



# Targeting the N-Terminus Domain of the Coronavirus Nucleocapsid Protein Induces Abnormal Oligomerization via Allosteric Modulation

Jia-Ning Hsu<sup>1†</sup>, Jyun-Siao Chen<sup>2†</sup>, Shan-Meng Lin<sup>1†</sup>, Jhen-Yi Hong<sup>1†</sup>, Yi-Jheng Chen<sup>2</sup>, U-Ser Jeng<sup>3</sup>, Shun-Yuan Luo<sup>2\*</sup> and Ming-Hon Hou<sup>1\*</sup>

<sup>1</sup>Institute of Genomics and Bioinformatics and Department of Life Sciences, National Chung Hsing University, Taichung, Taiwan, <sup>2</sup>Department of Chemistry, National Chung Hsing University, Taichung, Taiwan, <sup>3</sup>National Synchrotron Radiation Research Center, Department of Chemical Engineering, National Tsing Hua University, Hsinchu, Taiwan

## OPEN ACCESS

### Edited by:

Annalisa Pastore,  
King's College London,  
United Kingdom

### Reviewed by:

Paola Storici,  
Elettra Sincrotrone Trieste, Italy  
Peng Zhan,  
Shandong University, China

### \*Correspondence:

Shun-Yuan Luo  
syluo@dragon.nchu.edu.tw,  
Ming-Hon Hou  
mhho@nchu.edu.tw

<sup>†</sup>These authors have contributed  
equally to this paper and share first  
authorship

### Specialty section:

This article was submitted to  
Structural Biology,  
a section of the journal  
Frontiers in Molecular Biosciences

Received: 08 February 2022

Accepted: 17 March 2022

Published: 19 April 2022

### Citation:

Hsu J-N, Chen J-S, Lin S-M,  
Hong J-Y, Chen Y-J, Jeng U-S,  
Luo S-Y and Hou M-H (2022)  
Targeting the N-Terminus Domain of  
the Coronavirus Nucleocapsid Protein  
Induces Abnormal Oligomerization via  
Allosteric Modulation.  
Front. Mol. Biosci. 9:871499.  
doi: 10.3389/fmolb.2022.871499

Epidemics caused by coronaviruses (CoVs), namely the severe acute respiratory syndrome (SARS) (2003), Middle East respiratory syndrome (MERS) (2012), and coronavirus disease 2019 (COVID-19) (2019), have triggered a global public health emergency. Drug development against CoVs is inherently arduous. The nucleocapsid (N) protein forms an oligomer and facilitates binding with the viral RNA genome, which is critical in the life cycle of the virus. In the current study, we found a potential allosteric site (Site 1) using PARS, an online allosteric site predictor, in the CoV N-N-terminal RNA-binding domain (NTD) to modulate the N protein conformation. We identified 5-hydroxyindole as the lead via molecular docking to target Site 1. We designed and synthesized four 5-hydroxyindole derivatives, named P4-1 to P4-4, based on the pose of 5-hydroxyindole in the docking model complex. Small-angle X-ray scattering (SAXS) data indicate that two 5-hydroxyindole compounds with higher hydrophobic R-groups mediate the binding between N-NTD and N-C-terminal dimerization domain (CTD) and elicit high-order oligomerization of the whole N protein. Furthermore, the crystal structures suggested that these two compounds act on this novel cavity and create a flat surface with higher hydrophobicity, which may mediate the interaction between N-NTD and N-CTD. Taken together, we discovered an allosteric binding pocket targeting small molecules that induces abnormal aggregation of the CoV N protein. These novel concepts will facilitate protein-protein interaction (PPI)-based drug design against various CoVs.

**Keywords:** PPI-based drug design, n protein, allosteric modulator, COVID-19, MERS-CoV

**Abbreviations:** CoV, Coronaviruses; MERS-CoV N-NTD, N-terminal domain of the MERS-CoV nucleocapsid protein; FBDD, Fragment-based drug design; SARS, Severe acute respiratory syndrome; MERS, Middle East respiratory syndrome; COVID-19, Coronavirus disease 2019; N, Nucleocapsid; E, Envelope; M, Matrix; S, Spike; IDRs, Intrinsically disordered regions; NTD, N-terminal RNA-binding domain; CTD, C-terminal dimerization domain; RNP, Ribonucleoprotein; PPI, Protein-protein interaction; ALLINI's, Allosteric IN Inhibitors; IN, Integrase; DMF, Dimethylformamide; MR, Molecular replacement; SAXS, Small-angle X-ray scattering; Rg, Radii of gyration; Dmax, Maximum dimensions; EOM, Ensemble optimization method; IPTG, Isopropyl- $\beta$ -D-thiogalactopyranoside; PMSF, Phenylmethylsulfonyl fluoride; Ni-NTA, Nickel-nitrilotriacetic acid; NMR, Nuclear magnetic resonance; MES, 2-(N-morpholino) ethanesulfonic acid; PEG, Polyethylene glycol; HPLC, High-performance liquid chromatography.

## 1 INTRODUCTION

Recently emerging infections caused by coronaviruses (CoVs), including the severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS), and coronavirus disease 2019 (COVID-19), have led to a global public health emergency (Du Z. et al., 2020; Hui et al., 2020; Xu et al., 2020; Zhu et al., 2020). Interestingly, all the aforementioned CoVs belong to the  $\beta$  subgroup of the *Coronaviridae* family, which are characteristically single-stranded RNA viruses capable of circulating among mammals, such as bats, civet cats, and camels (Hui et al., 2020) and can cause respiratory illness in humans during epizootic spillovers. Therefore, the development of effective antiviral drugs and vaccines against these CoV infections is of the utmost importance to prevent their further spread (Ghosh et al., 2020; Khodadadi et al., 2020). The CoV genome contains four structural proteins: nucleocapsid (N), small envelope (E), matrix (M), and trimeric spike (S) glycoproteins, which are essential for virion assembly and function to complete the viral life cycle during infections. Among the structural proteins of CoVs, N proteins form a major structural component, are relatively evolutionarily conserved, and share the same modular organization (Chang et al., 2014) which consists of intrinsically disordered regions (IDRs): N-arm, C-arm, and two structural domains, including the N-terminal RNA-binding domain (NTD) and C-terminal dimerization domain (CTD) (Wootton et al., 2002; Chang et al., 2014). Dimeric N protein functions as a building block by binding to the viral RNA, forming a ribonucleoprotein (RNP) complex, a primary part of viral self-assembly, where subsequent viral replication and translation can proceed (Almazan et al., 2004; Chang et al., 2005; Jayaram et al., 2006; Zuniga et al., 2010). Furthermore, the N protein is also involved in regulating the host cell cycle and viral pathogenesis, ultimately facilitating production of the virus (McBride et al., 2014). These multifunctional characteristics of the N protein and its low rate of mutation make it a prominent target for the development of therapeutics against CoVs (Lin et al., 2014; Mori et al., 2015; Hu et al., 2017).

Two strategies can be used to inhibit the function of N proteins against CoVs. The first strategy is to develop antiviral agents that target the RNA-binding site of the N protein and specifically block RNP formation during viral replication. The second strategy is to affect the normal N protein oligomerization by inducing or inhibiting protein-protein interactions (PPIs) between N protein molecules. Modulation of PPIs can be achieved by the development of compounds that either bind to the PPI interaction surface, which directly affects the associated PPI (orthosteric modulators) or bind to the adjacent site of the PPI interface (allosteric modulators), which induces a conformational change to inhibit or enhance the PPIs of the target protein complexes (Thiel et al., 2012; Fischer et al., 2015; Modell et al., 2016; Petta et al., 2016). In general, allosteric modulators have the potential for greater subtype selectivity and do not need to compete with the bulky and relatively high-affinity PPI partners when compared to orthosteric ligands. Moreover, allosteric modulators possess no intrinsic

activity of their own, and thus, the side effects for clinical use could be reduced.

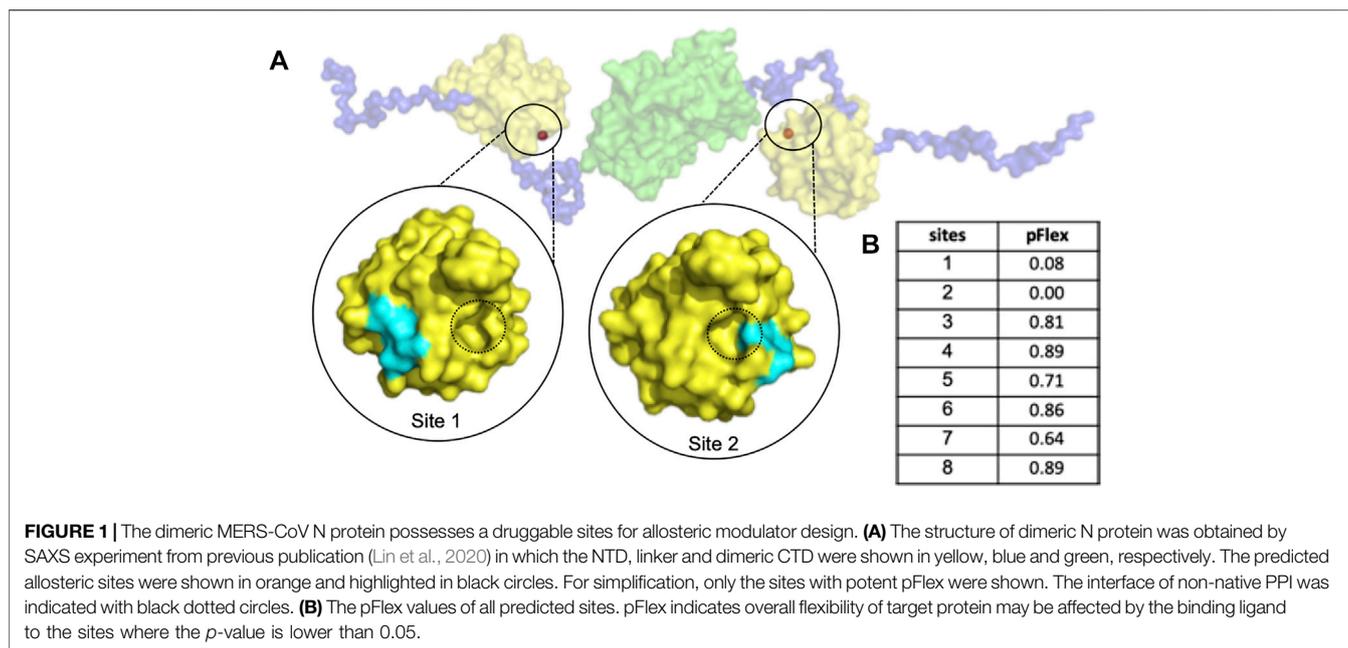
Most CoV N-NTD structures are folded in a monomeric conformation. In contrast, the CoV N-CTDs are always dimeric and responsible for N protein oligomerization via PPIs. In this study, we discovered one potential site, Site 1, in the MERS-CoV N-NTD with an online software, PARS, developed based on the normal-mode analysis method for designing an allosteric modulator. We utilized molecular docking technology to identify 5-hydroxyindole from an affordable 96-compound library designed by Muelle's group as the potential lead by targeting Site 1. We designed and synthesized four 5-hydroxyindole derivatives, named P4-1 to P4-4, based on the pose of 5-hydroxyindole in the docking model complex. Two 5-hydroxyindole derivatives, containing higher hydrophobic R-groups, exhibited the ability to elicit the oligomerization of N protein by small-angle X-ray scattering (SAXS) experiments. We also described the structure of MERS-CoV N-NTD complexed with P4 series compounds and revealed that higher hydrophobic R-group derivatives contribute to the hydrophobicity through the creation of flat interfaces between N-NTD and N-CTD. This unique approach based on targeting N-NTD to change the N protein structure by allosteric modulators can be potentially applied to discover novel drugs against CoV diseases, including SARS-CoV-2.

