Trends Phytochem. Res. 6(2) 2022 86-96

Trends in Phytochemical Research (TPR)

Journal Homepage: http://tpr.iau-shahrood.ac.ir

Original Research Article

Contribution of the lycotetraose moiety of α-tomatine to the interaction with the main proteases of coronaviruses PEDV and SARS-CoV-2

Gerard Vergoten[1] And Christian Bailly[2]*

[1]University of Lille, Inserm, INFINITE - U1286, Institut de Chimie Pharmaceutique Albert Lespagnol (ICPAL), Faculté de Pharmacie, 3 rue du Professeur Laguesse, BP-83, F-59006, Lille, France
[2]OncoWitan, Scientific Consulting Office, Lille (Wasquehal), 59290, France

ABSTRACT

We have investigated the interaction of the steroidal alkaloid tomatidine (TD) and its glycosylated derivative α-tomatine (αTM) with the 3C-like protease (3CLpro) of PEDV and with the main protease (Mpro) from the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) responsible for COVID-19, using molecular modeling. αTM can generate more stable complexes with both proteases than TD. The lycotetraosyl moiety of αTM largely occupies the binding cavity, pushing the steroid moiety outside of the groove, whereas with the aglycone TD fits into the site. The specific contribution of the tetrasaccharide was further investigated via the modeling of the unsubstituted lycotetraose molecule, and the glycoalkaloid demissine (DEM) bearing a lycotetraose unit and a different steroidal skeleton demissidine. The binding of DEM to the proteases is essentially driven by the carbohydrate moiety, projecting the aglycone outside of the binding cavity. This study shed light on the protein-binding capacity of the lycotetraose unit.

ARTICLE HISTORY

Received: 10 March 2022
Revised: 13 April 2022
Accepted: 10 May 2022
ePublished: 25 June 2022

KEYWORDS

Lycotetraose
Main protease
PEDV
SARS-CoV-2
Tomatidine
Tomatine

1. Introduction

The steroidal alkaloids tomatidine (TD) and α-tomatine (αTM), present to a large extent in the skin of green unripe tomatoes, serve to protect the plant and its fruits against attacks by viral, fungal and bacterial pathogens, and to deter herbivorous insects and parasites. The secondary metabolite αTM, a glycosylated derivative of the saponin aglycone TD (Fig. 1), accumulates to millimolar levels in vegetative tomato tissues. Both compounds display a large range of bioactivities, including antioxidant and anti-inflammatory properties, but αTM exhibits much higher cytotoxic effects than TD (Friedman, 2013, 2015). For example, αTM (IC_{50} = 3 ± 0.3 μM) is 200-times more cytotoxic to prostate PC3 cancer cells than the aglycone TD (IC_{50} = 599 ± 27 μM) (Choi et al., 2012). Certain tomato fungi overcome the basal defense barrier by the secretion of enzymes called tomatinases which degrade αTM into less fungitoxic compounds βTM and TD (Fig. 1) (Ökmen et al., 2013; Carere et al., 2017). The tolerance of fungi to αTM varies significantly from one strain to another. For example, the fungus *Neurospora crassa* is much more sensitive to αTM compared to βTM and TD (EC_{50} = 13, 122, 51 μM, respectively). In contrast, the pathogenic fungi *Alternaria alternata* and *A. solani* are practically insensitive to all three compounds (EC_{50} = > 1 mM) (Sandrock and Vanetten, 1998).

Both αTM and TD have been extensively investigated to define their nutritional and medicinal properties, and to delineate their mechanism of action (Bailly, 2021). Among their diverse activities, prominent antiviral effects have been noticed. αTM has been found to inactivate the herpes simplex virus (HSV), whereas TD was inactive against this virus. But the aglycone has revealed marked antiviral effects against other viruses, such as the dengue virus (DENV), chikungunya virus (CHIKV), and porcine epidemic diarrhea virus (PEDV).

✉ Corresponding author: Christian Bailly
Tel: +33 622 66 18 17; Fax: +33 622 66 18 17
E-mail address: christian.bailly@oncowitan.com, **doi: 10.30495/tpr.2022.1954759.1247**

