UNIVERSITÀ DEGLI STUDI FIRENZE

HR EXCELLENCE IN RESEARCH

DOTTORATO DI RICERCA
INTERNATIONAL DOCTORATE IN STRUCTURAL BIOLOGY

CYCLE XXXVIII

COORDINATOR Prof. Roberta Pierattelli

From Inhibition to Protein-Targeted Degradation: Structural Insights into SARS-CoV-2 and CVB3 Proteins for Antiviral Drug Development

Settore Scientifico Disciplinare CHIM/03

Doctoral Candidate
Dott. Alessia De Santis

Supervisor
Prof. Francesca Cantini

Coordinator
Prof. Roberta Pierattelli

November 2022 – October 2025

This thesis has been approved by the University of Florence, the University of Frankfurt and the Utrecht University

Universiteit Utrecht

Abstract

The COVID-19 pandemic has highlighted the urgent need for innovative antiviral strategies capable of addressing both existing and emerging viral threats. This PhD research focused on two complementary approaches for the development of antiviral therapeutics: the inhibition of viral proteases through small molecules and the exploitation of the targeted protein degradation (TPD) strategy. In particular, the work centred on the 3-Chymotrypsin-like Protease ($3CL^{Pro}$) of SARS-CoV-2, the 3C Protease ($3C^{Pro}$) of Coxsackievirus B3 (CVB3), and the Nsp13 helicase of SARS-CoV-2, enzymes that play essential roles in viral replication and represent validated drug targets.

The first part of this work investigated the inhibitory potential of gold-based metallodrugs such as Auranofin and its derivatives against SARS-CoV-2 $3CL^{Pro}$. Kinetic assays yielded Ki values in the low micromolar range (0.4–2.1 μM), while ESI-MS and X-ray crystallography confirmed adduct formation and identified gold coordination sites.

The second part explored the PROTAC (PROteolysis TArgeting Chimera) strategy as an innovative means to induce the degradation of viral proteases. Two PROTAC classes were designed and characterized, starting from respectively the Nirmatrelvir and Ibuzatrelvir nitrile inhibitors and the peptidomimetic inhibitor GC-376. Solution NMR spectroscopy, fluorescence binding assays, and analytical size-exclusion chromatography (SEC) were employed to monitor binary and ternary complex formation involving the viral protease, the PROTAC molecule, and the Celebron (CRBN) E3 ubiquitin ligase (CRBN-TBD and CRBN-midi). Evidence of productive ternary complex formation was obtained for the CVB3 $3C^{Pro}$:GC-376 PROTAC:CRBN-midi system, while steric hindrance likely prevented similar assembly with SARS-CoV-2 $3CL^{Pro}$.

A third research axis focused on the optimization of SARS-CoV-2 Nsp13 expression and purification. The implementation of zinc and iron supplementation increased soluble protein yield up to 0.5 mg/L of culture, enabling future studies on the enzyme's interaction with novel inhibitors within the PNRR Spoke 4 collaborative framework.

Finally, the study outlined the prospective use of cellular experiments, X-ray crystallization, computational modelling to characterize PROTAC-mediated ternary complexes. Collectively, the results provide molecular insights into inhibition and targeted degradation mechanisms, advancing the understanding of viral protease druggability and paving the way towards structure-guided antiviral design for future pandemics.

Table of contents

The present work is funded by the European Union – Next Generation EU, Mission 4 Component 2, Investment 1.5 – CUP B83C22003920001.

List of Figures

1. Chapter – Introduction

1.1. - Proteins as druggable targets

A 'druggable' target is a protein, peptide or nucleic acid with activity that can be modulated by a drug, which can consist of a small molecular weight chemical compound (SMOL) or a biologic compound (BIOL), such as an antibody, hormones or a recombinant protein [1]. SMOL drug targets mainly belong to the protein classes of enzymes, extracellular and nuclear receptors, ion channels, and transporters [2]. Identifying a good drug target is a challenging task. The process consists of a series of rigorous steps: target identification, target validation, druggability assessment and assayability evaluation. The properties for defining an ideal target include:

- The target plays a disease-modifying role and/or has a well-established function in disease pathophysiology.
- Target modulation is of limited relevance under normal physiological conditions or in unrelated diseases.
- If druggability is not evident, a 3D structure of the protein or of a close homolog should be available to assess it.
- The target is amenable to reliable assays, allowing high-throughput screening.
- Target expression is not uniformly distributed across tissues in the body.
- A disease- or target-specific biomarker is available to monitor therapeutic efficacy.
- Phenotype data support a favorable prediction regarding potential side effects.
- The target offers a favorable intellectual property landscape (minimal competition and freedom to operate).

