



Core fucose-specific *Pholiota squarrosa* lectin (PhoSL) as a potent broad-spectrum inhibitor of SARS-CoV-2 infection

Kazuhiko Yamasaki¹ , Naruhiko Adachi², Mya Myat Ngwe Tun³, Akihito Ikeda², Toshio Moriya², Masato Kawasaki², Tomoko Yamasaki¹, Tomomi Kubota¹, Izuru Nagashima⁴, Hiroki Shimizu⁴, Hiroaki Tateno⁴  and Kouichi Morita³

¹ Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

² Structural Biology Research Center, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), Tsukuba, Japan

³ Department of Virology, Institute of Tropical Medicine (NEKKEN), Nagasaki University, Nagasaki, Japan

⁴ Cellular and Molecular Biotechnology Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

Keywords

COVID-19; glycan-binding protein; molecular dynamics simulation; surface plasmon resonance; virus infection

Correspondence

K. Yamasaki, Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan
Tel: +(81)298619473
E-mail: k-yamasaki@aist.go.jp

(Received 23 February 2022, revised 6 August 2022, accepted 22 August 2022)

doi:10.1111/febs.16599

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike protein (S protein) is highly N-glycosylated, and a “glycan shield” is formed to limit the access of other molecules; however, a small open area coincides with the interface to the host’s receptor and also neutralising antibodies. Most of the variants of concern have mutations in this area, which could reduce the efficacy of existing antibodies. In contrast, N-glycosylation sites are relatively invariant, and some are essential for infection. Here, we observed that the S proteins of the ancestral (Wuhan) and Omicron strains bind with *Pholiota squarrosa* lectin (PhoSL), a 40-amino-acid chemically synthesised peptide specific to core-fucosylated N-glycans. The affinities were at a low nanomolar level, which were ~ 1000-fold stronger than those between PhoSL and the core-fucosylated N-glycans at the micromolar level. We demonstrated that PhoSL inhibited infection by both strains at similar submicromolar levels, suggesting its broad-spectrum effect on SARS-CoV-2 variants. Cryogenic electron microscopy revealed that PhoSL caused an aggregation of the S protein, which was likely due to the multivalence of both the trimeric PhoSL and S protein. This characteristic is likely relevant to the inhibitory mechanism. Structural modelling of the PhoSL–S protein complex indicated that PhoSL was in contact with the amino acids of the S protein, which explains the enhanced affinity with S protein and also indicates the significant potential for developing specific binders by the engineering of PhoSL.

Abbreviations

ACE2, angiotensin converting enzyme 2; CD209, cluster of differentiation 209; CD209C, CD209 antigen-like protein C; CLEC4G, C-type lectin domain family 4 member G; CLEC4M, C-type lectin domain family 4 member M; cryo-EM, cryogenic electron microscopy; DC-SIGN, dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin; DC-SIGNR, DC-SIGN-related protein; DLS, dynamic light scattering; FCS, foetal calf serum; FRIL, FcγR3 receptor-interacting lectin; GlcNAc, N-acetylglucosamine; GRFT, Griffithsin; K_D , dissociation constant; L-SIGN, liver/lymph node-specific intercellular adhesion molecule-3-grabbing non-integrin; MD, molecular dynamics; MEM, minimum essential medium; N-glycan, N-linked glycan; PDB, Protein Data Bank; PhoSL, *Pholiota squarrosa* lectin; PRNT, plaque reduction neutralisation test; qRT-PCR, quantitative reverse-transcription/real-time PCR; RBD, receptor-binding domain; RBM, receptor-binding motif; RP-HPLC, reversed-phase high performance liquid chromatography; S protein, spike protein; SARS-CoV, severe acute respiratory syndrome coronavirus; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SIGLEC1, sialic acid-binding immunoglobulin-like lectin 1; SPR, surface plasmon resonance.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infects human cells by binding the spike protein (S protein) on the virus surface with angiotensin-converting enzyme 2 (ACE2) on the host cell membrane [1–3]. The trimeric structure of the S protein in the prefusion state was determined via cryogenic electron microscopy (cryo-EM) [4]. The receptor-binding domain (RBD), which is responsible for contact with ACE2, adopts different conformations among the three monomers, with an up conformation for one monomer and a down conformation for the other two. However, on the virus surface, the number of up conformers varies from 0 to 2 among trimeric molecules [5].

The S protein is highly glycosylated and includes 22 N-linked glycans (N-glycans) per monomer attached to asparagine residues [6]. These N-glycans cover the protein surface, thus forming a “glycan shield” that limits the access to other molecules and to possibly interferes with the host’s immune system [6,7]. As an exception, a small open area was observed on one side of RBD. This area coincides with the ACE2 interface, which is only accessible to the RBD in the up conformation [4,8]. According to this limited accessibility, most of the neutralising antibodies bind to this open area [7,9]. Moreover, the mutations affecting the transmissibility and/or antigenicity e.g., Asn501Tyr, occur in this area, which may reduce the neutralisation effects of postvaccination serum [10]. Indeed, the Omicron variant, which contains multiple mutations in the above area, has caused a pandemic due to breakthrough infections despite booster vaccination [11].

In contrast, N-glycans have been found to facilitate the entry of viruses into the host cells. Some lectin receptors that recognise the N-glycans of the S protein behave as secondary receptors of SARS-CoV-2 infection, thereby assisting the ACE2-driven entry pathway [12,13]. These receptors include dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN; also known as CD209: cluster of differentiation 209), liver/lymph node-specific intercellular adhesion molecule-3-grabbing non-integrin (L-SIGN; also known as CLEC4M: C-type lectin domain family 4 member M or DC-SIGNR: DC-SIGN-related protein), and sialic acid-binding immunoglobulin-like lectin 1 (SIGLEC1; also known as sialoadhesin or CD169). Some neutralising antibodies, including the S309 antibody that recognises an N-glycan in the RBD [14], effectively blocked this lectin-directed pathway, instead of the ACE2-directed pathway [13].

Considering the contrasting accessibility of the N-glycans of the S protein relative to the surface of the proteins and the role of N-glycans in the secondary pathway, they may represent important drug development targets. Glycosylated residues are relatively invariant, and 18 of 22 glycans of SARS-CoV-2 are also conserved in SARS-CoV [15]. In addition, none of the 34 Omicron mutations affects the 22 N-glycosylation motifs [16]. Moreover, the two N-glycans within the RBD, Asn331 and Asn343, are considered necessary for infection because their deletion drastically reduced infectivity [17], which is presumably related to their role in the secondary pathway. Therefore, targeting N-glycans using glycan-binding molecules, including lectins, could represent a therapeutic strategy for most variants of the virus.

C-type lectins related to the above, namely, CD209 antigen-like protein C (CD209C) and C-type lectin domain family 4 member G (CLEC4G), were found to show high affinities with S protein and inhibitory effects on SARS-CoV-2 infection [18]. These activities have also been demonstrated for lectins of non-mammalian origins, namely, Flt3 receptor-interacting lectin (FRIL) isolated from hyacinth bean [19], and Griffithsin (GRFT) isolated from red alga [20]. To use these lectins as drugs, the off-target effects, including blood coagulation, should be reduced. This is especially important for the non-mammalian lectins, although it should also be considered for mammalian lectins at doses higher than the normal expression level. One solution would be to confer protein specificity to the lectin via engineering, which would allow them to be used at a low dose to inhibit SARS-CoV-2 infection. To achieve this, lectins should be able to form contact with the amino acids of the S protein. However, the lectins listed above would have a disadvantage because they recognise sugars at positions distal from the protein (Fig. 1A). In contrast, lectins that recognise positions close to the root asparagine are more suitable for such engineering.

A lectin from the mushroom *Pholiota squarrosa* (PhoSL) recognises the core fucosylation that occurs at the innermost *N*-acetylglucosamine (GlcNAc) (Fig. 1A) [21]. Its trimeric structure was determined by both our group and other researchers, and in this structure, a trisaccharide in the formula fucose(α 1–6)[GlcNAc(β 1–4)]GlcNAc is recognised in the binding pocket (Fig. 1B) [22–24]. In the S protein, more than half of the N-glycans are core-fucosylated [6,18]. Therefore, PhoSL is expected to form contacts with S protein amino acids and thereby acquire specificity to the protein by appropriate engineering. Moreover, the PhoSL monomer is a short peptide of only 40 amino acids and can therefore

