Fragment tailoring strategy to design of chemical entities as potential binders of novel corona virus main protease

Chinmayee Choudhury

1Department of Experimental Medicine and Biotechnology, PGIMER, Sector 12, Chandigarh-160012, India. Email: chinmayee.choudhury@gmail.com

Abstract

The recent pandemic of novel corona virus infections (COVID19) has put the world on serious alert. This is caused by a recent form of a positive sense RNA virus (nCoV) of coronaviridae family which is known to cause respiratory tract infections in humans. Absence of any specific drugs, vaccines or treatment measures for this deadly virus warrants intense research to design new chemical entities in order to inhibit the viral replication in human cells. The main protease of nCov (nCov-MP) cleaves the long polyprotein chains to release functional proteins required for replication of the virus and thus is a potential drug target. The current study employs state of art computational methods to design new molecules by linking molecular fragments which specifically bind to different constituent sub-pockets of the nCov-MP binding site. A huge library of 191678 fragments was screened against the binding cavity of nCov-MP and high affinity fragments binding to adjacent sub-pockets were tailored to generate new molecules. These newly formed molecules were further subjected to molecular docking, ADMET property filters and MMGBSA binding free energy calculations to select 17 best molecules (named as MP-In1 to Mp-In17), which showed interactions with the key binding site residues as the reference ligand. Nine out of these 17 molecules with better MMGBSA binding free energy than the reference molecule, were subjected to molecular dynamics simulations, which assessed the stabilities of their binding with nCov-MP. Eight molecules were found to form stable complexes with nCov-MP. These molecules can be further evaluated as potential starting points for nCov drug discovery.
Keywords: Covid-19, main protease, fragment-based drug discovery, docking, molecular dynamics
Introduction

The recent pandemic of novel corona virus infections (COVID19) has put the world on alert. This is caused by a positive sense RNA virus of *coronaviridae* family and *nidovirales* order which is known to cause respiratory tract infections in mammals including humans. A recent form of the virus, the novel coronavirus (nCoV) has emerged in china and has been named as SARS-CoV-2 because of the acute respiratory distress syndrome developing with these cases in infections which become severe with course of time. This is a zoonotic corona virus mediated disease which is third incidence after SARS and MERS [1]. The source has later been shown to have sequence homology as high as 96% with SARS-CoV of bats [2]. According to the latest WHO reports [3], 1773084 confirmed cases has been reported with 111652 deaths around a total of 166 countries which includes 9152 confirmed cases and 308 deaths in India. This pandemic is spreading exponentially and has become an issue of serious concern for the whole world. In the absence of any specific drugs and treatment measures, WHO is emphasizing on hand washing, personal protection, use of hand sanitizers and social distancing and isolation for prevention of spread of disease and contamination. This has proven effective in some countries to curtail the spread and they are still in phase 1 and 2 of the epidemic spread. However, in certain parts of the world it has become an issue of major and immediate concern due to advanced phases of epidemic [3]. Possible treatment strategies and methods have become urgent needs for the world. Various possible drug treatments have been used with some success and some negative studies. Repurposing the existing drugs as low hanging fruits is being used as the first strategy in search of a possible treatment for the disease. The various drugs tried so far are oseltamivir [4], systemic steroids in severe respiratory involvement which is still inconclusive [5], Lopinavir/Ritonavir with both positive [6] and equivocal results [7], chloroquine [8] phosphate/ Hydroxychloroquine [9]. Also some results have been seen with Ramdesivir [10, 11], Tocilizumab, RNA polymerase inhibitors like Favipiravir [11] and JAK-
STAT inhibitors like Baricitinib, Fedratinib and ruxolitinib [12]. Spike protein of the virus has been also been identified for vaccine development and studies are under process to develop a vaccine for the virus [13, 14]. The 32 kb long RNA genome of nCoV [15] codes for its structural proteins such as spike glycoprotein which facilitates the entry of the virus into the host cells through interaction with the host enzyme ACE2 [16], the nucleocapsid, envelope and other membrane proteins and the non-structural proteins such as the chymotrypsin like main protease [17] which cleaves the long polyprotein chains to release functional proteins required for replication. Thus, these proteins can be exploited as potential drug targets and hunting for new chemical entities (NCEs) with fewer side effects is the need of the hour to combat COVID19 [18]. However, successful discovery of NCEs will hugely depend on proper understanding of the structure, interactions and dynamics of validated targets and the unexplored potential of their binding sites to bind new chemotypes. Computational methods have become indispensable for infectious disease drug discovery in last few decades [19] not only to understand the drug-target interactions [19-21] and delineate the structure activity relationship of small druglike molecules [22-23], but also for screening huge chemical libraries providing a fast and less expensive alternative to the traditional high throughput screening [24-26]. Among a plethora of computational drug design strategies, fragment based de novo design of molecules has gained immense popularity [27-30]. Usually, fragment hits show very high binding affinity with the receptors pertaining to their size. As starting points, they provide profound opportunity for subsequent optimization leading to chemical entities with improved pharmacokinetic properties compared to molecules obtained as hits from high throughput screenings [27, 28]. Another major plus point of fragments is the low molecular complexity as compared to that of drug-like molecules, and thus reducing the search space to be explored [28]. Further, possibility of an exponentially huge number of combinatorial molecules by linking high affinity fragments ensures novel drug like chemotypes. So, with experimental high
throughput screenings, simultaneous development of new computational fragment screening strategies would surely prove useful to significantly reduce the number of molecules to be tested experimentally [31]. In this study, we have considered the main protease of the nCov (nCov-MP) as our target of interest for fragment-based design of new inhibitors [32]. As the protease binding pocket is a large cavity with three to four prominent sub pockets, it provides an interesting scope for screening fragments for these sub pockets and then linking them to design new molecules with optimal binding with the protease. With the arrival of the very first structure (6LU7) of this protein in PDB [32], several groups have come up with interesting strategies such as artificial intelligence based de novo design [33], repurposing existing drugs that can bind this protein [34] or virtually screening large chemical databases against this protein [35] to mention a few. Till date RCSB protein data bank (PDB) reports more than 80 structures of the SARS-Cov-2 MP binding more than 80 different fragment like molecules [36]. These fragments bind to a particular sub-pocket of the large binding pocket of SARS-Cov-2 main protease. These recent additions in PDB have further strengthened our idea of designing inhibitors through fragment linking.

