins the complex is inactive. At 60 °C S-MSP is mostly in the unbound form as the molten globular state is favoured. T-MSP does not need to undergo any conformational change to bind to PSII, so binding is still favourable at this temperature. The PSII intrinsic proteins are more flexible at this temperature, and the complex is active.

The significant implications of this hypothesis are two fold. Firstly it suggests that models of S-MSP based on the bound structure of T-MSP are valid so long as only the bound form of S-MSP is being considered. Secondly it raises the question; if T-MSP can function normally at 25 °C despite having a compact tertiary structure, what advantage does the molten globular structure have for S-MSP?

7.2 The importance of flexibility for MSP function

7.2.1 Summary of flexibility from these experiments

The results of the experiments presented in this thesis reveal a number of conditions that alter the flexibility of MSP. The unfavourable entropy change measured by ITC upon MSP binding to PSII suggests that it decreases in flexibility during this interaction. This change is greater in S-MSP than T-MSP (Table 5-2) and might be the formation of
tertiary structure in S-MSP. DSC studies show that addition of Ca\(^{2+}\) destabilises T-MSP (Figure 4-18), and indicate that Ca\(^{2+}\) can bind to an intermediate state with lower thermal stability (Table 4-9). Fusion of *E. coli* TRX to the N-terminus of T-MSP also slightly decreased the T\(_m\) (Table 4-6) despite the fusion protein restoring higher rates of activity than T-MSP or S-MSP (Table 3-1). Finally, the T\(_m\) of T-MSP is increased in the presence of sarcosine, but the ΔH\(_{cal}\) decreased (Table 4-10). This is hypothesised to arise from stabilisation of one part of the T-MSP structure, and conversion of another part to a more molten globular form, similar to that of S-MSP. The ANS binding data support this hypothesis, showing that T-MSP induces greater fluorescence in ANS when sarcosine is added (Figure 4-11). These experiments add to the large body of evidence that shows MSP is a highly flexible protein, and the extent of this flexibility can be modulated by a number of physiologically relevant conditions.

**7.2.2 Examples of other highly flexible proteins and their functions**

Natively unfolded or intrinsically disordered proteins have non-rigid structures under physiological conditions, and are biologically active. These can be divided into two groups: (*i*) natively unfolded proteins with a molten globular structure which have defined regions of secondary structure and a hydrophobic core; and (*ii*) natively unfolded proteins with properties of a random coil which have an extended structure in solution and have little or no secondary structure. The structure of MSP is considered to belong to the former group based on its high secondary structure measured by CD and FTIR, and its hydrophobic core measured by intrinsic fluorescence (Shutova et al. (2000) and references therein). For this reason random-coil like natively unfolded proteins will not be discussed here, but are reviewed in (Tompa 2002; Uversky 2002a, 2002b).

The molten globular protein conformation was first discovered as an intermediate state in protein folding en route to the final compact globular conformation typical of many proteins. Early experiments studied the equilibrium intermediate which can be stabilised by mild denaturing conditions, and is equivalent to the kinetic folding intermediate (Ptitsyn 1995a; Ptitsyn et al. 1995; Ptitsyn and Uversky 1994; Uversky et al. 1992). Recently it has been recognised that some proteins also have a molten globular conformation under physiological conditions. These natively molten globular
proteins have the same properties as the folding intermediate conformation, but are in their final folded state and are biologically active (Dunker et al. 2001).

Molten globular proteins have a number of defining physical properties: (i) a lack of rigid tertiary structure and more flexibility than the native form (ii) a high content of secondary structure (iii) a hydrophobic core and more compact than the denatured form. The structure of molten globular and natively unfolded proteins is reviewed in the following references (Privalov 1996; Ptitsyn 1996; Ptitsyn 1995b; Ptitsyn et al. 1995; Ptitsyn and Uversky 1994). The secondary structural content of molten globular proteins is apparent in their high ellipticity measured by UV CD, and the decreased tertiary structure by near UV CD (Bychkova et al. 1992; Bychkova et al. 1996; Kamiyama et al. 1999; Uversky et al. 1992). Their high degree of flexibility means molten globular proteins do not crystallise but can be measured by NMR, albeit with broader spectral peaks (Croy et al. 2004; Dunker et al. 2001). Flexibility also imparts a high susceptibility to protease digestion, and high rate of amide proton exchange (Croy et al. 2004; Dunker et al. 2001). Molten globular proteins also have an increased size due to looser packing. This manifests as a shorter retention time during size exclusion experiments, and an abnormally large radius measured by small-angle x-ray scattering (Dunker et al. 2001; Uversky et al. 1999; Uversky et al. 1992). Finally as discussed earlier in this work, binding of ANS is higher in molten globular proteins than compactly folded ones (Semisotnov et al. 1991), and their thermal unfolding is less cooperative and sometimes lacks an unfolding peak measured by DSC (Bychkova et al. 1992; Croy et al. 2004; Kamiyama et al. 1999; Poklar et al. 1997; Uversky et al. 1992), reviewed in detail in (Freire 1995b).

