***In silico* design of glyco-D,L-peptide antiviral molecules.**

Luigi Agostini^{1*}, Stefano Morosetti^{2‡}

¹School of Chemistry and Materials Engineering, Huizhou University, Huizhou 516007, China

*present address: Prince Mohammad Bin Fahd University, Al-Khobar 31952, Kingdom of Saudi Arabia

² Dipartimento di Chimica, Università di Roma La Sapienza, Piazzale A. Moro 5, 00185 Roma, Italy

‡Corresponding author. E-mail address: stefano.morosetti@uniroma1.it

Abstract

Background: most licensed antiviral drugs are nucleoside analogs. A recent research focuses on blocking a virus from entering the cells in the viral cell adsorption/entry stage. In this entry mechanism the glycans present on the viral surface play a fundamental role. Homochiral L-peptides acting this fusion mechanism have shown some inhibition of viral infection. Peptides with regularly alternating enantiomeric sequence (L,D-peptides) can assume structures, which are not accessible to the corresponding homochiral molecules. Further, L,D-peptides are less sensitive to the enzymatic digestion.

Aim: *in silico* design a L,D-peptide with a high affinity for the viral surface glycans, and consequently able to interfere with its fusion mechanism.

Methods: a 3 α ,6 α -Mannopentaose (3-6MP) molecule was used to simulate a viral surface glycan. Molecular Dynamics (MD) simulations of 3-6MP and D,L-peptide in water are performed using the force field AMBER12-GLYCAM06i. The binding constant was evaluated from trajectories. The D,L-peptide molecule was modified over the sequence, the length, the terminals and finally glycosylated to attain a very high binding constant value for 3-6MP.

In addition, the specific interaction between T lymphocyte CD4 glycoprotein and HIV envelope gp120 glycoprotein was studied through MD simulations between a D,L-peptide, bounded to a typical CD4 glycan, and a highly conserved HIV gp120 glycan.

Results: in the case of interaction with 3-6MP molecule, the very effective molecule obtained was H-D-Trp-L-Pro-D-Asn-L-Pro-D-Trp-L-Pro-D-Asn-L-Pro-OH where the Asn residues in position 3 and 7 are glycosylated with α -D-Mannopyranosyl-(1 $\rightarrow$ 3)-[α -D-mannopyranosyl-(1 $\rightarrow$ 6)]- α -D-mannopyranosyl-(1 $\rightarrow$ 4)-N-acetyl-beta-D-glucopyranosyl-(1 $\rightarrow$ 4)-N-acetyl-beta-D-glucopyranosyl-1-OH.

In the case of interaction with HIV envelope, the very effective molecule obtained, able to antagonize the CD4 glycoprotein, was H-D-Trp-L-Pro-D-Asn-L-Pro-D-Trp-L-Pro-D-Asn-L-Pro-OH where the Asn residue in position 3 is glycosylated with α -D-galactopyranosyl-(1 $\rightarrow$ 4)- N-acetyl-beta-D-glucopyranosyl-(1 $\rightarrow$ 2)- α -D-Mannopyranosyl-(1 $\rightarrow$ 3)- [α -D-galactopyranosyl-(1 $\rightarrow$ 4)- N-acetyl-beta-D-glucopyranosyl-(1 $\rightarrow$ 2)- α -D-Mannopyranosyl-(1 $\rightarrow$ 6)]- beta-D-mannopyranosyl-(1 $\rightarrow$ 4)-N-acetyl-beta-D-glucopyranosyl-(1 $\rightarrow$ 4)-N-acetyl-beta-D-glucopyranosyl-1-OH

Conclusion: these glycosylated D,L-peptide molecules are very promising representatives of a new class of antiviral agents.