Materials and Methods

Preparation of datasets

Crystal structure of nCov-MP: More than 80 different crystal structures of the nCov-MP bound to diverse ligands have been deposited in PDB till date. For this study, we have considered one 6LU7, the first one to be deposited in PDB which binds a potential peptide-based inhibitor N3 [32]. The Protein Preparation Wizard (PPW) module [37] of Schrödinger software package, version 2019-2 was used to pre-process the macromolecular structure downloaded from PDB. Missing hydrogens were added, water molecules beyond 5 Å of the active site were removed and appropriate bond orders were assigned to the structure. The
residues/side chains unresolved in some of the crystal structures were repaired with prime [38] module in the PPW the protonation states of the polar residues were optimized with the protassign module of PPW, which uses PROPKA to predict pKa values (pH 7.0±2.0) and side chain functional group orientations. The structure was then subjected to restrained minimization (cutoff RMSD 0.3 Å) with impref to avoid steric clashes. The prepared structure was further used for preparation of grids, molecular docking and molecular dynamics (MD) simulations.

**Fragment libraries:** Fragment structures were downloaded in the form of .sdf files from 4 different publicly available libraries, viz., Asinex fragments and building blocks [39], FCH group’s [40] ‘all purpose’ fragment library, fluorine fragment library, fragment like acids, fragment like amines, fragment like amino acids, high fsp³ fragment library, spiro fragments and FCH special selection of fragment library and ChemDiv [41] fragment library and. A consolidated set of 191678 unique fragments were verified for the ‘rule of 3’ [42] agreement and considered for fragment-based design. All the fragments were subjected to preparation in LigPrep [43], generating their ionization states at pH 7.0 (± 2.0) using Epik ionizer.

