Contemporary Approaches to the Discovery and Development of Broad-Spectrum Natural Product Prototypes for the Control of Coronaviruses

George S. Hanna, Yeun-Mun Choo, Ryan Harbit, Heather Paeth, Sarah Wilde, James Mackle, Jacopo-Umberto Verga, Bethany J. Wolf, Olivier P. Thomas, Peter Croot, James Cray, Courtney Thomas, Ling-Zhi Li, Gary Hardiman, Jin-Feng Hu, Xiaojuan Wang, Dharmeshkumar Patel, Raymond F. Schinazi, Barry R. O’Keefe, and Mark T. Hamann

Cite This: https://doi.org/10.1021/acs.jnatprod.1c00625

ABSTRACT: The pressing need for SARS-CoV-2 controls has led to a reassessment of strategies to identify and develop natural product inhibitors of zoonotic, highly virulent, and rapidly emerging viruses. This review article addresses how contemporary approaches involving computational chemistry, natural product (NP) and protein databases, and mass spectrometry (MS) derived target−ligand interaction analysis can be utilized to expedite the interrogation of NP structures while minimizing the time and expense of extraction, purification, and screening in BioSafety Laboratories (BSL)3 laboratories. The unparalleled structural diversity and complexity of NPs is an extraordinary resource for the discovery and development of broad-spectrum inhibitors of viral genera, including Betacoronavirus, which contains MERS, SARS, SARS-CoV-2, and the common cold. There are two key technological advances that have created unique opportunities for the identification of NP prototypes with greater efficiency: (1) the application of structural databases for NPs and target proteins and (2) the application of modern MS techniques to assess protein−ligand interactions directly from NP extracts. These approaches, developed over years, now allow for the identification and isolation of unique antiviral ligands without the immediate need for BSL3 facilities. Overall, the goal is to improve the success rate of NP-based screening by focusing resources on source materials with a higher likelihood of success, while simultaneously providing opportunities for the discovery of novel ligands to selectively target proteins involved in viral infection.

INTRODUCTION

The majority of life forms on Earth do not possess a complex immune system composed of specialized cells tasked with patrolling for pathogens, mounting a direct intervention, and adapting for future exposures. Instead, these plants, microbes, and animals rely on the biosynthesis of small molecules to control potential pathogens. These broad-spectrum small molecules can directly interact with macromolecules essential to the pathogen’s survival, thus controlling replication of the infectious organism.1,2 The antiviral nucleoside analogues spongosine, spongorthymidine and spongouridine are examples that were isolated from the Caribbean sponge Tectitethya crypta, in the 1950s. Their discovery inspired the development of thousands of related molecules, such as vidarabine, that have been employed with great clinical success in controlling viral infections caused by human immunodeficiency viruses (HIV), herpes simplex viruses, and hepatitis B/C viruses (HBV, HCV),...
HCV). More recently, molnupiravir, an orally available prodrug of N^4-hydroxycytidine, has emerged as a promising control for SARS-CoV-2 in humans by reducing transmission of the virus, although it may have mutagenic effects. With this history of success in mind, additional classes of antiviral agents would significantly benefit future responses to emerging viral pathogens and could be used to limit outbreaks of new viral strains from growing into pandemics. The methodological framework described in this review article is designed to utilize the growing body of information that has accumulated in the past century of NP isolation and characterization and apply it to contemporary drug discovery and development strategies. A critical opportunity exists to explore the structural diversity and complexity of NP libraries without having the actual molecule in hand through computational analysis. Artemisinin, an essential control for malaria infections and the basis for the 2015 Nobel Prize, exemplifies a NP prototype that has had a tremendous impact on human health and would not have been available through other approaches due to its unusual structural features. The approach outlined in this review article can be employed to address the need for new broad-spectrum natural product prototypes for pathogenic Betacoronavirus spp. as well as other families of viruses that present risks for future pandemic outbreaks.