## 2 RESULTS

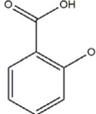
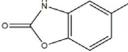
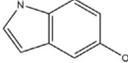
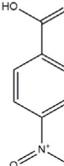
### 2.1 Discovery of the Protein Pocket of N-NTD for Drug Design

The abnormal aggregation of the N protein is induced by small molecules which manipulate a non-native PPI interaction that inhibits CoV replication (Lin et al., 2020); Thus, we proposed identifying a suitable cavity for the CoV N-NTD to induce abnormal N protein aggregation via allosteric modulation. To this end, we detected the pockets suitable for drug design using PARS, an online tool for allosteric site prediction based on the normal-mode analysis method (Panjkovich and Daura, 2014), with the dimeric structure of the N protein solved by SAXS experiments. Two sites on the protein surface were identified to have the potential for allosteric modulator design (Sites 1 and 2). Since site 2 is too close to the interface of non-native PPI, it may interfere with allosteric modulation; therefore, we selected site 1 for further design (**Figure 1**).

Docking studies were performed on the selected cavity using the Discovery Studio Client (v20.1.0.18287) with an assembled fragment library designed by Huschmann et al., characterized by a broad ligand diversity and high hit rate to the protein surface (Huschmann et al., 2016), to determine the starting chemical lead with high docking scores. Because low-molecular-weight ligands are considered to have the advantage of facilitating their optimization into potent compounds with drug-like properties (Murray and Rees, 2009), we excluded larger compounds. Thirty-two different structures with molecular weights (MWs) ranging from 150 to 250 Da were shortlisted. Second, the  $\pi$ -stacking interaction is commonly observed in protein-ligand complexes (Ferreira de Freitas and Schapira, 2017); therefore, we excluded ligands that



**TABLE 1** | Detailed information of the final fragments after screening procedures.

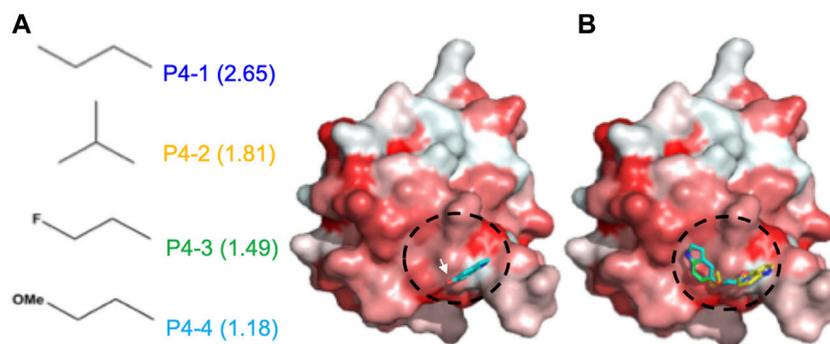
Molecule	Name	MW [Da]	tPSA	LigScore	logP (o/w)	Main Function
	Salicylic acid	138.12 g/mol	57.53	50.8536	1.27	Anti-inflammation
	chlorzoxazone	169.56 g/mol	38.33	57.729	1.8	Skeletal muscle relaxants
	5-Hydroxyindole	133.15 g/mol	36.02	52.7897	2	Anti-virus
	4-Nitrobenzoic acid	167.12 g/mol	83.12	50.4913	1.52	N/A

lacked aromatic rings in their structures. Third, the hydrophobic interaction plays a key role in protein aggregation; therefore, we excluded the hydrophilic fragments. Following the aforementioned criteria to exclude ligands, we narrowed down the potential ligands to the four fragments listed in **Table 1**. Among them, the 5-hydroxyindole scaffold stood out owing to its diverse biological activities and potential use in the medical industry (Kochanowska-Karamyan and Hamann, 2010; Ishikura et al., 2015). Furthermore, it has been used as a lead compound to develop antiviral agents, such as bufotenine (Vigerelli et al., 2020) and the antiviral drug for

influenza, arbidol (Herod et al., 2019; Du F. et al., 2020). Therefore, we chose 5-hydroxyindole (P4) for further experiments.

## 2.2 Design and Synthesis of 5-Hydroxyindole Derivatives by Targeting the Potential Allosteric Site of CoV N-NTD

Based on the pose of 5-hydroxyindole in the docking complex model (**Figure 2A**), we designed four derivative compounds with the modifications on the hydroxyl group at position C5, by which



**FIGURE 2** | Docking results of CoV N-NTD with 5-Hydroxyindole. **(A)** (left) The structures of each chemical moiety designed for 5-Hydroxyindole modification. The calculated  $m\text{LogP}$  values of each moiety were shown in brackets. (right) Surface representation of CoV N-NTD with the expected binding site of 5-Hydroxyindole, obtained by using the molecular docking. The surface was colored according to the hydrophobicity level at the protein surface. The chemical moieties were designed to add to the hydroxyl group of 5-Hydroxyindole (indicated by white arrow) to increase the hydrophobic contacts between each compound and the expected binding surface (indicated by black cycle) **(B)** Same as **(A)** except the 5-Hydroxyindole was replaced by P4-1, P4-2, P4-3, and P4-4 are shown in blue, yellow, green and cyan, respectively.

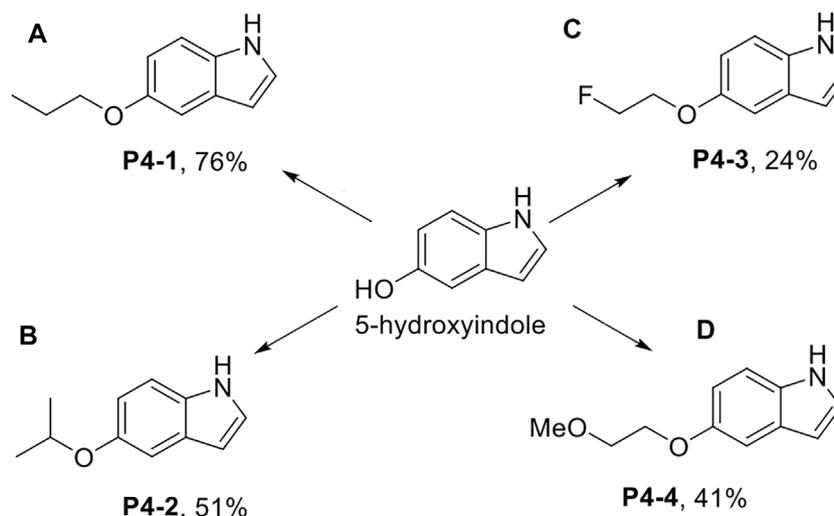
the modifications can extend to the predicted surface with higher binding affinity for allosteric modulation (**Scheme 1**). Furthermore, because hydrophobic interactions play an important role in the binding of allosteric modulators to their target proteins (Duan et al., 2019), we designed side chains with different degrees of hydrophobicity to strengthen these hydrophobic interactions. We named these four compounds as P4-1 to P4-4 and evaluated their abilities to modulate the protein oligomeric status. Molecular docking was performed again to evaluate the capacity of each compound to target the potent allosteric site. The results revealed that the P4-derived compounds fit well with the predicted surface (**Figure 2B**).

P4-1 to P4-4 were chemically synthesized. As shown in **Scheme 1**, we used 5-hydroxyindole and sodium hydride (NaH) with DMF as the solvent for deprotonation of the hydroxyl group and proceeded with 1-iodopropane to obtain