Folding or mis-folding intermediates. Many of the molten globular proteins studied are either intermediate states that occur during the folding of compact globular proteins, or are mis-folded conformations both of which are stabilised by the solvent conditions. Common solvent conditions that stabilise molten globular conformations include mild acidic pHs and moderate denaturant concentrations which partially unfold the protein. Osmolytes have also been used to force the protein to partially fold under denaturing conditions. Some chemically induced molten globular states of proteins are listed in Table 7-4.
Table 7-4. Proteins in chemically induced molten globular states.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Protein structure</th>
<th>Conditions</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-lactalbumin</td>
<td>α-helical with a β-turn</td>
<td>Low pH (3.6 and 2 respectively)</td>
<td>(Yutani et al. 1992a)</td>
</tr>
<tr>
<td>apo Retinol Binding protein</td>
<td>antiparallel β-sheet</td>
<td>Low pH (below 3)</td>
<td>(Bychkova et al. 1992)</td>
</tr>
<tr>
<td>β–lactamase</td>
<td>Central sheet and α-helices</td>
<td>pH 7, 0.8M Guanidine</td>
<td>(Semisotnov et al. 1991)</td>
</tr>
<tr>
<td>Cytochrome c</td>
<td>α-helical</td>
<td>Methanol concentrations above 40%</td>
<td>(Bychkova et al. 1996)</td>
</tr>
<tr>
<td>Cytochrome c</td>
<td>α-helical</td>
<td>Added sorbitol and NaCl to acid denatured protein</td>
<td>(Kamiyama et al. 1999)</td>
</tr>
<tr>
<td>Carbonic anhydrase B</td>
<td>β-sheet with some helix and loops</td>
<td>Low pH (3.6 and 2 respectively)</td>
<td>(Semisotnov et al. 1991; Uversky et al. 1992)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pH 3.6, pH 8, 1.9M Guanidine</td>
<td></td>
</tr>
<tr>
<td>Equinatoxin II</td>
<td>Tetramer of β–sandwich flanked by α-helices</td>
<td>Decreased pH</td>
<td>(Poklar et al. 1997)</td>
</tr>
</tbody>
</table>

The relevance of these chemically induced states for normal protein structure and function is a matter of debate. Alcohol-induced intermediates are thought to be a consequence of local protein folding induced by the solvent, rather than representing a real folding intermediate (Bhakuni 1998). In some proteins formation of the molten globule state at low pH is thought to mimic the conditions near the membrane surface. The negative electrostatic potential of the membrane attracts protons from the bulk solvent, creating a local pH decrease. This is of biological importance for proteins which need to undergo native to molten globule transitions to insert into membranes, or to release ligands at the membrane surface (Bychkova et al. 1996).

Natively molten globular proteins. Molten globular proteins occur naturally as folding intermediates when they are stabilised by chaperonin proteins such as groEL, while other chaperonins and ATP aid formation of the correct tertiary structure (Martin et al. 1991). Naturally occurring molten globular proteins also have biological functions, many of which involve transitions between the molten and fully globular states (Dunker et al. 2002). Some examples of natively molten globular proteins are given in Table 7-5, and shows these proteins have a variety of secondary structural elements and are found in both prokaryotes and eukaryotes.
In many cases molten globular proteins become fully structured upon binding to a ligand. An example is IκBα which is involved in protein-protein interactions in a signalling pathway. IκBα appears to have a molten globular structure in its native state, but becomes more structured upon binding to its target (Croy et al. 2004). One advantage of a protein undergoing a disorder to order transition upon binding to a ligand is that it imparts high specificity but low affinity. Conversely, if the protein retains a high degree of flexibility once it is bound to the ligand it can exhibit promiscuous binding which is useful when the protein functions to bind multiple partners. This appears to be the case for clusterin, a secreted protein which probably binds and clears cell debris. Clusterin has large molten globular regions which are thought to be involved in binding (Bailey et al. 2001).