### BROAD-SPECTRUM POTENTIAL FOR THE GENUS BETACORONAVIRUS

Whole genome sequencing shows SARS-CoV-2 has 79% similarity with SARS-CoV and 50% with MERS-CoV. However, the catalytic domains of the proteins responsible for RNA synthesis and replication are highly conserved (Figure 1). These proteins present attractive targets for the identification of drugs active across the entire genus Betacoronavirus. The application of computer-aided drug design tools allows for the inexpensive and rapid evaluation of NP structures in published databases and the scientific literature for in silico screening. Scaffolds that are potential drug leads can then be identified for structure–activity-relationship (SAR) analysis, isolation, or synthesis. Eleven coronavirus proteins regulating RNA synthesis and replication, i.e., SARS-CoV-2: nsp16-nsp10 (2′ O-MTase; PDB ID: 6W4H), 3CLPro (main protease; PDB ID: 6LU7), PLpro (papain-like protease; PDB ID: 7CJM), and RdRp (RNA-dependent RNA polymerase; PDB ID: 7BV2); SARS-CoV: nsp16-nsp10 (PDB ID: 3R24), 3CLPro (PDB ID: 3V3M), PLpro (PDB ID: 3E9S); RdRp (PDB ID: 6NUR); and MERS-CoV: nsp16-nsp10 (PDB ID: SYNB), 3CLpro (PDB ID: 4RSP), PLpro (PDB ID: 5W8U); have been identified as targets for broad-
spectrum therapeutics. Together, these molecular targets form a valuable screening panel for the identification of molecules with potential activity across species of the genus Betacoronavirus. The S-adenosylmethionine (SAM)-dependent methyltransferases (2′-O-MTase; nsp16 nsp10 protein) regulates the O-methylation in coronavirus, which is key for efficient viral RNA translation and evading the host’s immune system. The 3CLpro (also known as 3C-like protease) and RdRp are important proteins in mediating RNA synthesis and replication. PLpro plays an important role in RNA replication by cleaving the N-terminus of the replicase polyprotein. The crystal structures and substrate binding activities of these target proteins are freely available from the RCSB Protein Data Bank. An important aspect supporting the development of small-molecule inhibitors that target key enzymes from pathogenic members of Betacoronavirus, which include proteases, methyl transferases, and RNA-dependent polymerases, is that these drugs would provide potential broad-spectrum controls for SARS-CoV, SARS-CoV-2, and MERS as well as future infectious strains, thus providing options for in vivo prioritization. Overall, the goal is to narrow the number of molecules selected for more intensive in vitro evaluation with live pathogens. However, these two bottlenecks are addressed in the present framework through iterative in silico prioritization. Overall, the goal is to narrow the number of molecules selected for more intensive resource investment or the isolation of compounds for in vitro, and later in vivo, validation. However, two scenarios can arise that could slow progress, the first being that pure standards for candidate compounds are not readily obtainable. The second is a lack of access to the high-level BSL facilities required for in vitro evaluation with live pathogens. However, these two bottlenecks are addressed in the present framework through the proposed use of MS and molecular ion networking (MoIN). Advances in MS-based affinity screening and subsequent data analysis over the past decade can now facilitate the evaluation of ligand interactions beginning with complex mixtures of molecules, such as NP extracts. Molecular docking has been used successfully to identify and advance multiple generations of HIV integrase (e.g., raltegravir) and protease inhibitors (e.g., ampranavir) and to repurpose existing drugs for new diseases. Additionally, recent examples of MS-based ligand screening have demonstrated numerous methodologies
successfully identifying NP ligands for target proteins directly from extracts and have shed light on additional potential applications. Together, in silico molecular modeling and MS-based affinity validation can be harnessed to expedite the identification of new and promising natural products to control SARS-CoV-2 and future emerging members of the Betacoronavirus genus by targeting key proteins with significant homology across the entire genus. This framework can similarly be employed to address other pathogens of concern as they arise. The productivity of such efforts is enabled by utilizing the significant expansions of well-annotated NP libraries that now contain hundreds of thousands of well-defined extracts, fractions, and pure molecules from tens of thousands of species that can be analyzed by MS-based high-throughput screens (HTSS). These resources provide unique opportunities to analyze complex chemical mixtures in search of novel ligands for target proteins, quantify affinity of specific NPs with the target protein, and drive directed biological studies on promising compounds.

