The role of folic acid in the management of respiratory disease caused by COVID-19

Zahra Sheybani[1,9], Maryam Heydari Dokoohaki[2,9], Manica Negahdaripour[3,4], Mehdi Dehdashti[5], Hassan Zolghadr[6], Mohsen Moghadami[7], Seyed Masoom Masoompour[7], and Amin Reza Zolghadr[2,8*]

[1]Department of Internal Medicine, Aliasghar Hospital, Shiraz University of Medical Sciences, Shiraz, Iran.
[2]Department of Chemistry, Shiraz University, Shiraz, 71946-84795, Iran.
[3]Pharmaceutical Sciences Research Center, Shiraz University of Medical Sciences, Shiraz, Iran.
[4]Department of Pharmaceutical Biotechnology, School of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran.
[5]Microbiology laboratory, Moslemin Hospital, Shiraz, Iran.
[6]Medical School, Shiraz University of Medical Sciences, Shiraz, Iran.
[7]Non-Communicable Diseases Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
[8]Fars Science and technology Park, Shiraz, Iran.
[9]Z.Sh and M.H.D. are co-first authors.

Correspondence should be addressed to A. R. Z. (arzolghadr@shirazu.ac.ir); Tel: +98 713 613 7157, Fax: +98 713 646 0788, ORCID: 0000-0002-6289-3794 (A.R.Z).

1

Entrance of coronavirus into cells happens through the spike proteins on the virus surface, for which the spike protein should be cleaved into S1 and S2 domains. This cleavage is mediated by furin, which can specifically cleave Arg-X-X-Arg$\downarrow$ sites of the substrates. Furin, a member of proprotein convertases family, is moved from the trans-Golgi network to the cell membrane and activates many precursor proteins. A number of pathological conditions such as atherosclerosis, cancer, and viral infectious diseases, are linked with the impaired activity of this enzyme.

Despite the urgent need to control COVID-19, no approved treatment is currently known. Here, folic acid (folate), a water-soluble B vitamin, is introduced for the first time for the inhibition of furin activity. As such, folic acid, as a safe drug, may help to prevent or alleviate the respiratory involvement associated with COVID-19.

Coronaviruses, a family of Coronaviridae, can cause significant human pathologies such as respiratory tract infections in humans and other mammals.[1] Coronavirus infections are usually mild, but some beta coronaviruses including Middle East respiratory syndrome coronavirus (MERS-CoV) and severe acute respiratory syndrome coronavirus (SARS-CoV) may induce critical symptoms.[2,3]

In December 2019, an outbreak of lower respiratory tract infections was reported in Wuhan, China.[4] The pathogen was recognized as a novel RNA beta coronavirus, later named as SARS-CoV-2.[5] The infection caused by this virus, COVID-19, is declared by the World Health Organization (WHO) as a pandemic.[6] In view of SARS-CoV-2 novelty, further researches are required to obtain more insights about its pathogenesis.

Coronavirus (CoV) genome encodes four structural proteins, comprising spike (S), membrane (M), envelope (E), and nucleocapsid (N). The spike (S) protein of coronaviruses mediates receptor binding and fusion of the virus with the target cells.[7] Each class of coronavirus attaches to a specific cellular receptor to facilitate virus entrance into cells. Angiotensin-converting enzyme 2 (ACE2) and CD209L are shown responsible for SARS-CoV entrance.[8-9] It is reported that SARS-CoV-2 enters the respiratory tract by interacting with ACE2 receptor.[10] The spike protein comprises an amino (N)-terminal S1 subunit and a carboxyl (C)-terminal S2 subunit. The entrance of the virus is facilitated by cleavage of S protein to S1/S2 subunits. The S1 subunit binds to the ACE2 receptor, while the S2 site interacts with the cell membrane to mediate receptor-dependent endocytosis[11], as shown in Fig1a. The coronavirus spike protein is cleaved into S1 (receptor binding subunit) and S2 (membrane fusion subunit) by a proteolytic activation at the furin consensus motif $RRRR_{537}\downarrow S$ (R=arginine, $\downarrow$: cleavage site) in virus-infected cells. Additionally, the S2 subunit of the S protein is further cleaved at the second furin site ($RRRR_{690}\downarrow S$) in the infected cells expressing S constructs.[12-16] Mutations of one basic residue in the $RRRR_{690}\downarrow S$ motif and use of non-furin cleavable $PRRR\downarrow S$ sequence demonstrated that furin may play an important role in furin-dependent entry.[17] The working protease is a cellular proprotein convertase that circulates between plasma membrane, early endosome, and trans-Golgi network (TGN), by participation in endocytic and exocytic paths.[18,19] This proprotein convertase is a major candidate for processing the surface glycoproteins of pathogenic

viruses.[20,21] Furin can cleave precursor proteins with specific motifs to produce mature proteins with biological activity. The first (P1) and fourth (P4) amino acids at the N-terminus of the substrate cleavage site must be arginine "Arg-X-X-Arg ↓" (R-X-X-R, X= any amino acid, ↓: cleavage site). If the P2 position is basic lysine or arginine, the cleavage efficiency could be improved by about 10 times.[22] The results of a series of analyses have proposed that one of the important reasons for the high infectivity of COVID-19 is a redundant furin cut site in the virus spike protein.[23]