Keywords

Glyco-D,L-peptides; antiviral agents; D,L-peptides; L,D-peptides; AMBER; GLYCAM; GROMOS ; *in silico* design; inhibition of viral fusion mechanism; HIV gp120; CD4; HIV antiviral

Abbreviations

3-6MP: α -D-Mannopyranosyl-(1 $\rightarrow$ 3)-[α -D-mannopyranosyl-(1 $\rightarrow$ 3)-[α -D-mannopyranosyl-(1 $\rightarrow$ 6)]- α -D-mannopyranosyl-(1 $\rightarrow$ 6)]- β -D-mannopyranose

9glyco: α -D-galactopyranosyl-(1 $\rightarrow$ 4)- N-acetyl- β -D-glucopyranosyl-(1 $\rightarrow$ 2)- α -D-Mannopyranosyl-(1 $\rightarrow$ 3)- [α -D-galactopyranosyl-(1 $\rightarrow$ 4)- N-acetyl- β -D-glucopyranosyl-(1 $\rightarrow$ 2)- α -D-Mannopyranosyl-(1 $\rightarrow$ 6)]- β -D-mannopyranosyl-(1 $\rightarrow$ 4)-N-acetyl- β -D-glucopyranosyl-(1 $\rightarrow$ 4)-N-acetyl- β -D-glucopyranosyl-1-OH

CPK: atomic coloring scheme

FOR: formulated

GlcNAc: N-acetyl-glucosamine

GPU: Graphics Processing Unit

HB : hydrogen bond

MD: molecular dynamics

MW: molecular weight

N262glyco: α -D-Mannopyranosyl-(1 $\rightarrow$ 2)- α -D-Mannopyranosyl-(1 $\rightarrow$ 2)- α -D-Mannopyranosyl-(1 $\rightarrow$ 3)-[α -D-Mannopyranosyl-(1 $\rightarrow$ 6)-[α -D-Mannopyranosyl-(1 $\rightarrow$ 3)]- α -D-Mannopyranosyl-(1 $\rightarrow$ 6)]- β -D-mannopyranosyl-(1 $\rightarrow$ 4)-N-acetyl- β -D-glucopyranosyl-(1 $\rightarrow$ 4)-N-acetyl- β -D-glucopyranosyl-1-OH

penta: α -D-Mannopyranosyl-(1 $\rightarrow$ 3)-[α -D-mannopyranosyl-(1 $\rightarrow$ 6)]- β -D-mannopyranosyl-(1 $\rightarrow$ 4)-N-acetyl- β -D-glucopyranosyl-(1 $\rightarrow$ 4)-N-acetyl- β -D-glucopyranosyl-1-OH

PME: Particle Mesh Ewald

SNFG: Symbol Nomenclature for Glycans

TrpProAsn9glyco: H-D-Trp-L-Pro-D-Asn(-9glyco)-L-Pro-D-Trp-L-Pro-D-Asn-L-Pro-OH

TrpProAsnPenta: H-D-Trp-L-Pro-D-Asn(-Penta)-L-Pro-D-Trp-L-Pro-D-Asn(-Penta)-L-Pro-OH

Introduction

There is a continuous effort in the search for new antiviral agents, which aims to develop new and more effective drugs, capable of targeting an increasing number of diseases while, at the same time, limiting their side effects. The approved drugs have different mechanism of actions during the viral life cycle¹. A few of them are entry inhibitors. One of these (Enfuvirtide) is a polypeptide with 36 units. Many polypeptides were tested as antiviral^{2,3}. The viral fusion stage is often mediated by the glycoproteins present on its envelope⁴.

The oligo-L,D-peptides present structures which are not accessible to the corresponding homochiral peptides. In particular, by assuming local β conformations, they can realize helices of different diameter^{5,6}. Another property of these molecules with the peptide bond connecting D,L and L,D dimers is the decreased susceptibility to digestion by proteolytic enzymes compared with the LL dimers⁷. This results in their increased average life time in the organism.

The initial idea of our research was to find a short L,D-oligopeptide capable of binding the sugar moiety of the viral surface glycoproteins, so interfering with the fusion stage of the infection. For this purpose we performed different *in silico* Molecular Dynamics simulations (MD) on a system consisting of a L,D-peptide molecule and a oligosaccharide.

The chosen oligosaccharide is 3 α ,6 α -Mannopentaose (3-6MP) (α -D-Mannopyranosyl-(1 $\rightarrow$ 3)-[α -D-mannopyranosyl-(1 $\rightarrow$ 3)-[α -D-mannopyranosyl-(1 $\rightarrow$ 6)]- α -D-mannopyranosyl-(1-

>6)]-beta-D-mannopyranose) (fig.1). This is a common core structure for a large number of glycans present in glycoproteins⁸.

