Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Structural dynamics of COVID-19 main protease

Ashkan Shekaari, Mahmoud Jafari*

Department of Physics, K. N. Toosi University of Technology, Tehran 15875-4416, Iran

A R T I C L E   I N F O

Article history:
Received 27 April 2020
Revised 8 September 2020
Accepted 9 September 2020
Available online 9 September 2020

Keywords:
COVID-19 main protease
Internal dynamics
Normal modes

A B S T R A C T

Based on the importance of protease enzymes in functioning some viruses particularly coronaviridae, we have carried out an in silico investigation on the biologically important, yet unmapped phenomenon of activity and internal dynamics of COVID-19 main protease (Mpro) via applying finite-temperature all-atom molecular dynamics simulations. Temperature quench echoes generated by applying two successive cooling signals have therefore been analyzed in terms of the temperature-temperature correlation function of the protease within the harmonic approximation. An exponentially decaying brand of behavior has been found for the calculated echo depth values with increasing time, which has accordingly led to a much small dephasing time of about 150 fs, revealing a significant anharmonicity and therefore an overall structural stiffness for the COVID-19 main protease.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Over the last two decades, a large number of coronaviruses have so far been identified in their natural reservoir host (bats) since the outbreak of severe acute respiratory syndrome (SARS) in 2003 [1–4]. Recently, a new coronavirus identified as COVID-19 virus has also been discovered as the etiological agent responsible for the ongoing 2019–2020 viral pneumonia outbreak beginning in Wuhan, China [5–8]. Based on the fact that COVID-19 is an emerging situation, very few investigations have so far been carried out on the virus, and no targeted therapeutics or effective treatments are therefore available at present. A possible way to disrupt COVID-19 function, just like some other viruses such as the SARS coronavirus [9–14], is targeting its main protease (Mpro) as a key enzyme in mediating replication and transcription of the virus [15,16].

Although the crystal structure of COVID-19 Mpro (PDB ID: 6LU7 [17]) provides a solid basis for investigating the relationship between its structure and function, several questions regarding many dynamic aspects of its enzymatic mechanism such as substrate binding, orientation, catalysis, product release, and the way these processes take place have not yet been answered.

Proteins are soft materials and their biological functions at the most fundamental level are related to physical motions of their constituent atoms [18]. As a result, a detailed investigation on dynamics of COVID-19 Mpro could be necessary for a complete understanding of its function. The present work has therefore been devoted to inquiring into the activity and internal structural dynamics of COVID-19 Mpro within the normal mode picture of proteins. To this end, we have applied all-atom molecular dynamics (MD) simulations, as a widely used tool in studying structure and function of biological entities, and in unraveling the ways they interact with other bio/nano structures [19].

The present research has been organized as follows. Section 2 describes the applied computational setup and details. In Section 3 the obtained results are presented and being discussed. We finally conclude our results and discussion in Section 4.

2. Materials and method

All-atom MD simulations have been applied to investigate the internal structural dynamics of the COVID-19 Mpro as implemented in the NAMD computer software (version 2.13) [20] along with the VMD program (version 1.9.4) [21] to post-process the outputs. Initial atomic positions of the protease have been taken from the Protein Data Bank (PDB) with entry code 6LU7. The Mpro has been solvated in an orthorhombic water box with volume of $9.2 \times 10.7 \times 10$ nm$^3$ (Fig. 1) under periodic boundary conditions with a unit-cell padding of about 2 nm to decouple the weakest periodic interactions possible between the Mpro replicas. The Na and Cl ions have randomly been distributed in order for the whole system to be electrostatically neutral. The July 2018 update of CHARMM36 topology and parameter files [23,24] have been used as the force fields to calculate the potential energy surface with the standard TIP3P water model [25]. The switching and cutoff distances of about 1.0 and 1.2 nm have also been used for trun-
cating non-bonded van der Waals interactions, respectively. The particle-mesh Ewald (PME) [26] method with grid dimensions of 96 × 108 × 108 has been applied to deal with long-range interactions. The dielectric constant of the system has also been set to $\kappa_0 = 1.0$. The system has been minimized for 0.4 ns during which the convergences of total energy of the system ($E_{\text{Tot}}$), total energy of the Mpro ($E_{\text{prot}}$), number of hydrogen bonds, radius of gyration, root mean square deviation (RMSD), and salt bridge distances have accordingly been achieved, as illustrated in Fig. 2.