The key factors contributing to successful target identification consist of profound disease pathophysiological understanding and knowledge of its causative molecular mechanisms with the application of relevant target identification and target validation technologies (**Figure 1**) [1].
The number of potential protein targets could be larger than the number of the corresponding genes, since the presence of post-translational modifications and assembly of functional complexes; however, this is not likely to increase the number of specific sites modulated by a drug [3]. The molecular druggability of a target can be evaluated through several factors: (i) analysis of its three-dimensional (3D) structure, including any available co-complex data; (ii) identification of low-molecular-weight modulators and assessment of their drug-likeness and selectivity; and (iii) consideration of prior knowledge and internal experience with the target class. [4]. Combined

with projected timelines for lead generation, these elements allow prediction of the target's druggability as high, medium, or low.

Current approaches to evaluate protein druggability consist of methods exploring sequence-related properties of proteins as well as methods exploring the 3D-structures of protein [5,6]. Drug targets with known 3D-structures and potential drug targets with structures can be found in the Protein Data Bank (PDB) [7]. Knowing the 3D-structure is an advantage for a drug target evaluation because it enables the prediction of potential binding sites for SMOLs. Given the 3D structure of a protein, computational approaches have been developed to characterize binding pockets by geometrical and physicochemical criteria and assess their ability to accommodate low-molecular-weight compounds for ligand design [8].

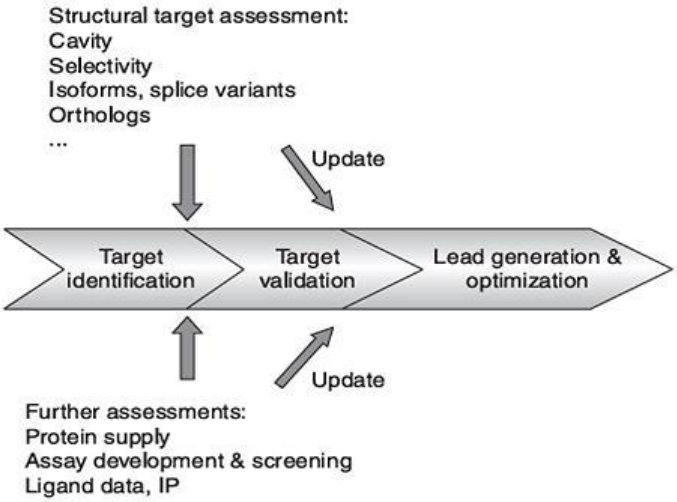


Figure 1. Structural druggability assessment in the drug discovery process (adapted from [6]).

Advances in structural and molecular biology, together with progress in biomolecular spectroscopic techniques, have greatly expanded our ability to determine 3D protein structures, now exceeding 100,000 entries [9]. Access to high-resolution structures of therapeutically relevant proteins enables the identification of binding cavities and supports the structure-based drug design (SBDD). Today, SBDD has become an essential component of both industrial drug discovery programs and academic research [10]. Thank to this approach, the development of protease inhibitor, Saquinavir [11] by Roche, has made HIV-AIDS a manageable and treatable disease. There are other drugs also, Indinavir [12], Nelfinavir [13], Amprenavir [14], Fosamprenavir [15], Tipranavir [16] discovered by SBDD from the same class that are approved by the FDA and are available in the market [17]. SBDD offers a more precise, efficient, and rapid approach to lead

discovery and optimization (**Figure 2**), as it directly exploits the 3D structure of the target protein together with molecular-level insights into the disease [18]. SBDD is an iterative process, which undergoes multiple cycles of iteration until optimization of the lead molecule is achieved. Among the relevant computational techniques, structure-based virtual screening (SBVS), molecular docking, and molecular dynamics (MD) simulations are the most common methods used in SBDD. These methods are widely applied to study binding energetics, characterize ligand-protein interactions, and assess conformational changes that occur during the docking process [19]. The fundamental step of SBDD process involves cloning of the target gene followed by the extraction, purification, and 3D structure determination of the protein. The 3D structures of all therapeutically important proteins are determined experimentally by integrative structure biology techniques such as NMR, X-ray Crystallography, or Cryo-Electron microscopy [20].

