Crystallographic and NMR evidence for flexibility in oligosaccharyltransferases and its catalytic significance

Asparagine-linked glycosylation (N-glycosylation) of proteins is widespread not only in eukaryotes, but also in archaea and some eubacteria. The entire process of protein N-glycosylation is divided into three steps: oligosaccharide donor synthesis, oligosaccharyl transfer reaction, and oligosaccharide processing (Fig. 1). The first step is the production of the oligosaccharide donor, the lipid-linked oligosaccharide (LLO). An oligosaccharide chain is assembled on a lipid-phospho carrier. In principle, the carrier is dolichol diphosphate in eukaryotes and archaea, and undecaprenol diphosphate in eubacteria, with the exceptions of halophilic and methanogenic archaea, which use dolichol monophosphate. The second step is the transfer of the oligosaccharide chain, preassembled on the lipid-phospho carrier, onto proteins. A membrane-bound enzyme, oligosaccharyltransferase (OST), catalyzes the formation of a covalent bond between the sugar residue of the reducing end of the oligosaccharide chain and the side-chain amide groups of the asparagine residues in proteins. The catalytic subunit of the OST enzyme is a polypeptide chain referred to as "STT3" (staurosporine and temperature sensitivity 3) in eukaryotes, "AglB" (archaeal glycosylation B) in archaea, and "PglB" (protein glycosylation B) in eubacteria, although they originated from a common ancestor. The eukaryotic OST is a multisubunit protein complex containing STT3, but the lower eukaryotic protozoan OST, and the archaeal and eubacterial OSTs, are single subunit enzymes consisting only of the STT3/AglB/PglB protein. In principle, the acceptor asparagine resides in the N-glycosylation sequon (Asn-X-Ser/Thr; X ≠ Pro), but glycosylation also occurs in atypical sequences, such as Asn-X-Cys, Asn-Gly-X, and Asn-X-Val (X ≠ Pro), albeit at very low frequencies. Eubacteria use an extended and thus more restrictive version of the sequon (Asp/Glu-X-Asn-Y-Ser/Thr; X, Y ≠ Pro) .

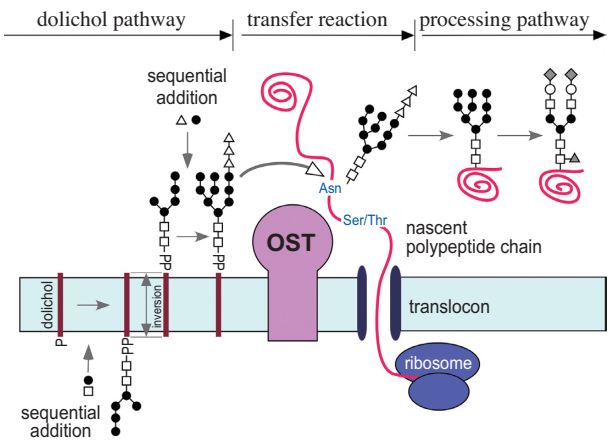

Fig. 1. Schematic representation of the N-glycosylation reaction. Oligosaccharyltransferase is the central enzyme of N-linked protein glycosylation.

The STT3/AglB/PglB proteins consist of an N-terminal multispan transmembrane region and a C-terminal globular domain. Multiple STT3/AglB/PglB proteins may be encoded in a single genome. As exemplified by the AglBs referred to in this study, *Archaeoglobus fulgidus* contains three paralogs. We discriminate among these archaeal paralogs with a letter plus an optional number, such as L (long) or S1 (short, number 1).

We previously determined the crystal structures of the C-terminal globular domains of *Pyrococcus furiosus* AglB-L [1], *Archaeoglobus fulgidus* AgB-S1 [2], and *Campylobacter jejuni* PglB [3]. The comparison of the three crystal structures unexpectedly revealed significant local variations in the conformations of about 25-residue segments in the C-terminal globular domains [2].

In this study, we have determined two crystal structures of the C-terminal globular domains of *P. horikoshii* AglB-L and *A. fulgidus* AglB-S2 [4]. The one-to-one structural comparison of the five crystal structures confirmed that the C-terminal globular domains of the AglB/PglB proteins contained the special plastic segment, and identified the resting state conformation of the plastic segment, free of crystal contact effects (Fig. 2).

We characterized the dynamic properties of the plastic segment in solution by ^{15}N NMR relaxation analyses. We selected the C-terminal domain of *A. fulgidus* AglB-S2 because it is the smallest among the three domains of life. Intriguingly, the mobile region contains the binding pocket for the recognition of the +2 Ser/Thr residue in the consensus sequence (Fig. 2). In agreement, the flexibility restriction forced by an engineered disulfide crosslink abolished the enzymatic activity, and its cleavage fully restored the activity (Fig. 3). These results suggest the necessity of multiple conformational states in the reaction. The dynamic nature of the Ser/Thr pocket could facilitate the efficient scanning of N-glycosylation sequons along nascent polypeptide chains.

