ChemMedChem

Communications
doi.org/10.1002/cmdc.202100156

Chemistry Europe
European Chemical Societies Publishing

Inspecting the Mechanism of Fragment Hits Binding on SARS-CoV-2 M^{pro} by Using Supervised Molecular Dynamics (SuMD) Simulations

Maicol Bissaro,[a] Giovanni Bolcato,[a] Matteo Pavan,[a] Davide Bassani,[a] Mattia Sturlese,*[a] and Stefano Moro*[a]

Computational approaches supporting the early characterization of fragment molecular recognition mechanism represent a valuable complement to more expansive and low-throughput experimental techniques. In this retrospective study, we have investigated the geometric accuracy with which high-throughput supervised molecular dynamics simulations (HT-SuMD) can anticipate the experimental bound state for a set of 23 fragments targeting the SARS-CoV-2 main protease. Despite the encouraging results herein reported, in line with those previously described for other MD-based posing approaches, a high number of incorrect binding modes still complicate HT-SuMD routine application. To overcome this limitation, fragment pose stability has been investigated and integrated as part of our in-silico pipeline, allowing us to prioritize only the more reliable predictions.

Fragment-based drug discovery (FBDD) has progressively established as a game-changing approach to navigate the chemical space in the drug discovery pipelines, both on academic and industrial early discovery stages.[1–3] By definition, fragments are low molecular weight organic molecules able to recognize a target of therapeutic interest in a mild affinity range and with a poor selectivity profile.[4] Intriguingly, the screening of small-sized fragment libraries in place of conventional larger ones has proven to provide better coverage of the chemical diversity and higher hit rates.[5,6] The identification of such weak binders, however, strictly depends on the implementation of biophysical screening techniques, such as X-Ray Crystallography (XRC), Nuclear Magnetic Resonance (NMR), surface plasmon resonance (SPR), or Thermal Shift Assay (TSA).[1,7,8] Anyway, broad differences exist among such methods and each of them suffers unique limitations in the challenging identification of reliable fragment; indeed the agreement in the hits identified is surprisingly limited.[9–11] Besides, only XRC and NMR offer the possibility to investigate the binding mode of weak binders. In light of this, structure-based computational strategies have increasingly gained appeal.[12–14] As highlighted in a recent review, during the last decade Molecular Dynamics (MD) simulations have been extensively applied also in the FBDD field, providing an atomistic insight on the fragment-receptor binding mechanisms, with a femtosecond temporal resolution.[15] From this perspective, we recently developed HT-SuMD, a computational protocol exploiting supervised MD simulations to perform the screening of a small fragments library in a competitive timescale.[16] The performance of the protocol in prioritizing the most promising fragment binders was compared with NMR-based screening, against the oncological protein target Bcl-x$_L$. Despite the notable agreement with NMR in identifying the most promising hits, the lack of structural data prevented the assessment of HT-SuMD accuracy in fragments binding mode prediction, which would represent a valuable set of information to guide the subsequent hit to lead (H2L) optimization steps. In this methodological study, we have therefore retrospectively investigated the accuracy of HT-SuMD simulations in reproducing the experimental binding mode of several fragment-protein complexes, exploiting the 3-C-like main protease (M^{pro}) of the novel SARS-CoV-2 coronavirus as a relevant case study. Following indeed the dramatic spread of the COVID-19 pandemic, a collaborative XRC fragment screening against the protein M^{pro} has timely offered to the scientific community valuable structural information to accelerate the rational design of new protease inhibitors.[17–19] For this validation study in detail, among the 71 fragments targeting the catalytic site of M^{pro} originally identified by the XRC screening, only the 23 presenting a reversible mechanism of recognition were taken into consideration, due to the impossibility of modeling covalent reactivity through classical molecular mechanics (MM) force fields.[20,21]

Results and Discussion

Characterization of fragment-receptor complexes

The high-quality M^{pro} crystallographic structures were collected from the Protein Data Bank database (PDB ID are reported in Table 1 of SI) and prepared by applying symmetric trans-

[a] M. Bissaro, G. Bolcato, M. Pavan, D. Bassani, Dr. M. Sturlese, Prof. S. Moro
Molecular Modeling Section (MMS)
Department of Pharmaceutical and Pharmacological Sciences
University of Padova
via Marzolo 5
35131 Padova (Italy)
E-mail: stefano.moro@unipd.it
mattia.sturlese@unipd.it