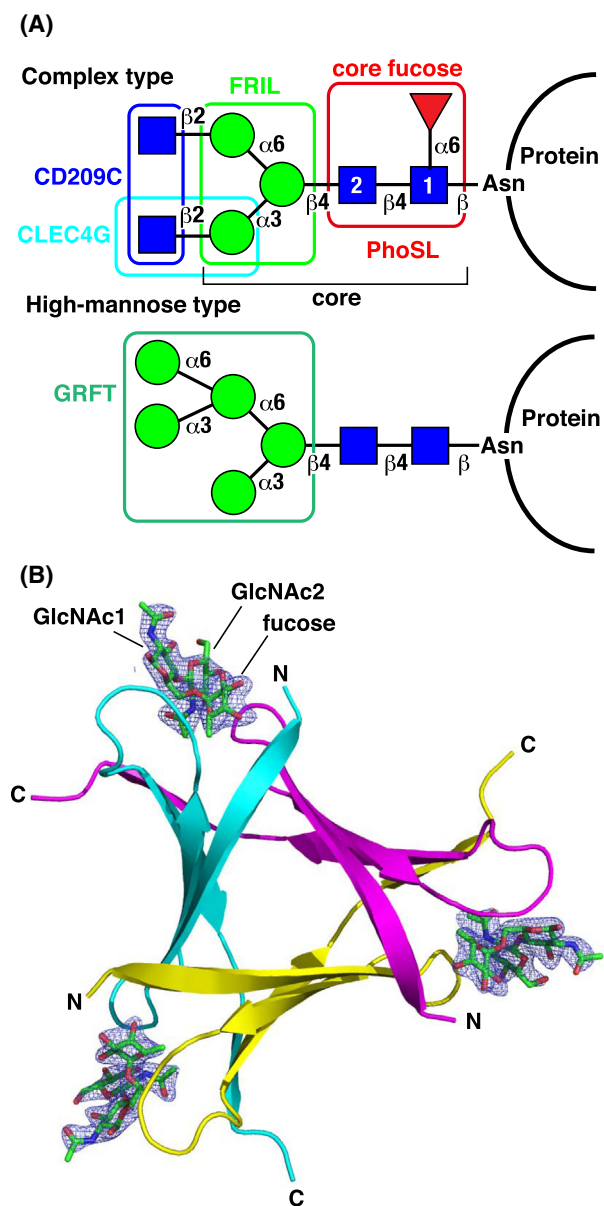


Fig. 1. Core fucosylation-specific lectin PhoSL. (A) Two typical forms of N-glycans that appear on the SARS-CoV-2 S protein [18]. Fucose, GlcNAc, and mannose are indicated by the red triangle, blue square, and green circle, respectively, and the types of connection are indicated. The moieties recognised by the respective lectins are enclosed. Two GlcNAcs recognised by PhoSL are numbered. (B) Crystal structure of PhoSL trimer in complex with trisaccharide fucose(α 1–6)[GlcNAc(β 1–4)]GlcNAc-OH, corresponding to the region indicated by the red enclosure in (A). The mFo–DFc omit maps contoured at 3.0σ are indicated for the trisaccharide, where the two GlcNAcs are numbered as in (A). The molecular graphic figure was created by PYMOL ver. 1.8 (Schrödinger, LLC, New York, NY, USA).

be chemically synthesised. Thus, the engineering of this peptide would be more flexible through the incorporation of unnatural amino acids.

Here, we investigated the therapeutic potential of PhoSL and demonstrated the strong affinities of PhoSL with the S proteins of both the ancestral (Wuhan) strain and Omicron variant. We found that PhoSL activity was involved in inhibiting infection by both SARS-CoV-2 strains. Cryo-EM analysis revealed that PhoSL aggregated with the S protein, and structural modelling of the association of PhoSL with the N-glycans at Asn331 or Asn343 in the RBD showed that PhoSL extensively interacted with the amino acids of the S protein, thus indicating its significant potential for engineering to confer specificity to the S protein.

Results

Binding of PhoSL to S protein

Among the 22 N-glycans of the SARS-CoV-2 S protein per monomer, 52% on average were revealed to be core-fucosylated [6] and thus should be specifically recognised by PhoSL (Fig. 1). Here, using surface plasmon resonance (SPR), we observed that PhoSL binds with the S protein (of the ancestral Wuhan strain, unless otherwise stated) immobilised on the sensor chip (Fig. 2). A simultaneous kinetic analysis of the sensorgrams revealed that the association and dissociation rate constants (k_a and k_d) were $4.9 \times 10^4 \text{ M}^{-1} \cdot \text{s}^{-1}$ and $4.5 \times 10^{-5} \text{ s}^{-1}$, respectively, and the equilibrium dissociation constant (K_D) was 0.91 nM (Fig. 2A; monomer concentration). Nevertheless, the errors in curve fitting are more serious than they appear because the association rates tend to be inaccurate and may easily be twice the true values (see that at $0.02 \mu\text{M}$). Because the affinity value in this analysis is based on the ratio of kinetic constants, we chose the equilibrium analysis (Fig. 2B), which is another standard method for evaluating affinity, and obtained K_D of $3.9 \pm 0.6 \text{ nM}$ (Fig. 2C). Note that we repeated injections to reach plateau, which was enabled by the very slow dissociation rate. In addition, for the Omicron-variant S protein, a similar affinity (K_D of $3.4 \pm 0.3 \text{ nM}$) was observed by the same equilibrium analysis. Note that the maximum number of PhoSL trimers bound to an S protein trimer was estimated to be ~ 8 , as calculated based on the immobilised amount and molecular weight ratios; moreover, the above analyses are based on a 1:1 binding scheme, which is enabled by assuming that the PhoSL trimers bind to the S protein independently but with similar affinities.

In addition, the core-fucose trisaccharide fucose(α 1–6)[GlcNAc(β 1–4)]GlcNAc-OH (Fig. 1A), competitively inhibits binding, with an IC_{50} value of $10.4 \pm 0.2 \mu\text{M}$

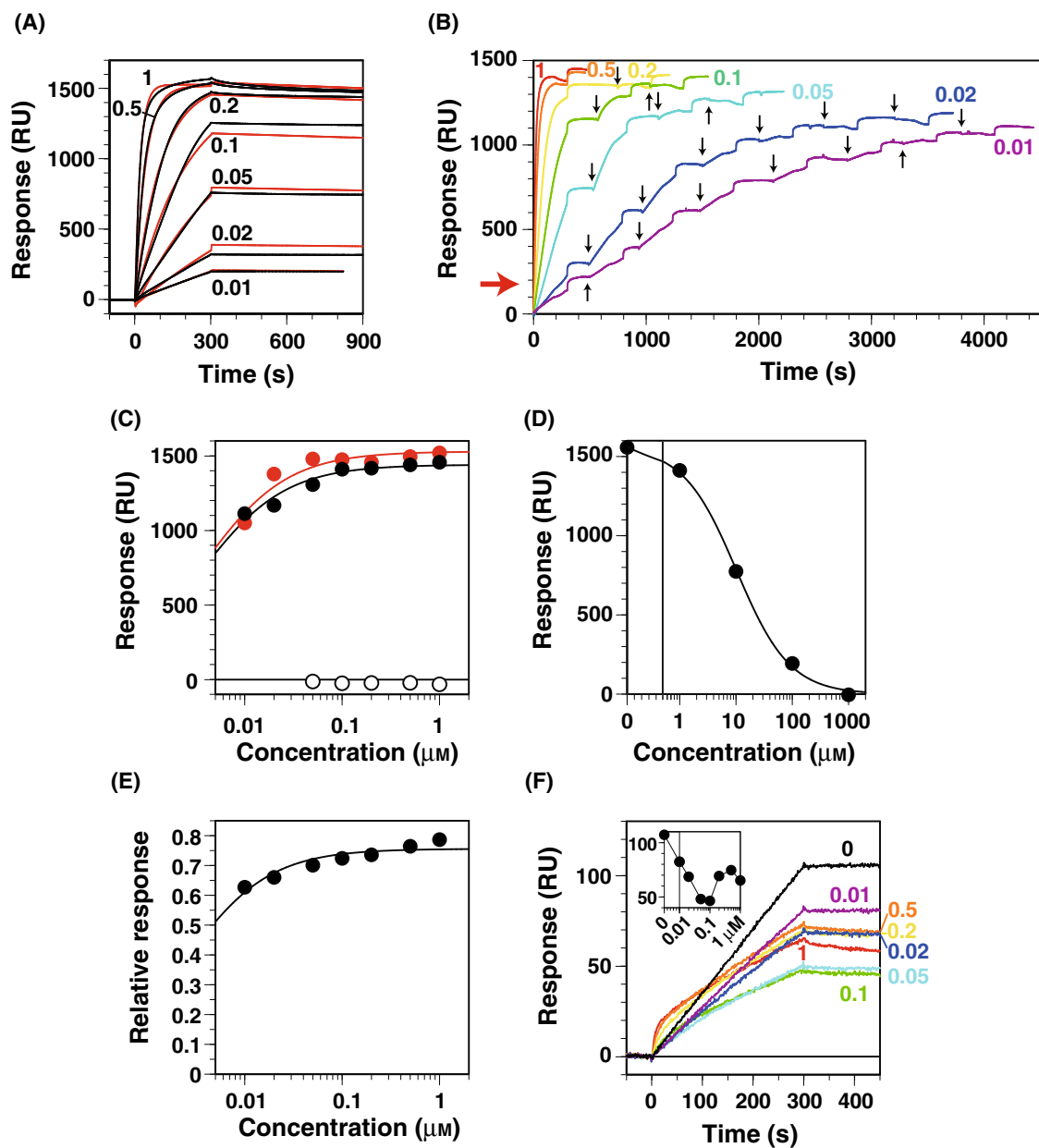


Fig. 2. Binding of PhoSL to the S protein. (A) SPR sensorgrams (the responses in the control flow-cell were subtracted) for the binding of PhoSL to S protein immobilised on the sensor chip. Respective concentrations of PhoSL as monomer in solutions are shown in μM . Simultaneous fitting curves were shown in red. (B) SPR sensorgrams for binding of the same molecules in (A) for the equilibrium analysis. Small arrows accompanying the sensorgrams indicate the start points (other than time 0) of sample injections for 300 s. A large red arrow on the vertical axis indicates the response value (177 RU) expected when one PhoSL trimer binds to one S protein trimer, as calculated from the immobilised amount and molecular weights. (C) Equilibrium response values at different protein concentrations. Closed black and red circles represent the binding profiles of PhoSL to the S proteins of Wuhan (same as in A, B) and Omicron strains, respectively, while open circles represent the binding of the sugar-binding-pocket mutant Trp28Thr-Val36Tle to the Wuhan S protein. A fitting curve to the simple 1 : 1 binding model is shown. (D) Inhibition of PhoSL–S protein binding by trisaccharide fucose(α 1-6)[GlcNAc(β 1-4)]GlcNAc-OH at a fixed concentration of PhoSL (0.1 μM). Note that the horizontal axis is linear at trisaccharide contents below 0.5 μM , while it is logarithmic at contents above 0.5 μM . A fitting curve to the Hill equation is shown. (E) Binding of PhoSL to the RBD of the S protein, which is shown as the equilibrium response values. The response values are divided by the expected values when one PhoSL trimer binds to one RBD molecule immobilised on the sensor chip. (F) Inhibition of S protein–ACE2 binding by PhoSL on the ACE2-immobilised sensor chip. Concentration of S protein was fixed (3 nM) while that of PhoSL was varied as shown in μM . The inset shows dependency of the response at 300 s on the concentration of PhoSL, where the horizontal axis is linear or logarithmic (below or above 0.01 μM , respectively).