Our aim is to suggest folic acid as a potential inexpensive, safe, and non-immunogenic drug candidate for the prevention or treatment of early stages of respiratory disease associated with COVID-19 (Fig. 1). Folic acid is a type of B vitamin normally found in foods such as spinach, broccoli, asparagus, dried beans, lentils, peas, and oranges. Folic acid helps the body produce and maintain new cells and also prevent changes to DNA that may lead to cancer. Noticeably, folic acid deficiency is associated with a variety of human malignancies, including colorectal cancer. The over-expression of folate receptors in the early stages of malignant cell formation can be due to folic acid deficiency. Besides, folate malnutrition can cause a high incidence of adenomatous polyps and premalignant lesions of the colon.[24] To this aim, the molecular dynamics (MD) simulations of the interactions of furin enzyme with folic acid and one of its active metabolites, folinic acid, was performed here for the first time to evaluate the interplay of these molecules with furin.
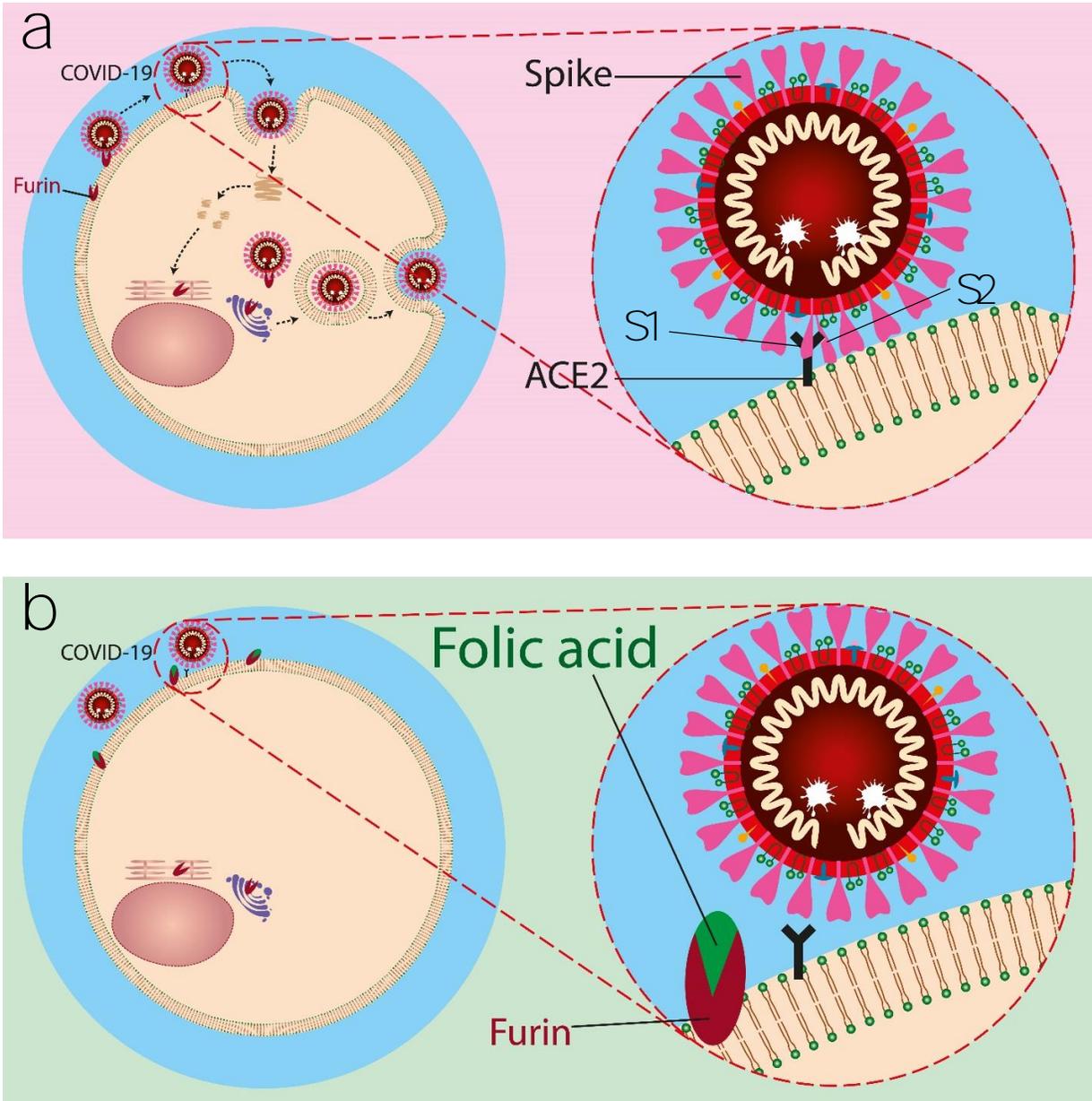
Fig. 1- A schematic representation of inhibitory action of folic acid. a, The mechanism of fusion and replication of COVID-19 virus. **b,** Inhibition of furin protein by folic acid.

