Molecular binding studies of anthocyanins with multiple antiviral activities against SARS-CoV-2

Precious Ayorinde Akinnusi*, Samuel Olawale Olubode and Wasiu Adeboye Salaudeen

Abstract

Background: The search for ideal drugs with absolute antiviral activity against SARS-CoV-2 is still in place, and attention has been recently drawn to natural products. Several molecular targets have been identified as points of therapeutic intervention. The targets used in this study include SARS-CoV-2 helicase, spike protein, RNA-dependent RNA polymerase, main protease, and human ACE-2. An integrative computer-aided approach, which includes molecular docking, pharmacophore modeling, and pharmacokinetic profiling, was employed to identify anthocyanins with robust multiple antiviral activities against these SARS-CoV-2 targets.

Result: Four anthocyanins (Delphinidin 3-O-glucosyl-glucoside, Cyanidin 3-O-glucosyl-rutinoside, Cyanidin 3-(p-coumaroyl)-diglucoside-5-glucoside), and Nasunin) with robust multiple inhibitory interactions were identified from a library of 118 anthocyanins using computer-aided techniques. These compounds exhibited very good binding affinity to the protein targets and moderate pharmacokinetic profiles. However, Cyanidin 3-O-glucosyl-rutinoside is reported to be the most suitable drug candidate with multiple antiviral effects against SARS-CoV-2 due to its good binding affinity to all five protein targets engaged in the study.

Conclusions: The anthocyanins reported in this study exhibit robust binding affinities and strong inhibitory molecular interactions with the target proteins and could be well exploited as potential drug candidates with potent multiple antiviral effects against COVID-19.

Keywords: SARS-CoV-2, Anthocyanins, Helicase, 3CL protease, RNA-dependent RNA polymerase, ACE-2, Molecular docking, Pharmacophore modeling, ADME/Tox

Background

Severe acute respiratory syndrome—coronavirus 2 (SARS-CoV-2)—is known to be a highly infectious beta-coronavirus that affects humans (Lu et al. 2020). The SARS-CoV-2 outbreak has been declared as a Public Health Emergency of International Concern (PHEIC) since 2020, and efforts have been intensified to discover antiviral drugs with potent activity against SARS-CoV-2 (Burki 2020).

Many viral protein molecules encoded by the viral genome play crucial roles in the virulence, transmissibility, and replication of SARS-CoV-2. They include the main protease (also known as the SARS-CoV-2 main protease enzyme or 3C protease) which plays a critical role in the SARS-CoV-2 life cycle by aiding the process of viral replication and transcription, the helicase protein (also called nonstructural protein 13) which plays a huge role in viral replication, RNA-dependent RNA polymerase, and the receptor-binding domain (RBD) of S-protein which binds strongly to the human angiotensin-converting enzyme 2 (ACE2) receptors among others (Lan et al. 2020; Jin et al. 2020; Newman et al. 2021).
The identification and development of compounds from natural origin, particularly plants, into potential drug candidates have shown encouraging and reassuring prospects that these compounds, including polyphenols, may be employed as therapeutic interventions against viral diseases (Mohammadi Pour et al. 2019). The polyphenols are plant secondary metabolites that structurally have more than one phenolic ring and different structural elements that bind these rings to one another. Different compounds, including flavonoids, belong to this category of secondary metabolites (Fakhar et al. 2021). Flavonoids are in turn classified into four subclasses, namely anthocyanins, flavonols, flavones, and flavanones (Castro-Acosta et al. 2016). Anthocyanins are glycosides of flavonoid with anthocyanidin C6-C3C6 skeleton, also called flavylium (2-phenylchromenylium) ion (Smeriglio et al. 2016). Several anthocyanin derivatives have been identified, and only six major anthocyanin derivatives are extensively dispersed, and they include pelargonidin, delphinidin, petunidin, cyanidin, peonidin, and malvidin (Prior and Wu 2006). Recent studies have shown that anthocyanins are potent antiviral agents, notably against influenza virus, and have the potentials as antivirals against SARS-CoV-2 (Mohammadi Pour et al. 2019; Fakhar et al. 2021; Messaoudi et al. 2021). Studies have also identified flavonoid derivatives with the same basic structure as anthocyanins having potential antiviral activity against SARS-CoV-2 main protease (Jo et al. 2020). However, the studies were restricted to one or two molecular targets. This study aims to identify anthocyanins with multiple effects against SARS-CoV-2.

