Acetamido-propanoic acid derived compounds as protease inhibitors to target coronaviruses

Short title: Novel coronavirus protease inhibitors

Arun HS Kumar[1] and Vikram Sharma[2]

[1]Veterinary Biosciences, School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland. [2]Head Proteomics Core, School of Biomedical Sciences, University of Plymouth, Plymouth, Devon PL6 8BU, United Kingdom.

Corresponding Author:

Dr. Arun HS Kumar, DVM, PhD.
Director Stemcology, Room 216, School of Veterinary Medicine, University College Dublin, Belfield, Dublin-04, Ireland. Phone: 0035317166230, Fax: 00353017166104,
email: arun.kumar@ucd.ie

Abstract

Background: Coronavirus infection as evolved into a major pandemic and is currently being treated using established antiviral agents developed for other similar viruses. Considering the frequent mutations rate in this virus, novel drugs will be necessary for effective treatment in future. Hence this study evaluated the potential of acetamido-propanoic acid derived compounds as viral protease inhibitors. **Materials and Methods**: Using cheminformatics approach novel acetamido-propanoic acid derived compounds were designed and their binding efficacy against the Coronavirus 2019 (Covid-19) protease was tested using Insilco pharmacology. **Results**: STGYC compounds had physico-chemical features favourable for developing them for potential clinical use. The binding efficacy of STGYC compounds against covid-19 protease was similar to that of favipiravir, which currently being reported to be effective in treating coronavirus infection. **Conclusion**: STGYC compounds shown favourable features to be further evaluated and developed as viral protease inhibitors.

Key Words: Protease inhibitors, Coronavirus, Covid-19, antiviral drugs,

Introduction

The last decade as witnessed major epidemics from coronavirus infections, which are known to cause respiratory and enteric symptoms.[1-3] Coronaviruses with their genome size ranging from 27-34 kilobases are the largest among the RNA viruses and are enveloped with a positive

sense single stranded RNA and nucleocapsid of helical symmetry.[3, 4] The coronaviruses enters the host cells by attaching to the cell surface receptor using its spike proteins, which is a key step in the coronavirus infectivity. In addition to the coronavirus spike proteins, its envelope, matrix, nucleocapsid and other non-structural proteins are involved in various stages of viral replication and pathogenesis.[1-5] Hence several of these viral proteins have been targeted to develop antiviral drugs. Several broad spectrum antiviral drugs are available to treat viral infections albite with varying efficacy (http://drugvirus.info/).[6, 7] Among these antiviral drug categories, the protease inhibitors are widely used and to best of our knowledge, protease inhibitors have been the first line of drugs to treat the current pandemic of the novel coronavirus of 2019 (Covid-19).[4, 5, 7] Hence in this work using the reported crystal structures of covid-19, an Insilco pharmacological approach was used to develop novel inhibitors of covid-19 protease.[8, 9]

Materials and Methods

The protein data bank (https://www.rcsb.org/) was screened for the reported crystal structures of covid-19 proteins. Of the nine reported crystal structures of covid-19 proteins until 15th March 2020, seven were shortlisted in our analysis. PDB files of the reported protein crystal structures were downloaded. VADAR version 1.8 (http://vadar.wishartlab.com/)[10] was used for the structural and stereochemical analysis of the protein structures. Using a cheminformatics approach novel Acetamido-propanoic acid derived compounds were designed and assessed in the ChemDraw software. Molecular docking was performed to reveal the interactions of novel ligands against the target protein i.e., Covid-19 protease using AutoDock Vina and the docked protein-ligand complex were visualised using the PyMOL v 1.8.2.0 software.[8, 9, 11]

Results

The seven Covid-19 protein structures analysed are listed in table 1. The molecular weight of the proteins and the best peptide sequences which may find application in antibody generation were identified in the VADAR screening. The structural analysis was focused at evaluating the relative proportion of helix, beta-sheet, coil, turns and the hydrogen-bonds (HBonds) (Figure 1). The total accessible surface area (ASA), ASA of the protein backbone and sidechains are reported in figure 2. The volume of the protein structures analysed in summarised in figure 3. A representative of acetamido-propanoic acid derived compounds is shown in figure 4a. These compounds are code named as STGYC compounds. The general physico-chemical characteristics of the STGYC compounds is summarised in table 2. Two of the STGYC compounds (STGYC 7126 and 7126A) were docked against the crystal structure of Covid-19 protease (PDB ID 6Y84) and their binding affinities are compared with that of the standard

compounds i.e., Favipiravir and Lopinavir/Ritonavir (Table 3). The ligand bound to its receptor site on the Covid-19 protease is shown in the figure 4B.