The ligand-furin docking was performed to predict the major binding sites of folic acid and folinic acid molecules in the active site of furin protein. In general, the interaction energies between protein and ligands are obtained by docking studies. The obtained binding free energies

were -161.6, and -159.2 kcal/mol for folic acid and folinic acid, respectively. The results showed that folic acid and folinic acid molecules interacted well with the active site residues of furin by formation of hydrogen bonds. Different atom sites of the two drug molecules established hydrogen bonding interactions with various amino acids of furin as shown in Fig. 2a. Interestingly, the binding sites of folinic acid and furin were clearly different. The interactions of Gly307, Glu271, Tyr313, Gln488, Ala532, Arg490, and Asp530 residues with hydrogen, oxygen, and nitrogen atoms of folic acid were dominant; whereas, strong hydrogen bonds were established through Ser311, Glu271, Arg490, Lys449, and Gln488 residues of folinic acid molecules. Moreover, while the H_{19} and H_{12} atoms of folic acid interacted substantially with Glu271, the H_2 and H_4 atoms of folinic acid formed hydrogen bonds with the Glu271 residue of furin. These findings proposed different orientation preferences of folic acid and folinic acid molecules in the binding site of furin.
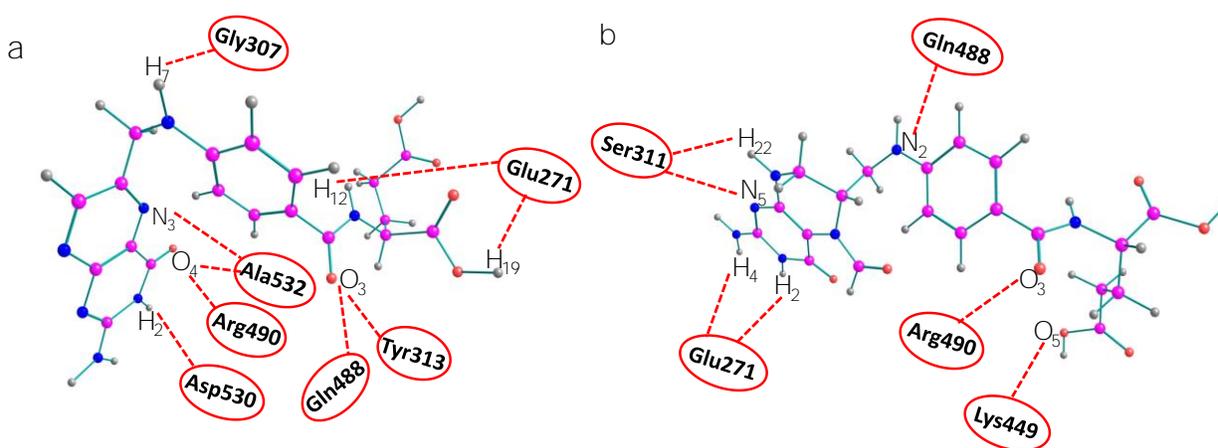
Fig. 2- Hydrogen bond interactions of drugs with furin. a, folic acid. **b,** folinic acid.