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cmdc.202100156

formation to each asymmetric unit, thus recreating the original functional dimer.[22] A visual inspection of the catalytic clefts has revealed how the 23 non-covalent fragments comprehensively explore most protease binding subsites (S1, S2, S3, and S1′) while providing decent coverage of chemical diversity. Besides, M^{pro} catalytic cleft is easily accessible from the bulk solvent and hence suitable to SuMD studies, as recently demonstrated for a couple of M^{pro} inhibitors.[23] The complexity, as well as the plasticity of the M^{pro} binding pocket, made this test case particularly challenging, the reason why an MD-based stability characterization of all the experimental-solved crystallographic complexes was performed, before investigating HT-SuMD accuracy in the fragment posing process. For this purpose, the AMBER14SB force field was combined with the general amber force field (GAFF) to parameterize respectively the protein biopolymers and the small organic fragments.[24,25] To ensure results robustness, 5 trajectories each 20 ns long were collected for all M^{pro}complexes, resulting in a total of 2.3 μs of conventional MD simulations. The content of information extrapolated from a single trajectory has been hence doubled by simply repeating the analysis against the two distal and independent catalytic sites of the homodimeric SARS-CoV-2 M^{pro}. To characterize the geometric stability of the experimental-solved fragment complexes the root-mean-square fluctuation (RMSF) of ligands heavy atoms has been chosen as a metric, then summarizing the results through a heatmap representation, as reported in Figure 1.

The colorimetric scale helps in differentiating those fragments which maintained the original binding mode during all the collected replicates (green color), from others undergoing a neat perturbation of the recognition modality (yellow color) or that even experience a spontaneous unbinding event, repetitively leaving the catalytic cleft (red color). Interestingly, a strong correlation was identified between the topological localization of the fragments and their $RMSF_{avg}$, with those ligands occupying the highly flexible S2 subsite also showing the more pronounced propensity in losing the experimental-solved binding mode. This information not only offers valuable insights for the H2L optimization phase but also opens up questions about the suitability of MD-based approaches for the posing of ligands characterized by such limited structural stability.

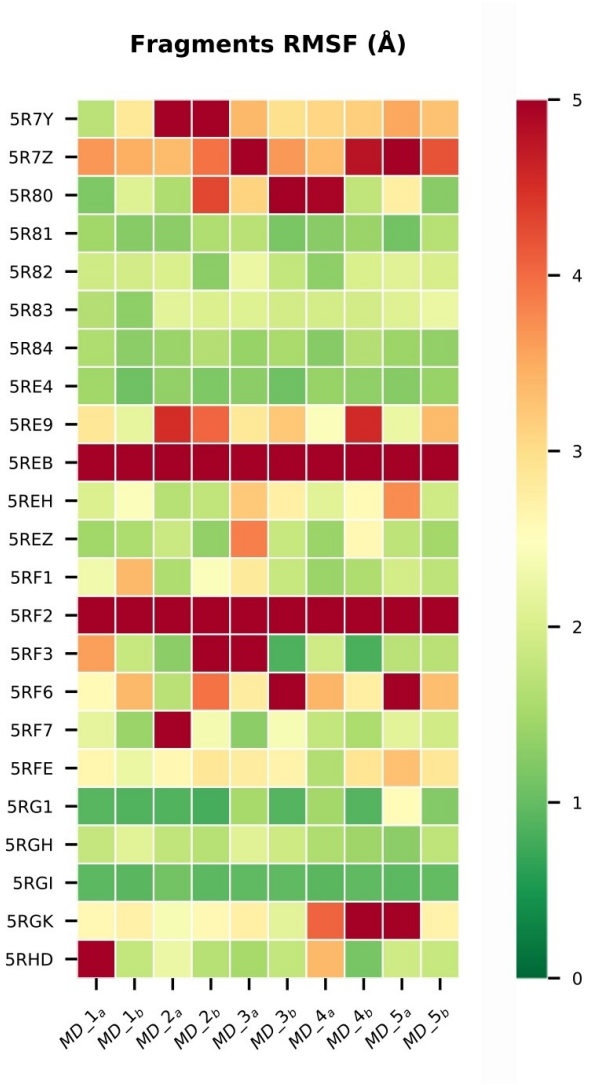

Figure 1. Fragment stability assessed by classical MD of the 23 crystallographic complexes. For each MD simulation collected (x-axis) starting from the crystallographic ligand-receptor complexes (y-axis), the pose stability value of the fragment is herein reported through a heatmap representation. The colorimetric scale, from green to red, quantitatively represents the RMSF computed for each ligand heavy atoms (0 to 5 Å scale). The MD simulation were carried out on each subunit of the M^{pro} functional dimer resulting in two set (labelled a and b) for each of the 5 runs.