Hydrogen bonds are identified within the hydrogen donor-acceptor distance of 3.0 Å with angle cutoff of 20.0°. The oxygen-nitrogen distance cutoff for salt bridges has also been set to 3.20 Å. The isothermal-isobaric (NPT) simulations have been performed at $T_S = 310$ K and 1.01325 bar using Langevin forces with the Langevin damping constant of about 2.5/ps along with the Nosé-Hoover Langevin piston pressure control. The integration time step is 1.0 fs. The reliability of our setup has also been tested to a great accuracy as described in Appendix A.

3. Results and discussion

For proteins, internal dynamics including a wide range of timescales from femtosecond bond stretching vibrations to deformations requiring seconds or more, has so far been of particular interest due to the importance of these motions in biological functioning [28–30]. Among them are picosecond motions, which exhibit a delocalized character that often involve harmonic displacements of a large number of atoms in different parts of the protein [31–34].

To probe the internal dynamics of the COVID-19 Mpro, we first have equilibrated the system during a NPT simulation at 310 K. We have used 3N − 6 degrees of freedom based on the fact that a N-atom protein has 3N − 6 internal normal modes after removing the six degrees of freedom associated with overall translation and rotation. The time dependence of the temperature auto-correlation function ($C_{TT}$) is then estimated so that the instantaneous temperature takes the form $T(t) = [(3N - 6)k_B]^{-1} \sum_{i=1}^{3N-6} \tilde{m}_i^2(t)$. From then on, the thermostat is removed and the next MD simulations are carried out in the microcanonical (NVE) ensemble. We indeed approximate the Mpro as a large collection of weakly interacting harmonic oscillators on the picosecond timescale for which only very short trajectories are required. These oscillators (normal modes) are synchronized through two successive quenching processes (cooling signals) as the required perturbations forcing the system to operate in a phase-coherent state in which normal modes oscillate in phase. At the beginning of each quenching process, zero velocities are assigned to all atoms of the Mpro, while their positions are left unchanged. The first velocity reassignment enforces the phase coherence for the oscillators decaying in time with a characteristic dephasing time $\tau_d$ due to frequency dispersion of the normal modes. The second reassignment after the delay time ($\tau$) serves as the probing signal, which through interference with the previously generated coherence of the normal modes, results in a resonance (echo) in the form of a sharp dip in $\langle C_{TT} \rangle$. Such a signal indeed tests the degree of coherence of the Mpro structure via examining both the depth of echo and the instant of time at which it takes place. These two are among the quantitative characteristics of the coherence associated to internal dynamics of COVID-19 Mpro.

Fig. 3 illustrates the normalized temperature auto-correlation function ($C_{TT}$) of COVID-19 Mpro calculated according to the formula

$$C_{TT}(t) = \frac{T(t)T(0) - (T(0))^2}{<T^2(t)> - <T(0)>^2}.$$  

Fitting an exponentially decaying function of the form $\exp(-t/\tau_0)$ to the calculated simulation data via a least-square method results in $\tau_0 = 73.736$ fs as the auto-correlation decay time of the COVID-19 Mpro.

Fig. 4 shows the two successive quench signals applied to the pre-equilibrated system for $\tau = 0.2$ ps as well as the related harmonic approximation prediction in a contrasting fashion.

The degree of agreement between our echo calculations and the theoretical prediction of the harmonic approximation, namely

$$T(t) \approx \frac{T_S}{4} \left[ 1 + \frac{T_1}{T_S} + \frac{2T_2}{T_S} - \left( 1 + \frac{T_1}{T_S} - 2\frac{T_2}{T_S} \right) C_{TT}(t - \tau) \right. \left. - 2\sqrt{\frac{T_1T_2}{T_S}} C_{TT}\left( |t - 2\tau| \right) \frac{1}{2} \left( 1 - \frac{T_1}{T_S} \right) C_{TT}(t - 2\tau) \right].$$  