Herein, a computational approach was adopted to investigate the therapeutic potentials of various anthocyanins against COVID-19. The targets include SARS-CoV-2 3CL protease, RNA-dependent RNA polymerase (RdRp), spike protein receptor-binding domain, helicase, and human ACE-2. An integrated computational approach was adopted to identify anthocyanins with potential multiple antiviral activities. The workflow includes molecular docking, ADME/Tox screening, and 3D pharmacophore modeling.

**Methods**

**Protein targets and ligand structures**

The SMILES structures of 118 anthocyanin derivatives were extracted from PhytoHub online database and converted to SDF using Marvin sketch graphical user interface. The compounds were imported into Maestro 11.5 on Schrodinger suite and prepared using the LigPrep functional tool. The 3D crystal structures of the target proteins, SARS-CoV-2 3CL protease (PDB ID: 7JT7), RNA-dependent RNA polymerase (PDB ID: 6M71), spike protein (PDB ID: 6LZG), helicase (PDB ID: 7NNG), and human ACE-2 (PDB ID: 6LZG) were also retrieved from an online repository (Protein Data Bank) and prepared subsequently for docking.

**Molecular docking**

The molecular docking procedure was carried out using the Glide script (Friesner et al. 2004) on Maestro 11.5. The compounds were docked into the prepared grid of the protein targets to identify compounds with potent inhibitory interactions with the proteins (Halgren 2009). The molecular docking procedure was initiated with the enzyme treated as a rigid body, while the ligand’s rotatable bonds were made to be flexible. The three levels of precision were employed in docking the compounds to the grids of the target proteins. Initially, high-throughput virtual screening (HTVS) was used to screen the compounds; subsequently, standard precision (SP) was employed to screen the highest-scoring compounds obtained from HTVS analysis. Finally, the more rigorous extra precision (XP) screened and scored the compounds. The top-scoring compounds were selected for post-docking analyses.

**ADMET analysis**

The ADME/Tox properties of the lead compounds, which include absorption, distribution, metabolism, excretion, and toxicity, were profiled using SwissADME and ProTox-II online servers. The pharmacological properties predicted include lipophilicity, water solubility, drug-likeness, bioavailability score, blood–brain barrier permeation, reaction with cytochrome p450 isoforms, LD50, carcinogenicity, and possible hepatotoxicity.

**Pharmacophore modeling**

The receptor–ligand complexes of the lead compounds were analyzed, and a hypothesis (E-pharmacophore) was generated using the phase interface of Schrodinger suite to highlight the major properties that actively contribute to the characteristic binding of the lead ligands to the active sites of the target proteins.

**Results**

One hundred and eighteen anthocyanins were analyzed to determine their potential inhibitory activity against five SARS-CoV-2 targets. Eight compounds with multiple binding affinities to the target proteins were identified. The docking scores of the top-scoring compounds of each target are presented in Table 1. The docking scores demonstrate the binding energy exhibited by the compounds in complex with the molecular targets analyzed. C1 (Delphinidin 3-O-glucosyl-glucoside) showed the highest binding energy when complexed with 3Cl protease. Similarly,
C2 (Cyanidin 3-O-glucosyl-rutinoside) had the highest docking score against the helicase protein and C3 (Cyanidin 3-(p-coumaroyl)-diglucoside-5-gluicoside) against RNA-dependent RNA polymerase. Furthermore, C3 (Cyanidin 3-(p-coumaroyl)-diglucoside-5-gluicoside) and C8 (Nasunin) had the highest docking scores in complex with SARS-CoV-2 spike protein and human ACE-2, respectively.