Among the three methods, X-ray crystallography is one of the most widely used techniques. To date, it counts more than 180,000 atomic structures of biological macromolecules and macromolecular assemblies deposited in the PDB. In most favorable cases, fragment libraries of a few hundred compounds can be screened using X-ray crystallography, delivering dozens of structures of complexes [21,22]. The aim of X-ray crystallography is to obtain a three-dimensional molecular structure from a crystal. A purified sample at high concentration is crystallized and the crystals are exposed to an X-ray beam [23]. The resulting diffraction patterns can then be processed, initially to yield information about the crystal packing symmetry and the size of the repeating unit that forms the crystal. This is obtained from the pattern of the diffraction spots. The intensities of the spots can be used to determine the "structure factors" from which a map of the electron density can be calculated [23]. However, many biomolecules are difficult to crystallize, or even cannot be crystallized because of factors such as intrinsically disordered regions and the heterogeneous conformational properties of samples [24]. Even though crystals of the selected target have been obtained, it might still be difficult to generate conditions for obtaining crystals/structures of the desired complexes [25].The mixture of free proteins and complex proteins in the crystal can cause poor electron density and thus prevent the ligand from inserting into the structure. Some differences were observed between structures of complexes obtained through soaking or co-crystallization [26]. As soaking implies the use of preformed crystals of the protein target, any conformational adaptation to ligand binding may be limited and the obtained structure might not reflect the binding modes that would be observed in the solution.

By contrast, NMR spectroscopy can determine 3D biomolecular structures in solution or in their solid state without the need to obtain crystals [27-29]. NMR spectroscopy has been particularly

useful for defining the conformation of proteins with low molecular weights as well as proteins with large intrinsically disordered regions [27,29-36]. NMR spectroscopy is the only method that allows direct monitoring of the conformational properties of a protein or its complex at the individual amino acid level (e.g. via recording ^{1}H, ^{15}N and ^{13}C chemical shifts) [24] and performs well for weak intermolecular interactions with dissociation constant (K_d) in the μM~mM range [37]. NMR spectroscopy allows monitoring of conformational changes of a protein or its complex upon the introduction of changes to the sample condition. Moreover, NMR spectroscopy has been exceedingly powerful in investigating dynamic properties of biomolecules [38-41]. Pharmaceutical NMR methodologies can be divided into two major categories: protein-based and ligand-based [42,43]. In the early stage of SBDD, a ligand-based approach is useful for the screening of hit ligands, whereas a protein-based approach results profitable for hit validation, such us structure determination of protein-ligand complexes [44,45]. Through the latter, ligand binding sites may be identified by chemical shift perturbation and mapped on the target protein [46]. Protein-based experiments require isotope labelling (^{15}N, ^{13}C) of a protein of interest, which can be accomplished by heterologous protein expression systems, using living host cells such as *Escherichia coli* and ^{13}C-enriched sugars and/or ^{15}N-enriched ammonium salts as carbon and nitrogen sources, respectively [42,47]. Uniform ^{13}C and/or ^{15}N labelling is necessary to assign NMR resonances of a target protein by performing traditional two-, three-, or higher-dimensional NMR measurements [42,47]. However, the difficulty of resonance assignment increases with increasing molecular weight and decreasing molecular tumbling speed, since these factors cause signal degeneration and line broadening [48]. Ligand-based methods are attractive because they are broadly applicable, impose few constraints on the composition of the target protein and do not require isotopic labelling of the protein or ligands [37]. Then, they can be used with an excess of ligand and a small concentration of protein and there is no upper limit size for the protein. However, ligands with very high affinities have low off-rates from the complex and can therefore score as non-binders [37].

Cryogenic electron microscopy (cryo-EM) is revolutionizing structural biology due to its unique capability of determining the structures of large protein complexes and assemblies. The atomic-resolution structure determination for proteins enabled by cryo-EM [49] allows us to understand the complex biological processes carried out by proteins as well as to identify potential therapeutic protein targets for drug discovery. Briefly, the method for sample preparation and analysis consists of applying a purified monodisperse protein sample onto an electron microscope grid, which acts as a support for the sample. The grids are manufactured to have a thin polymer layer with evenly