In some cases a transition between the globular holo-form and the molten apo-form allows controlled release of the ligand. An example of this is the structure-to-disordered transition of retinol-binding protein when it releases retinol A. This is thought to be in response to the increase in pH at the membrane surface, and functions to release retinol A to the target cells (Bychkova et al. 1992). Another example where proteins may become molten is during translocation across, or insertion into a membrane. Often the intermediate has a transit peptide which increases their molten globular character (Bychkova et al. 1988). This is the case of many proteins involved in toxin secretion. Some of these kill cells by forming pores that allow ion diffusion and de-polarise the membrane (Baldwin et al. 1996; Kristan et al. 2004; Merrill et al. 1990), others are chaperones that transport soluble toxins into the cell (Faudry et al. 2007; Housden et al. 2005).
Table 7-5. Examples of molten globular proteins with biological functions.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Structure</th>
<th>Function</th>
<th>Location</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retinol binding protein</td>
<td>(\beta)-barrel</td>
<td>Transports retinol A to cell surface receptors</td>
<td>Cytoplasmic, but functions near the cell membrane in eukaryotes</td>
<td>(Bychkova et al. 1992)</td>
</tr>
<tr>
<td>(\beta)-lactamase</td>
<td>Central sheet and (\alpha)-helices</td>
<td>Degrades antibiotics</td>
<td>Periplasmic space of bacteria</td>
<td>(Minsky et al. 1986)</td>
</tr>
<tr>
<td>Colicin E1</td>
<td>(\beta)-hairpin</td>
<td>Forms an ion-permeable pore</td>
<td>Cytoplasmic membrane, of target bacteria but must also traverse the outer membrane and periplasm of the secreting bacteria</td>
<td>(Merrill et al. 1990)</td>
</tr>
<tr>
<td>Tumor necrosis factor</td>
<td>(\beta)-structured trimer</td>
<td>Ion channel</td>
<td>Plasma membrane of target eukarotic cell.</td>
<td>(Baldwin et al. 1996)</td>
</tr>
<tr>
<td>Equinotoxin</td>
<td>Tetramer of (\beta)-sandwich flanked by (\alpha)-helices</td>
<td>Pore forming toxin</td>
<td>Plasma membrane of target cells</td>
<td>(Kristan et al. 2004)</td>
</tr>
</tbody>
</table>

7.2.3 Hypotheses of why MSP is flexible

*MSP interaction with PSII and release during photoinhibition.* In vivo a pool of unassembled MSP exists in the lumen that can bind to the intrinsic proteins upon assembly of the PSII complex. The disordered structure of MSP has been suggested to play a role in its resistance to aggregation, increasing its stability in the lumen (Wyman and Yocum 2005a). Flexibility may also be a requirement for the MSP to bind to PSII, since structural studies using FTIR and small angle x-ray scattering indicate that S-MSP undergoes structural changes upon binding (Hutchison et al. 1998b; Sachs et al. 2003; Svensson et al. 2004). NMR studies suggest that the loop regions of isolated T-MSP are less constrained solution than in the crystal structure (Nowaczyk et al. 2004). The ITC data presented in Chapter 5 indicates that upon binding to the PSII intrinsic proteins, both T-MSP and S-MSP undergo structural transitions which restrict their conformations, although the magnitude of this restriction is greater in the case of S-MSP. This indicates that MSP flexibility is necessary for binding of this subunit to the rest of the PSII complex.

The molten globular structure of isolated S-MSP might have a specific role in the release of MSP during photoinhibition and D1 turnover. So far in this work it has
been assumed that the presence of tertiary structure in T-MSP and lack of it in S-MSP is a consequence of one being from a thermophilic organism, and the other being from a mesophile. An alternate explanation could lie with the recently characterised GTP-dependent regulation of S-MSP during PSII turnover in plants (Lundin et al. 2007). This relies on GTP binding and hydrolysis, probably by a sequence which is unique to the MSP of higher plants, which results in the release of S-MSP from the PSII complex. The structure of MSP has low homology with the structure of other GTPases suggesting that S-MSP represents a novel plant-specific class of GTPase. The putative GTP-binding sequence is only conserved in higher plants, suggesting that this mode of regulation is not present in the PSII of other species (Lundin et al. 2007). The transition in S-MSP between the molten globular isolated form and the compactly folded bound form could be part of a mechanism regulating MSP binding to PSII. If the molten globular isolated form of S-MSP is stabilised by binding to GDP, this would present a mechanism for how GDP stimulates release of MSP from PSII. If this is the case, the lack of a cooperative transition in S-MSP (Figure 4-1) and the presence of a cooperative transition in T-MSP (Figure 4-6) might be due to a higher plant/cyanobacteria difference rather then a mesophile/thermophile difference. This could be tested by seeing if mesophilic cyanobacteria are molten globular or not; if the GTP mechanism is specific to higher plants, and if the molten globular state of S-MSP is a consequence of this mechanism mesostable cyanobacterial MSP is also expected to have tertiary structure as well.