Fragments posing through HT-SuMD

HT-SuMD protocol has been applied to investigate the binding mechanism of the 23 non-covalent fragments against the unliganded crystal structure of the SARS-CoV-2 M^{pro} (PDB ID 6YB7). As accurately described in the original paper, HT-SuMD manages the preparation, collection, and analysis of multiple SuMD simulations in an automatic modality, only requiring the binding pocket localization as initial information. SuMD, briefly, exploiting a tabu-like supervision algorithm that monitors in times variations in the ligand-protein binding site distances, could be considered an enhanced sampling approach improving the efficiency with which rare events, such as binding, are described.[26,27] For each fragment investigated, a solvated MD simulation box has been set up (a detailed description is reported on supplementary materials) and equilibrated after distancing the ligand at least 30 Å away from the protein catalytic cleft, to avoid premature intramolecular interactions. Also in this case, as an attempt to increase the robustness of the results, 10 SuMD replicates have been collected, resulting in a total of 6.3 μs of simulation time. The ensemble of 230 trajectories describing different fragment binding pathways has been later geometrically discretized through DBSCAN, a density-based clustering algorithm, which allows all the most populated ligand-protein states to emerge from the background noise.[28,29] In detail, a cluster is initialized if it contains at

least 25 similar fragments conformations, which therefore differ from each other by no more than 1.5 Å. Finally, each binding mode was qualitatively evaluated using the MM/GBSA approach to approximate the ligand-protein free energy of binding, thus allowing to perform a ranking of the predicted poses.[30] The accuracy of the predictions was assessed by comparing each cluster of fragment conformations identified with the respective crystallographic reference, computing the root-mean-square deviations (RMSD) of non-hydrogen atomic coordinates. The results obtained for the 23 M^{pro}crystallographic inhibitors have been extensively reported in the supplementary information (SI_HT-SuMD.xlsx) and graphically summarized in Figure 2, exploiting a colorimetric map to differentiate the correctness of the posing protocol. More specifically, for each fragment, the minimum RMSD ($RMSD_{min}$) and the average RMSD ($RMSD_{avg}$) values for the best cluster, i.e. the cluster closer to the crystallographic reference, were reported then comparing the predicted binding mode with the experimental one. The fragment posing exercise was considered correctly achieved if