Discussion

We report here novel acetamido-propanoic acid derived compounds which may potentially be useful has Covid-19 protease inhibitor. Although the binding affinity of STGYC compounds was lower than the lopinavir/ritonavir, the binding affinity was comparable to that of favipiravir, which is reported to be effective in treating covid-19 infection. The binding affinity and the physico-chemical properties of the STGY compounds were within the favourable range for them to further developed for clinical use. Considering the highly mutating nature of the coronaviruses including the covid-19,[2, 3] it is necessary to have alternatives such as the STGYC compounds as potential viral protease inhibitors. We have also included in this study the structural and stereochemical features of 7 covid-19 proteins, which are potentially targetable. In further studies we will look at the feasibility of STGYC compounds to target these alternate covid-19 targets using the Insilco described here.

In summary, STGYC compounds have favourable features to be developed as protease inhibitors for future clinical application.

Conflict of Interest

None

Acknowledgement

Research support from University College Dublin-Seed funding/Output Based Research Support Scheme (AHSK), Royal Society-UK (AHSK) and Stemcology (AHSK) is acknowledged.

References:

[1]. G. Grasselli, A. Pesenti, M. Cecconi, Critical Care Utilization for the COVID-19 Outbreak in Lombardy, Italy: Early Experience and Forecast During an Emergency Response, JAMA (2020).

[2]. A. Remuzzi, G. Remuzzi, COVID-19 and Italy: what next?, Lancet (2020).

[3]. D. The Lancet Infectious, COVID-19, a pandemic or not?, Lancet Infect Dis (2020).

[4]. G. Thomson, COVID-19: social distancing, ACE 2 receptors, protease inhibitors and beyond?, Int J Clin Pract (2020) e13503.

[5]. L.R. Baden, E.J. Rubin, Covid-19 - The Search for Effective Therapy, N Engl J Med (2020).

[6]. P.I. Andersen, A. Ianevski, H. Lysvand, A. Vitkauskiene, V. Oksenych, M. Bjoras, K. Telling, I. Lutsar, U. Dampis, Y. Irie, T. Tenson, A. Kantele, D.E. Kainov, Discovery and development of safe-in-man broad-spectrum antiviral agents, Int J Infect Dis (2020).

[7]. B. Cao, Y. Wang, D. Wen, et al., A Trial of Lopinavir-Ritonavir in Adults Hospitalized with Severe Covid-19, N Engl J Med (2020).
[8]. S.I. O'Donoghue, D.S. Goodsell, A.S. Frangakis, F. Jossinet, R.A. Laskowski, M. Nilges, H.R. Saibil, A. Schafferhans, R.C. Wade, E. Westhof, A.J. Olson, Visualization of macromolecular structures, Nat Methods 7(3 Suppl) (2010) S42-55.
[9]. D. Seeliger, B.L. de Groot, Ligand docking and binding site analysis with PyMOL and Autodock/Vina, J Comput Aided Mol Des 24(5) (2010) 417-22.
[10]. L. Willard, A. Ranjan, H. Zhang, H. Monzavi, R.F. Boyko, B.D. Sykes, D.S. Wishart, VADAR: a web server for quantitative evaluation of protein structure quality, Nucleic Acids Res 31(13) (2003) 3316-9.
[11]. O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J Comput Chem 31(2) (2010) 455-61.

Table 1. List and details of the covid-19 protein crystal structures evaluated

PDB file ID	**Details**	**M Wt**	**Antigenic Sequence**
5R80	Crystal Structure of COVID-19 main protease in complex with Z18197050	33522.64	TVNVLAWLYA PLTQDHVDIL VLDMCASLKE
6M03	Crystal structure of COVID-19 main protease in apo form	33797.96	GKVEGCMVQV GNVQLRVIGH VDTANPKTPK
6VW1	Structure of 2019-nCoV chimeric receptor-binding domain complexed with its receptor human ACE2	181354.5	STIEEQAKTF LFYQSSLASW VQNMNNAGDK
6VYO	Crystal structure of RNA binding domain of nucleocapsid phosphoprotein from SARS coronavirus 2	54752.61	IIWVATEGAL
6W01	Crystal Structure of NSP15 Endoribonuclease from SARS CoV-2 in the Complex with a Citrate	77760	GQVDLFRNAR LTVFFDGRVD KKPTETICAP
6W02	Crystal Structure of ADP ribose phosphatase of NSP3 from SARS CoV-2 in the complex with ADP ribose	35650.16	ADPIHSLRVC EDIQLLKSAY AVFDKNLYDK
6Y84	COVID-19 main protease with unliganded active site (2019-nCoV, coronavirus disease 2019, SARS-CoV-2)	33522.64	SGFRKMAFPS GKVEGCMVQV GNVQLRVIGH