**Structural changes during the S-state cycle.** The MSP binds near the [Mn₄Ca] cluster of PSII, and there are structural changes in the cluster during the S-cycle which could be transmitted through the proteins structure (Haumann et al. 2005a; Kulik et al. 2007). There is evidence that the MSP also changes conformation during the S-cycle: (i) isotopic labelling FTIR data shows that Asn Gln and Glu residues are perturbed during the S₁- S₂ transition (Sachs et al. 2003), (ii) the protease sensitivity of MSP changes depending on the oxidation state of the [Mn₄Ca] cluster (Hong et al. 2001), (iii) a reduction in the flexibility of MSP by cross-linking inhibits restoration of O₂ evolution to PSII samples, even though the cross-linked protein is able to rebind normally (Enami et al. 1998a). An increase in MSP flexibility has also been observed in response to pH and metal ions *in vitro*, and this is proposed *in vivo* to convert MSP from a closed to an
open active conformation during water oxidation, in response to a drop in pH in the lumen (Shutova et al. 2005).

Possible evolutionary relic. One group of proteins which commonly undergo transitions between molten globular and structured conformations are proteins associates with membranes. As described in detail in Section 7.3, it is possible that MSP evolved from a β-barrel protein in the outer bacterial membrane. This precursor protein could have been molten globular in solution, and the molten globular structure of MSP could just be a carry over from its previous function. It has been suggested that some functional proteins are natively disordered simply because many amino acid sequences will form disordered structures and that in the absence of a selective pressure to evolve into a globular form, they have remained unstructured (Williams et al. 2001)(Dunker et al. 2001). The same reasoning could apply to the MSP; if the MSP precursor was molten globular then it would remain that way unless there was a pressure to change. Adaptation to a thermophilic existence would represent such a pressure which would explain why T-MSP is not a molten globule.

7.3 \textit{G. violaceus} MSP as a ‘missing link’

\textit{G. violaceus} is a primitive contemporary cyanobacterium, which is believed to have diverged earliest during the evolutionary radiation of cyanobacteria (Nakamura et al. 2003). \textit{G. violaceus} exhibits a number of interesting features as a consequence of its position in the evolutionary tree. One of these features is that it lacks thylakoids. Rather than being inserted in the thylakoid membrane, the protein complexes of the light reactions of photosynthesis are inserted in the plasma membrane with the ‘lumenal’ portions oriented to the outside of the cell in the periplasmic space (Gugliemi 1981; Rippka 1974). Given that oxygenic photosynthesis evolved only once, \textit{G. violaceus} is a good candidate to examine ancestral PSII.

In Chapter 6, it was shown that recombinantly expressed G-MSP was unable to bind to the spinach PSII intrinsic proteins. The finding that the G-MSP is so divergent that it cannot bind to spinach PSII has important implications for the evolution of PSII. In their homology search of the MSP crystal structure from \textit{T. elongatus}, De Las Rivas and Barber (2004) found similarity to four proteins with a β-barrel fold. Three of the proteins, OmpX, streptavidin, PagP enzyme are found in the outer membranes of gram
negative bacteria, and the fourth, quinohemoprotein amine dehydrogenase (QHNDH) is found in the periplasmic space (see Koebnik et al. (2000), Sleytr et al. (2007) and Bishop (2005) for reviews). If, as predicted, the protein complexes of oxygenic photosynthesis arose in the plasma membrane of the cyanobacterial ancestor with the ‘lumenal’ components facing the periplasm, then the ‘lumenal’ portions of the intrinsic PSII proteins could have come into contact with soluble periplasmic proteins, or outer membrane proteins prior to their insertion into the outer membrane. These could have served as a starting point for the evolution of the extrinsic proteins, such as MSP from these β-barrel precursors. De Las Rivas and Barber (2004) note that structural homology was found only with the barrel portion of the proteins, and that there was no homology with the regions of interaction with PSII (De Las Rivas and Barber 2004). If the G-MSP does retain features of an intermediate stage in evolution from one of these proteins we would expect its binding site with PSII to show deviation from the MSP of more ‘advanced’ species. This prediction is consistent with our data indicating that the G-MSP binding site is sufficiently different that it cannot be interchanged with PSII intrinsic proteins from another species as most other MSP proteins can, and supports the idea that *G. violaceus* represents a missing link in the evolution of oxygenic photosynthesis.

### 7.3.1 The cyanobacterial periplasm and outer membrane

The cyanobacterium *G. violaceus*, lacks internal thylakoid membranes, and instead its PSII complex is embedded in the plasma membrane. This means that the extrinsic proteins, including the G-MSP must function in the periplasmic space outside the cytoplasmic membrane. Cyanobacterial cell walls have features of both gram negative and gram positive bacteria. Outside the plasma membrane is a layer of peptidoglycan which forms the cell wall; a characteristic of gram-positive bacteria. Outside the peptidoglycan layer, cyanobacteria have an outer membrane that defines the periplasmic space- a characteristic of gram negative strains. Outside this is the outer membrane (Figure 7-3) (Hoiczyk and Hansel 2000).
The outer membrane of gram negative bacteria is described as ‘leaky’ due to the presence of a large number of β-barrel porin proteins that allow passive diffusion of solutes into the periplasm (Nikaido 1992). Outer membrane preparations have been isolated from *Synechocystis* and the protein content has been analysed (Huang et al. 2004). This revealed several proteins with homology to plant chloroplast proteins, and many that are found in non-photosynthetic gram negative bacteria. Many of these are involved in transport, and most have a β-barrel fold.