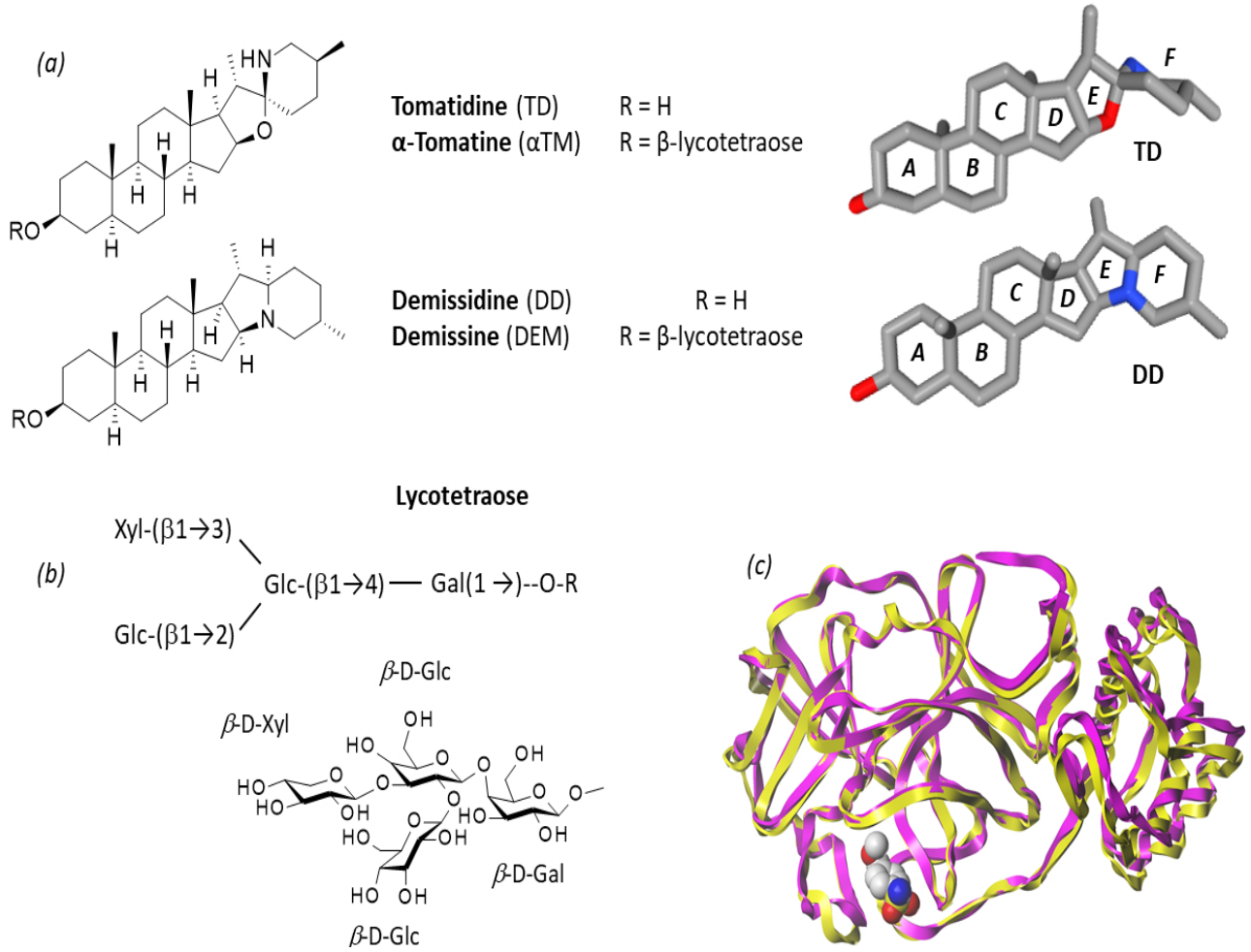

Fig. 1. (*a*) Structures of tomatidine (TD), α-tomatine (αTM), demissidine (DD), and demissine (DEM). Tridimensional hexacyclic models of TD and DD are presented. (*b*) Structure of the branched β-lycotetraose tetra-saccharide common to αTM and DEM (3-*O*-(β-D-glucopyranosyl-(1→2)-[β-D-xylopyranosyl-(1→3)]-β-D-glucopyranosyl-(1→4)-β-D-galactopyranoside). Cleavage of the lycotetraosyl unit by tomatinase releases TD and lycotetraose. (c) Superimposed ribbon structures of the two viral proteases: Mpro of SARS-CoV-2 with its co-crystalized ligand (PDB code: 5R80) and 3CLpro of PEDV (PDB code: 4XFQ), in green and cyan, respectively.

TD potently inhibits multiplication of serotype DENV-2 (EC_{50} = 0.82 ± 0.04 μM in Huh7 hepatocellular carcinoma cells, abolishing the production of virus particles during and after infection (Diosa-Toro et al., 2019). TD blocked virus multiplication at steps downstream of the virus cell entry and membrane fusion but prior to secretion of progeny virions. It is an interesting compound to combat Dengue virus, endemic in many tropical and sub-tropical countries. Similarly, the alkaloid was found recently to inhibit production of infectious CHIKV particles *in vitro*. TD reduces the infectious capacity of secreted viral particles, with a micromolar efficacy (EC_{50} = 1.3-3.8 μM) toward different CHIKV strains/genotypes (Troost et al., 2020; Troost-Kind et al., 2021). Then, TD has been identified as an inhibitor of PEDV replication in cultured Vero and IPEC-J2 cells *in vitro*. A direct inhibition of the 3CL protease of PEDV was evidenced both *in silico* and experimentally using purified recombinant 3CLpro enzyme. Biophysical measurements indicated that TD directly interacts with 3CLpro (K_d = 2.78 μM), in agreement with molecular modeling and molecular dynamic calculations which predicted that TD could bind to the active site of the protease. The effect was validated at the cellular level, using cells transfected with active or inactive mutants to confirm that the protease is indeed a target for the alkaloid (Wang et al., 2020a). In addition, the authors showed that TD can inhibit replication of other viruses, such as the transmissible gastroenteritis virus (TGEV), porcine reproductive and respiratory syndrome virus (PRRSV), encephalo myocarditis virus (EMCV) and seneca virus A (SVA), at least *in vitro* (Wang et al., 2020a). TD is now emerging as an interesting antiviral agent against the α-coronavirus PEDV which is the main causative agents of severe diarrhea in lactating pigs and suckling piglets.

These considerations prompted us to investigate and compare the potential binding of TD and its glycoside αTM to the PEDV 3C-like protease ($3CL^{pro}$), taking advantage of the crystallographic structure of the protein (Ye et al., 2016). Moreover, we studied the binding of these molecules to the main protease

(M^{pro}) of SARS-CoV-2 because these two coronavirus main proteases present a highly conserved structure and catalytic mechanism (Fig. 1). We reasoned that an inhibitor of PEDV-$3CL^{pro}$ like TD might be able to bind to SARS-CoV-2 M^{pro}. This has been observed with other compounds, such as the dipeptidyl inhibitor GC376 targeting both proteases (Wang et al., 2020b). We found that αTM can form stable complexes with the two proteases but in this case, it is the glycoside moiety of the steroidal alkaloid which plays a major role in the protein interaction. The specific contribution of the lycotetraose portion of the drug has been determined, via the analysis of the tetrasaccharide itself (β-D-glucopyranosyl-(1→2)-*O*-[β-D-xylopyranosyl-(1→3)]-*O*-β-D-glucopyranosyl-(1→4)-D-galactopyranosyl) and an analogue of αTM, the glycoalkaloid demissine, bearing the same carbohydrate but a distinct aglycone (Fig. 1). For the first time, our computational analysis reveals the contribution of the lycotetraose unit to the drug-protein interaction. The role of this oligosaccharide, that can be found in other natural products, is discussed.

2. Experimental

2.1. *In silico* molecular docking procedure

The tridimensional structures of the SARS-CoV-2 main protease in complex with a small molecule inhibitor (Z18197050) and the PEDV 3C-like protease were retrieved from the Protein Data Bank (www.rcsb.org) under the PDB codes 5R80 and 4XFQ, respectively (Ye et al., 2016; Douangamath et al., 2020). Docking experiments were performed with the GOLD software (GOLD 5.3 release, Cambridge Crystallographic Data Centre, Cambridge, UK). Before starting the docking procedure, the structure of the ligands has been optimized using a classical Monte Carlo conformational searching procedure as described in the BOSS software (Jorgensen et al., 2004).

With the 5R80 structure, based on shape complementarity criteria, the binding site for the ligands has been defined around amino acid residue Met165 (binding site for the ligand Z18197050). Given the important structural similarities between the two proteases, the position of the binding site is comparable in the 4XFQ structure and is centered on amino acid residue Glu187. Shape complementarity and geometry considerations are in favor of a docking grid centered in the volume defined by this amino acid. Within the binding site, side chains of specific amino acids have been considered as fully flexible. The flexible amino acids are (i) His41, Met49, Tyr54, Cys145, His164, Met165, Glu166, Phe181, Val186 and Asp187 for structure 5R80 and (ii) His41, Thr47, Tyr53, Cys144, Gln163, Leu164, Glu165, Glu185, Asp186 and Glu187 for structure 4XFQ. The ligand is always defined as flexible during the docking procedure. Up to 100 poses that are energetically reasonable were kept while searching for the correct binding mode of the ligand. The decision to keep a trial pose is based on ranked poses, using the PLP fitness scoring function (which is the default in GOLD version 5.3 used here (Jones et al., 1997)). The same procedure was used to establish molecular models for all compounds listed in Table 1. The empirical potential energy of interaction ΔE for the ranked complexes is evaluated using the simple expression ΔE(interaction) = E(complex) - (E(protein) + E(ligand)). For that purpose, the Spectroscopic Empirical Potential Energy function SPASIBA and the corresponding parameters were used (Vergoten et al., 2003; Lagant et al., 2004). Free energies of hydration (ΔG) were estimated using the MM/GBSA model in Monte Carlo simulations within the BOSS software (Jorgensen et al., 2004). The stability of the receptor-ligand complex is evaluated through the empirical potential energy of interaction (Vergoten et al., 2003; Lagant et al., 2004). The Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) procedure was used to evaluate free energies of hydration (Jorgensen et al., 2004) (within the Boss program (Jorgensen and Tirado-Rives, 2005)), in relation with aqueous solubility (Zafar and Reynisson, 2016). Molecular graphics and analysis were performed using Discovery Studio Visualizer, Biovia 2020 (Dassault Systèmes BIOVIA Discovery Studio Visualizer 2020, San Diego, Dassault Systèmes, 2020).