(Fig. 2D). In addition, Trp28Thr-Val36Tle (L-tert-leucine), which was designed to change the shape of the sugar-binding pocket according to our previous observations [24], did not show an affinity for the S protein (Fig. 2C). These results confirmed that PhoSL binds to the S protein via the recognition of core fucosylation.

The affinity of PhoSL for the S protein was ~ 1000 -fold higher than that for the various core-fucosylated glycans, as measured by frontal affinity chromatography (K_D of 2–10 μM ; concentration as monomer) [21]. We propose two schemes to account for this difference. The first scheme is the cooperative involvement of more than one N-glycan in binding with a single PhoSL trimer. The S protein possesses ~ 30 core-fucosylated N-glycans per trimer; however, only ~ 8 of the PhoSL trimers bind to the S protein on the sensor chip. Although the amine coupling method for the immobilisation masks a part of the molecular surface, including the covalently bonded Lys residues, two or more of core-fucosylated N-glycans may still be bound by the PhoSL trimer if the stereochemical restrictions allow the binding geometry. The second scheme is that preferable interactions are formed between PhoSL and the amino acids of the S protein, such as hydrogen bonds or hydrophobic interactions. As PhoSL recognises the core-fucose trisaccharide that includes the innermost GlcNAc (Fig. 1), some parts of PhoSL may closely approach the S protein amino acids.

We also observed the binding of PhoSL to the S protein RBD, which possesses two core-fucosylated N-glycans (Fig. 2E). The obtained K_D was 2.4 nM, which is slightly better than the whole S protein. The maximum number of PhoSL trimers bound to the RBD was ~ 0.8 , indicating that only one of the two N-glycans was available for PhoSL concurrently. In addition, a recent report indicated that some mutations of S protein decreased core fucosylation of the N-glycan at Asn343 to $\sim 85\%$ [25], which may have contributed to reducing this value; note that the S protein used here possesses several stabilising mutations (see [Materials and methods](#) section). Because the S protein trimer contains three RBDs, the contribution of the RBD to PhoSL binding should be approximately 1/3 that of the whole protein. Also, the high affinity with a single N-glycan strongly supports the second scheme for the affinity enhancement, i.e., preferable interactions with S protein amino acids.

We examined the inhibitory activity of PhoSL against the S protein–ACE2 interaction by SPR on the ACE2-immobilised sensor chip (Fig. 2F). We observed a clear drop in response dependently on the PhoSL concentration up to 50 nM, whereas it increased at

200 nM or higher, which was accompanied by an additional fast-binding phase and an accelerated dissociation. This behaviour indicates the switch of the molecule(s) that bind to ACE2 and is likely due to mixed effects of PhoSL: (a) direct or indirect interference with the S protein–ACE2 interaction via the S protein–PhoSL interaction, (b) binding of PhoSL with core-fucosylated N-glycans of ACE2, and (c) aggregation between PhoSL and S protein (see below). The situation is rather complicated, i.e., we do not tell whether PhoSL–ACE2 interaction simply increases the response or decreases the response by competing with the larger S protein. Moreover, aggregation may contribute to increasing the response initially by enlarging the effective molecular mass when S protein binds to ACE2, although it should interfere with the binding through reduction of the effective concentration of S protein when aggregation proceeds.

Neutralisation of SARS-CoV-2 infection by PhoSL

The above observations of the affinity between PhoSL and S protein at low nanomolar levels led us to expect its inhibitory activity against virus infection. In the plaque reduction neutralisation test (PRNT) using the primate Vero-E6 cells, we observed the neutralisation activity of PhoSL against both the Wuhan and Omicron strains of SARS-CoV-2 (Fig. 3). We did not observe plaques in either of the strains at 10 μM of PhoSL (100% inhibition). The analysis of concentration dependency yielded IC_{50} values of $0.37 \pm 0.17 \mu\text{M}$ and $0.36 \pm 0.05 \mu\text{M}$ (as a monomer) for the Wuhan and Omicron strains, respectively. These results strongly suggest that the inhibitory activity of PhoSL has broad-spectrum properties and is less sensitive to variations than most antibodies [10,11]. The slopes (Hill coefficients) of the curves were different for the two strains (3.0 ± 1.8 for Wuhan and 1.2 ± 0.1 for Omicron), and although the reason for this is unclear, it may be related to the mode of aggregation (see below).

The IC_{50} values of the weight concentration ($1.6 \mu\text{g}\cdot\text{mL}^{-1}$ for both the strains), which is typically used for comparisons among antibodies, were comparable to the value of $0.71 \mu\text{g}\cdot\text{mL}^{-1}$ obtained for FRIL by a similar PRNT assay [19]. The values were also comparable to the $0.42 \mu\text{g}\cdot\text{mL}^{-1}$ obtained for GRFT based on a quantitative reverse-transcription/real-time PCR (qRT-PCR) assay [20], and much stronger than the 50–100 $\mu\text{g}\cdot\text{mL}^{-1}$ for CLEC4G or CD209C obtained by qRT-PCR [18]. Therefore, PhoSL is among the most potent lectins for inhibition of SARS-CoV-2 infection.

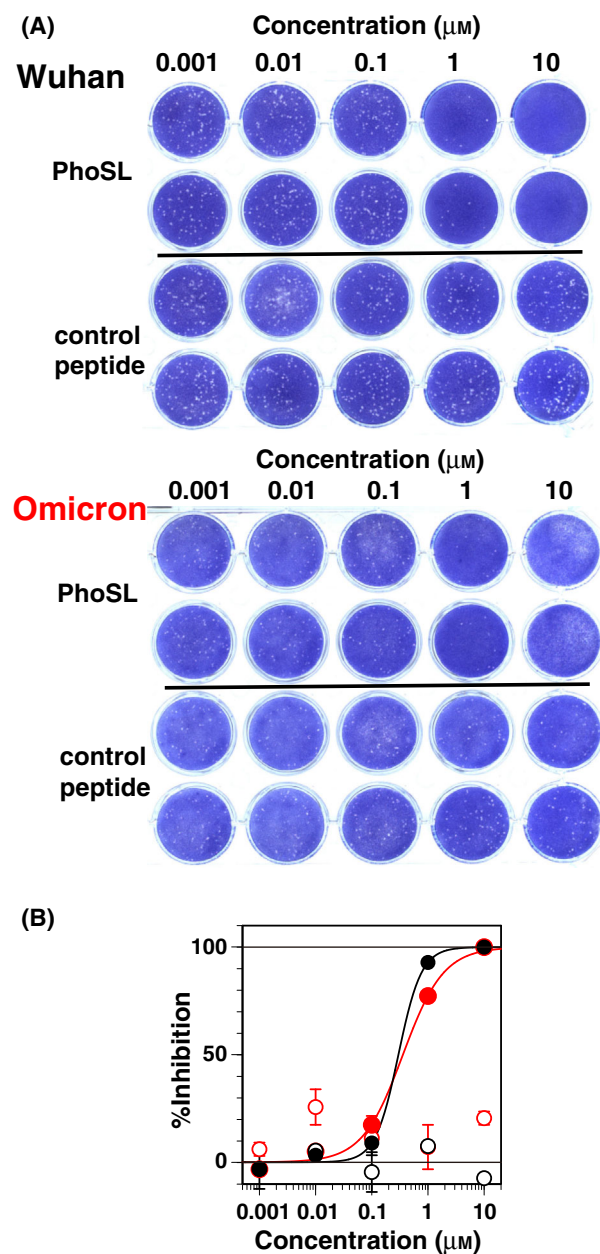


Fig. 3. Inhibition of SARS-CoV-2 infection by PhoSL. (A) The images of wells in the PRNT. Shown are duplicated wells for assays of infection by Wuhan and Omicron strains at different concentrations of PhoSL and a control 46mer peptide. (B) The inhibition rate as function of the concentrations of PhoSL and the control peptide (closed and open circles, respectively) for the Wuhan and Omicron strains (black and red, respectively). Error bars are from the deviation of plaque numbers between the duplicated wells. Indicated fitting curves to Hill equations are for the averaged numbers.

Aggregation of S protein trimer by PhoSL

To reveal the binding mode of PhoSL to the S protein, we applied cryo-EM experiments (Fig. 4) because the

molecular structure of the glycosylated S protein was determined by this method [4,5]. The S protein solution was quickly frozen on the grid, and an image was obtained that indicated the solution contained well-distributed particles (Fig. 4A, left). Two-dimensional classification analyses clearly showed that these particles possessed a pseudo threefold rotational symmetry (C_3), which is the characteristic of the S protein trimer (Fig. 4B).

In contrast, the image of the PhoSL-S protein mixture solution mostly lacked particles and instead showed aggregation (Fig. 4A, right). Although determining the structure of the complex became difficult, the observation provided important information about the effect of PhoSL on the S protein. Also in dynamic light scattering (DLS), we observed the aggregation of the S protein by PhoSL (Fig. 4C). Namely, it broadened the peak and decreased the average diffusion coefficient that is inversely proportional to the particle size. Note that the peak centred at $\sim 5 \times 10^{-8} \text{ cm}^2 \cdot \text{s}^{-1}$ in the right panel is likely due to the excess PhoSL, in equilibrium between association and dissociation.

The aggregation should be due to the multivalence of both trimeric PhoSL and the S protein, as shown in Fig. 4D. Considering that ~ 8 PhoSL trimers were bound to a single S protein trimer in the SPR experiment (Fig. 2B), the actual aggregation in solution would be complicated and irregular based on a three-dimensional arrangement. We emphasise that this aggregation should limit the access of other proteins, such as ACE2 or the lectin receptors, to the S protein.

