CRYSTALLOGRAPHIC AND NMR EVIDENCE FOR FLEXIBILITY IN OLIGOSACCHARYLTRANSFERASES AND ITS CATALYTIC SIGNIFICANCE

NYIRENDA, JAMES
Graduate School of Systems Life Science, Faculty of Science, Division of Structural Biology, Medical Institute of Bioregulation, Kyushu University

https://doi.org/10.15017/26441

出版情報：九州大学, 2012, 博士（システム生命科学）, 課程博士
バージョン：
権利関係：

CRYSTALLOGRAPHIC AND NMR EVIDENCE FOR FLEXIBILITY IN OLIGOSACCHARYLTRANSFERASES AND ITS CATALYTIC SIGNIFICANCE

by

JAMES NYIRENDA

MSc. Biochemistry, Hamdard University, New Delhi, India, 2003.
BSc. Chemistry, University of Zambia, Lusaka, Zambia, 2000.

AN ABSTRACT OF A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHYLOSOPHY (PhD.) IN SCIENCE.

Graduate School of Systems Life Science
Faculty of Science
Division of Structural Biology.
Medical Institute of Bioregulation,

KYUSHU UNIVERSITY
Kyushu, Fukuoka

2012

Signature of Author: _____

Date: _____

Abstract

Oligosaccharyltransferase (OST) is a membrane bound enzyme that catalyzes the transfer of an oligosaccharide to the asparagine residue in the sequon, Asn-X-Thr/Ser. Eukaryotic OST protein complex consists eight non identical subunits, and among them STT3 possesses the transferase activity. The equivalent to STT3 is a single subunit protein called PglB in Eubacteria, and AglB in Archaea. The primary sequences (600 to 1,000 residues) of the STT3/AglB/PglB proteins share a common architecture. The N-terminal part forms a multi-span transmembrane region and the C-terminal part forms a soluble, globular domain which contains a well-conserved, five-residue motif, WWDYG. Structural comparison of the crystal structures of the C-terminal globular domain of AglB from *Pyrococcus furiosus* and of PglB from *Campylobacter jejuni* revealed different conformations of the segment containing the WWDYG motif, raising a question about what was the true conformation of the segment without any crystal packing effects. As part of this research work, crystal structures of the C-terminal globular domain of AglB's from distant as well as close related organisms to *P.furiosus* were determined. Relevant to this study one close homolog, *Pyrococcus horikoshii* AglB, with sequence identity of about 70%, and one distant homolog, *Archaeoglobus fulgidus* AglB, with sequence identity of about 30% were selected. Comparison of the crystal structures with emphasis on the highly flexible region of the WWDYG motif was performed, and found a superimposable conformation of the WWDYG motif between the most distant pair: *A.fulgidus* AglB-S2 and *C.jejuni* PglB, even with a sequence overall similarity of less than 30%. ^{15}N NMR relaxation analysis studies to characterize the dynamic nature of OST using *A.fulgidus* AglB-S2 were performed. Intriguingly, the mobile region contains the binding pocket for +2 Ser/Thr residue in the N-glycosylation sequon. In agreement, the restriction of the flexibility forced by an engineered disulfide crosslink abolished the enzymatic activity, and its cleavage fully reversed the inactivation. These results suggest the multiple catalytic cycle and the essential involvement of a transient conformation in the reaction. It could be that the dynamic property of the Ser/Thr pocket facilitates the efficient scanning of N-glycosylation sequons along nascent polypeptide chains.

CRYSTALLOGRAPHIC AND NMR EVIDENCE FOR FLEXIBILITY IN OLIGOSACCHARYLTRANSFERASES AND ITS CATALYTIC SIGNIFICANCE

by

JAMES NYIRENDA

MSc. Biochemistry, Hamdard University, New Delhi, India, 2003.
BSc. Chemistry, University of Zambia, Lusaka, Zambia, 2000.

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHYLOSOPHY (PhD.) IN SCIENCE.

Graduate School of Systems Life Science
Faculty of Science
Division of Structural Biology
Medical Institute of Bioregulation,

KYUSHU UNIVERSITY
Kyushu, Fukuoka

2012

Supervised by:

Professor Daisuke Kohda (PhD.)

Table of Contents

List of Figures

List of Tables

Acknowledgements

I wish to greatly thank Professor Daisuke Kohda (PhD.) for accepting me as a research student in his laboratory and generously spending time advising, correcting and directing the path of this research in many ways. I wish to thank also the Graduate School of Systems Life Science of Kyushu University; Japan for enrolling me. Many thanks go to the Ministry of Science and Technology of the government of Japan (MEXT) for providing funding to carry out this research. My thanks also go to Dr. Takashi Saitoh, Dr. Akira Takano and Dr. Mayumi Igura for the valuable time they rendered in teaching and assisting me in data collection at the synchrotron as well as NMR machines and basic purification as well as crystallization protocols respectively. Oh! I still remember those days we would go and collect data over night at Tsukuba synchrotron facility. Many thanks go to the resource personnel at Tsukuba Photon Factory for allowing me to use the beam lines for x-ray diffraction studies. Many thanks also go to Dr. Nobuo Maita and Dr. Nobuo N. Noda for structure determination of *P.horikoshii* and *A.fulgidus* AglBs respectively. Thanks also go to Professor Fuyuhiko Inagaki (PhD.) for fruitful discussions and contributions to my research. I thank Dr. Yuzawa Satoru for going all the way to Spring8 synchrotron for preliminary data collection of *Pyrococcus abyssi*. Many thanks again go to the lab members for giving a hand here and there and providing support especially in the acquisition of reagents and also trying to do the interpretation from Japanese to English. I won't forget the young man Shunsuke Matsumoto, whom I fondly called Zendo, highly talented in recombinant DNA technology. I am grateful to my lab members Dr. Kouta Mayanagi, Dr. Atsushi Shimada, Mr. Daisuke Fujinami, Mr. Rei Matsuoka and the secretary Ms. Otsu Miki. Being married with three children, I wish to dedicate this thesis to my family and my parents Mr. and Mrs. Lot and Catherine Nyirenda. I know most of the time you wanted me to play with you, I would be found in the laboratory, sleep over and I have permanently missed those times you really wanted me to be by your side. All the same thank you for running this race by my side. My lovely wife Harriet for taking care of the children Khumbo-Mary, Chimwemwe-Gershom and Mapesho-Abigail and last but not the least, the Lord God Almighty Jehovah for giving me life, good health, strength, insight, wisdom and the patience when experiments yielded no reasonable data. Thank you all for being part of making this thesis.

Declaration

I, James Nyirenda being author of this thesis "Crystallographic and NMR Evidence for Flexibility in Oligosaccharyltransferases and its Catalytic Significance" for the award of Doctor of Philosophy (PhD.) in Science hereby declare that the work presented herein is a bonafide record of research work carried out by me under the supervision of Professor Daisuke Kohda (PhD.) and that to the best of my knowledge, no similar work has been reported before. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree.

Date…… Signed…….

James Nyirenda

This doctoral thesis has been examined by the following faculty of Kyushu University:

Prof. Yoshizumi Ishino (PhD.) _____

Prof. Mikita Suyama (PhD.) _____

Assoc. Prof. Kenji Inaba (PhD.) _____

Prof. Daisuke Kohda (PhD.) _____

Thesis supervisor

Abbreviations used in this thesis

AglB: Archaeal glycosylation B

ASH: Alignment of structural homologs

BLAST: Basic Local Alignment Search Tool

CBB: Coomassie Brilliant Blue

CNS: Crystallography and NMR System

COOT: Crystallographic Object-Oriented Toolkit

DTT: Dithiothreitol

EDTA: Ethylenediaminetetraacetic acid

GASH: Genetic-algorithm ASH

HEPES: 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid

HSQC: Heteronuclear single quantum coherence spectroscopy

MES: 2-(N-morpholino)ethanesulfonic acid

MOPS: 3-(N-morpholino)propanesulfonic acid

NMR: Nuclear Magnetic Resonance spectroscopy

NOE: Nuclear overhauser effect spectroscopy

OST: Oligosaccharyltransferase

PCR: Polymerase chain reaction

PDB: Protein data bank

PEG: Polyethylene glycol

PF BL-17A: Photon Factory Beam Line 17A (Tsukuba-Japan)

PglB: Protein glycosylation B (Eubacterial)

RMSD: Root mean square deviation

STT3: Staurosporine and temperature sensitive protein 3 (Eukaryotic)

TAE: Tris Acetate EDTA

TRIS: *Tris*(hydroxymethyl)aminomethane

Triton X-100: t-octylphenoxypolyethoxyethanol

Prior Publication

Much of this thesis work has been published in the paper (Nyirenda et al., 2012) and incorporates part of the work and contributions of various coauthors of the paper.

List of publications

1. James Nyirenda, Shunsuke Matsumoto, Takashi Saitoh, Nobuo Maita, Nobuo N. Noda, Fuyuhiko Inagaki, and Daisuke Kohda. Crystallographic and NMR Evidence for Flexibility in Oligosaccharyltransferases and its Catalytic Significance. Structure (2012), http://dx.doi.org/10.1016/j.str.2012.10.011

2. Shunsuke Matsumoto, Mayumi Igura, James Nyirenda, Masaki Matsumoto, Satoru Yuzawa, Nobuo Noda, Fuyuhiko Inagaki and Daisuke Kohda. Crystal Structure of the C-Terminal Globular Domain of Oligosaccharyltransferase from Archaeoglobus fulgidus at 1.75 A Resolution. Biochemistry **51**, 4157-66 (2012).

3. Maita Nobuo, Nyirenda James, Igura Mayumi, Kamishikiryo Jun & Kohda, Daisuke. Comparative structural biology of Eubacterial and Archaeal Oligosaccharyltransferases. *J Biol Chem* **285**, 4941-50 (2010).

Abstract

Oligosaccharyltransferase (OST) is a membrane bound enzyme that catalyzes the transfer of an oligosaccharide to the asparagine residue in the sequon, Asn-X-Thr/Ser. Eukaryotic OST protein complex consists eight non identical subunits, and among them STT3 possesses the transferase activity. The equivalent to STT3 is a single subunit protein called PglB in Eubacteria, and AglB in Archaea. The primary sequences (600 to 1,000 residues) of the STT3/AglB/PglB proteins share a common architecture. The N-terminal part forms a multi-span transmembrane region and the C-terminal part forms a soluble, globular domain which contains a well-conserved, five-residue motif, WWDYG. Structural comparison of the crystal structures of the C-terminal globular domain of AglB from *Pyrococcus furiosus* and of PglB from *Campylobacter jejuni* revealed different conformations of the segment containing the WWDYG motif, raising a question about what was the true conformation of the segment without any crystal packing effects. As part of this research work, crystal structures of the C-terminal globular domain of AglB's from distant as well as close related organisms to *P.furiosus* were determined. Relevant to this study one close homolog, *Pyrococcus horikoshii* AglB, with sequence identity of about 70%, and one distant homolog, *Archaeoglobus fulgidus* AglB, with sequence identity of about 30% were selected. Comparison of the crystal structures with emphasis on the highly flexible region of the WWDYG motif was performed, and found a superimposable conformation of the WWDYG motif between the most distant pair: *A.fulgidus* AglB-S2 and *C.jejuni* PglB, even with a sequence overall similarity of less than 30%. ^{15}N NMR relaxation analysis studies to characterize the dynamic nature of OST using *A.fulgidus* AglB-S2 were performed. Intriguingly, the mobile region contains the binding pocket for +2 Ser/Thr residue in the N-glycosylation sequon. In agreement, the restriction of the flexibility forced by an engineered disulfide crosslink abolished the enzymatic activity, and its cleavage fully reversed the inactivation. These results suggest the multiple catalytic cycle and the essential involvement of a transient conformation in the reaction. It could be that the dynamic property of the Ser/Thr pocket facilitates the efficient scanning of N-glycosylation sequons along nascent polypeptide chains.

Chapter 1 - Overview of N-linked glycosylation

Introduction

Asparagine linked (N-linked) glycosylation of proteins is a covalent modification that occurs in Eukarya, Archaea, and Eubacteria (Aebi et al., 2010). This essential post translation modification process of proteins has been widely studied in eukaryotes (Kelleher and Gilmore, 2006; Kelleher et al., 2003), using human and yeast (Knauer and Lehle, 1994; Lehle, 1992; Sharma et al., 1981; Tanner and Lehle, 1987), in Archaea (Abu-Qarn et al., 2007; Eichler, 2000; Igura and Kohda, 2011a; Igura and Kohda, 2011b; Konrad and Eichler, 2002; Yurist-Doutsch et al., 2010; Yurist-Doutsch et al., 2008), using *Haloferax volcanii* and *Pyrococcus furiosus*, and Eubacteria (Nothaft and Szymanski, 2010; Nothaft et al., 2010; Weerapana and Imperiali, 2006) and (Kowarik et al., 2006a), using *Campylobacter jejuni* as models. The enzyme that catalyzes the transfer of an oligosaccharide chain from the lipid-linked oligosaccharide donor to the asparagine residue in the consensus sequon, Asn-X-Thr/Ser (where X is any amino acid except Proline), is oligosaccharyltransferase (OST) (Karaoglu et al., 1997; Karaoglu et al., 1995; Kelleher and Gilmore, 1994; Kelleher and Gilmore, 2006; Kelleher et al., 2003). Eukaryotic OST is a protein complex consisting of eight non identical membrane protein subunits (Kelleher and Gilmore, 2006; Weerapana and Imperiali, 2006), and among them STT3 has been shown to possess the catalytic function (Igura and Kohda, 2011a; Igura and Kohda, 2011b; Igura et al., 2008; Kohda et al., 2007; Maita et al., 2010) and (Karaoglu et al., 1997; Kelleher and Gilmore, 2006; Nasab et al., 2008). The equivalent to STT3 is a single subunit protein called PglB (Larsen et al., 2004; Nothaft et al., 2010; Szymanski and Wren, 2005; Szymanski et al., 2002; Szymanski et al., 2003a; Szymanski et al., 2003b) in Eubacteria, and AglB (Igura et al., 2008) in Archaea.

Figure 1-1 shows a simplified cartoon rendering of asparagine or N-linked protein glycosylation across three domains of life namely Eukarya, Archaea and Eubacteria.

Figure 1-1 N-linked glycosylation across three domains of life

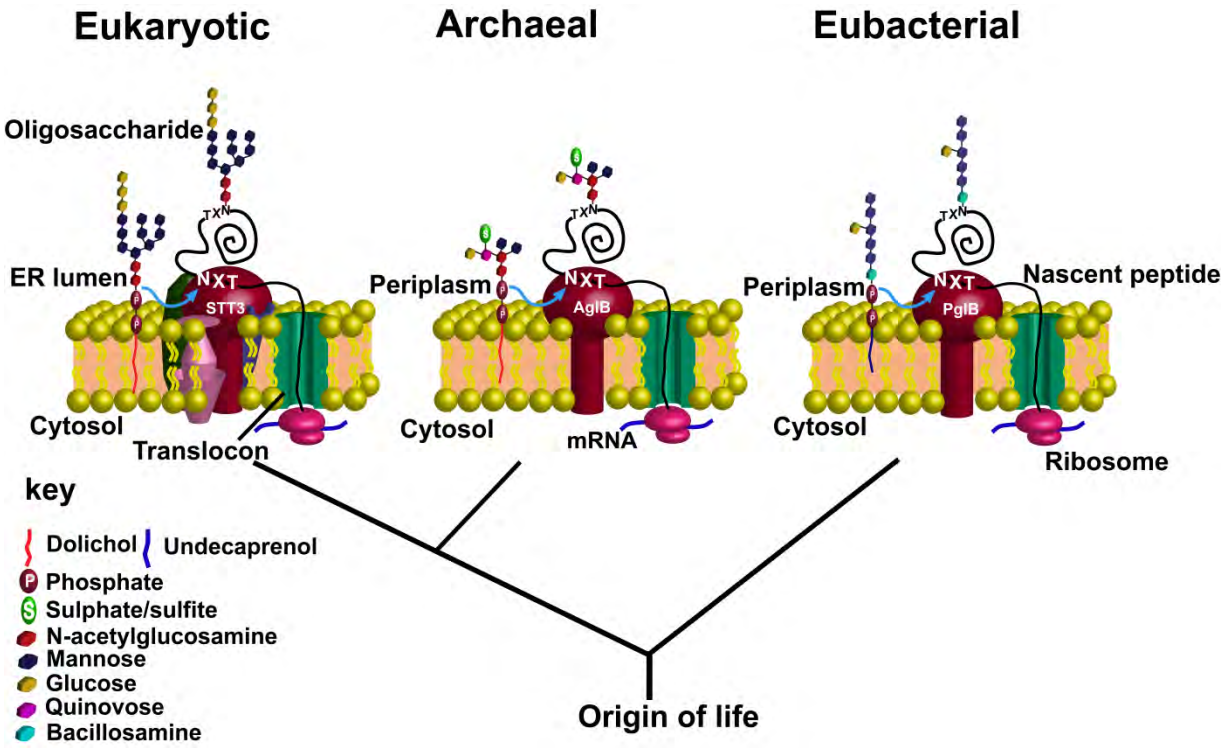

Figure 1-1 shows an artistic example of a typical glycosylation process. A nascent protein is glycosylated on the asparagine residue in the sequon N-X-T/S (where X is any amino acid except Proline) as it passes through the translocon near the OST complex. The enzyme responsible for transfer is STT3 in eukaryotes or the counter parts AglB in archaea or PglB in eubacteria. Eukaryotic OST occurs as a complex of eight non identical subunits while the archaeal and eubacterial OST's occur as single subunit membrane bound protein. A point worthy taking is that in eukaryotes, this process occurs in the lumen of the rough endoplasmic reticulum while in archaea and eubacteria, the process takes place in the periplasmic space.

The primary sequences (600 to 1,000 residues) of the STT3/AglB/PglB proteins share a common architecture. There is a multi-span transmembrane region in the N-terminal half of the primary sequence, and the C-terminal half forms a soluble, globular domain. Four short catalytic motifs were identified by alanine mutagenesis studies: A diacidic motif, D*X*D or E*X*D (*X* denotes any amino acid residue), is located in the first lumenal/extracellular loop of the N-terminal transmembrane region (Liu and Mushegian, 2003), and a second diacidic motif, G*XX*D*X*D or G*XX*D*X*E, has been recently identified in another loop of the same region(Igura and Kohda, 2011a). A well-conserved, five-residue motif, Tryptophan, Tryptophan, Aspartate, Tyrosine, Glycine (WWDYG), and another short motif, DK/DKi/MI, reside in the C-terminal globular domain (Igura et al., 2008; Maita et al., 2010; Matsumoto et al., 2012). The latter motif shows phylogenetically related variation (Maita et al., 2010); Eukaryotic STT3 proteins exclusively contain the DK motif, whereas Eubacterial PglB proteins only contain the MI motif. In contrast, Archaeal AglB proteins contain either the DK, DKi, or MI motif. Note that the DKi motif (a variant of the DK motif with a short loop *i*nsertion) was previously referred to as DM, but the structure determination of the AglB with the previously thought to possess the DM motif prompted the revision of the local sequence alignment, and renamed it as the DKi motif (Matsumoto et al., 2012).

The crystal structures of the C-terminal globular domain of AglB from *Pyrococcus furiosus* and that of PglB from *Campylobacter jejuni* were reported in the years 2008(Igura et al., 2008) and 2010(Maita et al., 2010), respectively . The first structure to be determined (*P.furiosus* AglB-L, L stands for *l*ong variant, arbitrally assigned as long and the other short since two sequences are in the database) raised the possibility that the unusual, *left-handed* 3_{10} helical conformation in the WWDYG motif was induced by the insertion of the side chain of a lysine 619 residue from another molecule in the crystal (Igura et al., 2007; Igura et al., 2008). This observation suggested a large plasticity of the segment containing the WWDYG motif (residues 491-584). This segment consists of 95 residues. Together with this segment, the N-terminal (6–7 residues) and the C-terminal (0–3 residues) were disordered in the native crystals. Then, the second structure of the *C.jejuni* PglB protein offered the opportunity for the structural comparison of the putative flexible segment. A close up view of the WWDYG motif segment comparison between *C.jejuni* PglB and *P.furiosus* AglB is shown in Figure 1-2.

Figure 1-2 WWDYG-motif segments of *C.jejuni* PglB and *P.furiosus* AglB respectively.

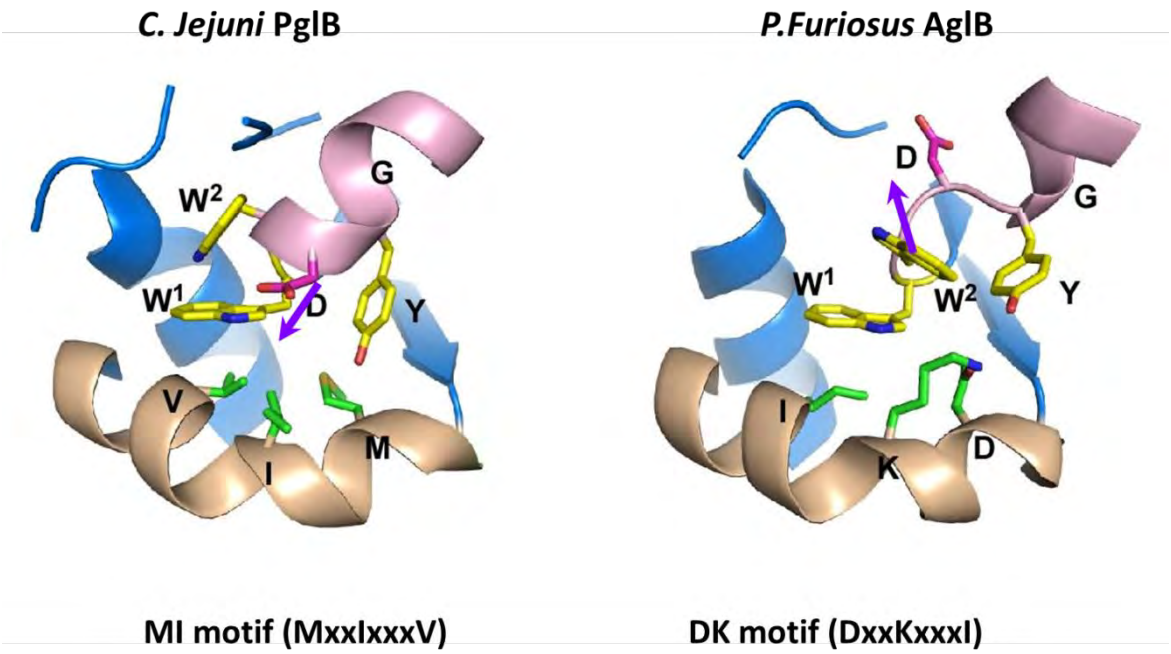

Figure 1-2 shows a Pymol cartoon rendition of part of the C-terminal globular domain. The consecutive five amino acid residue signature of all OST's, tryptophan, tryptophan, aspartate, glycine (WWDYG) are located in the N-terminal part of the pink helix. The brown kinked helix contains residues that form either the DK or the DKi or the MI motifs respectively. Tryptophan 1(W^1) has been shown to occupy the same relative position in all structures solved so far (Igura et al., 2008; Maita et al., 2010). The aspartate orients in different directions depicted by a purple arrow. Tryptophan number 2 orients in opposite directions in each structure.

The *C.jejuni* PglB structure (Figure 1-2) contained a more usual, right-handed 3_{10} helical conformation in the WWDYG motif region, leading to a different orientation of the α-helix after the WWDYG motif. The structural comparison confirmed the plasticity of the segment containing the WWDYG motif, and at the same time, raised a question about what was the true conformation of the segment without any crystal packing effects. This question is important because the WWDYG motif has been shown to be important in catalysis (Igura et al., 2008; Lizak et al., 2011; Yan and Lennarz, 2002), and its conformation is essential to understand the catalytic mechanism of the oligosaccharyltransferase. To address this question,

a comparative structural biology approach was adopted. One may expect that many orthologous proteins that share various extents of sequence identity will be crystallized in different crystal forms, and thus the deformation of very flexible segments will vary. If two structures are superimposable, the structure can be interpreted as a packing-free structure.

Here, the crystal structures of the C-terminal globular domain of AglBs from distant as well as close related organisms to *Pyrococcus furiosus* were determined. One close homolog, *Pyrococcus horikoshii* AglB-L, with sequence identity of about 70%, and one distant homolog, *Archaeoglobus fulgidus* AglB-S2, with sequence identity of about 30% were selected.

Working hypothesis

The diverse structural conformation presented by the two C-terminal domain structures of *P.furiosus* AglB-L(Igura et al., 2008) and *C.jejuni* PglB (Maita et al., 2010) provides a mobile or plastic region necessary for enzymatic activity.

Motivation

The unusual conformation of the WWDYG motif of *P.furiosus* OST probably due to the interaction with Lys619 (Figure 1-3) of another molecule in the crystal (Igura et al., 2007; Igura et al., 2008) as well as the different conformation of the WWDYG motif of *C.jejuni* OST(Maita et al., 2010) further suggested the presence of plasticity in the segment containing the WWDYG motif part of the active site. This observation instilled a lot of ambition in me and decided to test the hypothesis that this diverse conformation in the region containing the WWDYG motif is due to intrinsic flexibility of the segment and may be important to catalysis. Lysine 619 has been shown in Figure 1-3 with the side chain colored purple.

Figure 1-3 Distortion of the WWDYG motif by a protruding Lysine 619 residue

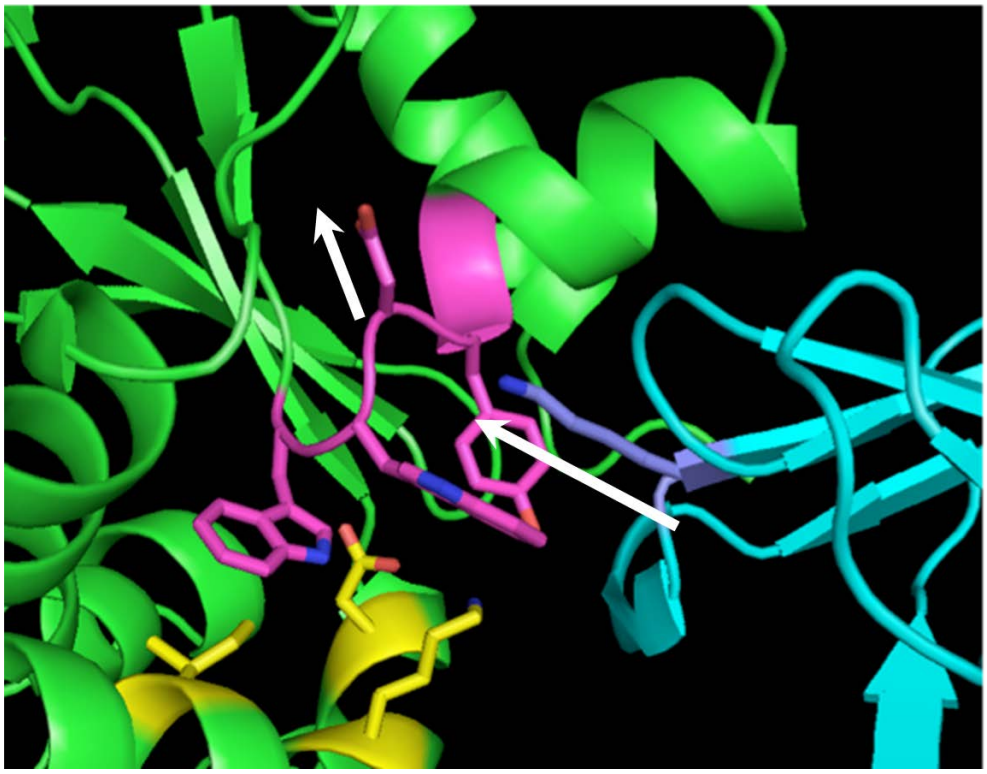

Figure 1-3 shows Lysine 619 (purple side chain) of one molecule (cyan) protruding into another molecule (green) thereby inducing a distortion of the WWDYG motif in the second molecule of the asymmetric unit (magenta). Here the aspartate side chain is moved in the opposite direction thereby distorting the WWDYG motif segment. This distorted structure is stabilized by the crystal packing observed in *P.furiosus* AglB-L crystals.(Igura et al., 2008)

Aims and objectives

The crystal structures of the globular domains of the OST's recently solved (*P.furiosus* AglB-L *and C.jejuni* PglB) showed a rather diverse overall native structural conformation. A mere amino acid (primary sequence) alignment over a diverse range of proteins rather produces results that are not in congruency with the overall native structures of proteins under investigation. A false identification of motifs can arise if a large sample size is not analyzed and more importantly if the fold is not even known.

In this research work, the major aim was to carry out a comparative study to assess the importance of the flexible or plastic region and make conclusions as to which structure resembles the thermodynamically stable state or one of the stable catalytic cycle states, free from crystal packing effects. To achieve this objective, x-ray crystallographic, NMR relaxation studies and engineered disulfide conformational restriction experiments were done on select candidate OST C-terminal globular domains as well as full length protein as models to help answer the hypothesis.

Rationale

To carry out this comparative study so that the aims and objectives of this research work were achieved, various constructs for the C-terminal globular domains of OST from archaea- *A.fulgidus and P.horikoshii* AglB's varying in both sequence identity and total polypeptide length, were made as select models in addition to the already solved structures of *P.furiosus* AglB-L and *C.jejuni* PglB. Spanning the two domains of life, Archaeal (*A.fulgidus, P.furiosus, P.horikoshii*)and Eubacterial (*C.jejuni*), respectively, the sample size as well as range was satisfied for further experimental works herein X-ray, NMR and disulfide bond restriction experimental works. The major question to be answered by these experiments was the importance of flexibility or plasticity in the segment encompassing the canonical and well preserved sequence of five consecutive amino acid residues namely Tryptophan-Tryptophan-Aspartate-Tyrosine-Glycine (WWDYG) herein called the WWDYG motif-segment that forms the Serine/Threonine binding pocket. Thus designed, the experiments were executed in like manner in order to help understand the variation in amino acid side chain orientation observed in crystal structures of *P.furiosus* AglB-L and *C.jejuni* PglB determined in earlier studies.