**NATURAL PRODUCT RESOURCES**

There are two key resources that are valuable for the described approach, databases of known NPs and NP extract repositories. Over the years a variety of NP databases have been developed and provide valuable information regarding chemical structures and published sources for a metabolite of interest. However, these databases cannot provide structures for molecules that have not yet been characterized, hence the need for complementary approaches involving MS-based protein–ligand interactions. A contemporary and notable example is the Collection of Open NPs (COCONUT), which has aggregated all open source NP structures and provides a searchable web interface that allows access to over 400 000 unique NP molecules with multiple data export options. Resources like these provide a valuable mechanism for utilizing the wealth of historical NP characterization efforts in modern computational screening approaches. In addition, accessible and well-annotated collection efforts coupled with automated extraction, fractionation, weighing, and preparation are allowing researchers from around the world and with diverse expertise to screen NP extract libraries for potential drug leads. Recent advances of such ongoing effort are best exemplified by the National Cancer Institute (NCI) of the United States Cancer Moonshot-funded NCI Program for Natural Product Discovery, which is nearly halfway to its goal of producing 1 million partially purified NP fractions that are freely available in 384-well plates. In addition to these efforts, a variety of similar NP libraries have been developed with differing foci on organisms (plants, marine invertebrates, or microorganisms) and geographic variation. Continued development of extract libraries, high-throughput NP chemistry approaches, and corresponding databases will greatly enhance the diversity of compounds available for screening to the antiviral drug discovery community. As these libraries continue to expand, it will be essential to maintain compatibility with modern screening technologies, such as MS-based protein–ligand interactions.

**MASS SPECTROMETRY-BASED SCREENING**

Analytical methods for evaluating the interactions between small molecules and biologically relevant proteins are critical for discerning a potential therapeutic compound’s utility and has been made easier by the wide commercial availability of major drug targets. Currently, MS ligand affinity and specificity screening methods are available to evaluate both covalent and noncovalent interactions, which is the result of decades of advances in “soft ionization techniques". Such methods have direct applications for testing proposed target proteins for the control of infectious diseases, establishing ligand binding, and providing information regarding binding affinities and the nature of the binding site interactions. Importantly, these analyses can be used both for established ligands before or after bioassay evaluation and for screening of compound libraries and NP extracts. This advantage facilitates the validation of in silico binding affinities of known NPs without having to isolate the compound from the source material when the pure compound is not commercially available, but the source organism, or a closely related species, is obtainable. These techniques can also be used as counter-screens for cell-based assays to reduce the likelihood of being misled by pan-assay interference compounds (PAINS) present in many screening libraries, thereby saving a great deal of time in hypothesis testing and for characterizing the nature of the protein–metabolite interaction. For example, the specificity of the protein binding can be evaluated by counter-screening any target ligands for their ability to bind to general nontarget proteins. It is important to realize that the elimination of NP molecules due to a PAINS assessment may be counter-productive in regard to eliminating molecules that could be easily chemically modified and optimized to eliminate off-target interference. The high sensitivity of the MS instrumentation lends itself valuable to reducing the amount of target protein and sample needed for analysis and greatly expanding the number of low-abundance molecules detectable. These characteristics mitigate two potential bottlenecks of this framework, the availability of target protein and the need for pre-MS fractionation of NP extracts. However, these aspects should be considered in order to reduce the cost of screening for difficult to express and produce proteins and increase the throughput by reducing the extent of NP extract fractionation and preparation required. Additionally, dereplication efforts are greatly aided by MoIN tools when coupled with MS-fragmentation pattern databases, such as the Global NP Social Molecular Networking (GNPS) Spectral Library, which can aid in putative structure identification by comparing spectra of known structural classes. Such computational approaches were utilized to identify the structures of aleutianamine and the karlotoxins, highly potent and selective NP inhibitors of pancreatic and lung cancer, respectively, and informed the discovery of their natural analogues. Additional biophysical approaches that are being adapted to NPs include differential scanning fluorimetry (DSF). DSF is a relatively inexpensive instrument and is useful in assessing binding interactions between small molecules and macromolecules including proteins, DNA, and RNA. In addition, DSF can distinguish PAINS from more desirable molecules. An assessment of interactions with DSF can be completed with as little as 7 μL. Importantly, by freeing researchers of the prerequisite of maintaining high-BSL facilities, these tools greatly increase the number of laboratories capable of screening for infectious disease targets, improving the global outlook for the development of emerging infectious disease controls.
CONCLUSION

