

BMB Reports – Manuscript Submission

Manuscript Draft

Manuscript Number: BMB-18-269

Title: Structural insights showing how arginine is able to be glycosylated by pathogenic effector proteins

Article Type: Perspective (Invited Only)

Keywords: Arginine glycosylation; Glycosyltransferase; Effector protein; NleB; SseK

Corresponding Author: Hyun-Soo Cho

Authors: Jun Bae Park¹, Youngki Yoo¹, Hyun-Soo Cho^{1,*}

Institution: ¹Department of Systems Biology and College of Life Science and Biotechnology, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea,

Structural insights showing how arginine is able to be glycosylated by pathogenic effector proteins

*Jun Bae Park, Youngki Yoo & Hyun-Soo Cho**

Department of Systems Biology and College of Life Science and Biotechnology, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea

*Corresponding author. E-mail: hscho8@gmail.com

E-mail address

hscho8@gmail.com (Hyun-Soo Cho), jbpark0323@yonsei.ac.kr (Jun Bae Park), think8989@gmail.com (Youngki Yoo)

Keywords

Arginine glycosylation, Glycosyltransferase, Effector protein, NleB, SseK

Abbreviations

NleB, Non-LEE encoded effector protein B; SseK, Salmonella SPI-2 effector K; UDP-GlcNAc, Uridine diphosphate N-acetylglucosamine; NMR, Nuclear magnetic resonance; EHEC, Enterohemorrhagic E. coli; EPEC, Enteropathogenic E. coli; DD, Death domain; HLH, Helix-loop-helix; MD, Molecular dynamics

Perspective to

Jun Bae Park and Young Hun Kim et al. (2018) Structural basis for arginine glycosylation of host substrates by bacterial effector proteins, Nature Communications, 9:4283, DOI: 10.1038/s41467-018-06680-6

Abstract

Glycosylation is one form of protein modification and plays a key role in protein stability, function, signaling regulation and even cancer. NleB and SseK are bacterial effector proteins and possess glycosyltransferase activity, even though they have different substrate preferences. NleB/SseKs transfer the GlcNAc sugar to an arginine residue of host proteins, leading to reduced NF- κ B-dependent responses. By combining X-ray crystallography, NMR, molecular dynamics, enzyme kinetic assays and *in vivo* experiments, we demonstrated that a conserved HEN (His-Glu-Asn) motif in the active site plays a key role in enzyme catalysis and virulence. The lid-domain regulates the opening and closing of the active site and the HLH domain determines the substrate specificity. Our findings provide evidence for the enzymatic mechanism by which arginine can be glycosylated by SseK/NleB enzymes.

Body

Effector proteins from pathogenic bacteria disrupt the host cell immune system critical for sustaining bacterial life. NleB is one of the effector proteins from EHEC or EPEC bacteria and possesses glycosyltransferase activity. TRADD, FADD, RIPK1 DD, TNFR1 DD and GAPDH are the glycosylation targets of NleB. These substrates play a key role in the apoptosis signaling pathway, but substrates glycosylated by NleB fail to participate in NF- κ B mediated apoptosis (Pearson J. S. *et al.* (2013) Nature 501, 247-251, Li S. *et al.* (2013) Nature 501, 242-246, Gao, X. *et al.* (2013) Cell Host Microbe 13, 87-99). Most substrates except GAPDH have a DD and the glycosylation target is the DD which participates in DD-DD interaction of the NF- κ B signaling cascade. This is how pathogens using NleB avoid the host's immune response. SseK from *Salmonella typhimurium* is the orthologue of NleB and has the same function as NleB. Interestingly, the glycosylation target is an arginine residue. Exactly how this mechanism works is poorly understood because the guanidine group has poor nucleophilicity.

In our recent study, using the X-ray crystallographic method we revealed the UDP

bound SseK1, SseK2 apo, UDP-GlcNAc bound SseK2, UDP bound SseK2 and NleB2 apo structures. We also demonstrated that the SseK1, SseK2 and NleB have highly similar structures and each consists of three domains; the catalytic domain, helix-loop-helix (HLH) domain and lid domain, which will be discussed independently.

Catalytic domain. In the catalytic domain, a canonical DxD motif (Asp-any residue-Asp) is located at the active site pocket (SseK1^{D223-x-D225}, SseK2^{D239-x-D241} and NleB2^{D218-x-D220}) which is involved in the coordination of a manganese ion and in the interaction with UDP-GlcNAc via a salt-bridge. The manganese ion is a cofactor to SseK/NleBs and acts as a bridge between the pyrophosphate of UDP-GlcNAc and DxD motif. Based on the UDP-GlcNAc bound SseK2 structure, pyrophosphate of UDP, Asp241 from DxD motif, Asn338, Ser340 and a water molecule coordinate the manganese ion and all of the residues are conserved in the SseK and NleB families.