Fig.1 Right: the 3-6MP molecule representing the glycan component of the glycoproteins. Left: The schematic representation of the molecule is enclosed in a rectangle. Outside the square the link with the protein portion is shown.

The D,L-peptide molecule initially chosen for the simulation is H-(D-Phe-L-Lys)₄-OH. The octapeptide can assume a β -helix conformation with 6.2 monomer per turn⁶. Our initial guess was that the helix would turn around the oligosaccharide (fig.2).

Fig.2. Left: H-(D-Phe-L-Lys)₄-OH in β -helix conformation with 6.2 monomer per turn. Middle: the helix turns round the oligosaccharide. Right: same as the middle, but with different colors for the oligosaccharide, Phe residues, Lys residues.

The rationale for the octapeptide chosen sequence was its solubility, obtained through the introduction of the charged Lys residues, and the stabilizing van der Waals interactions between Phe residues and the oligosaccharide.

The association of the two molecules was monitored performing MD simulations, through visual inspection of the trajectory by means of VMD 1.9.1⁹ program and following the molecules' center of gravity distance as a function of time.

Our simulations, however, showed that this system (fig.2) is unstable. The peptide molecule moves away from the oligosaccharide and evolves into a globular structure, where β -turns are present (fig. 3). This happens because the peptide sequences L,D and D,L can easily form β -turn conformations¹⁰. Therefore, the resulting association constant has a low value.

Fig. 3. Globular conformation, with hydrogen bond highlighted by dashed line in the β -turn. Left: CPK colors, right: peptide groups highlighted by different colors. The conformation is taken from a MD trajectory.

Given this result, we decided to modify the D,L-peptide molecule over the sequence, the length, the terminals and finally glycosylated in an attempt to attain a very high binding constant value. It is worth mentioning that glycopeptide antibiotic molecules have shown antiviral activity^{11,12}. Detailed Methods and Results are given below.

Methods

Force field : oligosaccharides are present in the system with the peptide molecule, hence the choice for the force field is AMBER12sb¹³ with GLYCAM06j^{14,15}.

Molecular Dynamics simulations : All-atom MD simulations were performed in water by using GROMACS 4.5.6 software package¹⁶. AMBER12sb and GLYCAM06j was obtained from the pertinent official sites^{17,18}. After minimization and 150 ps equilibration steps on a *NVT* ensemble and subsequent 150 ps equilibration steps on *NPT* ensemble, a 200 ns MD simulation was performed in a periodic box, with 1 Na⁺ and 1 Cl⁻ to simulate the ionic strength.

The following settings were used: An integration step of 2 fs with periodic boundary conditions at constant volume and temperature (T=310 K). Leap-frog integration scheme of the equations of motion. The default LINCS algorithm to constrain bonds involving H atoms. The Verlet cutoff scheme for neighbor searching, since parallel computing on GPUs was performed. Particle Mesh Ewald (PME) algorithm for long range electrostatics.

Three different simulations were performed for each peptide molecule, differing in box volume and/or relative position of peptide and 3-6MP molecules and/or initial peptide conformation.

Carbohydrate coordinates, topology and charges : they were obtained by the Carbohydrated Builder tool present in the GLYCAM-web site glycam.org. The molecule can be submitted in condensed Glycam notation¹⁹ or can be built specifying the monosaccharides and their linkages.

Files with the extensions rst7 for coordinates and parm7 for topology were obtained, in AMBER style.

The program glycam2gmx.pl^{20,21} can be used to obtain files in GROMOS stile. The command line is perl ./glycam2gmx.pl -prmtop *.parm7 -crd *.rst7 -outname *, which returns *.gro for coordinates and *.top for topology (* stands for the chosen names).

These files must be merged with the corresponding peptide.gro and peptide.top files obtained by the program pdb2gmx for the peptide molecule.

One coordinates file was obtained merging the carbohydrate coordinates in the peptide.gro file. They can be visualized by VMD program and translated through the program editconf (present in the GROMACS suite) if necessary.

One topology file was obtained by renaming the carbohydrate file *.top to *.itp and inserting the line #include “*.itp” in the peptide.top file. Duplicated records, atom types definitions and water topology must be deleted.