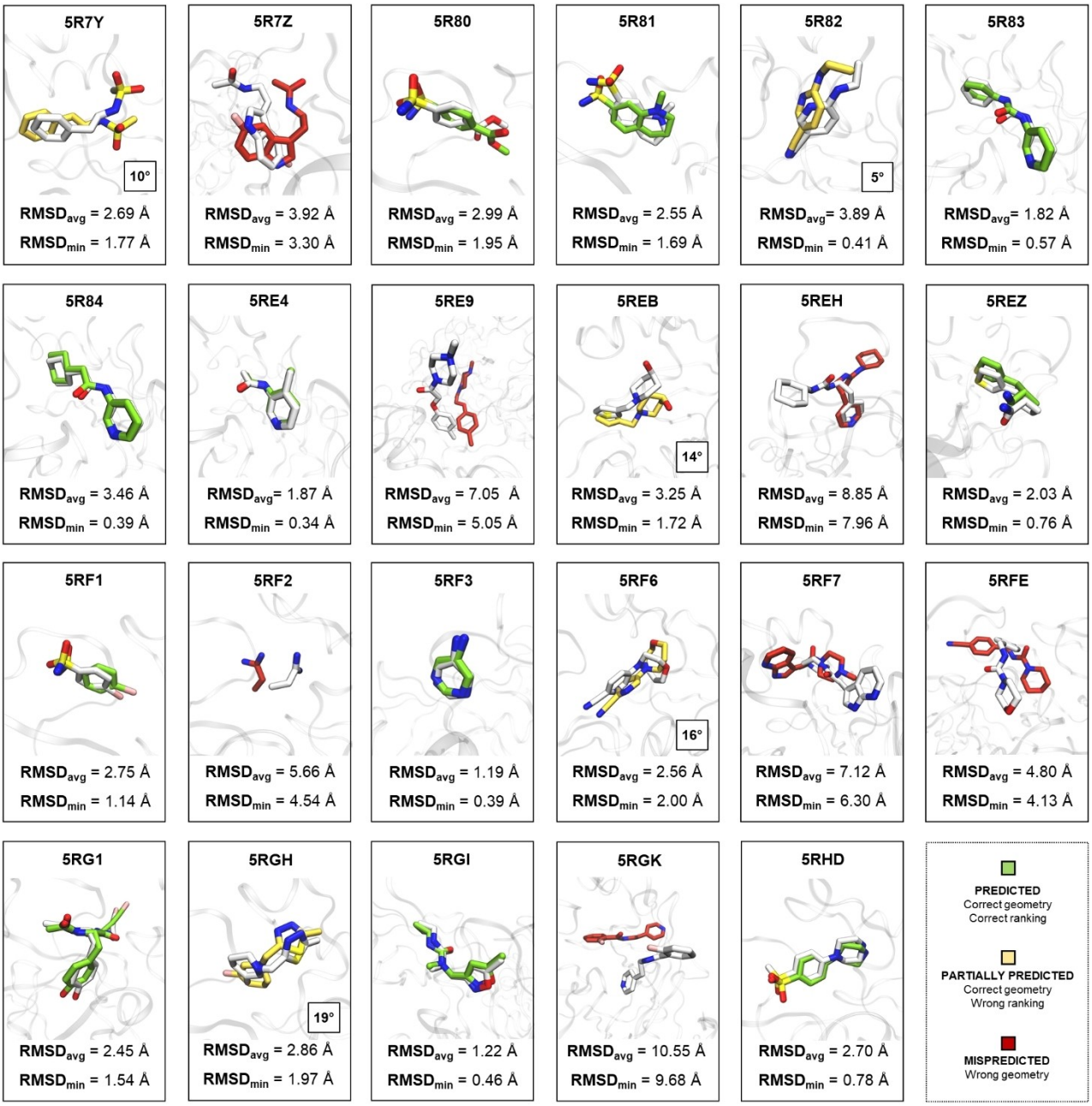

Figure 2. The results of the HT-SuMD posing protocol have been herein summarized. For each of the 23 fragments investigated the cluster of ligand conformations closes to the experimentally solved binding mode was reported, measuring the accuracy of the prediction through the $RMSD_{avg}$ and $RMSD_{min}$ values of the selected cluster. The crystallographic reference has been rendered in white color, while the HT-SuMD predicted binding modes have been differentiated in green, yellow, and red color, following the criteria described in the legend. In the case of partially predicted fragments, in which a good binding geometry was retrieved but erroneously ranked, the magnitude of the error has been underlined reporting the incorrect ranking position.

the $RMSD_{min}$ of the cluster selected falls below the cut-off value of 2 Å.

For 11 fragments out of 23, representing almost half of the considered cases, the protocol was able to identify and correctly rank the experimental binding mode (green-coloured molecules). Among these, the most noteworthy case is represented by the fragment with the PDB ID 5RGI, the only one targeting the S1′ subsite. HT-SuMD posing approach, fully exploring the conformational flexibility of the receptor, was able to reproduce the fragment crystallographic binding mode in an extremely accurate way, with an $RMSD_{min}$ value of 0.46 Å. This result is impressive since, in the unliganded M^{pro} structure chosen in this study, the S1′ pocket, due to a different orientation of the residues composing the catalytic dyad (H41 and C145), is initially inaccessible.

For the remaining 12 fragments, an in-depth analysis highlighted two orthogonal reasons underneath the HT-SuMD based posing failures. In 5 cases the MM/GBSA-based scoring method was unable to prioritize the experimental binding mode, even if it was exhaustively sampled by SuMD simulations (yellow-colored molecules). The incorrect ranking position was then reported in Figure 2 within a squared box, to underline the magnitude of the scoring error. This disagreement may be caused by limitations affecting the MM models, as errors in the fragments force field parameters or, more intriguingly, the crystallographic structures could capture only one of the possible accommodation states that the ligand can explore within the binding site.[31] In the other 7 cases instead, the experimental conformation was never sampled (red-colored molecules), suggesting possible MD-sampling issues that may be addressed by widening the number of SuMD replicates performed for each compound, however increasing the computational cost of our approach. The accuracy of HT-SuMD protocol, therefore, with 48% of correct binding mode predictions is greater than non-native docking-based protocols reported in the literature and in line with that of other MD-based fragments posing approaches.[32,33] It appears however evident how the posing of fragments still represents a tough pharmaceutical challenge, in particular, as suggested by Verdonk, for those characterized by a low-ligand efficiency (LE). Even our computational approach, in about half of the examined cases, fails to return a reliable result making its routine application very complex in a pharmaceutical drug discovery context.

To elucidate the applicability domain of HT-SuMD and better understand the limitations related to the implementation of MD-based protocols for the fragment binding modes prediction, we have therefore investigated if the fragment pose stability, a geometric-dynamic property, could impact the predictivity of our method. The fragment pose stability retraces the concept behind the structural stability criterion that has recently been discussed also by Barril's research group, as a complement to more traditional thermodynamic-based approaches in the identification of correct fragment-receptor binding mode.[34] HT-SuMD outcomes have therefore been compared, as reported in Figure 3, with the average values of atomic coordinates fluctuation (i.e. $RMSF_{avg}$) respectively

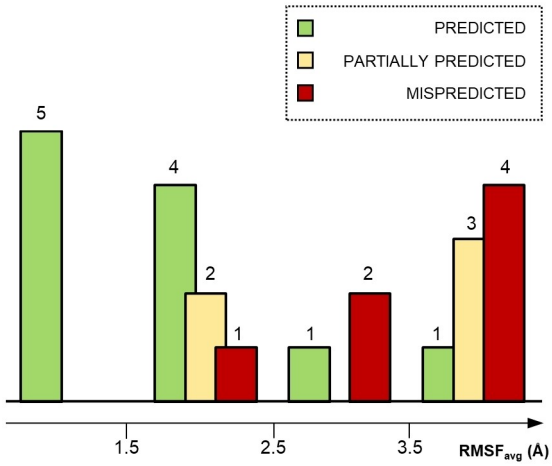

Figure 3. HT-SuMD predictions have been analyzed and related to the average fragment pose stability values ($RMSF_{avg}$) showed by each original crystallographic complex during the MD-based characterization study.

showed by each crystallographic fragment in the classical MD study previously discussed. Intriguingly, a clear pattern is noticeable since almost the totality of the correctly predicted binding modes (9/11) has been recovered for those fragments characterized by marked structural stability, with an $RMSF_{avg}$ value lower than 2.5 Å. Above this empirical cut-off, consistently most of the incorrect predictions concentrate, thus corroborating the existence of an inverse relationship linking together the stability of a crystallographic final state and the ability to correctly anticipate it through MD-based approaches, as our protocol configure.

