**Supporting information**

Investigating the Antiviral Therapeutic Potentialities of Marine Polycyclic Lamellarin Pyrrole Alkaloids as Promising Inhibitors for SARS-CoV-2 and Zika Main Proteases (Mpro)

Florbela Pereira\*, Loay Bedda Mohamed A. Tammam, Abdul-Kader Alabdullah, Reem K. Arafa\* and Amr El-Demerdash\*

aLAQV‑REQUIMTE, Department of Chemistry, NOVA School of Science and Technology, Universidade Nova de Lisboa, 2829‑516 Caparica, Portugal

bDrug Design and Discovery Laboratory, Helmy Institute for Medical Sciences, Zewail City of Science and Technology, Giza 12578, Egypt

c Biomedical Sciences Program, University of Science and Technology, Zewail City of Science and Technology, Giza 12578, Egypt

dDepartment of Biochemistry, Faculty of Agriculture, Fayoum University, Fayoum 63514, Egypt

e Crop Genetics Department, John Innes Centre, Colney, Norwich NR4 7UH, UK

f Division of Organic Chemistry, Department of Chemistry, Faculty of Sciences, Mansoura University, Mansoura 35516, Egypt

g Department ofBiochemistry and Metabolism, the John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK

\*Correspondences:

Amr El-Demerdash a\_eldemerdash83@mans.edu.eg, Amr.El-Demerdash@jic.ac.uk

Florbela Pereira florbela.pereira@fct.unl.pt

Reem K. Arafa rkhidr@zewailcity.edu.eg



**Figure S1**. Reported lamellarin pyrrole alkaloids (**1-22**)



**Figure S2**. Reported lamellarin pyrrole alkaloids (**23-39**)

****

**Figure S3.** The re-docking experiment of the positive control compound O6K against Mpro of SARS-CoV-2. The solved structure (PDB ID:6Y2G, Chain A) (blue cartoon with pink sticks for **O6K**) was superimposed on the docked complex (beige cartoon with sticks for O6K by element).

**Table S2**. Calculated binding free energies (∆GB, in kcal/mol) by molecular docking for the top 10 selected lamellarin derivatives and the positive control (**O6K)**, for each target, as well as their reported biological activities.

|  |  |  |
| --- | --- | --- |
| **Lamellarins** | **Molecular Docking****∆GB, in kcal/mola** | **Reported Biological Activities** |
| **SARS-CoV-2 Mpro** |  **Zika Mpro** |
| D (**7**) | --- | -8.13 | Potent inhibitor of Topoisomerase I (Ishibashi et al., 2002); cytotoxic activity against cancer cell lines e.g. HeLa, XC, Vero, MDCK (nM) (Ishibashi et al., 2002). |
| E (**9**) | -8.40 | --- | --- |
| G (**12**) | -8.57 | -8.47 | --- |
| H (**14**) | **-9.37** | -8.40 | Topoisomerase I (IC50 = 0.23 mM) (C. P. Ridley et al., 2002). |
| J (**16**) | -8.47 | -8.13 | --- |
| K (**17**) | **-9.40** | --- | Toxic against Topoisomerase I (Christian Bailly, 2004a). |
| L (**20**) | -8.67 | -8.17 | --- |
| S (**26**) | -9.00 | **-8.53** | --- |
| U (**30**) | -8.47 | -8.10 | --- |
| Z (**39**) | -9.30 | **-8.63** | --- |
| B 20-sulfate (**3**) | -8.40 | --- | --- |
| G 8-sulfate (**13**) | --- | -8.30 | --- |
| L 20-sulfate (**22**) | --- | -8.30 | --- |
| (**O6K**)b | -7.90 | -7.83 | --- |

a The lamellarin derivatives selected have a calculated ∆GB ≤ -8.4 kcal/mol and -8.1 kcal/mol for SARS-CoV-2 Mpro and Zika Mpro, respectively. b Positive Control.