Glycopeptide coordinates, topology and charges : they were obtained by the Glycoprotein Builder tool present in the GLYCAM-web site glycam.org. Step 1: the coordinates of the peptide under glycosylation are submitted in pdb²² format, without hydrogens and records CONNECT. Step 2: the ends are fixed. If the peptide is in zwitterionic form, CONTINUE must be chosen. Step 3: the glycan derivative is chosen by the interactive carbohydrate builder tool; the saccharide chain is selected, ending with aglycon –OH; this end will be linked to the peptide; after aglycon, branches can be introduced; term with DONE command; add glycan to glycosylation sites: N-linking and O-linking biologically likely sites are shown; SHOW ALL gives all sites with the solvent accessible surface area; select and continue; download current structure. Step 4: the structure is minimized. Step 5: download of the rst7 and parm7 files.

The program glycam2gmx.pl can be used to obtain files in GROMOS stile. In simulations of the glycopeptide with a carbohydrate molecule, its coordinates and topology will be merged in glycopeptide files by following the same procedure as before.

Association constant estimation : the distance between the molecules’ centers of gravity was calculated by the command g_dist for each MD simulation. When the intermolecular distance was lower than a certain value dependent upon the molecular weights (MW) of the molecules (in our case 1.5 nm) these molecules were considered associated. From the total time of the MD simulation (t_{tot}) is possible to extract a time when the molecules are considered associated (t_{ass}). Through the introduction of $f = t_{ass}/t_{tot}$, the fraction of association time, the association constant K_{ass} can be roughly calculated as:

$$K_{ass} = \frac{f * c}{(1 - f) * c * (1 - f) * c} = \frac{f}{(1 - f)^2} * c^{-1}$$

where the molar fraction is considered numerically equal to the time fraction, applying the ergodic postulate. The trajectories are of limited extension, and hence the approximation in constant evaluation.

The association constants reported were averaged over all trajectories calculated for the same molecular system. The error was calculated as half the range of values obtained.

Results

The peptide or glycopeptide sequences studied are given below together with their K_{ass} (in mol⁻¹) with 3-6MP molecule.

H-(D-Phe-L-Lys)₄-OH $K_{ass} = (1.65 \pm 0.31) * 10^2$

The distance vs time between the peptide and 3-6MP molecules in a single trajectory, is shown in fig. 4.

Fig.4 . Distance (nm) between H-(D-Phe-L-Lys)₄-OH and 3-6MP molecules, reported every 200 ps; the total time is 200 ns.

The effect of changing the terminal and amino acid sequence was explored with the following MD simulations, where FOR stands for the formulated NH₂, and Hyp is the hydroxyproline amino acid:

FOR-(D-Trp-L-Arg)₄-OH $K_{ass} = (2.01 \pm 0.89) * 10^2$

FOR-(D-Trp-L-Hyp)₄-OH $K_{ass} = (6.55 \pm 1.43) * 10^2$

The presence of the Hyp seems to have a moderate effect on increasing the binding constant.

In order to achieve the expected association constant, we introduced the following changes: 1) terminal ACE and NME were introduced to increase the number of intramolecular hydrogen bonds (HB) so as to stabilize the peptide helix conformation; 2) Pro or Hyp were introduced to impart conformational stiffness to the helix; 3) The D-aminoacid was replaced to try to introduce a range of hydrophobic and hydrogen bonding interactions.

All the following peptides behave like the peptide H-(D-Phe-L-Lys)₄-OH in spite of the changes introduced, as evidenced by the trajectory displayed and the K_{ass} obtained.

ACE-(D-Arg-L-Hyp)₄-NME $K_{ass} = (1.24 \pm 0.47) * 10^3$

ACE-(D-Trp-L-Hyp)₄-NME $K_{ass} = (1.36 \pm 0.88) * 10^3$

ACE-(D-Asn-L-Hyp)₄-NME $K_{ass} = (2.00 \pm 0.07) * 10^2$

The effect of the peptide's length was also studied through the following sequences:

Finally, a peptide with an overall negative charge, instead of positive, was studied:

All sequences shown do not exhibit a K_{ass} much higher than the initial peptide. The introduction of an amino acid D-Trp, D-Arg or the rigid conformation amino acid L-Hyp, increased the K_{ass} by a little less than an order of magnitude.