We demonstrated that comparative structural biology is an effective approach for analyzing the

structural and dynamic relationships of proteins. The 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 (Fig. 2). The comparative structural approach is also useful for the selection of the most suitable protein for certain purposes; the smallest AglB was selected for the NMR relaxation study, and the thermostable AglB was suitable for the custom-designed mutagenesis study, which required high protein stability to tolerate the boldly designed mutations (Fig. 3).

The isolated C-terminal domains are catalytically inactive, and thus we must know the three-dimensional structures of the full-length STT3/AglB/PglB proteins to elucidate the catalytic mechanism of OST. We determined the crystal structure of the full-length *A. fulgidus* AglB-L in two crystal forms at 2.5 Å (space group *C*2 in the presence of *n*-octyl-β-D-glucopyranoside) and 3.4 Å ($P4_32_12$ in lauryldimethylamine-oxide) resolutions [5]. X-ray diffraction data were collected at beamlines **BL44XU** and **BL32XU**. This is the second solved OST structure, after the eubacterial *C. lari* PglB determined by another group. The structural comparison of the two full-length OST proteins confirmed and extended the dynamic view of the N-glycosylation catalytic cycle, involving the plastic EL5 loop in the transmembrane region (Fig. 3), and the dynamic Ser/Thr binding pocket in the CC unit [5].

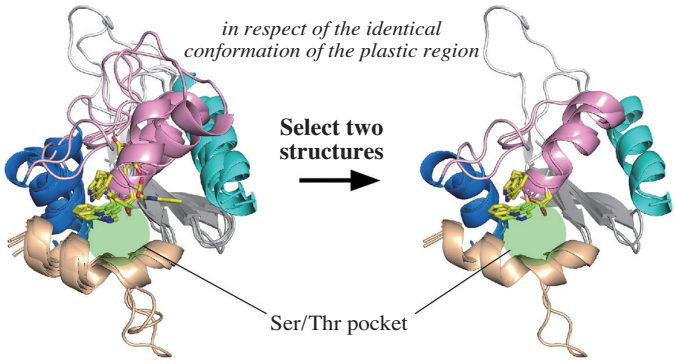

Fig. 2. Superposition of the CC structural unit in the five crystal structures of the C-terminal globular domains of the AglB/PglB proteins. The region that exhibits large plasticity includes the WWDYG motif (yellow side chains) and the following α-helical and loop regions (pink backbones). The right panel shows the two selected structures that share the identical conformation of the plastic region. This conformation is assumed to be the resting state conformation. The shaded circles indicate the position of the +2 Ser/Thr pocket.

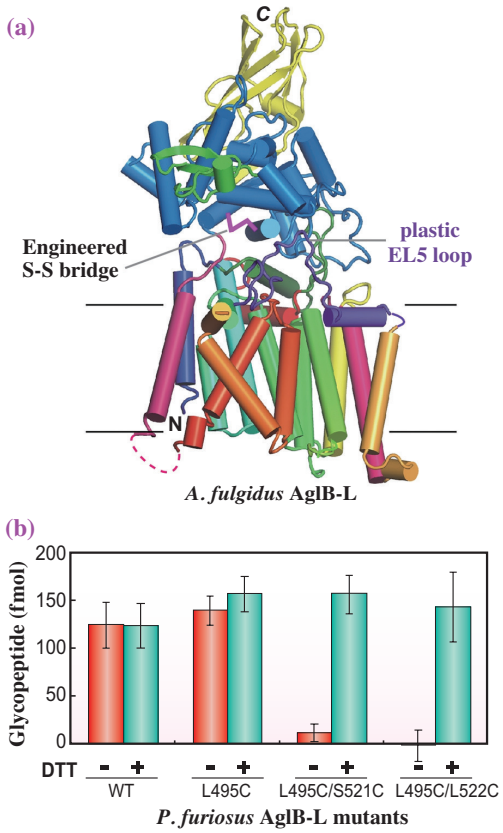

Fig. 3. Effects of the flexibility restriction in the C-terminal domain of *P. furiosus* AglB-L on OST activity. **(a)** The position of the engineered disulfide bond is shown by magenta sticks on the structure of the full-length *A. fulgidus* AglB-L protein. **(b)** Two double cysteine mutants (L495C/S521C and L495C/L522C) of the full-length *P. furiosus* AglB-L protein were constructed. As a negative control, a single cysteine mutant (L495C) was also generated. The *P. furiosus* AglB-L mutants bearing the engineered disulfide bond (red bars) had virtually no activity, whereas the same mutants without the disulfide crosslink after DTT treatment (blue bars) had completely restored activity.

James Nyirenda, Shunsuke Matsumoto
and Daisuke Kohda*

Medical Institute of Bioregulation, Kyushu University

*E-mail: kohda@bioreg.kyushu-u.ac.jp

References

[1] M. Igura *et al.*: EMBO J **27** (2008) 234.
[2] S. Matsumoto *et al.*: Biochemistry **51** (2012) 4157.
[3] N. Maita *et al.*: J. Biol. Chem. **285** (2010) 4941.
[4] *J. Nyirenda, S. Matsumoto, T. Saitoh, N. Maita, N.N. Noda, F. Inagaki and D. Kohda: Structure **21** (2013) 32.*
[5] S. Matsumoto *et al.*: Proc. Natl. Acad. Sci. USA **110** (2013) 17868.