The two most widely used methods to investigate protein-ligand stability and affinity are Molecular Dynamics (MD) and Monte Carlo (MC) simulations. Both methods use an empirical force field to control the total energy (MC, energy minimization) and forces (MD, Newton equations of motion). To use MD simulations confidently, a force field parameterized for dynamical properties is required. The development of a reliable and accurate force field for the conformational analysis is a concern. It requires accuracy of the force field over the whole potential surface, rather than in the region of the global minimum (Homans, 1990). The most used academic force fields (CHARMM, AMBER, GROMOS) do not exhibit the required vibrational spectroscopic quality. Determination of normal modes on a minimized protein structure using the above force fields can result in imaginary wavenumbers corresponding to maxima in the potential energy (transition states, mainly due to inadequate barriers to internal rotation). The spectroscopic SPASIBA force field has been specifically developed to provide refined empirical molecular mechanics force field parameters, as described in other studies (Vergoten et al., 2003; Meziane-Tani et al., 2006). For this reason, we preferred to use MC simulations rather than MD which requires a substantial increase in computer time to achieve the same level of convergence (Jorgensen and Tirado-Rives, 1996).

3. Results and Discussion

3.1. Molecular models of TD and αTM bound to proteases $3CL^{pro}$ and M^{pro}

A molecular docking analysis was performed to compare the interaction of TD and αTM with the protease $3CL^{pro}$ of PEDV, starting from the crystal structure of the protein (PDB: 4XFQ). A similar analysis was performed with the main protease (M^{pro}) of SARS-CoV-2 (PDB: 5R80). Molecular models were constructed with each ligand, and we calculated the empirical energy of interaction (ΔE) and free energy of hydration (ΔG), as reported in Table 1. The difference is considerable between the two

compounds, with large variations in both ΔE and ΔG values.

Table 1
Calculated potential energy of interaction (ΔE, kcal/mol) and free energy of hydration (ΔG, kcal/mol) for the interaction of the indicated natural products with the two viral proteases 3C-like protease ($3CL^{pro}$) of PEDV and the main protease (M^{pro}) of SARS-CoV-2.

Compounds	[a]CID	3CLpro (PEDV)		Mpro (SARS-CoV-2)	
		ΔE	ΔG	ΔE	ΔG
Tomatidine (TD)	65576	-48.3	-21.6	-44.3	-19.1
α-Tomatine (αTM)	28523	-95.7	-35.2	-89.9	-32.3
Lycotetraose	91860151	-81.2	-18.2	-71.8	-20.5
Demissine (DEM)	442975	-79.2	-34	-85	-26.6
Demissidine (DD)	101379	-44.5	-19.1	-45.1	-20.8
Pneumocandin B0	72475	-80.7	-26.1	-97.8	-45

[a]CID: compound identity number, as defined in PubChem (https://pubchem.ncbi.nlm.nih.gov)

The computed empirical energies of interaction are 2-fold more negative with αTM compared to TD and a similar behavior can be seen with the two proteases. For example, with M^{pro} of SARS-CoV-2 the calculated ΔG values was -89.9 kcal/mol with αTM compared to only -44.3 kcal/mol with TD. A similar difference was observed when using protease $3CL^{pro}$, the computed ΔG values was -95.7 kcal/mol with αTM compared to -48.3 kcal/mol with TD. In both cases, the ratio of ΔG values was about 0.5 ($\Delta G^{TD}/\Delta G^{\alpha TM}$). At first sight, this trend suggests that αTM can form much more stable complexes with both $3CL^{pro}$ and M^{pro} compared to TD. The ΔE values calculated with αTM are on the same level as those measured with the antifungal lipohexapeptide pneumocandin B0 which has been described as a ligand of $3CL^{pro}$ and PEDV replication inhibitor (Wang et al., 2020a). Nevertheless, this ligand exhibits a higher aptitude to form complexes with M^{pro} of SARS-CoV-2 compared to $3CL^{pro}$ of PEDV, whereas the reverse is observed with αTM. We could detect more than 25 molecular interactions between αTM and M^{pro}, including H-bonds, van der Waals contacts and (π)alkyl interactions (Fig. 2c). But beyond the calculated values, there are also significant differences in terms of compound orientations within the protein binding site. The protein cavity is centered on residues Glu187 and Met165, for $3CL^{pro}$ and M^{pro}, respectively. In the case of TD, it can be seen very clearly that the pentacyclic chromophore fits well into the cavity to establish van der Waals and (π)-alkyl contacts with the protein surface. On the opposite, in the case of αTM the site is essentially occupied by the glycoside moiety, not much by the aromatic chromophore which tends to project outside of the cavity, with a portion of the ligand completely free of any interaction (Fig. 2b). A similar trend was observed with the protease $3CL^{pro}$ of PEDV (*data not shown*). The pentacyclic core of TD lies into a protein groove, with the edge of the ligand approaching a protein β-sheet. The phenolic OH points toward the interior of the groove and the methyl-piperidine out. In the case of αTM, the protein groove is essentially occupied by the carbohydrate moiety, pushing the steroid core toward the exterior of the cavity, with the methyl-piperidine now completely accessible to the solvent. The carbohydrate portion of αTM inserts less deeply into the groove than the hexacyclic TD core. This can be seen from the analysis of the solvent accessible surface (SAS) in Fig. 2b. However, each of the four units of the lycotetraose moiety establishes H-bonds with the protein. The distinct orientations of the two products within the binding site of M^{pro} is illustrated in Fig. 3. The steroid core of TD fills the site whereas that of αTM protrudes outside of the cavity, occupied by the sugar residues. An overlapping view of the two molecules shows most clearly the difference of positioning between the two compounds (Fig. 3c).

From this first set of data, we concluded that TD is not a highly potent binder to $3CL^{pro}$ and M^{pro}. The capacity of the alkaloid to form stable complexes with the two proteases is well inferior to that of the natural product pneumocandin B0. The glycosylated analogue αTM displays a higher capacity to interact with the proteases, but the lycotetraose moiety profoundly affects the drug-protein interaction. Unexpectedly, the glycoside of aTM appears as the main protein interaction element.