spaced holes, and the protein molecules are suspended in these holes. The grid is flash frozen so the protein is trapped in vitreous ice. The protein molecules are ideally isolated from neighbouring molecules and in random orientations [50,51]. Single-particle analysis using programmes such as Relion [52] extracts just the individual molecules from the images, centres each particle and averages the molecules, first in two dimensions and then in three dimensions to eventually give a 3D density map into which an atomic model can be built. Cryo-EM allows structure determination of large and/or dynamic macromolecular assemblies in solution without the need to obtain crystals and, even more importantly, in their native state, including post-translational modifications; moreover, different conformational states can be resolved in one experiment. The reported resolution of cryo-EM structures has historically been lower than X-ray crystallography-generated structures [50]. However, recently published cryo-EM structures include biologically important and diverse structures that could probably not have been analyzed by X-ray crystallography, including large dynamic assemblies and membrane protein complexes, but also ever-smaller objects (the current reported limit being human haemoglobin (64 kDa) [53]) and structures at increasingly high resolution, including complexes with small molecules (the current reported limit being glutamate dehydrogenase at (1.8 Å) resolution [54]). Therefore, cryo-EM presently complements other structural biology techniques such as X-ray crystallography and NMR, with its strengths lying particularly in studying large structures of multi-megadalton size for which NMR imposes severe technical demands and for which crystallization is extremely difficult [51].

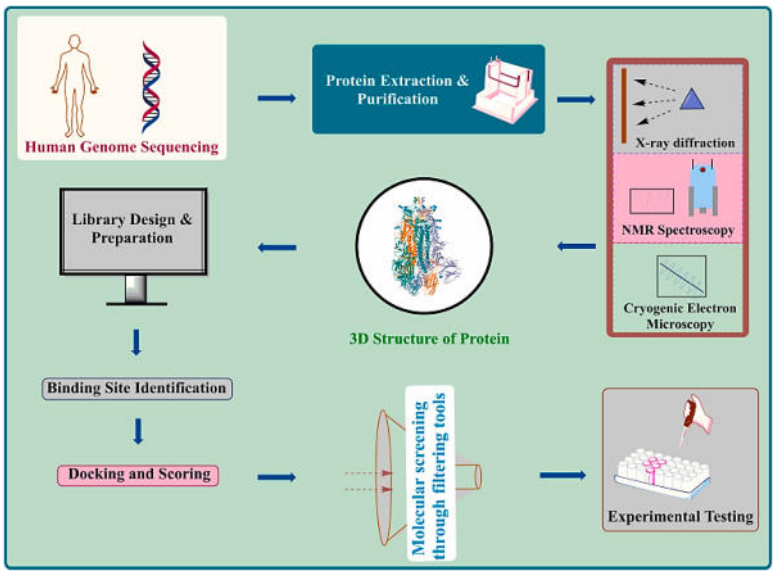


Figure 2. Workflow of structure-based drug design (SBDD) in the drug discovery process [55].

In this regard, structural biology could provide important contributions to characterize networks of druggable proteins, analyze protein-protein interactions (PPI) and protein ligand interactions, and protein post-translational modifications (necessary for target validation) both in vitro and in their native cellular environment. In fact, elucidation of interactions between protein complexes and their functional environment contributes to validating proteins or specific domains as drug targets.

In the case of infectious diseases, the target is usually a protein of the pathogen. In particular, viral pathogens appear constantly and have cost millions of human lives [56]. The National Institute of Allergy and Infectious Diseases of the United States of America defines emerging viruses as those "that have newly appeared in a population or have existed but are rapidly increasing in incidence or geographic range". Currently, nine viral diseases, including COVID-19, Ebola and Zika, are listed by WHO as prioritized diseases for research and development in emergency contexts. Structural biology appears to be fundamental in guiding the optimization of candidate compounds and the identification of targets, enabling the accelerated development of antiviral potential therapies for future pandemics. Within this frame, my PhD research activity was focused on the study of three viral target proteins, involved in SARS-CoV-2 (sections 1.2. and 1.7.) and CVB3 (section 1.3.) infections.

1.2. - SARS-CoV-2 $3CL^{Pro}$ as drug target

Coronaviruses cause upper respiratory tract illnesses and can infect multiple species. The first severe coronaviruses to impact humans were the severe acute respiratory syndrome-coronavirus (SARS-CoV), which emerged in 2002 [57-59] and the Middle East respiratory syndrome (MERS-CoV), which emerged in 2012 [60]. After the emergence of MERS-CoV, researchers expressed the importance of developing future therapeutic interventions for coronaviruses. The coronavirus which emerged in 2019 is a result of a new strain of severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2) [61]. Although vaccines are being pursued for COVID-19 as a possible therapeutic route, they can have limited effectiveness when strains continually change. Coronaviruses appear to be on a path of continual reemergence with new strains developing, and, as a result, other therapeutic interventions in addition to vaccines must be pursued.