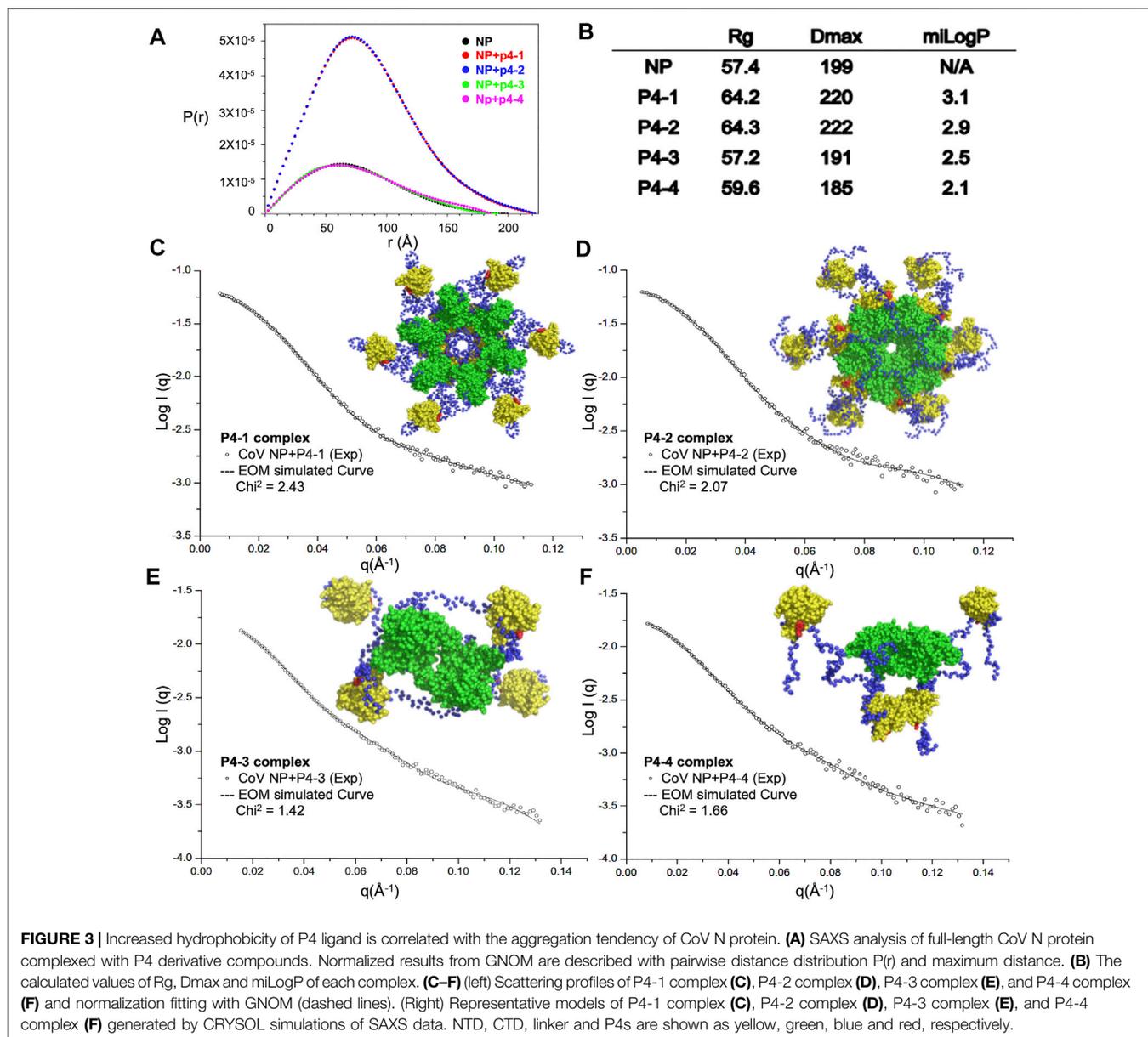
P4-1 at a yield rate of up to 76% (**Scheme 1A**). Similar to P4-1 production, P4-2 was obtained, followed by the reaction with 2-bromopropane at a yield rate of 51% (**Scheme 1B**). To generate P4-3, 5-hydroxyindole was reacted with 1-bromo-2-fluoroethane and potassium carbonate ( $\text{K}_2\text{CO}_3$ ) in acetone at a yield rate of 24% (**Scheme 1C**). As for P4-4, we referred to the procedures documented previously (Batista et al., 2008); 5-hydroxyindole was reacted with 1-iodo-2-methoxyethane and NaH in DMF to obtain P4-4 at a yield rate of 41% (**Scheme 1D**).

### 2.3 P4 Compounds Containing Hydrophobic R-Group Elicit Abnormal Oligomerization of CoV N-CTD by Targeting CoV N-NTD

To understand whether the designed modulators of N-NTD can elicit the abnormal aggregation of N protein, we purified full-

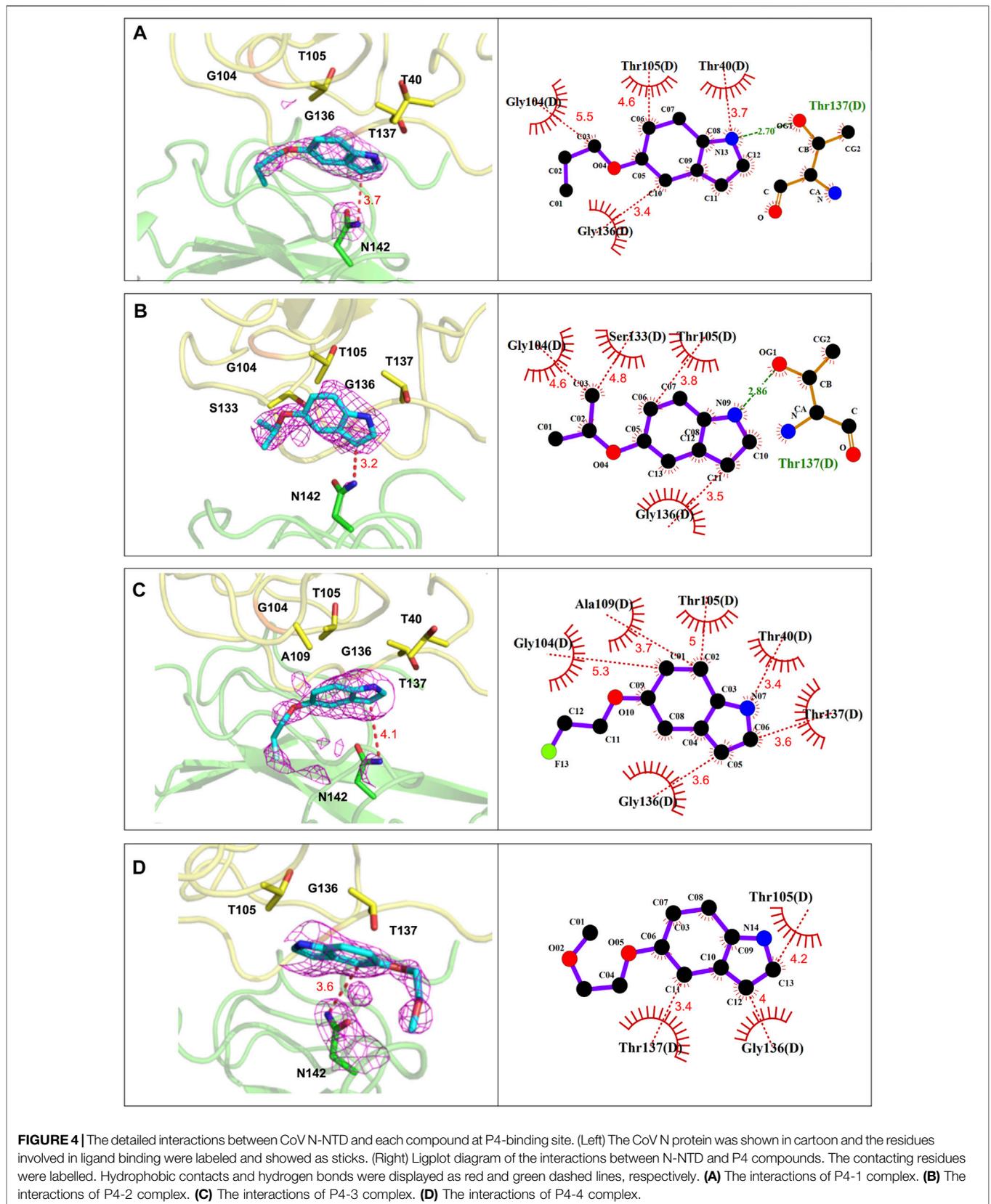


**Scheme 1** | Synthesis of P4-1 to P4-4. Reagent and conditions: **(A)** NaH, DMF, 25°C, 30 min; then 1-iodopropane, 25°C, 3 h; **(B)** 2-bromopropane, NaH, DMF, 60°C, 6 h then 2-bromopropane, NaH, 60°C, 3h; **(C)** 1-bromo-2-fluoroethane,  $\text{K}_2\text{CO}_3$ , acetone, reflux, 12 h; **(D)** 1-iodo-2-methoxyethane, NaH, DMF, 25°C, 3 h.



length MERS-CoV N protein and the quality of protein samples were pure in solution (**Supplementary Figure S1A**). We examined the solution conformation of full-length MERS-CoV N protein in the presence of each compound by SAXS (**Figure 3**). **Figure 3A** shows the fitted distance distribution function of the N protein with or without each compound. The addition of P4-1 and P4-2 resulted in an increase in the radius of gyration ( $R_g$ ) and maximum dimensions ( $D_{max}$ ) from 57 to 64 Å for P4-1 and from 199 Å to 220 Å for P4-2, respectively. In contrast, the addition of P4-3 and P4-4 did not change the value of  $R_g$  but caused a slight decrease in  $D_{max}$  from 199 to 191 (P4-3) and 199 to 185 Å (P4-4). The results revealed that the size of MERS-CoV N protein in solution increased upon binding to ligands with more hydrophobic R-groups. Next, we employed an ensemble optimization method (EOM) with SAXS data and structures of

the N-terminal domain (solved in this study) and C-terminal domain (CTD, PDB ID: 6G13) (Nguyen et al., 2019) to obtain the structural models for MERS-CoV N protein in complex with each ligand. The chosen ensemble fitted the data very well, and the representative structural models are shown in **Figures 3C–F**. In the presence of P4-1 and P4-2, the N proteins formed a dodecameric high-ordered structure, in which the CTD dimers aligned inward as a center ring. In each N protein dimer, the basic building block for CoV N protein, one NTD associated with the center ring formed by CTD, and the other NTD hung out of the ring in a monomer form (**Figures 3C,D**). However, the fitting conformation of the N protein with P4-3 or P4-4 was a tetramer, in which the two CTD dimers were placed at the center and the four NTD monomers were placed outside this center (**Figures 3E,F**). The conformation was similar to the previously solved



structure of MERS-CoV N protein, indicating that P4-3 and P4-4 compounds are unable to induce the formation of higher-order oligomerization of full-length N protein (**Supplementary Figure S2**) (Lin et al., 2020).

## 2.4 Crystal Structure of the MERS-CoV N-NTD Complexed With P4 Compounds

We applied X-ray crystallography to elucidate the detailed mechanism employed by P4 compounds, while providing insights into the structural features required for a P4 compound design. We purified MERS-CoV N-NTD and further confirmed these sample were homogenous in solution by size-exclusive chromatography (**Supplementary Figure S1B**). We determined the structures of MERS-CoV N-NTD in complex with each ligand by molecular replacement (MR) using the crystal structure of HCoV-OC43 N-NTD as the search model (Chen et al., 2013). The complex structures of N-NTD with P4 compounds were solved at a resolution of around 2.5 Å (**Supplementary Table S1**). As shown in **Supplementary Figure S2A**, the N-NTDs of all the complexes shared a similar structural core containing a five-stranded antiparallel  $\beta$ -sheet sandwiched between loops. The conformation was organized into a right-handed, fist-shaped structure conserved across various CoVs (Nations, 2020) (**Supplementary Figure S3B**). Additionally, the ligands were present alongside the protein (**Supplementary Figure S3A**). We then analyzed the detailed interaction between P4 compounds and N-NTD with LigPlot + suite (Laskowski and Swindells, 2011).