The international effort to produce high-resolution structural models of SARS-CoV-2 target proteins has allowed for the rapid computational screening of compounds from digital libraries and the scientific literature. Researchers have been rushing to find suitable antiviral treatments using existing approved drugs, which has led to the identification of prospects with in vitro activity, but thus far, few have provided a clinically significant control. Small-molecule therapeutic approaches have been highly successful for the control of HIV-1 and HCV and could be applied to control SARS-CoV-2 and other members of the Betacoronavirus genus. The framework described here provides a guide for the identification of NP prototypes for the generation of diverse classes of antivirals which could be utilized as a preemptive effort for the control of novel strains that evade vaccine coverage as well as viruses for which no effective vaccine exists. The emergence of threatening viral pathogens, including novel coronavirus strains and species, is inevitable; however, their progression to pandemic scale and the resultant loss of life and economic challenges are highly preventable. Guidelines and resources for antiviral development at NIAID are available at www.niaid.nih.gov/research/antiviral-development.

AUTHOR INFORMATION

Corresponding Author
Mark T. Hamann — Departments of Drug Discovery and Biomedical Sciences and Public Health, Colleges of Pharmacy and Medicine, Medical University of South Carolina, Charleston, South Carolina 29425, United States; Email: hamannm@musc.edu

Authors
George S. Hanna — Departments of Drug Discovery and Biomedical Sciences and Public Health, Colleges of Pharmacy and Medicine, Medical University of South Carolina, Charleston, South Carolina 29425, United States; orcid.org/0000-0001-7616-064X
Yeun-Mun Choo — Department of Chemistry, University of Malaya, Kuala Lumpur 50603, Malaysia
Ryan Harbit — College of Charleston, Charleston, South Carolina 29425, United States
Heather Paeth — Departments of Drug Discovery and Biomedical Sciences and Public Health, Colleges of Pharmacy and Medicine, Medical University of South Carolina, Charleston, South Carolina 29425, United States
Sarah Wilde — Department of Biology, Clemson University, Clemson, South Carolina 29631, United States
James Mackle — School of Biological Sciences & Institute for Global Food Security, Queens University, Belfast, Northern Ireland BT7 1NN, United Kingdom
Jacopo-Umberto Verga — School of Biological Sciences & Institute for Global Food Security, Queens University, Belfast, Northern Ireland BT7 1NN, United Kingdom
Bethany J. Wolf — Department of Public Health Sciences, Medical University of South Carolina, Charleston, South Carolina 29425, United States
Olivier P. Thomas — Marine Biodiscovery, School of Chemistry and Ryan Institute, National University of Ireland Galway, Galway H91Tk33, Ireland; orcid.org/0000-0002-5708-1409
Peter Croot — Irish Centre for Research in Applied Geoscience, Earth and Ocean Sciences and Ryan Institute, School of Natural Sciences, National University of Ireland, Galway, Galway H91Tk33, Ireland; orcid.org/0000-0003-1396-0601
James Gray — Department of Biomedical Education and Anatomy, College of Medicine and Division of Biosciences, College of Dentistry, The Ohio State University, Columbus, Ohio 43210, United States
Courtney Thomas — Department of Chemistry, South Carolina State University, Orangeburg, South Carolina 29117, United States
Ling-Zhi Li — School of Traditional Chinese Materia Medica, Shenyang Pharmaceutical University/SPU, Shenyang 110016, China; orcid.org/0000-0001-7616-064X
Gary Hardiman — School of Biological Sciences & Institute for Global Food Security, Queens University, Belfast, Northern Ireland BT7 1NN, United Kingdom; orcid.org/0000-0003-4558-0400
Jin-Feng Hu — School of Advanced Study, Zhejiang Provincial Key Laboratory of Plant Ecology and Conservation, Taizhou University, Zhejiang 318000, China; orcid.org/0000-0002-0367-1454
Xiaojuan Wang — School of Pharmacy, Lanzhou University, Lanzhou 730300, China