Table 2. General physico-chemical characteristics of the STGYC compounds

Molecular Formula:	C13H16FN5O6S2
Formula Weight:	421.4244
Composition:	C(37.05%) H(3.83%) F(4.51%) N(16.62%) O(22.78%) S(15.22%)
Molar Refractivity:	95.70 ± 0.3 cm^3
Molar Volume:	260.7 ± 3.0 cm^3
Parachor:	795.9 ± 4.0 cm^3
Index of Refraction:	1.655 ± 0.02
Surface Tension:	86.8 ± 3.0 dyne/cm
Density:	1.616 ± 0.06 g/cm^3
Polarizability:	37.93 ± 0.5 10-24cm^3
RDBE:	8
Monoisotopic Mass:	421.052601 Da
Nominal Mass:	421 Da
Average Mass:	421.4244 Da
M+:	421.052053 Da
M-:	421.05315 Da
[M+H]+:	422.059878 Da
[M+H]-:	422.060975 Da
[M-H]+:	420.044228 Da
[M-H]-:	420.045325 Da
Log P:	2.62 ± 0.86

Table 3. Binding affinity of the test and standard compounds.

Compounds	**Binding affinity (kcal/mol)**
Lopinavir/Ritonavir	-7.967 ± 0.608
Favipiravir	-5.978 ± 0.387
STGYC7126	-5.933 ± 0.250
STGYC7126A	-5.656 ± 0.336

Figure 1. Ramachandran plots of the seven protein crystal structures. Graph represents the relative proportion of helix, beta-sheet, coil, turns and the hydrogen-bonds (HBonds) in the protein structure.

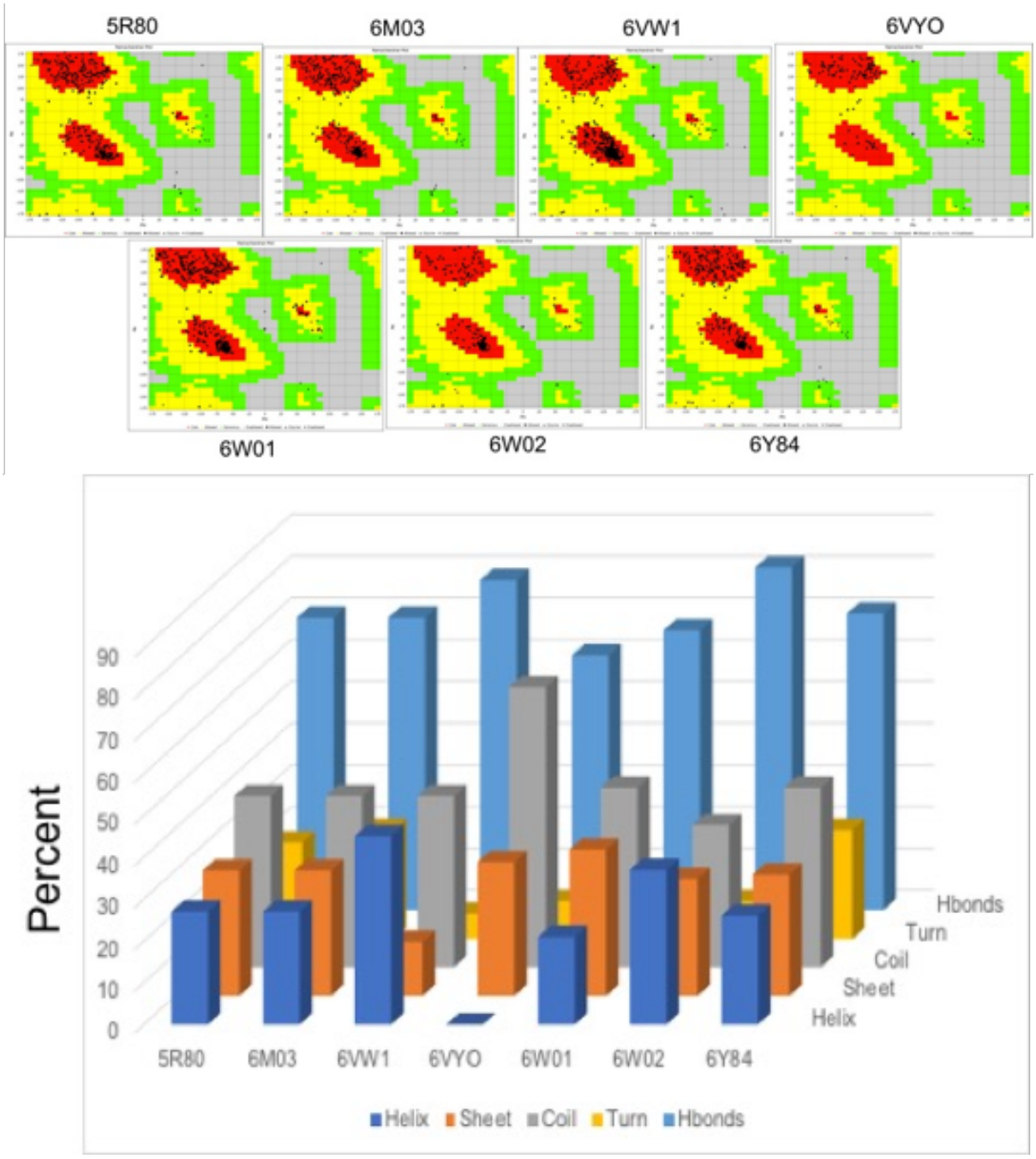

Figure 2: The total accessible surface area (ASA), ASA of the protein backbone and sidechains of the seven covid-19 proteins.