Fragment poses stability as a confidence metric

The relationship described above could therefore be exploited to drive the analysis and the interpretation of HT-SuMD results, providing an observable with which distinguish reliable binding modes predictions from decoys. To test this hypothesis, the results collected through HT-SuMD posing protocol were retrospectively evaluated simulating a real screening scenario, in which crystallographic references are not available. Hence, for each of the 23 M^{pro}fragments previously investigated through HT-SuMD, the binding mode with the lowest MM/GBSA score was blindly selected, regardless of whether or not it corresponds to the original experimental pose. Then, multiple classical MD simulations 20 ns long were started from the predicted final states, to characterize their relative fragment pose stability. Results of this study have been summarized in Figure 4, sorting the data concerning the $RMSF_{SuMD}$ values, or the average fluctuations of SuMD-predicted binding poses, computed on the fragment's heavy atoms.

A first interesting aspect to underline is how almost the totality of the correct binding modes anticipated by HT-SuMD (green-colored molecules) only undergoes a mild conformational perturbation during classical MD simulations, in agree-

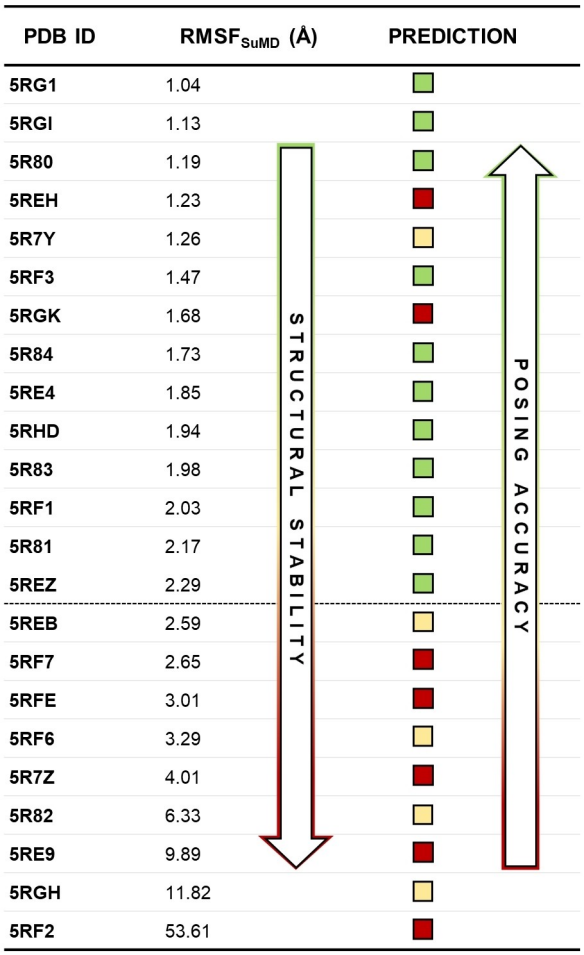

Figure 4. HT-SuMD predicted binding modes (i.e. the cluster of fragments conformations characterized by the lowest MM/GBSA value) have undergone an MD-based refinement step. The fragment poses stability of each prediction, measured as the RMSF$_{SuMD}$, has been exploited to rank HT-SuMD results, allowing in this way to efficiently prioritizing the correct binding modes at the expense of the incorrect ones. The dashed line delimits the empirical cut-off of 2.5 Å used to discriminate the reliability ofHT-SuMD posing prediction.

ment with the results described in the first part of the manuscript for the crystallographic complexes. On the contrary, incorrect binding mode (yellow and red-colored molecules) in most of the cases experience great lability when refined through MD simulations, sometimes even culminating in a spontaneous unbinding event of the fragment.

These observations corroborate the initial hypothesis, suggesting how a combination of HT-SuMD protocol for the posing of fragments with classical MD simulation for the refinement of results could represent an optimal operative pipeline, which allows overcoming some of the previously discussed methodological limitations. In this specific case indeed, the implementation of a geometric-dynamic property, namely the RMSF$_{SuMD}$, results extremely useful to qualitatively estimate the reliability of the in-silico predicted poses.