Peptide glycosylation was introduced. The further interactions with 3-6MP molecule, were expected to increase the K_{ass} .

GlcNAc stands for N-acetyl-glucosamine.

In consideration of these disappointing results, a more complex glycosylation was tried.

The glycoside attached to the peptide was: alpha-D-Mannopyranosyl-(1->3)-[alpha-D-mannopyranosyl-(1->6)]-beta-D-mannopyranosyl-(1->4)-N-acetyl-beta-D-glucopyranosyl-(1->4)-N-acetyl-beta-D-glucopyranosyl-1-OH (abbreviated as Penta). See fig. 5. It is the common motif in the high mannose and complex N-linked glycan chains of the glycoproteins.

Fig.5 The glycoside Penta attached to peptide. Upper left: scheme. Upper right: structure. Bottom: SNFG representation²³.

The molecule submitted to MD simulation was H-D-Trp-L-Pro-D-Asn(-Penta)-L-Pro-D-Trp-L-Pro-D-Asn(-Penta)-L-Pro-OH (abbreviated as TrpProAsnPenta). Where: Pro is present for its rigidity and Trp was introduced because of the higher K_{ass} obtained when present in the previously studied molecules. The molecule is shown in fig. 6,

Fig.6. H-D-Trp-L-Pro-D-Asn(-Penta)-L-Pro-D-Trp-L-Pro-D-Asn(-Penta)-L-Pro-OH. Up: scheme. Down: structure; the oligopeptide is shown in elongated structure and the single aminoacids in different colors for clarity.

In this case six different MD simulations were performed. The results are in all cases a stable bending between TrpProAsnPenta and 3-6MP after their initial matching into the simulation box. There can be different arrangements in the bending, but without dissociation. This demonstrates a very high K_{ass} (not measurable with the aforementioned method).

The results of a MD simulation are reported as the inter-molecular distance vs time (fig.7) and in (fig.8) is reported the structure of the association complex obtained.

Fig.7 . Distance (nm) between TrpProAsnPenta and 3-6MP molecules, reported every 200 ps; the total time is 200 ns.

Fig.8. 85 ns time structure, which represents the average of the association during the simulation; two representations are shown, one with different colors for different chemical units, the other with different colors for the 2 molecules, to locate them better.

Encouraged by these positive findings, we developed a simulation aimed to a specific case. The idea is to use for glycosylation the same cellular glycan chain preferentially exploited by the virus for the adhesion in the first stage of fusion, thus creating a molecular mimicry against the virus. Considering the specific case of HIV virus, this means using a glycan chain of the CD4 glycoprotein, which is involved in the interaction with the viral surface glycoproteins in the early stage of recognition⁴. Numerous different glycan chains are reported for this glycoprotein²⁴, with respectively: one, two or three branches. We chose the structure alpha-D-galactopyranosyl-(1->4)-N-acetyl-beta-D-glucopyranosyl-(1->2)- alpha-D-Mannopyranosyl-(1->3)- [alpha-D-galactopyranosyl-(1->4)- N-acetyl-beta-D-glucopyranosyl-(1->2)- alpha-D-Mannopyranosyl-(1->6)]- beta-D-mannopyranosyl-(1->4)-N-acetyl-beta-D-glucopyranosyl-(1->4)-N-acetyl-beta-D-glucopyranosyl-1-OH (abbreviated as 9glyco) reported in fig. 9 as our first representative, because it has the most recurrent monosaccharides for each given position, in the most common two branched motif.

Consequently, the glycopeptide used in the simulations was H-D-Trp-L-Pro-D-Asn(-9glyco)-L-Pro-D-Trp-L-Pro-D-Asn-L-Pro-OH (abbreviated as TrpProAsn9glyco).

Fig.9 The glycoside 9glyco attached to peptide. Upper left: scheme. Upper right: structure. Bottom: SNFG representation²³.