As shown in **Figure 4**, the indole ring and the side chain of P4-1 contributed to hydrophobic contacts with T40, T105, G136, T137, and G104 of N-NTD, respectively. In addition, one hydrogen bond was formed between the pyrrole ring of the indoline moiety and the T137 of N-NTD. The P4-2 complex shares a nearly identical interaction composition with P4-1, in which the indole ring and the side-chain contributed hydrophobic contacts by packing against T105, G136, G104, and S133 of N-NTD, respectively. The hydrogen bond between the pyrrole ring of the indoline moiety and the T137 of N-NTD was also detected in the P4-2 complex. Relative to P4-1 and P4-2, the interactions between P4-3 and P4-4 with N-NTD were mediated by only the hydrophobic contacts, contributed by T40, G104, T105, A109, G136, and T137 and T105, G136, and T137 of N-NTD for P4-3 and P4-4, respectively.

These structural analyses revealed that the side chains of P4 compounds did not strongly participate in protein interactions. Therefore, to understand why these four compounds have different capacities to induce the aggregation of whole N protein, we further examined their drug-binding position. We found that the P4-contacting site is a related hydrophobic and flat surface that accommodates P4 compounds and influences the position of the indole moiety based on the hydrophobicity of the R groups (**Supplementary Figure S4**). In accordance with this notion, the positions of the indole moiety of P4-1 and P4-2 almost overlapped, while the position of the indole moiety of P4-3 was far from that of P4-1 and P4-2. This resulted in the distances between the contacting residues of N-NTD and P4-1 (~3.98 Å)

and P4-2 (~3.91 Å) being moderately shorter than that of P4-3 (~4.1 Å). This also affected the distance between the pyrrole ring of each compound and the side chain of T137 of N-NTD, leading to an important hydrogen bond formed between N-NTD and P4-1, P4-2, but not P4-3 (**Supplementary Figure S4**). The strength of the hydrophobic contacts was largely determined by the interacting distance. In addition, the strength of hydrogen bonds is generally stronger than that of hydrophobic contacts (Ferreira de Freitas and Schapira, 2017), which implies that P4-1 and P4-2 bind to N-NTD more strongly than P4-3, which could explain why only P4-1 and P4-2 can elicit oligomerization. Accordingly, the side chain of P4-4, the most hydrophilic R group, was ejected by this hydrophobic environment and extended in the opposite direction to point to a hydrophilic cluster, which also largely changed the position of the P4-4 compound (**Supplementary Figure S4A**). Since our structural analysis showed that the P4-1 and P4-2 have more interactions with MERS-CoV N-NTD, we used the fluorescent quenching assay to confirm their binding affinities. The values of binding constant (Kd) for MERS-CoV-N-NTD to P4-1, P4-2 is found to be  $23.4 \pm 4.94 \mu\text{M}$  and  $13 \pm 4.24 \mu\text{M}$  respectively. Accordingly, titration of P4-4 to the protein showed no significant quenching suggesting weaker affinity (**Supplementary Figure S5**). These results suggest that the indole compounds could target a hydrophobic surface on N-NTD, in which the hydrophobicity of the R-group could affect the contact of indole compounds with the N protein, thus affecting the aggregation tendency of full-length N protein.

## 2.5 Ligand-Protein Surface Mediates the Interaction Between N-NTD and N-CTD

To uncover the mechanism of the allosteric modulation of full-length N protein by P4 compounds, we compared the differences in the surface properties of N-NTD upon binding of P4 compounds. We found that the 5-hydroxyindole moiety of P4 compound occupied one cavity on N-NTD to create a novel surface that we named as “ligand-protein surface” (**Supplementary Figure S4B**). We further found that, compared to those of P4-1 and P4-2, the side chains of P4-3 and P4-4 protruded out of the ligand-protein surface (**Supplementary Figure S6**). Based on the side-chain property of each compound, the results suggested that the newly created surfaces in the P4-1 and P4-2 complexes may be flatter and more hydrophobic than those of P4-3 and P4-4 (**Supplementary Figure S7**). The results of SAXS experiments indicated that one N-NTD was associated with the central aggregated N-CTD during ligand-induced modulating events. We thus speculated that the formation of the “ligand-protein surface” helps increase the interaction area of each P4-N-NTD complex to bind to N-CTD. To test this hypothesis, we examined the surface hydrophobicity of the centrally aggregated N-CTD in a higher-ordered N protein oligomer. As expected, in the solution model of full-length N protein with P4-1 and P4-2, the ligand-protein surface on N-NTD was close to the hydrophobic surface of the N-CTD ring (**Supplementary Figure S8**). Based on these observations, we

proposed that the binding of P4 affects the surface properties, including the flatness and hydrophobicity of N-NTD, which allows the association of N-NTD with N-CTD. This further influences the interaction behavior of N-CTD and eventually results in the higher ordered oligomerization of the N protein.

## 2.6 Application of P4 Compounds in Targeting the SARS-CoV-2 N-NTD

Since the N protein is highly conserved within various CoVs, we inferred that it could be possible to apply our strategy to other CoV inhibitions. Recently, SARS-CoV-2, a novel coronavirus closely related to SARS-CoV, has elicited a worldwide outbreak of pneumonia and poses a high risk to global public health (Hui et al., 2020; Zhu et al., 2020). Many structural studies have been performed to accelerate drug development against SARS-CoV-2 (Jin et al., 2020a; Jin et al., 2020b; Gao et al., 2020; Kang et al., 2020; Wang et al., 2020; Wu et al., 2020), including the crystal structure of SARS-CoV-2 N-NTD (Kang et al., 2020). To assess the possibility of P4 acting on SARS-CoV-2 N proteins and subsequently inhibiting viral replication, we compared the structure of the P4-1 complex with that of SARS-CoV-2 N-NTD within the P4 binding region. The amino acids of SARS-CoV-2, corresponding to the P4-1 interacting residues of MERS CoV N-NTD, have almost identical conformation compared to MERS CoV N-NTD (Figure 6A). The multiple sequence alignment of the N-NTD region of various  $\beta$ -CoVs also indicates that the residues involved in P4 binding are highly conserved (Figure 6B). This implies that P4 can induce higher-order oligomerization of the SARS-CoV-2 N protein, suggesting the possible development of indole derivatives in COVID-19 treatments.

## 3 DISCUSSION

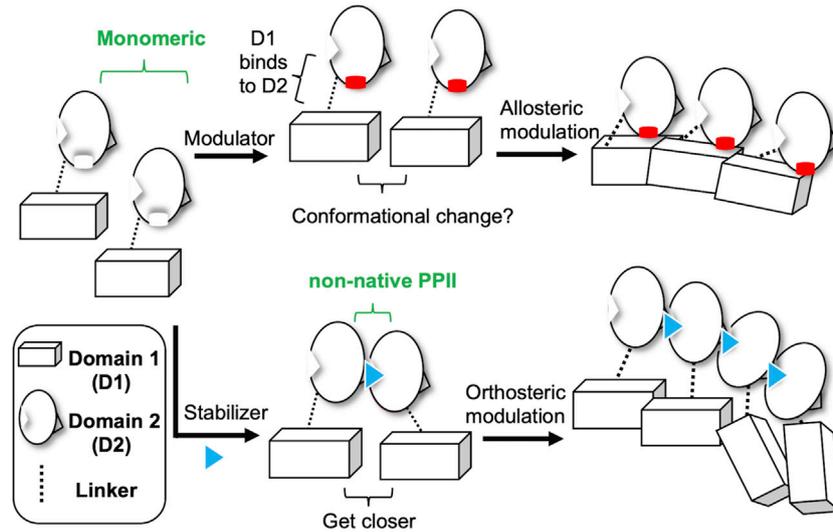
For the last two decades, there have been three successive outbreaks of new lethal CoVs such as SARS, MERS, and SARS-CoV-2. All these epidemics have significantly impacted human health and the economy. Interestingly, they are all characterized as the  $\beta$  subgroup of CoVs, and their RNA genomes are inherently inclined to have higher mutation rates. Therefore, the development of new anti-CoV agents is urgently needed to control the illnesses caused by these viruses.

The N protein is the most abundant structural protein of CoVs, which is responsible for the formation of the RNP complex to facilitate viral replication and translation (Almazan et al., 2004; Zuniga et al., 2010). The N protein is highly conserved within various CoVs and has a lower mutation rate. These properties make N protein a promising target for the development of an antiviral agents. Two strategies have been used to inhibit the N protein function (Cianci et al., 2012). One was to develop antiviral compounds targeting the RNA binding site of the N protein (Lin et al., 2014; Tarus et al., 2015) and the other was to block the normal N protein oligomerization (Gerritz et al., 2011) to prevent the formation of the RNP complex during viral replication. In our recent study, we identified manipulation of a non-native PPI of

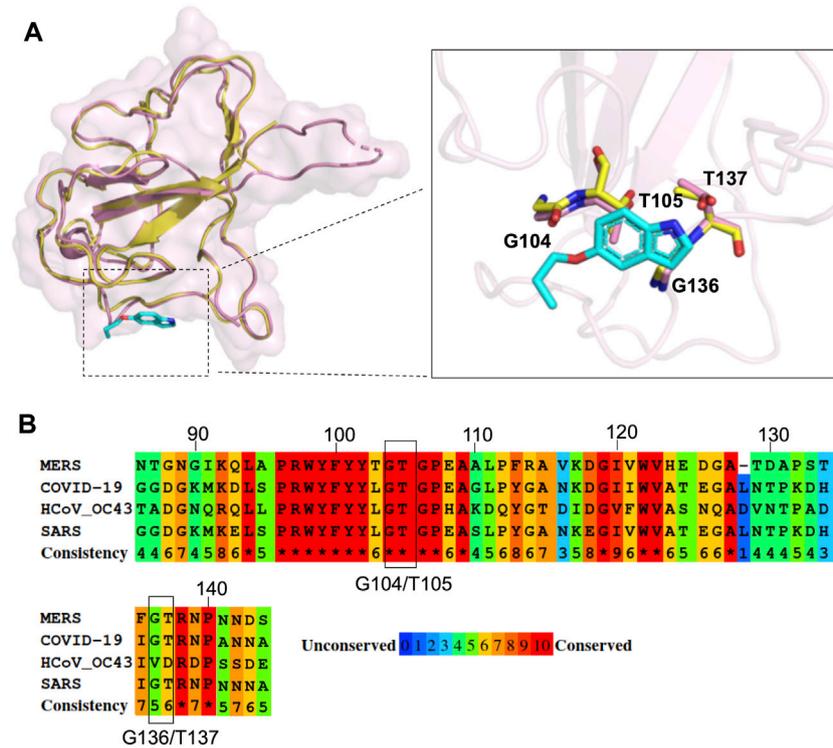
MERS-CoV N-NTD by a lead compound (P3), as being a novel strategy for designing antiviral drugs (WHO, 2020). These results established a model for a PPI stabilizer targeting N-NTD for antiviral drug development.