The outer membranes of cyanobacteria lack some common porin proteins found in gram negative bacteria and instead have the cyanobacteria specific SomA and SomB porin proteins (Hoiczyk and Hansel 2000). Porins all have similar structures, but poor sequence conservation between species; particularly in the loop regions which are variable (Hansel et al. 1998). SomA and SomB have been characterised and have 14 or 16 stand β-barrels which form the pore, and an N-terminal domain that extends into the periplasm and possibly connects the outer membrane and the peptidoglycan layer (Hansel et al. 1998; Hansel and Tadros 1998). SomA and SomB are larger proteins than bacterial porins, but have smaller conductances. OMP85 is a homologue of the higher plant TOC 75 (Reumann et al. 1999). TOC 75 forms a channel protein that is found in the chloroplast outer membrane of plants, and is involved in transport of nuclear encoded proteins into the chloroplast. In cyanobacteria OMP85 it forms an ion...
permeable channel, but the exact function is not known. GumB is a β-barrel protein that secretes polysaccharides in gram negative bacteria, and is localised to the outer membrane in *Synechocystis* (Huang et al. 2004). A homologue of the TolC efflux pump was identified in *Synechocystis*. This pump exports molecules from the cytoplasm to the cell exterior, by-passing the periplasm. In *E. coli* the structure of TolC shows it is a trimer with β-barrels that span the cytoplasmic and outer membranes, and an α-helical barrel crossing the periplasm (Huang et al. 2004). Some cyanobacteria also have active ABC (ATP-Binding Cassette) transporters proteins in their outer membranes, which are involved in nitrogen transport (Hoiczyk and Hansel 2000).

In addition to transport proteins, the protease/chaperone proteins HhoA and HtrA are associated with the periplasm and outer membrane respectively. HtrA is a homologue of DegP2, the protease involved in degradation of D1 in higher plants under photoinhibitory conditions, and could have a similar function in cyanobacteria (Huang et al. 2004).

### 7.3.2 The origins of MSP - evolution from a pore?

About 50% of the outer membrane of gram negative bacteria is made up of proteins. These include integral membrane proteins and lipoproteins attached to the membrane by their N-termini (Koebnik et al. 2000). The α-helical bundle transmembrane proteins are found in the cytoplasmic membrane, whereas β-barrel transmembrane proteins are generally found in the outer membrane (Koebnik et al. 2000).

Early predictions of the structure of MSP suggested homology with porin proteins (Kamiya and Shen 2003); however, the high resolution structures suggest greater similarity with the virulence factors OmpX and PagP, and with the biotin binding protein streptavidin which are all found in the outer membrane of gram negative bacteria, and with QHNDH, found in the periplasmic, space (De Las Rivas and Barber 2004). OmpX is an 8 stranded β-barrel which is elliptical in cross section, and has a large loop region extending to the cell exterior (Vogt and Schulz 1999). The crystal structure of OmpX was solved from *E. coli*, but its function has mostly been studied in the pathogenic bacteria *Yersinia enterocolitica* and *Salmonella typhimurium* (Heffernan et al. 1994). OmpX is a virulence factor that neutralises host cell defences, and the loop region might be involved in adhesion to host cells. PagP is an 8 stranded β-barrel with flexible loop regions connecting the β-strands (Hwang et al. 2002). PagP is found in the
outer membrane of pathogenic bacteria, and functions to transfer a palmitate chain onto the outer membrane lipid A. This modification protects the bacterium from the host immune response (Bishop 2005). QHNDH is a heterodimer. The $\alpha$-subunit, which has some homology with MSP (De Las Rivas and Barber 2004), has four domains i.e. domain 1 has two sub-domains with $c$-type cytochrome folds, domains 2, 3 and 4 are $\beta$-barrels. The $\beta$ subunit has a $\beta$-propeller structure, and the $\gamma$-subunit is mostly unstructured (Datta et al. 2001). QHNDH catalyses the oxidative deamination of primary aliphatic and aromatic amines, and is produced in the periplasm of *Paracoccus denitrificans* when n-butylamine or benzylamine are provided as carbon sources (Takagi et al. 1999).