The MD simulations of folate derivatives interaction with furin were also conducted to gain additional insight into the specific mechanism by which folic acid and folinic acid molecules can exert their potential inhibitory actions in the furin active sites of COVID-19 patients. Atoms labeling for the two drug molecules under study are shown in Fig 3a. The snapshots of the simulated systems after 100 ns (water molecules were removed for clarity) are depicted in Fig. 3b. The position and orientation of drug molecules in the simulations are in line with docking results. The intermolecular interactions between different residues of furin and drug molecules

were studied here. In this regard, the number of hydrogen bonds between folic acid and furin were obtained to be 3.5 ± 1.3 per time frame, relative to the value of 2.4 ± 1.5 hydrogen bonds per time frame for folinic acid. This is in accordance with the interaction free energies obtained by docking simulations.
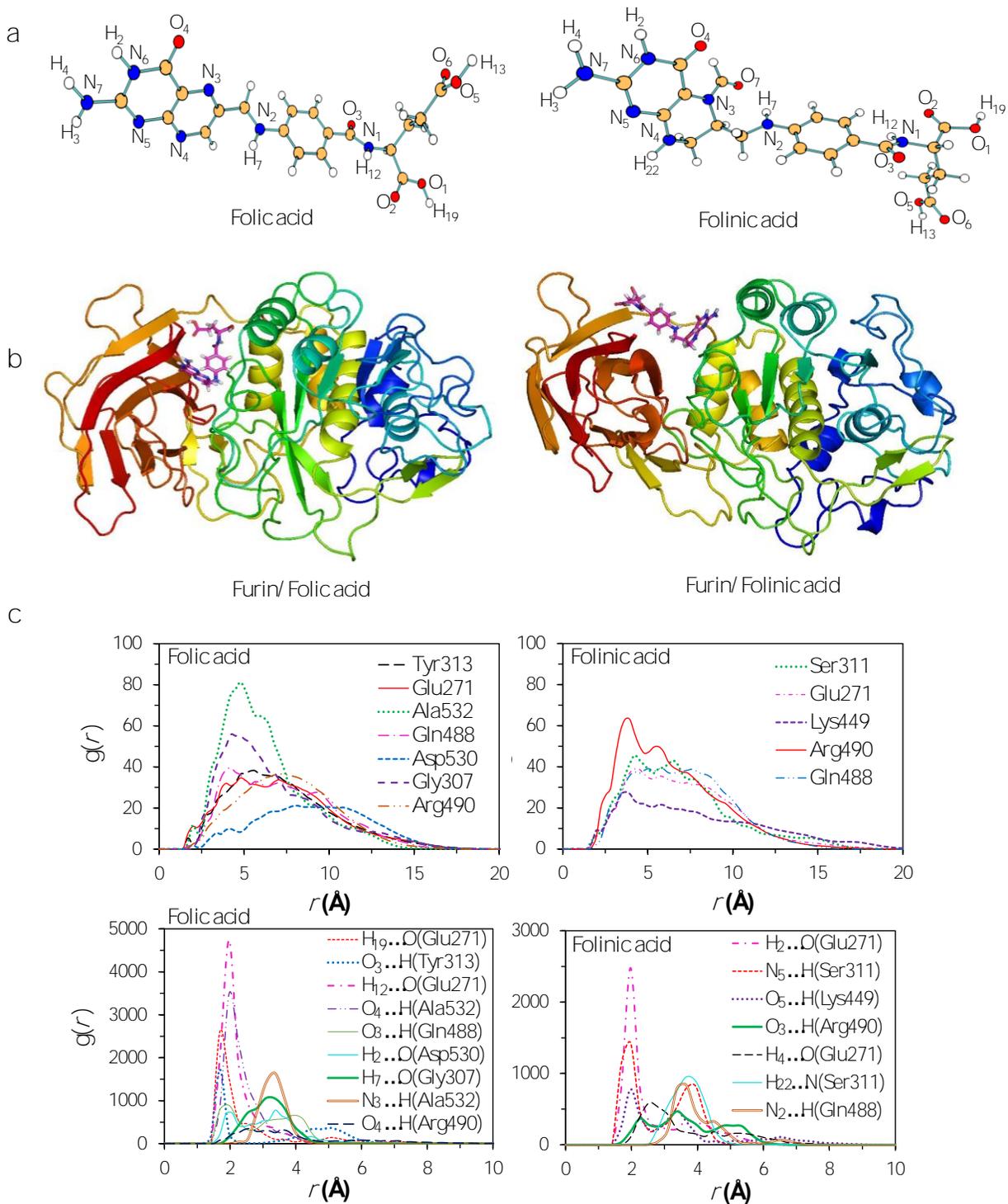
Fig. 3 | MD simulation results. a, The chemical structures of drug molecules with atoms labeling. **b,** One mode of drugs binding to the furin protein taken from a snapshot of the simulation at 100 ns. Water molecules were removed for clarity. **c,** Comparison of RDFs between the centers of mass of some residues of furin with drugs (top panel) and the atom…atom RDFs (bottom panel) for systems contain: folic acid (left panel), folinic acid (right panel).