**Fragment screening**

The prepared structure of 6LU7.pdb was directly used for receptor grid generation. ‘Receptor Grid Generation’ module of Schrödinger was utilized to define interaction grids for molecular docking keeping the centroid of the peptide like cocrystal-ligand as grid the centre. The size of the interaction grid was fixed to 14 Å for inner box and 20Å as outer box. The fragment library with 191678 fragments was subjected to docking calculations. Molecular docking calculations were performed using the Glide module of Schrödinger software [44] package with standard precision (SP) mode. 3 best poses were generated for each fragment. OPLS_2005 force field [45] was used for docking with all default parameters.
**Fragment joining and de novo compound design**

Top fragment hits with docking score < -7.00 were retained from each SP fragment-docking calculations for further design. The ‘combine fragments’ panel from the library design module was used for direct joining of the fragments prepositioned at different regions of the nCov-MP binding site to design new compounds. The panel joins the fragments by identifying feasible bonds that can be formed between the fragments. The angle between the bond directions were set to be 15°. The maximum distances between the two bonding atoms from different fragments were set to be 1Å, while the minimum distance between the centroids two fragments was set to be 2 Å. All bonds attached to hydrogen or halogens in a fragment were chosen for breaking and re-joining to another fragment. All atoms of the newly built molecule were subjected to minimization. 3 rounds of such fragment joining were performed where, in the first-round pairs of fragments were joined and in the next round the resultants of the first round were used as inputs to combine up to 4 fragments and so on.

**Molecular docking and binding energy estimation**

All the newly formed molecules were docked to the binding site of nCov-MP using Glide module of Schrodinger Suite. The same grid that was used for fragment screening was used for this docking too. Docking was performed in two sub-steps i.e., the SP docking and an extra precision (XP) docking [46]. 5 best docked poses were generated for each newly designed molecule and OPLS_2005 force field was used for docking with all default parameters. The resultant complexes of the molecules with nCov-MP were further submitted for binding energy estimation, where Molecular Mechanics-Generalized BornSurfaceArea (MM/GBSA) based binding free energy (ΔG_{bind}) were computed for the complexes using Prime module [38].

**Evaluation of physicochemical and pharmacokinetic properties and synthetic accessibility**
QuickProp module [47] of Schrodinger were employed to calculate the drug like properties and predict the physicochemical and pharmacokinetic (absorption, distribution, metabolism, excretion and toxicity) properties of all the new molecule hits selected in the previous section. A synthetic accessibility score was also predicted for each molecule from SwissADME server [48]. All the molecules which violate no druglikeliness rules were identified and then, 17 best molecules were selected based on their $(\Delta G_{\text{bind}})$. 

**Molecular dynamics (MD) simulations**

MD simulations were carried out on the complexes of nCov-MP with nine selected molecules using the Desmond MD simulation package (release 2018) of Schrodinger [49]. The OPLS_2005 [45] force field was employed for the protein-ligand complexes. Using the system builder tool of desmond the complexes were solvated in a cubical water box (TIP3P water model) keeping 12 Å buffer space in x, y and z dimensions. Each system was neutralized by adding appropriate counter ions and an ionic concentration of 0.15 M was maintained by adding Na+ and Cl− ions. The systems were minimized with 10000 steepest descent steps followed by gradual heating from 0 to 300 K, under NVT ensemble. The systems were thermally relaxed before the production run using Nose-Hoover Chain thermostat method for 5 ns and 5 ns of pressure relaxation with Martyna-Tobias-Klein barostat method. Finally, 50 ns production run under NPT ensemble was carried out for each system using a cutoff distance of 12 Å for non-bonded interactions. Simulation quality analysis, simulation event analysis, simulation interaction diagrams were used for trajectory analyses. **Figure 1** shows the overall workflow of the study.

(Insert Figure 1 here)

**Results and discussion**
**Binding site of nCov-MP and the experimental structures**

Crystal structure of nCov-MP reveals that the overall structure of nCov-MP is a combination of three domains [32]. The first and second domains (DI and DII) have an antiparallel β-barrel structure, where residue 8-100 comprise the DI and residues 102 to 184 form the DII. Residues 201 to 303 form the third domain (DIII) which is a combination of 5 α-helices and the connecting link between DII and DIII is a long loop (L1) formed by residues 185 to 200. The substrate binding site of nCov-MP is situated at the junction between DI and DII, which also extends up to L1. Figure S1 shows the domains and the overall secondary structure of the protein. Figure 2 shows the binding site sub-pockets and interactions of N3 with nCov-MP.