Structural modelling for interaction of PhoSL and S protein

Although the structure of the PhoSL-S protein complex was difficult to determine due to aggregation, we carried out computational modelling to obtain information on the binding mode (Fig. 5). The two glycosylation sites in the RBD (Asn331 and Asn343) were selected as the modelling sites because we observed the strong binding of PhoSL to at least one of them by SPR (Fig. 2E). At these sites, 98% of the N-glycans are core-fucosylated as shown in a report [6] ($\sim 100\%$ and 85% at Asn331 and Asn343, respectively, in another report [25]) and importantly, these N-glycans, especially that at Asn343, have been shown to be critical in infection [17].

The crystal structure of PhoSL in complex with the core-fucose trisaccharide determined in the present study (Fig. 1B) and the RBD (up conformation) in the cryo-EM structure of the S protein [4] were used to

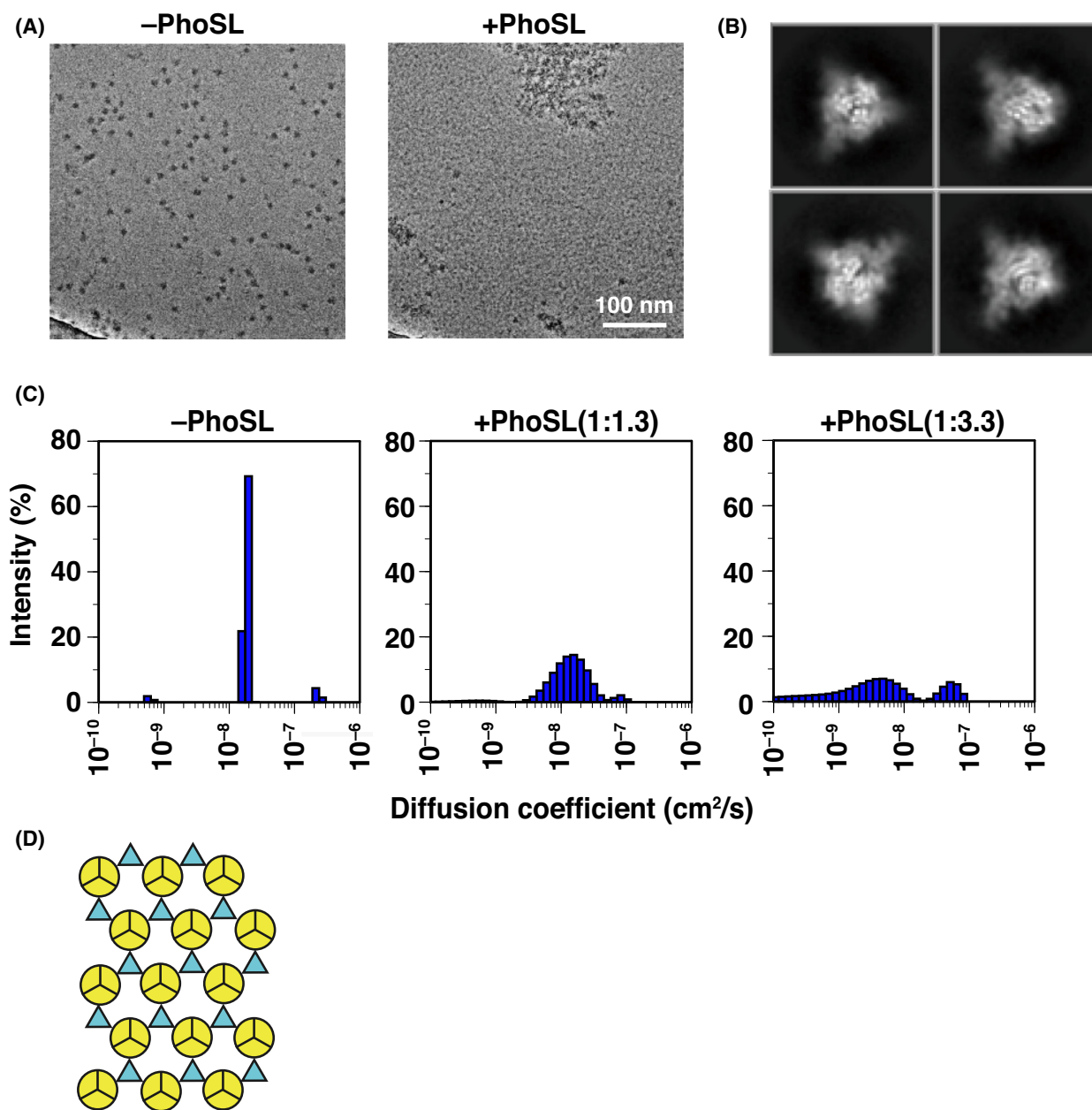


Fig. 4. S protein aggregation induced by PhoSL. (A) Cryo-EM images of the S protein (left) and S protein–PhoSL mixture (right). The scale bar is shown inside the right image. (B) Four representative images of the two-dimensional classification of the S protein particles. (C) DLS analysis of S protein in the absence (left) or presence of PhoSL at different mixing ratios (middle, right). (D) Schematic model of the aggregation of the S protein (circle) and PhoSL (triangle) formed in a thin ice layer on the grid chip.

generate the initial structure, where the innermost GlcNAc that was shared between the two structures was used for the fitting (Details are in the [Materials and methods](#) section). During 10 ns of the molecular dynamics (MD) simulation, PhoSL bound to the S protein stably, without a large change in relative orientation (Fig. 5A,B). In contrast, the RBD fluctuated

internally at the hinge between the two RBD subdomains. The motion at this hinge is relevant to the change between the up and down conformations. The positions of Asn331 and Asn343 are at a side of the RBD, which is different from the molecular interface with ACE2 (receptor binding motif, RBM; Fig. 5C,D), suggesting that PhoSL bound at these sites is unlikely

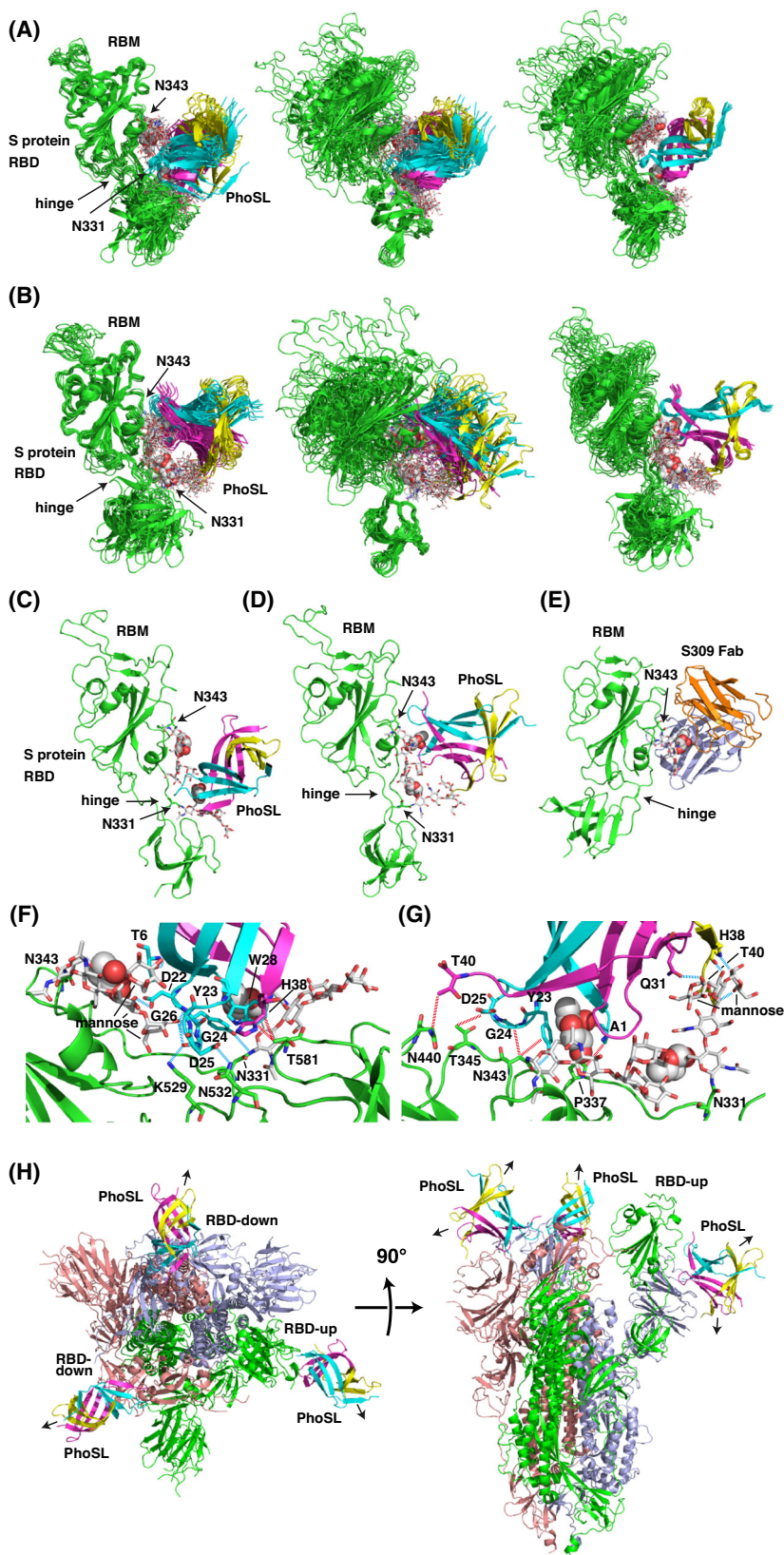


Fig. 5. Structural models for the binding of PhoSL to S protein. (A, B) Trajectory of the MD simulation of the complex of S protein RBD and PhoSL recognising core fucose of the N-glycan at Asn331 (A) or Asn343 (B). Structures extracted at every 1 ns from the trajectory of 10 ns are shown. From left to right, the upper and lower subdomains separated at the indicated hinge region, and PhoSL were used for the alignment of the extracted structures. (C, D) Representative structures for the MD simulation of the complex where PhoSL is bound primarily to the core-fucosylated N-glycan at Asn331 (C) or Asn343 (D). (E) Part (RBD) of the cryo-EM structure of the complex of the Fab fragment of the S309 antibody and the S protein (PDB: 6WPT) [14]. In (A–E), fucose and other sugars are shown as spheres and sticks, respectively, and the hinge region between two subdomains of the RBD is indicated. (F, G) Details of the molecular interface of the PhoSL–S protein complex formed at Asn331 (F) or Asn343 (G). Blue and red broken lines indicate hydrogen bonds (donor–acceptor distance < 3.5 Å) and hydrophobic contacts (C–C distance < 4.5 Å), respectively, except for those between PhoSL and the N-glycan at the primary recognition site. (H) Locations of PhoSL bound at the Asn343 glycosylation site in the overall structure of the S protein trimer with one up and two down conformations for the RBD [4]. The upper subdomain of the RBD that includes RBM was used for the fitting. The arrows indicate open sugar-binding pockets of the PhoSL trimers. The molecular graphics were created by PYMOL.

to compete with ACE2. This situation is similar to that for the neutralising antibody S309 [14], which recognises the core fucose of N-glycan at Asn343 (Fig. 5E).