**Table S3.** The detailed hydrogen bonds and hydrophobic interactions established upon docking the (**O6K**), and lamellarin derivatives (**1-39**) against the SARS-CoV-2 Mpro (PDB ID: 6LU7) chain A.

|  |  |  |
| --- | --- | --- |
| **Lamellarins** | **Common amino acids (aa)** | **Common aa number** |
| **H-bond residues** | **Hydrophobic interaction residues** |
| F (**11**), J (**16**), N (**24**), T 20-sulfate (**29**) | none | Gln189, Glu166, His163, His164, Leu141, Met165, Phe140 | 7 |
| B (**2**), B 20-sulfate (**3**), D triacetate (**8**) | Gly143 | Arg188, Asn142, Cys145, Gln189, Glu166, His41, His164, Met165, Pro168, Thr190 | 11 |
| I (**15**) | His41 | Arg188, Asn142, Cys145, Gln189, Glu166, Gly143, His164, Leu141, Met165, Phe140, Pro168, Thr190 | NAa |
| E (**9**), E 20-sulfate (**10**), L (**20**), U (**30**), U 20-sulfate (**31**), W (**34**), V 20-sulfate (**33**) | HIS163 | Arg188, Asn142, Asp187, Cys145, Gln189, Glu166, Gly143, His41, Met165 | 10 |
| G (**12**) | Asn142, His163 | Arg188, Asp187, Cys145, Gln189, Glu166, Gly143, His41, Met49, Met165 | NAa |
| D (**7**) | Asn142, Leu141 | Arg188, Asp187, Gln189, Glu166, His41, His163, His164, Met165, Phe140 | NAa |
| K triacetate (**19**), L triacetate (**21**), L 20-sulfate (**22**), N triacetate (**25**), T diacetate (**28**), X triacetate (**36**) | Cys145, His41 | Arg188, Asn142, Gln189, Glu166, Gly143, His164, Leu141, Met49, Met165, Phe140, Pro168 | 13 |
| K diacetate (**18**) | Cys145, Ser144 | Arg188, Asn142, Asp187, Gln189, Glu166, Gly143, His41, His164, Leu141, Met49, MetT165, Phe140, Tyr54 | NAa |
| Y 20-sulfate (**38**) | Cys145, Thr190 | Ala191, Gln189, Glu166, His163, His164, Leu141, Met165, Phe140, Pro168 | NAa |
| V (**32**) | Glu166, His163 | Arg188, Asn142, Asp187, Cys145, Gln189, Gly143, His41, His164, Leu141, Leu167, Met165 | NAa |
| A (**1**) | Gly143, Glu166 | Arg188, Asn142, Asp187, Gln189, His41, His164, Leu141, Leu167, Met165 | NAa |
| C (**4**) | Gly143, Leu141 | Arg188, Asn142, Asp187, Gln189, Glu166, His41, His163, His164, Met165, Phe140 | NAa |
| K (**17**)c | Gly143, Ser144 | Arg188, Asn142, Asp187, Cys145, Gln189, Glu166, His41, His163, His164, Leu141, Met49, Met165 | NAa |
| M (**23**) | Gly143, Thr190 | Arg188, Asn142, Cys145, Gln189, Glu166, His41, His164, Met165, Phe140, Pro168 | NAa |
| X (**35**) | Arg188, Gly143, Thr190 | Asn142, Cys145, Gln189, Glu166, His41, His164, Met165, Pro168 | NAa |
| G 8-sulfate (**13**) | Asn133, Leu287, Lys137 | Arg131, Asn238, Asp197, Asp289, Leu272, Leu286, Thr199, Tyr239, Val171 | NAa |
| S (**26**)c, Z (**39**)c | Cys145, Gly143, Ser144 | Asn142, Gln189, Glu166, His41, His163, His164, Leu141, Met49, Met165 | 12 |
| C diacetate (**5**), C 20-sulfate (**6**) | Cys145, Ser144, Tyr54 | Arg188, Asn142, Asp187, Gln189, Glu166, Gly143, His164, Leu141, Met49, Met165 | 13 |
| T (**27**), Y (**37**) | Gly143, His41, Thr190 | Arg188, Asn142, Asp187, Cys145, Gln189, Glu166, His164, Leu141, Met49, Met165, Ser144, Thr25 | 15 |
| H (**14**)3 | Asn142, Gly143, His163, Ser144 | Asp187, Cys145, Gln189, Glu166, His41, His164, Leu141, Met165 | NAa |
| **O6K**b | Gln189, Glu166, Tyr54 | Asn142, Asp187, Cys145, His41, His163, His164, Met49, Met165, Phe140, Thr190 | NAa |

a Not applicable. b Positive Control. c Best docking scored molecules.