The virus primarily affects the respiratory system causing flu-like illness with symptoms such as a cough, fever, and in more severe cases, difficulty breathing [62]. In addition to acute respiratory distress syndrome and respiratory failure, COVID-19 is now known to manifest systemic

inflammation, leading to sepsis, acute cardiac injury, and heart failure and multi-organ dysfunction in patients at high risk [63].

CoVs belong to the Coronaviridae family of order Nidovirales [64]. They have been classified into four genera that include α-, β-, γ-, and δ coronaviruses [65]. SARS-CoV-2 belongs to the genus Betacoronavirus (Beta-CoV) and is a single stranded positive-sense RNA (ssRNA(+)) ~ 29.9 Kb in length, composed of 13-15 (12 functional) open reading frames (ORFs) [66]. The SARS-CoV-2 genome shares about 82% sequence identity with SARS-CoV and MERS-CoV and >90% sequence identity for essential enzymes and structural proteins. This high level of the sequence revealed a common pathogenesis mechanism, thus, therapeutic targeting. [66,67].

The whole genome of SARS-CoV-2 encodes about 7096 residues long polyprotein which consists of many structural and non-structural proteins (NSPs). The nucleotide content of the viral genome is held mainly (~ 70%) by two non-structural proteins ORF1a and ORF1ab followed by structural proteins [66]. Polyproteins pp1a and pp1ab are encoded by ORFs 1a and 1b, where polyprotein pp1ab is encoded by the ribosomal frameshift mechanism of the gene 1b. These polyproteins are further processed by virally encoded proteinases and produce 16 NSPs, which are well conserved in all CoVs belonging to the same family [66]. The other 30% of the viral genome encodes 4 structural proteins (including spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins) and accessory proteins (3a, 6, 7a, 7b, 8, 9b), all essential for virion formation. The S protein has two structural subunits, where in the 'S1' subunit contains the cellular angiotensin-converting enzyme 2 (ACE2) receptor-binding domain (RBD) and the 'S2' subunit possesses structural elements required for cell membrane fusion. The M protein is a trans-membrane glycoprotein crucial for membrane fusion, whereas the E protein is required for virion assembly and morphogenesis [68]. The non-structural replicase proteins (pp1a: nsp1–nsp11 and pp1b: nsp12–nsp16) are involved in mRNA synthesis and replication, and the accessory proteins (3a, 3b, 6, 7a, 7b, 8 and 9b) participate in modulating host innate immunity and viral adaptation [69] (**Figure 3**).

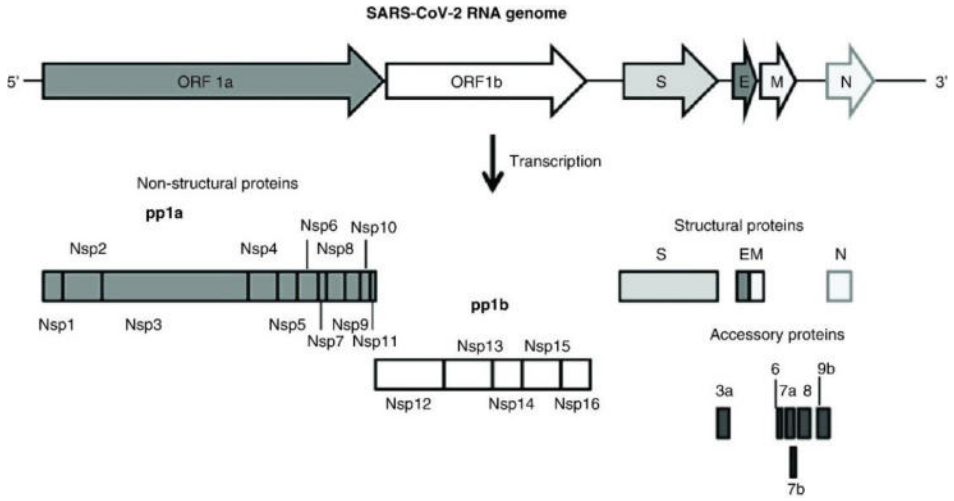


Figure 3. Genome organization of the SARS-CoV-2 [69].