Chapter 2 - X-ray crystallographic studies

Expression, purification and X-ray crystallography experiments

To answer the first question in the hypothesis, the first step in this research work was to design the experiment to generate desired recombinant target protein sample for x-ray crystallographic work. This objective was achieved by the following experimental methods employed.

Experimental methods and materials used

The DNA encoding the *P.horikoshii* AglB (O74088_PYRHO), codon-optimized for *Escherichia coli* expression, was synthesized and cloned into the pUC57 plasmid by GenScript (NJ, USA). A PCR product encoding the C- terminal globular domain region (residues Ala482- Glu976) of *P.horikoshii* AglB was sub cloned between the Nde I and Sal I sites of the pET41b+ vector (Novagen, Merck) using the In-Fusion PCR cloning kit (Clonetech, Takara Bio). The Forward Primer PCR2for_PyHoNde 5'-AAGGAGATATACATATGGCTCTGAAAAACACCG-3' and the reverse primer PCRrev_PyRAHOHIS 5'-CCGCAAGCTTGTCGACTCATTAGTGATGGTGATGGTGATGATGGTG-3' were used for amplification. The reverse primer was designed to amplify all the ten Histidines in the C-terminal flanked by a double stop codon TCATTA thus avoiding usage of the pET41b+ plasmid derived His tags (Figure 2-1).

Figure 2-1 PCR amplification of *Pyrococcus horikoshii* AglB-L

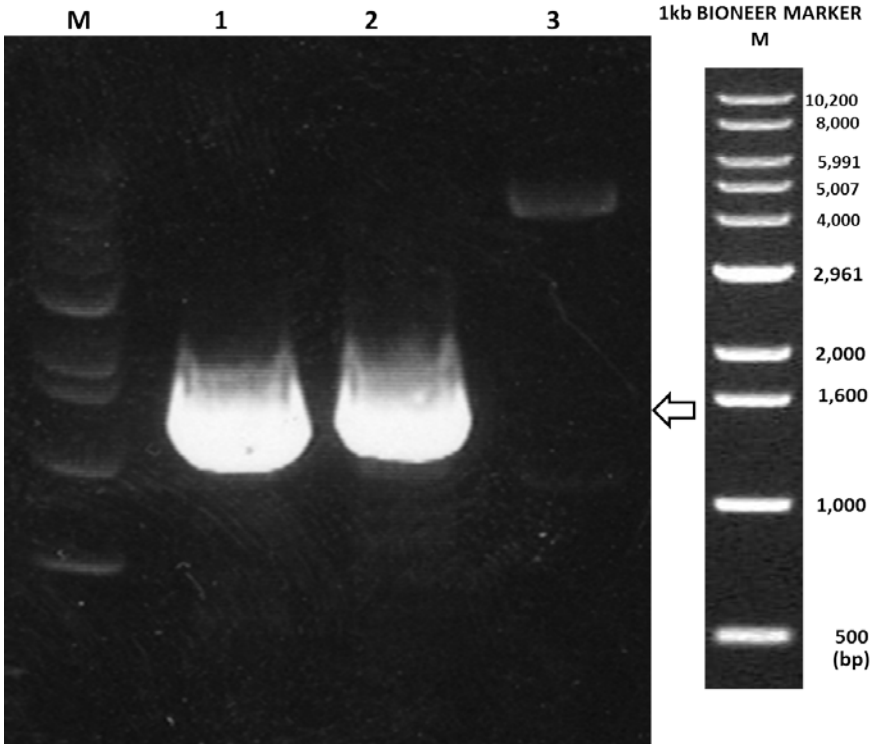

Figure 2-1 shows the PCR product of *P.horikoshii* AglB-L with 15bp N terminal and C terminal over hangs having sequence homology to the restriction digestion sites-Infusion technique (lanes 1 and 2). Lane 3 shows the double digested pET41b+ vector with Fast-Digest Nde I and Sal I restriction enzymes (Fermentas, using recommended procedures). The marker used was a 1kilo base BIONEER ladder shown on the far right. A 0.5% agarose gel preparation was used in recommended 1X TAE buffer for DNA electrophoretic works (100volts for 30min). The arrow head shows an apparent band shift of target DNA sequence of about 1485 base pairs.

After PCR amplification, infusion cloning technique and subsequent plating, positive colonies, selected by kanamycin resistance were picked up from Luria Broth (LB) agar plates and sub cultured for plasmid extraction (QIAGEN kit). Plasmid DNA was sequenced for confirmation using Big-dye version 3.1 within the Kyushu University facilities at Medical Institute of Bioregulation.

Correct sequenced plasmid DNA was used to transform *E.coli* BL21 (DE3) and plated on LB agar plates supplemented with 30mgl^{-1} kanamycin antibiotic as a selection marker. Resultant colonies were used for protein expression by pre culturing one colony in 5ml 2XYT

(SIGMA) media supplemented with 10µl of 30mg/ml kanamycin at 310 K for 5hours and transferring to 1 liter LB broth (SIGMA) media for main culture supplemented with 30mgl^{-1}. After absorbance at 600nm reached 0.5, protein was expressed by the addition of isopropyl-β-D-thiogalactopyranoside (Nacalai Tesque) to a final concentration of 0.5mM at 310 K in the *E. coli* BL21(DE3) Gold strain (Novagen, Merck) in LB medium. Cells were harvested by centrifugation at 8000xg for 15minutes at 269 K and resuspended in 30ml buffer A [50mM Tris-HCl buffer, pH 8.0, containing 150 mM NaCl, 2 mM $MgCl_2$, and 1 mM phenylmethylsulfonyl fluoride, supplemented with complete protease inhibitor mixture (ROCHE Applied Science) and 0.7 µl Benzonase 250 U/µl (Merck)], incubated on ice for 1 hour, and then sonicated using a Branson Sonifier model 250 on ice using the program "2 seconds on", "1 second off" at an amplitude of 25% for a total of 5 minutes. The C-terminal globular domain from *P.horikoshii* expressed exclusively as inclusion bodies, and collected as a pellet after centrifugation of the sonicate at 10,000 g for 30 min at 277 K.

Protein refolding

The pellet was washed with 15-20ml buffer B [50mM Tris-HCl, pH 8.0, 150mM NaCl, 2 mM EDTA, 2 mM DTT, and 1% Triton X-100] by vortexing and pipetting vigorously to solubilize the pellet. The suspension was centrifuged at 10,000g for 15minutes at 298 K. Triton wash was repeated three times each time centrifuging and collecting the pellet. Then, Triton X-100 was removed by washing the pellet using 20ml buffer B without Triton by repeating the vortexing, pipetting and centrifuging three times again and collecting the pellet each time. Finally, the pellet was solubilized in 20ml buffer B containing 8 M urea, vortexed and solubilized and incubated at room temperature with mild rotation for 1 hour. Then the solution was centrifuged at 10,000 g for 30 min at 298 K, and the insoluble materials were discarded. The urea solubilized supernatant was mixed with Ni-Sepharose resin (QIAGEN), pre-equilibrated with buffer C [20mM Tris-HCl, pH 8.0, 500 mM NaCl, and 8 M Urea] containing 20 mM imidazole. The mixture was incubated at 298 K with mild rotation for 1 hour, and then loaded onto an empty column (BIORAD Econo column). After washing the resin thoroughly with buffer C, the protein was eluted by using buffer C containing 300 mM imidazole, and 500 mM imidazole in a stepwise manner (Figure 2-2)

Figure 2-2 Nickel sepharose affinity tag purification of denatured *P.horikoshii* AglB-L

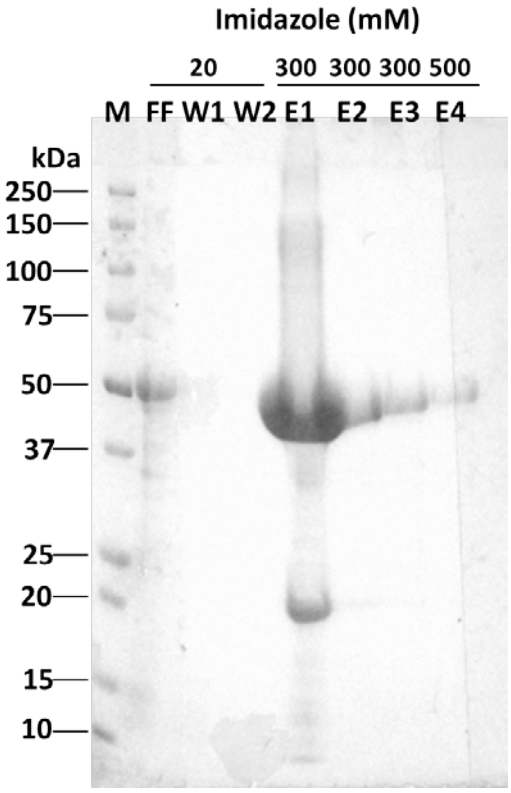

Figure 2-2 shows the elution pattern of the urea solubilized C-terminal globular domain of *P.horikoshii* AglB-L. M; Bio-Rad Precision Protein Standard, FF; flow through after incubation of supernatant, W1,W2; wash of the target protein on Nickel sepharose resin, E1 to E4; step wise elution of target with imidazole. The apparent molecular weight of the target protein was about 50kDa on the 10-20% acrylamide gel. The gel was stained using Quick CBB and partially destained using 50% methanol.

The fractions containing the target protein were pooled and concentrated. To refold the protein, the urea solubilized protein solution was directly injected onto a Superdex 200 HiLoad 26/60 gel filtration column (Amersham, GE healthcare), pre-equilibrated with buffer D [50 mM MES, pH 6.5, and 150 mM NaCl] and connected on an ACTA Purifier instrument (Amersham G.E). The programme was set up to elute 5ml fractions at a flow rate of 2.5ml/min maintaining pressure below 0.4 mega Pascal's (MPa). Fractions containing the refolded protein (Figure 2-3) were collected, and the protein was concentrated using an Amicon Ultra-15 centrifugal filter unit (Millipore).

Figure 2-3 Elution profile of *P.horikoshii* AglB-L using on-column refolding technique

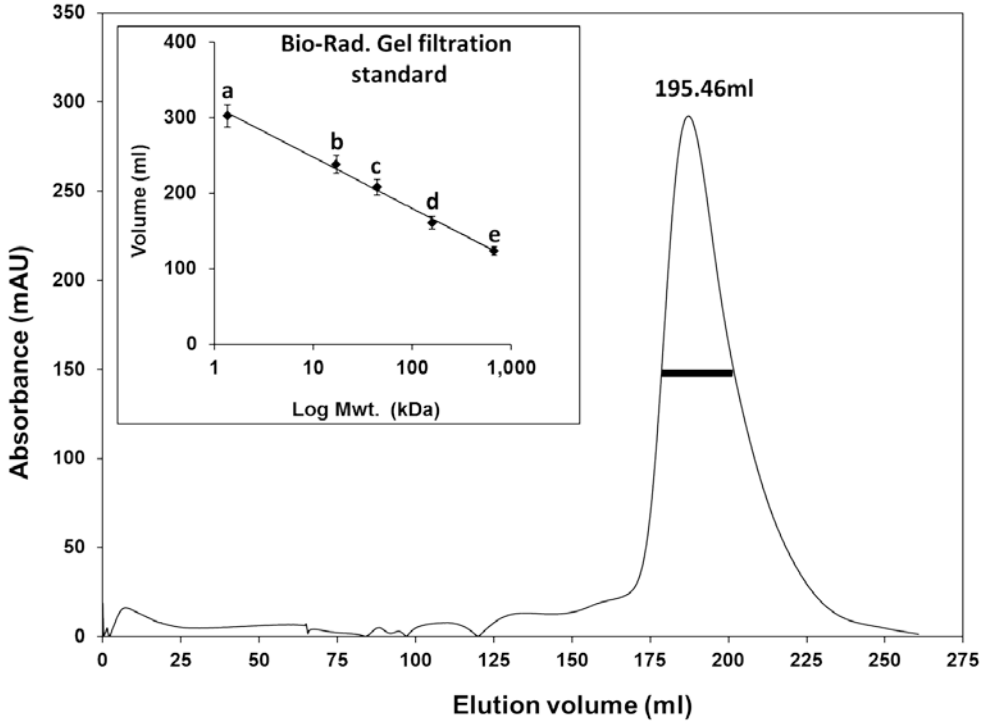

Figure 2-3 shows the elution pattern on a superdex 200 HiLoad 26/60 column. The insert shows the Bio-Rad protein standard run on the same column prior to sample elution. (a) Vitamin B 12; 1,350 Da (b) Myoglobin; 17,000 Da (c) Ovalbumin; 44,000 Da (d) γ-globulin; 158,000 Da and (e) Thyroglobulin; 670,000 Da. The horizontal black bar represents sample tubes run on SDS-PAGE (Figure 2-4) for checking purity and the top peaks were pooled, buffer exchanged on a superdex 200 10/300 GL column and concentrated using Amicon Ultra centrifuge tubes for crystallographic work. The target protein eluted at 195.46ml representing an apparent molecular weight of about 59kDa compared to the theoretical mass of about 57kDa.

Figure 2-4 Elution fractions of the C-terminal domain of *Ph*AglB-L as in Figure 2-3

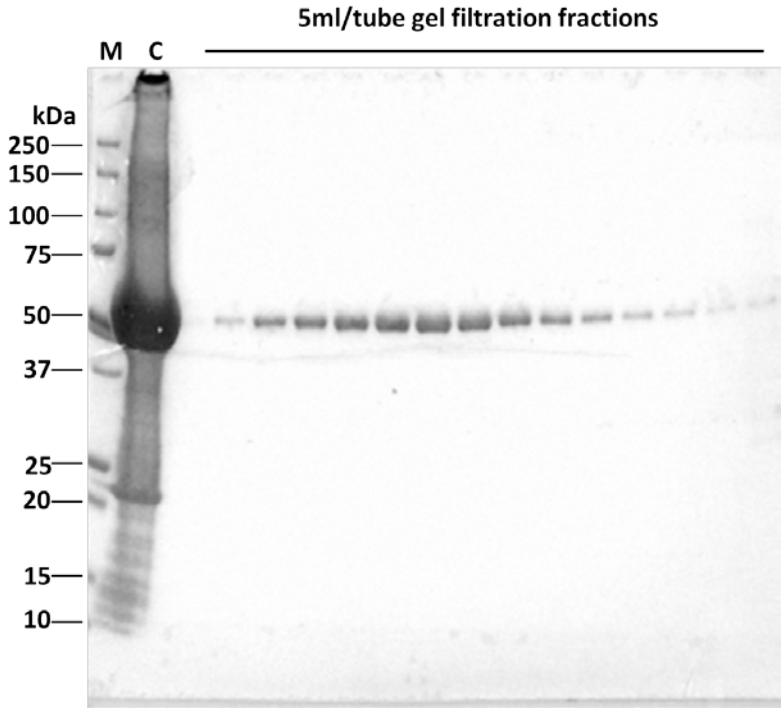

Figure 2-4 shows fractions of the C-terminal globular domain of *P.horikoshii* AglB-L (*Ph*AglB-L) represented by a horizontal bar in Figure 2-3 on the SDS-PAGE run at 23mAmperes and a total time of 75min. M; Bio-Rad marker, C; pooled crude target protein eluted by imidazole (Figure 2-2). The apparent molecular weight of the target protein was about 50kDa on the SDS-PAGE (10-20% acrylamide gel). A minor band at approximately 22 kDa was observed on the gel but was not incorporated in the crystal, vide infra. The gel was developed by Quick CBB stain and after rinsing with water was destained using 50% methanol as before.

The buffer was exchanged to buffer E [20 mM MES, pH 6.5] on a Superdex 200 10/300 GL (Amersham, GE healthcare). Fractions containing the protein were pooled and the protein was concentrated to 10 mg/ml for crystallization setup using the Hydra Plus One 96 well automated machine (Robin Instruments).

The genome of *A.fulgidus* contains three paralogs of AglB (AF_0380, AF_0329 and AF_0040 herein named as *A.fulgidus* AglB-L, *A.fulgidus* AglB-S1 and *A.fulgidus* AglB-S2 respectively i.e. *l*ong and *s*hort variant 1 and 2). A PCR product (from genomic DNA obtained from

NBRC 100126G, NITE Biological Resource Center Chiba, Japan) encoding the C-terminal globular domain (residues Glu433-Lys593) of *A.fulgidus* AglB-S2 (AF_0040) was cloned between the Nde1 and Sal1 sites of pET41b+ vector without a tag using Clonetech kit. Expression was done in *E. coli* BL21 Gold (DE3) cells in selenomethionyl core medium (Wako) derivative at 310 K by supplementing 50 mg l^{-1} L-selenomethionine and 30 mg l^{-1} kanamycin. Protein was expressed by the addition of isopropyl-β-D-thiogalactopyranoside at 310 K in the *E. coli* BL21(DE3) Gold strain (Novagen, Merck) to a final concentration of 0.5mM. Cells were harvested by centrifugation at 8000xg for 15minutes at 277 K and resuspended in 30ml TS buffer (50 mM Tris buffer, pH 8.0, 100 mM NaCl) and disrupted by sonication using a Branson Sonifier model 250 on ice using the program "2 seconds on", "1 second off" at an amplitude of 20% for a total of 5 minutes.

The protein was expressed in the supernatant and was purified by a sequential step of cation exchange with SP sepharose resin, cation exchange with Resource S (1ml), and size exclusion on a superdex75 10/300 GL column pre equilibrated with 50 mM MES, pH 5.5, 2 mM DTT, 100 mM NaCl. The protein was concentrated to 10 mg/ml in 20 mM MES, pH 5.5 as before for crystallization. All chromatography materials were purchased from GE Healthcare unless mentioned.

Crystallization and Structure Determination

The C-terminal domain of *Ph*AglB-L was crystallized in 0.1 M bis-Tris propane-HCl, pH 7.5, containing 0.2 M sodium citrate and 15% w/v PEG3350, at 293 K in hanging drops within 4 days (Figure 2-5). The C-terminal domain of *A.fulgidus* AglB-S2 (*Af*AglB-S2) was crystallized in 0.1 M MES-NaOH, pH 6.0, containing 0.1 M $MgCl_2$ and 10% w/v PEG3350, at 293 K in hanging drops within 1 day (Figure 2-6). Crystals were soaked in the reservoir solutions containing 20 % glycerol for *Ph*AglB-L and 20 % ethylene glycol for *Af*AglB-S2, for cryoprotection. Structure determination was performed by the molecular replacement method.

Data collection, phasing, and refinement statistics are summarized in Table 2-1. PyMol version 1.3 was used for graphic presentation (Schrödinger, LLC). Structural superposition was performed by the program GASH (Standley et al., 2005). The multiple sequence alignment was performed with the program MAFFT (Katoh and Toh, 2008).

Crystals of *Pyrococcus horikoshii* AglB-L

Figure 2-5 Crystals of *Pyrococcus horikoshii* AglB-L

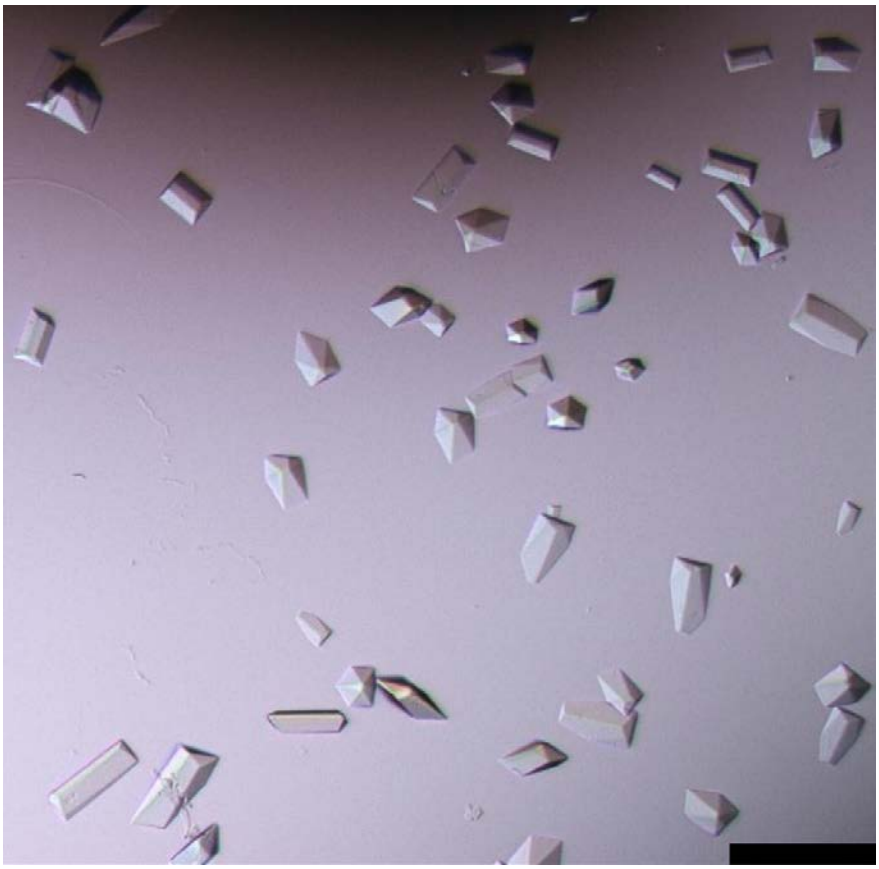

Figure 2-5 shows crystals of *P.horikoshii* AglB-L C-terminal globular domain. Crystals were grown in 0.2 M sodium citrate, 0.1 M Bis-Tris propane pH 7.5, 15% (w/v) PEG3350 at 293 K in the hanging drop within 4 days. Prior to collection of x-ray data sets, the crystals were soaked briefly in mother liquor containing 20% glycerol. X-ray data was collected at PF BL-17A. The horizontal bar represents 100μm.

Crystals of *Archaeoglobus fulgidus* AglB-S2

Figure 2-6 Crystals of *A.fulgidus* AglB-S2♣

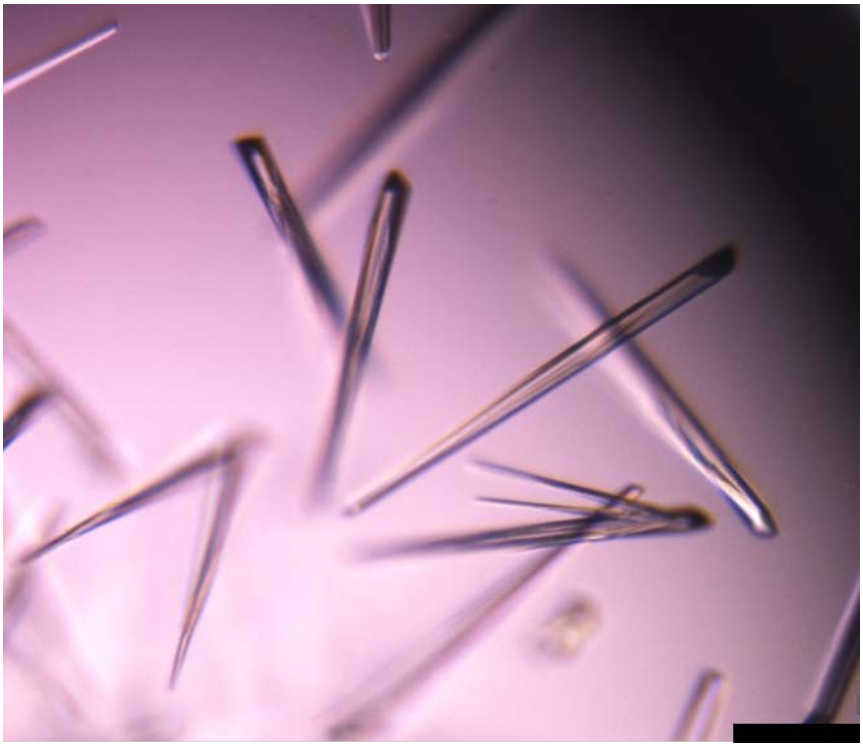

Figure 2-6 shows crystals of the C-terminal globular domain of *A.fulgidus* AglB-S2 in 0.1 M MES-NaOH, pH 6.0, containing 0.1 M $MgCl_2$ and 10% w/v PEG3350, at 293 K in hanging drops and grew within 1 day. The horizontal bar represents 100μm.

♣ Cloning, expression, purification and crystallization, courtesy of Matsumoto Shunsuke

Typical X-ray diffraction pattern of a protein crystal

Figure 2-7 Diffraction pattern of *P.horikoshii* AglB-L crystals

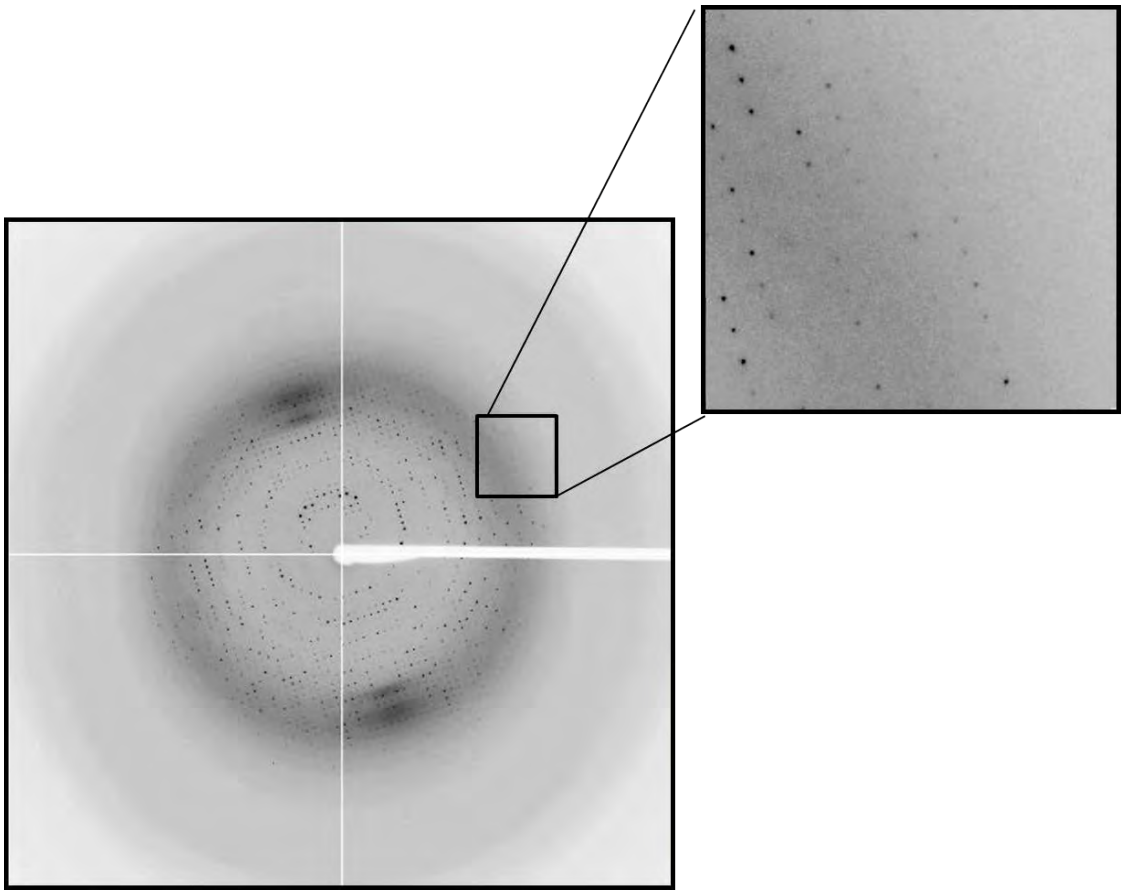

Figure 2-7 shows one frame image of the diffraction pattern obtained from crystals of *P.horikoshii* C-terminal domain. Crystals resolved to 2.7Å. The x-ray crystallographic data was obtained at the PF BL-17A.

Diffraction data processing

The diffraction data were processed using the program *HKL2000* (Otwinowski and Minor, 1997). The diffraction data of *Ph*AglB-L and *Af*AglB-S2 were processed to a resolution of 2.7 Å and 1.94 Å, respectively. The *Ph*AglB-L crystals contained two protein molecules per asymmetric unit (V_M = 3.3 Å^3 Da^{-1}, V_{solv} = 62.8 %), whereas the *Af*AglB-S2 crystals contained three protein molecules per asymmetric unit (V_M = 2.3 Å^3 Da^{-1}, V_{solv} = 46.2%). Structure determination was performed by molecular replacement using the program *MOLREP* (Vagin and Teplyakov, 2010) and using the C-terminal globular domain of *P.furiosus* AglB-L (*Pf*AglB-L) structure (PDB entry: 2ZAG) and *Af*AglB-S1 (PDB entry: 3VGP) as search models for *P.horikoshii* AglB-L and *A.fulgidus* AglB-S2 respectively. Further manual model building and refinement calculations were performed with the programs *COOT* (Emsley and Cowtan, 2004) and *REFMAC*(Murshudov et al., 1997) for *Ph*AglB-L, and *COOT* (Emsley and Cowtan, 2004)and *CNS* (Brunger et al., 1998) for *Af*AglB-S2.
Data collection, phasing and refinement statistics are summarized in Table 2-1.

The final PyMol cartoon rendering of the solutions from refinement data are shown in Figure 2-8 for the C-terminal globular domains of *P.horikoshii* AglB-L and *A.fulgidus* AglB-S2 respectively.

Accession numbers for the crystal structures

The atomic coordinates and structure factors of the C-terminal globular domains of *Ph*AglB-L and *Af*AglB-S2 have been deposited in the Protein Data Bank, with the accession codes 3VU1 and 3VU0, respectively.