Observing the ranking reported in Figure 3, as the structural stability of the HT-SuMD predicted binding mode decreases, a worsening in posing accuracy occurs contextually. Intriguingly, also, in this case, 2.5 Å configure as a valuable empirical threshold which allows us to prioritize all the 11 correct fragment binding mode predictions. However, it is worth noting how the same cut-off is also responsible for the incorporation of three false positives, predictions characterized by remarkable structural stability, but which are nevertheless geometrically far from the crystallographic reference. For what concerns the fragment belonging to the PDB ID 5R7Y complex, HT-SuMD protocol has probably prioritized a metastable binding mode anticipating the experimental one, that has been nevertheless sampled through MD simulations but incorrectly scored by MM/GBSA. In the other two cases (PDB ID 5REH and 5RGK) the misprediction affects two fragments sharing a similar structure and interactivity. In the specific case of the 5REH complex, the HT-SuMD posing protocol has prioritized an alternative binding mode in which the pyridine portion of the fragment is correctly predicted, reproducing the key hydrogen bond interaction with H163 residue, while the remaining flexible portion is erroneously accommodated in the subsite S2 causing, as indicated in Figure 2, the high RMSD value of the cluster. This aspect is particularly interesting in the FBDD context, considering how the mild affinity profile characterizing these compounds could determine multiple recognition modes.

Conclusion

The elucidation of fragment binding modes in the early stages of FBDD campaigns still represents a tough medicinal chemistry task, which can be mitigated by the concomitant application of in-silico approaches. In this work, we have therefore investigated the geometric accuracy with which our recently developed computational protocol can reproduce experimentally solved fragment-receptor complexes. For this purpose, the XRC structures of 23 non-covalent fragments targeting SARS-CoV-2 M^{pro}, a pharmaceutical hot target in this actual COVID-19 pandemic, were exploited. HT-SuMD, as summarized in Figure 5, samples for each fragment multiple binding trajectories (Box 1), which are subsequently geometrically discretized through DBSCAN clustering and energetically evaluated using the MM/GBSA approach (Box 2). Our methodology was able to recover and prioritize in almost half of the cases taken into consideration (48%) the original fragment bound geometry, with an accuracy comparable to that described for other MD-based posing approaches.

Intriguingly, a clear correlation has been identified between HT-SuMD posing accuracy and the stability of the respective crystallographic complexes, with most of the correct binding modes predictions retrieved for those fragments characterized by a low RMSF$_{avg}$. In light of this aspect, a refinement step of HT-SuMD results through classical MD simulations has become an integrative part of our posing protocol (Figure 5– Box 3). More specifically, the structural stability of the predicted binding mode, i.e. the RMSF$_{SuMD}$, has been exploited and validated as a metric to qualitatively estimate the reliability of each single in-silico prediction. In this way, it was possible to effectively

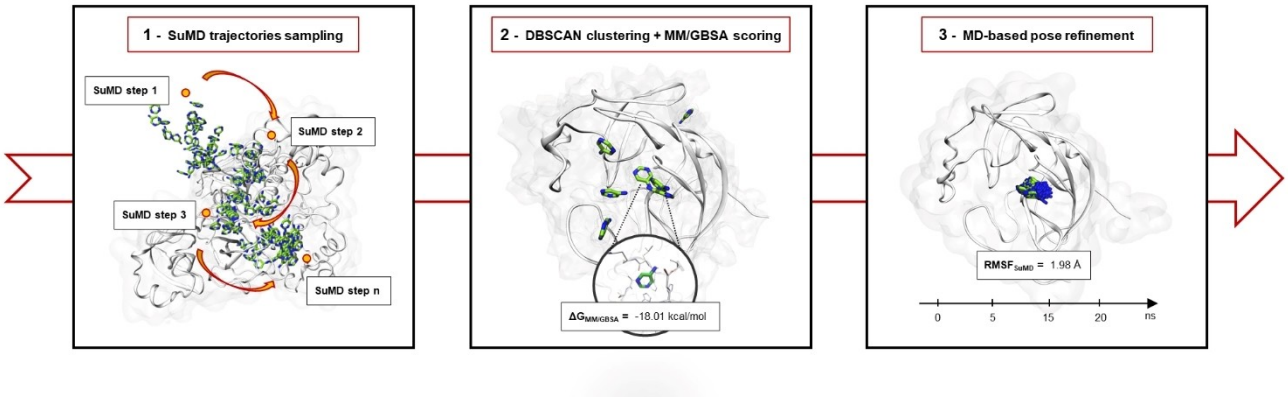

Figure 5. HT-SuMD protocol for the posing of fragments mainly consists in three operative steps, that are respectively summarized in this graphical workflow. In detail, supervised MD simulations are exploited to sample multiple binding trajectories for all the fragments analyzed (1), then DBSCAN clustering algorithm allows to identify of the most populated ligand conformation, which is energetically evaluated using MM/GBSA scoring method (2). The in-silico predicted binding modes finally undergo an MD-based refinement step, using the $RMSF_{SuMD}$ as a metric to qualitatively characterize the posing reliability.

rank and prioritize the 11 correct HT-SuMD binding poses while discharging the ones characterized by a marked instability that was mainly revealed as incorrect predictions. This concept is exemplified in Video1 (supplementary information), reporting how MM/GBSA, a thermodynamic-based approach, fails in distinguishing a correct form and incorrect fragment binding pose, while the subsequent MD refinement steps allow highlighting a marked difference between the two different predictions, in terms of $RMSF_{SuMD}$.