could now be seen, where $C_{TT}(t)$ in (Eq. (1) [36]) is approximated by $\exp(-t/\tau_0)$, and $T_1 = T_2 = 0$ because of quenching. Moreover, the average temperature of the system after applying the first and second cooling signals drops to approximately one half (from 0 to $\tau$), and to one third (from $\tau$ to $4\tau$) of the initial equilibrium temperature ($T_S$). This is due to the fact that each particle has a potential energy and immediately begins to move again after each quench, converting part of its potential energy back to kinetic energy exactly according to the equipartition theorem. Such a small timescale of order $\tau = 0.2$ ps is enough for the Mpro to return to equilibrium, as the figure shows. At $t = 2\tau$, the echo of the applied probing signal is seen in the form of a relatively sharp dip in the diagram of normalized temperature ($T(t)/T_S$). For $t > \tau$, a clear deviation from the equipartition theorem is also observed, indicating that the system is very sensitive to such external perturbations.

To estimate the dephasing time ($\tau_d$) of the Mpro, we have calculated the temperature echo for 21 values of the delay time ranging from 0.05 to 2.0 ps every 0.05 ps. The echo depth values have accordingly been calculated via subtracting the minimum temperature of the system from the average value after the second quench as shown in Fig. 5.

Fitting $\exp(-t/\tau_d)$ to the calculated echo depth values results in a characteristic dephasing time of about $\tau_d = 151.638$ fs as an inherent property of the Mpro. The smallness of this value could be inferred from the fact that it is indeed around the timescale of hydrogen bond vibrations in proteins, revealing that the COVID-19 Mpro has a high degree of structural stiffness as a physical property.
Fig. 2. Convergence diagrams of (a) total energy of the system ($E_{tot}$), (b) total energy of the M$^{pro}$ ($E_{prot}$), (c) number of hydrogen bonds, (d) radius of gyration, (e) RMSD, and (f) salt bridge distances during minimization—plotted by Gnuplot [27]. The convergence values are $-362797.8$ kcal/mol, $-3430.353$ kcal/mol, $138.228$ nm, and $0.8575$ Å, respectively, except those of the salt bridges, which are within the range of $\approx 3.0 - 5.0$ Å.

Fig. 3. Normalized temperature-temperature correlation function (red squares) and the exponentially decaying fit function (blue) of the form $\exp(-t/\tau_0)$ with a heavy tail, leading to the auto-correlation decay time of about $\tau_0 = 73.736$ fs obtained via a nonlinear least-square (NLLS) fit using Xmgrace (version 5.1.25) [35]. Deviation from the harmonic approximation $[\eta(\tau)]$ is also evident and is calculated at $t = 2\tau$ (the echo point) according to

$$\eta(\tau) = \left[ \frac{T(2\tau)}{T_s} \right]_{\text{simulation}} - \left[ \frac{T(2\tau)}{T_s} \right]_{\text{theory}}.$$  \hspace{1cm} (2)

as illustrated in Fig. 6 for different values of $\tau$.

Fig. 4. The first (synchronization) and the second (probing) quench signals applied to the M$^{pro}$ calculated for $\tau = 0.2$ ps over the intervals $[0, \tau]$ and $[\tau, 4\tau]$, respectively. The related temperature echo is detected at $2\tau$ in the form of a relatively sharp dip (red circle) in the normalized temperature $T(t)/T_s$. Deviation from the harmonic approximation is evident within $[\tau, 4\tau]$.

It is seen that as $\tau$ increases, deviation from theory [Eq. (1)] becomes larger, though according to a limiting behavior, which asymptotically approaches to about 73.