3.2. Molecular models of protease-bound lycotetraose and demissine

We set out to investigate further this aspect, via the modeling of two additional molecules: the unsubstituted lycotetraose unit representing the glycoside portion of αTM, and the natural product demissine (DEM). This latter compound is a steroidal glycoalkaloid bearing the same β-lycotetraosyl moiety as αTM, but with a different steroidal skeleton known as demissidine (Fig. 1). DEM, initially isolated from *Solanum* species (Osman et al., 1976), is structurally similar to solanidine (5-dehydrodemissidine), the aglycon of solanine which is the main alkaloid isolated from the crop potato (*Solanum tuberosum*). Like αTM, DEM is an insecticidal glycoalkaloid (Kozukue et al., 2008; Jared et al., 2016). Models of lycotetraose and DEM bound to the protein $3CL^{pro}$ are presented in Fig. 4. The models obtained with M^{pro} are shown in Fig. 5. The calculated ΔE/ΔG values are

given in Table 1. With both proteins, the unsubstituted tetrasaccharide fits well into the protein binding site, inserting its xylosyl(β1→3) and glucosyl(β1→2) residues inside the narrow cavity. The good fit results in a highly favorable ΔE value with $3CL^{pro}$ (-81.2 kcal/mol) very similar to that measured with DEM (-79.2 kcal/mol). For lycotetraose, the calculated ΔE value is little inferior with M^{pro} (-71.8 kcal/mol) compared to the value with $3CL^{pro}$, a trend comparable to the effect observed with αTM. There is no doubt that the β-lycotetraosyl moiety, common to αTM and DEM, is the main driver of their interaction with the two proteases. The carbohydrate moiety of DEM inserts into the cavity and projects the steroidal portion outside of the groove. This can be clearly seen with both $3CL^{pro}$ and M^{pro} (Fig. 4 and Fig. 5). The situation with DEM is entirely like αTM. In both cases, the protease binding is driven by the lycotetraose unit whereas the steroidal unit is essentially evicted from the binding site. The unsubstituted glycoside inserts more deeply into the site than the steroid-substituted glycoside (Fig. 4b). A superimposed model of αTM and lycotetraose shows the similar positioning of the tetrasaccharide unit in the protein cavity (Fig. 3d). Altogether, the data indicate that binding of the glycoalkaloids αTM and DEM to the two proteases is essentially driven by their common glycoside moiety. The lycotetraose unit could be considered as a protein sensor. Attached to the TD core, this tetrasaccharide provides a protein interaction system and αTM can apparently form very stable complexes with $3CL^{pro}$ of PEDV.

3.3. Lycotetraose: a protein recognition element

The steroidal alkaloid TD is an inhibitor of PEDV replication in cultured cells and its activity has been attributed to a direct binding to and inhibition of the 3CL protease of the porcine virus (Wang et al., 2020a). Our *in silico* analysis indicates that the binding of TD to the protein $3CL^{pro}$ can take place but it is much less favorable than that observed with the reference compound pneumocandin B0 which has also been shown to inhibit PEDV replication (Wang et al., 2020a). The glycosylated analogue αTM is apparently better equipped than TD to form stable complexes with both the $3CL^{pro}$ and M^{pro} proteins. It will be interesting to investigate the anti-PEDV and anti-SARS-CoV-2 effects of this common tomato glycoalkaloid.

Interestingly, we discovered that the lycotetraose unit of αTM can function as a protein recognition element. This glycoside unit clearly drives the interaction of αTM and its analogue demissine (DEM) with the protein. It is therefore not just a solubilizing agent but also a bioactive entity, likely contributing to the diverse pharmacological properties of the saponin αTM. TD is essentially non-cytotoxic whereas αTM displays marked antiproliferative activities against cancer cell lines. The glycoside moiety of αTM is essential to the anticancer activity and partial deletion of the tetra-saccharide, to produce tri-(β1/β2-TM), di-(Y-TM) or mono-saccharides (δ-TM) strongly reduces the anti-proliferative effect (Friedman et al., 2009; Choi et al., 2012). The lycotetraose is a key element of αTM and this is the reason why some fungi have developed a strategy to hydrolyze the glycoside unit of αTM, with specific tomatinase enzymes (Ökmen et al., 2013; Carere et al., 2017).

4. Concluding remarks

Our study shed light on the potential protein-binding role of the lycotetraose moiety. This unit can be found in different natural products, such as αTM and DEM but also other products including (i) uttroside B, an anticancer saponin isolated from *Solanum nigrum* Linn (Nath et al., 2016), (ii) esculeosides A and B-1, bidesmosidic saponins from the ripe fruits of tomato (*Solanum lycopersicum* L.) (Nohara et al., 2010; Manabe et al., 2013), (iii) Karataviosides J and K, bisdesmosidic steroidal glycosides from the bulbs of *Allium karataviense* (Kuroda et al., 2015), (iv) lycoperosides F-G-H isolated from tomato fruits (*Lycopersicon sculentum*) (Yahara et al., 2004) and a few other spirostane-type glycosides (Vasquez et al., 1997). In all cases, the lycotetraose unit has been considered essential to express a strong cytotoxic activity (Ikeda et al., 2003). The lycotetraoside is available synthetically (Takeo et al., 1984; Jones et al., 2015) and can be also prepared by enzymatic hydrolysis of αTM with a recombinant endoglycosidase tomatinase (Woods et al., 2004). It can be linked to cholesterol to prepare lycotetraosyl-cholesterol hybrid compounds with anticancer properties (Ikeda et al., 2006). This tetrasaccharide also contributes to the marked antimicrobial effects of αTM against pathogenic protozoa that infect both humans and animals (Liu et al., 2016). It is therefore important to better understand its intrinsic activities. Our computational analysis suggests a direct role for lycotetraose as a protease binding element. This tetrasaccharide deserves further studies, in particular as a regulatory element of the PEDV 3C-like protease.

In addition, our work reinforces the idea that plant glycosides can provide useful tools to investigate interaction with viral proteases. Here we focused on αTM and its lycotetraosyl glycoside. Other plant glycosides shall be considered further as well, such as saikosaponin A and liquiritin which can antagonize the SARS-CoV-2 spike glycoprotein-binding partner DC-SIGN (Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin) (Gao et al., 2021). For a long time, glycoconjugates of alkaloids have been considered for the design of novel antiviral agents. The glycoconjugation can serve to improve water solubility but also target recognition and drug-protein complex stability (Wu et al., 2014). The use of the lycotetraosyl glycoside as a protein-binding element shall be encouraged.

Abbreviations

$3CL^{pro}$, coronavirus 3C-like protease of PEDV; M^{pro}, main protease of SARS-CoV-2; TD, tomatidine; αTM, α-tomatine; DEM, demissine.