Despite these preliminary encouraging results, which must be certainly consolidated with further case studies, an improvement in the fragment posing accuracy is however still desirable. From this perspective, the ever-increasing computing power that will be available in the next years coupled with the continuous optimization of the conformational sampling algorithm, as well as the force fields model used, could pave the way for the development of more accurate fragment posing protocols, that could massively impact many in-silico FBDD pipelines.

Acknowledgements

MMS lab is very grateful to Chemical Computing Group, OpenEye, and Acellera for the scientific and technical partnership. MMS lab gratefully acknowledges the support of NVIDIA Corporation with the donation of the Titan V GPU, used for this research. This scientific work has been financially supported by MIUR (PRIN2017, No. 2017MT3993) and by Fondazione Cariparo (An Integrated Strategy for the Fast Discovery of SARS-CoV-2 Main Protease (M^{pro}) Inhibitors, No. 55812).

Conflict of Interest

The authors declare no conflict of interest.

Keywords: FBDD · SBDD · Molecular Dynamics · SARS-CoV-2 · Posing

[1] C. W. Murray, D. C. Rees, *Nat. Chem.* **2009**, *1*, 187–192.
[2] D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, H. Jhoti, *Nat. Rev. Drug Discovery* **2016**, *15*, 605–619.
[3] C. Jacquemard, E. Kellenberger, *Expert Opin. Drug Discovery* **2019**, *14*, 413–416.
[4] F. Giordanetto, C. Jin, L. Willmore, M. Feher, D. E. Shaw, *J. Med. Chem.* **2019**, acs.jmedchem.8b01855.
[5] R. J. Hall, P. N. Mortenson, C. W. Murray, *Prog. Biophys. Mol. Biol.* **2014**, *116*, 82–91.
[6] P. J. Hajduk, J. Greer, *Nat. Rev. Drug Discovery* **2007**, *6*, 211–219.
[7] B. J. Davis, S. D. Roughley, *Annu. Rep. Med. Chem.* **2017**, *50*, 371–439.
[8] R. Ma, P. Wang, J. Wu, K. Ruan, *Molecules* **2016**, *21*, DOI 10.3390/molecules21070854.
[9] B. J. Davis, D. A. Erlanson, *Bioorg. Med. Chem. Lett.* **2013**, *23*, 2844–2852.
[10] J. Schiebel, N. Radeva, S. G. Krimmer, X. Wang, M. Stieler, F. R. Ehrmann, K. Fu, A. Metz, F. U. Huschmann, M. S. Weiss, U. Mueller, A. Heine, G. Klebe, *ACS Chem. Biol.* **2016**, *11*, 1693–1701.
[11] J. Wielens, S. J. Headey, D. I. Rhodes, R. J. Mulder, O. Dolezal, J. J. Deadman, J. Newman, D. K. Chalmers, M. W. Parker, T. S. Peat, M. J. Scanlon, *J. Biomol. Screening* **2013**, *18*, 147–159.
[12] C. Sheng, W. Zhang, *Med. Res. Rev.* **2013**, *33*, 554–598.
[13] J. Mortier, C. Rakers, R. Frederick, G. Wolber, *Curr. Top. Med. Chem.* **2012**, *12*, 1935–1943.
[14] L. R. de Souza Neto, J. T. Moreira-Filho, B. J. Neves, R. L. B. R. Maidana, A. C. R. Guimarães, N. Furnham, C. H. Andrade, F. P. Silva, *Front. Chem.* **2020**, *8*, 93.
[15] M. Bissaro, M. Sturlese, S. Moro, *Drug Discovery Today* **2020**, DOI 10.1016/j.drudis.2020.06.023.
[16] F. Ferrari, M. Bissaro, S. Fabbian, J. De Almeida Roger, S. Mammi, S. Moro, M. Bellanda, M. Sturlese, *J. Enzyme Inhib. Med. Chem.* **2021**, *36*, 1–14.
[17] "Coronavirus disease 2019 (COVID-19) Situation Report – 43," can be found under https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200303-sitrep-43-covid-19.pdf?sfvrsn=2c21c09c_2, n.d.
[18] A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer, P. Lukacik, C. D. Owen, E. Resnick, C. Strain-Damerell, A. Aimon, P. Ábrányi-Balogh, J. Brandão-Neto, A. Carbery, G. Davison, A. Dias, T. D. Downes, L. Dunnett, M. Fairhead, J. D. Firth, S. P. Jones, A. Keeley, G. M. Keserü, H. F. Klein, M. P. Martin, M. E. M. Noble, P. O'Brien, A. Powell, R. N. Reddi, R. Skyner, M. Snee, M. J. Waring, C. Wild, N. London, F. von Delft, M. A. Walsh, *Nat. Commun.* **2020**, *11*, 1–11.
[19] T. C. M. Consortium, H. Achdout, A. Aimon, E. Bar-David, H. Barr, A. Ben-Shmuel, J. Bennett, M. L. Bobby, J. Brun, S. BVNBS, M. Calmiano, A. Carbery, E. Cattermole, J. D. Chodera, A. Clyde, J. E. Coffland, G. Cohen,