Interestingly, the binding mode of the uracil group of UDP-GlcNAc is slightly different, despite structure similarity between SseK1 and SseK2. In SseK1, two tryptophan residues (Trp51^{SseK1} and Trp331^{SseK1}) interact with the uracil group via π - π stacking interaction. In SseK2, Trp65 and Phe203 interact with the same orientation, even though Trp347 (corresponding to Trp331^{SseK1}) is conserved.

The HEN (His-Glu-Asn) motif is located in the active site pocket of the catalytic domain and this motif plays a crucial role in enzyme catalysis and virulence. Combining results of the *in vitro* enzyme kinetic PARP cleavage assay, NF- κ B expression assay and mouse CFU assay, catalytic efficacy and bacterial virulence were dramatically decreased when the HEN motif was substituted with an alanine residue. For this reason, we suggest that the HEN motif is the potential drug target for inhibition of NleB/SseK enzymes.

Based on the STD-NMR, MD simulation and *in vitro* enzyme kinetic assays results, we demonstrated that the WR motif (Trp-Arg, this arginine is different from the glycosylation target arginine) of the acceptor substrate is important for interaction with the enzyme. This WR motif is highly conserved in TRADD, FADD and GAPDH. When the WR motif of GAPDH is substituted with alanine, the binding affinity of the enzyme-substrate and the enzyme efficacy is decreased. Binding affinity and enzyme efficacy of the double mutant are synergistically decreased, compared to the single substitution. The WR motif is close to the HEN motif of the enzyme in the docking model of the acceptor arginine to the UDP-GlcNAc bound SseK2 structure and this agrees with the STD-NMR data.

Lid domain. Comparing the apo structure and UDP-GlcNAc bound structure of SseK2, we demonstrated that the lid domain regulates the open-closed state of the active site pocket. An unambiguous electron density map of lid domain was found in the UDP-GlcNAc bound SseK2 structure while in the apo structure this was not the case. The open state of SseK2 has a wider active site for binding the UDP-GlcNAc substrate and after binding of UDP-GlcNAc, the lid domain is covered up and fixed to the active site to shield the water and/or other donor-substrate in order to increase enzyme efficiency. Substrate-dependent ordering of the flexible loop (lid domain) is common between the other glycosyltransferase and SseK2, who share similar conformational changes, as well. Furthermore, such a conformational change is known to help assist in acceptor substrate binding (induced fit model). The structure of SseK2 allows for the plausibility of an induced fit model. However, supportive data is needed to demonstrate it more clearly.

HLH domain. Interestingly, NleB/SseKs have different substrate preferences despite the fact that the NleB/SseK families having similar sequences and structures (El Qaidi, S. *et al.* (2017) J Biol Chem 292. 11423-11430). For example, SseK2 glycosylates FADD, but SseK1 does not. Recently, our study suggested that the donor substrate specificity is chosen by the HLH domain of the enzyme. The protruded HLH domain is not a

common domain of other glycosyltransferases. According to the long Gaussian accelerated molecular dynamics simulation and the crystal structure analysis results, the HLH domain is more flexible than other domains. Furthermore, the structural dynamics are different between SseK1 and SseK2. A comparison between molecular docking models of FADD DD-SseK1 complex and FADD DD-SseK2 complex shows that the molecular dynamics of the HLH domain is highly different between SseK1 and SseK2. Put together, we suggested that the HLH domain regulates substrate preference.

Reaction mechanism. According to the stereochemistry of the anomeric bonding of donor sugar, glycosyltransferase can be classified into two groups, namely a retaining or an inverting enzyme. By NMR analysis, we demonstrated that the SseK1 is a retaining-glycosyltransferase. Because we failed to reveal the crystal structure of the acceptor substrate bound SseK or NleB, the sugar configuration data is important for predicting the position of acceptor arginine. Combined with the structural aspect of HEN motif, MD simulation, NMR and *in vitro* enzyme kinetic assays result, we suggest that SseK has a front face S_Ni mechanism (Diagram 1).

In conclusion, to better understand how arginine can be glycosylated, we revealed that the crystal structures of NleB/SseKs. SseK1 and SseK2 are similar, but have different UDP-GlcNAc binding modes. Using state-of-the-art methods, we demonstrated the importance of the HEN and WR motifs. The HLH domain is involved in glycosylation target selection and the lid domain regulates the open-close state of the enzyme's active site. SseK1 has a front face S_Ni mechanism for transferring the GlcNAc sugar. These data suggest potential sites for drug development against enteric pathogens.