Because the two coordinates of PhoSL that recognise the core fucose of N-glycan at Asn331 or Asn343 overlap with each other (Fig. 5C,D), two PhoSL molecules are unlikely to bind at the both sites concurrently, which well explains the maximum number of binding molecules of < 1.0 per RBD observed with SPR (Fig. 2E). Alternatively, one PhoSL molecule may interact with the both N-glycans of the RBD, where it binds at the core fucose of N-glycan at Asn331 and also forms hydrogen bonds with the mannose moiety of the N-glycan at Asn343, and *vice versa* (Fig. 5F,G). In addition, PhoSL interacts with the amino acids of the S protein through hydrogen bonds and hydrophobic contacts. Despite the dynamical fluctuations, the interactions with glycan and amino acids were observed throughout the simulation, which explains the enhanced affinity with the S protein observed by SPR compared to the affinities with isolated glycans [21]. In contrast, the simultaneous binding of two core fucoses of a single S protein trimer molecule by two sugar-binding pockets of a PhoSL trimer (the first scheme in the SPR section) appears to be geometrically unlikely in the current model. Among the interactions, those by Asp22, Tyr23, Gly24, and His38 of PhoSL bound to the core fucose at Asn331 and those by Tyr23 and His38 of PhoSL bound to that at Asn343 are relatively stable and observed in 90% or more of the structures extracted from the trajectory.

Sixteen of the 34 mutation sites of the Omicron-variant S protein were located within the RBD (Fig. 6). Most of them are at the interface with ACE2 (RBM), which may be related to the enhanced transmissibility. However, Gly339Asp and Asn440Lys were located at the interface with PhoSL bound to Asn343 in the model (Fig. 6B). Although these mutations do not appear to interfere with the binding, the close proximity with Tyr23 and Thr40 of PhoSL provides clues for designing specific binders to the S protein of the variant.

To evaluate the mode of aggregation, we placed PhoSL molecules on the entire structure of the S protein trimer (Fig. 5H). When PhoSL binds to the N-glycans at Asn343 of the three chains, as shown in Fig. 5D, open sugar-binding pockets may point to the outside, thus allowing for the binding of other S protein molecules. Because more PhoSL molecules should

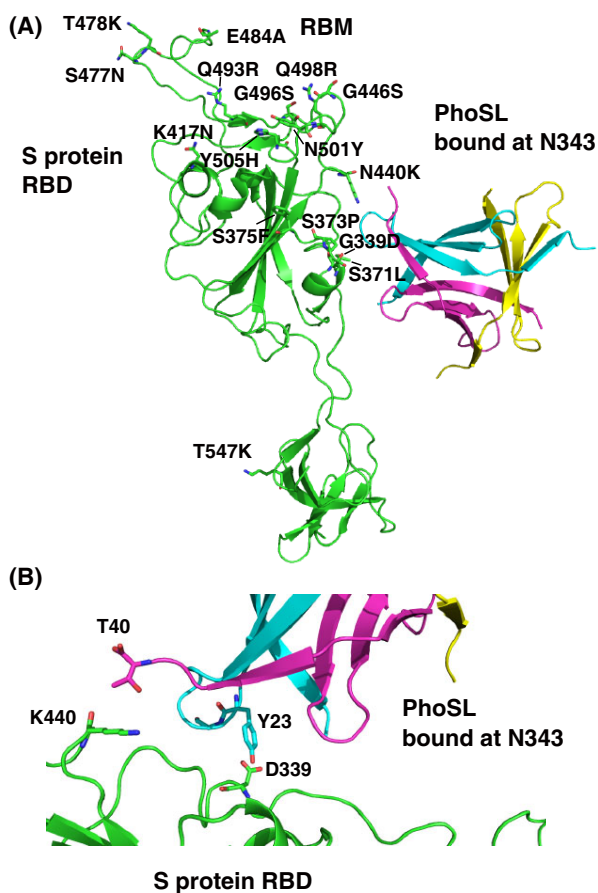


Fig. 6. The Omicron mutations indicated on the present model of the complex of S protein RBD and PhoSL. (A) The Omicron mutations occurring in RBD, as shown in the complex model with PhoSL bound at Asn343 (orientation is similar to Fig. 5D). (B) The two Omicron mutations occurring at the molecular interface with PhoSL bound at Asn343 (orientation is similar to Fig. 5G). The graphics were created by PYMOL.

bind to N-glycans at other sites, as shown by SPR, geometry for aggregation is likely to be more flexible.

Here we performed the MD simulation of the RBD in the up conformation in complex with PhoSL. Although the RBD in this conformation is relatively separated from the other parts of the trimeric S protein (Fig. 5H), we cannot exclude a possibility that the trimeric assembly may affect the dynamics and binding mode. This is more critical to the other two RBDs apparently fixed in the down conformation, which may show significantly limited dynamics at the hinge and consequently altered binding modes (especially for binding at Asn331 located near the hinge).

Discussion

Mechanisms for infection inhibition by PhoSL

The most straightforward explanation for the mechanism underlying the inhibition of SARS-CoV-2 infection is aggregation of the S protein. This process gathers the S protein molecules on the virus surface or causes aggregation of the virus itself by linking S proteins on the same or different virus particles. Consequently, the aggregation limits the accessibility of S protein molecules or virus particles to host cell receptors and reduces their effective concentrations. The orientation of the S protein on the virus surface is highly flexible, and the angles are varied by up to 60° [5], which enables the two-dimensional aggregation on the virus surface. In addition, aggregation of virus particles by lectin was observed for the tetrameric FRIL, which showed neutralising activities against influenza and SARS-CoV-2 (aggregation was shown for the former) [19].

An alternative mechanism is that PhoSL directly interferes with the binding of the S protein and ACE2. In the SPR experiment, PhoSL induced initial decrease in the response of S protein–ACE2 binding (Fig. 2F). However, due to the modelled geometry of PhoSL bound at Asn313 or Asn343 of the RBD, which showed different molecular interfaces with ACE2 (RBM) (Fig. 5C,D), we suppose that this mechanism is unlikely. Neutralising antibody S309 binds core fucose at Asn343, which is similar to that in the PhoSL model (Fig. 5E), and it does not compete with the binding of ACE2 to the S protein [14]. Nevertheless, aggregation of S protein by PhoSL should limit the access of ACE2 to the S protein, as described above. We may also raise a possibility that PhoSL interferes with the binding of the S protein and ACE2, indirectly via inducing a conformational change in the trimeric assembly.

PhoSL also appears to be able to interfere with the binding of the S protein to lectin receptors (DC-SIGN, L-SIGN, and SIGLEC1) [12,13]. S309 is involved in the inhibition of this secondary pathway instead of competing with ACE2 for binding to the S protein [13]. With similar binding sites considered for PhoSL, that is, the core fucose at Asn343 (Fig. 5D,E), we can hypothesise that PhoSL also has inhibitory activity against this pathway. These lectin receptors recognise the outer regions of N-glycans. Namely, DC-SIGN and L-SIGN preferentially bind to high mannose-type N-glycans (Fig. 1A, lower), and the former recognises fucosylation at the outer GlcNAc of the complex-type N-glycan (outermost in Fig. 1A, upper) [26]. Because

high mannose-type N-glycans are not core-fucosylated [18] and PhoSL does not recognise fucosylation other than core fucosylation with the α 1–6 linkage [21], these lectins and PhoSL likely do not compete with each other. SIGLEC1 recognises sialic acids even outside of the outermost GlcNAc (disaccharide sialic acid–galactose is attached to the GlcNAc) [27].

Nevertheless, DC-SIGN and L-SIGN also bind GlcNAc₂Man₃ (the five outmost sugars of the complex type N-glycan in Fig. 1A), for which the crystal structures of the complexes are available [Protein Data Bank (PDB) entries 1K9I and 1K9J for DC-SIGN and L-SIGN, respectively] [28]. When we examined these structures, the PhoSL–S protein complex model, and the S309–S protein complex structure (Fig. 5C–E), these lectin receptors and PhoSL or S309 obviously compete with each other for the same N-glycan, with steric clashes occurring when bound concurrently (data not shown). Therefore, similar to the S309 antibody, it is possible that PhoSL directly interferes with the secondary pathway of the virus infection. This supposition is valid when applied to the cells expressing lectin receptors, which may not be the case in the present study.

Potential of PhoSL as an antiviral drug

In the present study, we revealed that PhoSL had an antiviral effect on both the Wuhan and Omicron strains of SARS-CoV-2 at the submicromolar level. This finding is based on the strong affinity of PhoSL with the S protein at a low nanomolar level, which causes aggregation, likely due to the multivalence of both PhoSL and the S protein. We emphasise that PhoSL is likely to exert a broad-spectrum effect on SARS-CoV-2 variants, probably because the N-glycosylation sites are well conserved, which represents a significant advantage over the neutralising antibodies.