Table 2-1 Summary of data collection, phasing and refinement statistics

	*Ph*AglB-L	*Af*AglB-S2
Data collection statistics		
Beamline	PF BL-17A	PF BL-17A
Wavelength (Å)	0.9800	0.9788
Oscillation range (º)	180	170
Space group	$P2_12_12_1$	$P3_1$
Cell parameters (Å)	a = 83.47, b = 94.84, c = 186.35	a =b = 111.21, c = 36.71
No. of molecules in AU	2	3
Resolution range (Å)a	40.0 - 2.7 (2.75-2.7)	50.0 - 1.94 (1.97-1.94)
Observed reflections	282,413	201,663
Unique reflections	41,657	37,593
Completeness (%)a	99.33 (89.8)	100.0 (100.0)
R_{merge}(I)a,b	0.064 (0.294)	0.097 (0.465)
I / σ (I)	22.8 (4.0)	28.1 (4.7)
Refinement statistics		
Resolution range (Å)	30-2.70	36.4 - 1.94
No. of protein atoms	7,852	3,844
No. of water / ion molecules	80 / 4	367 / 3
R/R_{free}	0.171/0.214	0.188/0.219
rmsd from ideality bond length (Å)	0.011	0.005
angles (º)	1.494	1.10
Ramachandran plot (%)c		
Favored region	96.0	97.6
Allowed region	3.8	2.4
Outlier region	2	0

[a] $R_{merge}(I) = (\Sigma\Sigma|I_i - \langle I \rangle|)/\Sigma\Sigma I_i$, where I_i is the intensity of the *i*th observation and $\langle I \rangle$ is the mean intensity. Values in parentheses refer to the outer shell.

[b] rmsd, root mean square deviation.

[c] Calculated using the program RAMPAGE.

Figure 2-8 Crystal Structures of the C-terminal Domains of Two Archaeal AglB Proteins

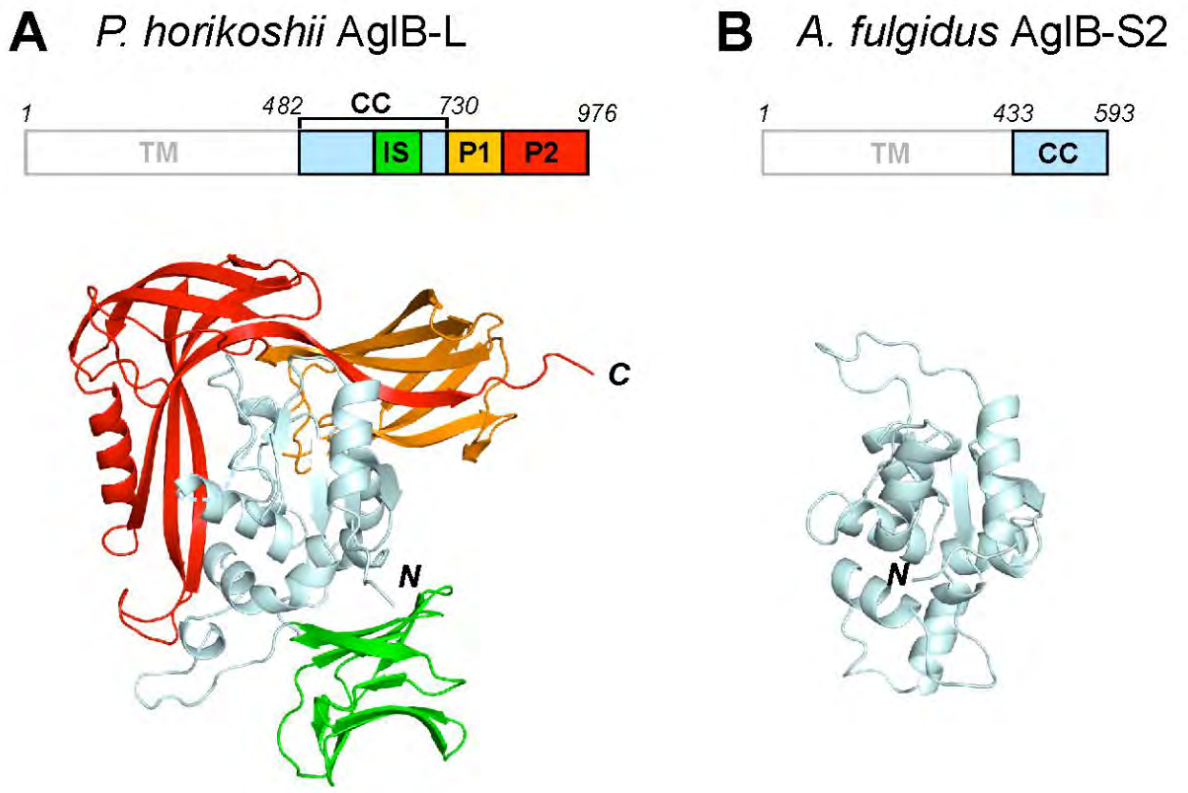

Figure 2-8 shows the (A) Overall structure of the C-terminal domain of *Pyrococcus horikoshii* AglB-L. The domain organization is schematically shown: TM, transmembrane; CC, central core; IS, insertion; P1, peripheral 1; P2, peripheral 2. The TM region, which was not included in the structure determination, is outlined in gray. (B) Overall structure of the C-terminal domain of *Archaeoglobus fulgidus* AglB-S2. The C-terminal domain consists of the CC structural unit alone.

Topological representation of PglB and AglB proteins

Two structures, *A.fulgidus* AglB-S2 and *P.horikoshii* AglB-L were compared using the known topology of the Eubacterial OST from *C.lari* PglB and are shown in Figure 2-9. All the three structures differ in size and domain structures of their respective C terminal globular domains. The N-terminal transmembrane domains of the Archaeal counterparts were just inferred using the orientation of *C.lari* PglB structure.

Figure 2-9 Topology of Archaeal OST domains with respect to Eubacterial OST

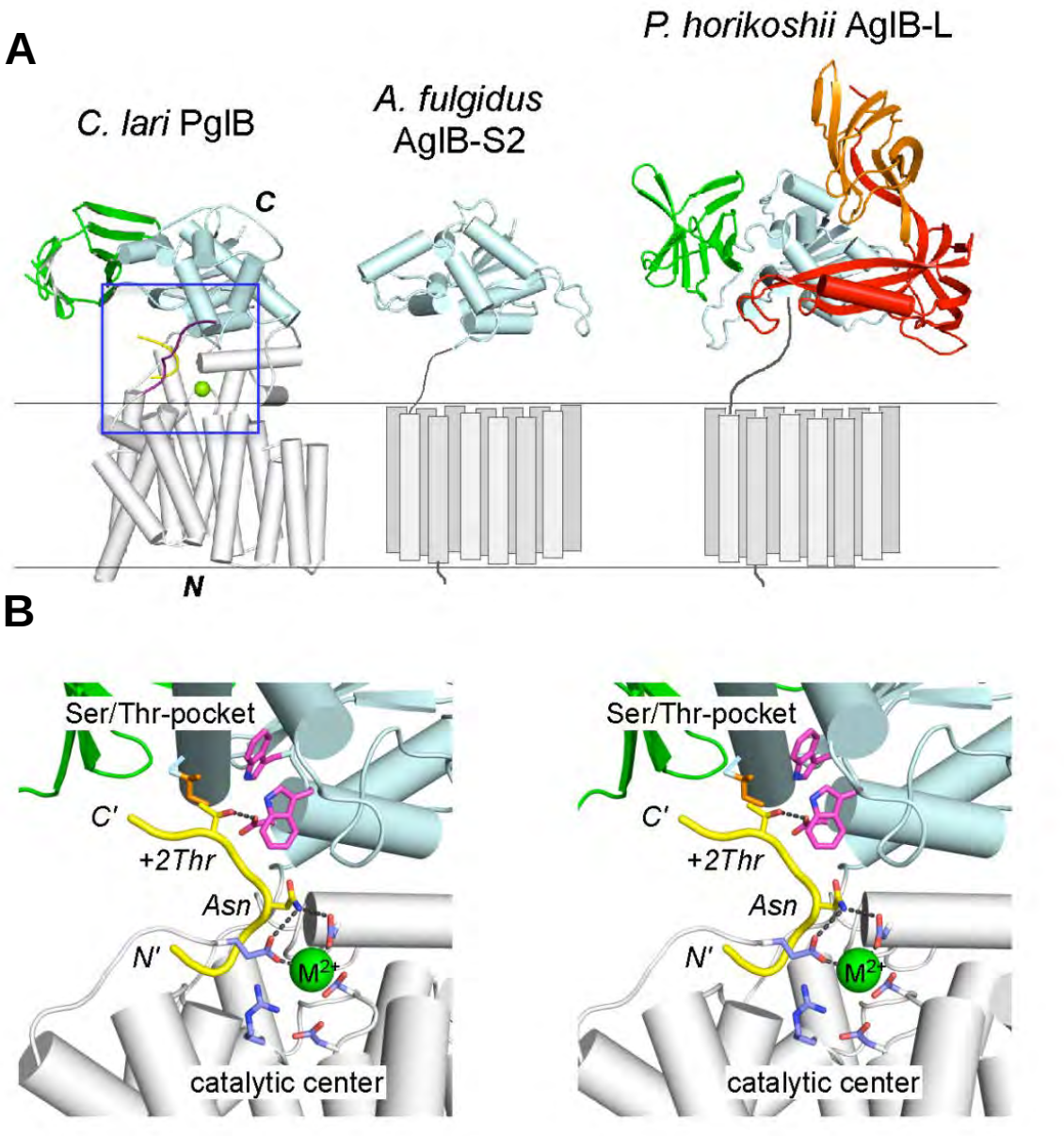

Figure 2-9 (A) shows the overall protein architectures of the AglB/PglB proteins. The full-length PglB structure from *Campylobacter lari* (left, PDB entry 3RCE) consists of the N-terminal transmembrane (TM) region and the C-terminal globular domain. The TM region contains the external loop 5 (*purple*), which was assumed to be involved in the enzymatic function. The C-terminal domain consists of the CC (*palecyan*) and IS (*green*) structural units. The C-terminal domains from *Af*AglB-S2 (center) and *Ph*AglB-L (right) were aligned with that of *Cl*PglB. The undetermined structures of the TM region of the two Archaeal AglBs are schematically depicted as a bundle of gray rectangles. The horizontal lines indicate the hypothetical position of the membranes. (B) Stereo view of the close-up of the active site (blue box, in A). The EL5 loop was not drawn for clarity. The bound peptide (*yellow* tube) contains an acceptor Asn side chain and a Thr side chain at the +2 position. Hydrogen bonds (dashed *black* lines) and interactions with a divalent metal cation (*green* ball) were regarded as a possible mechanism of amide nitrogen activation of the Asn side chain (Lizak et al., 2011). The side chain of the +2 Thr residue is recognized by a binding pocket formed by the Trp-Trp-Asp segment (*magenta* side chains) of the WWDYG motif, and Ile (*orange* side chain) of the DK/MI motif.

Selection of a crystal-contact free structure pair

To achieve this objective, a total of six structures, *P.furiosus* AglB-L(Igura et al., 2008), *P.horikoshii* AglB-L(this study), *A.fulgidus* AglB-S1(Matsumoto et al., 2012), *A.fulgidus* AglB-S2(this study), *C.jejuni* PglB(Maita et al., 2010) and *C.lari* PglB(Lizak et al., 2011) were compared to check for regions whose side chains in one protein molecule had contacts on the WWDYG portion of the second nearest neighbor in the crystal asymmetric unit. The results are shown in Figure 2-10. Only those structures which had side chains outside the 5Å length were selected as crystal contact free structures to represent the thermodynamically stable state of the OST enzyme. The rest having side chains within the 5Å limit with respect to contact with the WWDYG motif segment were deemed having undesirable crystal contact effects. Those selected as representatives of thermodynamically stable states were *A.fulgidus* AglB-S2, *C.jejuni* PglB and *C.lari* PglB respectively.

Figure 2-10 Close-up Views of the Crystal Contact Sites Involving the WWDYG Motif

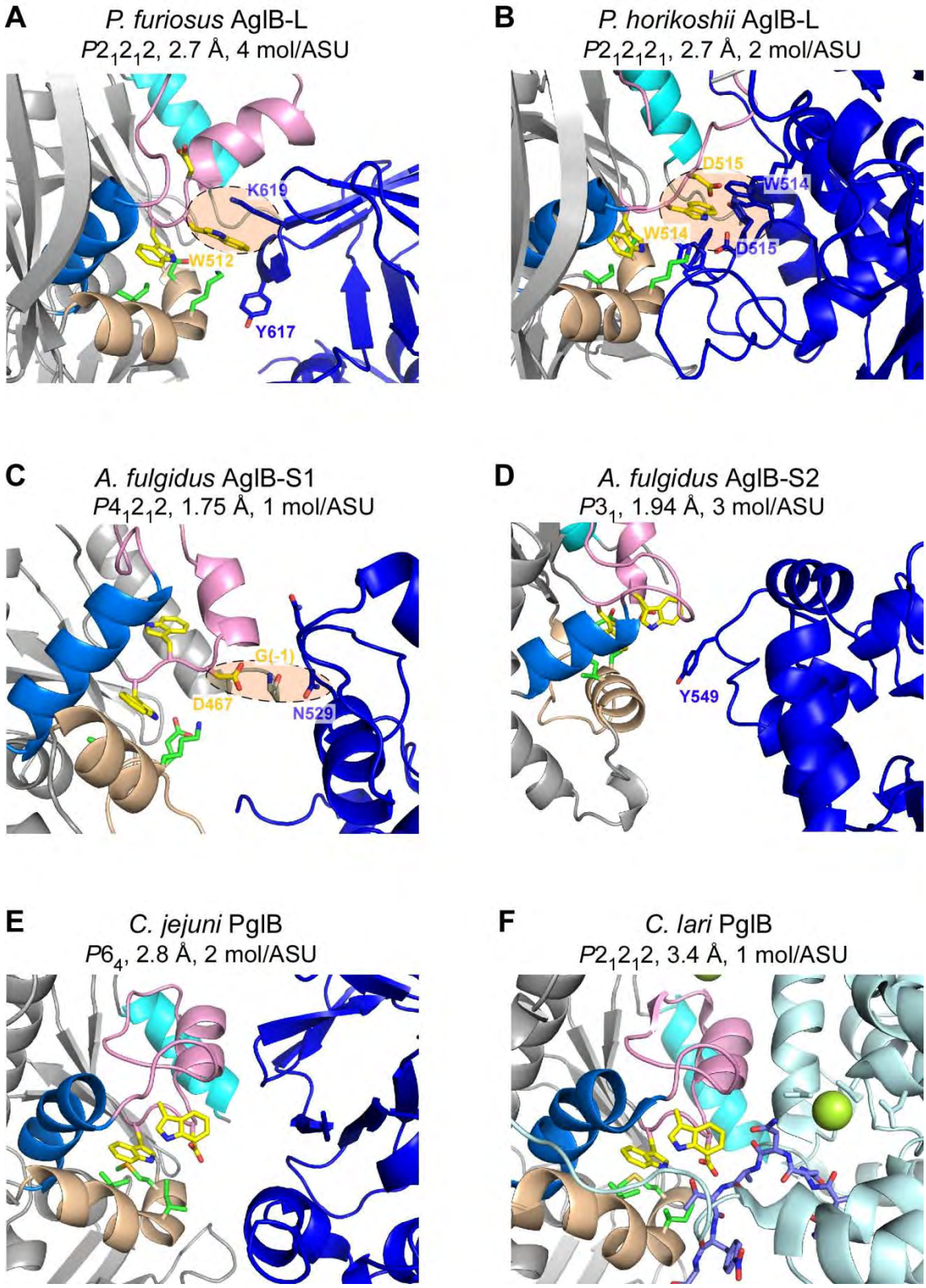

Figure 2-10 shows various crystal contacts for the structures of the C-terminal globular domain of select OSTs. The THL segment (= turn + α-helix + loop) is colored *pink*, and the kinked helix bearing the DK/MI motif is *light brown*. The side chains of the Trp-Trp-Asp part of the WWDYG motif are colored *yellow*, and those of the three signature residues of the DK/MI motif are *green*. The side chains of the amino acid residues that are located within 5 Å of the Trp-Trp-Asp part of the WWDYG motif in a neighboring protein molecule are colored *deep blue* and labeled. The shaded ovals show the clash site between the two neighboring molecules. The space group, resolution, and number of molecules in the asymmetric unit are shown. (C) The N-terminal Gly residue at position -1 was directly involved in the crystal contacts in the *Af*AglB-S1 structure. (F) The bound acceptor peptide is represented by the blue stick model, and the bound divalent metal ion is depicted as a *yellow-green* sphere in the *Cl*PglB structure. Panels (A-C) show structures distorted by crystal contact effects, and (D-F) show crystal contact-free structures.

Discussion

The C-terminal globular domains of *P.horikoshii* AglB-L, and *A.fulgidus* AglB-S2 were solved to resolutions of 2.7Å and 1.94Å respectively using the molecular replacement method(Vagin and Teplyakov, 2010). Overall, the structure of *P.horikoshii* does not differ from that of *P.furiosus* on a domain by domain basis if viewed from a framework point of view and having had solved the problem of distortion of the lysine residue of which *P.horikoshii* lack (rather possesses Arginine 621 at the structurally equivalent position to Lysine 619), it was concluded that the WWDYG-motif segment has variable conformations. The crystal structure of *A.fulgidus* AglB-S2 (this study) together with *A.fulgidus* AglB-S1 reported by Matsumoto et al 2012 also revealed that not only is side chain variability present but the C-terminal domain of select OST enzymes contain variable motifs as a new motif in S1 and S2 was identified as the DKi (Matsumoto et al., 2012) motif as mentioned earlier. Taken together, a comparison of globular domains reveals that globally, the framework is more or less conserved in evolution while the details at side chain level reveal subtle structural variability not merely as a crystal packing phenomenon but probably due to inherent structural plasticity within the WWDYG motif segment. Following such observations, it was proposed that this plasticity phenomenon observed may have important catalytic functions conferred upon OST in various domains of life.

Chapter 3 - Comparative structural analysis

Packing effect on various OST's

The aim of this work was to analyze structural data with respect to multiple structural alignments to answer the question of presence or absence of crystal packing effects and contacts. This was achieved by using Bioinformatics software MAFFT(Katoh and Toh, 2008) and GASH(Standley et al., 2005). Five crystal structures of the C-terminal globular domains of *P.furiosus* AglB-L(Igura et al., 2008), *C.jejuni* PglB(Maita et al., 2010), *P.horikoshii* AglB-L (this study), *A.fulgidus* AglB-S1(Matsumoto et al., 2012) and S2 (this study) respectively were successfully solved in the same laboratory. Comparing all the structures revealed many differences when the region of interest, the WWDYG motif segment was considered in detail. As a result, this work led to a more rigorous comparative study using information on the crystals.

Comparative structural analysis is proving to be an important tool in structural biology as some new motifs become more vivid and more elucidated than mere amino acid sequence alignment. In total, five AglB/PglB structures (all solved from one laboratory) and one structure from a different laboratory (Lizak et al., 2011) were used for detailed comparison. A holistic comparison of the globular domains with emphasis on the highly flexible region of the WWDYG motif was performed, and found a highly superimposable conformation of the WWDYG motif segment between the most distant pair: *A.fulgidus* AglB and *C.jejuni* PglB, even with an overall sequence similarity of less than 20%. Recently, full length oligosaccharyltransferase OST from *Campylobacter lari* PglB was reported and using a pairwise superpositioning of the globular domains of *C.jejuni* PglB and *C.lari* PglB(Lizak et al., 2011) gave a perfect match over the main chain and side chains of the two structures due to high sequence identity (52%) and incidentally *C.lari* PglB(Lizak et al., 2011) C-terminal domain overlapped with *A.fulgidus* AglB-S2 in what was termed a high superpositioning reminiscent of a thermodynamically stable state. From these comparison studies it was proposed that these structural conformations of *C.jejuni* PglB, *A.fulgidus* AglB and the *C.lari* PglB represent the conformation in the resting or thermodynamically stable state of part of the catalytic site of the OST enzyme, free from the crystal packing artifacts.

Summary of Multiple Structural Alignment of five Representative Structures

To gain further understanding of the crystal contacts effects reported earlier (Igura et al., 2008; Maita et al., 2010), a multiple structural alignment of the C- terminal globular domains of OST's including the full length Eubacterial OST that were solved was done by using the programme GASH (Standley et al., 2005) a software for structural alignment (http://pdbj.org/ash/index.html) and pairwise sequence identity was calculated by using the arithmetic mean sequence length as the denominator (May, 2004). The results are presented in Table 3-1, Figures 3-1, 2 and 3 below.

Table 3-1 RMSD and aligned chain length of select C-terminal globular domains of OST's

Pair	RMSD	Align. length	Pair	RMSD	Align. length	Pair	RMSD	Align. length
Pf:Ph	**0.99**	**470**	*Ph:Cj*	3.03	150	*Cj:Af*S1	2.19	110
Pf:Cj	3.05	153	*Ph:Af*S1	2.52	129	***Cj:Af*S2**	**2.03**	**133**
*Pf:Af*S1	2.59	130	*Ph:Af*S2	2.23	123	***AfS1:Af*S2**	**1.93**	**149**
*Pf:Af*S2	2.45	132	***Cj:Cl***	**1.15**	**258**			

The various combinations were carried out and those highlighted had rmsd values of less than 2.00 showing a high degree of sequence identity. With the exception of the *Cj:Af*S2 having an rmsd value of 2.03, the rest had values above 2.20 and such combinations were not considered for pair wise structural comparison studies.

PyMol Cartoon Representation of Pairwise Structural Alignment

For the structures solved, 5 structures of the C-terminal globular domains from *P. furiosus* AglB-L, *P.horikoshii* AglB-L, *A.fulgidus* AglB-S1, *A.fulgidus* AglB-S2 and *C.jejuni* PglB were from the same laboratory and 1 full length structure (*C.lari* PglB) was solved in a different laboratory(Lizak et al., 2011). A pairwise comparison using the C-terminal globular domains is illustrated in Figures 3-1 (high sequence identity pairs), Figure 3-2 (low sequence identity pair) and Figure 3-3 (moderately high sequence identity pair) respectively. It was unexpected to find that structures whose sequence identity was as low as 19% could show a high degree of structural superposition free from crystal contact effects (Figure 3-2). What was anticipated was that high sequence identity structures (Figures 3-1 A and B) would show a high degree of superposition but was not so with exception of the eubacterial pair (Figure 3-3). Note that only the C-terminal globular domains from the eubacterial organisms was aligned as the structure of *C.jejuni* PglB is only the C-terminal domain while that of *C.lari* PglB is full length.

Figure 3-1 Pairwise comparison of high sequence identity structures

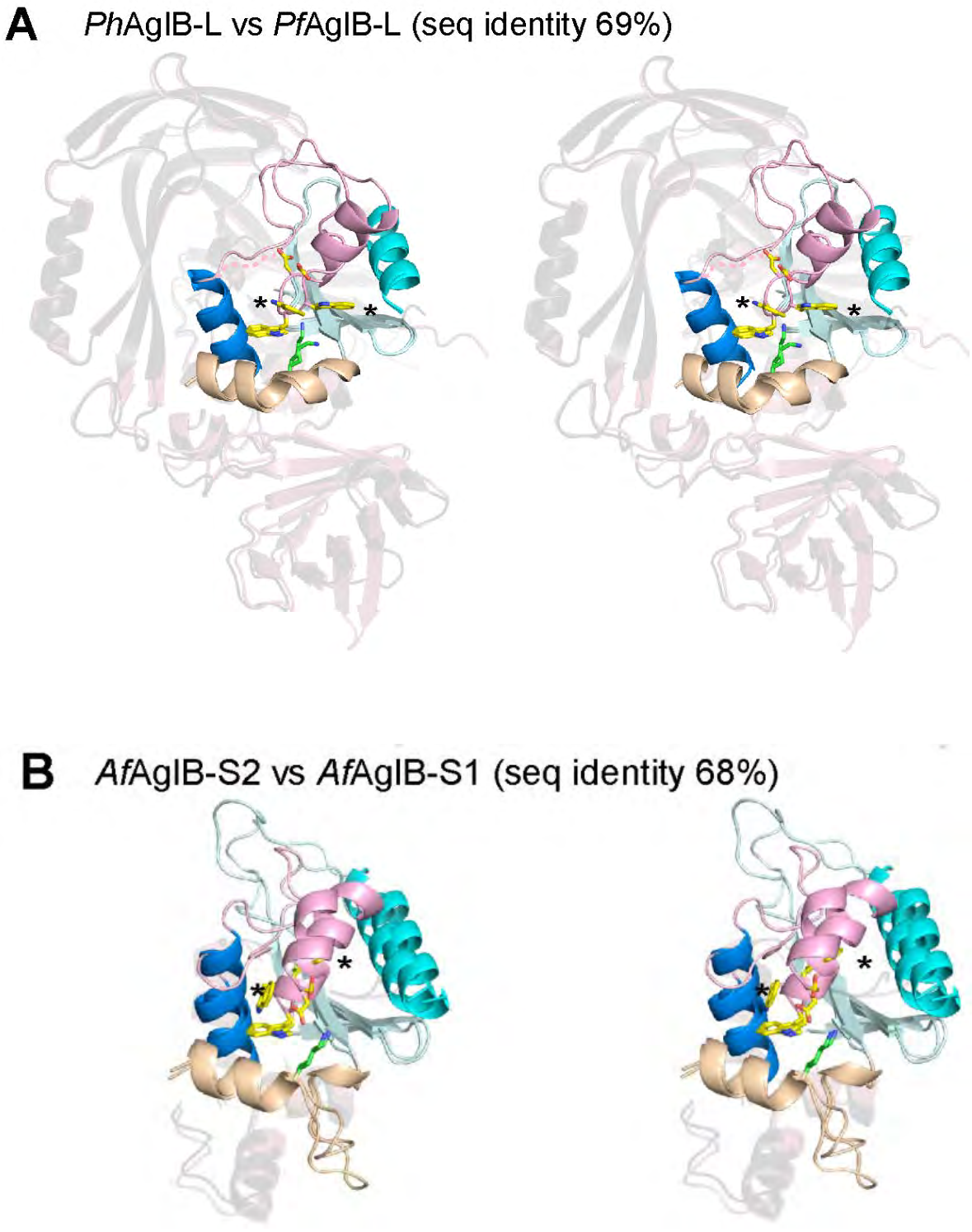

Figure 3-1 shows structural comparisons as indicated in Table 3-1. (A) Stereo view of superimposition of the *Ph*AglB-L and the *Pf*AglB-L structures. The root-mean-square deviation (rmsd) of the structural superposition is 0.99 Å over 470 aligned Cα atoms. The CC structural units are highlighted while the other units are transparent. The THL segment (= turn + α-helix + loop) is colored *pink*, and the kinked helix bearing the DK motif is *light brown*. The side chains of the Trp-Trp-Asp part of the WWDYG motif are colored *yellow*, and that of the Lys residue of the DK motif is *green*. The asterisks mark the second Trp residue of the WWDYG motif. (B) Superimposition of the *Af*AglB-S2 and the *Af*AglB-S1 structures. The rmsd is 1.93 Å over 149 aligned Cα atoms. The high sequence identity structures rather had a low superposition in the WWDYG motif segment a situation not anticipated as high sequence identity was postulated to equally translate into high degree of superposition of the structural elements.

Figure 3-2 Comparison of two OST's from different domains of life, *Cj*PglB and *Af*AglB-S2

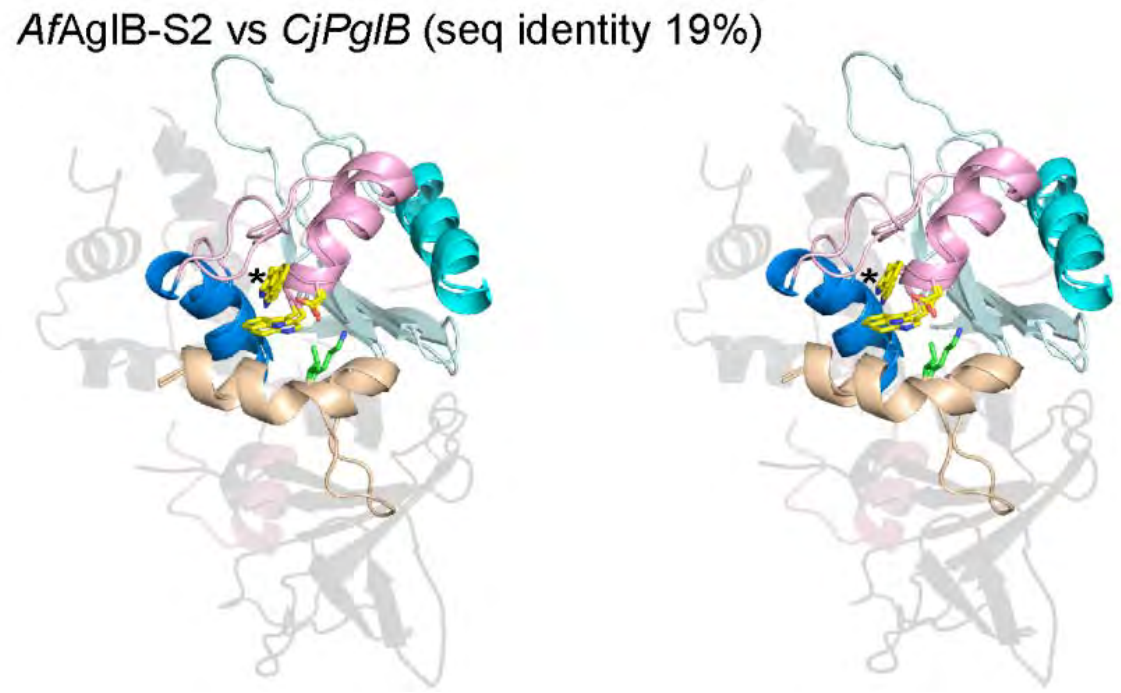

Figure 3-2 shows structural comparison of OST's from different domains of life i.e. the superimposition of the *Af*AglB-S2 and the *Cj*PglB structures. The rmsd is 2.03 Å over 133 aligned Cα atoms. The side chain of the Ile residue of the MI motif in PglB is colored *green*. The lysine from *Af*AglB-S2 is colored green with the nitrogen colored blue. The second tryptophan labeled with an asterisk as well as the other residues in the WWDYG motif segment highly superimposed contrary to initial thought that the two proteins being from different domains of life may possess completely different folds. With the exception of the insertion loop in the DK motif (Matsumoto et al., 2012), the rest of the secondary structural elements superimposed. Note that the C-terminal globular domain of *A.fulgidus* AglB-S2 lacks the Beta barrel insertion domain structure identified in *P.furiosus (Igura et al., 2008)*, *P.horikoshii* (this study) and *C.jejuni* (Maita et al., 2010)and *C.lari* (Lizak et al., 2011) OST's respectively.