(Insert Figure 2 here)

As of now 88 different crystal structures of nCov-MP have been reported to PDB, most of which are from the PanDDA analysis group depositions [36], where each of the structures binds a unique ligand at different cavities all over the protein structure. Some of these small fragment-like ligands binding to the substrate binding site, mostly occupy one or two sub-pockets of the huge substrate binding cavity. A quick look at the positioning of the peptide-based ligand N3 in 6LU7.pdb reveals that, different constituent fragments of N3 comfortably occupy almost all sub-pockets of the huge binding site exploiting the potential of the binding cavity to bind bigger molecules. N3 makes H-bonds with G143, H164, E166 and Q189. The MMGBSA binding energy of N3 was calculated to be -77.36 kcal/mol. Our study is inspired by these crystal structures as we attempt to computationally screen a huge fragment library against the crystal structure 6LU7 and then combine the best fragment hits from different sub-pockets to design new molecules.

**Fragment screening and linking**
A huge fragment library consisting of 191678 fragments was constructed from the publicly available Asinex, FCH and ChemDiv fragment libraries. These fragments were then screened against the energy grid created by keeping the N3 molecule as the grid centre and the dimensions of the inner and outer grid boxes were kept as large as 14 Å and 20 Å to cover all the sub-pockets inside and adjacent to the main binding pocket. Glide SP protocol was applied initially for fast screening of the fragments, which returned 40805 fragment hits with docking scores ranging from -9.79 to -5.55. 1974 top hits with SP docking score < -7 were chosen for further study. The nearest atoms of two fragments binding to different sub-pockets and are that are pre-positioned with respect to each other were joined to form a new molecule. The potential bonds that can be formed between two fragments placed in the adjacent sub-pockets were identified. These potential new bonds were identified by based on i) the distance between the atom that remains in one fragment and the atom that leaves in the other fragment must be less than 1 Å after adjustment of the bond lengths in each fragment to the ideal bond length for the new bond, ensuring the right alignment of fragments, ii) the angle between the bond directions should be less than 15° for a right rotational alignment of the fragments and iii) the distance between the fragment centroids should be below 2 Å to make sure that the fragments do not occupy the same location in the receptor cavities (Figure 1). However, the fragments can have some overlapping regions and both internal and peripheral (H and halogen) bonds were considered for breaking to form a bond with another fragment. Once such potential new bonds are identified, the fragments were linked in three rounds. In the first round, pairs of fragments were joined and in the next rounds of joining, the resultant molecules from the first round are considered for joining based on the above criteria and so on. The fragments, having no atoms as close as 1 Å, but are lying in adjacent binding pockets are still considered for linking by introducing methylene groups to each fragment, and then if the bond formation criteria are satisfied, the two fragments were joined by maximum two methylene linkers. The minimum
and maximum fragments to be joined were set to be 2 and 4 respectively. The fragments were randomly sampled in several non-redundant trials in order to manage the huge number of combinations. The number of such trials were set to be 20 for this study. Once the fragments are joined, the resultant structures are subjected to energy minimization restraining all heavy atoms in the fragments except for the linker atoms with a restraint of 100 kcal/mol. The fragment linking process taking 1974 selected high scoring fragments pre-positioned in different subpackets generated 487 novel molecules which were further screened using several levels of screening filters.

**Screening of newly designed molecules**

The 487 newly formed molecules were subjected to further screening using four different levels of filters. XP docking scores were used as the first level of screening. XP docking of calculations were performed using the same huge interaction grid generated for fragment docking in order to provide enough space for several conformers of the newly formed molecules to access all the sub-pockets and place themselves in binding site in a conformation that is energetically most favourable. The XP docking scores ranged from -14.13 to -6.997 for these molecules. The top scoring 172 molecules with Glide XP docking score < -10 were then subjected to the second level of filter i.e., calculation of ADMET properties with QuickProp module of Schrodinger which predicts many significant and pharmacologically relevant properties to estimate the drug likeliness of a given molecules. One can compare certain properties of a particular molecule with the given ranges of those of 95% of known drugs. Also, QuickProp can identify the presence of 30 types of reactive functional groups that may cause false positives during virtual screening studies. The important properties that are calculated and can be compared with the ranges of known drugs are MW, dipole, IP, EA, SASA, FOSA, FISA, PISA, WPSA, PSA, volume, #rotor, donorHB, accptHB, glob, QPpolrz, QPlogPC16,
QPlogPoct, QPlogPw, QPlogPo/w, logS, QPLogKhsa, QPlogBB, #metabol, etc. The descriptions of all these descriptors are listed in List S1. We have prioritized our screened compounds based on the number of property or descriptor values that fall outside the 95% range of similar values for 95% of known drugs (#stars) calculated by QuickProp, hence a smaller #stars suggests that a molecule is more drug-like than molecules with more #stars. We screened all the compounds that have passed our previous filters which have #star as 0. Thus, we obtained a total of 83 molecules which violate no drug likeness rules. The next level filter used for our screening was ΔGbind. The ΔGbind of N3 with nCov-MP was calculated to be -77.36 kcal/mol. Hence, we set a cut-off of -70.00 to select 17 molecules, which were named as MP-In1 to MP-In17. Table 1 lists the important ADMET properties of the selected molecules while Figure S2 shows the Molecular recognition interactions between MP-In1-MP-In17 and nCov-MP after molecular docking.