To use PhoSL as an antiviral drug, however, we need to reduce the off-target effects on core-fucosylated N-glycans in other proteins or cells. Among the lectins that show antiviral activity, PhoSL recognises the position closest to the glycan root. Consequently, numerous interactions between the amino acids of PhoSL and S protein were observed in the models (Fig. 5F,G). By introducing mutations in PhoSL to enhance these interactions or generate additional interactions, we suppose that the specificity to the S protein increases, which should enhance the antiviral effect, regardless of the mechanisms discussed in the previous section. Considering that the IC₅₀ values of representative neutralising antibodies, such as

S309 ($0.08 \mu\text{g}\cdot\text{mL}^{-1}$) [14] and S2X259 ($0.14 \mu\text{g}\cdot\text{mL}^{-1}$) [29], are 1/10–1/20 of PhoSL ($1.6 \mu\text{g}\cdot\text{mL}^{-1}$), enhancing the antiviral effect by 10–20-fold should make PhoSL a potential lead molecule.

PhoSL is a 40 amino acid-long peptide, and its chemical synthesis enables the incorporation of a variety of unnatural amino acids. For example, alteration of Tyr23 to 4-carboxyphenylalanine would reinforce hydrogen bonds by the charge-assisting effect [30], and alteration of Gly24 or Gly26 to L- or D-amino acids with side chains would generate new interactions with the S protein (Fig. 5F,G). Moreover, alteration of Tyr23 to 4-aminophenylalanine and alteration of Thr40 to Asp or Glu would form an electrostatic attraction with Asp339 and Lys440 of the Omicron-variant S protein, respectively (Fig. 6B). Thus, PhoSL has a large degree of freedom with regard to chemical modifications, which is similar to the lead optimisation process in the development of small molecule drugs. In addition, a chemically synthesised material must have merits in terms of the cost of production and storage compared with biologically produced materials, such as antibodies or other lectins. At the same time, PhoSL retains its macromolecular characteristics after trimerisation, thus forming sugar-binding pockets to accept the parts of N-glycans for strict recognition. Therefore, we believe that this study provides a novel and feasible modality for the development of antiviral drugs as well as diagnostic agents.

Materials and methods

Peptide synthesis

Peptides were synthesised by a microwave-assisted solid-phase peptide synthesis [31], using a peptide synthesiser, MWS-1000 (EYELA, Tokyo, Japan). Fmoc-Ala-OH, Fmoc-Asn(Trt)-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Val-OH, Fmoc-Asp(OtBu)-(Dmb)Gly-OH and Fmoc-Asp(OtBu)-Thr(psiMe, Mepro)-OH were purchased from Merck (Darmstadt, Germany) and Watanabe Chemical Industries LTD. (Hiroshima, Japan) and used as mono amino acids synthons or dipeptide synthons. Fmoc-Thr(tBu)-Alko-PEG Resin and Fmoc-Val-Alko-PEG Resin were purchased from Watanabe Chemical Industries LTD. Other reagents were purchased from Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan), Fujifilm Wako Pure Chemical Corp. (Osaka, Japan) and Sigma-Aldrich (St. Louis, MO, USA) and were used without further purification.

A PhoSL 40mer peptide (APVPVTKLVCDGDTYKC-TAYLDYGDGKWVAQWDTAVFHHT) without terminal modifications was chemically synthesised as below; Fmoc-Thr(tBu)-Alko-PEG Resin (150 mg , $0.21 \text{ mmol}\cdot\text{mg}^{-1}$) was allowed to swell in DMF for a period of hours. To the filtered resin was added 20% piperidine in DMF (ca. 1.5 mL) and the mixture was shaken at 50°C under microwave irradiation for 3 min, followed by washing with DMF. The resins were, then, treated with corresponding amino acids synthon (5 equivalents for amino acid synthons or 1.5 equivalent for dipeptide synthons), COMU (1 equivalent for synthons), and DIEA (1 equivalent for synthons) in DMF (ca. 1 mL) at 50°C under microwave irradiation for 5–10 min and then washed with DMF. Subsequently, an acetyl capping reaction was carried out using 13 mm HOBt in $\text{Ac}_2\text{O}/\text{DIEA}/\text{DMF}$ (4.75/2.25/93.0 v/v/v, ca. 1.5 mL) for 3 min at room temperature, followed by washing with DMF. These three reactions, namely deprotection of the Fmoc group, the condensation of amino acids synthons and the acetyl capping, were repeated sequentially to form peptide chain. The obtained peptidyl resin was washed with methylene chloride, dried, and then to which a cleavage cocktail (TFA/ $\text{H}_2\text{O}/\text{EDT}/\text{TIS} = 94/2.5/2.5/1$) (ca. 6 mL) was added. The mixture was shaken at room temperature for 2 h to release peptides from the resin and deprotection at peptide side chains. The resin was filtered and washed with TFA three times. Combined filtrates were concentrated by streaming of nitrogen gas. The resulting residue was precipitated by addition of ice-cold diethyl ether to give a white solid. The solid was washed with ice-cold diethyl ether twice, dissolved in 20% acetonitrile aq. and lyophilized. The obtained crude material was purified by reversed-phase high performance liquid chromatography (RP-HPLC) [conditions; Column: Inertsil ODS-3 ($250 \times 20 \text{ mm}$), Eluent A: water containing 0.1% TFA, eluent B: acetonitrile containing 0.1% TFA, Eluent (A/B = 75/25) was employed, then the ratio of eluent B was increased lineally from 25% to 30% over 60 min with a flow rate of $6.0 \text{ mL}\cdot\text{min}^{-1}$, column oven at 40°C , monitored by UV absorption at 220 nm] to give the targeted PhoSL peptide (55%).

As a control for the infection assay, hepatitis B virus surface antigen Pre-S1 peptide (genotype C) (GTNL SVPNPLGFFPDHQLDPAFGANSNPDWDFNPNKDH WPEANQV: 46mer) also without terminal modification was chemically synthesised as below; Fmoc-Val-Alko-PEG Resin (250 mg , $0.24 \text{ mmol}\cdot\text{mg}^{-1}$) was allowed to swell in DMF for a period of hours. To the filtered resin was added 20% piperidine in DMF (ca. 2 mL) and the mixture was shaken at 50°C under microwave irradiation for 2 min, followed by washing with DMF. The resins were, then, treated with corresponding amino acids synthon (5 equivalents for amino acid synthons), COMU (1 equivalent for synthons), and DIEA (2 equivalents for synthons) in DMF (ca. 1 mL) at 50°C under microwave irradiation for

5–10 min and then washed with DMF. Subsequently, an acetyl capping reaction was carried out using 13 mM HOBt in Ac₂O/DIEA/DMF (4.75/2.25/93.0 v/v/v) (ca. 2 mL) for 1 min at room temperature, followed by washing with DMF. These three reactions, namely deprotection of the Fmoc group, the condensation of amino acids synthons and the acetyl capping, were repeated sequentially to form peptide chain. The obtained peptidyl resin was treated with the same as the synthesis PhoSL. The obtained crude material was purified by RP-HPLC [conditions; Column: Inertsil WP300-C8 (250 × 10 mm), Eluent A: water containing 0.1% TFA, eluent B: acetonitrile containing 0.1% TFA, Eluent (A/B = 90/10) was employed, then the ratio of eluent B was increased linearly from 10% to 40% over 40 min with a flow rate of 4.0 mL·min⁻¹, column oven at 40 °C, monitored by UV absorption at 220 nm] to give the targeted control peptide (21%).

The synthesised peptides were analysed by HPLC and mass spectra (MS). By HPLC [column: InertSustain C18 for PhoSL or Inertsil WP-300 C18 for the control peptide, 4.6 mm × 250 mm; eluate: acetonitrile with 0.1% TFA, gradient increase from 30% to 45% over 15 min for PhoSL or 20% to 50% over 20 min for the control peptide; column oven temperature: 30 °C; flow rate: 1 mL·min⁻¹, detector: 220 nm] we evaluated the purities of both over 99% (Rt 11.0 min for PhoSL, 15.9 min for the control peptide). MALDI-TOF MS spectra were recorded on a MALDI-8020 (Shimadzu, Kyoto, Japan) in linear positive ion mode with *a*-cyano-4-hydroxycinnamic acid (CHCA, 10 mg·mL⁻¹ in 50% acetonitrile solution containing 0.1% TFA) as matrix. MS data were obtained as average masses, namely PhoSL: *m/z* calcd for C₂₀₂H₂₉₆N₄₈O₆₁S₂ [M + H]⁺ 4437.985 (avg.), found 4438.129; Control peptide: *m/z* calcd for C₂₃₁H₃₂₁N₆₃O₇₂ [M + H]⁺ 5133.478 (avg.), found 5133.943.

The inactive double mutant Trp28Thr-Val36Tle (L-tert-Leu) of PhoSL with purity > 95% by HPLC was purchased from BEX Co., Ltd. (Tokyo, Japan).

Crystallography

The PhoSL trimer was purified by gel filtration using a Superdex 75 10/300 column (Cytiva, Marlborough, MA, USA), after which the concentration was evaluated by absorbance at 280 nm. Crystals of the complex of PhoSL and trisaccharide fucose(α1-6)[GlcNAc(β1-4)]GlcNAc-OH (Tokyo Kasei, Tokyo, Japan) were obtained for the mixture (0.15 μL each) of a sample solution [5.5 mg·mL⁻¹ PhoSL (1.2 mM as a monomer) and 1.86 mM of the trisaccharide in 3.5 mM HEPES (7.5)] and a mother liquor solution [4.3 M sodium formate (pH 7.0) (Hampton, Aliso Viejo, CA, USA)] by hanging-drop vapour diffusion against 50 μL of the above mother liquor. After incubation at 20 °C for 4 days, rod-like crystals with a length of 0.1 mm and width of 0.01 mm were formed.