Figure 3-3 Comparison of Eubacterial OST's *Cj-Cl*PglB

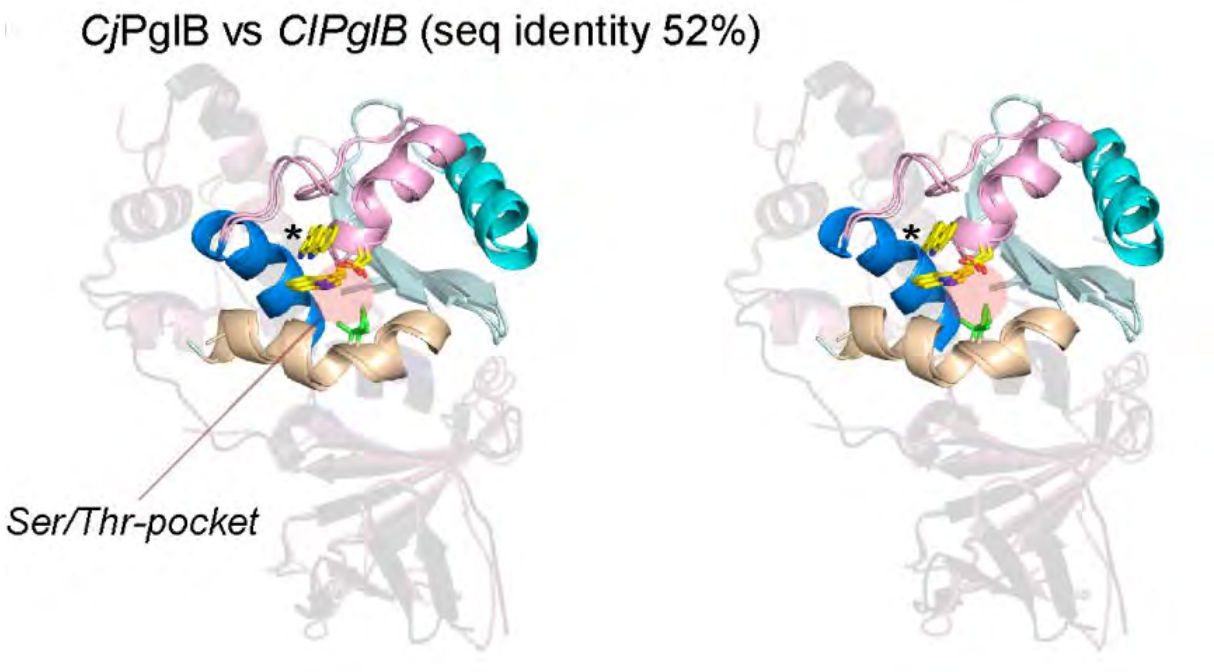

Figure 3-3 shows structure based alignment of two Eubacterial OST's *Campylobacter jejuni* and *Campylobacter lari* PglB's. With a sequence identity of about 50%, the two structures are highly superimposed. The rmsd is 1.15 Å over 258 aligned Cα atoms. The shaded circle indicates the location of the Ser/Thr pocket, formed between the WWDYG motif and the MI motif. Note that the structure from *C.lari* PglB(Lizak et al., 2011) is full length while that of *C.jejuni* PglB(Maita et al., 2010) is for the C-terminal globular domain. This observation of a high degree of superposition in the residues forming the WWDYG motif segment was anticipated due to fairly high sequence identity of the eubacterial OST's.

Discussion

The most distant pair, that of *C.jejuni* PglB (eubacteria) and *A.fulgidus* AglB-S2 (archaea) was found to be highly superposable even with very low sequence identity (~less than 20%) and consequently deemed this pair as the most thermodynamically stable or resting form of the catalytic conformation of OST free from crystal packing effects. From the crystal structures solved so far, it can be inferred that even among the same family of enzymes, structural differences may be observed if a large enough sample size is used for comparative studies as employed in this study. Crystal structures from closely and distantly related but same family of enzymes were used to solve the problem of crystal packing and deducing which of the structures and pairs was consistent with crystallographic rules about crystal packing artifacts observed in this field of science. The consistent pair was that of a Eubacteria, *C.jejuni* PglB and an Archaeon, *A.fulgidus* AglB-S2. Other reasons why certain structures possessed the clash in their crystals around the WWDYG-motif segment could be attributed to differences in the environment(Chopra et al., 2008) , source of origin and polypeptide length among others (Carugo and Argos, 1997; Eyal et al., 2005; Krissinel, 2010; Zhang et al., 1995) or simply the inherent intrinsic plasticity of the molecular structure under observation.

Chapter 4 - NMR relaxation studies

Expression and labeling of samples for NMR experiments

The main objective of NMR relaxation studies was to offer evidence of flexibility of the region under study in solution. To achieve this purpose, various sample preparations of the C-terminal globular domain of *A.fulgidus* AglB-S2 were made for each detailed task discussed here forth.

Sample preparation

For NMR studies, *A.fulgidus* AglB-S2 (AF_0040) was sub cloned in pET47b+ between the Sma1 and Sal1 sites retaining the N terminal 6x His tag.[^{13}C,^{15}N] or ^{15}N uniformly labeled protein expressed in the supernatant and was purified by elution on the Ni-sepharose column under native conditions using buffer containing 20mM NaH_2PO_4,150mM NaCl,200mM Imidazole at pH7.0. Fractions containing target protein were pooled and further purified on a superdex 200 10/300 GL column equilibrated with 20mM NaH_2PO_4, 150mM NaCl, 2mM DTT. Peak top fractions were collected and buffer exchanged into 20mM MOPS pH7.0, 50mM NaCl, 2mM DTT. Protein was concentrated to 1.0 mM and a total volume of 260µl supplemented with 10% D_2O was loaded onto a 5.0 mm Shigemi tube and NMR spectra taken at 308 K on a Bruker 600MHz or 700MHz spectrometer.

Backbone Peak Assignment experiments

[^{13}C,^{15}N] uniformly labeled protein was expressed in M9 media supplemented with 2.0g/L ^{13}C-glucose and 1.0g/L ^{15}N- ammonium chloride(Isotec) as the sole source of carbon and nitrogen respectively. Pre culture was prepared in 5.0ml 2X YT media and cultured at 310K for 3hr 30min and 1.0% (v/v) was used to inoculate 1.0L M9 media supplemented with 30.0µg/L kanamycin. Growth was followed at the same temperature and when $O.D_{600nm}$ reached 0.3 absorbance units, temperature was reduced to 291 K before induction with Isopropyl β-D-1-thiogalactopyranoside (Nacalai tesque) to a final concentration of 0.5mM. After 20h, cells were harvested by centrifugation at 5000xg at 277 K for 15min. The pellet was washed in 20ml wash buffer (20mM sodium phosphate pH7, 150mM NaCl) and then resuspended in 40ml sonication buffer [20mM sodium phosphate pH7, 150mM NaCl , 1mM

Benzylsulfonyl Fluoride (Wako),2mM DTT supplemented with complete cocktail inhibitor (Roche)]. The sonicate was centrifuged at 10000xg at 277 K for 30min and pellet was discarded. The supernatant was applied on a Bio-Rad open column loaded with (1.5ml 50% resin/L culture) prewashed High Performance Ni-sepharose (GE Healthcare) equilibrated with binding/wash buffer [20mM sodium phosphate pH7, 150mM NaCl supplemented with 10mM imidazole]. The resin was washed with abundant wash buffer to remove unbound material. Target protein was eluted with elution buffer [20mM sodium phosphate pH7, 150mM NaCl] supplemented with 200mM and 300mM imidazole respectively.

Quality of the elution fractions was checked by SDS_PAGE on a gel with a gradient of 10-20% acrylamide. Fractions containing 95% target protein were pooled and applied on a superdex 200 10/300 GL column (Amersham) equilibrated with [20mM sodium phosphate pH7, 150mM NaCl, 2mM DTT]. The protein eluted as a single peak and fractions containing the peak tops were pooled and concentrated using an Amicon 10,000 mwco centrifuge tube.

For preparing an NMR sample, protein was exchanged into 20 mM MOPS pH7, 50mM NaCl, 2.0mM DTT buffer. A total of 400µl protein 1.0mM, supplemented with D_2O to a final concentration of 10% was added into a shigemi tube and measurements were taken on a Bruker 600MHz instrument fitted with a TXI cryo probe. ^{15}N 2D-HSQC spectra were taken at 308 K and processed using NMRpipe on the DEL workstation and assignment of the backbone peaks was performed using the program *Sparky* (http://www.cgl.ucsf.edu/home/sparky/).

Spectral Assignment

Spectral assignments of backbone signals were performed with standard multidimensional NMR experiments, HNCACB, CBCA(CO)NH, HN(CA)CO and HNCO coupled with ^{1}H, ^{15}N -[Valine, Tryptophan ,Lysine, Alanine and Tyrosine] amino acid specific labeled 2D-HSQC's to help in disambiguation.. Chemical shifts were referenced to internal DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid), Appendix A. The ^{1}H,^{15}N 2D-HSQC spectrum of *A.fulgidus* AglB-S2 is shown in Figure 4-1.

^{15}N relaxation experiments were acquired at 600 and 700 MHz on Bruker Avance spectrometers. Pulse sequences used were the gradient sensitivity enhancement version

(Farrow et al., 1994). R_1 was determined from data with relaxation delays of 30, 80, 150, 300, 500, and 800 ms, and R_2 with delays of 57.6, 86.4, 115.2, 144.0, 172.8, and 201.6 ms. In the heteronuclear ^{1}H-^{15}N NOE experiment, the ^{1}H saturation period was 3.0 s during the relaxation delay. Uncertainties were estimated to be 10% of the measurement at both fields. NMR spectra of *A.fulgidus* AglB-S2 were recorded at 308 K on a Bruker Avance600 spectrometer equipped with a TXI cryoprobe.

Figure 4-1 ^{1}H,^{15}N-HSQC spectrum of *A.fulgidus* AglB-S2 at 600MHz, 308 K.

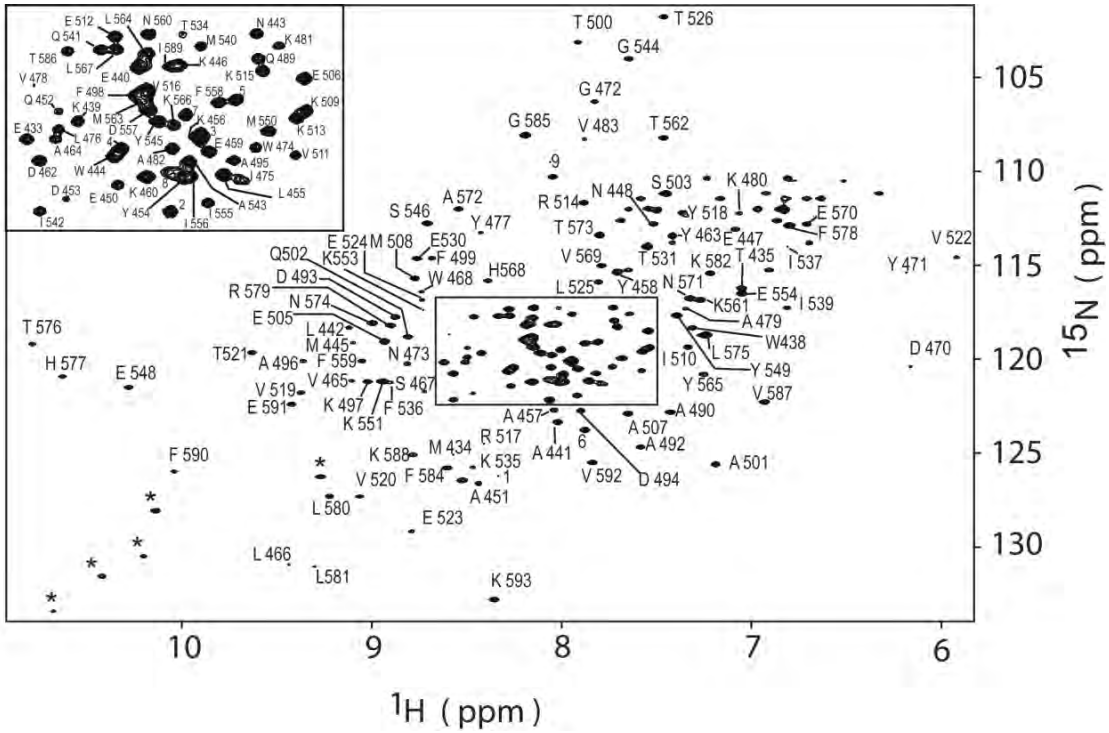

Figure 4-1. The 600 MHz ^{1}H-^{15}N 2D HSQC Spectrum of *A.fulgidus* AglB-S2 was recorded at 308 K and peaks were assigned by different 2D spectra as well as amino acid specific labeled ^{15}N HSQC spectra. The side-chain ^{1}H-$^{15}N^{\varepsilon 1}$ cross peaks of the five tryptophan residues are marked with asterisks. The cross peaks numbered 1 to 9 are plasmid-derived, and were assigned to Ala10, Ala11, Leu12, Glu13, Val14, Leu15, Phe16, Gln17, and Gly19, respectively.

R_1, R_2 and NOE Relaxation Parameters in Protein NMR

Two major sources of relaxation of the ^{15}N nuclear spins in proteins comprise dipolar coupling with the protons and the associated chemical shift anisotropy (CSA) of the NH interaction. The movement of the NH bond is characterized by the spectral density function $J(\omega)$. This is in turn related to the three relaxation parameters, R_1, R_2 and NOE given by;

$$R_1 = (d^2/4)[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H - \omega_N)] + c^2 J(\omega_N) \quad 1$$

$$R_2 = (d^2/8)[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H - \omega_N)] + c^2/6[4J(0) + 3J(\omega_N)] + R_{ex} \quad 2$$

$$NOE = 1 + (d^2/4R_1)(\gamma_H/\gamma_N)[6J(\omega_H - \omega_N) - J(\omega_H - \omega_N)] \quad 3$$

where $d = (\mu_0 h\gamma_N\gamma_H/8\pi^2)\langle r_{NH}^{-3}\rangle$, $c = \omega_N\Delta\sigma/\sqrt{3}$, μ_0 is the permeability of free space, h is Planck's constant, γ_H and γ_N are the gyromagnetic ratios of ^1H and ^{15}N,respectively. For the nitrogen- hydrogen bond; $r_{NH} = 1.02$ Å; ω_H and ω_N are the larmor frequencies of hydrogen and nitrogen respectively and $\Delta\sigma = -160$ ppm is the chemical shift anisotropy for ^{15}N nuclei in helical polypeptides.

R_1 is the longitudinal or spin-lattice relaxation rate constant$\left(R_1 = {}^1/_{T_1}\right)$, 4

R_2 is the transverse or spin-spin relaxation rate constant $\left(R_2 = {}^1/_{T_2}\right)$ and 5

NOE is the nuclear overhauser effect.

$$J(\omega) = \frac{2}{5}\left[\frac{S^2\tau_m}{1+(\omega\tau_m)^2} + \frac{(1-S_f^2)\tau_f'}{1+\left(\omega\tau_f'\right)^2} + \frac{(S_f^2-S^2)\tau_S'}{1+\left(\omega\tau_f'\right)^2}\right], \quad 6$$

where $\tau_f' = \tau_f\tau_m/(\tau_f + \tau_m)$, $\tau_s' = \tau_s\tau_m/(\tau_s + \tau_m)$, where τ_m is the overall rotational correlation time of the molecule, τ_f is the effective correlation time for internal motions on a fast time scale ($\tau_f < 100$ to200 ps). τ_s is the effective correlation time for internal motions on a slow time scale ($\tau_f < \tau_s < \tau_m$). The term $R_{ex,}$ is the rate of chemical exchange between

two conformational states and describes the loss of transverse magnetization due to chemical or conformational exchange (Mandel et al., 1995).

The relaxation data collected at 600 and 700MHz for *A.fulgidus* AglB-S2 are shown in Figure 4-2. The residues at both fields having R_2 values of 20 (600MHz) or 25 (700MHz) and above were inferred to contain the chemical exchange rate R_{ex} as explained earlier.

Figure 4-2 Plot of relaxation parameters NOE, R_1 and R_2 at 600 and 700MHz.

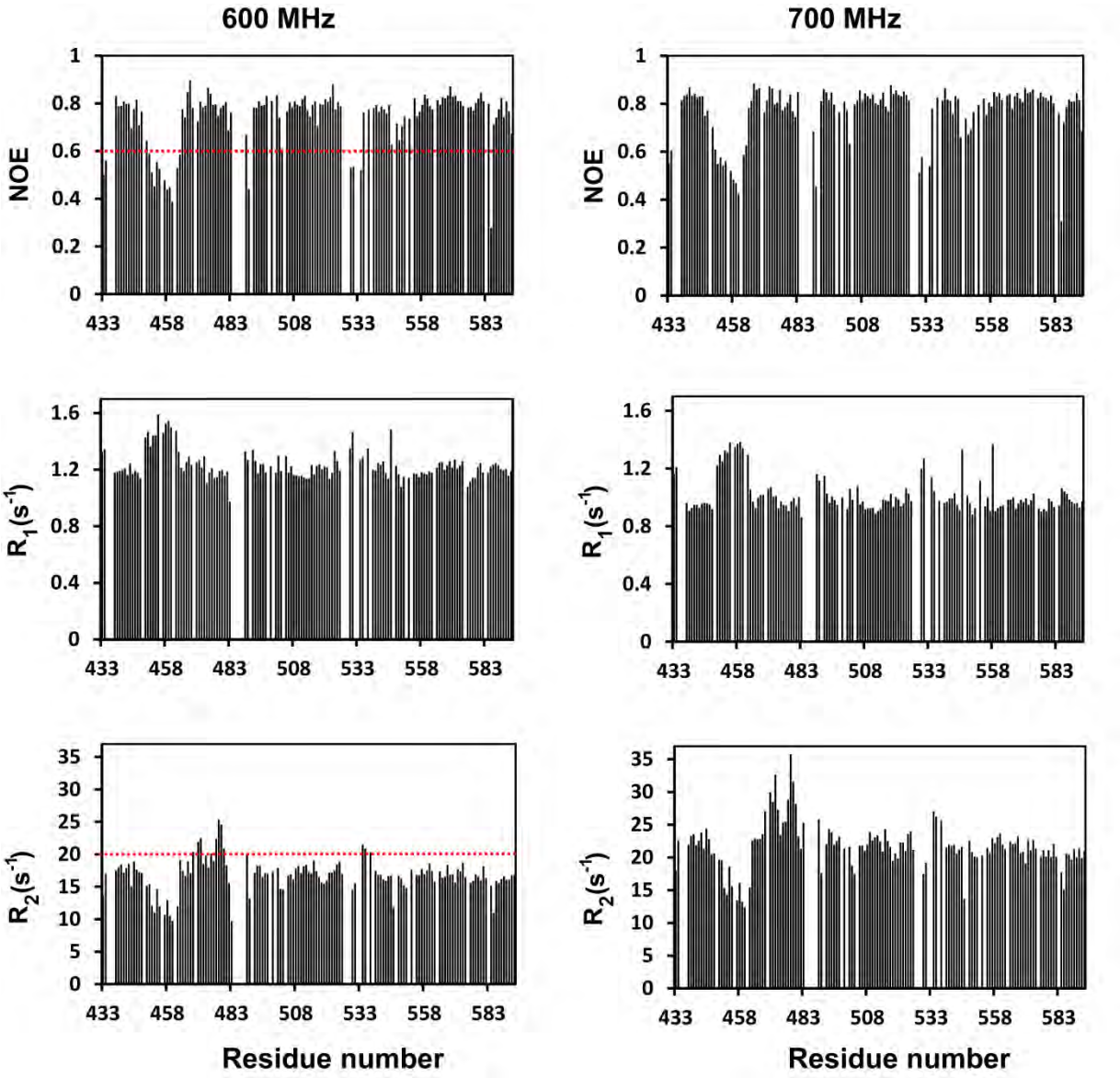

Figure 4-2 shows plots of relaxation parameters NOE, R_1 and R_2 verses residue number for the C-terminal globular domain of *A.fulgidus* AglB-S2 collected at two magnetic fields, 600 and 700 MHz respectively. The red dotted lines for NOE (0.6 and below) and R_2 (20 and above) for the data collected at 600 MHz show the cut off points for residues experiencing fast internal motion at the picosecond to nanosecond timescale and domain motion at the microsecond to millisecond time scale respectively. These residues having peak intensities of 0.6 and below, and peak intensities of 20 and above were removed for estimating the global diffusion tensor vide infra. The data used for plotting this information is tabulated as Appendices B and C respectively and was used for model free analysis.

Backbone Dynamics

[^{15}N] uniformly labeled protein was expressed in M9 media supplemented with 4.0g/L glucose and 1.0g/L ^{15}N- ammonium chloride (Isotec) as the sole source of carbon and nitrogen respectively. Sample was prepared as before and measurements taken on the same instrument.

Using ^{15}N relaxation, data for backbone dynamics for three relaxation parameters; spin-lattice relaxation rate constants (R_1), spin-spin relaxation rate constants (R_2) and steady-state heteronuclear ^{1}H-^{15}N NOE's were collected by taking measurements at 308 K at 600MHz and 700MHz respectively.

Of the total assigned peaks, 139 peaks were positively quantified and used for relaxation studies. The relaxation data for the three parameters are described herein within the framework of the model-free formalism and data was processed using the in-house Red Hat-Linux platform on the DELL workstation. In order to apply model free analysis method, the error in the relaxation results was fixed at 10% to avoid a more stringent condition. Initial data input files were formatted as per requirement of Fast Model free program(Cole and Loria, 2003). Input parameters were derived by first estimating the global rotational time (τ_m) using the program r2r1_tm hosted at (http://www.palmer.hs.columbia.edu/software/diffusion.html). A seed of 101 amino acids which were free from high R_2 (values greater than 20) and low NOE (values less than 0.6) were used for estimating τm at 600MHz (Figure 4-2). After estimating τm, diffusion tensors were estimated using the programs r2r1_diffusion and quadric_diffusion at Arthur Palmers website (http://www.palmer.hs.columbia.edu/software/diffusion.html). The data is shown in Tables 4-1 and 4-2 respectively.

Table 4-1 r2r1_diffusion results of *A.fulgidus* AglB-S2 at 600MHz (101 amino acids).

Model	D_{iso}(1/s)	$D_{\parallel}/D_{\perp}$	χ^2(red)	F
Isotropic	4.045 e7	-	817.81	-
Axial	3.539 e7	1.400±0.074	758.83	3.59

Table 4-2 Quadric_diffusion results of *A.fulgidus* AglB-S2 at 600MHz (101 amino acids).

Model	D_{iso}(e-7/s)	$D_{\parallel}/D_{\perp}$	$2D_{zz}/(D_{xx}+D_{yy})$	D_{xx}/D_{yy}	χ^2(red)	F
Isotropic	1.425	-	-	-	10.47	-
Axial	1.434	1.174±0.012	-	-	8.09	10.78
anisotropic	1.432		1.165±0.013	0.948±0.014	8.14	0.70

Model Free Analysis

Among the 157 non-proline residues in the C-terminal domain of *A.fulgidus* AglB-S2, a total of 132 residues were used for the Lipari-Szabo-type model-free analysis. The remaining 25 residues were not used, due to low peak intensities and peak overlapping. The relaxation parameter values obtained at 600 and 700 MHz ^{1}H frequencies, and the coordinates of the crystal structure of *A.fulgidus* AglB-S2 were formatted as described (Tjandra et al., 1995; Tjandra et al., 1996). The initial parameter values of the global diffusion tensor were estimated from the R_2/R_1 ratio for a subset of residues (101 residues) in a rigid part of the molecule, using the programs *Pdbinertia* (ver. 1.11), *r2r1_tm*, and *quadric_diffusion* (ver. 1.13, A. G. Palmer III, Columbia University, available at http://www.palmer.hs.columbia.edu/software/diffusion.html). An axially symmetric diffusion

tensor was selected by comparison with the isotropic diffusion and fully anisotropic diffusion tensors, on the basis of a partial F-test (Table 4-3). The overall shape of the *Af*AglB-S2 molecule was approximated by a prolate ellipsoid with a diffusion tensor ratio of $D_{parallel}$ to $D_{perpendicular}$, $D_{\parallel}/D_{\perp} = 1.2$.

Local motional models for the backbone ^{1}H-^{15}N bonds were selected as described (Mandel et al., 1995). The R_1, R_2, heteronuclear NOE values, and the axially symmetric diffusion tensor were input in the setup FAST-Modelfree GUI (Cole and Loria, 2003), and the program was executed using the default set of parameters, except the cutoff value for S^2 was set to 0.5. After consecutive iterations, four motional models were adopted to describe the internal motion of each ^{1}H-^{15}N bond vector: Model 1: (S^2), Model 2: (S^2, τ_e), Model 3: (S^2, R_{ex}), Model 4: (S^2, τ_e, R_{ex}). The square of the generalized order parameter (S^2) describes the degree of spatial restriction of a ^{1}H-^{15}N bond vector in a molecular reference frame, the correlation time for fast internal motion (τ_e) is the effective correlation time for the internal fast motion of a ^{1}H-^{15}N vector, and the chemical exchange-induced relaxation rate (R_{ex}) is a phenomenological exchange term that accounts for chemical exchange processes that contribute to the decay of transverse magnetization during the CPMG pulse train. These local motional parameters are summarized in Figure 4-3. The final optimized parameters of the rotational diffusion model were $\tau_m = 12.07 \pm 0.08$ ns, and $D_{\parallel}/D_{\perp} = 1.21 \pm 0.05$. The S^2 values ranged from 0.598 ± 0.039 to 0.983 ± 0.027, with an average of 0.938 ± 0.034. The residues that contained R_{ex} were of interest in this study. Ten residues were fit to Model 3: Trp468, Asp470, Tyr471, Asn473, Tyr477, Val478, Ala479, Lys480, Lys535, and Ile537. Four residues were fit to Model 4: Met434, Tyr463, Gln489, and Thr534. The R_{ex} values ranged from 4.33 ± 1.6 s^{-1} for Asn473 to 10.17 ± 1.8 s^{-1} for Val478. A plot of the model free results is shown in Figure 4-3 A with a structural map of the values on 4-3 B respectively.

Results from Model Free analysis.

Results from the estimation of the global correlation time τ_m using r2r1_tm showed that the molecule tumbles with a correlation time of 11.6 ns which was in good agreement with the results from the more rigorous quadric_diffusion programme as shown in Table 4-3.

Table 4-3 Rotational diffusion models using 600MHz quadric_diffusion derived data.

Model [a]	τ_m (ns) [b]	$D_{\|}/D_{\perp}$	D_x/D_y	χ^2_{red}	F [d]	P [e ▽]
Isotropic	11.70 ± 0.02	-	-	10.47	-	-
Axially Symmetric	11.62 ± 0.02	1.174 ± 0.012	-	8.09	10.78	< 0.0001
Anisotropic	11.64 ± 0.02	1.165 ± 0.013 [c]	0.948 ± 0.014	8.14	0.70	0.50

[a] Evaluated by the program *quadric_diffusion,* using the relaxation data obtained at 600 MHz.

[b] Effective correlation time, $\tau_m = 1/6D_{iso}$. $D_{iso}=(D_{\|} + 2D_{\perp})/3$.

[c] $D_{\|}/D_{\perp} = 2D_z/(D_x + D_y)$.

[d] The significance of the reduction in χ^2_{red} for model selection between the two nested models. The two degrees of freedom for the comparison of the Isotropic and Axially Symmetric models are $F_{4\text{-}1,101\text{-}4} = F_{3,97}$, and those for the comparison of the Axially Symmetric and Anisotropic models are $F_{6\text{-}4,\ 101\text{-}6} = F_{2,95}$, where the number of parameters are 1, 4, and 6 for the Isotropic, Axially Symmetric, and Anisotropic models, respectively. The number of residues used for the model selection was 101.

[e] Probability that the reduction of χ^2_{red} is randomly achieved.

Propagation of error expression $\tau_{m\,iso} = (6D_{iso})^{-1}$ and $\frac{\delta\tau_{m\,iso}}{\tau_{m\,iso}} = \sqrt{\left(\frac{-1\delta D_{iso}}{D_{iso}}\right)^2}$ was used to calculate error in $\tau_{m\,iso}$.