Insert Table 1 here

nCov-MP residues which participated in H-bonds or salt bridge interactions with the 17 selected ligands are mostly F140, L141, G143, S144, C145, H163, E166, Q189, T190. Interestingly, all these residues make salt bridge interactions with E166 and only MP-In7 makes a π-π stacking with H41. Apart from making H-bond and salt bridge interactions with residues of one or two sub-pockets, these molecules occupy the other sub-pockets by shape complementarity and hydrophobic contacts. We also observed a fused tricyclic fragment (SMILES: Cn1c2CCCCc2c2cccccc12) which occurs in four of the top 17 screened molecules, which makes factorable contacts with the sub-pocket present at the interface between DI, DII and L1, surrounded by residues C44 to M49 of D1, P168 to H172 of DII and F185 to Q192 of L1. The various components of the XP docking score and the MMGBSA binding energies of the 17 molecules have been given in Figure 3.
Nine (MP-In1 to MP-In9) out of the 17 molecules were found to have \( \Delta G_{\text{bind}} \) less than -75.00 kcal/mol, which were further subjected to molecular dynamics simulations to verify the stabilities of the interactions. Figure 4 shows the molecular interactions and binding pocket occupancy of these 9 molecules with nCov-MP.

**Understanding the stabilities of MP-In (1-9) and nCov-MP complexes**

MD simulation is a technique of apt choice in order to estimate the stability of the identified MP-In(1-9) and nCov-MP interactions under dynamical conditions. It also significantly enhances strong binding of ligands with the target [50]. The generated 9 complexes were submitted to MD simulations for 50 ns in an aqueous environment to study the evolution of these systems with respect to time. Various analyses were carried out on the MD trajectories to evaluate the stabilities of the complexes. The Root Mean Square Deviation (RMSD) was used to measure the average change in displacement of the whole protein-ligand complexes and the ligands solely for all 5000 frames in the trajectory with respect to the first frame. Figure 5 shows the RMSD of the protein (Orange) and ligands (Grey) in the nine systems in the last 1000 frames i.e., 10 ns.

RMSD of the protein in all systems indicated that the simulations have equilibrated, the fluctuations towards the end of the simulation are around some thermal average structures. Changes in RMSD values of all the protein-ligand complexes were of the order of 1-3 Å, which normally acceptable for small, globular proteins. The ligand RMSD values were calculated when the protein-ligand complex is first aligned on the protein backbone of the reference and
then the RMSD of the ligand heavy atoms were measured. As the RMSD of the ligands with respect to the protein observed for complexes with MP-In1, MP-In2, MP-In3, MP-In4, MP-In8 and MP-In9 were not significantly larger in the last 10 ns, indicate that these ligands bind stably inside the binding pockets. Though for complexes with MP-In5, and MP-In7, RMSD of the ligands with respect to the protein in the last 10 ns of simulation was slightly higher, due to the torsional flexibilities of some bonds attaching the -CF$_3$ and the pyridine groups in MP-In5 and the long aliphatic tail and the bond attaching the piperazine moiety in MP-In7. However, these molecules made several stable interactions with the active site residues and did not show a tendency to diffuse out of the cavity. The complex with MP-In6 showed significantly higher ligand RMSD as compared to the protein RMSD profiles, indicating their tendency to diffuse away from its initial binding site. As shown in Figure 6, MP-In6 did not occupy all the binding sub-pockets of the binding site too. We tried to understand these differences in stabilities of the complexes through the protein ligand contacts or interactions. Protein interactions with the ligand monitored throughout the simulation has been given in Figure 6. Interactions that occur more than 30% of the simulation time in each trajectory through 0 to 50 ns is shown in Figure S3.