J. Cole, A. Contini, L. Cox, M. Cvitkovic, A. Dias, A. Douangamath, S. Duberstein, T. Dudgeon, L. Dunnett, P. K. Eastman, N. Erez, M. Fairhead, D. Fearon, O. Fedorov, M. Ferla, H. Foster, R. Foster, R. Gabizon, P. Gehrtz, C. Gileadi, C. Giroud, W. G. Glass, R. Glen, I. Glinert, M. Gorichko, T. Gorrie-Stone, E. J. Griffen, J. Heer, M. Hill, S. Horrell, M. F. D. Hurley, T. Israely, A. Jajack, E. Jnoff, T. John, A. L. Kantsadi, P. W. Kenny, J. L. Kiappes, L. Koekemoer, B. Kovar, T. Krojer, A. A. Lee, B. A. Lefker, H. Levy, N. London, P. Lukacik, H. B. Macdonald, B. MacLean, T. R. Malla, T. Matviiuk, W. McCorkindale, S. Melamed, O. Michurin, H. Mikolajek, A. Morris, G. M. Morris, M. J. Morwitzer, D. Moustakas, J. B. Neto, V. Oleinikovas, G. J. Overheul, D. Owen, R. Pai, J. Pan, N. Paran, B. Perry, M. Pingle, J. Pinjari, B. Politi, A. Powell, V. Psenak, R. Puni, V. L. Rangel, R. N. Reddi, S. P. Reid, E. Resnick, M. C. Robinson, R. P. Robinson, D. Rufa, C. Schofield, A. Shaikh, J. Shi, K. Shurrush, A. Sittner, R. Skyner, A. Smalley, M. D. Smilova, J. Spencer, C. Strain-Damerell, V. Swamy, H. Tamir, R. Tennant, A. Thompson, W. Thompson, S. Tomasio, A. Tumber, I. Vakonakis, R. P. van Rij, F. S. Varghese, M. Vaschetto, E. B. Vitner, V. Voelz, A. von Delft, F. von Delft, M. Walsh, W. Ward, C. Weatherall, S. Weiss, C. F. Wild, M. Wittmann, N. Wright, Y. Yahalom-Ronen, D. Zaidmann, H. Zidane, N. Zitzmann, *bioRxiv* **2020**, 2020.10.29.339317.

[20] M. De Vivo, M. Masetti, G. Bottegoni, A. Cavalli, *J. Med. Chem.* **2016**, *59*, 4035–4061.

[21] V. Salmaso, S. Moro, *Front. Pharmacol.* **2018**, *9*, DOI 10.3389/fphar.2018.00923.

[22] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–242.

[23] G. Bolcato, M. Bissaro, M. Pavan, M. Sturlese, S. Moro, *Sci. Rep.* **2020**, *10*, 20927.

[24] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

[25] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157–1174.

[26] D. Sabbadin, S. Moro, *J. Chem. Inf. Model.* **2014**, *54*, 372–376.

[27] A. Cuzzolin, M. Sturlese, G. Deganutti, V. Salmaso, D. Sabbadin, A. Ciancetta, S. Moro, *J. Chem. Inf. Model.* **2016**, *56*, 687–705.

[28] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, *J. Mach. Learn. Res.* **2011**.

[29] X. Ester, M. Kriegel, H. P. Sander, J. Xu, *Proc. 2nd Int. Conf. Knowl. Discov. Data Min.* **1996**, *96*, 226–231.

[30] D. M. Y. P. A. K. D. A. Case, K. Belfon, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham, III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, G. Giambasu, M. K. Gilson, H. Gohlke, A. W. Goetz, R. Harris, S. Izadi, S. A. Izmailov, K. Kasavajhala, A. Kovalenko, R. Krasny, T, **2020**.

[31] D. L. Mobley, K. A. Dill, *Structure* **2009**, *17*, 489–498.

[32] M. L. Verdonk, I. Giangreco, R. J. Hall, O. Korb, P. N. Mortenson, C. W. Murray, *J. Med. Chem.* **2011**, *54*, 5422–5431.

[33] N. M. Lim, M. Osato, G. L. Warren, D. L. Mobley, *J. Chem. Theory Comput.* **2020**, *16*, 2778–2794.

[34] M. Majewski, X. Barril, *J. Chem. Inf. Model.* **2020**, *60*, 1644–1651.

Manuscript received: March 4, 2021
Accepted manuscript online: April 2, 2021
Version of record online: May 6, 2021