The glycoprotein CD4 interacts with the HIV glycoprotein gp120, which can relate through the exposure of different glycans⁴. The glycan at position Asn262 in HIV gp120 is chosen for the interaction with TrpProAsn9glyco, because it has the best resolved structure in a recent quaternary complex with CD4, and its presence is essential for the fusion mechanism²⁵. The high mannose glycan used in the simulations is alpha-D-Mannopyranosyl-(1->2)-alpha-D-Mannopyranosyl-(1->2)-alpha-D-Mannopyranosyl-(1->3)-[alpha-D-Mannopyranosyl-(1->6)-[alpha-D-Mannopyranosyl-(1->3)]- alpha-D-Mannopyranosyl-(1->6)]-beta-D-mannopyranosyl-(1->4)-N-acetyl-beta-D-glucopyranosyl-(1->4)-N-acetyl-beta-D-glucopyranosyl-1-OH with the OH replacing the connection to Asn262 position. It is here abbreviated as N262glyco and reported in fig.10.

Fig.10 The glycoside N262glyco. Upper left: scheme. Upper right: structure. Bottom: SNFG representation²³.

MD simulations were performed over TrpProAsn9glyco and N262glyco in water. The results are in all cases a stable bending between TrpProAsn9glyco and N262glyco after their initial matching into the simulation box. Similar to the previous case, here too, we found a very high K_{ass} (not measurable with the aforementioned method).

The results of one MD simulation are reported (fig.11) showing the structures of the association complex obtained:

Fig.11. 73.8 ns time structure, which represents the average of the association during the simulation; two representations are shown, one with different colors for different chemical units, the other with different colors for the 2 molecules, to locate them better.

It is worth mentioning that the molecules' relative positions, in the 'association state', are consistent with a hypothetical interaction between CD4 and gp120 glycoproteins, as indicated by the opposite positions of the glycosides binding points, which are the peptide and the GlcNAc terminals, respectively. These findings are illustrated in Fig. 12.

Fig.12. Interaction between the glycans 9glyco and N262glyco. Left: in vivo; the glycoproteins involved are indicated. Center: model interaction; the binding of 9glyco to D,L peptide is indicated. Right: the representative structure of the association obtained through MD simulation; red: N262glyco, blu: D,L peptide 9glyco. The position of D,L peptide is highlighted. The arrow indicates the GlcNAc bonded to gp120 in vivo.

Conclusions

Our results showed that oligo-D,L-peptides can achieve a stable association with a glycan, only when glycosylated. Further, the glycosylation is effective only if the glycan side chains are suitably long.

Actually H-D-Asn-L-Pro-D-Asn(-GlcNAc)-L-Pro-D-Asn(-GlcNAc)-L-Pro-D-Asn(-GlcNAc)-L-Pro-OH which has three monomeric glycosylation sites, does not achieve a stable association with 3-6MP molecule.

H-D-Trp-L-Pro-D-Asn(-Penta)-L-Pro-D-Trp-L-Pro-D-Asn(-Penta)-L-Pro-OH realizes *in silico* the goal of a high bending with the molecule 3-6MP, which is the common core structure of a large number of glycans present in glycoproteins.

H-D-Trp-L-Pro-D-Asn(-9glyco)-L-Pro-D-Trp-L-Pro-D-Asn-L-Pro-OH realizes *in silico* the goal of a high bending with the molecule N262glyco, where 9glyco is a common glycosylation of CD4 and N262glyco a conserved glycan chain of gp120, so to realize a molecular mimicry against the HIV virus.

If our *in silico* studies will be confirmed by *in vivo* tests, this array of compounds will represent an entire new class of anti-viral drugs, with promising therapeutic potential.

A possible synthetic route for these molecules is presented in Appendix.

The great variability, achievable through modifications of the peptide sequence, length, endings and through its glycosylation, allows the possibility of fine tuning logP, solubility and interaction specificity.

Finally, it is possible to use for glycosylation the same cellular glycan chain preferentially exploited by the virus for the adhesion in the first stage of fusion, thus creating a molecular mimicry against the virus. In the case of HIV virus, we have used a glycosylic chain of CD4 glycoprotein⁴, involved in interaction with the surface glycoproteins of the virus in the early stage of recognition.

In future work, it will be extremely interesting to test less common glycosylations such as Arginine N-glycosylation²⁶ and Arginine N-rhamnosylation²⁷.

Appendix

Synthetic route

The molecules H-D-Trp-L-Pro-D-Asn(-Penta)-L-Pro-D-Trp-L-Pro-D-Asn(-Penta)-L-Pro-OH and H-D-Trp-L-Pro-D-Asn(-9glyco)-L-Pro-D-Trp-L-Pro-D-Asn-L-Pro-OH can be synthesized following these steps.