The radial distribution functions (RDFs) between furin's residues and drug molecules are shown in Figs. 3c. For clarity, visual inspection of interaction sites are also demonstrated in Fig. 4 (see Fig. 3a for atom's labels). To obtain statistically reliable structural data, RDFs are calculated by averaging over trajectories of long production runs. The RDFs between the center of mass of folic acid and different amino acids of furin indicates that Ala532, Tyr313, and Glu271 interacted with folic acid with a relatively high probability and small dynamics both at short and long distances. In this case, the atom…atom RDFs demonstrated that position of the first peak obtained for H_{19}…O(Glu271) and O_3…H(Tyr313) were smaller than that of O_4…H(Ala532). Interestingly, MD simulations showed that the distance of O_3 atom of folic acid with Tyr313 and Gln488 hydrogen atoms were calculated to be 1.66 Å, and 1.96 Å, respectively (see Fig. 3a for atom labeling and Fig. 4a). The main interaction of O_4 atom of folic acid with Ala532, was located at 1.97 Å. The RDF peaks in Fig. 3c (right panel) showed that the folinic acid tended positioning slightly towards Ser311, Glu271, Arg490, Gln488, and Lys499 at short and long ranges. As shown in Fig. 4b, the main atomic interactions in this case were H_2…O(Glu271), and N_5…H(Ser311), which were located at 1.97 Å.

More details can be obtained from the spatial distribution function of the hydrogen bonding between the furin's amino acids and different atom sites of the drug molecules by the calculation of combined radial/angular distribution function (CDF) as a powerful tool for defining hydrogen bond criteria. Fig. 4c (up panel) manifests the most favorable hydrogen bonding interaction between the polar hydrogen atoms of the Tyr313 and the O_3 atom of carbonyl group of folic acid in the angle range $170 < \theta < 180$ at the distance of around 1.66 Å. Particularly, the CDF of Fig. 4c (down panel) indicates that there are interactions between the hydroxyl group of Ser311 and the folinic acid with the angle range $170 < \theta < 180$ and 1.97 Å distance. These interactions occurred when the drug molecules tilted substantially to directly interact with furin. Based on these findings, the folic acid…furin interactions were more probable than folinic acid.

9
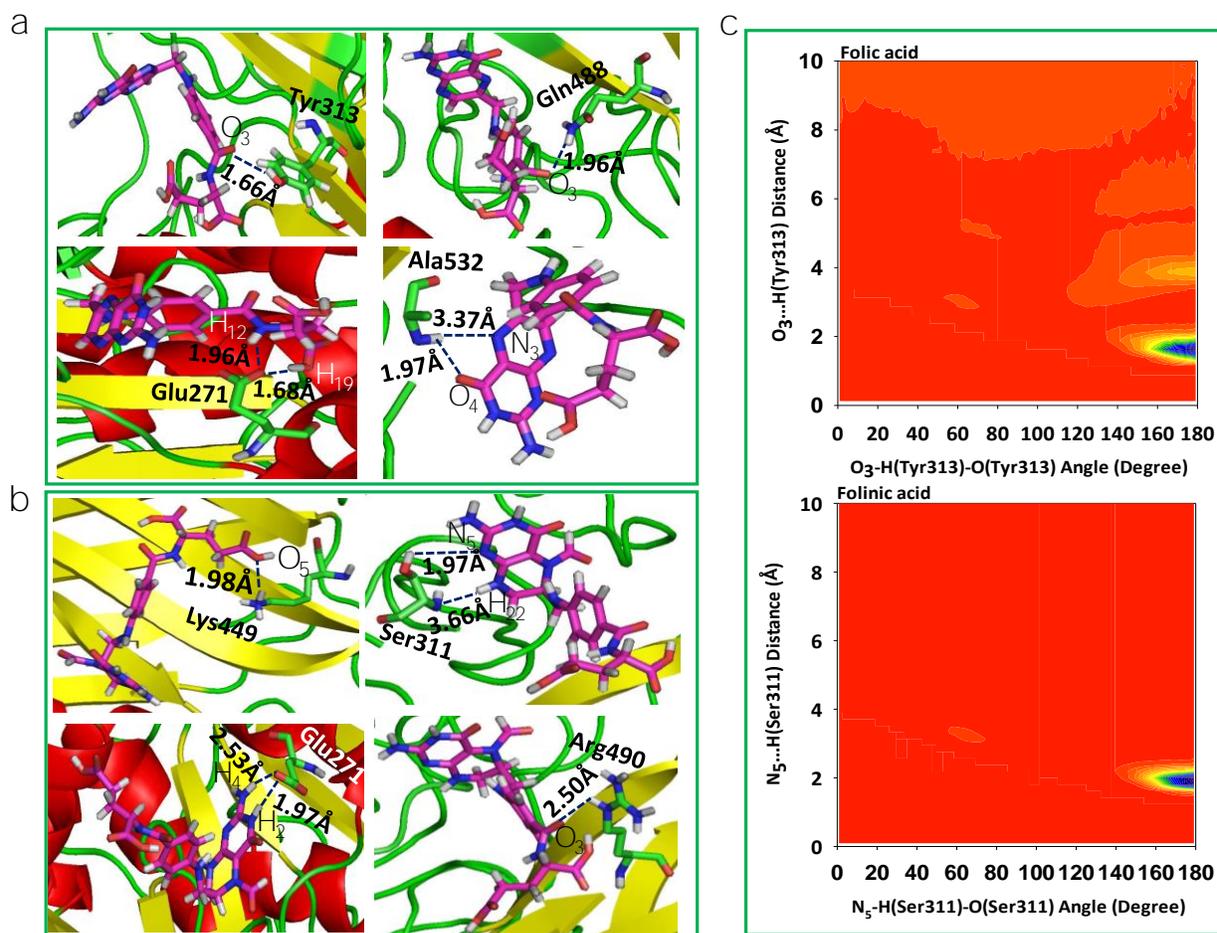
Fig. 4 | The main interaction of drugs with different residues of furin. a, Folic acid. **b,** Folinic acid. Drug molecules and residues are represented with sticks. For drugs, the red, blue, pink and white atom sites represent the oxygen, nitrogen, carbon, and hydrogen atoms. For amino acids of furin, the carbon atoms are colored in green. **c,** CDF of (up panel) furin-folic acid, (down panel) furin-folinic acid systems.