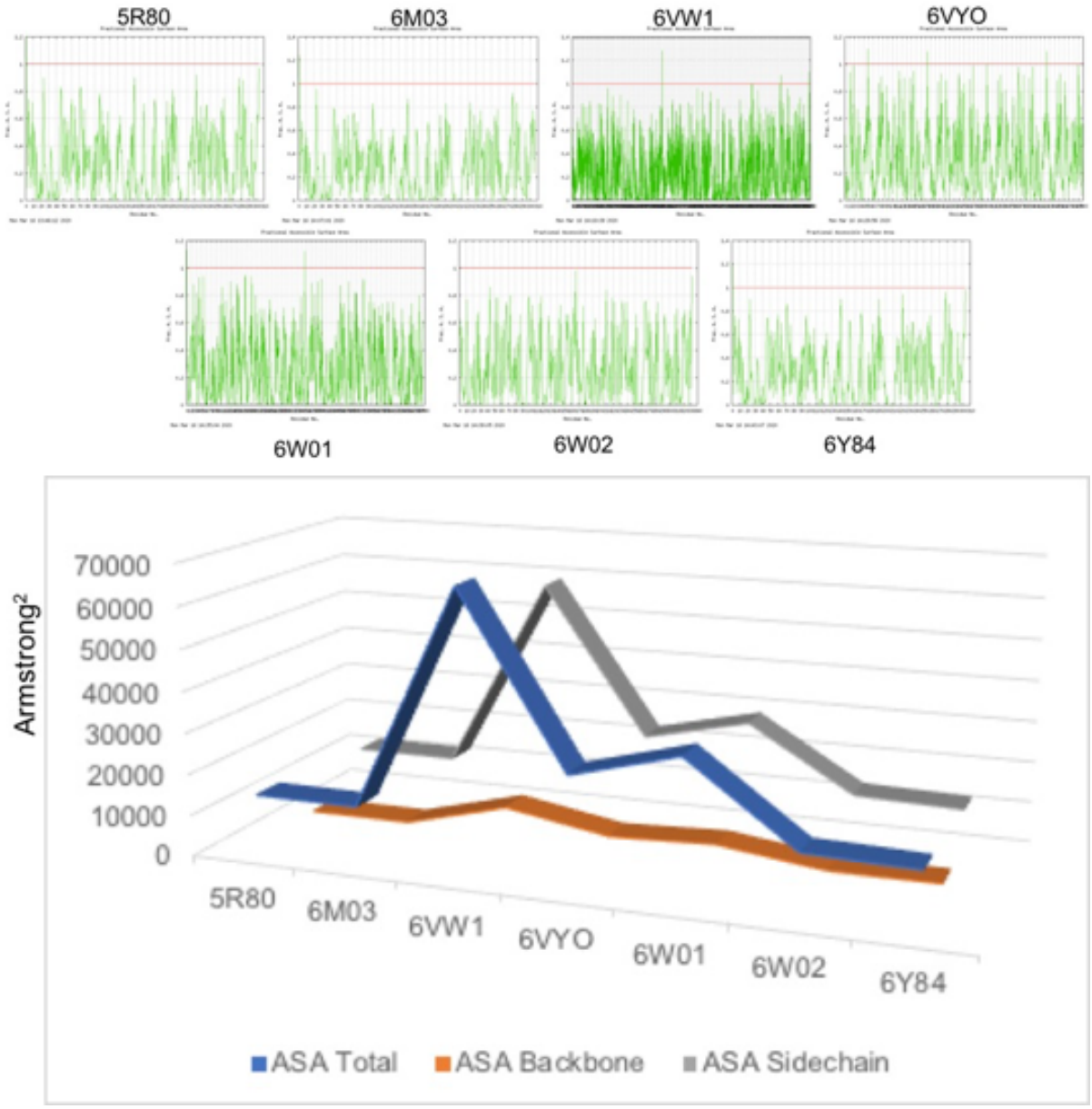

Figure 3: Volume of the seven covid-19 proteins.

Figure 4: Representative structure of the STGYC compounds (A). STGYC compounds and the standard drugs (Favipiravir and Lopinavir/Ritonavir) shown bound to its receptors on the covid-19 protease (B).