The peptide can be prepared manually by conventional solid phase chemistry on Fmoc-L-Pro-Wang resin according to published protocols²⁸. Fmoc-Xaa-OH will be used for couplings.

The Fmoc-Asparagine containing N-acetylglucosamine can be prepared^{29,30} or purchased.

The resulting glycosylated Asn can be reacted with oligosaccharides by using glycosyltransferase making the sugar chain elongation possible^{30,31}.

The test of the Arginine N-glycosylation and Arginine N-rhamnosylation can be obtained, using the protected α - and β -rhamnosylated arginine amino acids, following the procedures reported in 26 and 27 references.

References

- 1) De Clercq E, Li G. Approved antiviral drugs over the past 50 years. Clin Microbiol Rev. 2016; 29:695–747. <http://dx.doi.org/10.1128/CMR.00102-15>
- 2) Qureshi A, Thakur N, Himani T, Kumar M. AVPdb: a database of experimentally validated antiviral peptides targeting medically important viruses. Nucl. Acids Res. 2014; 42 (D1): D1147-D1153. <https://doi.org/10.1093/nar/gkt1191>.
- 3) <http://crdd.osdd.net/servers/avpdb/index.php>
- 4) White JM, Delos SE, Brecher M, Schornberg K. Structures and Mechanisms of Viral Membrane Fusion Proteins: Multiple Variations on a Common Theme. Crit Rev Biochem Mol Biol. 2008 ; 43(3): 189–219. <https://doi.org/10.1080/10409230802058320>
- 5) Novelli F, De Santis S, Morosetti S, Titubante M, Masci G, Scipioni A. Peptides with regularly alternating enantiomeric sequence: From ion channel models to bioinspired nanotechnological applications. Peptide Science. 2018; e24043. <https://doi.org/10.1002/pep2.24043>
- 6) De Santis P, Morosetti S, Rizzo R. Conformational Analysis of Regular Enantiomeric Sequences. Macromolecules. 1974; 7: 52. <https://doi.org/10.1021/ma60037a011>
- 7) Stadler RH, Lineback DR. Process-Induced Food Toxicants: Occurrence, Formation, Mitigation, and Health Risks, 2009. John Wiley & Sons (see chapter 6.2 Dietary Significance of Processing-Induced D-Amino Acids. M. Friedman)
- 8) Chemical Biology of Glycoproteins. 2017. Ed. Zhongping Tan, Lai-Xi Wang. Royal Society of Chemistry.
- 9) Humphrey W, Dalke A, Schulten K. VMD – Visual Molecular Dynamics. J. Molec. Graphics. 1996; 14.1: 33-38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)
- 10) De Santis S, Novelli F, Sciubba F, Casciardi S, Sennato S, Morosetti S, Scipioni A. Tunable length nanotubes from a self-assembling pH and thermosensitive linear L,D-peptide - polymer conjugate. J. Colloid Interface Sci. 2018; submitted
- 11) Balzarini J, Pannecouque C, De Clercq E, Pavlov AY, Printsevskaya SS, Miroshnikova OV, Reznikova MI, Preobrazhenskaya MN. Antiretroviral Activity of Semisynthetic Derivatives of Glycopeptide Antibiotics. J. Med. Chem. 2003; 46: 2755-2764. <https://doi.org/10.1021/jm0300882>