Discussion

Furin enzyme is associated with a great number of pathologies, including bacterial and viral infections, cancer, and metastasis. Hence, this protein is extremely considered as a drug target.[25,26] A characteristic feature of furin in the protease family is its very limited reactivity toward typical covalent inhibitors due to spatial restrictions.[27] Previous studies have documented that furin could promote the activation of coronavirus by sequence-specific cleavage of the spike

protein. Furin cleaves a wide variety of protein precursors in the preferred consensus motif (RXR(K)R/R (R=arginine, K= lysine, X= any amino acid).[12,28]

Therefore, furin protein appears to be a promising target for the infection treatment. The present study identified folic acid as a furin-binding protein by using an all atom MD simulation study. It is well known that folate receptors are mainly expressed in lungs and kidney in normal conditions.[29] Interestingly, the ACE2 receptors are also mostly expressed in lung. Recent studies have proposed that furin inhibition can have a substantial role in the prevention of COVID-19 infection progress.[30] Folic acid is small and stable over a broad range of temperatures and *pH* values, and it retains its ability to bind to the folate receptor after conjugation with drugs or diagnostic markers.[31,32] The present study introduces the ability of folic acid to interact and inhibit furin proprotein.

In this study, structural parameters such as radial distribution functions and combined radial-angular distribution functions were used to analyze the intermolecular interactions between folic acid (or folinic acid) with furin protein. The RDF presents the probability of finding a particle at a certain distance from another reference particle, therefore, contains the information of the average nearest neighbors' distance.[33] The first peaks of simulated RDFs were positioned at very short distances with very high probabilities, which could be attributed to the strong intermolecular interaction between folic acid molecule and furin enzyme. The combined distribution functions confirmed the hydrogen bond character of folic acid-furin interactions. The results indicated that the interactions between folinic acid and furin were high, but substantially lower than that of folic acid. In this way, folic acid could block the access of COVID-19 spikes to furin and prevent the cell entry and consequently turn-over of the virus.

In summary, our results suggest that folic acid could be used to inhibit the furin enzyme. The association of folic acid with furin would affect the structure of the protein and consequently interfere with its proteolytic capability. Thus, folic acid, as a safe drug, could be useful in the prevention or management of COVID-19-associated respiratory disease in the early stages of the disease.

References

1. Richman, D. D., Whitley, R. J. & Hayden, F. G. *Clinical Virology*. (ASM Press, Washington, 2016) (John Wiley & Sons).

2. Drosten, C., Günther, S., Preiser, W., Van Der Werf, S., Brodt, H. R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R. A. & Berger, A. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med*. **348**, 1967-1976 (2003).

3. Assiri, A., McGeer, A., Perl, T. M., Price, C. S., Al Rabeeah, A. A., Cummings, D. A., Alabdullatif, Z. N., Assad, M., Almulhim, A., Makhdoom, H. & Madani, H. Hospital outbreak of Middle East respiratory syndrome coronavirus. *N. Engl. J. Med*. **369**, 407-416 (2013).

4. Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X. & Cheng, Z. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*, **395**, 497-506 (2020).

5. Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N. & Bi, Y. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*, **395**, 565-574 (2020).