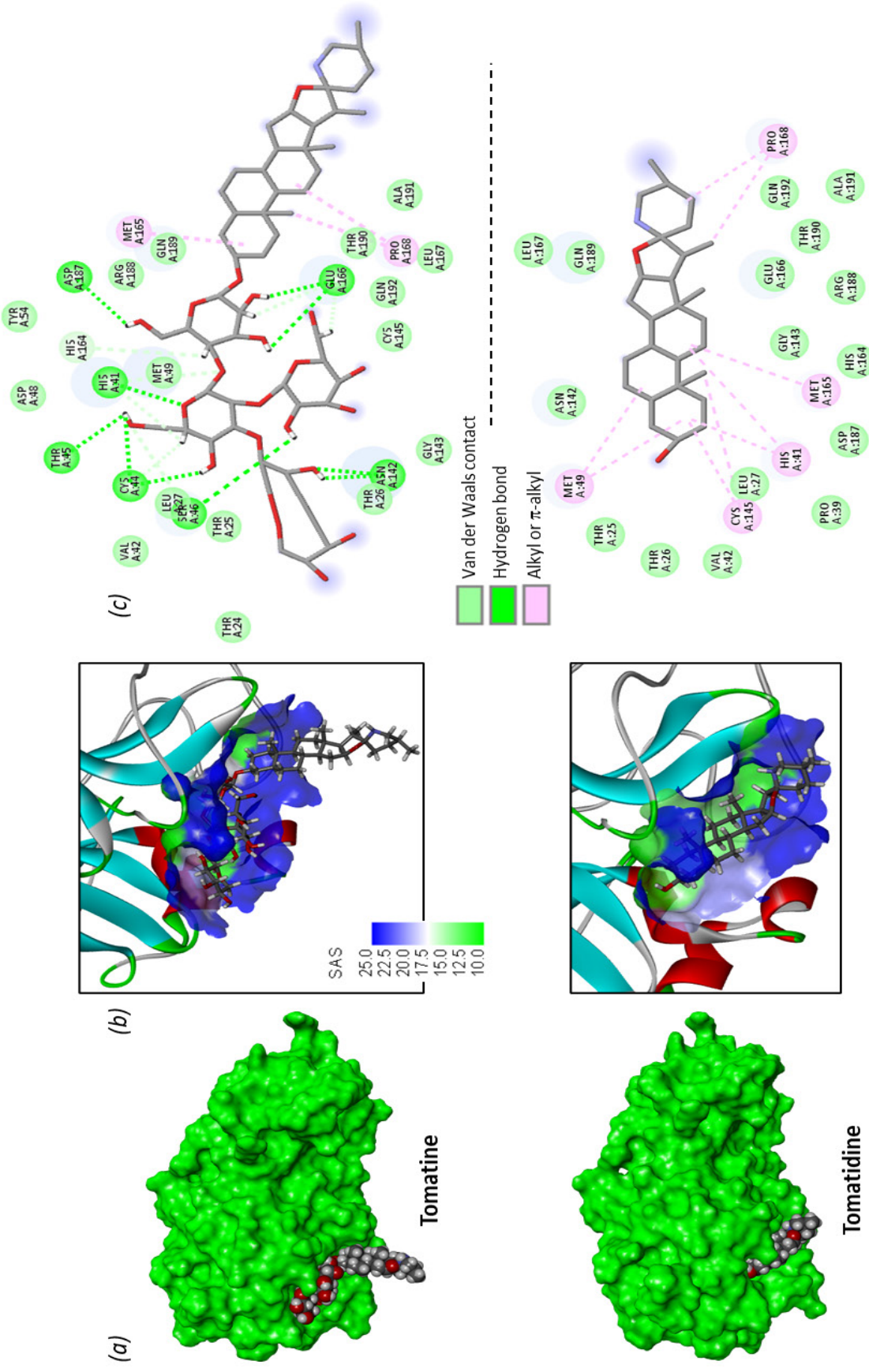

Fig. 2. Molecular models of α-tomatine and tomatidine bound to M^{pro} of SARS-CoV-2. Panel (*a*) shows the ligand bound to the active site of the protein (green). Panel (*b*) shows a detailed view of the compound in the binding site, with the solvent-accessible surface (SAS) defined. Panel (*c*) refers to the binding map contacts for each ligand, with the indicated color code.