- 12) Colson P, Raoult D, Fighting viruses with antibiotics: an overlooked path. *International Journal of Antimicrobial Agents*. 2016; 48: 349–352.
<https://doi.org/10.1016/j.ijantimicag.2016.07.004>
- 13) Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci*. 2013; 3: 198-210.
<https://doi.org/10.1002%2Fwcms.1121>
- 14) Kirschner KN, Yongye AB, Tschampel SM, Daniels CR, Foley BL, Woods RJ. Glycam06. *J. Comput. Chem*. 2008; 29: 622-655. <https://doi.org/10.1002%2Fjcc.20820>
- 15) <http://glycam.org/docs/help/2014/07/02/which-protein-force-fields-are-compatible-with-glycam/>
- 16) Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput*. 2008; 4 (2): 435–447. doi:10.1021/ct700301q.
- 17) http://www.gromacs.org/Downloads/User_contributions/Force_fields
- 18) <http://glycam.org/docs/forcefield/all-parameters/>
- 19) <http://glycam.org/docs/custombuilders/condensed-notation/>
- 20) Wehle M, Vilotijevic I, Lipowsky R, Seeberger PH, Varon Silva D, Santer M. Mechanical Compressibility of the Glycosylphosphatidylinositol(GPI) Anchor Backbone by Independent Glycosidic Linkages. *J Am Chem Soc*. 2012;134(46):18964–18972.
<https://doi.org/10.1021/ja302803r>
- 21) <https://listserv.uga.edu>
- 22) <http://www.wwpdb.org/documentation/file-format-content/format33/v3.3.html>
- 23) Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, Stanley P, Hart G, Darvill A, Kinoshita T, Prestegard JJ, Schnaar RL, et al. Symbol Nomenclature for Graphical Representation of Glycans. *Glycobiology* 2015; 25:1323-1324.
<https://doi.org/10.1093/glycob/cwv091>
- 24) <https://glyconnect.expasy.org/browser/proteins/587>
- 25) Kong L, Wilson IA, Kwong PD. Crystal structure of a fully glycosylated HIV-1 gp120 core reveals a stabilizing role for the glycan at Asn262. *Proteins*. 2015; 83(3): 590–596.
<https://doi.org/10.1002/prot.24747>
- 26) Pan M, Li S, Li X , Shao F, Liu L, Hu HG. Synthesis of and Specific Antibody Generation for Glycopeptides with Arginine N-GlcNAcylation. *Angew. Chem. Int. Ed*. 2014; 53: 1 – 6. <https://doi.org/10.1002/anie.201407824> .
- 27) Wang S, Corcilius L, Sharp PP, Rajkovic A, Ibba M, Parker BL, Payne RJ. Synthesis of rhamnosylated arginine glycopeptides and determination of the glycosidic linkage in bacterial elongation factor P. *Chem. Sci*. 2017; 8: 2296. DOI:
<https://doi.org/10.1039/c6sc03847f> .
- 28) Atherton E, Sheppard RC. *Solid Phase Synthesis—A Practical Approach* (Oxford: IRL) 1989
- 29) Haneda K, Inazu T, Yamamoto K, Kumagai H, Nakahara Y, Kobata A. Transglycosylation of intact sialo complex-type oligosaccharides to the *N*-acetylglucosamine moieties of glycopeptides by *Mucor hiemalis* endo- β -*N*-acetylglucosaminidase. *Carbohydr. Res*. 1996; 292: 61. [https://doi.org/10.1016/S0008-6215\(96\)91025-3](https://doi.org/10.1016/S0008-6215(96)91025-3)
- 30) Yamamoto K. Chemo-enzymatic synthesis of bioactive glycopeptide using microbial endoglycosidase. *J Biosci Bioeng*. 2001;92(6):493-501. [https://doi.org/10.1016/S1389-1723\(01\)80307-8](https://doi.org/10.1016/S1389-1723(01)80307-8)
- 31) Schuster M, Wang P, Paulson JC, Wong CH. Synthesis of glycopeptides and oligosaccharides. *J. Am. Chem. Soc*. 1994; 116: 1135.
<http://dx.doi.org/10.1021/ja00082a047>

The structures reported in fig. 1-3, 5-6, and 8 are obtained by RasWin v.2.7.5

This research was supported by the “Departments of Excellence-2018” Program (*Dipartimenti di Eccellenza*) of the Italian Ministry of Education, University and Research, DIBAF-Department of University of Tuscia, Project “Landscape 4.0 – food, wellbeing and environment”, through the provision of free computing time on NARTEN Cluster Facility, given by Prof. Nico Sanna.

Supplementary material

The file pept_penta_T6AO.mpg shows the association of fig.8 in rotation.

The file pept9glyco_N262.mpg shows the association of fig.11 in rotation.

They have been obtained as follows. The program MoluCAD shows the rotation of the molecule imported in pdb format (<https://www.kinematics.com/products/molucad.php>). The image screen is recorded by means of the program CamStudio in avi format (<https://camstudio.org/>). The avi format is converted in mpg format by means of the online converter (<https://www.online-convert.com/>)