

Рациональные подходы создания высокоизбирательного и эффективного ингибитора главной протеазы SARS-CoV-2

Булыгин А.^{1*}, Кузнецов Н.^{1, 2}

¹ Институт химической биологии и фундаментальной медицины СО РАН, Новосибирск, Россия

² Факультет естественных наук, Новосибирский государственный университет, Новосибирск, Россия

* abulygin@niboch.nsc.ru

Ключевые слова: SARS-CoV-2; главная протеаза; молекулярное моделирование; ингибитор

Мотивация и цель: Пандемия COVID-19, вызванная коронавирусом SARS-CoV-2, стала третьим случаем коронавирусной инфекции. Вместе с другими двумя вирусами, SARS-CoV и MERS-CoV, вирус SARS-CoV-2 является высокопатогенным и имеет уровень смертности более 1 %. Более того, стало понятным, что коронавирус часто мутирует, что снижает эффективность доступных вакцин и вынуждает регулярно создавать новые.

Вирусные полимеразы и протеазы являются подходящими мишенями для лекарств прямого действия. На данный момент существует три одобренных противокоронавирусных лекарства: ремдесивир [1], молнупиравир [2] и нирматрелвир [3]. Однако оба лекарства, направленных против полимеразы, не рекомендованы для использования из-за незначительной эффективности в случае с ремдесивиром [4] и потенциального увеличения частоты мутаций вируса в случае с молнупиравиром [2]. Нирматрелвир, ингибитор главной протеазы (M^{pro}), стал более успешным, его начали выпускать в форме таблеток. Он имеет хорошую биодоступность – доля усваиваемого препарата равна примерно 50 %, но растворимость соединения находится на умеренном уровне (около 1 мг/мл) [3]. Таким образом, по-прежнему остается актуальным создание высокоизбирательного и эффективного ингибитора протеазы, который можно принимать перорально.

В данной работе мы провели глубокий литературный обзор *in vitro* и *in vivo* исследований ингибиторов M^{pro} и создали набор возможных ингибиторов и их частей, которые предположительно позволяют достичь всех необходимых свойств, а именно высокого сродства к вирусному ферменту, избирательности, биодоступности и растворимости.

Методы и алгоритмы: Молекулярно-динамическое моделирование было проведено в пакете программ GROMACS 2020.6 в сочетании с силовым полем AMBER ff99SB-ILDN [5]. Свободные энергии связывания комплексов были вычислены сочетанием методов молекулярной механики и Пуассон-Больцмановской площади поверхности. Для вычисления этим методом был использован скрипт “g_mmpbsa” для формата файлов GROMACS [6]. Онлайн-сервис SwissADME был использован для расчета лекарственных характеристик: всасываемости, распределения, метаболизма и выделения [7].

Результаты: Было проведено МД-моделирование комплексов основной протеазы коронавируса с несколькими известными ингибиторами с различным уровнем эффективности: PF-0835231 [8], нирматрелвир, MPI3 [9], GC373 [10], боцепревир и телапревир [11]. Анализ образуемых водородных связей между данными

ингибиторами и аминокислотными остатками активного центра M^{pro} показал, что наибольшее среднее по времени значение водородных связей имеет PF-0835231 – 6.6 связи, тогда как максимально возможное для него значение равно 9. Энергия связывания PF-0835231 в активном центре фермента равна -141 ± 13 кДж/моль. Благодаря глубокому проникновению в связывающий центр фермента, наилучшую энергию связывания (-189 ± 14 кДж/моль) имеет соединение МРІЗ. Однако IC_{50} для МРІЗ (8 нМ) хуже, чем для PF-0835231 (0.3 нМ), что может объясняться высокой гибкостью молекулы МРІЗ, снижающей эффективность образования комплекса с ферментом. Недостатком высокоэффективного ингибитора PF-0835231 является его неудовлетворительная растворимость (менее 0.1 мг/мл), которая может быть повышена фосфорилированием молекулы. Однако такая модификация приводит к практически полной потере биодоступности [3]. Согласно полученным данным, нирматрелвир является третьим по энергии связывания (-138 ± 13 кДж/моль), имея $IC_{50} = 3$ нМ. Таким образом, можно заключить, что свободная энергия связывания соединений в активном центре фермента на уровне -140 кДж/моль и ниже может обеспечить эффективное ингибирование протеазы, однако индивидуальные структурные особенности ингибитора могут иметь существенное значение для создания эффективного и селективного препарата.