The physicochemical properties of the identified compounds are presented in Table 2. The predicted log p value which represents the lipophilicity ranged from −3.68 to 0.42 with C7 having the highest value and C2, the lowest. The consensus log p value is the arithmetic mean of five different predictive models of lipophilicity, namely iLOGP, XLOGP3, WLOGP, MLOGP, and Silicos-IT log P. C7, which had the highest value, has the highest probability of gastrointestinal absorption.

Consequently, the predicted Silicos-IT log SW which measures the water solubility of the compounds ranged from -3.23 to 2.12 with C2 being the most water soluble and C7, the least soluble. Water solubility contributes substantially to the movement of small molecular weight compounds in systemic circulation.

The amino acid interaction of the top-scoring compounds with the active sites of the target proteins are presented in Table 3. Covalent hydrogen bond is the major form of interaction observed in the characteristic binding of the compounds to each protein target.

The drug-likeness and bioavailability scores of the top-scoring compounds are presented in Table 4. All the compounds have a bioavailability score of 0.17, and all violated 3 of the Lipinski rule-based filter of drug-likeness. Similarly, all the test compounds violated 1 of the Veber rule-based filters except C5 and C8 with 2 violations. Drug-likeness and bioavailability score predicts the likelihood of a compound being a drug candidate.

The SwissADME predicted pharmacokinetic profiles of the compounds are presented in Table 5. All the compounds have the same toxicity profile, an LD50 of 5000 mg/kg BW, and are not carcinogenic neither do they have a probability to cause harm to the liver (Fig. 1).

Figure 2 shows the pharmacophore model of the top-scoring compounds in each protein target. The model shows that aromatic ring and hydrogen bond (donor and acceptor) are the major forms of interactions that
### Table 3  2D Amino acid interaction of the top-scoring compounds

| Targets                          | Compounds |
|---------------------------------|-----------|
| 3Cl Pro                         | C1        | C8        |
| Helicase                       | C4        | C3        |
| RNA-dependent RNA polymerase    |           |           |
| Spike protein RBD               | C3        | C2        | C6        |
| ACE 2                           | C8        | C3        |
contribute to the binding of the compounds to the protein targets.

Pharmacophore model

Discussion

The global spread of SARS-CoV-2 requires urgent and novel therapeutic discoveries given the high failure rate of traditional drug discovery methods (Mirabelli et al. 2021). Attentions have been constantly drawn to natural compounds due to their relatively lower side effects (Lin et al. 2017). In light of this, different anthocyanins have been previously shown to be good inhibitors of SARS-CoV-2 molecular targets (Fakhar et al. 2021; Messaoudi et al. 2021).

Eight anthocyanins with impressive binding data were identified from 118 anthocyanins screened, and they were found to show robust binding affinity to the protein targets. Hydrogen bonding is the major form of interaction observed in the binding data. In all categories, the identified anthocyanins have higher binding affinity than the standard drugs: remdesivir and hydroxychloroquine. C1, C2, C3, C5, and C8 showed impressive binding affinity to SARS-CoV-2 3CL protease. The protein functions in the maturation of viral polyprotein and is essential for the completion of SARS-CoV-2 life cycle (Kan et al. 2005). Therefore, inhibiting this protein is a proven therapeutic option in curbing coronavirus disease. All the compounds exhibited good binding affinity to 3CL protease but C1, C2, C3, C5, and C5 have greater binding than the other compounds. Similarly, C1, C2, C3, C4, and C8 showed excellent binding affinity to SARS-CoV-2 helicase. The helicase protein, also called non-structural protein 13 (NSP 13), has been identified as a target for antivirals due to its role in viral replication (Newman et al. 2021). Of all compounds tested, C4 had the highest docking score (−11.72) and the pharmacophore model