These initial values were then used as part of the input information in the Model Free program implemented in Fast Model Free graphical user interface (GUI)

▽ Calculated using the online resources at http://graphpad.com/quickcalcs/PValue1.cfm

Figure 4-3 Model free analysis of the C-terminal globular domain of *A.fulgidus* AglB-S2

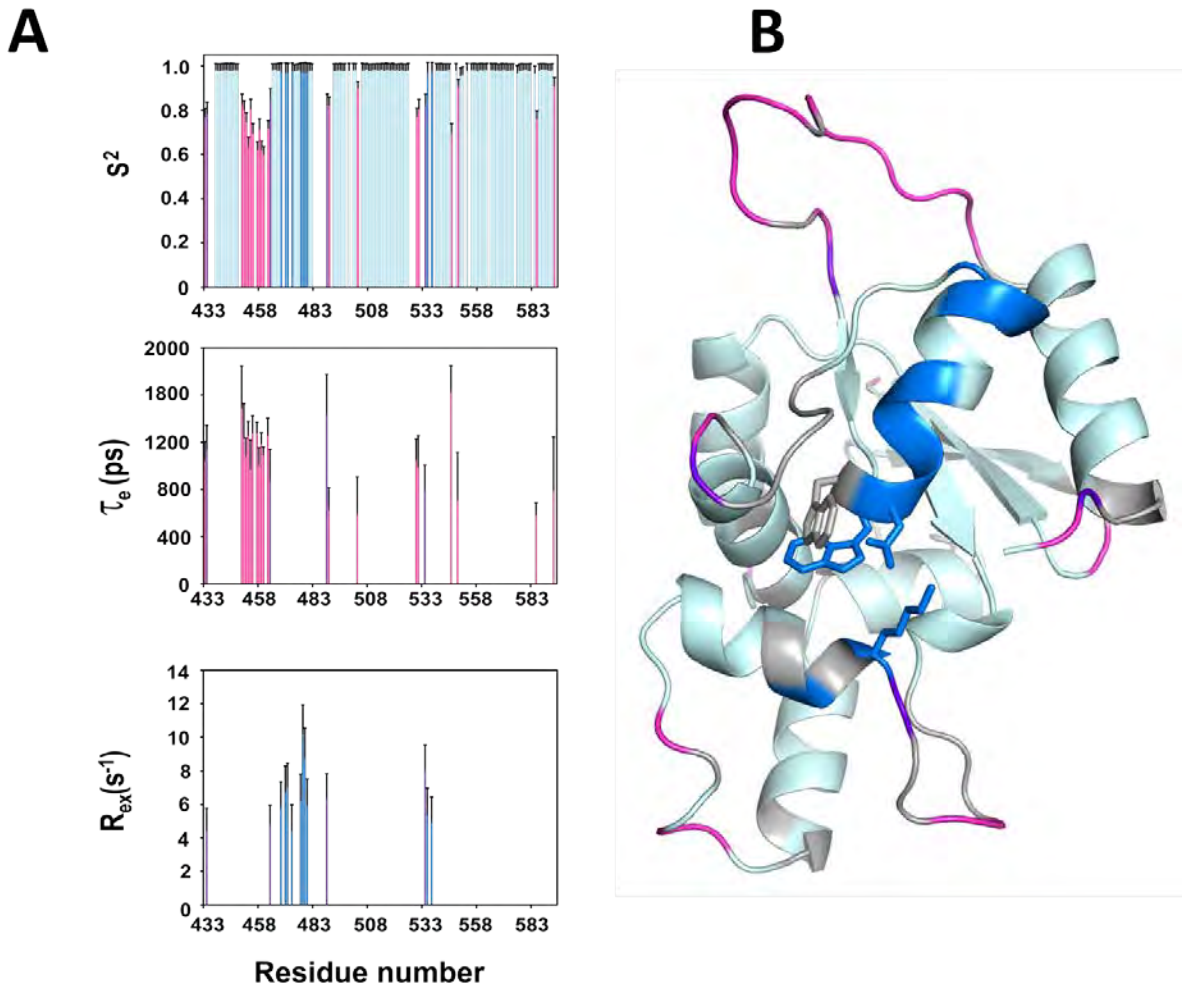

Figure 4-3 shows order parameters as a function of residue number for the C-terminal globular domain of *A.fulgidus* AglB-S2. (A) The top left panel shows the generalized order parameter S^2 which is a measure of mobility of the N-H bond. The middle left panel shows the internal correlation time in the ps-ns timescale (τ_e),and the lower left panel shows the chemical exchange contribution (R_{ex}) all plotted as functions of the residue number in the protein sequence. The bars are color-coded according to the model selection: Model 1 (S^2), *cyan*; Model 2 (S^2, τ_e), *pink*; Model 3 (S^2, R_{ex}), *blue*; Model 4 (S^2, τ_e, R_{ex}), *purple*. (B) shows a plot of the model free parameters in (A) on the structure of the C-terminal globular domain of *A.fulgidus* AglB-S2. Residues with no information were colored gray. The data was generated using Model version 4.16(Mandel et al., 1995) implemented in Fast Model free version 1.01(Cole and Loria, 2003).

Constant-time CPMG Relaxation Dispersion analysis of *A.fulgidus* AglB-S2 at 600MHz,308 K.

NMR Relaxation Measurements

A 0.3mM sample of uniformly ^{15}N labeled sample of *A.fulgidus* AglB-S2 was used to collect data for relaxation dispersion experiments at 600MHz and 308K as before. The ^{15}N R_2 relaxation dispersion data were recorded using the CPMG (Carr-Purcell-Meiboom-Gill) pulse train with a constant time of T_{CPMG} = 40 ms (Tollinger et al., 2001). A series of 2D ^{1}H-^{15}N HSQC-based spectra were recorded with varying ν_{CPMG} frequencies (25, 50 × 2, 75, 100, 125, 150, 175 ×2, 200, 325, 400 × 2, and 500 Hz). Data processing, peak picking, intensity integration, and curve fitting were performed using the program NMRPipe (Delaglio et al., 1995).

Dynamics of a protein may be defined as no exchange (model 1 (Bieri and Gooley, 2011)) where the effective relaxation rate R_2, $R_2^{eff} = R_2^0$ 7

where R_2^0 is the effective relaxation rate at infinite νCPMG.

The effective relaxation rate, R_2^{eff} , was extracted from the cross peaks according to:

$$R_2^{eff} = \frac{1}{T_{CPMG}} \ln \frac{I_o}{I_{\nu CPMG}} \quad 8$$

where T_{CPMG} is the constant delay time=40 ms, I_o and $I_{\nu CPMG}$ are the peak intensities of the reference spectrum $T_{CPMG} = 0$ and peaks at the respective CPMG frequency.

In the case of model 2 (2-state fast-limit exchange (Bieri and Gooley, 2011)), the effective relaxation rate constant R_2, R_2^{eff} , is given by the expression:

$$R_2^{eff} = R_2^0 + \frac{\Phi}{k_{ex}}\left[1 - \frac{4\nu CPMG}{k_{ex}} tanh\left(\frac{k_{ex}}{4\nu CPMG}\right)\right] \quad 9$$

where $\Phi = p_a * p_b * \delta\omega^2$, 10

p_a and p_b are the populations of the major and minor state respectively and $\delta\omega$ is the chemical shift difference between states. In expression 10, the population p_b and $\delta\omega$ cannot be separately determined, so only Φ, R_2^0 and k_{ex} can be extracted. The above condition is for a two state exchange model as outlined by Bieri and Gooley et al, 2011.

Effective transverse ^{15}N relaxation rate, R_2^{eff}, as a function of the CPMG frequency, ν_{CPMG}, was calculated using the peak intensity in the reference spectrum recorded with T_{CPMG} = 0 ms and the peak intensity at ν_{CPMG} in the program *NESSY* version 12.2.1 (Bieri and Gooley, 2011). The programme was executed with default values unchanged. R_2^{eff} dispersion profiles of 132 cross peaks were fitted to no exchange model or two-site fast exchange model. Model selection was based on Akaike information criteria with second order correction for small sample sizes (AICc). Sixty-two residues were better fitted to the two-site fast exchange model. For subsequent analysis, only residues having a difference in R_2^{eff} greater than 4 s^{-1} between the lowest and highest CPMG frequencies were retained. Thus, 24 residues (M434, K446, E447, N448, Q452, Y454, D462, Y463, V465, S467, W468, N473, Y477, A482, V483, A492, D494, R514, V519, E523, L525, T534, K535, and K582) were selected as Model 2 for plotting (Figure 4-4A) and mapping on the structure (Figure 4-4 B). From the 24 residues, 7 residues (S467, W468, N473, Y477, A482, T534, K535), contained in the THL segment and the kinked helix, were clustered and simultaneously fit to the two-site fast exchange model (model 2), and a global k_{ex} was calculated (1,834 ± 88 s^{-1}). Data collected for the 7(cluster analysis) and 24 (model 2) amino acids is presented in Table 4-4 and Appendix E-1 for constant-time CPMG relaxation dispersion analysis respectively.

Results from Relaxation Dispersion by NESSY Cluster analysis.

Out of 62 residues assigned to model 2 by the programme NESSY, 24 residues were ultimately used to be representative of 2-state fast-limit motion and of these residues, those found in the THL and DK motif segments were further clustered as these were thought to have a concerted motion within the molecule and since they were residues of interest to this study. After analysis as shown in Table 4-4, a single k_{ex} value was obtained to represent motion in the plastic region encompassing the Ser/Thr binding pocket.

Table 4-4 Cluster analysis of THL-DKi motif residues by global fit to calculate single k_{ex}

Residue		R_2 (s^{-1})	err	k_{ex} (s^{-1})	err	Φ $(ppm/s)^2$	err (10^{-12})	χ^2
SER	467	21.31	0.42	1833.59	87.95	0.00112	7.45	21.05
TRP	468	24.93	0.32	1833.59	87.95	0.00063	4.66	18.17
ASN	473	21.75	0.22	1833.59	87.95	0.00081	4.40	31.18
TYR	477	24.21	0.35	1833.59	87.95	0.00089	5.92	23.47
ALA	482	17.37	0.22	1833.59	87.95	0.00101	4.95	106.06
THR	534	26.40	0.53	1833.59	87.95	0.00097	8.26	61.15
LYS	535	24.96	0.32	1833.59	87.95	0.00113	6.20	124.82

Residues in the turn helix loop (THL) and the invariant DKi motif segment were fit to model 2 for cluster analysis. A global k_{ex} of 1834±88 (s^{-1}) was obtained implying motion of about 0.5ms within the μs-ms time scale as suggested by model free analysis.

Figure 4-4 Constant time CPMG relaxation dispersion analysis of *A.fulgidus* AglB-S2

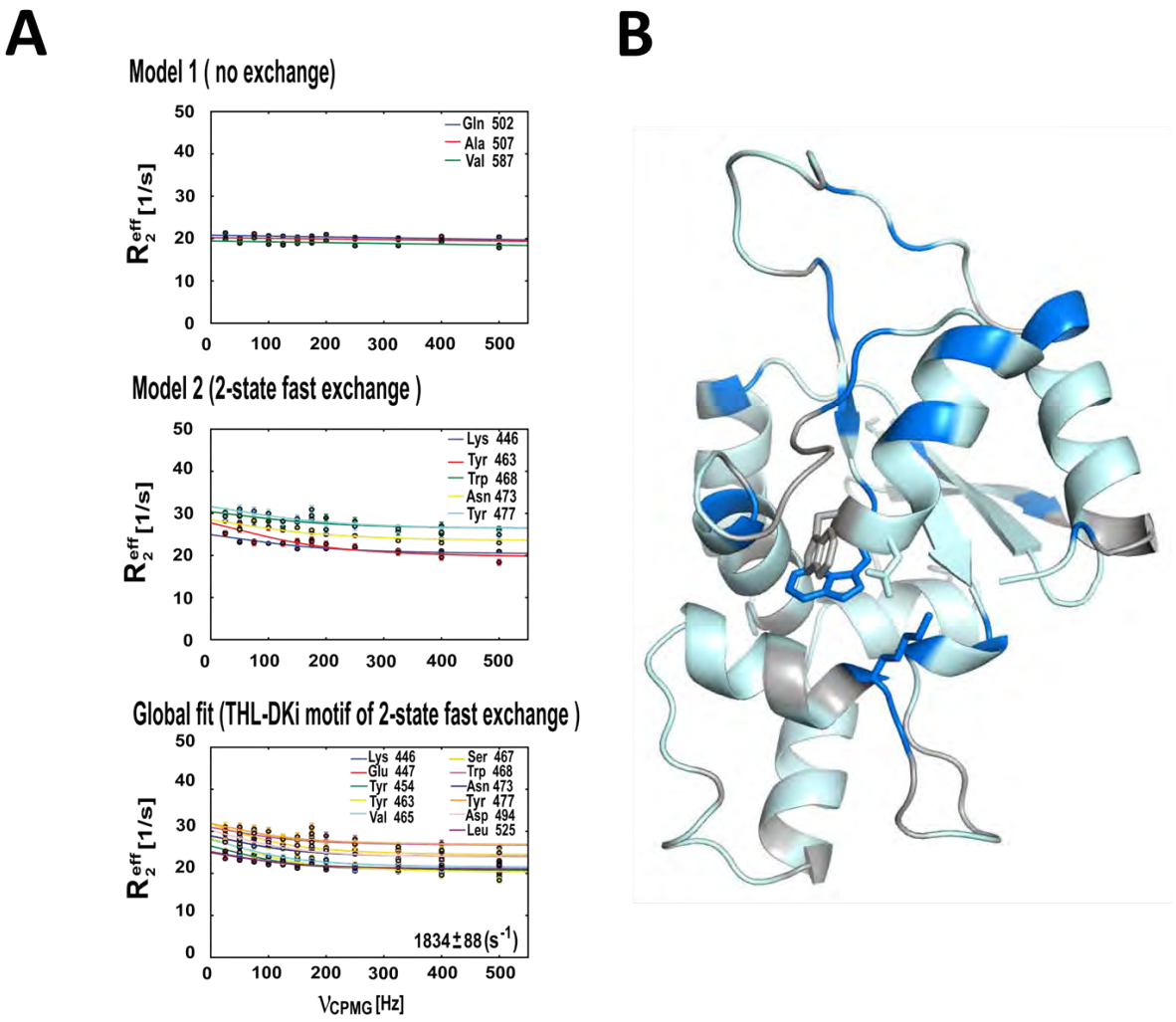

Figure 4-4 shows the plots of effective R_2 (R_2^{eff}), values at; (A) top left panel, infinite ν_{CPMG} (no exchange), middle panel, residues experiencing 2-state fast-limit exchange and bottom panel, representative residues covering the THL-DK motif segment experiencing a concerted dynamic motion that contributes to the formation and collapse of the Ser/Thr binding pocket of the active site of the enzyme with a global k_{ex} of 1834±88 (s^{-1}). (B) The information for model 1 ; no exchange (*cyan*) and model 2; 2 state fast exchange motion (*blue*) was plotted on the structure of *A.fulgidus* AglB-S2 shown on the right panel. Residues with no information were colored *gray* as in Figure 4-3.

Figure 4-5 Multiple structural alignment of the C-terminal globular domains of OSTs, *Cj, Pf, Ph, Af*S1*and Af*S2 and comparison with NMR results of *A.fulgidus* AglB-S2

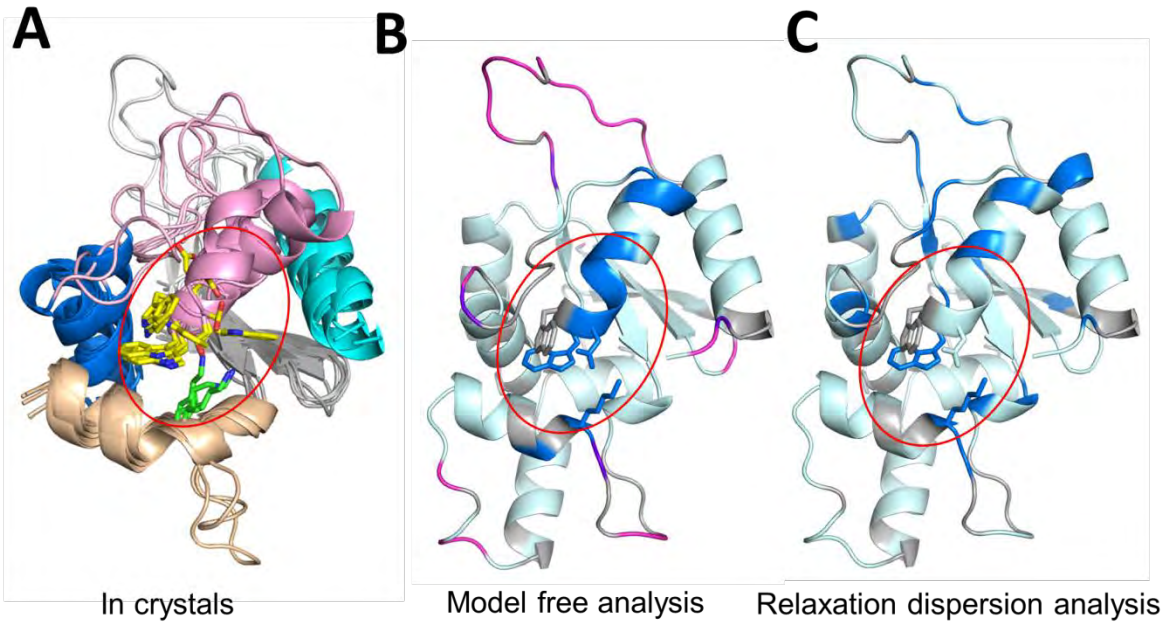

Figure 4-5 (A) shows the superposition of the CC structural unit in the five crystal structures of the C-terminal globular domains of the AglB/PglB proteins (*Af*AglB-S1, *Af*AglB-S2, *Cj*PglB, *Pf*AglB-L, and *Ph*AglB-L). (B and C) show maps of NMR relaxation experiments results as in Figures 4-3 and 4-4. The red oval shows the plastic (X-ray) or alternatively called mobile region (NMR). The same color code is used as in Figures 3-1 to 3-3 that is the WWDYG motif side chains in yellow and the side chains forming the DK/DKi/MI motif are shown in green. Structural superpositioning was done by GASH(Standley et al., 2005) software. The structures generally superimpose in the main frame but small variations in the WWDYG motif segment were of interest to this study. The overall structure of the *Af*AglB-S2 was overlaid in *pale cyan,* as a visual aid**.**

Discussion

The values of the reduced chi squared value (χ^2_{red}) and F statistic in Table 4-3 were used to discriminate between isotropic and axially symmetric and axially symmetric and anisotropic models respectively. The results were as expected according to the hypothesis and deducing from the crystal structure topology, the solution data was in good agreement. Using these results, the axially symmetric model was chosen to represent the global and local motions of the molecule as defined by the tensors. Results were analyzed using the Model Free software version 4.16 and the plots are shown in Figure 4-3. Curve fitting of representative model 1, 2 and global fit residues by NESSY version 12.2.1 were plotted and shown in Figure 4-4. From these it was apparent that there was observable motion on the micro to millisecond time scale and such motions were important in contributing to the catalysis. A comparison of the crystal structures and results obtained from NMR studies (Fig. 4-5) were in agreement that the region under study was plastic in the crystal structures and in solution mobility was observed as well. Overall, the results got from solution NMR studies indeed shed more light on the inherent flexibility exhibited by the archaeon oligosaccharyltransferase. The chemical shifts and the results of the model-free analysis of the C-terminal globular domain of *Af*AglB-S2 have been deposited in the Biological Magnetic Resonance Data Bank, with the accession number 18477 and have been tabulated as Appendices A, B, C and D for Chemical shifts, Relaxation analysis R_1, R_2, NOE at two fields and backbone dynamics recorded at two fields using Model free formalism approach respectively.

Chapter 5 - Disulfide bond engineering

Restriction of Flexibility by an Engineered Disulfide Bond

The aim of this experiment was to provide biochemical evidence of the importance of flexibility of the WWDYG-motif segment (observed in NMR experiments) to support catalysis of the oligosaccharyltransferase.

To achieve this objective, *P.furiosus* AglB-L was selected for protein disulfide bond engineering because it has the unique property of being thermostable and is able to tolerate mutations which otherwise would be lethal in many species. This is shown in Figure 5-1; the disulfide was introduced by site directed mutagenesis using the TOYOBO mutagenesis kit. Sequence alignment of *C.lari* PglB and *P.furiosus* AglB-L (Figure 5-1 A and B) was used to infer the desired position of the disulfide bond far enough from the enzymes active site not to interfere with activity but proximal enough to induce restriction of flexibility of the WWDYG-motif segment within the C-terminal globular domain. Figure 5-2 shows a close up view of the position of inferred disulfide bonds.

Figure 5-1 Engineering a disulfide bond in *P.furiosus* AglB-L using *C.lari* PglB as a template

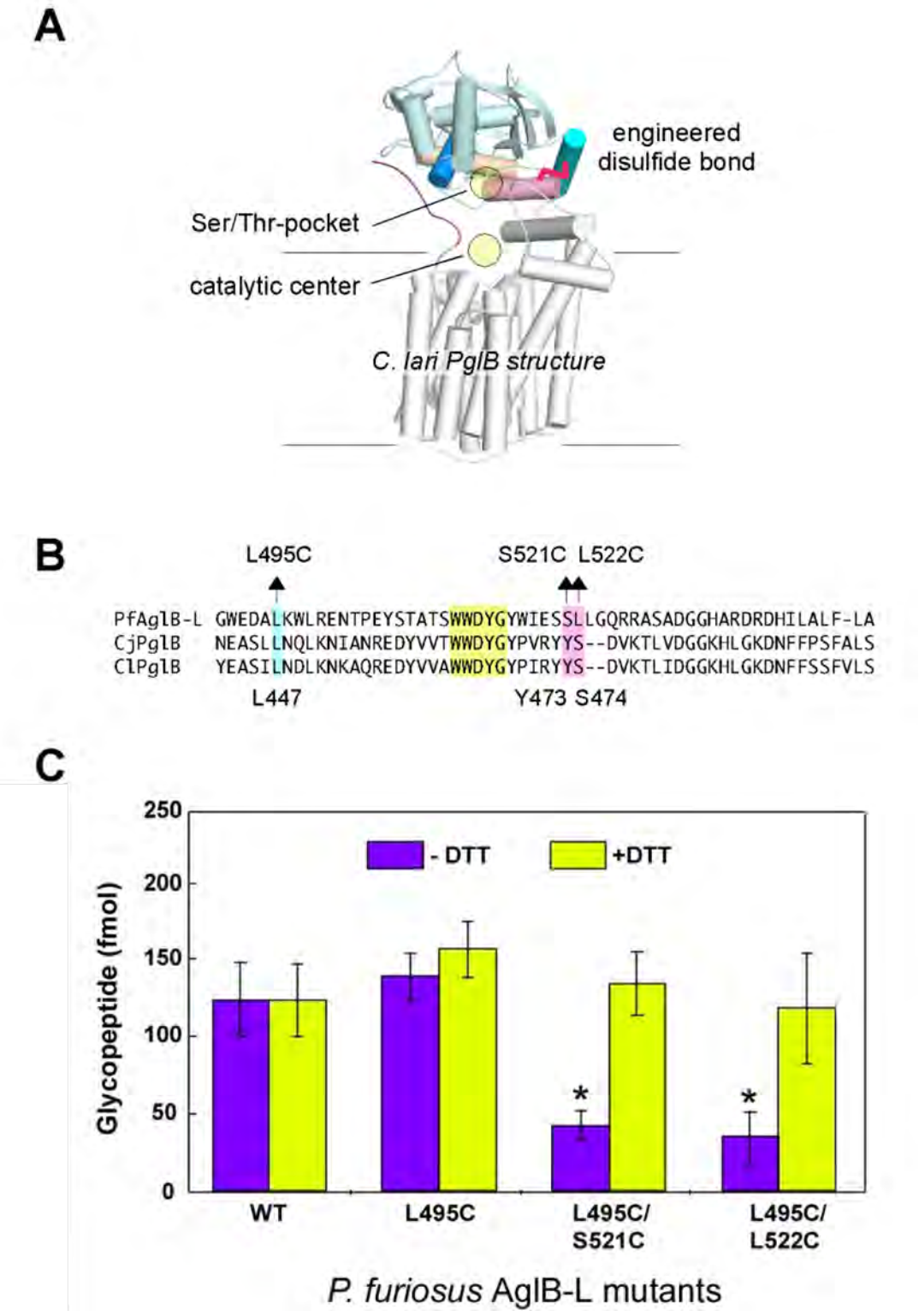

Figure 5-1 (A) Schematic diagram of the crystal structure of the full-length *C.lari* PglB (PDB entry, 3RCE). Magenta sticks represent the position of the engineered disulfide bond. The color of the α-helices is the same as in Figures 3-1 to 3-3. The two shaded circles indicate the position of the +2 Ser/Thr pocket in the C-terminal globular domain and that of the catalytic center in the N-terminal transmembrane region. (B) Multiple sequence alignment and the cysteine mutations for linking the two α-helices. The positions of the mutations to cysteine residues in *Pf*AglB-L were determined by a sequence alignment, with *Cl*PglB as the template. (C) Specific OST activity of the double cysteine mutants of *Pf*AglB-L, with and without the preincubation with 5 mM DTT, before correcting for disulfide bond content. The addition of DTT cleaved the engineered disulfide bond. The specific OST activities of the wild type (WT) and the single cysteine mutant are also shown, as a control. The specific activity (S) was calculated by using the equation $S = A/P$, where A is the mean OST activity, and P is the mean amount of protein. The standard deviation (σ_S) of the specific activity was calculated by using the propagation of error equation, $(\sigma_S/S)^2=(\sigma_A/A)^2+(\sigma_P/P)^2$. The error bars represent the mean ± S.D. calculated from three independent protein quantifications and three independent OST assays. Data were compared to the wild type was assessed using an unpaired two-tailed (Welch's) t-test, assuming unequal variance. P-values were 0.022 and 0.017 for the L495C/S521C and L495C/L522C mutants, respectively.

Experimental methods employed

Experiments were done on the wild type and the cysteine-single and double mutants as follows:

Preparation of Disulfide-Stiffened AglB and OST Assay

The full-length *P.furiosus* AglB-L mutants containing an engineered disulfide bond were generated using a KOD mutagenesis kit (Toyobo). The C-terminal His-tagged mutants in the *E. coli* membrane fractions were partially purified using a Ni-Sepharose column, in the same manner as the wild type (Igura and Kohda, 2011a). All buffer solutions used for the purification were pH 7.2 and did not contain any reducing agents, to facilitate the disulfide bond formation. The determination of the protein amount and the OST assay were performed

as described (Igura and Kohda, 2011a; Matsumoto et al., 2012), except for the following minor modifications: Under the conditions to maintain the engineered disulfide bond, the proteins were pre-incubated without DTT for 1 h at 310 K, and the assay reaction mixture lacked DTT. Under the conditions to cleave the disulfide bond, the proteins were incubated with 5 mM DTT for 1 h at 310 K, and the reaction mixture contained 2 mM DTT. The number of sulfhydryl groups was determined by the maleimide-conjugated polyethylene glycol (Mal-PEG) alkylation method (Makmura et al., 2001). Proteins were TCA-precipitated, washed with acetone, and dissolved in 100 mM Tris-HCl, pH 7.5, 1% SDS, and 1 mM Mal-PEG (NOF Corp). Alkylated *Pf*AglB-L proteins were fractionated by SDS–PAGE and visualized by western blotting with anti-His tag antibodies.

Quantification of protein and oligopeptides was done using pixel intensity on the LAS 3000(Fuji) instrument (Figure 5-3) using the Green LED's in Pro Mode Green (520nm EPI) and using the following settings, 575DF20 filter and 0.85 iris.

Figure 5-2 Close-up Views of the Insertion Site of the Engineered Disulfide Bond

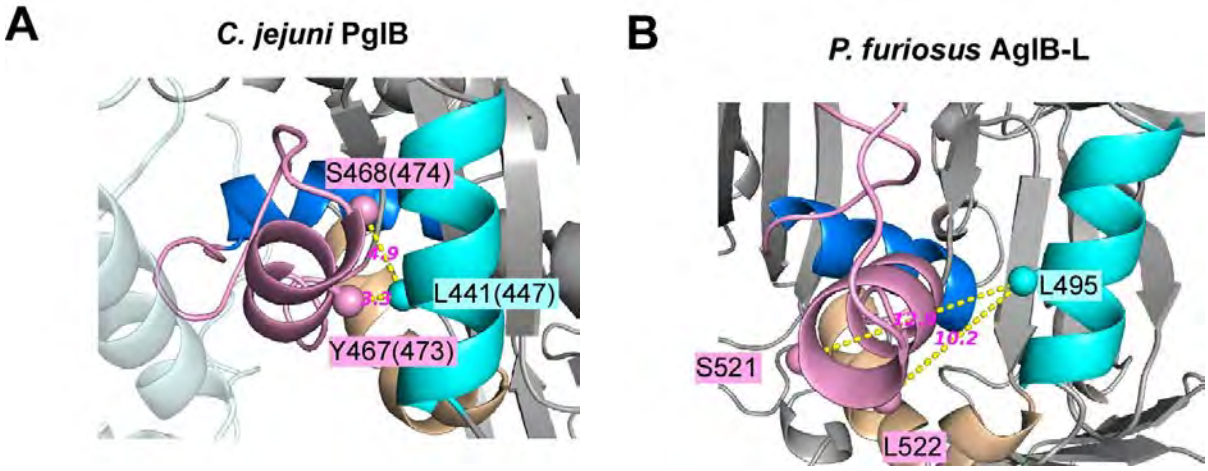

Figure 5-2 *Pink* and *cyan* spheres represent the C_β carbons of the selected side chains, and *yellow* dotted lines indicate the distance in Å between them, (A) in the *C.jejuni* PglB structure, and (B) in the *P.furiosus* AglB-L structure. The *C.jejuni* PglB was used for illustration, because of its higher resolution than the *C.lari* PglB structure. The residue numbers in parentheses are those for *C.lari* PglB. Note that the C_β-C_β distances are too long (> 10 Å) to form a disulfide bridge in the crystal structure of *Pf*AglB-L, due to the artificial displacement of the *pink* α-helix by the crystal contact effects. The successful disulfide bond formation (Figure 5-3) indicated that the conformation of the THL segment in the *Pf*AglB-L in solution was similar to those found in the *Af*AglB-S2, *Cj*PglB, and *Cl*PglB crystals.