Анализ полученных данных позволяет предположить, что ингибитор M^{pro} должен обладать пятью главными качествами, придающими высокий уровень связывания и избирательности (рис. 1): 1) заместитель S0 должен быть небольшим; 2) заместитель S1 должен быть похож на глутамин, имея одного донора и одного акцептора водородной связи; 3) боковая цепь S2 должна быть гидрофобной и иметь длину в три-четыре С-С связи; 4) для повышения избирательности по отношению к M^{pro} участок S2 должен иметь структуру боковой цепи, жестко связанную с остовом (например, бициклопролин), или должен быть жестким и объемным заместителем S4 (например, индол, содержащий заместитель в четвертом положении); 5) заместитель S4 должен повторять форму кармана S4 связывающего центра. При этом значительных ограничений структуры заместителя в положении S3 не было обнаружено.

Формулирование таких критериев для эффективного ингибитора позволяет пойти двумя путями улучшения свойств предложенных ранее соединений. Один из них заключается в создании аналогов PF-0835231, обладающих хорошей растворимостью без потери эффективности связывания. Для этого в ходе данной работы были проанализированы аналоги PF-0835231, несущие различные модификации в положениях S0, S1 и S4, затем для всех модифицированных соединений рассчитывали растворимость с помощью SwissADME и наконец проводили моделирование комплекса созданных ингибиторов с ферментом. Лучший результат показало соединение, в котором в качестве S4-заместителя был фторированный индол, а в качестве S1 – остаток сукцинимид (рис. 2). Кроме того, гидроксиметилкетонная группа, отвечающая за образование ковалентного комплекса с ферментом, была заменена на альдегидную (на рис. 2 выделена

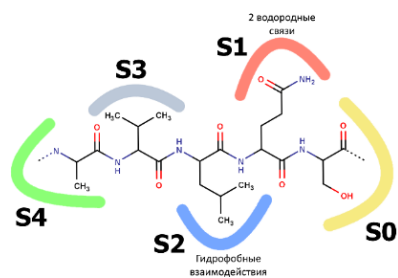


Рис. 1. Общая структура субстрата M^{pro}

кружком), так как в литературе было показано, что альдегидный вариант обладает большей биодоступностью [12]. Итоговое соединение имело энергию связывания -156 ± 13 кДж/моль, что на 15 кДж/моль лучше значения для PF-0835231. При этом количество образуемых водородных связей в активном центре не изменилось относительно PF-0835231, расчетная растворимость улучшилась примерно в два раза (с 0.1 до 0.2 мг/мл), а биодоступность стала высокой согласно расчету SwissADME. Данные характеристики позволяют предположить, что предложенное соединение может обладать улучшенной эффективностью действия в сравнении с исходным ингибитором PF-0835231. Еще одним способом улучшения растворимости любого рассмотренного ингибитора оказалась замена пирролидона в положении S1 на имидазолидон (см. рис. 2). Согласно SwissADME, это изменение ведет к улучшению растворимости в 3–5 раз в зависимости от алгоритма расчета растворимости. Данная модификация большинства рассмотренных соединений, по результатам моделирования, не приводит к изменениям ни в положении ингибиторов в активном центре, ни в количестве образуемых водородных связей, ни в величине энергии связывания.

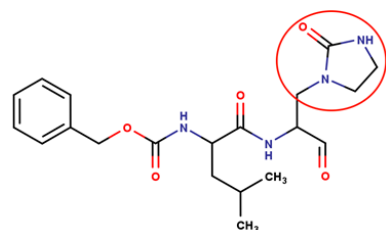


Рис. 2. Пример структуры соединения с имидазолидоном (выделен кружком) в качестве S1

Второй путь улучшения известных ингибиторов заключается в локальном изменении структуры ингибитора для увеличения эффективности связывания с сохранением хорошей растворимости. Перспективной основой для работы в данном направлении может выступить молекула нирматрелвира, поскольку ее растворимость уже находится на приемлемом уровне (1 мг/мл). Дальнейшее увеличение этого параметра может сопровождаться нежелательным снижением биодоступности, которая не может быть предсказана точно имеющимися в данный момент алгоритмами расчета. Поэтому мы решили попытаться улучшить связывание нирматрелвира с ферментом. Остаток бициклопролина, входящий в нирматрелвир и повышающий избирательность к M^{pro} (пункт 4 критериев эффективности), открывает возможность использования широкого набора S4-заместителей. Наилучшие результаты в качестве заместителя в положении S4 показал остаток тиафена (рис. 3), приводя к понижению свободной энергии связывания на 42 кДж/моль, до -180 ± 15 кДж/моль, в то время как растворимость модифицированного соединения согласно расчету SwissADME практически не изменилась.