Insert Figure 6 here

These interactions were categorized into 4 types: Hydrogen Bonds, Hydrophobic, Ionic and Water Bridges. As shown in Figure 6 and Figure S3, the residues G143, S144, C145, E166, Q189, T190, Q192 mostly make stable H-bonds with the ligands. Hydrogen-bonding properties of the molecules in drug design is considered important because of their strong influence on drug specificity, metabolism and adsorption. The hydrophobic contacts shown in the figure include π-Cation; π-π; and Other, non-specific interactions which generally hydrophobic amino acids and aromatic or aliphatic groups on the ligands. The residues H41, M49, M165, L167 and P168 mostly formed hydrophobic contacts with the hydrophobic fragments of MP-In(1-9).
H41 was also shown to form $\pi$-$\pi$ stacking with the aromatic rings of MP-In7 and MP-In8 (Figure S3). Ionic or polar interactions, between two oppositely charged atoms were mostly shown by E166 while T26, N142 formed stable water bridges i.e., hydrogen-bonded protein-ligand interactions mediated by a water molecule in complexes with MP-In-8 and MP-In9. It was observed that, for the complex with MP-In6, the fractions of the above interactions were significantly low (<0.5), as it lost all interactions except with E166 and N142, which might be the cause of its unstable binding and tendency to diffuse out of the cavity. Hence, all the other eight molecules, except MP-In6 might be considered as potential inhibitors of nCov-MP based on their stable molecular interactions, good binding energy comparable to the reference ligand and good shape complementarity and binding pocket occupancies.

**Summary**

The current study attempts to design new molecules by tailoring fragments that bind to various sub-pockets of the binding sites of nCov-MP. A huge library of publicly available molecular fragments was screened against nCov-MP binding site in order to obtain sub-pocket specific fragments. These fragments prepositioned in adjacent binding sub-pockets were linked to form new molecules. these new molecules were further screened against nCov-MP using extra precision docking, ADMET and druglike filters and MMGBSA free energy of binding to find 17 potential molecules, out of which nine showed better binding energies as compared to the reference inhibitor. MD simulations were run on these nine molecules to ensure the stabilities of their binding and interactions and eight of them showed stable binding through various stable molecular interactions such as H-bonding, salt bridges, hydrophobic contacts and water bridged H-bonds. We propose these eight molecules as potential starting points for experimentalists working on nCov-MP inhibitor design.

**Acknowledgements**
CC thanks Department of science and Technology for financial assistance in the form of DST-INSPIRE Faculty award, Dr. Anuradha Chakravarti, Head, Department of Experimental Medicine and Biotechnology, PGIMER, Chandigarh for providing required infrastructure and Schrodinger for providing short term licences for some of the modules.

References


[33] Bung, Navneet, Sowmya Ramaswamy Krishnan, Gopalakrishnan Bulusu, and Arijit Roy. "De novo design of new chemical entities (NCEs) for SARS-CoV-2 using artificial intelligence." (2020).


[35] Pendyala, Brahmaiah, and Ankit Patras. "In silico Screening of Food Bioactive Compounds to Predict Potential Inhibitors of COVID-19 Main protease (Mpro) and RNA-dependent RNA polymerase (RdRp)." (2020).