Internal dynamics of the M$^{pro}$ therefore exhibits a weak harmonicity with a significant deviation from theoretical predictions within the normal-mode description of proteins. The coherence of atomic motions in the COVID-19 M$^{pro}$ decays on such a small
timescale of about \( \tau_d \approx 151.638 \text{ fs} \) through nonlinear contributions to forces between the constituent atoms. The reason why the calculated echo depth values (Fig. 5), or the behavior of normalized temperature (Fig. 4), are not so exact as those predicted by the harmonic picture reveals that the system has a weak harmonicity, and this is largely due to the existence of anharmonic contributions arising from torsional, electrostatic, and van der Waals interactions that contribute to the potential acting within the COVID-19 MPro. The temperature of the system at the echo time \( t = 2\tau \) is weighted the most by those modes that have the most total energy, i.e., those that reach their classical turning points on harmonic potential wells described by the harmonic approximation. As Fig. 5 shows, the echo depth decreases exponentially with \( \tau \), indicating that the temperature of the system at the echo point becomes closer to its average value after the second cooling pulse applied. Therefore, one possible explanation for decrease in the echo depth is that the more the value of \( \tau \), the less the number of modes that reach their classical turning points at \( t = 2\tau \), which accordingly reduces the echo depth. On the other hand, that temperature at the echo point does not go completely to zero is due to the small contribution of those modes that have non-zero kinetic energy at \( t = \tau \). One, however, cannot establish a correlation between decrease in the echo depth on one hand, and increase in the contribution of these modes as \( \tau \) increases on the other, because at the quench time, their contribution to kinetic energy is exactly zero. Another possible explanation for decrease in the echo depth can then be provided so that there exists a considerable deviation from the harmonic picture and therefore the potential well each particle within the MPro experiences significantly deviates from being exact parabola as the value of delay time increases.

4. Conclusions

In summary, we have investigated the activity and internal dynamics of the COVID-19 MPro within the normal-mode picture of proteins using all-atom MD simulations. We therefore have generated temperature quench echoes by applying two successive quench signals and have then examined them in terms of the temperature auto-correlation function within the harmonic approximation. Analyzing the exponentially decaying trend of the calculated echo depth values with increasing time has led to a much small dephasing time of about 150 fs, exhibiting accordingly a significant anharmonicity and therefore a total structural solidity for the COVID-19 MPro.

Declaration of Competing Interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome. We also confirm that the manuscript has been read and approved by all named authors, and the order of authors listed in the manuscript has been approved by all named authors.

Appendix A. Temperature and energy distributions

We have tested the reliability of our computational setup during two simulations within NVE and NPT ensembles, and then by monitoring the probability distributions of the instantaneous temperature and kinetic energy of the system as illustrated in Fig. A.7.

As it is seen, the temperature distribution [Fig. A.7(a)] is clearly Gaussian (normal) in accordance with the central limit theorem, and could then be described by a Gaussian function of the form

\[
P(T - T_0) = c \exp \left[ \frac{-(T - T_0)^2}{2\sigma^2} \right].
\]

so that \( T = 2E_k/3k_B \) is the instantaneous temperature according to the equipartition theorem, \( E_k \) is the total kinetic energy of the system, \( \sigma^2 \) is the variance, \( k_B \) is the Boltzmann constant, and \( c \) is a normalization constant. Fitting Eq. (A.1) to the calculated simulation data results in \( c = 0.6 \), \( T_0 = 308.55 \text{ K} \), and \( \sigma^2 = 0.85 \text{ K}^2 \), being in agreement with \( T_0 = 310 \text{ K} \) and \( \sigma^2 = 2T^2 / 3N = 0.69 \text{ K}^2 \) for the total number \( N = 92586 \) of atoms of the system. The distribution of kinetic energy [Fig. A.7(b)] is a Maxwell-Boltzmann (MB) distribution, and fitting a function of the form

\[
f(E_k) = 2 \sqrt{\frac{E_k}{K_B T}} (k_B T)^{-3/2} \exp \left( -\frac{E_k}{k_B T} \right)
\]

to the simulation data accordingly leads to \( k_B T = 0.612421 \text{ kcal/mol} \), or equivalently \( T = 308.28 \text{ K} \), which is in a remarkable agreement with \( T_0 = 310 \text{ K} \). See Appendix B for the proof of Eq. (A.2).

Appendix B. The proof of Eq. (A.2)

We have

\[
f(E_k) dE_k = f(p) d^3p,
\]

where \( d^3p \) is the infinitesimal phase-space volume of momenta corresponding to the energy interval \( dE_k \). From the spherical symmetry of the energy-momentum dispersion relation, \( E_k = |p|^2 / 2m \), this can be expressed in terms of \( dE_k \):

\[
d^3p = 4\pi |p|^2 dE_k = 4\pi m \sqrt{2mE_k} dE_k.
\]
Plugging (B.2) back into (B.1), and expressing everything in terms of $E_k$, results in

$$f(E_k)dE_k = \frac{(2\pi mk_B T)^{-3/2}}{4\pi m^{3/2}2mE_k} \exp\left(-\frac{E_k}{k_BT}\right) = 2\sqrt{\frac{m}{\pi k_BT}}(k_BT)^{-3/2} \exp\left(-\frac{E_k}{k_BT}\right).$$

CRediT authorship contribution statement

Ashkan Shekara: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing - review & editing. Mahmoud Jafarid: Project administration, Supervision.