Figure 5-3 Protein quantification, OST activity and disulfide bond content of mutants

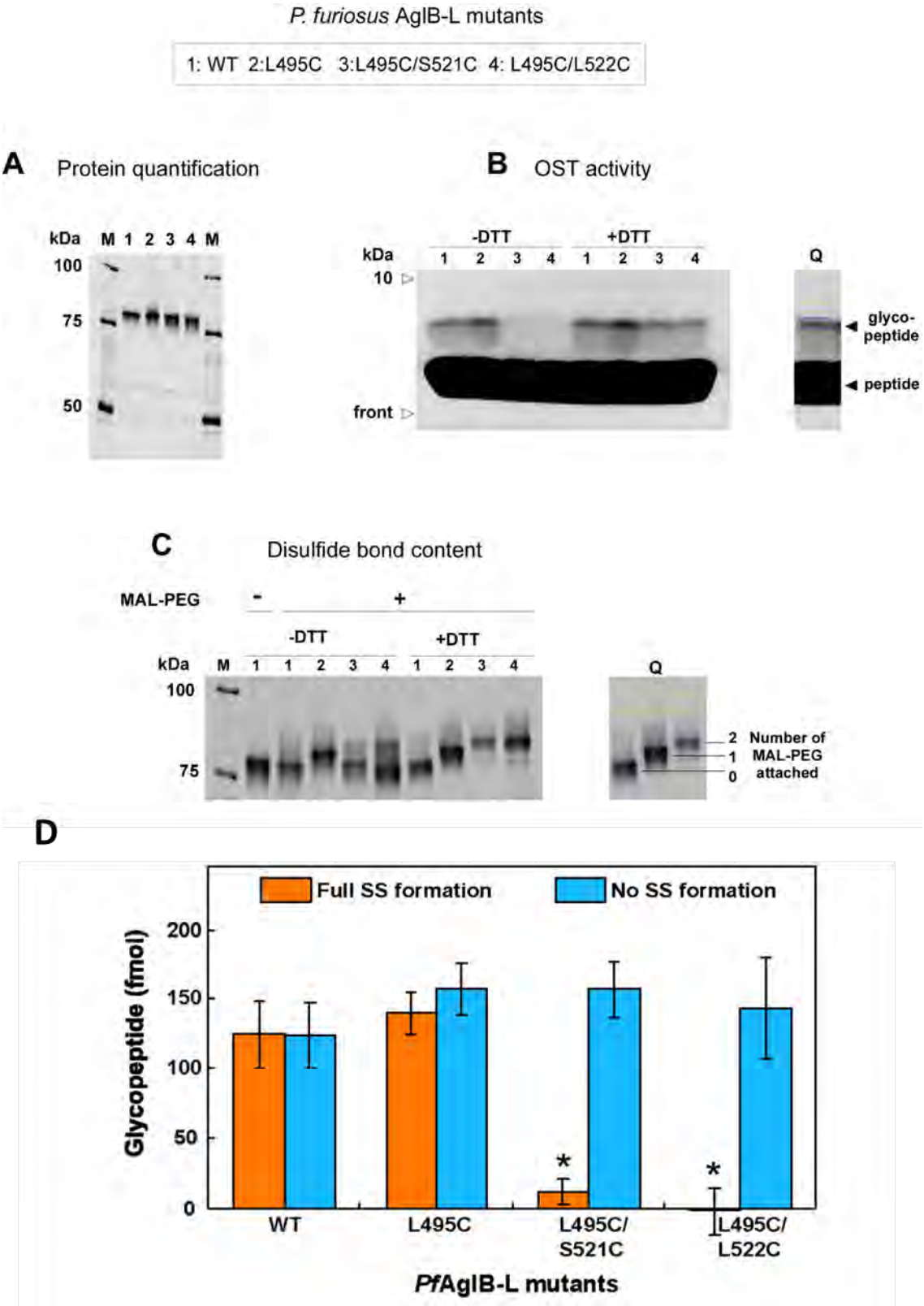

Figure 5-3 Protein Quantification, OST Activity, and Disulfide Bond Content of the *Pf*AglB-L Mutants. The images show results of the experiment. Lane 1, wild type; lane 2, L495C; lane 3, L495C/S521C; lane 4, L495C/L522C; lane M, molecular weight markers; lane Q, area of intensity integral. (A) Western blot image for quantitative measurement of the protein amount. TCA (trichloroacetic acid) precipitation was performed before SDS-PAGE(Matsumoto et al., 2012). The fluorescence of the secondary antibodies was detected at 800 nm. The His-tagged protein standards (6xHis Protein Ladder, QIAGEN) were used as internal standards. (B) SDS-PAGE image for the determination of the OST activity. The fluorescence of the TAMRA dye attached to the N-terminus of the peptide was detected at 575 nm. The positions of the MW markers and the tracking dye front are indicated. +DTT means the incubation with 5 mM dithiothreitol (DTT) before the assay. The pixel intensity of the glycopeptide bands (*blue* closed curve) was integrated and corrected by subtracting the isometric background (*yellow* closed curve), and was converted to the molar amount using the published calibration curve (Kohda et al., 2007). (C) The disulfide bond content was determined by counting the number of free sulfhydryl groups. +MAL-PEG means the reaction with maleimide-polyethylene glycol (MAL-PEG, Mwt = 2,000 Da). +DTT means the incubation in the presence of 5 mM DTT before the TCA precipitation and the reaction with MAL-PEG. The pixel intensity of the band corresponding to *Pf*AglB-L (*blue* rectangles) was integrated and corrected by subtracting the isometric background (*yellow* rectangles). Note that the wild-type *Pf*AglB-L contains one naturally occurring disulfide bond between Cys637 and Cys657 in the IS structural unit in the C-terminal globular domain, but this disulfide bond was stable in the presence of 5 mM DTT and didn't react with MAL-PEG, and thus could be safely ignored in the present experiments. This was justified by the absence of a band shift of the wild-type *Pf*AglB-L after the MAL-PEG reaction, irrespective of the preincubation with DTT. (D) Specific OST activities of the double cysteine mutants of *Pf*AglB-L, with full formation and without formation of the engineered disulfide bond. The specific OST activities of the wild type (WT) and the single cysteine mutant are also shown, as controls. The error bars represent the mean ± S.D. calculated from three independent protein quantifications, three independent OST assays, and three independent disulfide bond content determinations. Data were compared to the wild type and was assessed using an unpaired two-tailed t-test, assuming unequal variance. P-values were 0.0074 and 0.0024 for the L495C/S521C and L495C/L522C mutants, respectively.

Discussion

The engineering of the disulfide bond in *P.furiosus* AglB-L was very intuitive and provided classical biochemical evidence for presence of a mobile region and indeed the necessity of collapsing and reformation of the Ser/Thr binding pocket in the activity of oligosaccharyltransferase. Although the crystal structure of *P.furiosus* AglB-L experienced a distortion in the N-terminal part of the helix encompassing the WWDYG motif segment (Igura et al., 2008; Maita et al., 2010), transfer of the inferred disulfide bond to *P.furiosus* using residue positions from the full length structure of *C.lari* PglB was in practice feasible. With these results it was conclusive to say that there is indeed flexibility in the region of the active site of oligosaccharyltransferase and that this flexibility first observed as a mere crystal contact phenomenon but later proved to be inherent plasticity of the residues forming part of the active site was important in maintaining the collapse and reformation of the Ser/Thr binding pocket identified by Lizak et al, 2011. Mere introduction of cysteine mutations did not affect protein stability but rather the actual formation of the disulfide bond greatly impaired activity which was only restored upon disulfide bond cleavage by DTT.

Chapter 6 - Summary of Results

Results

The following summarizes the findings according to the research carried out and reported in this thesis in a sequential fashion.

Overall Structures of the C-terminal Globular Domains of *P.horikoshii* AglB-L and *A.fulgidus* AglB-S2

Among the catalytic subunits of oligosaccharyltransferases, *Pyrococcus horikoshii* AglB-L (*Ph*AglB-L) is one of the largest, and *Archaeoglobus fulgidus* AglB-S2 (*Af*AglB-S2) is one of the smallest, in the three domains of life (Matsumoto et al., 2012). In this study, the crystal structures of the C-terminal globular domains of *Ph*AglB-L and *Af*AglB-S2 at 2.7 Å and 1.94 Å resolutions, respectively (Table 2-1) were determined. The C-terminal domain of the *Pyrococcus* structure consists of four structural units, called CC (*c*entral *c*ore), IS (*i*nsertion), P1 (*p*eripheral *1*) and P2 (*p*eripheral *2*), whereas that of the *Archaeoglobus* structure comprises only the CC structural unit (Figures 2-8 A and B). The CC unit mainly consists of α-helices. The IS unit is an antiparallel β-barrel-like structure inserted into the CC unit. The P1 and P2 form peripheral structures, mainly consisting of β-strands, and surround the CC unit. The spatial orientation of the C-terminal globular domains relative to the lipid membranes was modeled on the full-length structure of *C.lari* PglB as a template (Figure 2-9 A). As observed, the IS, P1 and P2 structural units in the *Ph*AlgB-L structure, as well as the IS unit in *Cl*PglB, do not undergo severe steric hindrance to the lipid membranes. The close up view of the active site for *C.lari* PglB is shown in Figure 2-9 B. Six crystal structures in their respective asymmetric units were analyzed for presence or absence of crystal contacts (Figure 2-10).

The newly determined structures were compared with the previous structures of the C-terminal domain: one orthologous protein from the same genus *Pyrococcus* (Igura et al., 2008), and one paralogous protein from the same species *Archaeoglobus fulgidus* (Matsumoto et al., 2012). Reflecting the high sequence identities of about 70%, the overall structures

superimposed well in each pair (Figures 3-1 A and B). However, a closer examination of the structures revealed significant local conformational differences in the CC structural unit. The region that exhibits large plasticity includes the WWDYG motif (*yellow* side chains) and the following α-helical and loop regions (*pink* backbones). Although the first Trp residue of the WWDYG motif occupies equivalent positions with the same side-chain orientation, the second Trp residue, marked by asterisks, adopts totally different side chain orientations in each pair. The same phenomenon occurs for the side chains of the Asp and Tyr residues in the WWDYG motif. Such large conformational differences between highly homologous proteins are rarely observed. One rational explanation is a crystallographic artifact due to the molecular contacts between neighboring molecules in the crystal lattice. In fact, direct inter-molecular interactions were identified in these crystals (Figure 2-10 A, B and C).

Conformation of the Turn-Helix-Loop Segment Free from the Crystal Contact Effects

For convenience, the plastic region was re defined as the "Turn-Helix-Loop" segment and referred to as such hereafter. The Turn-Helix-Loop (THL) segment consists of a turn structure including the WWDYG motif, the following α-helix, and a loop structure. This loop most likely interacts with the innermost monosaccharide residue of the N-glycan moiety of the substrate LLOs (Lizak et al., 2011). A method to find an undisturbed conformation of the THL segment, free from the crystal contact effects was searched. One possible method is to collect many crystal structures of homologous proteins, and then identify an isomorphic pair. In all possible combinations, the *Af*AglB-S2 structure determined in this study has the same THL segment conformation as that in the *Campylobacter jejuni* PglB structure, even though they only share 19 % sequence identity (Figure 3-2). This is clearly exemplified by the same side-chain orientation of the second Trp residue in the WWDYG motif. In addition, the conformation of the THL segment in the structure of the *Campylobacter jejuni* PglB protein is identical to that in the *Campylobacter lari* PglB protein (Figure 3-3). This is not surprising, given the high sequence identity (52 %) between the two *Campylobacter* PglB's, but it

indicates that the C-terminal globular domain is structurally independent of the N-terminal transmembrane region. In accordance with the superimposable conformation of the THL segment, no obvious crystal contacts involving the WWDYG motif are present between neighboring molecules in both the *Af*AglB-S2 and *Cj*PglB crystals (Figure 2-10 D and E). In the *C.lari* PglB crystal, the WWDYG motif is sequestered within the protein molecule and does not contact the other molecules (Figure 2-9 B and Figure 2-10 F). Instead, the WWDYG motif interacts with a peptide substrate, but the peptide binding does not induce any notable conformational change in the C-terminal globular domain. In summary, irrespective of the sequence identities, the conformations of the THL segment in the three crystal structures, *Af*AglB-S2, *Cj*PglB, and *Cl*PglB (Figures 2-10 E to F), are identical, but those in the other crystal structures, *Af*AglB-S1, *Ph*AglB-L, and *Pf*AglB-L (Figures A to C), are different from one another, due to the distortion by the crystal contact effects. Taken together, these results indicated that the THL segment has intrinsic structural plasticity. The superimposable conformation most likely represents the resting state and peptide substrate-bound conformations of the plastic THL segment of the OST enzymes.

NMR Evidence for the Mobility of the Turn-Helix-Loop Segment in Solution

Although crystallographic plasticity is interpretable as flexibility in a protein molecule, clear experimental evidence in a monomeric state in solution is necessary. For this purpose, samples of ^{15}N-labeled C-terminal domains of *Af*AglB-S1 and *Af*AglB-S2 , were prepared and measured their ^{1}H-^{15}N HSQC spectra. *Af*AglB-S2 (161 residues + N-terminal His tag) was selected for further NMR analyses because its spectrum was clear enough for residue assignment . Most of the main chain ^{1}H-^{15}N cross peaks in the HSQC spectrum were assigned by standard triple resonance experiments (Figure 4-1). Two ^{15}N spin relaxation rates, R_1 and R_2, and the heteronuclear ^{1}H-^{15}N nuclear Overhauser effects (NOEs) at two static magnetic field strengths measurements were taken (Figure 4-2).

According to the Lipari-Szabo-type model-free formalism, the dynamics of individual ^{15}N spins in a protein can be separated into the overall tumbling of a protein molecule and the

internal motions of each ^{1}H-^{15}N bond within a protein molecule (Lipari and Szabo, 1982a; Lipari and Szabo, 1982b). The model-free analysis consists of two steps: estimation of the rotational diffusion model for an entire protein molecule, and selection of a motional model for each ^{15}N spin (Palmer, 1997). The overall rotational tumbling is described by the effective correlation time, τ_m. An axially symmetric diffusion model for the rotational tumbling of the *Af*AglB-S2 molecule was adopted based on the results of the r2r1_tm program and the statistical test by the program quadric_diffusion (Tables 4-1, 2 and 3). The τ_m value calculated from the crystal structure (11.6 ns) was in good agreement with that obtained in the model-free analysis (12.1 ns), indicating the monomeric state of the *Af*AglB-S2 molecule in solution, at the concentrations used in the NMR experiments. In the treatment of internal local motions, the square of the generalized order parameter (S^2) and the effective correlation time for the internal motion (τ_e) characterize the amplitude and the timescale of internal motion for each ^{15}N spin in the ps-ns timescale, respectively. The chemical exchange-induced relaxation rate (R_{ex}) is a phenomenological term to account for the contribution of conformational dynamics in the μs-ms timescale to the R_2 term. Local motional models for ^{1}H-^{15}N bonds were selected on a statistical basis (Mandel et al., 1995): Model 1, (S^2); Model 2, (S^2, τ_e); Model 3, (S^2, R_{ex}); Model 4, (S^2, τ_e, R_{ex}), in which the terms required to account for the motions are listed in parentheses. Among the 157 non-proline residues of *Af*AglB-S2, 96 residues were assigned to Model 1, 20 residues to Model 2, 10 residues to Model 3, and 4 residues to Model 4. Two residues, Gly472 and Val483, could not be fit to any model. The plots of S^2, τ_e, and R_{ex} as a function of residue number are shown in Figure 4-3 A, and the residues are colored according to the selected models on the crystal structure in Figure 4-3 B.

For the majority of the residues in the structure, the near unity S^2 values indicated that their ^{1}H-^{15}N bond vectors are firmly fixed on the *Af*AglB-S2 structural framework (*cyan* in Figure 4-3 B). This means that the *Af*AglB-S2 molecule behaves mainly as a rigid body in solution. Some residues, mainly in the loop regions, exhibit fast internal motion with motional correlation times on the order of picosecond–nanosecond (ps-ns), as revealed by the necessity of the τ_e term for the description of their dynamic properties (*pink* and *purple* in Figure 4-3 B). Other residues on the turn and α-helical regions in the THL segment have slow

conformational transitions on the μs-ms timescale, as shown by the necessity of the R_{ex} term (*blue* and *purple* in Figure 4-3 B). The majority of these residues are located on the WWDYG motif and the following α-helix. When compared to the crystal structure, these R_{ex} residues are excellently coincident with the plastic THL segment (Figure 4-5 A), which demonstrates the correctness of the interpretation of the conformational plasticity in crystals as a sign of flexibility. A plot of R_2^{eff} against ν_{CPMG} (Figure 4-4 A) gave characteristic curves indicative of model 1 (top panel Figure 4-4A), model 2 (middle panel Figure 4-4A) and representative residues of the THL contributing to the concerted motion (Figure 4-4A bottom panel) respectively.

Estimation of the Timescale of the Protein Fluctuation

^{15}N R_2 relaxation dispersion measurement provided the information on the timescales of the dynamic fluctuations of proteins (Loria et al., 1999). R_2^{eff} dispersion profiles of 132 cross peaks (same set as with Model free analysis) were fitted to Model 1(no exchange) or Model 2 (two-site fast exchange) using the program *NESSY* (Bieri and Gooley, 2011). Twenty-four residues were selected as Model 2 and mapped on the *Af*AglB-S2 structure (Figure 4-4 B). The distribution of the relaxation dispersion Model 2 is consistent with that obtained by the model-free analysis (Figure 4-3 A and B). It was assumed that motion in the plastic region was concerted, and the ^{15}N spins in the plastic region may experience the same exchange of states. From the 24 residues (Appendix E-1), 7 residues (Table 4-4 and Figure 4-4A bottom panel) contained in the THL segment and the kinked helix were simultaneously fit to Model 2, to obtain a single exchange constant of $1,834 \pm 88\ s^{-1}$.

Design of Conformationally Restricted Mutants using a Disulfide Bond

To assess the contribution of the dynamic nature to the enzymatic activity, effects of a conformational restriction in the C-terminal domain on the OST activity were investigated. With reference to the crystal structure of the full-length *C.lari* PglB (Lizak et al., 2011), a search for an appropriate location for a disulfide cross-link that could restrict the flexibility in the C-terminal domain (Figure 5-1 A) was done. The engineered disulfide bond must be

carefully designed to avoid direct interference with the enzymatic activity. The disulfide bond was placed at a site distant from both the catalytic site in the transmembrane region and the Ser/Thr pocket in the C-terminal domain. For efficient conformational restriction, the disulfide bond should connect two rigid structures, such as α-helices. With these considerations, positions of two cysteine's, one (L447) in the most N-terminal α-helix in the C-terminal globular domain, and the other (Y473 or S474) in the α-helix in the THL segment (Figures 5-1 A and B) were selected as such. In the crystal structures of *Cj*PglB and *Cl*PglB, the side chains at the two candidate positions are close enough to form a disulfide bond (Figure 5-2 A). In this study, *Pyrococcus furiosus* AglB-L was chosen for the generation of conformationally restricted mutants, because it has been intensively studied both structurally and enzymatically (Igura and Kohda, 2011a; Igura et al., 2007; Igura et al., 2008) in the same laboratory where this experimental work was performed. Also considered was the extreme stability of the protein from the hyperthermophilic organism. Although the THL segment in the *Pf*AglB-L crystal structure was distorted by the crystal contact effects (Figure 5-2 B), the design of the engineered disulfide bond can be transferred from the *Campylobacter* sequences to the *Pyrococcus* sequence through the structure aided sequence alignment (Figure 5-1 B). Figure 5-1 C shows results of the OST assay of the mean of three independent experiments before correcting for disulfide content.

Conformationally Restricted AglB is Inactive

To achieve this, two double cysteine mutants (L495C/S521C and L495C/L522C) of the full-length *Pf*AglB-L protein were constructed. As a negative control, a single cysteine mutant (L495C) was also generated. The wild type and three mutants were expressed in *E. coli* membrane fractions, and were partially purified by Ni-affinity resin after solubilization in the presence of 1% n-dodecyl-β-D-maltopyranoside. The protein amount was quantified by western blotting, using anti-His tag antibodies and fluorescently-labeled secondary antibodies (Figure 5-3 A). The OST activity was measured by the PAGE method (Kohda et al., 2007) (Figure 5-3 B). The percentage of disulfide bond formation was estimated by counting the

number of sulfhydryl groups, using the Mal-PEG alkylation method (Makmura et al., 2001) (Figure 5-3 C). The chemical modification of a sulfhydryl group with maleimide-polyethylene glycol (Mwt = 2,000 Da) causes the mobility shift of proteins in SDS-PAGE. Results showed that the two cysteine residues formed a disulfide bond in about 75 % (78.5 ± 1.5 % for L495C/S521C, 74.5 ± 0.5 % for L495C/L522C) of the double cysteine mutants. After an incubation with a reducing agent, dithiothreitol (DTT), the percentage of disulfide bond formation decreased to about 15 % (15.1 ± 0.8 % for L495C/S521C, 17.2 ± 5.0 % for L495C/L522C). The formation and cleavage of the engineered disulfide bond were not perfect, but they allowed a quantitative investigation of the effects of the disulfide cross-link on the *Pf*AglB-L mutant activity.

The specific activity of the double cysteine mutants, as well as the wild type and the single cysteine mutant (Figure 5-1 C) were independently determined using the same protocols. The disulfide cross-linked AglB mutants were unable to catalyze the N-glycan transfer reaction, and importantly, this inhibitory effect was reversed upon cleavage of the disulfide bond. This was evident even before the correction of the disulfide bond formation percentage (Figure 5-1C), but was almost perfect after the correction (Figure 5-3 D). The *Pf*AglB-L mutants bearing the engineered disulfide bond (*orange* bars) had virtually no activity, whereas the same mutants without the disulfide cross-link (*blue* bars) had completely restored activity. This conclusion is not absolutely dependent on the particular design of the engineered disulfide bond, because two different disulfide bonds, C495-C521 and C495-C522, provided the same results. The full recovery of activity under reducing conditions clearly demonstrates that the cysteine substitution is not relevant, and the disulfide bond crosslinking is solely responsible for the inactivation effect. For confirmation, the wild type and the single cysteine mutant displayed full activity under both redox conditions.

Chapter 7 - Overall Discussion

It is now widely accepted that the dynamics of proteins contribute to their various functions, including enzyme catalysis (Daniel et al., 2003; Kern et al., 2005). They cover timescales ranging from 10^{-15} to > 1 s. Of particular interest is the global dynamics of a protein molecule, because such collective motions of large regions within a protein molecule are often required for substrate binding, catalytic conversion, product release, and enzymatic activity control. A polypeptide chain is intrinsically flexible, but in a native protein structure, it is folded and densely packed. Thus, collective motions involving many amino acid residues are inevitably slow on the μs-ms timescale, which is close to the timescales of enzyme reactions (Agarwal, 2005; Palmer et al., 2001). Although many experimental techniques have been used, the analyses of such collective, large movements in a protein are not always straightforward. Protein crystallography provides direct evidence for conformational changes by comparing structures in the presence and absence of ligands, but it does not reveal minor, transiently formed conformations. This research work presented herew4 offers a unique means for the detection of collective large motions in the C-terminal globular domain of the OST enzymes by crystallography and their verification by solution NMR.

To answer many of the questions asked, multiple comparisons of orthologous structures from different organisms and paralogous structures within an organism were performed. The six structures of the C-terminal globular domains of the AglB/PglB proteins(Igura et al., 2008; Lizak et al., 2011; Maita et al., 2010; Matsumoto et al., 2012), including the two new structures in this study, share sequence identities ranging from 20 % to 70 %. The common structure is the approximately 150-residue structural unit called CC. The presence of significant local conformational differences in the CC structural unit, regardless of the sequence identities, was totally unexpected (Figures 3-1, 3-2 and 3-3). The conformationally variable segment consists of three consecutive structural elements, turn – α-helix – loop, and is referred to as the THL segment. Conformational variations due to crystal packing effects are a common phenomenon: some portions of a protein, such as the N- and C-termini and long loops, often show conformational polymorphisms in crystals. However, in this study, focus was only on the local structural plasticity, because the THL segment contains

the highly conserved WWDYG motif. Thus, it was expected that the conformational plasticity would be relevant to the enzyme functions.

NMR analyses were performed to assess the flexibility of the plastic region in the C-terminal globular domain of *Af*AglB-S2. The observation that some residues contained high values of the R_2 term at both fields suggested the presence of the characteristic chemical exchange constant (R_{ex}) in such values (Figure 4-2). The model-free analysis generated an appropriate motional model for each ^{1}H-^{15}N bond. In this study, the residues of interest are those that contain the R_{ex} term; *i.e.,* Model 3 and Model 4 (colored *blue* and *purple* in Figures 4-3 A and B). The necessity of the R_{ex} term to describe the ^{15}N spin relaxation suggested that the conformational exchange process occurs on the micro to millisecond (μs-ms) timescale. The majority of these residues are located in the THL segment. In addition, the residues with R_{ex} terms are located on the kinked helix close to the WWDYG motif, although the kinked helix showed little plasticity in the crystals. This suggested that part of the kinked helix is in concerted motion with the THL segment in solution. Importantly, the interface between the WWDYG motif and the kinked helix forms the +2 Ser/Thr-binding pocket (Figure 2-9 B). The relaxation dispersion analysis suggested that the timescale of the dynamics of the Ser/Thr-binding pocket was about 0.5 milliseconds (Figure 4-4 A and B). Since many biochemical events occur on this timescale, the involvement of the conformational fluctuation in the OST activity is highly suggestive. The comparison of NMR and x-ray crystallographic data showed a high correlation as shown in Figure 4-5 A, B and C respectively.

To test this hypothesis, experiments to restrict the mobility in the C-terminal globular domain of the full-length *Pf*AglB-L were performed. Disulfide-bond mediated stiffening was previously used to lock the intramolecular conformations, as well as the inter-domain or inter-subunit spatial arrangements of proteins (Gill et al., 2009; Horan and Noller, 2007; Palle et al., 2008; Peske et al., 2000). Double cysteine mutants were designed to incorporate a disulfide bond in the C-terminal globular domain of *P.furiosus* AglB-L. Importantly, the resting state can bind a substrate peptide without conformational changes, as revealed by the indistinguishable structures between *Cj*PglB in the apo form and *Cl*PglB in the complex state with a substrate peptide. Thus, it was reasonable enough to assume that the disulfide crosslink that locked the structure in the resting state did not affect the OST activity, if the conformational fluctuations of the THL segment were not relevant to the catalytic activity.

However, the results obtained were totally opposite to what was assumed: the OST activity of the engineered AglB was abolished in the presence of the disulfide bond (Figures 5-1 C and 5-3 D). This inhibitory effect is completely reversible upon disulfide bond cleavage, confirming that the effects are due to the crosslinking. These results clearly indicated that the single resting-state conformation is not sufficient, and another conformation must be involved in the catalytic reaction. Considering that the mobile region contains the Ser/Thr pocket, this pocket must transiently collapse during the catalytic cycle.

The report on the *C.lari* PglB crystal structure suggested that the binding of the +2 Ser/Thr residues in the glycosylation sequon to the Ser/Thr binding pocket is a prerequisite for the proper spatial positioning of the acceptor Asn side chain in the same sequon into a narrow "porthole" of the PglB protein (Gilmore, 2011). This enables simultaneous interactions between the side-chain amido group and the two side-chain carboxy groups (Figure 2-9 B), and its precise coordination geometry forces the rotation of the C-N bond of the side-chain amido group, leading to the amide nitrogen activation (Lizak et al., 2011). The C-terminal domains reported here lack the catalytic acidic residues, and therefore could not be used to test the hypothesis put forward by Lizak et al., but it is reasonable to assume that the collapse of the Ser/Thr pocket immediately disrupts the precise geometry necessary for the amide nitrogen activation. The report on the *C.lari* PglB crystal structure also suggested a three-step catalytic cycle, involving the conformational changes of the external loop 5 in the transmembrane region. Although the C-terminal globular domain of the PglB protein was assumed to be static, their model is fully compatible with the dynamic Ser/Thr pocket revealed in this study.

Finally, what functional advantage does the dynamic Ser/Thr pocket add to the OST enzymatic activity? The dynamic nature of the Ser/Thr binding pocket was considered to be an essential feature for the OST enzymes to scan for N-glycosylation sequons efficiently along nascent polypeptide chains. The OST enzyme has evolved to search all potential N-glycosylation sites, by recognizing the hydroxyl group-containing amino acid residues at the +2 position using the Ser/Thr pocket, and then releasing the N-glycan-modified sequons quickly to start the search for the next sequon. This property is required to steadily glycosylate two neighboring sequons along the polypeptide chain emerging from a translocon in the membranes, while coupled with ribosomal protein synthesis. The Eubacterial *Campylobacter jejuni* PglB was shown to primarily work as a post-translational modification enzyme

(Kowarik et al., 2006a). Thus, although the basic catalytic mechanism is probably conserved in all of the OST enzymes, the sequon recognition mechanism might not be the same. In fact, the *C.jejuni* sequon was extended to Asp/Glu-X-1-Asn-X-Ser/Thr, where both X-1 and X are any amino acid except for proline (Chen et al., 2007; Kowarik et al., 2006b). A conserved arginine residue in the transmembrane region, R331 in the *Cl*PglB structure and the corresponding R328 in *Cj*PglB, interacts with the acidic residue at the -2 position (Lizak et al., 2011); thus, PglB might have more static sequon recognition.

Chapter 8 - Conclusion and future works

In conclusion, findings in this research work have demonstrated that comparative structural biology is a viable and easy visual approach for the structure-function relationship study of proteins. Structural comparison of the closely and distantly related AglB and PglB proteins led to the unexpected discovery of the mobile region in the C-terminal globular domain. Also, a comparative structural approach is useful for the selection of the most suitable protein for special purposes: The smallest AglB (that from *A.fulgidus* AglB-S2) was used for the NMR relaxation studies with the C-terminal globular domain having a weight of about 20kDa, and for disulfide bond studies *P.furiosus* AglB-L was selected for the custom-designed mutagenesis study that required the high stability of the protein to tolerate bold designs of mutations and due to its biochemistry well established in the laboratory where this research was done and its transferase activity is much higher and more specific than the current models in the laboratory.

Overall, the aims and objectives of this study were achieved and it was shown that plasticity in the WWDYG motif segment in the C-terminal globular domain of archaeal oligosaccharyltransferases as revealed by the comparison of X-ray crystal structures is essential for the activity. The presence of residues with high (20 and above) values of R_2 (Figure 4-2) highly suggested domain motion at the micro to millisecond time scale due to presence of the phenomenological exchange constant R_{ex}. The observation of the mobile region in the C-terminal globular domain of an archaeon OST from *A.fulgidus* AglB-S2 by solution NMR (Figures 4-3 and 4-4) analyzed by Model free software as well as constant-time Curr Purcell Meiboom Gill (CPMG) relaxation dispersion experiment analysis by NESSY (Figures 4-4 A and B) software which allowed for clustering of residues within the helix-turn-loop region and the DKi motif which make part of the acceptor peptide binding site known as the Ser/Thr binding pocket proved the observations in crystal structures.

Residues in the WWDYG motif segment and the DKi motif segment were deemed to have concerted motion with a rate of about 1,834 ±88 (s^{-1}) or about 0.5ms on the microsecond to millisecond time scale and this was shown to be the case in the study carried out and reported in this thesis. A comparison of the data obtained from NMR studies and x-ray

crystallographic studies (Figure 4-5) showed a high correlation between the two experiments. The X-ray and NMR results were complemented with disulfide bond engineering (Figures 5-1 to 5-3 using *P.furiosus* AglB-L) and this allowed for the general conclusion that in solution, dynamics of the Ser/Thr binding pocket at the microsecond to millisecond motions are observed and that these contribute to the overall activity. This was shown that restriction of mobility by incorporating a disulfide bond, led to abolishment of activity while the reduction of the same restored activity. A point worthy noting is that though the cysteine mutations were introduced, such did not impair protein stability as mere formation or reduction of the disulfide bond was the sole contributor to abolishment of activity or complete reversal upon treatment with reducing agent DTT respectively.