[40] http://fchgroup.net/fragment-libraries.php

[41] https://www.chemdiv.com/


**Table1.** ADMET properties* of the 17 selected molecules.

<table>
<thead>
<tr>
<th>Name</th>
<th>#stars</th>
<th>QPlogP C16</th>
<th>QPlogP oct</th>
<th>QPlogPw</th>
<th>QPlogK hsa</th>
<th>QPlogH ERG</th>
<th>QPlogP o/w</th>
<th>QPlogK p</th>
<th>QPlogB B</th>
<th>#metab</th>
<th>QPlogS</th>
<th>CNS</th>
<th>FOSA</th>
<th>FISA</th>
<th>Synthetic feasibilityy</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP-In1</td>
<td>0</td>
<td>15.27</td>
<td>25.32</td>
<td>12.57</td>
<td>1.10</td>
<td>-5.16</td>
<td>5.31</td>
<td>-4.35</td>
<td>-0.57</td>
<td>5</td>
<td>-6.33</td>
<td>1</td>
<td>440.80</td>
<td>119.48</td>
<td>4.98</td>
</tr>
<tr>
<td>MP-In2</td>
<td>0</td>
<td>16.82</td>
<td>28.47</td>
<td>15.97</td>
<td>0.43</td>
<td>-5.70</td>
<td>4.06</td>
<td>-4.64</td>
<td>-0.87</td>
<td>4</td>
<td>-6.27</td>
<td>0</td>
<td>562.96</td>
<td>131.24</td>
<td>5.58</td>
</tr>
<tr>
<td>MP-In3</td>
<td>0</td>
<td>15.25</td>
<td>27.87</td>
<td>15.26</td>
<td>0.84</td>
<td>-6.97</td>
<td>4.03</td>
<td>-7.18</td>
<td>-0.20</td>
<td>5</td>
<td>-4.92</td>
<td>0</td>
<td>315.15</td>
<td>150.09</td>
<td>4.3</td>
</tr>
<tr>
<td>MP-In4</td>
<td>0</td>
<td>15.04</td>
<td>27.54</td>
<td>19.16</td>
<td>-0.09</td>
<td>-6.35</td>
<td>1.51</td>
<td>-5.81</td>
<td>-1.81</td>
<td>6</td>
<td>-3.22</td>
<td>-2</td>
<td>332.10</td>
<td>211.19</td>
<td>4.6</td>
</tr>
<tr>
<td>MP-In5</td>
<td>0</td>
<td>15.70</td>
<td>26.50</td>
<td>14.47</td>
<td>0.79</td>
<td>-6.87</td>
<td>4.76</td>
<td>-4.29</td>
<td>-0.54</td>
<td>8</td>
<td>-5.79</td>
<td>0</td>
<td>151.48</td>
<td>140.40</td>
<td>4.31</td>
</tr>
<tr>
<td>MP-In6</td>
<td>0</td>
<td>14.79</td>
<td>28.35</td>
<td>19.53</td>
<td>0.08</td>
<td>-6.52</td>
<td>1.29</td>
<td>-7.70</td>
<td>-0.93</td>
<td>6</td>
<td>-1.89</td>
<td>-1</td>
<td>317.85</td>
<td>190.63</td>
<td>4.85</td>
</tr>
<tr>
<td>MP-In7</td>
<td>0</td>
<td>17.92</td>
<td>31.04</td>
<td>18.74</td>
<td>0.10</td>
<td>-3.80</td>
<td>1.76</td>
<td>-5.21</td>
<td>-1.85</td>
<td>8</td>
<td>-5.09</td>
<td>-2</td>
<td>219.90</td>
<td>216.91</td>
<td>4.82</td>
</tr>
<tr>
<td>MP-In8</td>
<td>0</td>
<td>17.12</td>
<td>30.53</td>
<td>17.10</td>
<td>0.78</td>
<td>-6.64</td>
<td>3.23</td>
<td>-8.80</td>
<td>-1.21</td>
<td>4</td>
<td>-4.23</td>
<td>-2</td>
<td>288.03</td>
<td>238.17</td>
<td>3.9</td>
</tr>
<tr>
<td>MP-In9</td>
<td>0</td>
<td>16.65</td>
<td>29.50</td>
<td>21.06</td>
<td>0.03</td>
<td>-6.20</td>
<td>1.85</td>
<td>-7.47</td>
<td>-1.16</td>
<td>5</td>
<td>-2.22</td>
<td>-2</td>
<td>340.75</td>
<td>195.09</td>
<td>4.6</td>
</tr>
<tr>
<td>MP-In10</td>
<td>0</td>
<td>15.16</td>
<td>22.74</td>
<td>12.72</td>
<td>0.78</td>
<td>-7.40</td>
<td>4.43</td>
<td>-3.70</td>
<td>-0.78</td>
<td>6</td>
<td>-5.76</td>
<td>-1</td>
<td>437.09</td>
<td>104.38</td>
<td>4.85</td>
</tr>
<tr>
<td>MP-In11</td>
<td>0</td>