**Table S4**. The detailed hydrogen bonds and hydrophobic interactions established upon docking the (**O6K**), and lamellarin derivatives (**1-39**) against the Zika Mpro (PDB ID: 5H4I) chains A and B.

|  |  |  |
| --- | --- | --- |
| **Lamellarins** |  **Common amino acids (aa)** | **Common aa number** |
| **H-bond residues** | **Hydrophobic interaction residues** |
| B 20-sulfate (**3**),N triacetate (**25**),U 20-sulfate (**31**) | none | Ala132, Asp129, Gly151, Gly153, Ser135, Tyr130, Tyr161, Val155 | 8 |
| T 20-sulfate (**29**) | Ala132, Gly153, Tyr161 | Asn152, Asp83a, HisS51, Phe84a, Pro131, Tyr130, Val155 | NAb |
| H (**14**)d | Asn152, Gly151, Gly153, Pro131, Tyr130 | Ala132, Ser135, Tyr161, Val155 | NAb |
| K diacetate (**18**) | Asp129 | Ala132, Asn152, Asp75, Asp83a, Gly151, Gly153, His51, Phe84a, Tyr130, Tyr161, Val154, Val155 | NAb |
| E 20-sulfate (**10**) | Asp129, His51, Lys54, Ser135 | Ala132, Asp83a, Tyr161, Val155 | NAb |
| A (**1**), C (**4**), I (**15**), V (**32**) | Asp129, His51, Ser135, Tyr130 | Ala132, Asp83a, Gly151, Tyr161, Val155 | 9 |
| E (**9**), F (**11**) | Asp129, Val155 | Ala132, Asn152, Asp83a, Gly151, His51, Ser135, Tyr130, Tyr161, Val154 | 11 |
| V 20-sulfate (**33**) | Asp83a | Ala132, Asp129, Gly151, His51, Pro131, Ser135, Tyr130, Tyr161, Val154, Val155 | NAb |
| D (**7**), N (**24**) | Asp83a, Asp129, His51, Ser135 | Ala132, Gly151, Gly153, Tyr130, Tyr161, Val154, Val155 | 11 |
| K (**17**) | Asp83a, Asp129, His51, Ser135, Tyr130 | Ala132, Gly151, Gly153, Tyr161, Val154, Val155 | NAb |
| J (**16**), L (**20**), G (**12**)d, S (**26**)d, Z (**39**)d | Asp83a, Asp129, Ser135 | Ala132, Gly151, His51, Tyr130, Tyr161, Val154, Val155 | 10 |
| G 8-sulfate (**13**), U (**30**) | Asp83a, Ser135 | Ala132, Asn152, Asp129, GLY151, His51, Tyr130, Tyr161, Val154, Val155 | 11 |
| M (**23**), X (**35**) | Asp83a, Ser135, Tyr130 | Ala132, Asn152, Asp129, Gly151, Tyr161, Val154, Val155 | 10 |
| B (**2**), W (**34**) | Asp83a, Ser135, Val155 | Ala132, Asn152, Asp129, Gly153, Tyr130, Tyr161, Val154 | 10 |
| Y 20-sulfate (**38**) | Gly151 | Ala132, Asn152, Asp129, Gly151, Tyr130, Tyr161, Val154 | NAb  |
| **O6K**c | Asp83a, Asn152, Gly153, His51, Ser135, Tyr161 | Ala132, Gly151, Lys54, Ser81a, Trp50, Val36, Val72 | NAb |

a From chain A. b Not applicable. c Positive Control. d Best docking scored molecules.



**Figure S4**: SARs studies for most promising lamellarins compounds against SARS-Cov-2 Mpro based on their binding affinities values and compered to positive control (**OK6**)



**Figure S5**: SARs studies for most promising lamellarins compounds against Zika Mpro based on their binding affinities values and compered to positive control (**OK6**)