The structural similarity of MSP to outer membrane and periplasmic proteins raises the possibility MSP actually evolved from an outer membrane protein. Pore-forming transport proteins seem a good candidate as they are common in all outer membranes. In this hypothetical scenario, which is summarised in Figure 7-4, the PSII precursor intrinsic proteins in the ancestral cell were localised in the cytoplasmic membrane. The portions of the protein which are oriented towards the lumen and that bind the [Mn$_4$Ca] cluster in modern organisms would have protruded into the periplasmic space and faced the outer membrane. The PSII precursor had a ‘naked’ 4Mn cluster that was ligated by the intrinsic proteins, but not shielded from the periplasm by extrinsic proteins. The outer membrane of modern cyanobacteria and gram negative bacteria includes many $\beta$-barrel proteins involved in transport. Some of these proteins are observed undergo a globular-to-molten transition upon insertion into the membrane (Bychkova et al. 1988). The ancestral cell would also have had transporters proteins in its outer membrane, and these probably inserted by the same method. This presents a mechanism for how the MSP precursor might have become associated with the PSII precursor complex. It is possible that some of these $\beta$-barrel outer membrane proteins bound non-specifically to the PSII precursor intrinsic proteins while in their molten globular transition state during insertion into the membrane. If this interaction proved beneficial to PSII precursor function it would have been selected for, and eventually evolved into the specific interaction observed today. Genetic drift and neutral mutations that preserved the structure and function of the MSP precursor would have accumulated over time, causing the primary sequence to diverge from that of the original outer-membrane protein. *G. violaceus* diverged, preserving some ancient features including
location of PSII in the cytoplasmic membrane. Other oxygenic photosynthetic organisms evolved thylakoids and the MSP diverged further, with changes to its binding site, and evolution of the product/substrate channel. Higher plants diverged, and GTPase activity evolved in higher plants.

Figure 7-4. Hypothetical scenario for evolution of MSP from a β-barrel outer membrane protein. The ancestral cell had an organisation similar to that of *G. violaceus* with the PSII-precursor complex embedded in the cytoplasmic membrane. Before evolution of the extrinsic proteins, the naked 4Mn cluster ([Mn₄Ca] precursor) would be exposed to the periplasm, which has a ‘leaky’ connection to the cell exterior through the porin proteins. (A) Porin protein precursors are transported into the periplasm through the outer membrane. This mechanism could involve removal of a transit peptide. (B) The porin adopts a molten globular conformation in the periplasmic space. (C) The molten globular porin inserts into the outer membrane in the normal manner. (D) The molten globular porin protein forms non-specific interactions with the naked 4Mn cluster. This provides a more stable hydrophobic environment for the 4Mn cluster, and increases PSII function. Variants which mis-traffic porin proteins are highly selected for and a specific binding site develops. Genetic drift and neutral mutations preserving the structure and function accumulate over time, and the primary sequence loses homology with the porin protein. Adaptation of the modern substrate/product channels increases the rate of water splitting. GTPase activity evolves in higher plants.

Recently Raymond and Blankenship (2008) used a new algorithm to compare the geometry and binding site of the [Mn₄Ca] cluster of PSII with the binuclear Mn cluster of Mn catalase (Raymond and Blankenship 2008). One of the most interesting findings is that the protein bulk from the Mn catalase overlaps with the protein bulk of the extrinsic subunits in the *T. elongatus* PSII. Raymond and Blankenship (2008) note
that this is ‘consistent with the idea of a periplasmic protein that could donate electrons to a primitive photosystem’. Mn catalase is found in the periplasm of a number of bacteria, and functions to protect the cell from oxidative damage (Alyamani et al. 2007; Klotz and Hutcheson 1992; Pacello et al. 2008; Short and Blakemore 1989). In this paper Raymond and Blankenship (2008) hypothesized that the four-Mn cluster evolved from the two-Mn enzyme first by providing ligands to the two Mn ions, which allowed it to function independently from the bound Mn catalase, and second by continuing to bind Mn catalase, so that two Mn ions were bound intrinsically, and two extrinsically. It follows from this that eventually all four ions would become ligated entirely by the intrinsic proteins. Once this happened, the Mn binding site would no longer be available as a point of interaction between the PSII-precursor intrinsic proteins and the Mn catalase ‘subunit’, and the affinity between the two would decrease. The absence of the Mn catalase ‘subunit’ would leave a naked four-Mn cluster. At this point, binding of other protein subunits that sequestered the cluster in a hydrophobic environment would be favourable, even if such binding was non-specific. This model is consistent with the observation that the protein bulk of the extrinsic proteins and Mn-catalase overlap in the Raymond and Blankenship model. It is also consistent with the experimental observation in Chapter 3 that introduction of protein bulk in the form of E. coli TRX served to restore activity in the absence of the small extrinsic proteins, even if the interaction was non-specific. The higher similarity of MSP to virulence factors compared to MSP and porins does not preclude pore-forming proteins being the precursor of MSP. Many membrane proteins have yet to be structurally characterized and these could include pore proteins with high homology to MSP. This is particularly true if the nearest homologue is also highly flexible like MSP, in which case it may not be able to form crystals for study, and will not appear on a similarity structure search.