X-ray diffraction was measured at the BL-17A beamline in the Photon Factory, KEK (Tsukuba, Japan). Crystal parameters were obtained by processing the data using xds [32] (Table 1). The diffraction phase was determined by molecular replacement using the PHASER module [33] of the CCP4 program package [34]. A structure generated by merging free PhoSL (PDB: 6A86; chains A, B, and C) [24] and the bound trisaccharide structure in our previous MD simulation [24] was used as the template model. Refinement and model building were performed using the REFMAC5 [35] and COOT [36] modules in CCP4. The co-ordinates and structure factor data were deposited in PDB under accession ID 7VU9.

Surface plasmon resonance

Surface plasmon resonance measurements for the S protein-PhoSL interactions were performed using a Biacore X apparatus (Cytiva) at 298 K. The Wuhan-strain S protein (Val16–Pro1213 containing Arg683Ala, Arg685Ala, Phe817Pro, Ala892Pro, Ala899Pro, Ala942Pro, Lys986Pro, and Val987Pro mutations and a His-Tag at the C-terminus; produced in HEK293 cells; AcroBiosystems, Newark, DE, USA) or the Omicron-variant S protein (Val16–Pro1213 containing Ala67Val, His69–Val70 deletion, Thr95Ile, Gly142Asp, Val143–Tyr145 deletion, Asn211 deletion, Leu212Ile, Arg214–Glu-Pro-Glu-Asp215 insertion, Gly339Asp, Ser371Leu, Ser373Pro, Ser375Phe, Lys417Asn,

Table 1. Data collection and refinement statistics in crystallography of the complex of PhoSL trimer and fucose(α1-6)[GlcNAc(β1-4)]GlcNAc-OH.

Crystallographic data	
Space group	C222
Unit cell	
<i>a/b/c</i> (Å)	126.7/127.1/39.4
<i>α/β/γ</i> (°)	90.0/90.0/90.0
Wavelength (Å)	1.000
Resolution range (outer shell) (Å)	44.87–2.15 (2.28–2.15)
Total reflections	115 974
Unique reflections	17 730
Completeness (outer shell) (%)	99.5 (99.3)
<i>R</i> _{merge} (outer shell) (%)	8.7 (66.1)
Average <i>I</i> /σ(<i>I</i>) (outer shell)	9.18 (1.81)
Refinement	
<i>R</i> _{work} / <i>R</i> _{free} (%)	17.9/24.5
RMSD from ideal values	
Bond length (Å)	0.010
Bond angle (°)	1.77
Ramachandran plot	
Favoured (%)	100.0
Average <i>B</i> -factors (Å ²) (number of atoms)	
Protein	41.8 (1902)
Glycan	42.2 (234)
Formate	41.2 (18)
Water	44.6 (154)

Asn440Lys, Gly446Ser, Ser477Asn, Thr478Lys, Glu484Ala, Gln493Arg, Gly496Ser, Gln498Arg, Asn501Tyr, Tyr505His, Thr547Lys, Asp614Gly, His655Tyr, Asn679Lys, Pro681His, Asn764Lys, Asp796Tyr, Asn856Lys, Gln954His, Asn969Lys, and Leu981Phe mutations in addition to the above stabilisation mutations and a His-tag for the Wuhan protein; HEK293; AcroBiosystems) was immobilised on the sensor chip CM5 (Cytiva) for 5500–7500 resonance units (RU) by the amine coupling method. HBS-EP+ buffer [10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, and 0.05% surfactant P20] was used as the running buffer. The same buffer solutions containing wild-type or mutant PhoSL at concentrations of 10 nM–1 μ M were injected into the flow cells at 10–20 μ L min⁻¹ for 5 min. At lower concentrations, the injection was repeated until the reaction reached a plateau. The molecular weight of the glycosylated S proteins was estimated to be both 180 kDa per monomer, according to the SDS/PAGE and multiangle light scattering data shown in the vendor's sheets.

Inhibition of S protein-PhoSL binding by trisaccharide fucose(α 1–6)[GlcNAc(β 1–4)]GlcNAc-OH was also analysed in the same buffer system with a constant concentration of PhoSL (100 nM) and 0–1 mM of the trisaccharide.

For the binding of RBD (Arg319–Phe541, Val367Phe, C-terminal His-Tag; HEK293; AcroBiosystems) and PhoSL, the former was immobilised on the Sensor Chip NTA (Cytiva) at ~2000 RU. The running buffer was 10 mM HEPES (pH 7.4), 150 mM NaCl, 50 μ M EDTA, and 0.005% surfactant P20. The molecular weight of the glycosylated protein is 34 kDa (on average), as shown in the vendor's sheet.

For inhibition of S protein-ACE2 (Gln18–Ser740 with a C-terminal His-Tag; HEK293; Thermo Fisher Scientific, Waltham, MA, USA) binding by PhoSL, ACE2 was immobilised on the Sensor Chip NTA (Cytiva), where the same running buffer as above for RBD was used for analysis. The concentration of the S protein was fixed to 3 nM where that of PhoSL was varied from 0 to 1 μ M.

The relationships between the equilibrium response values (y) at different concentrations of analyte (PhoSL) (x) were fitted using an in-house FORTRAN90 program to a simple 1 : 1 binding equation, $y = a/(1 + b/x)$, where a and b are fitting variables that represent the maximal response at the infinite analyte concentration (R_{\max}) and dissociation constant (K_D), respectively. Similarly, the relationships between the equilibrium responses (y) at different concentrations of trisaccharide (x) were fitted to the Hill equation, $y = a/(1 + (x/b)^c)$, where a , b , and c are fitting variables that represent the responses without inhibition, IC_{50} , and Hill coefficient, respectively. For the simultaneous kinetic fitting, vendor's BIAEVALUATION software ver. 3 (Cytiva, Marlborough, MA, USA.) was used. To estimate the uncertainty level for key parameters, experiments were duplicated or triplicated.

Virus inhibition assay

To evaluate the neutralisation activity, a PRNT was performed in a biosafety level 3 (BSL-3) laboratory at Nagasaki University. Briefly, the wild-type PhoSL peptide and the control 46mer peptide dissolved in 20 mM Tris buffer (pH 7.5) containing 1 mM DTT were serially diluted 10-fold (from 10 μ M to 1 nM) in minimum essential medium (MEM) supplemented with 2% foetal calf serum (FCS). An equal volume of 100 PFU/200 μ L/well virus (TY-WK-521/2022, Wuhan strain and TY38-873, BA.1, Omicron strain of SARS-CoV-2) was mixed with diluted peptides and incubated at 37 °C for 60 min. Each mixture was inoculated onto a Vero-E6 cell monolayer in 24-well plates in duplicate. After incubation at 37 °C for 60 min, the infected cells were overlaid with 1.25% methylcellulose 4000 in 2% FCS in MEM and incubated for 5 days. The plates were washed with PBS (–) to remove the methylcellulose, fixed with 4% paraformaldehyde solution for overnight at room temperature, rinsed, and stained with crystal violet.

Using an in-house FORTRAN90 program, plaque numbers (y) as a function of concentrations of PhoSL (x) were fitted to the Hill equation, $y = a/(1 + (x/b)^c)$, where variables a , b , and c are the background plaque numbers, IC_{50} , and Hill coefficient, respectively. Calculations were carried out against all combinations of the duplicated data, to evaluate errors in the parameters.

Cryo-EM analysis

For cryo-grid preparation, 3 μ L of the S protein sample (3.2 μ M in the absence or presence of 138 μ M PhoSL, both the concentrations are for monomer) dissolved in 5 mM Tris (8.0) buffer containing 50 mM NaCl and 0.02% NaN₃ was applied onto a holey carbon grid (Quantifoil, Cu, R1.2/1.3, 300 mesh). The grid was rendered hydrophilic by a 30 s glow-discharge in air (11 mA current) with PIB-10 (Vacuum Device Inc., Ibaraki, Japan). The grid was then blotted for 5 s (blot force 15) at 18 °C and 100% humidity before being flash-frozen in liquid ethane using Vitrobot Mark IV (Thermo Fisher Scientific).

For automated data collection, 437 micrographs were acquired using a Talos Arctica G2 microscope (Thermo Fisher Scientific) operating at 200 kV in nanoprobe mode and EPU software. Movie micrographs were collected at a resolution of 4 k \times 4 k using a Falcon 3EC direct electron detector (electron counting mode) at a nominal magnification of 120 000 (0.88 \AA -pixel⁻¹). Fifty movie fractions were recorded at an exposure of 1.00 electrons per \AA^2 per fraction, which corresponded to a total exposure of 50 e⁻ \cdot \AA^{-2} . The defocus steps used were –1.0, –1.5, –2.0, and –2.5 μ m. The movies were processed using MOTIONCOR2 [37], GCTF [38], and RELION3.1 [39].

Dynamic light scattering

Measurements were performed on DynaPro NanoStar instrument (Wyatt Technology, Santa Barbara, CA, USA) equipped with a 658-nm GaAs laser, using disposable 4 μ L-volume cells (Wyatt). Solutions contained 3.0 μ M S protein (as monomer) in the absence or presence of 4.0 or 10.0 μ M PhoSL (also as monomer) in the same buffer as Cryo-EM experiment.

Molecular modelling

The structural models for PhoSL–S protein interactions at Asn331 or Asn343 sites were produced by MD simulation, essentially as previously performed [24]. The initial models were generated by docking the crystal structure of PhoSL complexed with trisaccharide fucose(α 1–6)[GlcNAc(β 1–4)]GlcNAc-OH (Fig. 1B), where the bound waters were eliminated, and a region including RBD (Val320–Ser591) in the cryo-EM structure of the glycosylated S protein (PDB entry: 6XKL; chain A, the RBD-up conformer) that exhibits two GlcNAcs (the inner GlcNAcs numbered in Fig. 1A) at the Asn331 and Asn343 sites. Among them, GlcNAc1 in the two structures were made overlapped, under the condition that the proteins do not cause steric clashes. After docking, the three mannoses were attached to GlcNAc2 to complete the N-glycan core structure.