6. World Health Organization. Coronavirus disease (COVID-19) outbreak. Preprint at https://www.who.int (2020).

7. Bosch, B. J., van der Zee, R., de Haan, C. A. & Rottier, P. J. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J. Viro.* **77**, 8801-8811 (2003).

8. Hofmann, H., Pyrc, K., van der Hoek, L., Geier, M., Berkhout, B. & Pöhlmann, S. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc. Natl. Acad. Sci*. **102**, 7988-7993 (2005).

9. Delmas, B., Gelfi, J., L'Haridon, R., Sjöström, H. and Laude, H. Aminopeptidase N is a major receptor for the enteropathogenic coronavirus TGEV. *Nature* **357**, 417-420 (1992).

10. Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L. & Chen, H. D. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, **579**, 270–273 (2020).

11. Li, W., Moore, M. J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M. A., Somasundaran, M., Sullivan, J. L., Luzuriaga, K., Greenough, T. C. & Choe, H. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450-454 (2003).

12. Yamada, Y. & Liu, D. X. Proteolytic activation of the spike protein at a novel RRRR/S motif is implicated in furin-dependent entry, syncytium formation, and infectivity of coronavirus infectious bronchitis virus in cultured cells. *J. Viro.* **83**, 8744-8758 (2009).

13. Belouzard, S., Chu, V. C. & Whittaker, G. R. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc. Natl. Acad. Sci*. **106**, 5871-5876 (2009).

14. Wang, H., Yang, P., Liu, K., Guo, F., Zhang, Y., Zhang, G. & Jiang, C. SARS coronavirus entry into host cells through a novel clathrin-and caveolae-independent endocytic pathway. *Cell Res.* **18**, 290-301 (2008).

15. Gallagher, T. M. & Buchmeier, M. J. Coronavirus spike proteins in viral entry and pathogenesis. *Virology* **279**, 371-374 (2001).

16. Liu, Z., Xiao, X., Wei, X., Li, J., Yang, J., Tan, H., Zhu, J., Zhang, Q., Wu, J. & Liu, L. Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. *J. Med. Virol.* 1-7 (2020).

17. Yamada, Y., Liu, X. B., Fang, S. G., Tay, F. P. & Liu, D. X. Acquisition of cell–cell fusion activity by amino acid substitutions in spike protein determines the infectivity of a coronavirus in cultured cells. *PloS one* **4**, e6130 (2009).

18. Bosshart, H., Humphrey, J., Deignan, E., Davidson, J., Drazba, J., Yuan, L., Oorschot, V., Peters, P. J. & Bonifacino, J. S., The cytoplasmic domain mediates localization of furin to the trans-Golgi network en route to the endosomal/lysosomal system. *J. Cell Biol.* **126**, 1157-1172 (1994).

19. Vidricaire, G., Denault, J. B. & Leduc, R. Characterization of a secreted form of human furin endoprotease. *Biochem. Biophys. Res. Commun*. **195**, 1011-1018 (1993).

20. Basak, A., Zhong, M., Munzer, J. S., CHRÉTIEN, M. & SEIDAH, N. G. Implication of the proprotein convertases furin, PC5 and PC7 in the cleavage of surface glycoproteins of Hong Kong, Ebola and respiratory syncytial viruses: a comparative analysis with fluorogenic peptides. *Biochem J.* **353**, 537-545 (2001).

21. Feliciangeli, S. F., Thomas, L., Scott, G. K., Subbian, E., Hung, C. H., Molloy, S. S., Jean, F., Shinde, U. & Thomas, G. Identification of a *pH* sensor in the furin propeptide that regulates enzyme activation. *J. Biol. Chem.* **281**, 16108-16116 (2006).

22. Henrich, S., Cameron, A., Bourenkov, G. P., Kiefersauer, R., Huber, R., Lindberg, I., Bode, W. & Than, M. E. The crystal structure of the proprotein processing proteinase furin explains its stringent specificity. *Nat. Struct. Mol. Biol.* **10**, 520-526 (2003).

23. Li, H., Wu, C., Yang, Y., Liu, Y., Zhang, P. Wang, Y., Wang, Q., Xu, Y., Li, M., Zheng, M., Chen, L. Furin, a potential therapeutic target for COVID-19. (chinaXiv:202002.00062v1); DOI: 10.12074/202002.00062, Preprint at http://www.chinaxiv.org/abs/202002.00062 (2020).

24. Giovannucci, E., Stampfer, M. J., Colditz, G. A., Rimm, E. B., Trichopoulos, D., Rosner, B. A., Speizer, F. E. & Willett, W. C. Folate, methionine, and alcohol intake and risk of colorectal adenoma. *J. Natl. Cancer Inst.*, **85**, 875-883 (1993).