Funding
G.S.H. acknowledges support from NIH/NCCIH 1F31AT011158-01 for Fellowship funding. M.T.H. acknowledges support from the South Carolina SmartState Program and The Abney Foundation. G.H. acknowledges support from NIH U54MD010706, U01DA045300, and QB start-up funds. J.M. was supported by an NI Department for the Economy (DeI) Research Studentship, J.U.V. was the recipient of an ERASMUS Student Mobility Traineeship. Part of this project (Grant-Aid Agreement Nos. PBA/MB/16/01 and PD/OC/19/02/01) was carried out with the support of the Marine Institute and is funded under the Marine Research Program by the Irish Government. R.F.S. is supported in part by the Center for AIDS Research at Emory University (P30-AI-050409).

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Gogineni, V.; Schinazi, R. F.; Hamann, M. T. Chem. Rev. 2015, 115 (18), 9655–9706.
(46) Vu, H.; Pham, N. B.; Quinn, R. J. J. Biomol. Screening 2008, 13 (4), 265−275.
(47) Fu, Y.; Luo, J.; Qin, J.; Yang, M. J. Pharm. Biomed. Anal. 2019, 168, 189−200.
(48) Ren, C.; Bailey, A. O.; VanderPorten, E.; Oh, A.; Phung, W.; Mulvihill, M. M.; Harris, S. F.; Liu, Y.; Han, G.; Sandoval, W. Anal. Chem. 2019, 91 (1), 903−911.
(49) Thomford, N. E.; Senthebane, D. A.; Rowe, A.; Munro, D.; Seele, P.; Maroyi, A.; Dzobo, K. Natural Products for Drug Discovery in the 21st Century: Innovations for Novel Drug Discovery. Int. J. Mol. Sci. 2018, 19 (6), 1578.
(50) Simithy, J.; Fuanta, N. R.; Alturki, M.; Hobrath, J. V.; Balunas, M. J.; Bunsko, D.; Neto, F. C.; Castaño-Espriu, L.; Chang, C.; Clark, T. N. ACS Cent. Sci. 2019, 5 (11), 1824−1833.
(51) Grkovic, T.; Akee, R. K.; Thornburg, C. C.; Trinh, S. K.; Britt, J. R.; Harris, M. J.; Evans, J. R.; Kang, U.; Ensel, S.; Henrich, C. J. ACS Chem. Biol. 2020, 15 (4), 1104−1114.
(52) Simon, R. P.; Winter, M.; Kleiner, C.; Wehrle, L.; Karnath, M.; Ries, R.; Zeeb, M.; Schnapp, G.; Fiegen, D.; Häbe, T. T.; Runge, F.; Bretscheider, T.; Li, S.; Shui, W. Curr. Opin. Biotechnol. 2020, 64, 24−31.
(53) Nothias, L.-F.; Nothias-Esposito, M. l.; da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Cachet, X. ACS Chem. Biol. 2017, 12 (10), 2644−2651.
(54) Waters, A. L.; DiPaola, R. S.; Dutch, R. E. Science 2020, 368 (6493), 829−830.
(55) Yoshimoto, F. K. Protein J. 2020, 39 (3), 198−216.
(56) Yin, W.; Mao, C.; Luan, X.; Shen, D. D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; Chang, S.; Xie, Y. C.; Tian, G.; Jiang, H. W.; Tao, S. C.; Shen, J.; Jiang, Y.; Jiang, H.; Xu, Y.; Zhang, S.; Zhang, Y.; Xu, H. E. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. Science 2020, 368, 4338−4344.