In this study, we designed P4 compounds to elicit high-order oligomerization of N protein via a mechanism different from that of P3. Among all types of PPI modulators, the allosteric modulator is defined as a group of ligands that bind to the sites on a protein spatially far from the functional sites and induce a conformational change that influences the binding of the protein to its interacting partner(s) (Petta et al., 2016; Duan et al., 2019). Our crystal structures of P4 complexes revealed that P4 acts on a surface remote from the non-native PPI of the N-NTD dimer. The thermal-stability experiments also implied that P4 did not stabilize this non-native dimer (Supplementary Figure S9). Furthermore, the solution conformations of full-length N protein in the presence or absence of P4 compounds showed a significant difference, in which P4 largely altered the interacting behavior of N-CTD, resulting in the abnormal oligomerization of the N protein. Moreover, according to the solved structures of allosteric modulators in complex with their target proteins, the allosteric modulators always contain a hydrophobic scaffold accompanied by certain polar interactions (Duan et al., 2019). For instance, the crystal structure of CD40L in complex with its allosteric inhibitor, BIO8898, revealed that the binding pocket of BIO8898 is 80% hydrophobic and 5% polar (Silvian et al., 2011). In addition, the complex structure of hCdc34 and its allosteric stabilizer, CC0651, revealed that the contact between CC0651 and hCdc34 was dominated by hydrophobic interactions, and only one hydrogen bond was formed in this complex (Ceccarelli et al., 2011). These examples imply that the interaction of allosteric modulators may require the majority of hydrophobic contacts with the help of certain critical hydrogen bonds to strengthen the ligand-protein interaction, thereby affecting the overall conformations of the target proteins. Our structural analyses revealed that the interactions between N-NTD and P4 compounds were mainly mediated by hydrophobic contacts. Furthermore, the formation of one critical hydrogen bond plays an important role in this regulation. However, the detailed mechanism of P4-induced aggregation of N proteins remains unproven. Based on the above-mentioned definitions and properties of P4 compounds, we inferred that P4 functions through allosteric modulation.

Allosteric ligands possess a unique mechanism; they bind to the region far from the orthosteric site and thus do not need to compete with their PPI partners. This can reduce the dosage of allosteric modulators and offers them a promising feature in PPI-based drug development. However, due to the limited knowledge of allosteric regulation in PPI manipulation, discovering allosteric PPI modulators is still highly challenging. Similar to P4 compounds, most allosteric PPI modulators were discovered serendipitously. One famous example of allosteric PPI modulators against viruses is the allosteric IN inhibitor (ALLINI) (Wilson et al., 2019), a major class of novel inhibitors of HIV-1 integrase (IN), initially designed to inhibit the interaction of IN and its cellular



**FIGURE 5** | Two different mechanisms of small-molecules targeting CoV N-NTD for inducing abnormal oligomerization. Allosteric modulator binds to the CoV N-NTD to create a hydrophobic flat surface that effect the oligomeric tendency of N-CTD, which induces the abnormal oligomerization of N protein. Whereas, N-NTD stabilizer acts on the non-native interface of CoV N-NTD dimer, which reduces the distance between targeted N proteins and eventually results in the aggregation of whole N protein.



**FIGURE 6** | Structural comparison of MERS CoV N-NTD with that of SARS-CoV-2 at P4 binding surface. **(A)** The structure of P4-1:MERS CoV N-NTD complex (yellow) was aligned to SARS-CoV-2 N-NTD (PDB: 6M3M), the interacting residues are highlighted in right box. **(B)** Protein sequence alignment of N-NTD of various CoVs. The conservation scoring was calculated by PRALINE and indicated as 0 (least conserved) to 10 (most conserved). The residues involved in P4-1 binding were indicated by black box.

substrate, but was later proven to promote IN multimerization, thus inhibiting IN activity. This suggests that allosteric modulators of protein oligomerization can be a strategically

used to develop antiviral agents, and the results from our study could serve as the basis for allosteric PPI modulator design against CoV N protein.

In the past decade, fragment-based drug discovery (FBDD) has become a powerful tool for the discovery of lead compounds against pathogens. X-ray crystallography also plays a profound role in both drug discovery and fragment advancement. Recently, some drugs have been discovered using the FBDD approach combined with X-ray crystallography; these drugs are currently under safety and effectiveness evaluation in clinical trials (Wyatt et al., 2008; Howard et al., 2009). Hence, FBDD may be used as a powerful tool to facilitate drug development. To consolidate this approach, more detailed correlations between ligands and targets via structural techniques need to be included to enhance their progress. We also investigated the chemical properties required for the therapeutic effects of P4 compounds targeting the CoV N-NTD. Two chemical properties are crucial for the development of allosteric PPI modulators targeting CoV N-NTD. First, the indoline moiety, which not only serves as a hydrophobic core for acting on the targeting surface, but also provides an N-H site for hydrogen bond formation between the ligand and N-NTD. Second, a hydrophobic branching moiety is required to fine-tune the final position of the indoline core to initiate an interaction between the ligand and N-NTD. This information can be further utilized for new drug development via the FBDD approach.

Finally, based on our recent studies and current results, we propose two different strategies of small molecules targeting CoV N-NTD to induce abnormal aggregation (Figure 5). One is to target the non-native interface of the N-NTD dimer, which functions as a stabilizer to drive the N protein to get closer, thus resulting in oligomerization of the N protein. The other is to act on one N-NTD, which serves as an allosteric modulator that affects the oligomerization behavior of N-CTD. Our results show that compounds P4-1 or P4-2 can induce the abnormal aggregation of the CoV N protein (Figure 3). We also found that P4-1 or P4-2 may act on other CoV N proteins, such as SARS-CoV-2, in addition to MERS-CoV, based on the results of structural comparison and sequence alignment (Figure 6).

Taken together, our current results provide the structural features of small molecules capable of eliciting high-order oligomerization of CoV N protein through an allosteric mechanism. Although it is not possible at this stage to demonstrate the ability of P4-1 and P4-2 to inhibit the potential propagation of coronaviruses, our study highlights an important strategy for manipulating the PPIs by targeting viral proteins. Finally, these findings will accelerate the development of novel anti-CoV agents based on the manipulation of PPIs of the CoV N proteins.

## 4 MATERIALS AND METHODS

### 4.1 Cloning, Protein Expression, and Purification

MERS-CoV N-NTD expression and purification were performed according to previously described methods (Wang et al., 2015). The cDNA fragments of MERS-CoV N39-165 protein were cloned into a pET-28a expression vector (Merck, Darmstadt, Germany) containing a histidine tag-encoding sequence. The

pET-28a/MERS-CoV N-NTD construct was transformed into *Escherichia coli* BL21 (DE3) pLysS cells and the cells were grown in Luria Bertani (LB) broth containing 50 µg/ml kanamycin at 37°C until the optical density (OD) reached 0.6–0.8 at 600 nm. However, the full-length MERS CoV N protein was cloned into a pET-21a expression vector (Merck, Darmstadt, Germany) containing a histidine tag-encoding sequence. The transformed cells were grown in LB broth containing 50 µg/ml ampicillin for 5 h at 200 rpm and 37°C. Protein expression was induced by adding 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), followed by incubation at 10°C at 200 rpm for 24 h. The bacteria were harvested via centrifugation (8,000 rpm, 12 min, 4°C), and the bacterial pellets were resuspended in the lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM sodium chloride (NaCl), 15 mM imidazole, 1 mM phenylmethylsulfonyl fluoride (PMSF)) and sonicated on ice. The soluble protein was then obtained from the supernatant after centrifugation (13,000 rpm, 40 min, 4°C). The MERS-CoV N-NTD protein was purified using a nickel-nitrilotriacetic acid (Ni-NTA) column (Merck, Darmstadt, Germany) with an elution gradient ranging from 15 to 250 mM of imidazole. Pure fractions were collected and dialyzed against a low-salt buffer and concentrated using a filter device with a 3 kDa cutoff membrane (Merck, Darmstadt, Germany). Protein concentrations were determined using the Bradford method (Wang et al., 2015).

### 4.2 Compound Design

To start the fragment search, the crystal structure of MERS-CoV N-NTD was used as a template to target the B-site of the non-native dimerization interface with a PPI fragment library and the LibDock software. According to the filtering criterion described in the main text, the chosen small molecule would serve as a starting fragment and further be applied to similarity searches using the ZINC database. The derivatives of the selected compound with the desired R-group modification were further synthesized as described below.

### 4.3 Synthesis of P4 Compounds

#### 4.3.1 5-(Propoxy)-1H-Indole (P4-1)

Sodium hydride (66 mg, 1.65 mmol) and 5-hydroxyindole (200 mg, 1.50 mmol) were dissolved in 2.4 ml of dimethylformamide (DMF) and stirred at 25°C for 30 min, followed by the addition of 1-iodopropane (0.16 ml, 1.65 mmol) and stirred at 25°C for 3 h. The reaction mixture was then extracted with EtOAc (15 × 3 ml) and combined organic layers, which were washed with brine, dried over magnesium sulfate (MgSO<sub>4</sub>), and concentrated. The residue was purified by column chromatography on silica gel to obtain compound **P4-1** (199 mg, 76%) as a brown oil.  $R_f$  0.29 (EtOAc/Hex = 1/10);  $[\alpha]_D^{25} +30.33$  (c 0.1, DCM); IR (NaCl)  $\nu$  3,443, 3,039, 2,860, 1,625 cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.05 (s, 1H), 7.27 (t, J = 8.0 Hz, 1H), 7.18 (s, 1H), 7.11 (d, J = 2.4 Hz, 1H), 6.89 (d, J = 2.4 Hz, 1H), 6.47 (m, 1H), 3.97 (t, J = 8.0 Hz, 2H), 1.83 (m, 2H), 1.06 (t, J = 8.0 Hz, 3H); <sup>13</sup>C NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  153.2, 130.9, 128.1, 124.9, 112.6, 111.7, 103.4, 101.8, 70.4, 22.6, 10.4; HRMS (ESI, M + Na<sup>+</sup>). The calculated MW for C<sub>11</sub>H<sub>13</sub>NONa was 198.0895, whereas the actual value was 198.0897. The <sup>1</sup>H and <sup>13</sup>C

nuclear magnetic resonance (NMR) spectral data of compound P4-1 were consistent with previously published literature (Unzue et al., 2016).