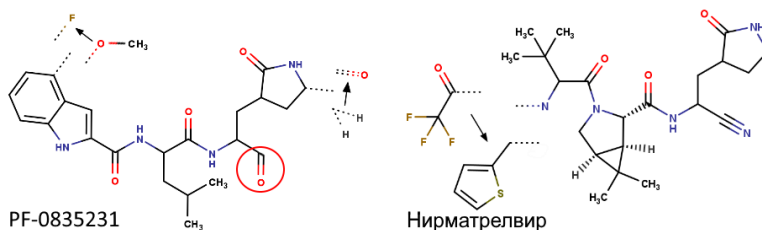


Рис. 3. Структуры PF-0835231 и нирматрелвира с обозначением наилучших модификаций

Вывод: В данной работе методами молекулярного моделирования проведен структурно-функциональный анализ ингибиторов главной протеазы коронавируса и предложены способы улучшения растворимости и эффективности связывания соединений на основе структурно-модифицированных аналогов известных ингибиторов.

Финансирование: Исследование выполнено в рамках государственного задания ИХБФМ СО РАН (№ 121112900214-2).

Approaches to rational design of a highly selective and potent SARS-CoV-2 main protease inhibitor

Bulygin A.^{1*}, Kuznetsov N.^{1, 2}

¹ Institute of Chemical Biology and Fundamental Medicine, SB RAS, Novosibirsk, Russia

² Department of Natural Sciences, Novosibirsk State University, Novosibirsk, Russia

* abulygin@niboch.nsc.ru

Key words: SARS-CoV-2; main protease; inhibitor; molecular modelling; rational design

Motivation and Aim: COVID-19 caused by the SARS-CoV-2 is the third case of coronavirus pandemics. Along with the other two, SARS-CoV and MERS-CoV, it is highly pathogenic and has mortality rates more than 1 %. In addition, it became apparent that coronavirus has high mutability that decrease the effectiveness of available vaccines and makes necessary to regularly develop new ones.

The viral polymerase and proteases are applicable as direct action targets of anticoronavirus drugs. Currently there are three approved anti-coronavirus drugs: remdesivir [1], molnupiravir [2] and nirmatrelvir [3]. However, both polymerase-directed drugs are not recommended for wide use because of negligible efficacy in the case of remdesivir [4] or potentially increased mutability of the virus in the case of molnupiravir [2]. Nirmatrelvir, the inhibitor of the main protease (M^{pro}), has become more successful being produced in tablet form. It has good bioavailability (about 50 %) but only moderate solubility under 1 mg/ml [3]. Thus, there is still a need in a highly selective and potent M^{pro} inhibitors, that can be administered orally.

In this work we conducted a deep review of *in vitro* and *in vivo* studies of M^{pro} inhibitors and designed a set of parts of possible inhibitors, which presumably achieved all necessary properties, namely binding ability, selectivity, bioavailability and solubility.

Methods and Algorithms: Molecular dynamics simulations were performed with GROMACS 2020.6 and the AMBER ff99SB-ILDN force field [5]. Binding free energies of complexes were calculated by combining MD data with the molecular mechanic/Poisson-Boltzmann surface area method. The “g_mmpbsa” script for GROMACS data [6] was used to calculate the free energies of MM-PBSA. SwissADME online tool was used for prediction of ADME properties [7].

Results: MD simulations of several known inhibitors with different inhibition efficiency were conducted. The chosen inhibitors were: PF-0835231 [8], nirmatrelvir, MPI3 [9], GC373 [10], boceprevir and telaprevir [11]. The analysis of the H-bonds formed between the inhibitors and M^{pro} showed that the highest time average number of H-bonds has PF-0835231 – 6.6 bonds with a maximum of 9. MPI3 has the best binding energy (-189 ± 14 kJ/mol) thanks to penetration deep into the active site, PF-0835231 is on the second place (-141 ± 13 kJ/mol). However, MPI3 has worse IC_{50} than that of PF-

0835231 (8 nM against 0.3 nM), that may be explained by the high flexibility with a lot of rotational degrees of freedom that is considered as a bad factor for binding. On the other hand, PF-0835231 has an inappropriate solubility (less than 0.1 mg/ml), that can be increased by phosphorylation, which leads to an extremely low oral bioavailability [3]. Nirmatrelvir takes third place in binding with binding energy of -138 ± 13 kJ/mol, while its $IC_{50} = 3$ nM.