| Compounds | Lipinski #violations | Veber #violations | Bioavailability score |
|-----------|----------------------|------------------|----------------------|
| C1        | 3                    | 1                | 0.17                 |
| C2        | 3                    | 1                | 0.17                 |
| C3        | 3                    | 1                | 0.17                 |
| C4        | 3                    | 1                | 0.17                 |
| C5        | 3                    | 2                | 0.17                 |
| C6        | 3                    | 1                | 0.17                 |
| C7        | 3                    | 1                | 0.17                 |
| C8        | 3                    | 2                | 0.17                 |

| Compounds | BBB permeant | Pgp substrate | CYP1A2 inhibitor | CYP2C19 inhibitor | CYP2C9 inhibitor | CYP2D6 inhibitor | CYP3A4 inhibitor | Log Kp (cm/s) |
|-----------|--------------|---------------|------------------|-------------------|------------------|------------------|------------------|--------------|
| C1        | No           | No            | No               | No                | No               | No               | No               | −11.8        |
| C2        | No           | Yes           | No               | No                | No               | No               | No               | −13.12       |
| C3        | No           | Yes           | No               | No                | No               | No               | No               | −12.98       |
| C4        | No           | Yes           | No               | No                | No               | No               | No               | −13.73       |
| C5        | No           | Yes           | No               | No                | No               | No               | No               | −13.39       |
| C6        | No           | No            | No               | No                | No               | No               | No               | −11.84       |
| C7        | No           | No            | No               | No                | No               | No               | No               | −9           |
| C8        | No           | No            | No               | No                | No               | No               | No               | −13.12       |

| Compounds | LD50 (mg/kg) | Toxicity class | Hepatotoxicity | Carcinogenicity |
|-----------|--------------|----------------|----------------|-----------------|
| C1        | 5000         | 5              | −              | −               |
| C2        | 5000         | 5              | −              | −               |
| C3        | 5000         | 5              | −              | −               |
| C4        | 5000         | 5              | −              | −               |
| C5        | 5000         | 5              | −              | −               |
| C6        | 5000         | 5              | −              | −               |
| C7        | 5000         | 5              | −              | −               |
| C8        | 5000         | 5              | −              | −               |

C1 = Delphinidin 3-O-glucosyl-glucoside, C2 = Cyanidin 3-O-glucosyl-rutinoside, C3 = Cyanidin 3-(p-coumaroyl)-diglucoside-5-glucoside, C4 = Cyanidin 3-O-xylosyl-rutinoside, C5 = Cyanidin 3-(sinapoyl)-diglucoside-5-glucoside, C6 = Petunidin 3-O-rutinoside, C7 = Malvidin 3-O-(6′-caffeoyl-glucoside), C8 = Nasunin
showed that hydrogen bond and aromatic rings contribute substantially to its binding to NSP13.

All the test compounds (Fig. 1) exhibited good molecular binding affinity to RNA-dependent RNA polymerase (RdRp) also named NSP 12. The docking scores of the compound ranged from $-12.73$ to $-8.91$ with C3 having the highest binding affinity to the protein. C3 interacted with seven amino acids in the active site of NSP12 (THR319, SER255, ASP390, THR393, THR394, SER397, and ASN459).

Only C2 ($-10.07$) showed a good binding affinity to the receptor-binding domain of the viral spike protein required for viral entry into the host’s cell. Finally, all the compounds exhibited robust binding affinity ($-11.42$ to $-13.9$) to the fifth protein target: human angiotensin-converting enzyme 2 (ACE2) which is the receptor for SARS-CoV-2 entry into the human host.

Considering the heat map of the docking scores (Fig. 3), it can be adjudged that C2 is the most suitable multitarget antiviral agent against SARS-CoV-2 among...
Fig. 2: Pharmacophore model of top-scoring compounds. A = C1-3C1 protease complex, B = C4-Helicase complex, C = C3-RNA-dependent RNA polymerase complex, D = C2-Spike protein RBD complex, E = C8-Ace 2 complex.
the compounds reported. Similarly, C1, C3, and C8 also showed promising inhibitory potential against the target proteins.