The infection cycle of SARS-CoV-2 begins with the interaction of the spike (S) protein expressed on virion particles and the angiotensin-converting enzyme II (ACE2) of the host cell. Upon entry, the viral RNA associates with host ribosomes, resulting in the translation of the two large co-terminal polyproteins that are processed by proteolytic enzymes, 3C-like Protease ($3CL^{Pro}$) and Papain-like Protease (PL^{Pro}) [66]. They cleave the C-terminus of the polypeptide chain at 11 sites and the N-terminus of the polypeptide chain at three sites, respectively [70]. The cleavage products include structural proteins and some important non-structural proteins, such as RNA-dependent RNA polymerase (RdRp) and helicase [71].

$3CL^{Pro}$ (also named main protease, M^{pro} or NSP5) is encoded by ORF1a and with more cleavage sites than PLpro, it has always been an attractive non-structural protein for the development of drugs targeting coronavirus. SARS-CoV-2 $3CL^{Pro}$ is a dimeric cysteine protease consisting of 306 amino acid residues. Each monomer of the protease contains two regions: the N-terminal catalytic dyad region and the C-terminal region. The N-terminal catalytic dyad is located in a cleft between chymotrypsin-like domains I (residues 8–101) and II (residues 102–184) (**Figure 4**) [71]. Domains I and II consist of six-stranded antiparallel-barrel structures, which together resemble the architecture of chymotrypsin and the other 3C proteases found in picornaviruses [72]. The C-terminal region (domain III) contains five alpha-helices, and the globular folds formed by these helices are the basis for maintaining the structure of the active dimer of $3CL^{Pro}$ [73]. Meanwhile, domain III (residues 201–303) is connected to domain II through a long flexible loop (residues 185–200), making the structure more flexible [74]. It is noteworthy that the two monomers are inserted into the C-terminal domain III at approximately right angles to each other. Therefore, any deflection of the C-terminal domain III of the dimer will affect its activity. Dimerization of the

enzyme is necessary for catalytic activity, because the N-finger (residues 1-7) of each of the two protomers interacts with Glu166 of the other protomer and thereby helps shape the S1 pocket of the substrate-binding site. Glu166 in SARS-CoV-2 forms critical interactions with the N-terminal finger residues of the heterologous monomer. Deletion of the first few residues in the N-terminal finger of SARS-CoV results in significant loss in enzymatic activity and disruption of dimerization [75-77]. Dimerization involves other specific intermolecular interactions between two protomers such as the salt bridge between N-terminus and domain III of another protomer and the specific electrostatic and hydrophobic interactions between the two different domains III [78].

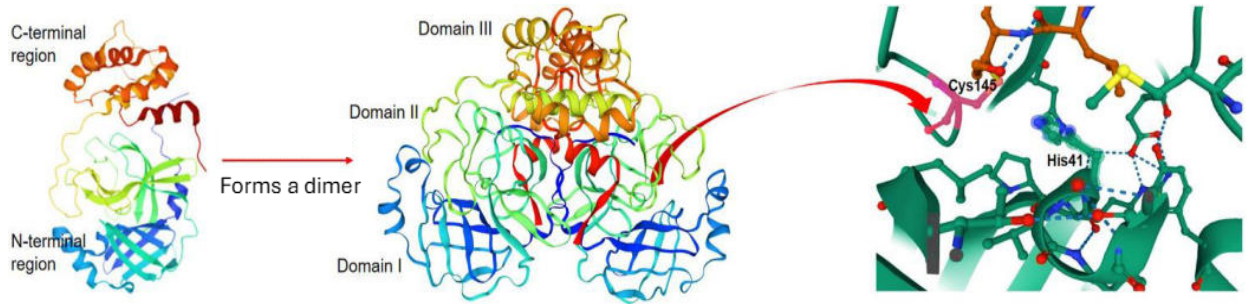


Figure 4. SARS-CoV-2 $3CL^{Pro}$ 3D structure (monomer and dimer) and the N-terminal catalytic dyad (His41-Cys145) (adapted from [71]).