Taking together all the knowledge it could be stated that M^{pro} inhibitor should have five main properties (Fig. 1): (1) small S0-residue; (2) glutamine-like S1-residue with one H-bond acceptor and one H-bond donor; (3) S2-residue with the length of 3 or 4 C-C bonds; (4) a bulky part of a backbone (like a bicycloproline) or bulky and rigid S4-residue (like an indole with substituents) for selectivity; (5) S4-residue repeating the shape of the enzyme's S4-pocket for binding.

Formulating these criteria allows us to go two ways to improve known inhibitors. First, we tried to improve the solubility of PF-0835231 without decrease in binding. The best result showed a molecule, in which we used fluoride indole as S4-substituent and succinimide as S1 (Fig. 2). We also changed S0-substituent intended for covalent binding ("warhead") from hydroxymethylketone to an aldehyde because it was shown that the inhibitors with aldehyde warhead have better bioavailability [12]. The final compound has binding energy of -156 ± 13 kJ/mol that is by 15 kJ/mol less than that of PF-0835231. Number of H-bonds didn't change and SwissADME service showed that these changes improved the solubility twofold (from approximately 0.1 to 0.2 mg/ml) and bioavailability became high.

Another way to improve the solubility of any inhibitor is to replace a pyrrolidone S1-moiety by an imidazolidone one (Fig. 2). According to SwissADME this change makes the solubility 3–5 times higher depending on the solubility calculation algorithm. We modeled most of the compounds with this modification. Results showed that there are no changes in position, binding energy or a number of H-bonds for all tested compounds.

The second way to improve known inhibitors is to improve binding ability while maintaining the solubility on the same level. Nirmatrelvir is a good basis for this. This compound already has an acceptable solubility (about 1 mg/ml), but, on the other hand, further increase of this parameter may lead to bioavailability decrease, which cannot be well predicted yet. Thus, we made an attempt to improve binding of nirmatrelvir. Bicycloproline, which confers selectivity by itself, opens up the possibility of using much wider range of S4-substituents. Best results showed the compound with a small aromatic thiophene as S4 (Fig. 3). It has binding energy of -180 ± 15 kJ/mol that is 42 kJ/mol lower than that of nirmatrelvir, while its solubility almost didn't change.

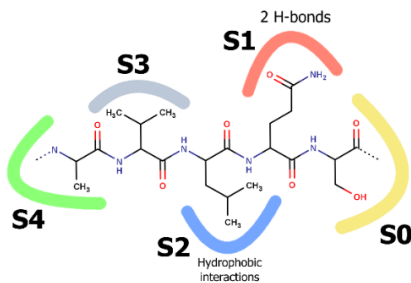


Fig. 1. The general structure of M^{pro} substrate

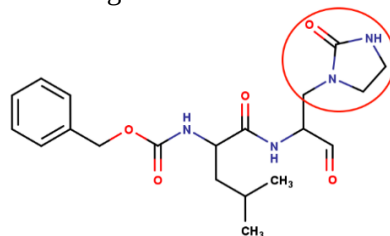


Fig. 2. Example structure of a compound with an imidazolidone S1-substituent

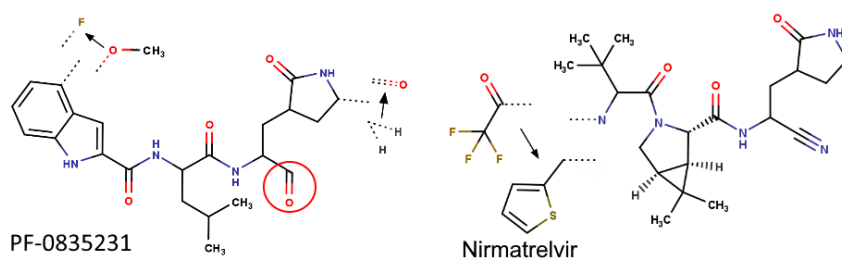


Fig. 3. Structures of PF-0835231 and nirmatrelvir with modifications that give the greatest improvement

Conclusions: Thus, in this work we conducted structure-function analysis of SARS-CoV-2 main protease inhibitors using molecular modeling methods and propose approaches of improving the solubility and binding ability of compounds based on structurally modified analogues of known inhibitors.