To summarise there are four lines of evidence that suggest MSP could have evolved from a pore. First the location: the PSII precursor very likely existed in the plasma membrane with parts of the intrinsic proteins facing the outer membrane which contains many β-barrel proteins that could have been precursors. Second the structure: the barrel portion of MSP has homology with outer membrane and periplasmic proteins from bacteria. The loop regions of such barrels are often divergent in bacteria (Koebnik et al. 2000). The loop regions of G-MSP are also variable, shown in Chapter 6. Third the function: pore proteins are involved in solute transport. MSP is putatively involved
Chapter 7. General discussion

in water and/or proton transport. Fourth a mechanism: many β-barrel pore proteins are known to proceed via a molten globular intermediate during export through the cytoplasmic membrane, and insertion into the outer membrane. It is possible such a partially folded state could interact with and be stabilised by the PSII precursor intrinsic en route to the outer membrane. S-MSP is a molten globule in its non-PSII bound state in solution.

7.4 Conclusions and future directions

The findings presented here have extended the current knowledge of the structure and function of the MSP subunit of PSII. They also add to the ever increasing pool of information about natively molten globular proteins and their dynamics. The principles highlighted in this work are: (i) The importance of flexibility in the structure of MSP, and the likelihood that a transition from molten globular to regular globular structure occurs upon interaction with PSII in some species (ii) The differences in the MSP between thermophilic and mesophilic organisms, including a possible mechanism for the thermostability of PSII in *T. elongatus* (iii) The arrangement of cellular components in *G. violaceus* and the inability of its MSP protein to bind PSII from other organisms suggests a mechanism for evolution of MSP from an outer membrane protein. These ideas are important for two areas of active research into the water splitting capability of PSII: investigation of the proposed product/substrate channel involving parts of MSP and artificial photosthes.

Investigation of a product/substrate channel. Recent crystal structures of PSII from *T. elongatus* have identified potential channels for entry of the substrate water and exit of the products O₂ and H⁺. Two different sets of channels have been proposed (Ho and Styring 2007; Murray and Barber 2007), both of which involve different residues from MSP. An obvious strategy for investigating the existence of such a channel is to mutate these residues and examine the effect of this on water exchange rates in PSII. These mutations could be made in *Synechocystis* which would allow assembly of PSII, including the mutated MSP, to occur in the cell, before extraction of the thylakoid membranes. Alternatively mutated MSP could be expressed recombinantly and reconstituted into MSP-depleted PSII complexes after PSII has been extracted from the cell. Investigations using the latter method are currently being undertaken. Another
potential strategy is to use G-MSP as a ‘natural mutant’, since it lacks the conserved residues proposed to form the MSP-portion of the channel. Although G-MSP cannot bind to spinach PSII, it may have enough homology to cross reconstitute with cyanobacterial PSII proteins.

Artificial photosynthesis. In nature one of the by-products of oxygenic photosynthesis, O₂, allowed evolution of aerobic organisms and arguably shaped life on earth as it exists today. Now amid increasing concerns about climate changes caused by burning of fossil fuels, some researchers are looking again to the by-products of oxygenic photosynthesis for a solution. If protons from water oxidation can be captured and reduced to H₂, this could provide an alternative source of energy. Burning of H₂ is both clean, emitting water as its end product and renewable, and using only water and energy from the sun to produce (Levin et al. 2004). Some researchers are focusing on H₂ production from whole cells; particularly Chlamydomonas reinhardtii, unicellular green algae which produces photosynthetically derived H₂ under nutrient depleted conditions (Hankamer et al. 2007; Melis 2007; Melis et al. 2007). Another approach is to engineer systems that mimic the water splitting capability of PSII. Such systems would have more versatility than a cell-based method and might avoid the problems associated with low H₂ production rates and scale-up of bioreactors. A molecular water splitting machine will almost certainly have to be protein-based to tune the electron transport properties of the redox active cofactors and ensure directionality of electron transfer. The protein system will need to be far simpler than the >20 polypeptide chains that comprise PSII. It will also need to bind light activated pigments similar to chlorophyll, include a metal site for storage of charge equivalents, and include Tyr side chains to transfer electrons between the two. It will also probably be heterodimeric. A promising system has used bacterioferritin, a homodimeric α-helical Fe-binding protein, and it to carry out light-driven electron transfer (Wydrzynski et al. 2007). It is hoped that this system can be developed to transfer electrons from the metal centre to a bound quinine, and ultimate be used as the basis for an H₂-evolving bio-catalyst (Wydrzynski et al. 2007). The properties of the natural system investigated in the work described here will need to be take into account when engineering artificial photosystems. Any artificial water splitting apparatus will need to deal with substrate delivery and product removal, and flexibility of the protein matrix will likely play an important role in the whole system.
References


Barber J (2006b) Structure of Photosystem II and its implications for understanding the water splitting reaction it catalyses. BBA-Bioenergetics 36-36.


quantitative assembly into Photosystem II and for high rates of oxygen evolution activity. *Biochemistry* 37, 14230-14236.