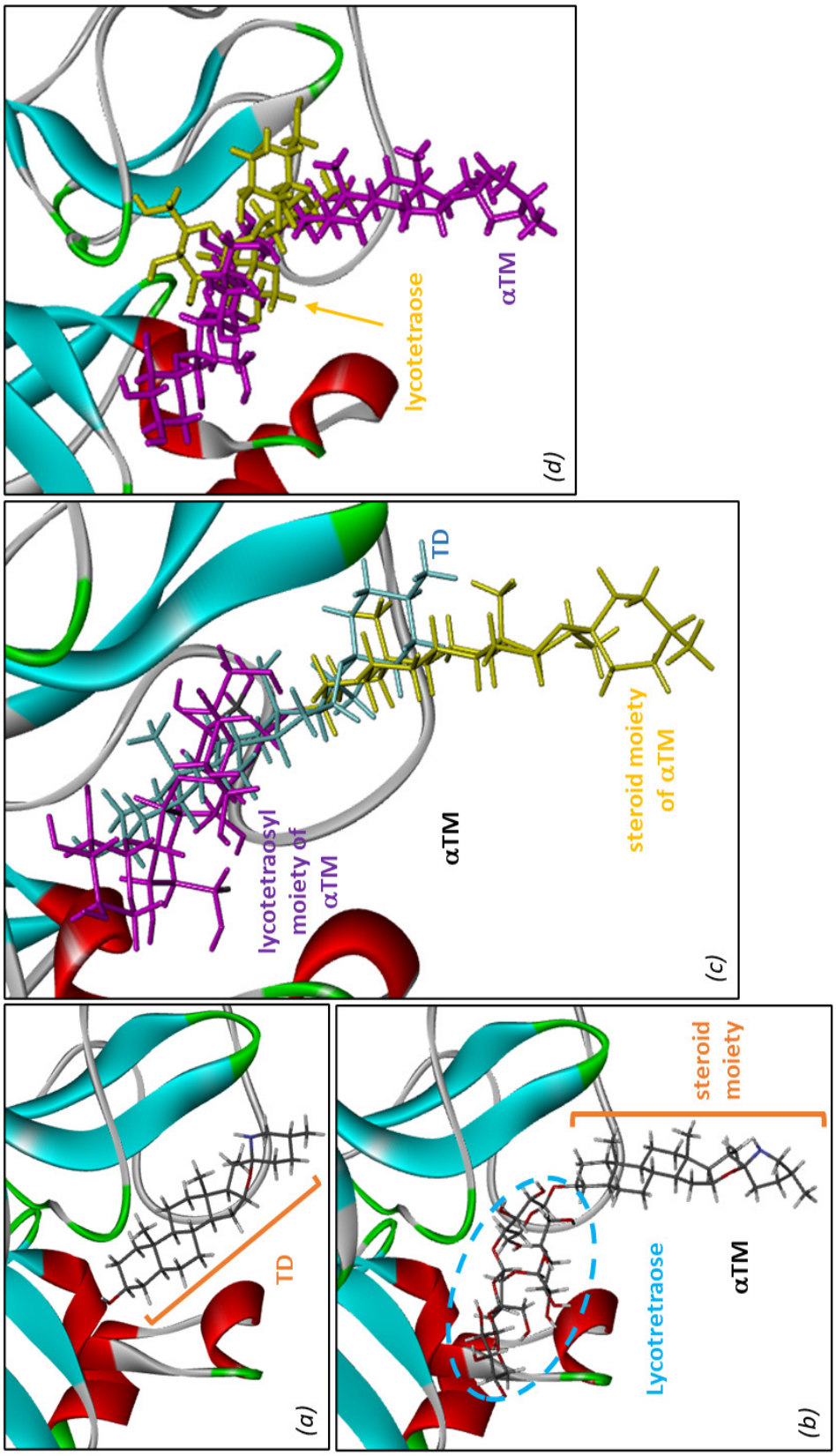

Fig. 3. A detailed view of tomatidine (TD) and tomatine (αTM) bound to the active site of M^{pro} (SARS-CoV-2) to show the different location of the compounds. (*a*) The TD aglycone inserts into the active site, whereas in the case of αTM (*b*), the site is essentially occupied with the lycotetraose unit pushing the steroid moiety outside of the site. Superimposed models of (*c*) αTM and TD and (*d*) αTM and lycotetraose, showing the distinct orientations of the molecules in the active site.

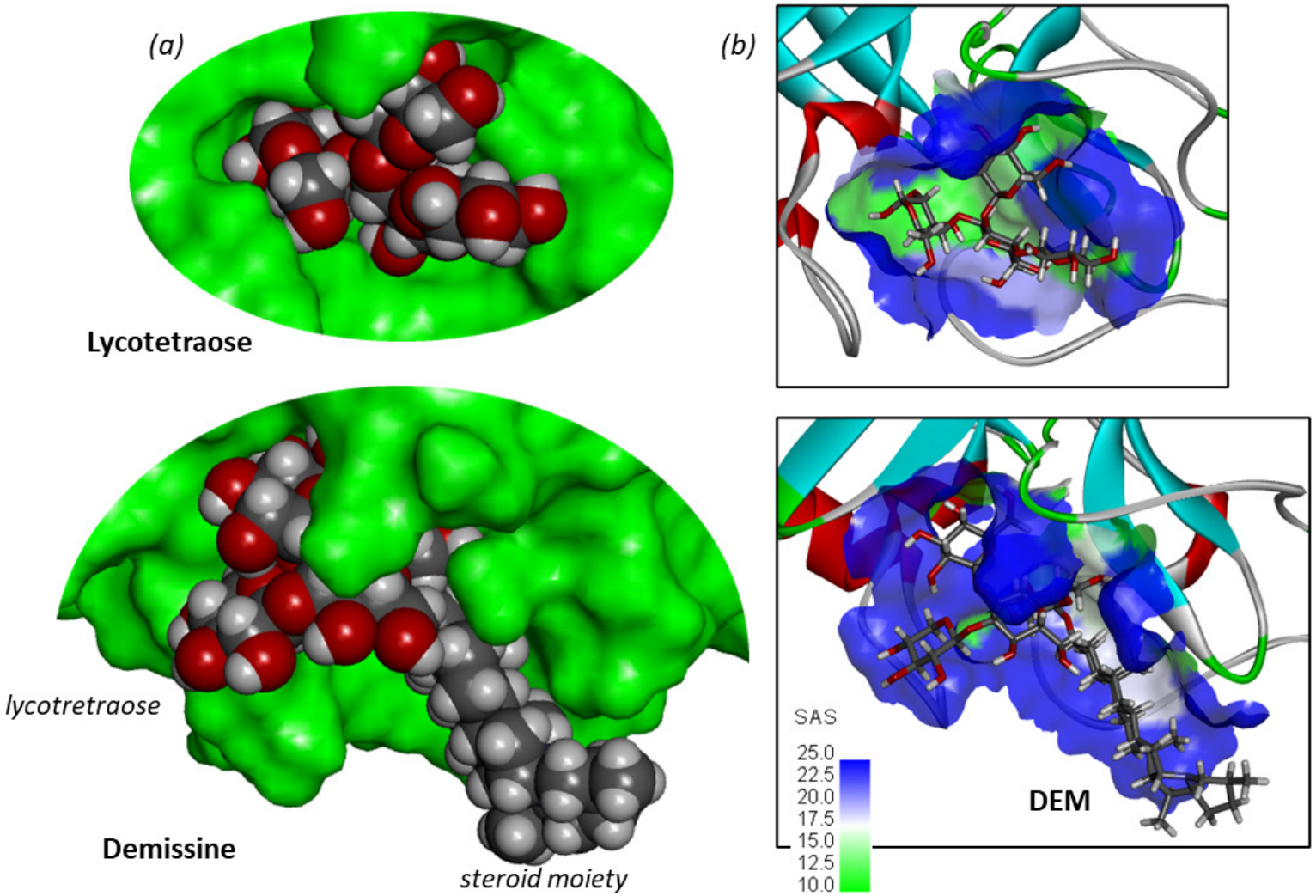

Fig. 4. Molecular models of lycotetraose and demissine bound to 3CL^{pro} of PEDV. Panel (*a*) shows the ligand bound to the active site cavity of the protein (green). Panel (*b*) shows a detailed view of the compound in the binding site, with the solvent-accessible surface (SAS) defined. The lycotetraose unit inserts well into the active site, whereas the steroid unit of desmissine (DEM) protrudes essentially toward the outside of the site.