Taken together, the experimental evidence provided an insight in the future work on elucidating the mechanism of catalysis by oligosaccharyltransferases in general as well as transfer of the knowledge of crystallization techniques gathered from Archaeal and Eubacterial OST's to eukaryotic STT3 complex. Many congenital disorders of glycosylation (CDG's) as a result of under glycosylation or lack of it will be understood more broadly. Also cases involving evasion of human immunity by some viruses which utilize the N-linked glycosylation pathway as a means of protection will be elucidated and provide more insight on how to design therapeutics that are effective at combating various diseases.

Appendix A - Chemical shifts of *A.fulgidus* AglB-S2 at 600MHz, 308 K.

Table A-1 Chemical shifts for *A.fulgidus* AglB-S2 at 600MHz, 308K.

Residue		N	H	Residue		N	H	Residue		N	H
ALA	10	126.385	8.295	ALA	479	117.477	7.311	ILE	542	122.325	8.531
ALA	11	122.337	8.022	LYS	480	112.398	7.027	ALA	543	120.983	7.952
LEU	12	120.277	7.904	LYS	481	117.758	7.606	GLY	544	104.178	7.609
GLU	13	120.597	8.215	ALA	482	120.598	8.016	TYR	545	119.845	8.071
VAL	14	119.251	7.771	VAL	483	108.45	7.841	SER	546	112.936	8.668
LEU	15	123.932	7.838	GLN	489	118.107	7.684	GLU	548	121.671	10.242
PHE	16	119.659	7.964	ALA	490	122.985	7.389	TYR	549	117.838	7.356
GLN	17	121.225	8.033	ALA	491	111.606	8.677	MET	550	120.117	7.644
GLY	18	110.46	8.008	ALA	492	124.842	7.545	LYS	551	121.339	8.906
GLU	433	120.341	8.58	ASP	493	117.925	8.837	LYS	553	117.534	8.691
MET	434	125.259	8.532	ASP	494	122.911	7.861	GLU	554	116.674	7.011
THR	435	116.377	7.013	ALA	495	120.921	7.78	ILE	555	122.099	7.88
TRP	438	118.499	7.274	ALA	496	120.27	9.323	ILE	556	121.376	7.942
LYS	439	119.844	8.383	LYS	497	121.362	8.982	ASP	557	119.537	8.103
GLU	440	118.352	8.143	PHE	498	119.129	8.156	PHE	558	119.324	7.837
ALA	441	123.518	7.983	PHE	499	114.792	8.643	PHE	559	120.268	9.013
LEU	442	118.488	9.08	THR	500	103.293	7.877	ASN	560	117.428	8.107
ASN	443	117.424	7.691	ALA	501	125.762	7.149	LYS	561	117.012	7.229
TRP	444	120.824	8.243	GLN	502	118.974	8.771	THR	562	108.387	7.426
MET	445	119.296	9.059	SER	503	111.355	7.412	MET	563	119.263	8.121
LYS	446	118.233	7.988	GLU	505	119.225	8.893	LEU	564	117.983	8.115
GLU	447	113.251	7.045	GLU	506	118.664	7.505	TYR	565	120.978	7.213
ASN	448	112.971	7.476	ALA	507	123.066	7.612	LYS	566	119.949	8.011
LEU	449	118.887	7.216	MET	508	115.864	8.734	LEU	567	117.86	8.232
GLU	450	121.599	8.228	LYS	509	119.546	7.497	HIS	568	115.987	8.351
ALA	451	126.781	8.4	ILE	510	119.53	7.296	VAL	569	115.192	7.75
GLN	452	119.571	8.458	VAL	511	120.782	7.54	GLU	570	112.971	6.67
ASP	453	121.982	8.429	GLU	512	117.501	8.239	ASN	571	116.931	7.287
TYR	454	121.378	7.968	LYS	513	119.761	7.537	ALA	572	112.176	8.504
LEU	455	121.313	7.819	ARG	514	111.842	7.845	THR	573	113.564	7.761
LYS	456	120.277	7.92	LYS	515	118.459	7.667	ASN	574	118.246	8.958
ALA	457	122.89	8	VAL	516	119.01	8.114	LEU	575	118.878	7.202
TYR	458	115.541	7.667	ARG	517	124.652	8.371	THR	576	119.365	10.748
GLU	459	120.683	7.872	TYR	518	112.381	7.324	HIS	577	121.09	10.59
LYS	460	121.39	8.112	VAL	519	121.949	9.335	PHE	578	113.051	6.76
ASP	462	120.932	8.531	VAL	520	127.475	9.025	ARG	579	118.371	8.862
TYR	463	113.649	7.378	THR	521	119.814	9.592	LEU	580	127.471	9.184
ALA	464	120.349	8.47	VAL	522	114.75	5.881	LEU	581	131.214	9.262
VAL	465	121.317	9.066	GLU	523	129.332	8.752	LYS	582	115.596	7.178
LEU	466	131.096	9.397	GLU	524	117.009	8.699	PHE	584	125.961	8.564
SER	467	121.884	8.683	LEU	525	116.059	7.768	GLY	585	108.241	8.151
TRP	468	116.576	8.7	THR	526	101.945	7.425	THR	586	117.907	8.423
ASP	470	120.557	6.125	GLU	530	114.811	8.723	VAL	587	122.443	6.894
TYR	471	115.554	6.146	THR	531	114.173	7.51	LYS	588	125.259	8.746
GLY	472	106.459	7.788	THR	534	117.445	7.978	ILE	589	118.332	8.012
ASN	473	120.406	8.774	LYS	535	125.919	8.431	PHE	590	126.151	10.002
TRP	474	120.564	7.694	PHE	536	121.409	8.857	GLU	591	122.568	9.384
ILE	475	121.45	7.755	ILE	537	114.167	6.771	VAL	592	125.67	7.801
LEU	476	120.064	8.458	ILE	539	117.43	6.774	LYS	593	132.963	8.317
TYR	477	113.432	8.39	MET	540	117.769	7.907				
VAL	478	118.858	8.552	GLN	541	117.876	8.294				

Appendix B - Relaxation parameters of *A.fulgidus* AglB-S2 at 600MHz,308 K.

Table B-1 Relaxation parameters of *A.fulgidus* AglB-S2 at 600MHz, 308 K.

Residue		R1 600	R2 600	NOE 600
GLU	433	1.321 ± 0.132	13.569 ± 1.357	0.497 ± 0.050
MET	434	1.338 ± 0.134	16.976 ± 1.698	0.559 ± 0.056
TRP	438	1.173 ± 0.117	17.385 ± 1.739	0.829 ± 0.083
LYS	439	1.181 ± 0.118	17.891 ± 1.789	0.785 ± 0.079
GLU	440	1.187 ± 0.119	18.339 ± 1.834	0.788 ± 0.079
ALA	441	1.191 ± 0.119	17.083 ± 1.708	0.805 ± 0.081
LEU	442	1.203 ± 0.120	17.764 ± 1.776	0.795 ± 0.080
ASN	443	1.156 ± 0.116	18.416 ± 1.842	0.796 ± 0.080
TRP	444	1.239 ± 0.124	14.982 ± 1.498	0.693 ± 0.069
MET	445	1.168 ± 0.117	18.757 ± 1.876	0.774 ± 0.077
LYS	446	1.188 ± 0.119	17.495 ± 1.750	0.812 ± 0.081
GLU	447	1.17 ± 0.117	17.16 ± 1.716	0.706 ± 0.071
ASN	448	1.134 ± 0.113	17.04 ± 1.704	0.763 ± 0.076
GLU	450	1.421 ± 0.142	15.041 ± 1.504	0.641 ± 0.064
ALA	451	1.462 ± 0.146	15.287 ± 1.529	0.587 ± 0.059
GLN	452	1.357 ± 0.136	12 ± 1.200	0.508 ± 0.051
ASP	453	1.437 ± 0.144	10.897 ± 1.090	0.45 ± 0.045
TYR	454	1.439 ± 0.144	14.597 ± 1.460	0.552 ± 0.055
LEU	455	1.587 ± 0.159	11.858 ± 1.186	0.523 ± 0.052
ALA	457	1.456 ± 0.146	10.566 ± 1.057	0.476 ± 0.048
TYR	458	1.52 ± 0.152	12.8 ± 1.280	0.436 ± 0.044
GLU	459	1.54 ± 0.154	10.413 ± 1.041	0.446 ± 0.045
LYS	460	1.495 ± 0.150	9.65 ± 0.965	0.384 ± 0.038
ASP	462	1.472 ± 0.147	11.889 ± 1.189	0.526 ± 0.053
TYR	463	1.321 ± 0.132	19.027 ± 1.903	0.582 ± 0.058
ALA	464	1.21 ± 0.121	17.316 ± 1.732	0.774 ± 0.077
VAL	465	1.187 ± 0.119	16.604 ± 1.660	0.738 ± 0.074
LEU	466	1.249 ± 0.125	18.853 ± 1.885	0.845 ± 0.085
SER	467	1.286 ± 0.129	16.983 ± 1.698	0.894 ± 0.089
TRP	468	1.229 ± 0.123	20.265 ± 2.027	0.778 ± 0.078
ASP	470	1.248 ± 0.125	21.837 ± 2.184	0.722 ± 0.072
TYR	471	1.261 ± 0.126	22.379 ± 2.238	0.805 ± 0.081
GLY	472	1.212 ± 0.121	18.408 ± 1.841	0.781 ± 0.078
ASN	473	1.291 ± 0.129	19.712 ± 1.971	0.787 ± 0.079
TRP	474	1.103 ± 0.110	17.876 ± 1.788	0.863 ± 0.086
ILE	475	1.174 ± 0.117	19.99 ± 1.999	0.837 ± 0.084
LEU	476	1.206 ± 0.121	18.872 ± 1.887	0.791 ± 0.079
TYR	477	1.136 ± 0.114	22.314 ± 2.231	0.794 ± 0.079
VAL	478	1.141 ± 0.114	25.254 ± 2.525	0.745 ± 0.075
ALA	479	1.186 ± 0.119	24.484 ± 2.448	0.778 ± 0.078
LYS	480	1.193 ± 0.119	20.746 ± 2.075	0.788 ± 0.079
LYS	481	1.152 ± 0.115	18.195 ± 1.820	0.804 ± 0.080
ALA	482	1.182 ± 0.118	15.444 ± 1.544	0.684 ± 0.068
VAL	483	0.969 ± 0.097	9.623 ± 0.962	0.758 ± 0.076

Table B-2 Relaxation parameters of *A.fulgidus* AglB-S2 at 600MHz,308 K, continued.

Residue		R_1 600	R_2 600	NOE 600
GLN	489	1.325 ± 0.133	20.02 ± 2.002	0.665 ± 0.067
ALA	490	1.263 ± 0.126	13.082 ± 1.308	0.437 ± 0.044
ALA	492	1.333 ± 0.133	17.037 ± 1.704	0.781 ± 0.078
ASP	493	1.259 ± 0.126	18.132 ± 1.813	0.778 ± 0.078
ASP	494	1.169 ± 0.117	18.164 ± 1.816	0.806 ± 0.081
ALA	495	1.235 ± 0.124	16.41 ± 1.641	0.788 ± 0.079
ALA	496	1.236 ± 0.124	16.965 ± 1.697	0.791 ± 0.079
LYS	497	1.176 ± 0.118	16.909 ± 1.691	0.827 ± 0.083
PHE	499	1.218 ± 0.122	17.189 ± 1.719	0.807 ± 0.081
ALA	501	1.171 ± 0.117	17.82 ± 1.782	0.832 ± 0.083
GLN	502	1.295 ± 0.130	14.567 ± 1.457	0.737 ± 0.074
SER	503	1.186 ± 0.119	14.458 ± 1.446	0.606 ± 0.061
GLU	505	1.291 ± 0.129	16.642 ± 1.664	0.763 ± 0.076
GLU	506	1.174 ± 0.117	16.791 ± 1.679	0.802 ± 0.080
ALA	507	1.218 ± 0.122	16.026 ± 1.603	0.781 ± 0.078
MET	508	1.157 ± 0.116	17.677 ± 1.768	0.806 ± 0.081
LYS	509	1.155 ± 0.116	18.137 ± 1.814	0.792 ± 0.079
ILE	510	1.147 ± 0.115	16.994 ± 1.699	0.786 ± 0.079
VAL	511	1.152 ± 0.115	18.049 ± 1.805	0.816 ± 0.082
GLU	512	1.14 ± 0.114	18.3 ± 1.830	0.827 ± 0.083
LYS	513	1.13 ± 0.113	17.154 ± 1.715	0.768 ± 0.077
ARG	514	1.135 ± 0.114	16.92 ± 1.692	0.742 ± 0.074
LYS	515	1.228 ± 0.123	18.963 ± 1.896	0.792 ± 0.079
VAL	516	1.163 ± 0.116	17.312 ± 1.731	0.806 ± 0.081
ARG	517	1.222 ± 0.122	16.409 ± 1.641	0.703 ± 0.070
TYR	518	1.235 ± 0.124	15.568 ± 1.557	0.795 ± 0.080
VAL	519	1.196 ± 0.120	15.376 ± 1.538	0.792 ± 0.079
VAL	520	1.216 ± 0.122	15.864 ± 1.586	0.815 ± 0.082
THR	521	1.212 ± 0.121	17.097 ± 1.710	0.805 ± 0.081
VAL	522	1.131 ± 0.113	17.07 ± 1.707	0.824 ± 0.082
GLU	523	1.171 ± 0.117	17.404 ± 1.740	0.877 ± 0.088
GLU	524	1.327 ± 0.133	18.287 ± 1.829	0.772 ± 0.077
LEU	525	1.256 ± 0.126	18.744 ± 1.874	0.803 ± 0.080
THR	526	1.188 ± 0.119	16.882 ± 1.688	0.785 ± 0.079
GLU	530	1.349 ± 0.135	14.422 ± 1.442	0.529 ± 0.053
THR	531	1.461 ± 0.146	15.432 ± 1.543	0.534 ± 0.053
THR	534	1.26 ± 0.126	21.388 ± 2.139	0.518 ± 0.052
LYS	535	1.283 ± 0.128	20.788 ± 2.079	0.758 ± 0.076
ILE	537	1.346 ± 0.135	20.182 ± 2.018	0.771 ± 0.077
ILE	539	1.195 ± 0.120	17.34 ± 1.734	0.778 ± 0.078
MET	540	1.189 ± 0.119	16.707 ± 1.671	0.791 ± 0.079
GLN	541	1.244 ± 0.124	16.715 ± 1.672	0.764 ± 0.076
ILE	542	1.227 ± 0.123	16 ± 1.600	0.783 ± 0.078
ALA	543	1.255 ± 0.126	15.799 ± 1.580	0.771 ± 0.077
GLY	544	1.17 ± 0.117	16.56 ± 1.656	0.755 ± 0.076
TYR	545	1.131 ± 0.113	16.628 ± 1.663	0.793 ± 0.079

Table B-3 Relaxation parameters of *A.fulgidus* AglB-S2 at 600MHz,308 K, continued.

Residue		R_1 600			R_2 600			NOE 600		
SER	546	1.48	±	0.148	11.702	±	1.170	0.624	±	0.062
GLU	548	1.223	±	0.122	16.613	±	1.661	0.714	±	0.071
TYR	549	1.164	±	0.116	16.152	±	1.615	0.643	±	0.064
MET	550	1.075	±	0.108	15.129	±	1.513	0.701	±	0.070
LYS	551	1.144	±	0.114	14.689	±	1.469	0.743	±	0.074
LYS	553	1.137	±	0.114	17.503	±	1.750	0.734	±	0.073
ILE	555	1.169	±	0.117	16.716	±	1.672	0.82	±	0.082
ILE	556	1.166	±	0.117	16.69	±	1.669	0.743	±	0.074
ASP	557	1.144	±	0.114	17.661	±	1.766	0.762	±	0.076
PHE	558	1.178	±	0.118	16.871	±	1.687	0.791	±	0.079
PHE	559	1.169	±	0.117	17.395	±	1.740	0.835	±	0.084
ASN	560	1.162	±	0.116	18.526	±	1.853	0.818	±	0.082
LYS	561	1.182	±	0.118	17.325	±	1.733	0.785	±	0.079
THR	562	1.172	±	0.117	15.705	±	1.571	0.772	±	0.077
LEU	564	1.211	±	0.121	17.208	±	1.721	0.81	±	0.081
TYR	565	1.246	±	0.125	16.256	±	1.626	0.792	±	0.079
LYS	566	1.249	±	0.125	16.464	±	1.646	0.823	±	0.082
LEU	567	1.193	±	0.119	18.199	±	1.820	0.818	±	0.082
HIS	568	1.226	±	0.123	16.722	±	1.672	0.83	±	0.083
VAL	569	1.258	±	0.126	16.809	±	1.681	0.868	±	0.087
GLU	570	1.209	±	0.121	15.552	±	1.555	0.825	±	0.083
ASN	571	1.266	±	0.127	17.639	±	1.764	0.83	±	0.083
ALA	572	1.202	±	0.120	17.245	±	1.725	0.807	±	0.081
THR	573	1.225	±	0.123	18.619	±	1.862	0.807	±	0.081
ASN	574	1.258	±	0.126	16.414	±	1.641	0.786	±	0.079
THR	576	1.076	±	0.108	15.471	±	1.547	0.78	±	0.078
HIS	577	1.111	±	0.111	15.785	±	1.579	0.782	±	0.078
PHE	578	1.138	±	0.114	16.819	±	1.682	0.768	±	0.077
ARG	579	1.136	±	0.114	16.434	±	1.643	0.799	±	0.080
LEU	580	1.214	±	0.121	15.889	±	1.589	0.817	±	0.082
LEU	581	1.242	±	0.124	18.073	±	1.807	0.843	±	0.084
LYS	582	1.175	±	0.118	16.23	±	1.623	0.806	±	0.081
PHE	584	1.172	±	0.117	15.085	±	1.509	0.795	±	0.080
GLY	585	1.216	±	0.122	10.825	±	1.083	0.275	±	0.028
THR	586	1.231	±	0.123	15.777	±	1.578	0.709	±	0.071
VAL	587	1.238	±	0.124	15.474	±	1.547	0.737	±	0.074
LYS	588	1.226	±	0.123	16.112	±	1.611	0.773	±	0.077
ILE	589	1.2	±	0.120	16.482	±	1.648	0.82	±	0.082
PHE	590	1.19	±	0.119	15.951	±	1.595	0.741	±	0.074
GLU	591	1.2	±	0.120	15.981	±	1.598	0.806	±	0.081
VAL	592	1.152	±	0.115	16.665	±	1.667	0.764	±	0.076
LYS	593	1.185	±	0.119	16.659	±	1.666	0.67	±	0.067

Appendix C - Relaxation parameters of *A.fulgidus* AglB-S2 at 700MHz,308 K

Table C-1 Relaxation parameters of *A.fulgidus* AglB-S2 at 700MHz, 308 K.

Residue		R_1 700	R_2 700	NOE 700
GLU	433	1.16 ± 0.116	17.895 ± 1.790	0.546 ± 0.055
MET	434	1.204 ± 0.120	22.55 ± 2.255	0.603 ± 0.060
TRP	438	0.958 ± 0.096	21.853 ± 2.185	0.811 ± 0.081
LYS	439	0.903 ± 0.090	23.175 ± 2.318	0.829 ± 0.083
GLU	440	0.921 ± 0.092	23.43 ± 2.343	0.836 ± 0.084
ALA	441	0.943 ± 0.094	21.776 ± 2.178	0.865 ± 0.087
LEU	442	0.942 ± 0.094	22.428 ± 2.243	0.831 ± 0.083
ASN	443	0.924 ± 0.092	23.699 ± 2.370	0.838 ± 0.084
TRP	444	0.951 ± 0.095	21.218 ± 2.122	0.823 ± 0.082
MET	445	0.958 ± 0.096	24.331 ± 2.433	0.828 ± 0.083
LYS	446	0.953 ± 0.095	22.683 ± 2.268	0.829 ± 0.083
GLU	447	0.946 ± 0.095	20.401 ± 2.040	0.747 ± 0.075
ASN	448	0.915 ± 0.092	20.52 ± 2.052	0.769 ± 0.077
GLU	450	1.216 ± 0.122	19.572 ± 1.957	0.701 ± 0.070
ALA	451	1.291 ± 0.129	19.46 ± 1.946	0.606 ± 0.061
GLN	452	1.247 ± 0.125	15.257 ± 1.526	0.545 ± 0.055
ASP	453	1.319 ± 0.132	14.208 ± 1.421	0.574 ± 0.057
TYR	454	1.307 ± 0.131	18.516 ± 1.666	0.541 ± 0.054
LEU	455	1.375 ± 0.124	15.512 ± 1.396	0.557 ± 0.056
ALA	457	1.343 ± 0.121	13.303 ± 1.197	0.516 ± 0.052
TYR	458	1.366 ± 0.123	15.994 ± 1.439	0.48 ± 0.048
GLU	459	1.382 ± 0.124	13.178 ± 1.186	0.466 ± 0.047
LYS	460	1.339 ± 0.121	12.284 ± 1.106	0.422 ± 0.042
ASP	462	1.293 ± 0.116	15.285 ± 1.376	0.584 ± 0.058
TYR	463	1.05 ± 0.095	22.496 ± 2.025	0.623 ± 0.062
ALA	464	0.965 ± 0.087	22.817 ± 2.054	0.78 ± 0.078
VAL	465	0.92 ± 0.083	22.635 ± 2.037	0.809 ± 0.081
LEU	466	0.992 ± 0.089	22.703 ± 2.043	0.881 ± 0.088
SER	467	1.013 ± 0.101	23.458 ± 2.111	0.855 ± 0.086
TRP	468	1.015 ± 0.102	26.957 ± 2.426	0.862 ± 0.086
ASP	470	1.052 ± 0.105	29.856 ± 2.687	0.76 ± 0.076
TYR	471	1.068 ± 0.107	28.391 ± 2.555	0.81 ± 0.081
GLY	472	1 ± 0.100	32.538 ± 2.928	0.87 ± 0.087
ASN	473	1.006 ± 0.101	27.172 ± 2.445	0.86 ± 0.086
TRP	474	0.919 ± 0.092	23.373 ± 2.104	0.794 ± 0.079
ILE	475	0.966 ± 0.097	25.269 ± 2.274	0.801 ± 0.080
LEU	476	0.942 ± 0.094	25.336 ± 2.280	0.854 ± 0.085
TYR	477	0.939 ± 0.094	28.714 ± 2.584	0.77 ± 0.077
VAL	478	0.9 ± 0.090	35.67 ± 3.210	0.783 ± 0.078
ALA	479	0.967 ± 0.097	31.437 ± 2.829	0.8 ± 0.080
LYS	480	0.99 ± 0.099	28.075 ± 2.527	0.836 ± 0.084
LYS	481	0.938 ± 0.094	23.115 ± 2.080	0.765 ± 0.077
ALA	482	0.997 ± 0.090	21.222 ± 1.910	0.742 ± 0.074
VAL	483	0.859 ± 0.077	25.206 ± 2.269	0.846 ± 0.085

Table C-2 Relaxation parameters of *A.fulgidus* AglB-S2 at 700MHz, 308 K, continued.

Residue		R_1 700	R_2 700	NOE 700
GLN	489	1.157 ± 0.104	25.719 ± 2.315	0.681 ± 0.068
ALA	490	1.111 ± 0.100	17.474 ± 1.747	0.452 ± 0.045
ALA	492	1.144 ± 0.103	21.91 ± 2.191	0.808 ± 0.081
ASP	493	1.02 ± 0.092	24.219 ± 2.422	0.859 ± 0.086
ASP	494	0.958 ± 0.086	23.729 ± 2.373	0.845 ± 0.085
ALA	495	1.003 ± 0.090	21.822 ± 2.182	0.812 ± 0.081
ALA	496	0.977 ± 0.088	22.403 ± 2.240	0.844 ± 0.084
LYS	497	0.942 ± 0.094	23.083 ± 2.308	0.794 ± 0.079
PHE	499	0.995 ± 0.100	21.27 ± 2.127	0.76 ± 0.076
ALA	501	0.916 ± 0.092	21.578 ± 2.158	0.807 ± 0.081
GLN	502	1.053 ± 0.105	18.648 ± 1.865	0.77 ± 0.077
SER	503	0.981 ± 0.098	17.387 ± 1.739	0.63 ± 0.063
GLU	505	1.073 ± 0.107	21.699 ± 2.170	0.795 ± 0.080
GLU	506	0.946 ± 0.095	21.719 ± 2.172	0.81 ± 0.081
ALA	507	0.967 ± 0.097	20.918 ± 2.092	0.853 ± 0.085
MET	508	0.913 ± 0.091	21.85 ± 2.185	0.81 ± 0.081
LYS	509	0.918 ± 0.092	23.782 ± 2.378	0.84 ± 0.084
ILE	510	0.918 ± 0.092	22.535 ± 2.028	0.829 ± 0.083
VAL	511	0.922 ± 0.092	22.959 ± 2.066	0.818 ± 0.082
GLU	512	0.883 ± 0.088	23.284 ± 2.096	0.837 ± 0.084
LYS	513	0.899 ± 0.090	22.321 ± 2.009	0.799 ± 0.080
ARG	514	0.914 ± 0.091	20.823 ± 1.874	0.792 ± 0.079
LYS	515	0.981 ± 0.098	24.215 ± 2.179	0.814 ± 0.081
VAL	516	0.973 ± 0.097	22.436 ± 2.019	0.843 ± 0.084
ARG	517	0.97 ± 0.097	21.508 ± 1.936	0.785 ± 0.079
TYR	518	1.021 ± 0.102	19.346 ± 1.741	0.767 ± 0.077
VAL	519	0.93 ± 0.093	20.535 ± 2.054	0.874 ± 0.087
VAL	520	0.997 ± 0.100	19.712 ± 1.971	0.838 ± 0.084
THR	521	0.983 ± 0.098	22.177 ± 2.218	0.852 ± 0.085
VAL	522	0.935 ± 0.094	22.165 ± 2.217	0.833 ± 0.083
GLU	523	0.952 ± 0.095	20.884 ± 2.088	0.829 ± 0.083
GLU	524	1.059 ± 0.106	23.514 ± 2.351	0.848 ± 0.085
LEU	525	1.023 ± 0.102	23.889 ± 2.389	0.834 ± 0.083
THR	526	0.969 ± 0.097	21.043 ± 2.104	0.81 ± 0.081
GLU	530	1.193 ± 0.119	17.383 ± 1.738	0.51 ± 0.051
THR	531	1.264 ± 0.126	19.097 ± 1.910	0.575 ± 0.058
THR	534	1.135 ± 0.114	26.921 ± 2.692	0.536 ± 0.054
LYS	535	1.037 ± 0.104	26.14 ± 2.614	0.776 ± 0.078
ILE	537	0.971 ± 0.097	25.554 ± 2.555	0.823 ± 0.082
ILE	539	0.957 ± 0.096	21.438 ± 2.144	0.81 ± 0.081
MET	540	0.965 ± 0.097	21.885 ± 2.189	0.863 ± 0.086
GLN	541	0.988 ± 0.099	21.641 ± 2.164	0.813 ± 0.081
ILE	542	0.987 ± 0.099	21.882 ± 2.188	0.808 ± 0.081
ALA	543	1.027 ± 0.103	20.512 ± 2.051	0.756 ± 0.076

Table C-3 Relaxation parameters of *A.fulgidus* AglB-S2 at 700MHz, 308 K, continued.

Residue		R_1 700	R_2 700	NOE 700
GLY	544	0.943 ± 0.094	21.087 ± 2.109	0.829 ± 0.083
TYR	545	0.906 ± 0.091	21.538 ± 2.154	0.817 ± 0.082
SER	546	1.327 ± 0.133	13.579 ± 1.358	0.657 ± 0.066
GLU	548	1.007 ± 0.101	22.501 ± 2.250	0.734 ± 0.073
TYR	549	0.954 ± 0.095	20.693 ± 2.069	0.668 ± 0.067
MET	550	0.877 ± 0.088	20.052 ± 2.005	0.689 ± 0.069
LYS	551	0.922 ± 0.092	19.959 ± 1.996	0.761 ± 0.076
LYS	553	1.113 ± 0.111	20.208 ± 2.021	0.791 ± 0.079
ILE	555	0.934 ± 0.093	21.355 ± 2.136	0.819 ± 0.082
ILE	556	0.995 ± 0.100	20.633 ± 2.063	0.752 ± 0.075
ASP	557	0.899 ± 0.090	22.835 ± 2.284	0.802 ± 0.080
PHE	558	1.363 ± 0.136	22.08 ± 2.208	0.785 ± 0.079
PHE	559	0.899 ± 0.090	23.042 ± 2.304	0.845 ± 0.085
ASN	560	0.921 ± 0.092	23.568 ± 2.357	0.828 ± 0.083
LYS	561	0.936 ± 0.094	21.863 ± 2.186	0.84 ± 0.084
THR	562	0.939 ± 0.094	21.18 ± 2.118	0.812 ± 0.081
LEU	564	0.981 ± 0.098	22.327 ± 2.233	0.83 ± 0.083
TYR	565	0.98 ± 0.098	21.937 ± 2.194	0.837 ± 0.084
LYS	566	0.995 ± 0.100	22.094 ± 2.209	0.774 ± 0.077
LEU	567	0.916 ± 0.092	23.105 ± 2.311	0.827 ± 0.083
HIS	568	0.952 ± 0.095	20.609 ± 2.061	0.842 ± 0.084
VAL	569	0.98 ± 0.098	20.766 ± 2.077	0.818 ± 0.082
GLU	570	0.961 ± 0.096	18.97 ± 1.897	0.804 ± 0.080
ASN	571	0.988 ± 0.099	22.616 ± 2.262	0.864 ± 0.086
ALA	572	0.944 ± 0.094	20.988 ± 2.099	0.841 ± 0.084
THR	573	0.976 ± 0.098	22.59 ± 2.259	0.845 ± 0.085
ASN	574	1.023 ± 0.102	21.277 ± 2.128	0.856 ± 0.086
THR	576	0.918 ± 0.092	20.117 ± 2.012	0.82 ± 0.082
HIS	577	0.895 ± 0.090	21.046 ± 2.105	0.844 ± 0.084
PHE	578	0.914 ± 0.091	20.038 ± 2.004	0.826 ± 0.083
ARG	579	0.899 ± 0.090	21.024 ± 2.102	0.822 ± 0.082
LEU	580	0.986 ± 0.099	19.95 ± 1.995	0.81 ± 0.081
LEU	581	0.967 ± 0.097	21.981 ± 2.198	0.831 ± 0.083
LYS	582	0.929 ± 0.093	20.042 ± 2.004	0.798 ± 0.080
PHE	584	0.938 ± 0.094	17.654 ± 1.765	0.756 ± 0.076
GLY	585	1.059 ± 0.106	14.977 ± 1.348	0.307 ± 0.031
THR	586	1.036 ± 0.104	20.519 ± 1.847	0.719 ± 0.072
VAL	587	1.02 ± 0.102	20.347 ± 1.831	0.784 ± 0.078
LYS	588	0.98 ± 0.098	19.599 ± 1.764	0.814 ± 0.081
ILE	589	0.965 ± 0.097	21.146 ± 1.903	0.806 ± 0.081
PHE	590	0.952 ± 0.095	19.913 ± 1.792	0.807 ± 0.081
GLU	591	0.957 ± 0.096	21.262 ± 1.914	0.841 ± 0.084
VAL	592	0.924 ± 0.092	19.805 ± 1.782	0.812 ± 0.081
LYS	593	0.969 ± 0.097	20.806 ± 2.081	0.681 ± 0.068

Appendix D - Backbone dynamics of *A.fulgidus* AglB-S2 at 600 and 700MHz, using Model free Analysis

Table D-1 Backbone dynamics of *A.fulgidus* AglB-S2.