Funding: This study was supported by the Russian state-funded project for ICBFM SB RAS (No. 121112900214-2).

Список литературы/References

1. Siegel D., Hui H.C., Doerffler E., Clarke M.O. et al. Discovery and Synthesis of a Phosphoramidate Prodrug of a Pyrrolo[2,1- f] [Triazin-4-Amino] Adenine C-Nucleoside (GS-5734) for the Treatment of Ebola and Emerging Viruses. *J Med Chem.* 2017;60:1648-1661. doi 10.1021/acs.jmedchem.6b01594
2. Cox R.M., Wolf J.D., Plemper R.K. Therapeutically Administered Ribonucleoside Analogue MK-4482/EIDD-2801 Blocks SARS-CoV-2 Transmission in Ferrets. *Nat Microbiol.* 2021;6:11-18. doi 10.1038/s41564-020-00835-2
3. Owen D.R., Allerton C.M.N., Anderson A.S., Aschenbrenner L., Avery M., Berritt S., Boras B., Cardin R.D. et al. An Oral SARS-CoV-2 M^{pro} Inhibitor Clinical Candidate for the Treatment of COVID-19. *Science.* 2021;374:1586-1593. doi 10.1126/science.abl4784
4. Sheahan T.P., Sims A.C., Zhou S., Graham R.L., Pruijssers A.J., Agostini M.L., Leist S.R., Schäfer A., Dinno K.H. et al. An Orally Bioavailable Broad-Spectrum Antiviral Inhibits SARS-CoV-2 in Human Airway Epithelial Cell Cultures and Multiple Coronaviruses in Mice. *Sci Transl Med.* 2020;12(541):eabb5883. doi 10.1126/scitranslmed.abb5883
5. Abraham M.J., Murtola T., Schulz R., Páll S., Smith J.C., Hess B., Lindahl E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX.* 2015;1-2:19-25. doi 10.1016/j.softx.2015.06.001
6. Kumari R., Kumar R., Lynn A. G_mmpbsa – A GROMACS Tool for High-Throughput MM-PBSA Calculations. *J Chem Inf Model.* 2014;54:1951-1962. doi 10.1021/ci500020m
7. Daina A., Michielin O., Zoete V. SwissADME: A Free Web Tool to Evaluate Pharmacokinetics, Drug-Likeness and Medicinal Chemistry Friendliness of Small Molecules. *Sci Rep.* 2017;7:42717. doi 10.1038/srep42717
8. Hoffman R.L., Kania R.S., Brothers M.A., Davies J.F. et al. Discovery of Ketone-Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic Treatment of COVID-19. *J Med Chem.* 2020;63:12725-12747. doi 10.1021/acs.jmedchem.0c01063
9. Yang K.S., Ma X.R., Ma Y., Alugubelli Y.R., Scott D.A., Vatansever E.C., Drelich A.K., Sankaran B. et al. A Quick Route to Multiple Highly Potent SARS-CoV-2 Main Protease Inhibitors**. *Chem Med Chem.* 2021;16:942-948. doi 10.1002/cmdc.202000924
10. Vuong W., Fischer C., Khan M.B., van Belkum M.J., Lamer T., Willoughby K.D., Lu J., Arutyunova E. et al. Improved SARS-CoV-2 M^{pro} Inhibitors Based on Feline Antiviral Drug GC376: Structural Enhancements, Increased Solubility, and Micellar Studies. *Europ J Med Chem.* 2021;222:113584. doi 10.1016/j.ejmech.2021.113584
11. Anson B.J., Chapman M.E., Lendy E.K., Pshenychnyi S., D'Aquila R.T., Satchell K.J.F., Mesecar A.D. Broad-Spectrum Inhibition of Coronavirus Main and Papain-like Proteases by HCV Drugs. *Res Square.* 2020. doi 10.21203/rs.3.rs-26344/v1
12. Qiao J., Li Y.-S., Zeng R., Liu F.-L., Luo R.-H. et al. SARS-CoV-2 M^{pro} Inhibitors with Antiviral Activity in a Transgenic Mouse Model. *Science.* 2021;371:1374-1378. doi 10.1126/science.abf1611