Sequence alignment revealed that the $3CL^{Pro}$-CoV-2 shares 96% similarity with SARS, which is highly conservable among CoVs [79]. Meanwhile, the structure of these proteins (in apo forms) was also found to be very similar to the low R.M.S.D. value (0.459 Å) from the structural superimposition. Notably, the sequence alignment showed that there are 12 non-conserved residues (out of 306 total residues) in $3CL^{Pro}$-CoV-2. Nevertheless, mutational works on these residues indicated that none of the mutations on these residues (T35V, A46S, S65N, L86V, R88K, S94A, H134F, K180N, L202V, A267S, T285A, I286L) seriously affected the catalytic activity of $3CL^{Pro}$-CoV-2. This indicated that the non-conserved residues have no major roles in the enzymatic activity of $3CL^{Pro}$-CoV-2 [80]. Interestingly, polar interactions via a hydrogen bond between two Thr285 of each monomer were observed in SARS-CoV, but not in SARS-CoV-2 [81]. The absence of this residue in $3CL^{Pro}$-CoV-2, nevertheless, has no effect on its dimerization and the activity of this protein. Among conserved residues, there are 15 residues (His41, Met49, Gly143, Ser144, His163, His164, Met165, Glu166, Leu167, Asp187, Arg188, Thr190, Ala191, Gln192) responsible for the substrate binding. Moreover, 10 conserved residues (Arg4, Ser10, Gly11, Glu14, Asn28, Ser139, Phe140, Ser147, Glu290, Arg298) are responsible for the enzyme dimeric structure. The $3CL^{Pro}$, is not only the most cysteine protease conserved in structure, but also in its function in all known CoVs [82].

The substrate binding site of the $3CL^{Pro}$ is composed by six subsites (S1'-S5), labelled using the Schechter and Berger nomenclature [83], which corresponds to P1'-P5 positions of the recognition sequence. The amino acid species recognized by coronavirus $3CL^{Pro}$ at the S1', S1, S2 and S4 are specific, and this phenomenon is the same as the substrate specificity of picornavirus 3C protease [72]. The enzyme has a recognition cleavage sequence of Leu-Gln↓Ser-Ala-Gly (↓ marks the cleavage site), in which the Gln forms hydrogen bonds with residues in the S1 subsite and a small amino acid (Ser, Ala, or Gly) occupies the S1' subsite [81]. The absolute dependence of the virus on the correct function of this protease, along with the absence of a homologous human protease, makes $3CL^{Pro}$ one of the most pursued targets for the development of specific protease inhibitors [79,84]. S1' is an important site of covalent warhead action, which contains the Cys145-His41 catalytic dyad. Cys145 works as a nucleophile, while His41 acts as the proton donor/acceptor, participating in the catalytic process. The S1 and S2 binding subsites are two deep pockets [71]. An oxyanion ring, which is mainly formed by the main chain nitrogen of Gly143, Ser144, and Cys145, exists on the S1 subsite, that is an indispensable functional component of $3CL^{Pro}$ [85]. In the S1 subsite, the imidazole side chain of conserved histidine residue interacts with the carboxamide side-chain of P1; this reaction is generally accepted to be conducive of specificity for glutamine residue at P1 [86,87]. Strong hydrogen bonds between Gln-P1 and His163, Gln-P1 and Phe140 ensure suitable binding of the substrate into the S1 site [86]. The S2 subsite is a flexible hydrophobic cavity that favors hydrophobic amino acids, such as leucine and the absence of any electron density was found to indicate the presence of ordered water [88]. On the C-terminal side of the substrate, the P1' position is occupied by a small residue like Ala, Ser and Gly residues which directly interact with the S1' shallow subsite through van der Waals interactions [89]. As a shallow hydrophobic pocket, the S4 subsite forms a narrow cavity and it always includes smaller amino acids (Ala, Val, Pro, and Thr). The S1'-S1-S2-S4 core sites include Leu27, His41, Met49, Ile51, Tyr54, Phe140, Leu141, Ser144, Cys145, Tyr161, His163, Met165, Glu166, Leu167, His172, Asp187, Thy190, and Ala191 residues [73,90] (**Figure 5**). In addition, His163 and Glu166 are essential amino acids that recognize substrate Gln-P1 and participate in hydrogen bond formation [77]. Then, the flipped Asn142 acts as a half-gate to cover the S1 subsite. Met49 protrudes from the S2 subsite resulting in an upward shift of Gln189, which forms the other half of the gate and interacts with the P2 position of the substrate or inhibitor. Thus, Asn142, Met49, and Gln189 are considered to be substrate binding induced gate regulation switches at S1 and S2 subsites [88].