25. Seidah, N. G. & Prat, A. The biology and therapeutic targeting of the proprotein convertases. *Nat. Rev. Drug Discov.* **11**, 367-383 (2012).

26. Jiao, G. S., Cregar, L., Wang, J., Millis, S. Z., Tang, C., O'Malley, S., Johnson, A. T., Sareth, S., Larson, J. & Thomas, G. Synthetic small molecule furin inhibitors derived from 2, 5-dideoxystreptamine. *Proc. Natl. Acad. Sci.* **103**, 19707-19712 (2006).

27. Dahms, S. O., Arciniega, M., Steinmetzer, T., Huber, R. & Than, M. E. Structure of the unliganded form of the proprotein convertase furin suggests activation by a substrate-induced mechanism. *Proc. Natl. Acad. Sci.* **113**, 11196-11201 (2016).

28. Millet, J. K. & Whittaker, G. R. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proc. Natl. Acad. Sci.* **111**, 15214-15219 (2014).

29. Parker, N., Turk, M. J., Westrick, E., Lewis, J. D., Low, P. S. & Leamon, C. P. Folate receptor expression in carcinomas and normal tissues determined by a quantitative radioligand binding assay. *Anal. Bioanal. Chem.* **338**, 284-293 (2005).

30. Couture F, Kwiatkowska A, Dory YL, Day R. Therapeutic uses of furin and its inhibitors: a patent review. *Expert Opinion on Therapeutic Patents.* **25**, 379-96 (2015).

31. Mansoori, G. A., Brandenburg, K. S. & Shakeri-Zadeh, A. A comparative study of two folate-conjugated gold nanoparticles for cancer nanotechnology applications. *Cancers* **2**, 1911-1928 (2010).

32. Zwicke, G. L., Ali Mansoori, G. & Jeffery, C. J. Utilizing the folate receptor for active targeting of cancer nanotherapeutics. *Nano Rev.* **3**, 18496 (2012).

33. Zahn, S., Wendler, K., Delle Site, L. & Kirchner, B. Depolarization of water in protic ionic liquids. *Phys. Chem. Chem. Phys.* **13**, 15083-15093 (2011).

Methods

Computational details. In this study, docking of two folate analogs, folic acid and folinic acid against human furin was performed using molegro virtual docker (MVD) version 6.0 software.[34] An Xray crystal structure of human furin was used for docking studies taken from the Protein Data Bank (PDB code 5MIM) accessed at the URL (http://www.rscb.org/pdb) with reasonable resolution (≤ 1.9 Å). The initial coordinates of folic acid and folinic acid were obtained by DFT calculations at the B3LYP/6-311+G** level of theory implemented in Gaussian software (see Fig. 3a).[35] Two drugs were docked against furin protein and 10 independent runs were performed with the guided differential evolution algorithm. In case of furin-drug complexes, the program generally identified five different binding sites. Among these five predicted cavities, the one with the volume of 75.776 Å^3 was selected as the potential binding site for investigation. The Moldock score of the best drug-furin complex of folic acid, and folinic acid were -140.4, and -136.9

14

kcal/mol, respectively. The best configurations from the above-mentioned procedures were selected as the initial structures for MD simulation in aqueous media. In this case, the MD simulations were performed with the GROMACS 4.5.4 program using the GROMOS96 53A6 force field employing periodic boundary conditions in three dimensions.[36] The simulations consisted of two different systems: (I) furin-folic acid in water, (II) furin-folinic acid in water. The partial atomic charges were calculated by using natural population analysis as implemented in Gaussian 09 program. All simulations were performed in the presence of water by using simple point charge (SPC) model.[37] The net charge of the systems was neutralized by addition of the same amount of sodium or calcium ions. After energy minimization, an equilibration NVT run has been carried out over 500 ps, while restraining the position of furin by force constant of 1000 kJ mol^{-1} nm^{-2} to their initial position. The systems were equilibrated for 500 ps using an NPT ensemble. The production runs were performed for 100 ns at 310 K.

References

34. Naeem, S., Hylands, P. & Barlow, D. Docking studies of chlorogenic acid against aldose redutcase by using molgro virtual docker software. *J. Appl. Pharm. Sci.* **3**, 13 (2013).

35. Frisch, M. J. E. A., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, 4. B., Petersson, G. E., Nakatsuji, H. Gaussian 09. (2009).

36. Hess, B., Kutzner, C., Van Der Spoel, D. & Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput*. **4**, 435-447 (2008).

37. Berendsen, H. J. C., Grigera, J. R. & Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **91**, 6269-6271 (1987).

Acknowledgements