Residue		2^o	Status	Model	S^2			τ_e(ps)			$R_{ex}(s^{-1})$			Γ_i
GLU	433			2	0.767	±	0.041	1042	±	165				2.293
MET	434			4	0.783	±	0.052	1133	±	214	4.393	±	1.373	1.220
THR	435		ND											
MET	436	α_1	NI											
ASP	437	α_1	NI											
TRP	438	α_1		1	0.979	±	0.031							2.644
LYS	439	α_1		1	0.977	±	0.032							4.142
GLU	440	α_1		1	0.977	±	0.032							4.252
ALA	441	α_1		1	0.979	±	0.031							2.634
LEU	442	α_1		1	0.980	±	0.031							3.545
ASN	443	α_1		1	0.980	±	0.032							4.736
TRP	444	α_1		1	0.978	±	0.032							10.542
MET	445	α_1		1	0.976	±	0.036							6.593
LYS	446	α_1		1	0.979	±	0.032							3.703
GLU	447	α_1		1	0.980	±	0.029							5.832
ASN	448	α_1		1	0.979	±	0.030							2.405
LEU	449	$L_{\alpha1\beta A}$	ND											
GLU	450	$L_{\alpha1\beta A}$		2	0.826	±	0.046	1486	±	358				3.761
ALA	451	$L_{\alpha1\beta A}$		2	0.798	±	0.045	1234	±	295				6.514
GLN	452	$L_{\alpha1\beta A}$		2	0.745	±	0.043	1077	±	158				1.810
ASP	453	$L_{\alpha1\beta A}$		2	0.639	±	0.040	1256	±	120				2.695
TYR	454	$L_{\alpha1\beta A}$		2	0.803	±	0.046	967	±	250				7.932
LEU	455	$L_{\alpha1\beta A}$		2	0.693	±	0.047	1277	±	145				3.446
LYS	456	$L_{\alpha1\beta A}$	ND											
ALA	457	$L_{\alpha1\beta A}$		2	0.619	±	0.039	1273	±	97				3.267
TYR	458	$L_{\alpha1\beta A}$		2	0.712	±	0.048	1004	±	149				4.947
GLU	459	$L_{\alpha1\beta A}$		2	0.615	±	0.046	1182	±	96				3.397
LYS	460	$L_{\alpha1\beta A}$		2	0.598	±	0.039	1088	±	75				4.384
PRO	461	$L_{\alpha1\beta A}$	NI											
ASP	462	$L_{\alpha1\beta A}$		2	0.716	±	0.038	1253	±	150				1.856
TYR	463	$L_{\alpha1\beta A}$		4	0.853	±	0.045	878	±	262	4.761	±	1.172	0.308
ALA	464	β_A		1	0.980	±	0.029							6.364
VAL	465	β_A		1	0.982	±	0.027							6.537
LEU	466	β_A		1	0.979	±	0.031							7.488
SER	467	α_2		1	0.978	±	0.034							8.500
TRP	468	α_2		3	0.970	±	0.043				5.804	±	1.535	2.425
TRP	469	α_2	NI											
ASP	470	α_2		3	0.967	±	0.047				6.694	±	1.601	8.639
TYR	471	α_2		3	0.969	±	0.043				6.979	±	1.484	3.300
GLY	472	α_2												
ASN	473	α_2		3	0.976	±	0.038				4.336	±	1.637	6.516
TRP	474	α_2		1	0.975	±	0.031							4.223
ILE	475	α_2		1	0.976	±	0.032							9.584
LEU	476	α_2		1	0.980	±	0.030							8.944
TYR	477	α_2		3	0.970	±	0.044				6.210	±	1.573	2.584
VAL	478	α_2		3	0.971	±	0.043				10.170	±	1.762	3.742
ALA	479	α_2		3	0.967	±	0.049				8.737	±	1.812	2.224
LYS	480	$L_{\alpha2\alpha3}$		3	0.970	±	0.042				5.883	±	1.616	2.600
LYS	481	$L_{\alpha2\alpha3}$		1	0.978	±	0.034							5.696
ALA	482	$L_{\alpha2\alpha3}$		1	0.978	±	0.031							10.826
VAL	483	$L_{\alpha2\alpha3}$	NM											
VAL	484	$L_{\alpha2\alpha3}$	NI											
CYS	485	$L_{\alpha2\alpha3}$	NI											

Table D-2 Backbone dynamics of *A.fulgidus* AglB-S2 continued.

Residue		2°	Status	Model	S^2			τ_e(ps)			$R_{ex}(s^{-1})$			Γ_i
ASN	486	$L_{\alpha2\alpha3}$	NI											
ASN	487	$L_{\alpha2\alpha3}$	NI											
PHE	488	$L_{\alpha2\alpha3}$	NI											
GLN	489	$L_{\alpha2\alpha3}$		4	0.823	±	0.049	1442	±	336	6.332	±	1.484	0.703
ALA	490	$L_{\alpha2\alpha3}$		2	0.820	±	0.038	628	±	187				3.124
ALA	491	$L_{\alpha2\alpha3}$	Low											
ALA	492	α_3		1	0.976	±	0.033							9.080
ASP	493	α_3		1	0.978	±	0.032							8.420
ASP	494	α_3		1	0.980	±	0.029							5.372
ALA	495	α_3		1	0.979	±	0.032							4.196
ALA	496	α_3		1	0.976	±	0.033							4.366
LYS	497	α_3		1	0.978	±	0.032							4.293
PHE	498	α_3	Low											
PHE	499	α_3		1	0.977	±	0.034							3.967
THR	500	α_3	Low											
ALA	501	$L_{\alpha3\alpha4}$		1	0.980	±	0.031							1.924
GLN	502	$L_{\alpha3\alpha4}$		1	0.981	±	0.029							6.675
SER	503	$L_{\alpha3\alpha4}$		2	0.896	±	0.033	588	±	319				0.695
GLU	504	α_4	NI											
GLU	505	α_4		1	0.979	±	0.033							8.102
GLU	506	α_4		1	0.979	±	0.030							2.687
ALA	507	α_4		1	0.983	±	0.028							2.912
MET	508	α_4		1	0.977	±	0.031							2.418
LYS	509	α_4		1	0.976	±	0.032							4.115
ILE	510	α_4		1	0.981	±	0.028							3.620
VAL	511	α_4		1	0.982	±	0.027							4.014
GLU	512	α_4		1	0.979	±	0.031							3.903
LYS	513	α_4		1	0.979	±	0.030							3.475
ARG	514	α_4		1	0.982	±	0.029							3.035
LYS	515	$L_{\alpha4\beta B}$		1	0.976	±	0.034							9.348
VAL	516	$L_{\alpha4\beta B}$		1	0.981	±	0.032							4.046
ARG	517	$L_{\alpha4\beta B}$		1	0.981	±	0.029							7.151
TYR	518	β_B		1	0.979	±	0.031							3.100
VAL	519	β_B		1	0.981	±	0.031							2.664
VAL	520	β_B		1	0.979	±	0.029							1.483
THR	521	β_B		1	0.981	±	0.027							3.752
VAL	522	β_B		1	0.978	±	0.031							2.739
GLU	523	α_5		1	0.980	±	0.030							1.788
GLU	524	α_5		1	0.978	±	0.030							9.865
LEU	525	α_5		1	0.976	±	0.032							7.653
THR	526	$L_{\alpha5\alpha6}$		1	0.979	±	0.032							2.386
VAL	527	$L_{\alpha5\alpha6}$	NI											
ALA	528	$L_{\alpha5\alpha6}$	NI											
PRO	529	$L_{\alpha5\alpha6}$	NI											
GLU	530	$L_{\alpha5\alpha6}$		2	0.768	±	0.042	1050	±	178				3.084
THR	531	$L_{\alpha5\alpha6}$		2	0.801	±	0.047	988	±	266				7.105
ASN	532	$L_{\alpha5\alpha6}$	NI											
LYS	533	$L_{\alpha5\alpha6}$	NI											
THR	534	$L_{\alpha5\alpha6}$		4	0.822	±	0.049	793	±	216	7.987	±	1.547	0.662
LYS	535	α_6		3	0.970	±	0.046				5.335	±	1.631	4.134
PHE	536	α_6	Low											
ILE	537	α_6		3	0.972	±	0.044				4.903	±	1.522	2.936
ILE	538	α_6	NI											
ILE	539	α_6		1	0.979	±	0.031							3.149
MET	540	α_6		1	0.979	±	0.031							3.288

Table D-3 Backbone dynamics of *A.fulgidus* AglB-S2 continued.

Residue		2^o	Status	Model	S^2			τ_e(ps)			$R_{ex}(s^{-1})$	Γ_i^*
GLN	541	α_6		1	0.978	±	0.033					4.178
ILE	542	α_6		1	0.979	±	0.030					4.231
ALA	543	α_6		1	0.977	±	0.032					5.184
GLY	544	$L_{\alpha6\alpha7}$		1	0.980	±	0.029					3.064
TYR	545	$L_{\alpha6\alpha7}$		1	0.983	±	0.026					2.138
SER	546	$L_{\alpha6\alpha7}$		2	0.697	±	0.043	1622	±	224		1.433
PRO	547	$L_{\alpha6\alpha7}$	NI									
GLU	548	$L_{\alpha6\alpha7}$		1	0.978	±	0.032					9.501
TYR	549	$L_{\alpha6\alpha7}$		2	0.907	±	0.032	709	±	404		1.879
MET	550	$L_{\alpha6\alpha7}$		1	0.953	±	0.039					10.350
LYS	551	$L_{\alpha6\alpha7}$		1	0.958	±	0.038					5.083
ASN	552	α_7	NI									
LYS	553	α_7		1	0.978	±	0.033					7.246
GLU	554	α_7	Low									
ILE	555	α_7		1	0.980	±	0.030					2.000
ILE	556	α_7		1	0.981	±	0.030					5.567
ASP	557	α_7		1	0.979	±	0.031					3.788
PHE	558	α_7		1	0.974	±	0.035					14.307
PHE	559	α_7		1	0.981	±	0.030					3.433
ASN	560	α_7		1	0.979	±	0.032					3.919
LYS	561	α_7		1	0.981	±	0.028					3.272
THR	562			1	0.980	±	0.031					3.211
MET	563	α_8	Low									
LEU	564	α_8		1	0.980	±	0.031					3.862
TYR	565	α_8		1	0.981	±	0.029					3.987
LYS	566	α_8		1	0.977	±	0.034					4.874
LEU	567	α_8		1	0.980	±	0.028					4.983
HIS	568	α_8		1	0.978	±	0.030					1.755
VAL	569	α_8		1	0.979	±	0.032					2.706
GLU	570	α_8		1	0.982	±	0.027					1.186
ASN	571	$L_{\alpha8\beta C}$		1	0.975	±	0.035					5.092
ALA	572	$L_{\alpha8\beta C}$		1	0.978	±	0.030					1.798
THR	573	$L_{\alpha8\beta C}$		1	0.979	±	0.030					4.871
ASN	574	$L_{\alpha8\beta C}$		1	0.980	±	0.029					3.802
LEU	575	$L_{\alpha8\beta C}$	ND									
THR	576	$L_{\alpha8\beta C}$		1	0.964	±	0.038					2.361
HIS	577	$L_{\alpha8\beta C}$		1	0.976	±	0.032					2.844
PHE	578	β_C		1	0.978	±	0.032					1.225
ARG	579	β_C		1	0.981	±	0.028					1.854
LEU	580	β_C		1	0.982	±	0.027					2.441
LEU	581	β_C		1	0.978	±	0.031					4.034
LYS	582	β_C		1	0.977	±	0.033					1.384
ASN	583	β_C	NI									
PHE	584	$L_{\beta C\beta D}$		1	0.964	±	0.035					2.263
GLY	585	$L_{\beta C\beta D}$		2	0.760	±	0.035	585	±	101		6.260
THR	586	$L_{\beta C\beta D}$		1	0.978	±	0.031					9.608
VAL	587	β_D		1	0.983	±	0.027					5.436
LYS	588	β_D		1	0.982	±	0.028					2.026
ILE	589	β_D		1	0.982	±	0.027					3.006
PHE	590	β_D		1	0.980	±	0.028					2.913
GLU	591	β_D		1	0.978	±	0.030					3.113
VAL	592	β_D		1	0.980	±	0.029					1.477
LYS	593			2	0.909	±	0.038	784	±	462		1.880

* Γ_i is the sum-of-squared-residuals for the *i* th spin

Appendix E - Constant-time CPMG relaxation dispersion of *A.fulgidus* AglB-S2 residues assigned model 2 by NESSY at 600MHz, 308 K

Table E-1 Representative residues fit to 2-site fast-limit exchange by NESSY ver. 12.2.1

Residue		Model	R_2 (s^{-1})			k_{ex} (s^{-1})			R_{ex} exp^1($Rads^{-1}$)♦		
MET	434	2	24.92	±	3.25	1545.20	±	1194.90	7.47	±	3.00
LYS	446	2	21.11	±	0.05	500.00	±	26.05	4.53	±	0.12
GLU	447	2	17.67	±	0.59	2370.78	±	299.44	6.49	±	0.56
ASN	448	2	19.70	±	0.11	841.75	±	64.24	4.49	±	0.15
GLN	452	2	22.39	±	0.34	500.00	±	89.13	10.16	±	0.79
TYR	454	2	17.88	±	0.28	1904.47	±	121.57	7.42	±	0.25
ASP	462	2	16.02	±	0.07	500.00	±	21.78	6.79	±	0.18
TYR	463	2	16.14	±	1.03	1976.25	±	287.65	10.76	±	0.95
VAL	465	2	21.19	±	0.25	776.72	±	100.05	6.67	±	0.38
SER	467	2	23.04	±	0.54	1158.43	±	188.51	7.60	±	0.51
TRP	468	2	22.76	±	3.28	3044.36	±	1678.44	6.89	±	3.23
ASN	473	2	19.54	±	2.08	2812.90	±	857.99	8.33	±	2.02
TYR	477	2	15.71	±	2.87	4687.64	±	992.11	15.16	±	2.82
ALA	482	2	17.63	±	0.23	1731.15	±	90.05	7.61	±	0.20
VAL	483	2	34.88	±	1.96	1452.06	±	479.54	11.56	±	1.70
ALA	492	2	20.08	±	0.12	821.79	±	50.77	6.19	±	0.16
ASP	494	2	15.10	±	2.45	3624.43	±	723.85	13.92	±	2.40
ARG	514	2	20.41	±	0.05	907.46	±	25.46	5.05	±	0.06
VAL	519	2	20.17	±	0.11	547.95	±	42.99	6.15	±	0.30
GLU	523	2	24.14	±	0.26	679.63	±	141.27	5.17	±	0.49
LEU	525	2	21.27	±	0.03	500.00	±	17.70	4.18	±	0.07
THR	534	2	28.71	±	0.29	500.00	±	80.80	8.42	±	0.64
LYS	535	2	23.20	±	1.17	2336.55	±	371.29	10.42	±	1.11
LYS	582	2	18.72	±	0.06	510.31	±	33.19	4.57	±	0.19

♦ In NMR, rates are measured in rad s^{-1} and so the conversion from rad s^{-1} to frequency, is 1 rad s^{-1}= 0.159Hz

References

Abu-Qarn, M., S. Yurist-Doutsch, A. Giordano, A. Trauner, H.R. Morris, P. Hitchen, O. Medalia, A. Dell, and J. Eichler, 2007. Haloferax volcanii AglB and AglD are involved in N-glycosylation of the S-layer glycoprotein and proper assembly of the surface layer. J Mol Biol 374: 1224-36.

Aebi, M., R. Bernasconi, S. Clerc, and M. Molinari, 2010. N-glycan structures: recognition and processing in the ER. Trends Biochem Sci 35: 74-82.

Agarwal, P.K., 2005. Role of protein dynamics in reaction rate enhancement by enzymes. J Am Chem Soc 127: 15248-56.

Bieri, M., and P.R. Gooley, 2011. Automated NMR relaxation dispersion data analysis using NESSY. BMC Bioinformatics 12: 421.

Brunger, A.T., P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, and G.L. Warren, 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 54: 905-21.

Carugo, O., and P. Argos, 1997. Protein-protein crystal-packing contacts. Protein Sci 6: 2261-3.

Chen, M.M., K.J. Glover, and B. Imperiali, 2007. From peptide to protein: comparative analysis of the substrate specificity of N-linked glycosylation in C. jejuni. Biochemistry 46: 5579-85.

Chopra, G., C.M. Summa, and M. Levitt, 2008. Solvent dramatically affects protein structure refinement. Proc Natl Acad Sci U S A 105: 20239-44.

Cole, R., and J.P. Loria, 2003. FAST-Modelfree: a program for rapid automated analysis of solution NMR spin-relaxation data. J Biomol NMR 26: 203-13.

Daniel, R.M., R.V. Dunn, J.L. Finney, and J.C. Smith, 2003. The role of dynamics in enzyme activity. Annu Rev Biophys Biomol Struct 32: 69-92.

Delaglio, F., S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6: 277-93.

Eichler, J., 2000. Novel glycoproteins of the halophilic archaeon Haloferax volcanii. Arch Microbiol 173: 445-8.

Emsley, P., and K. Cowtan, 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126-32.

Eyal, E., S. Gerzon, V. Potapov, M. Edelman, and V. Sobolev, 2005. The limit of accuracy of protein modeling: influence of crystal packing on protein structure. J Mol Biol 351: 431-42.

Farrow, N.A., R. Muhandiram, A.U. Singer, S.M. Pascal, C.M. Kay, G. Gish, S.E. Shoelson, T. Pawson, J.D. Forman-Kay, and L.E. Kay, 1994. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. Biochemistry 33: 5984-6003.

Gill, M.B., P. Vivithanaporn, and G.T. Swanson, 2009. Glutamate binding and conformational flexibility of ligand-binding domains are critical early determinants of efficient kainate receptor biogenesis. J Biol Chem 284: 14503-12.

Gilmore, R., 2011. Structural biology: Porthole to catalysis. Nature 474: 292-3.

Horan, L.H., and H.F. Noller, 2007. Intersubunit movement is required for ribosomal translocation. Proc Natl Acad Sci U S A 104: 4881-5.
Igura, M., and D. Kohda, 2011a. Selective control of oligosaccharide transfer efficiency for the N-glycosylation sequon by a point mutation in oligosaccharyltransferase. J Biol Chem 286: 13255-60.
Igura, M., and D. Kohda, 2011b. Quantitative assessment of the preferences for the amino acid residues flanking archaeal N-linked glycosylation sites. Glycobiology 21: 575-83.
Igura, M., N. Maita, T. Obita, J. Kamishikiryo, K. Maenaka, and D. Kohda, 2007. Purification, crystallization and preliminary X-ray diffraction studies of the soluble domain of the oligosaccharyltransferase STT3 subunit from the thermophilic archaeon Pyrococcus furiosus. Acta Crystallogr Sect F Struct Biol Cryst Commun 63: 798-801.
Igura, M., N. Maita, J. Kamishikiryo, M. Yamada, T. Obita, K. Maenaka, and D. Kohda, 2008. Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. EMBO J 27: 234-43.
Karaoglu, D., D.J. Kelleher, and R. Gilmore, 1997. The highly conserved Stt3 protein is a subunit of the yeast oligosaccharyltransferase and forms a subcomplex with Ost3p and Ost4p. J Biol Chem 272: 32513-20.
Karaoglu, D., S. Silberstein, D.J. Kelleher, and R. Gilmore, 1995. The Saccharomyces cerevisiae oligosaccharyltransferase: a large hetero-oligomeric complex in the endoplasmic reticulum. Cold Spring Harb Symp Quant Biol 60: 83-92.
Katoh, K., and H. Toh, 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform 9: 286-98.
Kelleher, D.J., and R. Gilmore, 1994. The Saccharomyces cerevisiae oligosaccharyltransferase is a protein complex composed of Wbp1p, Swp1p, and four additional polypeptides. J Biol Chem 269: 12908-17.
Kelleher, D.J., and R. Gilmore, 2006. An evolving view of the eukaryotic oligosaccharyltransferase. Glycobiology 16: 47R-62R.
Kelleher, D.J., D. Karaoglu, E.C. Mandon, and R. Gilmore, 2003. Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties. Mol Cell 12: 101-11.
Kern, D., E.Z. Eisenmesser, and M. Wolf-Watz, 2005. Enzyme dynamics during catalysis measured by NMR spectroscopy. Methods Enzymol 394: 507-24.
Knauer, R., and L. Lehle, 1994. The N-oligosaccharyltransferase complex from yeast. FEBS Lett 344: 83-6.
Kohda, D., M. Yamada, M. Igura, J. Kamishikiryo, and K. Maenaka, 2007. New oligosaccharyltransferase assay method. Glycobiology 17: 1175-82.
Konrad, Z., and J. Eichler, 2002. Protein glycosylation in Haloferax volcanii: partial characterization of a 98-kDa glycoprotein. FEMS Microbiol Lett 209: 197-202.
Kowarik, M., S. Numao, M.F. Feldman, B.L. Schulz, N. Callewaert, E. Kiermaier, I. Catrein, and M. Aebi, 2006a. N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. Science 314: 1148-50.
Kowarik, M., N.M. Young, S. Numao, B.L. Schulz, I. Hug, N. Callewaert, D.C. Mills, D.C. Watson, M. Hernandez, J.F. Kelly, M. Wacker, and M. Aebi, 2006b. Definition of the bacterial N-glycosylation site consensus sequence. EMBO J 25: 1957-66.
Krissinel, E., 2010. Crystal contacts as nature's docking solutions. J Comput Chem 31: 133-43.
Larsen, J.C., C. Szymanski, and P. Guerry, 2004. N-linked protein glycosylation is required for full competence in Campylobacter jejuni 81-176. J Bacteriol 186: 6508-14.

Lehle, L., 1992. Protein glycosylation in yeast. Antonie Van Leeuwenhoek 61: 133-4.
Lipari, G., and A. Szabo, 1982a. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. J Am Chem Soc 104: 4559-4570.
Lipari, G., and A. Szabo, 1982b. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. J Am Chem Soc 104: 4546-4559.
Liu, J., and A. Mushegian, 2003. Three monophyletic superfamilies account for the majority of the known glycosyltransferases. Protein Sci 12: 1418-31.
Lizak, C., S. Gerber, S. Numao, M. Aebi, and K.P. Locher, 2011. X-ray structure of a bacterial oligosaccharyltransferase. Nature 474: 350-5.
Loria, J.P., M. Rance, and A.G. Palmer, 3rd, 1999. A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. J Am Chem Soc 121: 2331-2332.
Maita, N., J. Nyirenda, M. Igura, J. Kamishikiryo, and D. Kohda, 2010. Comparative structural biology of eubacterial and archaeal oligosaccharyltransferases. J Biol Chem 285: 4941-50.
Makmura, L., M. Hamann, A. Areopagita, S. Furuta, A. Munoz, and J. Momand, 2001. Development of a sensitive assay to detect reversibly oxidized protein cysteine sulfhydryl groups. Antioxid Redox Signal 3: 1105-18.
Mandel, A.M., M. Akke, and A.G. Palmer, 3rd, 1995. Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme. J Mol Biol 246: 144-63.
Matsumoto, S., M. Igura, J. Nyirenda, M. Matsumoto, S. Yuzawa, N. Noda, F. Inagaki, and D. Kohda, 2012. Crystal structure of the C-terminal globular domain of oligosaccharyltransferase from Archaeoglobus fulgidus at 1.75 A resolution. Biochemistry 51: 4157-66.
May, A.C., 2004. Percent sequence identity; the need to be explicit. Structure 12: 737-8.
Murshudov, G.N., A.A. Vagin, and E.J. Dodson, 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53: 240-55.
Nasab, F.P., B.L. Schulz, F. Gamarro, A.J. Parodi, and M. Aebi, 2008. All in one: Leishmania major STT3 proteins substitute for the whole oligosaccharyltransferase complex in Saccharomyces cerevisiae. Mol Biol Cell 19: 3758-68.
Nothaft, H., and C.M. Szymanski, 2010. Protein glycosylation in bacteria: sweeter than ever. Nat Rev Microbiol 8: 765-78.
Nothaft, H., X. Liu, D.J. McNally, and C.M. Szymanski, 2010. N-linked protein glycosylation in a bacterial system. Methods Mol Biol 600: 227-43.
Otwinowski, Z., and W. Minor, 1997. Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods in Enzymology 276: 307-326.
Palle, K., L. Pattarello, M. van der Merwe, C. Losasso, P. Benedetti, and M.A. Bjornsti, 2008. Disulfide cross-links reveal conserved features of DNA topoisomerase I architecture and a role for the N terminus in clamp closure. J Biol Chem 283: 27767-75.
Palmer, A.G., 3rd, 1997. Probing molecular motion by NMR. Curr Opin Struct Biol 7: 732-7.
Palmer, A.G., 3rd, C.D. Kroenke, and J.P. Loria, 2001. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. Methods Enzymol 339: 204-38.

Peske, F., N.B. Matassova, A. Savelsbergh, M.V. Rodnina, and W. Wintermeyer, 2000. Conformationally restricted elongation factor G retains GTPase activity but is inactive in translocation on the ribosome. Mol Cell 6: 501-5.

Sharma, C.B., L. Lehle, and W. Tanner, 1981. N-Glycosylation of yeast proteins. Characterization of the solubilized oligosaccharyl transferase. Eur J Biochem 116: 101-8.

Standley, D.M., H. Toh, and H. Nakamura, 2005. GASH: an improved algorithm for maximizing the number of equivalent residues between two protein structures. BMC Bioinformatics 6: 221.

Szymanski, C.M., and B.W. Wren, 2005. Protein glycosylation in bacterial mucosal pathogens. Nat Rev Microbiol 3: 225-37.

Szymanski, C.M., D.H. Burr, and P. Guerry, 2002. Campylobacter protein glycosylation affects host cell interactions. Infect Immun 70: 2242-4.

Szymanski, C.M., S.M. Logan, D. Linton, and B.W. Wren, 2003a. Campylobacter--a tale of two protein glycosylation systems. Trends Microbiol 11: 233-8.

Szymanski, C.M., F.S. Michael, H.C. Jarrell, J. Li, M. Gilbert, S. Larocque, E. Vinogradov, and J.R. Brisson, 2003b. Detection of conserved N-linked glycans and phase-variable lipooligosaccharides and capsules from campylobacter cells by mass spectrometry and high resolution magic angle spinning NMR spectroscopy. J Biol Chem 278: 24509-20.

Tanner, W., and L. Lehle, 1987. Protein glycosylation in yeast. Biochim Biophys Acta 906: 81-99.

Tjandra, N., S.E. Feller, R.W. Pastor, and A. Bax, 1995. Rotational diffusion anisotropy of human ubiquitin from 15N NMR relaxation. J Am Chem Soc 117: 12562-12566.

Tjandra, N., P. Wingfield, S. Stahl, and A. Bax, 1996. Anisotropic rotational diffusion of perdeuterated HIV protease from 15N NMR relaxation measurements at two magnetic fields. J Biomol NMR 8: 273-84.

Tollinger, M., N.R. Skrynnikov, F.A. Mulder, J.D. Forman-Kay, and L.E. Kay, 2001. Slow dynamics in folded and unfolded states of an SH3 domain. J Am Chem Soc 123: 11341-52.

Vagin, A., and A. Teplyakov, 2010. Molecular replacement with MOLREP. Acta Crystallogr D Biol Crystallogr 66: 22-5.

Weerapana, E., and B. Imperiali, 2006. Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. Glycobiology 16: 91R-101R.

Yan, Q., and W.J. Lennarz, 2002. Studies on the function of oligosaccharyl transferase subunits. Stt3p is directly involved in the glycosylation process. J Biol Chem 277: 47692-700.

Yurist-Doutsch, S., H. Magidovich, V.V. Ventura, P.G. Hitchen, A. Dell, and J. Eichler, 2010. N-glycosylation in Archaea: On the coordinated actions of Haloferax volcanii AglF and AglM. Mol Microbiol.

Yurist-Doutsch, S., M. Abu-Qarn, F. Battaglia, H.R. Morris, P.G. Hitchen, A. Dell, and J. Eichler, 2008. AglF, aglG and aglI, novel members of a gene island involved in the N-glycosylation of the Haloferax volcanii S-layer glycoprotein. Mol Microbiol 69: 1234-45.

Zhang, X.J., J.A. Wozniak, and B.W. Matthews, 1995. Protein flexibility and adaptability seen in 25 crystal forms of T4 lysozyme. J Mol Biol 250: 527-52.