First, 5-hydroxyindole (500 mg, 3.76 mmol) was dissolved in 5 ml of DMF along with 2-bromopropane (0.42 ml, 4.51 mmol) and sodium hydride (300 mg, 7.51 mmol), and stirred at 60°C. After 6 h, 2-bromopropane (0.71 ml, 7.51 mmol) and sodium hydride (300 mg, 7.51 mmol) were added and it was stirred at 60°C for 3 h. The reaction mixture was extracted with EtOAc (20 × 3 ml), and the combined organic layers were washed with brine and dried over MgSO<sub>4</sub>. The residue was purified by column chromatography on silica gel to give the desired compound P4-2 (5-Isopropoxy-1H-indole) (334 mg, 51%) as a yellow oil. *R*<sub>f</sub> 0.44 (EtOAc/Hex = 1/4).  $[\alpha]_D^{25} +57.59$  (*c* 0.1, DCM); IR (NaCl)  $\nu$  3,443, 3,039, 2,860, 3,039, 2,860, 1,625 cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.05 (s, 1H), 7.28 (d, *J* = 8.8 Hz, 1H), 7.18 (t, *J* = 2.8 Hz, 1H), 7.14 (d, *J* = 2.4 Hz, 1H), 6.86 (dd, *J* = 4.4, 1.2 Hz, 1H), 6.47 (m, 1H), 4.52 (m, 1H), 1.36 (d, *J* = 4.0, 6H), 1.35 (s, 3H); <sup>13</sup>C NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  152.0, 131.2, 128.3, 124.9, 114.4, 111.6, 106.4, 102.2, 71.4, 22.2; HRMS (ESI, M + Na<sup>+</sup>). The calculated MW for C<sub>11</sub>H<sub>13</sub>NONa was 198.0895, whereas the actual MW was 198.0898. The <sup>1</sup>H and <sup>13</sup>C NMR spectral data of compound P4-2 were consistent with previously published literature (Taydakov et al., 2011).

#### 4.3.2 5-(2-Fluoroethoxy)-1H-indole (P4-3)

First, 2-bromoethanol (1.14 ml, 16.00 mmol) was added dropwise to diethylaminosulfur trifluoride (2.10 ml, 15.90 mmol) in diglyme (10 ml) at -50°C. The reaction mixture was heated to 25°C and stirred for 1 h. The most volatile portion was distilled at 70°C, and the residue was used in the next step of the reaction without further purification. Then, 5-hydroxyindole (200 mg, 1.50 mmol) in acetone (2 ml) was added to the solution of 1-bromo-2-fluoroethane (286 mg, 2.25 mmol) and K<sub>2</sub>CO<sub>3</sub> (415 mg, 3.00 mmol) in DMF. The reaction mixture was heated under reflux and stirred for 6 h and then extracted with EtOAc (15 × 3 ml). The combined organic layers were washed with brine and dried over MgSO<sub>4</sub>. The residue was purified by column chromatography on silica gel to give the desired compound P4-3 (66 mg, 24% in 2 steps) as a yellow oil. *R*<sub>f</sub> 0.35 (EtOAc/Hex = 1/4).  $[\alpha]_D^{25} +22.27$  (*c* 0.1, DCM); IR (NaCl)  $\nu$  3,443, 3,039, 2,860, 1,625 cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.08 (s, 1H), 7.30 (d, *J* = 8.8 Hz, 1H), 7.20 (t, *J* = 2.8 Hz, 1H), 7.13 (d, *J* = 2.4 Hz, 1H), 6.91 (dd, *J* = 4.4, 1.2 Hz, 1H), 6.49 (m, 1H), 4.84 (t, *J* = 4.4 Hz, 1H), 4.72 (t, *J* = 4.4 Hz, 1H), 4.30 (t, *J* = 4.0 Hz, 1H), 4.23 (t, *J* = 4.0 Hz, 1H); <sup>13</sup>C NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  124.9, 113.0, 111.8, 103.9, 102.4, 83.1, 81.4, 68.2; HRMS (ESI, M + Na<sup>+</sup>). The calculated MW for C<sub>10</sub>H<sub>10</sub>FNONa was 202.0644, whereas the actual MW was 202.0690. The <sup>1</sup>H and <sup>13</sup>C NMR spectral data for compound P4-3 were consistent with previously published literature (Wang et al., 2017).

#### 4.3.3 5-(2-Methoxy-ethoxy)-1H-indole (P4-4)

To a solution of sodium iodide (2.71 g, 18.1 mmol) in acetone (15.6 ml), 2-chloroethyl methyl ether (0.78 ml, 8.60 mmol) was added. The mixture was then heated in reflux and stirred. After 16 h, the mixture was filtered and concentrated. The residue was

used in the next reaction step, without further purification. Sodium hydride (66 mg, 1.65 mmol) was dissolved in 1.2 ml of DMF, and the mixture was slowly added with 5-hydroxyindole (200 mg, 1.50 mmol) in DMF (1.2 ml) and stirred at 25°C. After 30 min, 2-methoxyethyl iodide (363 mg) was added to the mixture and stirred continuously at 25°C for 3 h. The resulting mixture was washed with brine and the aqueous phase was extracted with EtOAc (5 × 50 ml). The combined organic layers were dried over MgSO<sub>4</sub>, filtered, and concentrated under reduced pressure. The residue was purified by column chromatography on silica gel to give the desired compound P4-4 (117 mg, 41% in 2 steps) as a brown oil. *R*<sub>f</sub> 0.60 (EtOAc/Hex = 1/1).  $[\alpha]_D^{25} +21.30$  (*c* 0.1, DCM); IR (NaCl)  $\nu$  3,443, 3,039, 2,860, 1,625 cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.07 (s, 1H), 7.28 (d, *J* = 8.0 Hz, 1H), 7.19 (d, *J* = 2.4 Hz, 1H), 7.13 (d, *J* = 2.4 Hz, 1H), 6.93 (m, 1H), 6.48 (m, 1H), 4.17 (m, 2H), 3.77 (m, 2H), 3.29 (s, 3H); <sup>13</sup>C NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  153.1, 131.1, 128.1, 124.9, 112.9, 111.7, 103.6, 102.2, 71.2, 68.0, 59.0; HRMS (ESI, M + Na<sup>+</sup>). The calculated MW for C<sub>11</sub>H<sub>13</sub>NO<sub>2</sub>Na was 214.0844, whereas the actual MW was 214.0841.

## 4.4 Crystallization and Data Collection

Each MERS-CoV N-NTD:P4 complex was crystallized using the sitting-drop vapor diffusion method. The crystals of MERS-CoV N-NTD:P4 complex were obtained from a solution of 5 mg/ml MERS-CoV N-NTD, 25 mM Tris-HCl (pH 7.5), 75 mM NaCl, 140 mM 2-(N-morpholino) ethanesulfonic acid (MES) (pH 5.5), 75 mM ammonium sulfate ((NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>), 29% polyethylene glycol (PEG) 3,350, 2 mM sodium bromide (NaBr), and 2 mM P4 ligand equilibrated against 300  $\mu$ L of reservoir solution consisting of 280 mM MES (pH 5.5), 150 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 58% PEG 3350, and 4 mM NaBr at 20°C. The crystals were harvested after 3 weeks. Diffraction datasets for the MERS-CoV N-NTD in complex with P4-1~P4-3 were collected at beamline 05A1 of the Taiwan Photon Source (TPS) of the National Synchrotron Research Center (NSRRC; Hsinchu City, Taiwan). Diffraction of the MERS-CoV N-NTD:P4-4 complex was performed at the beamline 15A1 of Taiwan Light Source (TLS). All the crystals were mounted directly on the loops without use of cryoprotectant.

## 4.5 Structural Determination and Refinement

All diffraction data were processed and scaled using the HKL2000 software. The phases were solved by molecular replacement (MR) in PHENIX (version 1.10.1) (Adams et al., 2002) using HCoV-OC43 N-NTD (PDB ID: 4J3K) as the search model. According to the electron density map, the structures were adjusted and refined using WinCoot (version 0.8.4) (Emsley et al., 2010) and PHENIX until the R-factor and R-free values decreased. The crystallographic and refinement statistics of the MERS-CoV N-NTD/P4 ligands complexes are listed in Table 1. The refined structures were visualized using PyMOL (version 1.8) (Kudlicki et al., 2007). The hydrophobic interactions between MERS-CoV N-NTD and P4 ligands were analyzed using LigPlot<sup>+</sup> (Laskowski and Swindells, 2011). The distances of hydrophobic contacts between N-NTD and each compound were defined as

the length between the nearest atom of the contacting residues and each compound, which were calculated using PyMOL.

## 4.6 Small-Angle X-Ray Scattering Experiments

SAXS data were collected using a monochromatic X-ray beam ( $\lambda = 0.828 \text{ \AA}$ ) at the BL23A beamline coupled with a high-performance liquid chromatography (HPLC) system fitted with an Agilent-Bio SEC-3 300  $\text{\AA}$  column (Agilent Technologies, Santa Clara, United States) at the NSRRC (Hsinchu City, Taiwan). Protein samples (44  $\mu\text{M}$  of MERS-CoV N) were incubated with ligands (440  $\mu\text{M}$ ) in a buffer containing 50 mM Tris-HCl (pH 8.5) and 150 mM NaCl for 1 h on ice. Then, a 300  $\mu\text{L}$  aliquot was injected into the column at a flow rate of 0.02 ml/min and directed into a quartz capillary (2 mm diameter). For SAXS measurement, with X-rays and a sample-to-detector distance of 2.5 m, the scattering vector  $q$ , defined as  $q = (4\pi/\lambda) \sin\theta$ , with the scattering angle  $2\theta$  during a 30 s exposure, 36 frames of all samples were collected and merged, and analyzed to determine the initial  $R_g$  in PRIMUS (Version 3.1). The  $D_{\max}$  and  $P(r)$  distance distributions were calculated from the experimental scattering curve using GNOM (version 4.1). EOM was used with the EMBL Hamburg web interface (Petoukhov et al., 2012). The crystal structures of the CTD domain of MERS-CoV N protein (PDB ID: 6G13) (Nguyen et al., 2019), MERS-CoV N-NTD (PDB ID: 4UD1) (Papageorgiou et al., 2016) and MERS-CoV N-NTD/P4 ligands (solved in this study) were used as rigid bodies in EOM analysis (Kelley et al., 2015). For EOM analysis, numerous models were generated at the beginning (structural pool). The EOM, fitting the experimental scattering curve with a linear combination, was selected from the structural pool. The presented conformations of P4 complexes were chosen because their ensemble-generated curves fit best to the experimental results of SAXS.