**ADMET profile**

The predicted value of log P measures the lipophilicity of the compounds. For better accuracy, the arithmetic mean of five different models of the partition coefficient of n-octanol to water was adopted as the log P in this study. An oral drug candidate must be sufficiently lipophilic to enable it to cross the intestine into the systemic circulation. C7 had the highest (0.42) log P value and is predicted to be the most lipophilic of all compounds tested. The value of Silicos-IT Log SW as predicted by SwissADME represents the degree of solubility in water. C2 was found to be the most water-soluble. Drugs are transported to the cells that need them through the hydrophilic systemic circulation; therefore, a drug candidate must be sufficiently hydrophilic to aid its transport in the systemic circulation.

The pharmacokinetic screening of the compounds showed that they all have a bioavailability score of 0.17. Abbots Bioavailability Score is the likelihood of a compound having greater than 10% bioavailability in rats or measurable Caco-2 permeability (Martin 2005).

All the compounds are predicted to be non-substrates of the CYP isoforms engaged and would not elicit a drug–drug interaction. Cytochrome P450 is a family of highly similar enzymes that play a big role in the metabolism and excretion of various compounds. Studies have suggested that about 50–90% of biologically active compounds are substrates of five isoforms of the superfamily (CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4) (Diana et al. 2017), and inhibition of the activity of these enzymes can cause a drug–drug response (Huang et al. 2008).

**Conclusions**

Delphinidin 3-O-glucosyl-glucoside, Cyanidin 3-O-glucosyl-rutinoside, Cyanidin 3-(p-coumaroyl)-diglucoside-5-glucoside), and Nasunin exhibit robust binding data to five SARS-CoV-2 molecular targets; therefore, these compounds are recommended for further analyses as they could be explored as potential antiviral agents with multiple targets against SARS-CoV-2.

**Abbreviations**

ACE: Angiotensin-converting enzyme; CoV: Coronavirus; CYP: Cytochrome P450; NSP: Non-structural protein; P-gp: Permeability glycoprotein; RBD: Receptor-binding domain; RdRp: RNA-dependent RNA polymerase; SARS: Severe acute respiratory syndrome.

**Acknowledgements**

Not applicable.

**Author contributions**

PAA conceptualized the study, edited the manuscript, and carried out the post-docking analyses, SOO carried out the molecular docking procedure and edited the manuscript, and WAS wrote the manuscript. All authors have read and approved the manuscript.

**Funding**

Not applicable.
Availability of data and materials
The datasets generated are available from the corresponding author on request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 15 December 2021  Accepted: 30 March 2022
Published online: 13 April 2022

References

Burki TK (2020) Coronavirus in China. Lancet Respir Med 8(3):238. https://doi.org/10.1016/s2213-2600(20)30056-4

Castro-Acosta ML, Lenihan-Geels GN, Corpe CP, Hall WL (2016) Berries and anthocyanins: promising functional food ingredients with postprandial glycaemia-lowering effects. Proc Nutr Soc 75(S):342–355. https://doi.org/10.1017/PNS00065116000240

Daina A, Michielin O, Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep 7(1):1–13. https://doi.org/10.1038/srep42717

Fakhari Z, Faramarzi B, Pacifico S, Faramarzi S (2021) Anthocyanin derivatives as potent inhibitors of SARS-CoV-2 main protease: an in-silico perspective of therapeutic targets against COVID-19 pandemic. J Biomol Struct Dyn 39(16):6171–6185. https://doi.org/10.1080/07391102.2020.1801510

Fan K, Wei P, Peng Q, Chen S, Huang C, Ma L, Lai I (2004) Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like protease. J Biol Chem 279(3):1637–1642. https://doi.org/10.1074/jbc.M310875200

Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. I. Method and assessment of docking accuracy. J Med Chem 47(7):1739–1749. https://doi.org/10.1021/jm0306430