Betts S D, Ross, J R, Hall, K U, Pichersky, E, and Yocum, C F (1996a) Functional reconstitution of Photosystem II with recombinant manganese-stabilizing proteins containing mutations that remove the disulfide bridge. *BBA-Bioenergetics* 1274, 135-142.


Betts S D, Ross, J R, Pichersky, E, and Yocum, C F (1997b) Mutation Val235ala weakens binding of the 33-kDa manganese stabilizing protein of Photosystem II to one of two sites. *Biochemistry* 36, 4047-4053.


Dau H, and Haumann, M (2005) Considerations on the mechanism of photosynthetic water oxidation - dual role of oxo-bridges between Mn ions in (i) redox-
potential maintenance and (ii) proton abstraction from substrate water. 

*Photosynth Res* 84, 325-331.


Frankel L K, Cruz, J A, and Bricker, T M (1999) Carboxylate groups on the manganese-stabilizing protein are required for its efficient binding to Photosystem II. Biochemistry 38, 14271-14278.


References

Photosystem II including water oxidation based on the structure of Photosystem II. *Philos T Roy Soc B* 357, 1337-1344.

Fukada H, and Takahashi, K (1998) Enthalpy and heat capacity changes for the proton dissociation of various buffer components in 0.1 m potassium chloride. *Proteins* 33, 159-166.


Ghanotakis D F, Babcock, G T, and Yocum, C F (1984b) Structural and catalytic properties of the oxygen-evolving complex - correlation of polypeptide and manganese release with the behavior of $\mathbf{z}^+$ in chloroplasts and a highly resolved preparation of the PS-II complex. *Biochim Biophys Acta* 765, 388-398.


**References**


Hutchison R S, Betts, S D, Yocum, C F, and Barry, B A (1998b) Conformational changes in the extrinsic manganese stabilizing protein can occur upon binding to the Photosystem II reaction center: An isotope editing and FTIR study. *Biochemistry* 37, 5643-5653.


Leuschnern C, and Bricker, T M (1996) Interaction of the 33 kDa extrinsic protein with Photosystem II: Rebinding of the 33 kDa extrinsic protein to Photosystem II membranes which contain four, two, or zero manganese per Photosystem II reaction center. *Biochemistry* 35, 4551-4557.


Miyao M, and Murata, N (1984a) Calcium-ions can be substituted for the 24-kDa polypeptide in photosynthetic oxygen evolution. Febs Lett 168, 118-120.


two genes encoding extrinsic 33-kDa proteins in Photosystem II. *Febs Lett* 523, 138-142.


Popelkova H, Im, M M, and Yocum, C F (2002b) N-terminal truncations of manganese stabilizing protein identify two amino acid sequences required for binding of the eukaryotic protein to Photosystem II and reveal the absence of one binding-related sequence in cyanobacteria. *Biochemistry* 41, 10038-10045.


References


van den Burg B, and Eijsink, V G H (2002) Selection of mutations for increased protein

Photosystem II the light-driven water: Plastoquinone oxidoreductase (T.

light-driven water: Plastoquinone oxidoreductase (T. Wydrzynski, and K. Satoh,

Structural basis for difference in heat capacity increments for Ca^{2+} binding to

*Biophys Chem* 115, 115.

Vieille C, and Zeikus, G J (2001) Hyperthermophilic enzymes: Sources, uses, and
molecular mechanisms for thermostability. *Microbiology And Molecular
Biology Reviews* 65, 1-1.

escherichia coli reveals possible mechanisms of virulence. *Structure With
Folding & Design* 7, 1301-1309.


polypeptide of the oxygen-evolving complex of Photosystem-II is a putative
calcium-binding protein and is encoded by a multi-gene family in pea. *Plant Mol

protein implicated in alzheimer's disease and learning, is natively unfolded.
*Biochemistry* 35, 13709-13715.

interactions between peptides and lipid membranes. *Peptide-Lipid Interactions*
52, 31-56.

Williams P D, Pollock, D D, and Goldstein, R A (2001) Evolution of functionality in
lattice proteins. *J Mol Graph Model* 19, 150-156.

Witt H T (2005). Photosystem II: Structural elements, the first 3D crystal structure and
References