## 4.7 Protein Sequence Alignment and Structure Prediction

Protein sequence alignment of N-NTDs of various CoVs was performed using PRALINE. Sequences of all CoV N proteins have been deposited in Genbank with ACCESSION numbers of ASU90457.1, QHO62115.1, QBP84763.1 and AYW99827.1 for MERS-CoV, SARS-CoV-2, HCoV-OC43, and SARS-CoV, respectively. For the dimeric SARS-CoV-2 N-NTD, the SARS-CoV-2 N-NTD (PDB ID: 6M3M) structures were superimposed onto each monomer of the MERS-CoV N-NTD dimer, and the PDB file was exported using the PyMOL.

## REFERENCES

- Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., et al. (2002). PHENIX: Building New Software for Automated Crystallographic Structure Determination. *Acta Crystallogr. D Biol. Cryst.* 58 (Pt 11), 1948–1954. doi:10.1107/s0907444902016657
- Almazán, F., Galán, C., and Enjuanes, L. (2004). The Nucleoprotein Is Required for Efficient Coronavirus Genome Replication. *J. Virol.* 78 (22), 12683–12688. doi:10.1128/JVI.78.22.12683-12688.2004

## 4.8 Accession Numbers

PDB ID: 6LNN (N:P4-1 complex), 7DYD (N:P4-2 complex), 6LZ6 (N:P4-3 complex) and 6LZ8 (N:P4-4 complex).

SASBD ID: SASDNF6 (N:P4-1 complex), SASDNG6 (N:P4-2 complex), SASDNH6 (N:P4-3 complex), SASDNI6 (N:P4-4 complex).

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## AUTHOR CONTRIBUTIONS

J-NH: Methodology, Validation, Investigation, Data curation. S-ML: Methodology, Validation, Investigation, Writing—Original Draft. Y-JC, J-SC, and S-YL: Methodology—Compound design and Synthesis. U-SJ: Methodology. J-YH: Writing—review and editing. M-HH: Conceptualization, Methodology, Validation, Writing—review and editing, Resources, Supervision, Project administration, Funding acquisition.

## FUNDING

This work was supported by Ministry of Science and Technology (MOST) (MOST 109-2311-B-005-007-MY3, MOST 109-2628-M-005-001-MY4, MOST 109-2327-B-005-005) and ENgineering in Agriculture Biotech Leadership (ENABLE) (ENABLE Center 109ST001C).

## ACKNOWLEDGMENTS

We would like to thank the staff of the National Synchrotron Radiation Research Center (Taiwan) for data collection.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.871499/full#supplementary-material>

- Andradebatista, I. A. S. D., Gonçalves, M. I. D. A., Singh, A. K., Hackmann, E. R. K., and Santoro, A. I. R. (2008). Quantitative Determination of Dimethylaminoethanol in Cosmetic Formulations by Nuclear Magnetic Resonance Spectroscopy. *J. AOAC Int.* 91 (6), 1303–1308. doi:10.1093/jaoac/91.6.1303
- Ceccarelli, D. F., Tang, X., Pelletier, B., Orlicky, S., Xie, W., Plantevin, V., et al. (2011). An Allosteric Inhibitor of the Human Cdc34 Ubiquitin-Conjugating Enzyme. *Cell* 145 (7), 1075–1087. doi:10.1016/j.cell.2011.05.039
- Chang, C.-k., Hou, M.-H., Chang, C.-F., Hsiao, C.-D., and Huang, T.-h. (2014). The SARS Coronavirus Nucleocapsid Protein - Forms and Functions. *Antiviral Res.* 103, 39–50. doi:10.1016/j.antiviral.2013.12.009

- Chang, C.-k., Sue, S.-C., Yu, T.-h., Hsieh, C.-M., Tsai, C.-K., Chiang, Y.-C., et al. (2005). The Dimer Interface of the SARS Coronavirus Nucleocapsid Protein Adapts a Porcine Respiratory and Reproductive Syndrome Virus-like Structure. *FEBS Lett.* 579 (25), 5663–5668. doi:10.1016/j.febslet.2005.09.038
- Chen, I.-J., Yuann, J.-M. P., Chang, Y.-M., Lin, S.-Y., Zhao, J., Perlman, S., et al. (2013). Crystal Structure-Based Exploration of the Important Role of Arg106 in the RNA-Binding Domain of Human Coronavirus OC43 Nucleocapsid Protein. *Biochim. Biophys. Acta (Bba) - Proteins Proteomics* 1834 (6), 1054–1062. doi:10.1016/j.bbapap.2013.03.003
- Cianci, C., Gerritz, S. W., Deminie, C., and Krystal, M. (2013). Influenza Nucleoprotein: Promising Target for Antiviral Chemotherapy. *Antivir. Chem. Chemother.* 23 (3), 77–91. doi:10.3851/IMP2235
- Du, Z., Zhou, Q., Sun, W., Yang, C., Wu, C., Wang, L., et al. (2020a). 5-Hydroxyindole-Based EZH2 Inhibitors Assembled via TCCA-Catalyzed Condensation and Nenitzescu Reactions. *Molecules* 25 (9), 2059. doi:10.3390/molecules25092059
- Du, Z., Wang, L., Cauchemez, S., Xu, X., Wang, X., Cowling, B. J., et al. (2020b). Risk for Transportation of Coronavirus Disease from Wuhan to Other Cities in China. *Emerg. Infect. Dis.* 26 (5), 1049–1052. doi:10.3201/eid2605.200146
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and Development of Coot. *Acta Crystallogr. D Biol. Cryst.* 66 (Pt 4), 486–501. doi:10.1107/S0907444910007493
- Ferreira de Freitas, R., and Schapira, M. (2017). A Systematic Analysis of Atomic Protein-Ligand Interactions in the PDB. *Med. Chem. Commun.* 8 (10), 1970–1981. doi:10.1039/c7md00381a
- Fischer, G., Rossmann, M., and Hyvönen, M. (2015). Alternative Modulation of Protein-Protein Interactions by Small Molecules. *Curr. Opin. Biotechnol.* 35, 78–85. doi:10.1016/j.copbio.2015.04.006
- Gao, Y., Yan, L., Huang, Y., Liu, F., Zhao, Y., Cao, L., et al. (2020). Structure of the RNA-dependent RNA Polymerase from COVID-19 Virus. *Science* 368 (6492), 779–782. doi:10.1126/science.abb7498
- Gerritz, S. W., Cianci, C., Kim, S., Pearce, B. C., Deminie, C., Discotto, L., et al. (2011). Inhibition of Influenza Virus Replication via Small Molecules that Induce the Formation of Higher-Order Nucleoprotein Oligomers. *Proc. Natl. Acad. Sci. U.S.A.* 108 (37), 15366–15371. doi:10.1073/pnas.1107906108
- Ghosh, A. K., Brindisi, M., Shahabi, D., Chapman, M. E., and Mesecar, A. D. (2020). Drug Development and Medicinal Chemistry Efforts toward SARS-Coronavirus and Covid-19 Therapeutics. *ChemMedChem* 15 (11), 907–932. doi:10.1002/cmdc.202000223
- Herod, M. R., Adeyemi, O. O., Ward, J., Bentley, K., Harris, M., Stonehouse, N. J., et al. (2019). The Broad-Spectrum Antiviral Drug Arbidol Inhibits Foot-And-Mouth Disease Virus Genome Replication. *J. Gen. Virol.* 100 (9), 1293–1302. doi:10.1099/jgv.0.001283
- Howard, S., Berdini, V., Boulstridge, J. A., Carr, M. G., Cross, D. M., Curry, J., et al. (2009). Fragment-based Discovery of the Pyrazol-4-Yl Urea (AT9283), a Multitargeted Kinase Inhibitor with Potent aurora Kinase Activity. *J. Med. Chem.* 52 (2), 379–388. doi:10.1021/jm800984v
- Hu, Y., Sneyd, H., Dekant, R., and Wang, J. (2017). Influenza A Virus Nucleoprotein: A Highly Conserved Multi-Functional Viral Protein as a Hot Antiviral Drug Target. *Ctmc* 17 (20), 2271–2285. doi:10.2174/1568026617666170224122508
- Hui, D. S., I Azhar, E., Madani, T. A., Ntoumi, F., Kock, R., Dar, O., et al. (2020). The Continuing 2019-nCoV Epidemic Threat of Novel Coronaviruses to Global Health - the Latest 2019 Novel Coronavirus Outbreak in Wuhan, china. *Int. J. Infect. Dis.* 91, 264–266. doi:10.1016/j.ijid.2020.01.009
- Huschmann, F. U., Linnik, J., Sparta, K., Ühlein, M., Wang, X., Metz, A., et al. (2016). Structures of Endothiapepsin-Fragment Complexes from Crystallographic Fragment Screening Using a Novel, Diverse and Affordable 96-compound Fragment Library. *Acta Cryst. Sect F* 72 (Pt 5), 346–355. doi:10.1107/s2053230x16004623
- Ishikura, M., Abe, T., Choshi, T., and Hibino, S. (2015). Simple Indole Alkaloids and Those with a Nonrearranged Monoterpenoid Unit. *Nat. Prod. Rep.* 32 (10), 1389–1471. doi:10.1039/c5np00032g
- Jayaram, H., Fan, H., Bowman, B. R., Ooi, A., Jayaram, J., Collisson, E. W., et al. (2006). X-ray Structures of the N- and C-Terminal Domains of a Coronavirus Nucleocapsid Protein: Implications for Nucleocapsid Formation. *J. Virol.* 80 (13), 6612–6620. doi:10.1128/JVI.00157-06
- Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., et al. (2020a). Structure of Mpro from SARS-CoV-2 and Discovery of its Inhibitors. *Nature* 582, 289–293. doi:10.1038/s41586-020-2223-y
- Jin, Z., Zhao, Y., Sun, Y., Zhang, B., Wang, H., Wu, Y., et al. (2020b). Structural Basis for the Inhibition of SARS-CoV-2 Main Protease by Antineoplastic Drug Carmofur. *Nat. Struct. Mol. Biol.* 27, 529–532. doi:10.1038/s41594-020-0440-6
- Kang, S., Yang, M., Hong, Z., Zhang, L., Huang, Z., Chen, X., et al. (2020). Crystal Structure of SARS-CoV-2 Nucleocapsid Protein RNA Binding Domain Reveals Potential Unique Drug Targeting Sites. *bioRxiv* 2003, 977876. doi:10.1101/2020.03.06.977876
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The PyMol Web portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* 10 (6), 845–858. doi:10.1038/nprot.2015.053
- Khodadadi, E., Maroufi, P., Khodadadi, E., Esposito, I., Ganbarov, K., Esposito, S., et al. (2020). Study of Combining Virtual Screening and Antiviral Treatments of the Sars-CoV-2 (Covid-19). *Microb. pathogenesis* 146, 104241. doi:10.1016/j.micpath.2020.104241
- Kochanowska-Karamyan, A. J., and Hamann, M. T. (2010). Marine Indole Alkaloids: Potential New Drug Leads for the Control of Depression and Anxiety. *Chem. Rev.* 110 (8), 4489–4497. doi:10.1021/cr900211p
- Kudlicki, A., Rowicka, M., and Otwinowski, Z. (2007). The Crystallographic Fast Fourier Transform. Recursive Symmetry Reduction. *Acta Cryst. Sect A* 63 (Pt 6), 465–480. doi:10.1107/S0108767307047411
- Laskowski, R. A., and Swindells, M. B. (2011). LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery. *J. Chem. Inf. Model.* 51 (10), 2778–2786. doi:10.1021/ci200227u
- Lin, S.-M., Lin, S.-C., Hsu, J.-N., Chang, C.-k., Chien, C.-M., Wang, Y.-S., et al. (2020). Structure-Based Stabilization of Non-native Protein-Protein Interactions of Coronavirus Nucleocapsid Proteins in Antiviral Drug Design. *J. Med. Chem.* 63 (6), 3131–3141. doi:10.1021/acs.jmedchem.9b01913
- Lin, S.-Y., Liu, C.-L., Chang, Y.-M., Zhao, J., Perlman, S., and Hou, M.-H. (2014). Structural Basis for the Identification of the N-Terminal Domain of Coronavirus Nucleocapsid Protein as an Antiviral Target. *J. Med. Chem.* 57 (6), 2247–2257. doi:10.1021/jm500089r
- McBride, R., van Zyl, M., and Fielding, B. (2014). The Coronavirus Nucleocapsid Is a Multifunctional Protein. *Viruses* 6 (8), 2991–3018. doi:10.3390/v6082991
- Modell, A. E., Blosser, S. L., and Arora, P. S. (2016). Systematic Targeting of Protein-Protein Interactions. *Trends Pharmacol. Sci.* 37 (8), 702–713. doi:10.1016/j.tips.2016.05.008
- Mori, M., Kovalenko, L., Lyonnais, S., Antaki, D., Torbett, B. E., Botta, M., et al. (2015). Nucleocapsid Protein: A Desirable Target for Future Therapies against HIV-1. *Curr. Top. Microbiol. Immunol.* 389, 53–92. doi:10.1007/82\_2015\_433
- Murray, C. W., and Rees, D. C. (2009). The Rise of Fragment-Based Drug Discovery. *Nat. Chem* 1 (3), 187–192. doi:10.1038/nchem.217
- Nguyen, T. H. V., Lichère, J., Canard, B., Papageorgiou, N., Attoumani, S., Ferron, F., et al. (2019). Structure and Oligomerization State of the C-Terminal Region of the Middle East Respiratory Syndrome Coronavirus Nucleoprotein. *Acta Cryst. Sect D Struct. Biol.* 75 (Pt 1), 8–15. doi:10.1107/S2059798318014948
- Ni, D., Lu, S., and Zhang, J. (2019). Emerging Roles of Allosteric Modulators in the Regulation of Protein-protein Interactions (PPIs): A New Paradigm for PPI Drug Discovery. *Med. Res. Rev.* 39, 2314–2342. doi:10.1002/med.21585
- Panjikovich, A., and Daura, X. (2014). PARS: a Web Server for the Prediction of Protein Allosteric and Regulatory Sites. *Bioinformatics* 30 (9), 1314–1315. doi:10.1093/bioinformatics/btu002
- Papageorgiou, N., Lichère, J., Baklouti, A., Ferron, F., Sévajol, M., Canard, B., et al. (2016). Structural Characterization of the N-Terminal Part of the MERS-CoV Nucleocapsid by X-ray Diffraction and Small-Angle X-ray Scattering. *Acta Cryst. Sect D Struct. Biol.* 72 (Pt 2), 192–202. doi:10.1107/S2059798315024328
- Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., et al. (2012). New Developments in the ATSAS program Package for Small-Angle Scattering Data Analysis. *J. Appl. Cryst.* 45 (Pt 2), 342–350. doi:10.1107/S0021889812007662
- Petta, I., Lievens, S., Libert, C., Tavernier, J., and De Bosscher, K. (2016). Modulation of Protein-Protein Interactions for the Development of Novel Therapeutics. *Mol. Ther.* 24 (4), 707–718. doi:10.1038/mt.2015.214