Halgren TA (2009) Identifying and characterizing binding sites and assessing druggability. J Chem Inf Model 49(2):377–389. https://doi.org/10.1021/ci800324m

Huang SM, Strong JM, Zhang L, Reynolds KS, Nallani S, Temple R, Lesko LJ (2008) New era in drug interaction evaluation: US Food and Drug Administration update on CYP enzymes, transporters, and the guidance process. J Clin Pharmacol 48(6):662–670. https://doi.org/10.1177/009127007312153

Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, Yang H (2020) Structure of Mpro from COVID-19 virus and discovery of its inhibitors. bioRxiv. https://doi.org/10.1101/2020.05.27.110127

JZ (2021) Morphological cell profiling of SARS-CoV-2 infection identifies drug repurposing candidates for COVID-19. Proc Natl Acad Sci. https://doi.org/10.1073/pnas.2011012

Kan B, Wang M, Jing H, Xu H, Jiang X, Yan M, Liang W, Zheng H, Wan K, Liu Q, Cui B, Xu Y, Zhang E, Wang H, Ye J, Li G, Li M, Cui Z, Qi X, Chen K, Du L, Gao K, Zhao YT, Zou XZ, Feng YJ, Gao YF, Hai R, Yu D, Guan Y, Yu J (2005) Molecular evolution analysis and geographic investigation of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal market and on farms. J Virol 79(18):11892–11900. https://doi.org/10.1128/JVI.79.18.11892-11900.2005

Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, Tan W (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet 395(10224):565–574. https://doi.org/10.1016/s0140-6736(20)30251-8

Martin YC (2005) A bioavailability score. J Med Chem 48(9):3164–3170. https://doi.org/10.1021/jm0492002

Messaoudi O, Gouzi H, EL-Hoshoudy AN, Benaceur F, Patel C, Goswami D, Ben-dahou M (2021) Berries anthocyanins as potential SARS-CoV-2 inhibitors targeting the viral attachment and replication; molecular docking simulation. Egypt J Pet 30(1):33–43. https://doi.org/10.1016/j.ejpe.2021.01.001

Mirabelli C, Worthing JW, Zhang CJ, McCarty SM, Fursmidt R, Pretto CD, Sexton JZ (2021) Morphological cell profiling of SARS-CoV-2 infection identifies drug repurposing candidates for COVID-19. Proc Natl Acad Sci. https://doi.org/10.1073/pnas.2011012

Mohammadi Pour P, Fakhri S, Asgary S, Farzamz M, Echeverria J (2019) The signaling pathways, and therapeutic targets of antiviral agents: focusing on the antiviral approaches and clinical perspectives of anthocyanins in the management of viral diseases. Front Pharmacol 10:1207. https://doi.org/10.3389/fphar.2019.01207

Newman JA, Douangamath A, Yazdani S, Yosaatmadja Y, Aimon A, Brandao Neto J et al (2021) Structure, mechanism and crystallographic fragment screening of the SARS-CoV-2 NSP13 helicase. BioRxiv. https://doi.org/10.1101/2021.03.15.453526

Pinto D, Park YJ, Beltramello M, Walls AC, Tortorici MA, Bianchi S, Corti D (2020) Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV-2 antibody. Nature 583(7815):290–295. https://doi.org/10.1038/s41389-fphar.2019.01207

Prior RL, Wu X (2006) Anthocyanins: structural characteristics that result in unique metabolic patterns and biological activities. Free Radical Res 40(10):1014–1028. https://doi.org/10.1080/10715760600758522

Smeriglio A, Barreca D, Bellocco E, Trombeta D (2016) Chemistry, pharmacology and health benefits of anthocyanins. Phytother Res 30(8):1265–1286. https://doi.org/10.1002/ptr.5642

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen journal and benefit from:

► Convenient online submission
► Rigorous peer review
► Open access: articles freely available online
► High visibility within the field
► Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com