Acknowledgements

This study was supported by the Grants from the National Research Foundation of Korea (NRF) funded by the Korean government (MEST) (NRF- 2016R1A2B2013305, 2016R1A5A1010764 and 2017M3A9F6029755, and the Strategic Initiative for

Microbiomes in Agriculture and Food funded by Ministry of Agriculture, Food and Rural Affairs (918012-4).

Figure legend

Diagram 1. Schematic diagram of the reaction mechanism of SseK and NleB families. Arginine glycosylation is catalysed by SseK from *S. typhimurium* or NleB from pathogenic *E.coli* using UDP-GlcNAc. The HEN (His-Glu-Asn) motif is catalytically important and this motif is highly conserved in SseK and NleB families. Combined with computer simulation and crystal structures, β -phosphate (β -P) of UDP, histidine and glutamate stabilize and/or activate the acceptor arginine. The proposed mechanism provides an insight of how arginine is able to be glycosylated by SseK and NleB families.

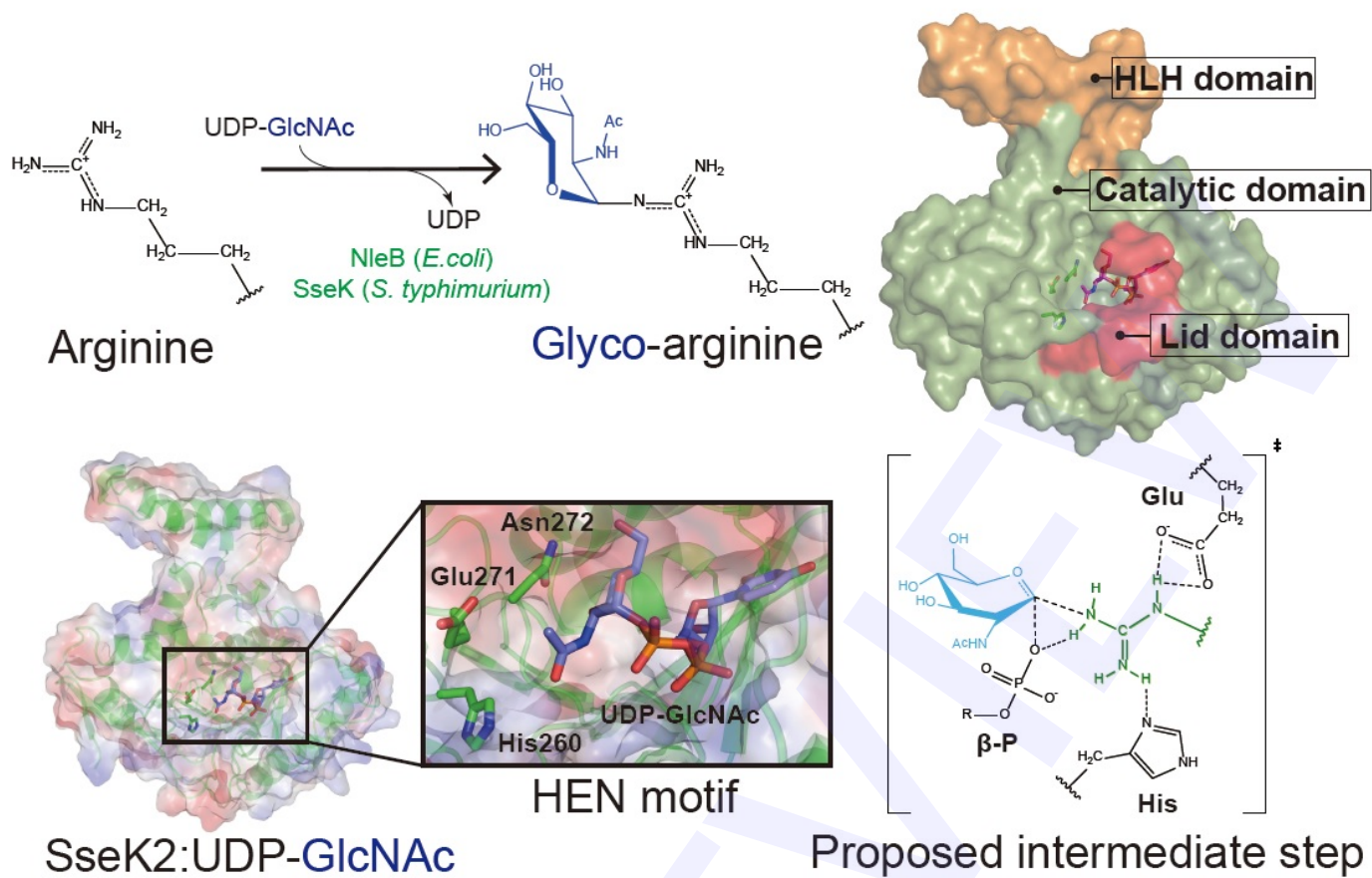


Fig. 1.