Using the LEaP module of the AMBER 2018 package [40], the parameter and coordinate files were generated in the Amberff14SB [41] and GLYCAM_06j-1 [42] force fields via the following processes: (a) protons were attached to the protein and glycans, (b) the net charge of the total system was neutralised by the addition of sodium ions, and (c) a box of TIP3P water molecules was set to a thickness of 15 Å.

Energy minimisation was performed in a stepwise manner, namely, for all protons in the system, all atoms except for those of the protein backbone and sugar ring, and all atoms in the system. Preparatory MD calculations were performed in a stepwise manner: 60 ps at 100 K under a constant pressure of 1 atm with positional restraints on atoms except for waters and sodium ions, which were reduced by six steps (300, 100, 20, 5, 2, and 1 kcal·mol⁻¹·Å⁻²); 100 ps by essentially the same protocol as the preceding one, excluding the positional restraints; 100 ps each by essentially the same protocol, at 200 K and at 300 K; and 200 ps at 300 K in the constant volume mode for equilibrium. Then, the production run for 10 ns at 300 K was carried out also in the constant volume mode. During the MD calculations, bonds involving hydrogen were constrained by the SHAKE protocol [43], and a cutoff distance of 8 Å was applied for the noncovalent interactions. The time step was 2 fs for the production run and 1 fs for the other runs.

The CPPTRAJ module [44] was used to analyse the trajectory, where the representative structures were selected as the

lowest root mean square distances from the average coordinates of the non-proton atoms over the production run.

Acknowledgements

We thank Prof K. Hirayama of Nagasaki University for organisation of this collaboration and Prof T. Senda of KEK for organising the cryo-EM facility. We also thank Dr A. Kameyama of AIST for help with the mass spectroscopy measurements, Dr H. Y. Yu for help with the DLS measurements, and Drs M. Ikemoto, Y. Kato, Y. Onishi, and Y. Hagihara of AIST for the encouragement and support. The cryo-EM data were obtained at the KEK Laboratory via the Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS) program of the Japan Agency for Medical Research and Development (AMED) under grant number JP21am0101071 (Prof Toshiya Senda of KEK) and support number 2986 (KY). The X-ray diffraction data were also obtained at KEK, and MD simulations were performed at the Information Technology Center, Nagoya University. This project was supported by a Grant-in-Aid for Scientific Research (B) of the Japan Society for the Promotion of Science (JSPS) under grant number 18H02397 (KY, TK, HS, and HT), and also by the internal Grant-L program of AIST (KY, TK, HS, and HT).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

KY, NA, MMNT, MK, HS, HT, and KM designed and supervised the experiments. KY, NA, MMNT, MK, TY, TK, and IN performed the experiments. KY, NA, MMNT, AI, and TM analysed the data. KY, NA, MMNT, and HS wrote the paper.

Peer Review

The peer review history for this article is available at <https://publons.com/publon/10.1111/febs.16599>.

Data availability statement

Most data are included in the manuscript, whereas those not shown will be provided upon request. The crystal structure of PhoSL in complex with the core-fucose trisaccharide is available in Protein Data Bank (7VU9).

References

- Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. 2020;**181**:271–80.e8.
- Monteil V, Kwon H, Prado P, Hagelkruys A, Wimmer RA, Stahl M, et al. Inhibition of SARS-CoV-2 infections in engineered human tissues using clinical-grade soluble human ACE2. *Cell*. 2020;**181**:905–13.e7.
- Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell*. 2020;**181**:894–904.e9.
- Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 2020;**367**:1260–3.
- Ke Z, Oton J, Qu K, Cortese M, Zila V, McKeane L, et al. Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature*. 2020;**588**:498–502.
- Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science*. 2020;**369**:330–3.
- Grant OC, Montgomery D, Ito K, Woods RJ. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci Rep*. 2020;**10**:14991.
- Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*. 2020;**367**:1444–8.
- Ejemel M, Li Q, Hou S, Schiller ZA, Tree JA, Wallace A, et al. A cross-reactive human IgA monoclonal antibody blocks SARS-CoV-2 spike-ACE2 interaction. *Nat Commun*. 2020;**11**:4198.
- Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol*. 2021;**19**:409–24.
- Kuhlmann C, Mayer CK, Claassen M, Maponga T, Burgers WA, Keeton R, et al. Breakthrough infections with SARS-CoV-2 omicron despite mRNA vaccine booster dose. *Lancet*. 2022;**399**:625–6.
- Amraei R, Yin W, Napoleon MA, Suder EL, Berrigan J, Zhao Q, et al. CD209L/L-SIGN and CD209/DC-SIGN act as receptors for SARS-CoV-2. *ACS Cent Sci*. 2021;**7**:1156–65.
- Lempp FA, Soriaga LB, Montiel-Ruiz M, Benigni F, Noack J, Park YJ, et al. Lectins enhance SARS-CoV-2 infection and influence neutralizing antibodies. *Nature*. 2021;**598**:342–7.
- Pinto D, Park YJ, Beltramello M, Walls AC, Tortorici MA, Bianchi S, et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature*. 2020;**583**:290–5.
- Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*. 2020;**181**:281–92.e6.
- Cui Z, Liu P, Wang N, Wang L, Fan K, Zhu Q, et al. Structural and functional characterizations of infectivity and immune evasion of SARS-CoV-2 omicron. *Cell*. 2022;**185**:860–71.e13.
- Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, et al. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell*. 2020;**182**:1284–94.e9.
- Hoffmann D, Mereiter S, Oh YJ, Monteil V, Elder E, Zhu R, et al. Identification of lectin receptors for conserved SARS-CoV-2 glycosylation sites. *EMBO J*. 2021;**40**:e108375.
- Liu YM, Shahed-Al-Mahmud M, Chen X, Chen TH, Liao KS, Lo JM, et al. A carbohydrate-binding protein from the edible lablab beans effectively blocks the infections of influenza viruses and SARS-CoV-2. *Cell Rep*. 2020;**32**:108016.
- Ahan RE, Hanifehnezhad A, Kehribar ES, Oguzoglu TC, Foldes K, Ozcelik CE, et al. A highly potent SARS-CoV-2 blocking lectin protein. *ACS Infect Dis*. 2022;**8**:1253–64.
- Kobayashi Y, Tateno H, Dohra H, Moriwaki K, Miyoshi E, Hirabayashi J, et al. A novel core fucose-specific lectin from the mushroom *Pholiota squarrosa*. *J Biol Chem*. 2012;**287**:33973–82.
- Yamasaki K, Yamasaki T, Tateno H. The trimeric solution structure and fucose-binding mechanism of the core fucosylation-specific lectin PhoSL. *Sci Rep*. 2018;**8**:7740.
- Cabanettes A, Perkams L, Spies C, Unverzagt C, Varrot A. Recognition of complex core fucosylated N-glycans by a mini lectin. *Angew Chem Int Ed Engl*. 2018;**57**:10178–81.
- Yamasaki K, Kubota T, Yamasaki T, Nagashima I, Shimizu H, Terada RI, et al. Structural basis for specific recognition of core fucosylation in N-glycans by *Pholiota squarrosa* lectin (PhoSL). *Glycobiology*. 2019;**29**:576–87.
- Kuo CW, Yang TJ, Chien YC, Yu PY, Hsu SD, Khoo KH. Distinct shifts in site-specific glycosylation pattern of SARS-CoV-2 spike proteins associated with arising mutations in the D614G and alpha variants. *Glycobiology*. 2022;**32**:60–72.
- Guo Y, Feinberg H, Conroy E, Mitchell DA, Alvarez R, Blixt O, et al. Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nat Struct Mol Biol*. 2004;**11**:591–8.
- Kelm S, Pelz A, Schauer R, Filbin MT, Tang S, de Bellard ME, et al. Sialoadhesin, myelin-associated glycoprotein and CD22 define a new family of sialic

- acid-dependent adhesion molecules of the immunoglobulin superfamily. *Curr Biol.* 1994;**4**:965–72.
- 28 Feinberg H, Mitchell DA, Drickamer K, Weis WI. Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science.* 2001;**294**:2163–6.
- 29 Tortorici MA, Czudnochowski N, Starr TN, Marzi R, Walls AC, Zatta F, et al. Broad sarbecovirus neutralization by a human monoclonal antibody. *Nature.* 2021;**597**:103–8.
- 30 Gilli G, Gilli P. Towards an unified hydrogen-bond theory. *J Mol Struct.* 2000;**552**:1–15.
- 31 Merrifield RB. Solid phase peptide synthesis. 1. Synthesis of a tetrapeptide. *J Am Chem Soc.* 1963;**85**:2149–54.
- 32 Kabsch W. Xds. *Acta Crystallogr D Biol Crystallogr.* 2010;**66**:125–32.
- 33 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Cryst.* 2007;**40**:658–74.
- 34 Potterton E, Briggs P, Turkenburg M, Dodson E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr.* 2003;**59**:1131–7.
- 35 Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr.* 1997;**53**:240–55.
- 36 Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* 2010;**66**:486–501.
- 37 Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng YF, Agard DA. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods.* 2017;**14**:331–2.
- 38 Zhang K. Gctf: real-time CTF determination and correction. *J Struct Biol.* 2016;**193**:1–12.
- 39 Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJH, Lindahl E, et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife.* 2018;**7**:e42166.
- 40 Case DA, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham TE III, Cruzeiro VWD, et al. AMBER 2018. San Francisco: University of California, San Francisco; 2018.
- 41 Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput.* 2015;**11**:3696–713.
- 42 Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, et al. GLYCAM06: a generalizable biomolecular force field. *Carbohydrates. J Comput Chem.* 2008;**29**:622–55.
- 43 Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical-integration of cartesian equations of motion of a system with constraints – molecular-dynamics of N-alkanes. *J Comput Phys.* 1977;**23**:327–41.
- 44 Roe DR, Cheatham TE III. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput.* 2013;**9**:3084–95.