<td>13.42</td>
<td>25.89</td>
<td>17.92</td>
<td>-0.23</td>
<td>-4.55</td>
<td>2.02</td>
<td>-4.58</td>
<td>-0.69</td>
<td>5</td>
<td>-3.38</td>
<td>0</td>
<td>358.08</td>
<td>130.79</td>
<td>5.1</td>
</tr>
<tr>
<td>MP-In12</td>
<td>0</td>
<td>15.24</td>
<td>26.47</td>
<td>16.37</td>
<td>0.22</td>
<td>-3.09</td>
<td>2.85</td>
<td>-5.22</td>
<td>-1.03</td>
<td>6</td>
<td>-4.07</td>
<td>-2</td>
<td>604.43</td>
<td>146.03</td>
<td>4.6</td>
</tr>
<tr>
<td>MP-In13</td>
<td>0</td>
<td>16.10</td>
<td>31.73</td>
<td>22.05</td>
<td>0.00</td>
<td>-5.34</td>
<td>1.83</td>
<td>-7.50</td>
<td>-0.78</td>
<td>4</td>
<td>-2.09</td>
<td>-1</td>
<td>361.04</td>
<td>183.14</td>
<td>5.59</td>
</tr>
<tr>
<td>MP-In14</td>
<td>0</td>
<td>16.72</td>
<td>34.43</td>
<td>20.69</td>
<td>0.42</td>
<td>-6.68</td>
<td>2.12</td>
<td>-9.43</td>
<td>-1.61</td>
<td>5</td>
<td>-4.13</td>
<td>-2</td>
<td>393.44</td>
<td>264.88</td>
<td>5</td>
</tr>
<tr>
<td>MP-In15</td>
<td>0</td>
<td>16.75</td>
<td>31.15</td>
<td>17.16</td>
<td>0.75</td>
<td>-6.54</td>
<td>2.99</td>
<td>-8.12</td>
<td>-1.10</td>
<td>6</td>
<td>-4.09</td>
<td>-2</td>
<td>585.43</td>
<td>192.52</td>
<td>4.98</td>
</tr>
<tr>
<td>MP-In16</td>
<td>0</td>
<td>13.19</td>
<td>25.58</td>
<td>20.28</td>
<td>-0.30</td>
<td>-4.21</td>
<td>0.79</td>
<td>-6.43</td>
<td>-1.52</td>
<td>7</td>
<td>-1.87</td>
<td>-2</td>
<td>366.70</td>
<td>209.63</td>
<td>5.35</td>
</tr>
<tr>
<td>MP-In17</td>
<td>0</td>
<td>14.24</td>
<td>26.21</td>
<td>12.71</td>
<td>0.99</td>
<td>-6.05</td>
<td>4.26</td>
<td>-6.25</td>
<td>-1.18</td>
<td>4</td>
<td>-6.38</td>
<td>-2</td>
<td>396.43</td>
<td>195.27</td>
<td>4.42</td>
</tr>
</tbody>
</table>

*Description of all the fields are listed in List S1 [47]. Values ranges from 1 to 10, a value of 1 indicates easily synthesizable, while 10 indicates very difficult to synthesize.
Figures

Figure 1. Overall workflow of the study
Figure 2. Positioning of N3 in the binding site sub-pockets and molecular interactions of N3 with nCov-MP, from the structure 6LU7.
Figure 3. Various components of the XP docking score and the MMGBSA binding energies (Kcal/mol) of the 17 selected molecules with MMGBSA dG_bind < -70 kcal/mol.
Figure 4. Molecular interactions and binding pocket occupancy of these 9 selected molecules with better MMGBSA binding energy with nCov-MP as compared to N3
Figure 5. RMSD of the protein (Orange) and ligands (Grey) in the nine complexes of MP-In(1-9) with nCov-MP in the last 1000 frames (10 n) of MD simulation. Each frame represents a snapshot saved at 10 ps interval.
Figure 6. Protein–Ligand Contacts monitored throughout the simulation. The stacked bar charts are normalized over the course of the trajectory; e.g., a value of 0.7 suggests that 70% of the simulation time the specific interaction is maintained. Values over 1.0 are possible when a protein residue makes multiple contacts of same subtype with the ligand.