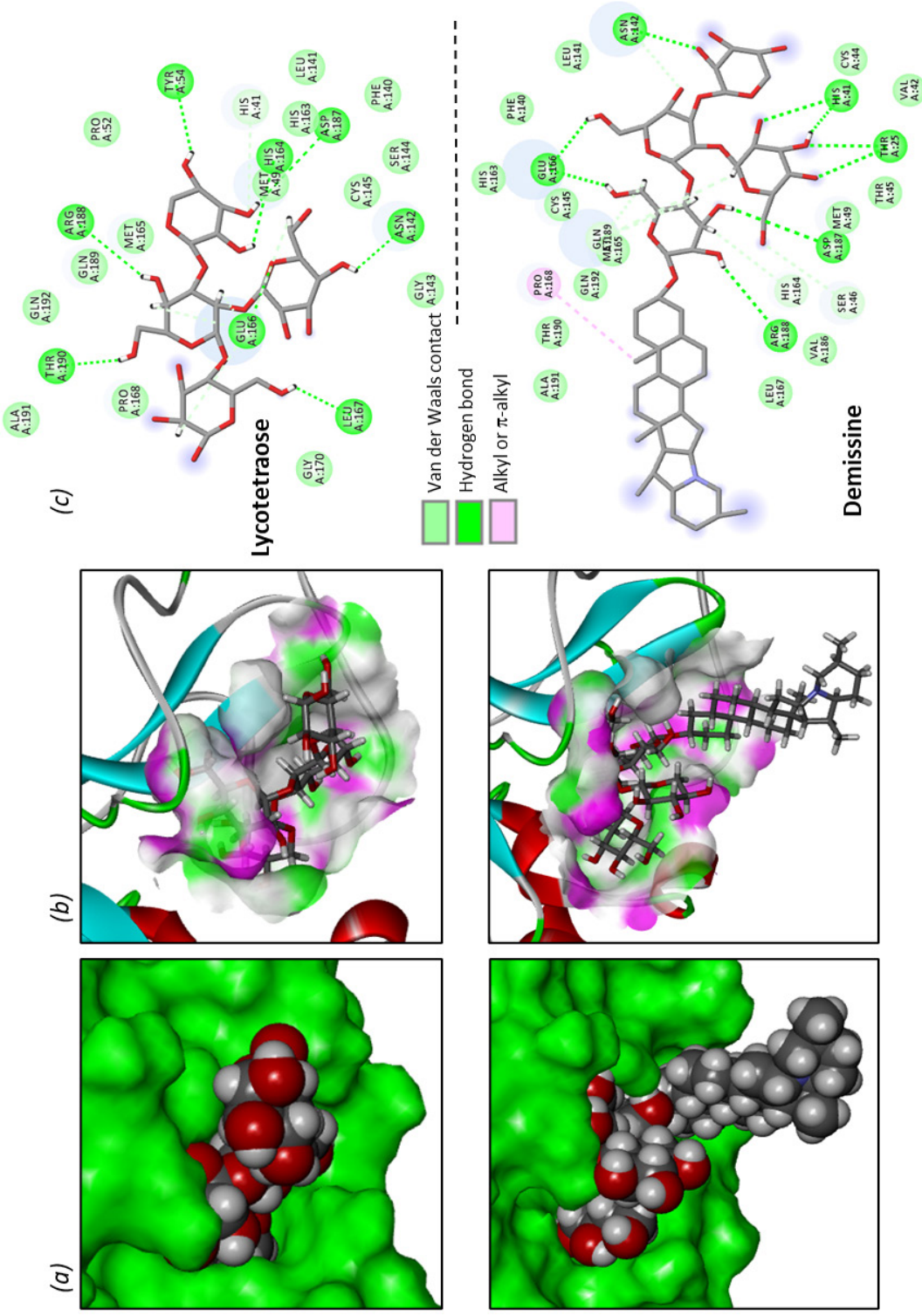

Fig. 5. Molecular models of lycotetraose and demissine bound to M^{pro} of SARS-CoV-2. Panel (*a*) shows the ligand bound to the active site cavity of the protein (green). Panel (*b*) shows a detailed view of the compound in the binding site, with the H-bond donor and acceptor groups available within the cavity. Panels (*c*) refers to the binding map contacts for each ligand, with the indicated color code.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare that there is no conflict of interest.

References

Bailly, C., 2021. The steroidal alkaloids alpha-tomatine and tomatidine: Panorama of their mode of action and pharmacological properties. Steroids 176, 108933.

Carere, J., Benfield, A.H., Ollivier, M., Liu, C.J., Kazan, K., Gardiner, D.M., 2017. A tomatinase-like enzyme acts as a virulence factor in the wheat pathogen *Fusarium graminearum*. Fungal Genet. Biol. 100, 33-41.

Choi, S.H., Ahn, J.B., Kozukue, N., Kim, H.J., Nishitani, Y., Zhang, L., Mizuno, M., Levin, C.E., Friedman, M., 2012. Structure-activity relationships of alpha-, beta(1)-, gamma-, and delta-tomatine and tomatidine against human breast (MDA-MB-231), gastric (KATO-III), and prostate (PC3) cancer cells. J. Agric. Food Chem. 60, 3891-3899.

Diosa-Toro M., Troost B., van de Pol D., Heberle A.M., Urcuqui-Inchima S., Thedieck K., Smit J.M., 2019. Tomatidine, a novel antiviral compound towards dengue virus. Antiviral Res. 161, 90-99.

Douangamath, A., Fearon, D., Gehrtz, P., Krojer, T., Lukacik, P., Owen, C.D., Resnick, E., Strain-Damerell, C., Aimon, A., Ábrányi-Balogh, P., Brandão-Neto, J., Carbery, A., Davison, G., Dias, A., Downes, T.D., Dunnett, L., Fairhead, M., Firth, J.D., Jones, S.P., Keeley, A., Keserü, G.M., Klein, H.F., Martin, M.P., Noble, M.E.M., O'Brien, P., Powell, A., Reddi, R.N., Skyner, R., Snee, M., Waring, M.J., Wild, C., London, N., von Delft, F., Walsh, M.A., 2020. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. Nat. Commun. 11, 5047.

Friedman, M., Levin, C.E., Lee, S.U., Kim, H.J., Lee, I.S., Byun, J.O., Kozukue, N., 2009. Tomatine-containing green tomato extracts inhibit growth of human breast, colon, liver, and stomach cancer cells. J. Agric. Food Chem. 57, 5727-5733.

Friedman, M., 2013. Anticarcinogenic, cardioprotective, and other health benefits of tomato compounds lycopene, alpha-tomatine, and tomatidine in pure form and in fresh and processed tomatoes. J. Agric. Food Chem. 61, 9534-9550.

Friedman, M., 2015. Chemistry and anticarcinogenic mechanisms of glycoalkaloids produced by eggplants, potatoes, and tomatoes. J. Agric. Food Chem. 63, 3323-3337.

Gao, M., Li, H., Ye, C., Chen, K., Jiang, H., Yu, K., 2021. Glycan epitopes and potential glycoside antagonists of DC-SIGN involved in COVID-19: *in silico* study. Biomolecules 11, 1586.

Homans, S.W., 1990. A molecular mechanical force field for the conformational analysis of oligosaccharides: comparison of theoretical and crystal structures of Man alpha 1-3Man beta 1-4GlcNAc. Biochemistry 29, 9110-9118.

Ikeda, T., Tsumagari, H., Honbu, T., Nohara, T., 2003. Cytotoxic activity of steroidal glycosides from solanum plants. Biol. Pharm. Bull. 26, 1198-201.

Ikeda, T., Yamauchi, K., Nakano, D., Nakanishi, K., Miyashita, H., Ito, S.I., Nohara, T., 2006. Chemical trans-glycosylation of bioactive glycolinkage: synthesis of an α-lycotetraosyl cholesterol. Tetrahedron Lett. 47, 4355-4359.

Jared, J.J., Murungi, L.K., Wesonga, J., Torto, B., 2016. Steroidal glycoalkaloids: chemical defence of edible African nightshades against the tomato red spider mite, *Tetranychus evansi* (Acari: Tetranychidae). Pest Manag. Sci. 72, 828-836.

Jones, G., Willett, P., Glen, R.C., Leach, A.R., Taylor, R., 1997. Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol. 267, 727-748.