References

[1] W. Li, Z. Shi, M. Yu, W. Ren, C. Smith, J.H. Epstein, H. Wang, G. Crameri, Z. Hu, H. Zhang, J. Zhang, J. McEarchen, H. Field, P. Daszak, B.T. Eaton, S. Zhang, L.F. Wang. Bats are natural reservoirs of SARS-like coronaviruses. Science 310 (2006) 676–679.

[2] X.-Y. Ge, J.-L. Li, X.-L. Yang, A.A. Chmura, G. Zhi, J.H. Epstein, J.K. Mazer, B. Hu, W. Zhang, C. Peng, Y.-J. Zhang, C-M. Luo, B. Tan, N. Wang, Y. Zhi, G. Crameri, S.-Y. Zhang, L.-F. Wang, P. Daszak, Z.L. Shi. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor, Nature 503 (2013) 535–538.

[3] L. Yang, Z. Wu, X. Ren, F. Yang, G. He, J. Jiang, J. Dong, L. Sun, Y. Zhi, J. Du, S. Zhang, Q. Jin. Novel SARS-like betacoronaviruses in bats, China, 2011. Emerg. Infect. Dis. 19 (2013) 989–991.

[4] B. Hu, L.-F. Zeng, X.-L. Yang, X.-Y. Ge, W. Zhang, B. Li, J.-Z. Xie, X.-R. Shen, Y.-Z. Zhang, N. Wang, D.-S. Luo, X.-S. Zheng, M.-N. Wang, P. Daszak, L.-F. Wang, J. Cui, Z.L. Shi. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS Coronavirus, PLoS Pathog. 13 (2017) e1006698.

[5] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Gao, G.F. Gao, W. Tan, A novel coronavirus from patients with pneumonia in China, 2019, N. Engl. J. Med. 382 (2020) 727–733.

[6] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K.S.M. Leung, E.H.Y. Lau, J.Y. Wong, X. Xiong, X. Niang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Zhang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, R. Liu, Z. Ma, Y. Zhang, G. Shi, T.Y.T. Lam, J.T.W. Gao, B.J. Cowling, B. Yang, G.M. Leung, F. Feng, Early transmission dynamics in Wuhan, China of novel coronavirus–infected pneumonia, N. Engl. J. Med. 382 (2020) 1199–1207.

[7] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, Z. Zhang, H.-R. Shi, Y. Zhi, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, R.-X. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.L. Shi. A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 579 (2020) 270–273.

[8] F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, M.-L. Yuan, Y.-L. Zhang, F.-H. Dai, Y. Liu, Q.-M. Wang, J.-J. Zheng, L. Xu, E.C. Holmes, V.Z. Zhang, A new coronavirus associated with human respiratory disease in China, Nature 579 (2020) 265–269.

[9] T.A. John, C.J. Kuo, H.P. Hsieh, Y.C. Wang, K.K. Huang, C.P.C. Lin, P.F. Huang, X. Chen, P.H. Liang, Evaluation of metal-conjugated compounds as inhibitors of 3CL protease of SARS-cov, FEDS Lett. 574 (2004) 116–120.

[10] X.W. Zhang, Y.L. Ya, Exploring the binding mechanism of the main proteinase in SARS-associated coronavirus and its implication to anti-SARS drug design, Bioorg. Med. Chem. 12 (2004) 2219–2223.
[32] B. Brooks, M. Karplus, Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor, Proc. Natl. Acad. Sci. U. S. A. 80 (1983) 6571–6575.

[33] M. Levitt, C. Sander, P.S. Stern, Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme, J. Mol. Biol. 181 (1985) 423.

[34] T. Ichiye, M. Karplus, Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations, Proteins 2 (1987) 236.

[35] http://plasma-gate.weizmann.ac.il/Grace/.

[36] D. Xu, K. Schulten, O.M. Becher, M. Karplus, Temperature quench echoes in proteins, J. Chem. Phys. 103 (1995) 3112–3123.