- Silvian, L. F., Friedman, J. E., Strauch, K., Cachero, T. G., Day, E. S., Qian, F., et al. (2011). Small Molecule Inhibition of the TNF Family Cytokine CD40 Ligand through a Subunit Fracture Mechanism. *ACS Chem. Biol.* 6 (6), 636–647. doi:10.1021/cb2000346
- Tarus, B., Bertrand, H., Zedda, G., Di Primo, C., Quideau, S., and Slama-Schwok, A. (2015). Structure-based Design of Novel Naproxen Derivatives Targeting Monomeric Nucleoprotein of Influenza A Virus. *J. Biomol. Struct. Dyn.* 33 (9), 1899–1912. doi:10.1080/07391102.2014.979230
- Taydakov, I. V., Dutova, T. Y., Sidorenko, E. N., and Krasnoselsky, S. S. (2011). Convenient Modification of the Leimgruber-Batcho Indole Synthesis: Reduction of 2-nitro- $\beta$ -pyrrolidinostyrenes by the FeCl<sub>3</sub>-Activated Carbon-N<sub>2</sub>h<sub>4</sub>-h<sub>2</sub>o System. *Chem. Heterocycl Comp.* 47, 425–434. doi:10.1007/s10593-011-0776-2
- Thiel, P., Kaiser, M., and Ottmann, C. (2012). Small-molecule Stabilization of Protein-Protein Interactions: an Underestimated Concept in Drug Discovery? *Angew. Chem. Int. Ed.* 51 (9), 2012–2018. doi:10.1002/anie.201107616
- Unzue, A., Zhao, H., Lolli, G., Dong, J., Zhu, J., Zechner, M., et al. (2016). The “Gatekeeper” Residue Influences the Mode of Binding of Acetyl Indoles to Bromodomains. *J. Med. Chem.* 59 (7), 3087–3097. doi:10.1021/acs.jmedchem.5b01757
- Vigerelli, H., Sciani, J. M., Pereira, P. M. C., Lavezo, A. A., Silva, A. C. R., Collaço, R. C. O., et al. (2020). Bufotenine, a Tryptophan-Derived Alkaloid, Suppresses the Symptoms and Increases the Survival Rate of Rabies-Infected Mice: the Development of a Pharmacological Approach for Rabies Treatment. *J. Venom. Anim. Toxins Incl. Trop. Dis.* 26, e20190050. doi:10.1590/1678-9199-jvatitd-2019-0050
- Wang, L., Ye, J., He, Y., Deuther-Conrad, W., Zhang, J., Zhang, X., et al. (2017). 18 F-Labeled Indole-Based Analogs as Highly Selective Radioligands for Imaging Sigma-2 Receptors in the brain F-Labeled Indole-Based Analogs as Highly Selective Radioligands for Imaging Sigma-2 Receptors in the Brain. *Bioorg. Med. Chem.* 25 (14), 3792–3802. doi:10.1016/j.bmc.2017.05.019
- Wang, Q., Wu, J., Wang, H., Gao, Y., Liu, Q., Mu, A., et al. (2020). Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. *Cell* 182, 417–428. doi:10.1016/j.cell.2020.05.034
- Wang, Y.-S., Chang, C.-k., and Hou, M.-H. (2015). Crystallographic Analysis of the N-Terminal Domain of Middle East Respiratory Syndrome Coronavirus nucleocapsid Protein. *Acta Cryst. Sect F* 71 (Pt 8), 977–980. doi:10.1107/S2053230X15010146
- WHO (2020). World Health Organization Situation Report.
- Wilson, T. A., Koneru, P. C., Rebersburg, S. V., Lindenberger, J. J., Kobe, M. J., Cockroft, N. T., et al. (2019). An Isoquinoline Scaffold as a Novel Class of Allosteric HIV-1 Integrase Inhibitors. *ACS Med. Chem. Lett.* 10 (2), 215–220. doi:10.1021/acsmchemlett.8b00633
- Wootton, S. K., Rowland, R. R. R., and Yoo, D. (2002). Phosphorylation of the Porcine Reproductive and Respiratory Syndrome Virus Nucleocapsid Protein. *J. Virol.* 76 (20), 10569–10576. doi:10.1128/jvi.76.20.10569-10576.2002
- Wu, Y., Wang, F., Shen, C., Peng, W., Li, D., Zhao, C., et al. (2020). A Noncompeting Pair of Human Neutralizing Antibodies Block COVID-19 Virus Binding to its Receptor ACE2. *Science* 368, 1274–1278. doi:10.1126/science.abc2241
- Wyatt, P. G., Woodhead, A. J., Berdini, V., Boulstridge, J. A., Carr, M. G., Cross, D. M., et al. (2008). Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (AT7519), a Novel Cyclin Dependent Kinase Inhibitor Using Fragment-Based X-ray Crystallography and Structure Based Drug Design. *J. Med. Chem.* 51 (16), 4986–4999. doi:10.1021/jm800382h
- Xu, X., Chen, P., Wang, J., Feng, J., Zhou, H., Li, X., et al. (2020). Evolution of the Novel Coronavirus from the Ongoing Wuhan Outbreak and Modeling of its Spike Protein for Risk of Human Transmission. *Sci. China Life Sci.* 63, 457–460. doi:10.1007/s11427-020-1637-5
- United Nations (2020). *COVID-19 to Slash Global Economic Output by \$8.5 Trillion over Next Two Years*. Available at: <https://www.un.org/en/desa/covid-19-slash-global-economic-output-85-trillion-over-next-two-years>.
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., et al. (2020). A Novel Coronavirus from Patients with Pneumonia in china, 2019. *N. Engl. J. Med.* 382, 727–733. doi:10.1056/NEJMoa2001017
- Zúñiga, S., Cruz, J. L. G., Sola, I., Mateos-Gómez, P. A., Palacio, L., and Enjuanes, L. (2010). Coronavirus Nucleocapsid Protein Facilitates Template Switching and Is Required for Efficient Transcription. *J. Virol.* 84 (4), 2169–2175. doi:10.1128/JVI.02011-09

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Hsu, Chen, Lin, Hong, Chen, Jeng, Luo and Hou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.