Jones, N.A., Nepogodiev, S.A., Field, R.A., 2015. Efficient synthesis of methyl lycotetraoside, the tetrasaccharide constituent of the tomato defence glycoalkaloid α-tomatine. Org. Biomol. Chem. 3, 3201-3206.

Jorgensen, W.L., Tirado-Rives, J., 1996. Monte Carlo versus molecular dynamics for conformational sampling. J. Phys. Chem. 100, 14508-14513.

Jorgensen, W.L., Tirado-Rives, J., 2005. Molecular modeling of organic and biomolecular systems using BOSS and MCPRO. J. Comput. Chem. 26, 1689-1700.

Jorgensen, W.L., Ulmschneider, JP, Tirado-Rives, J., 2004. Free energies of hydration from a generalized Born model and an ALL-atom force field. J. Phys. Chem. B 108, 16264-16270.

Kozukue, N., Yoon, K.S., Byun, G.I., Misoo, S., Levin, C.E., Friedman, M., 2008. Distribution of glycoalkaloids in potato tubers of 59 accessions of two wild and five cultivated Solanum species. J. Agric. Food Chem. 56, 11920-11928.

Kuroda, M., Ori, K., Takayama, H., Sakagami, H., Mimaki, Y., 2015. Karataviosides G-K, five new bisdesmosidic steroidal glycosides from the bulbs of *Allium karataviense*. Steroids 93, 96-104.

Lagant, P., Nolde, D., Stote, R., Vergoten, G., Karplus, M., 2004. Increasing normal modes analysis accuracy: The SPASIBA spectroscopic force field introduced into the CHARMM program. J. Phys. Chem. A 108, 4019-4029.

Liu, J., Kanetake, S., Wu, Y.H., Tam, C., Cheng, L.W., Land, K.M., Friedman, M., 2016. Antiprotozoal effects of the tomato tetrasaccharide glycoalkaloid tomatine and the aglycone tomatidine on mucosal trichomonads. J. Agric. Food Chem. 64, 8806-8810.

Manabe, H., Fujiwara, Y., Ikeda, T., Ono, M., Murakami, K., Zhou, J.R., Yokomizo, K., Nohara, T., 2013. Saponins esculeosides B-1 and B-2 in Italian canned tomatoes. Chem. Pharm. Bull. (Tokyo) 61, 764-767.

Meziane-Tani, M., Lagant, P., Semmoud, A., Vergoten, G., 2006. The SPASIBA force field for chondroitin sulfate: vibrational analysis of D-glucuronic and N-acetyl-D-galactosamine 4-sulfate sodium salts. J. Phys. Chem. A 110, 11359-11370.

Nath, L.R., Gorantla, J.N., Thulasidasan, A.K., Vijayakurup, V., Shah, S., Anwer, S., Joseph, S.M., Antony, J., Veena, K.S., Sundaram, S., Marelli, U.K., Lankalapalli, R.S., Anto, R.J., 2016. Evaluation of uttroside B, a saponin from *Solanum nigrum* Linn, as a promising chemotherapeutic agent against hepatocellular carcinoma. Sci. Rep. 6,

36318.
Nohara, T., Ono, M., Ikeda, T., Fujiwara, Y., El-Aasr, M., 2010. The tomato saponin, esculeoside A. J. Nat. Prod. 73, 1734-1741.
Ökmen, B., Etalo, D.W., Joosten, M.H.A.J., Bouwmeester, H.J., de Vos, R.C.H., Collemare, J., de Wit, P.J.G.M., 2013. Detoxification of α-tomatine by *Cladosporium fulvum* is required for full virulence on tomato. New Phytol. 198, 1203-1214.
Osman, S.F., Herb, S.F., Fitzpatrick, T.J., Sinden, S.L., 1976. Commersonine, a new glycoalkaloid from two *Solanum* species. Phytochemistry 15, 1065-1067.
Sandrock, RW, Vanetten, HD., 1998. Fungal sensitivity to and enzymatic degradation of the phytoanticipin alpha-tomatine. Phytopathology 88, 137-143.
Takeo, K., Nakaji, T., Shinmitsu, K., 1984. Synthesis of lycotetraose. Carbohydrate Res. 133, 275-287.
Troost, B., Mulder, L.M., Diosa-Toro, M., van de Pol, D., Rodenhuis-Zybert, I.A., Smit, J.M., 2020. Tomatidine, a natural steroidal alkaloid shows antiviral activity towards chikungunya virus *in vitro*. Sci. Rep. 10, 6364.
Troost-Kind, B., van Hemert, M.J., van de Pol, D., van der Ende-Metselaar, H., Merits, A., Borggrewe, M., Rodenhuis-Zybert, I.A., Smit, J.M., 2021. Tomatidine reduces chikungunya virus progeny release by controlling viral protein expression. PLoS Negl. Trop. Dis. 15, e0009916.
Vázquez, A., González, G., Ferreira, F., Moyna, P., Kenne, L., 1997. Glycoalkaloids of *Solanum commersonii* Dun. ex Poir. Euphytica 95, 195-201.
Vergoten, G., Mazur, I., Lagant, P., Michalski, J.C., Zanetta, J.P., 2003. The SPASIBA force field as an essential tool for studying the structure and dynamics of saccharides. Biochimie 85, 65-73.
Wang, P., Bai, J., Liu, X., Wang, M., Wang, X., Jiang, P., 2020. Tomatidine inhibits porcine epidemic diarrhea virus replication by targeting 3CL protease. Vet. Res. 51,136.
Wang, Y.C., Yang, W.H., Yang, C.S., Hou, M.H., Tsai, C.L., Chou, Y.Z., Hung, M.C., Chen, Y., 2020. Structural basis of SARS-CoV-2 main protease inhibition by a broad-spectrum anti-coronaviral drug. Am. J. Cancer Res. 10, 2535-2545.
Woods, K., Hamilton, C.J., Field, R.A., 2004. Enzymatic liberation of lycotetraose from the *Solanum* glycoalkaloid alpha-tomatine. Carbohydr. Res. 339, 2325-2328.
Wu, M., Han, G., Meng, C., Wang, Z., Liu, Y., Wang, Q., 2014. Design, synthesis, and anti-tobacco mosaic virus (TMV) activity of glycoconjugates of phenanthroindolizidines alkaloids. Mol. Divers. 18, 25-37.
Yahara, S., Uda, N., Yoshio, E., Yae, E., 2004. Steroidal alkaloid glycosides from tomato (*Lycopersicon esculentum*). J. Nat. Prod. 67, 500-502.
Ye, G., Deng, F., Shen, Z., Luo, R., Zhao, L., Xiao, S., Fu, Z.F., Peng, G., 2016. Structural basis for the dimerization and substrate recognition specificity of porcine epidemic diarrhea virus 3C-like protease. Virology 494, 225-235.
Zafar, A., Reynisson, J., 2016. Hydration free energy as a molecular descriptor in drug design: A feasibility study. Mol. Inform. 35, 207-214.