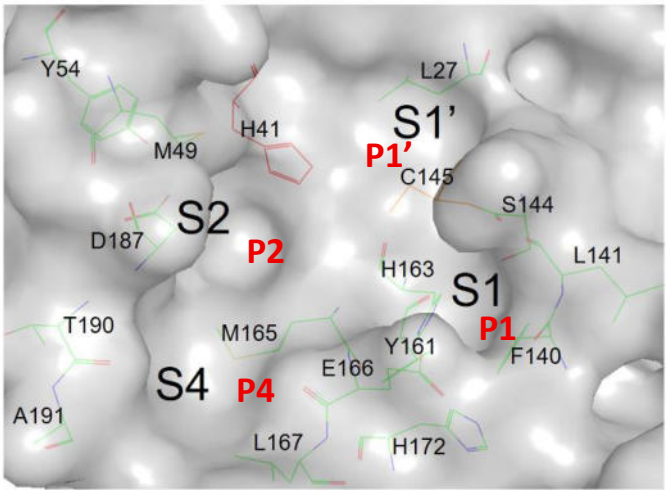


Figure 5. The substrate-binding pocket of SARS-CoV-2 3CL^{Pro} (adapted from [71]).

Unlike the traditional catalytic triad of chymotrypsin, the Cys145-His41 dimer constitutes the SARS-CoV-2 3CL^{Pro} catalytic site, which is located between S1' and S1 subsites [90,91]. During the SARS-CoV-2 3CL^{Pro} catalytic process, the protonation state of the imidazole group of His41 leads to the formation of an ionized H41+—C145 dimer [92]. The deprotonated Cys145-thiol is highly nucleophilic and attacks the carbonyl carbon of the scissile bond in the substrate peptide plane. After nucleophilic attack, the N-terminal substrate fragment is released from the enzyme, and a thioester is formed between the enzyme's sulfhydryl group and the carbonyl group at the substrate's cleavage site. Next, the thioester is hydrolyzed by nucleophilic attack of a deprotonated water molecule, the corresponding C-terminal substrate fragment is released, and free enzyme is regenerated [93,94] (**Figure 6**). The water molecule maintains the protonation state of His41 and forms hydrogen bonds with the side chains of His41, His164, and Asp187 to stabilize their positions [95].

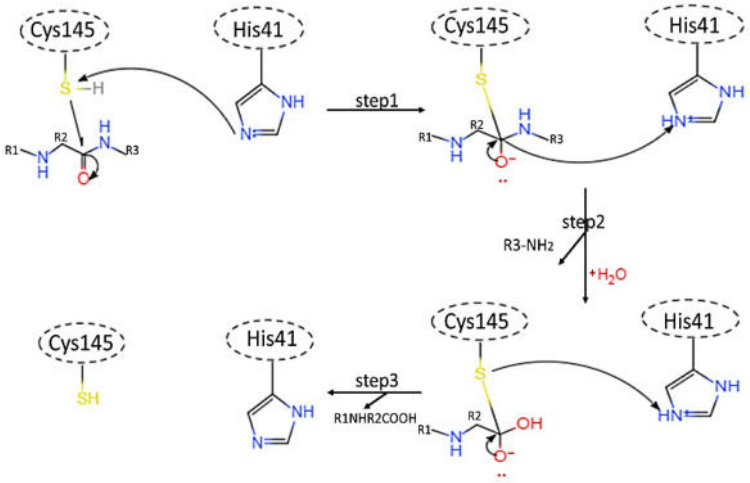


Figure 6. SARS-CoV-2 3CL^{Pro} catalytic mechanism (adapted from [71]).

1.3. - CVB3 $3C^{Pro}$ from Enterovirus

$3CL^{Pro}$ of SARS-CoV-2 exhibits a high degree of structural similarity across 3CL proteases in all genera of Coronaviruses as well as across 3C proteases belonging to the Enterovirus genus within the large Picornavirus family [96,97]. Indeed, the name of '3C-like protease' refers to the similarities between this $3CL^{Pro}$ protease and the other 3C proteases found in enteroviruses and picornaviruses, namely their similar core structural homology and substrate specificities [72]. Infection with some of these viruses can lead to serious outcomes and cause clinical disease much more frequently than coronaviruses. As an example, Coxsackievirus of group B3 (CVB3), a cardiotropic virus, is among numerous known etiologies for myocarditis [98]. This serotype of enterovirus has been long identified as one of the leading causes of viral myocarditis, a condition that may proceed to acute as well as chronic heart failure [99,100]. No effective therapeutic strategy for the prevention and treatment of these diseases is nowadays available, and enterovirus disease has a medical impact, with newborn infants and young children being at risk for septic-like disease [101]. With regard to tissue permissiveness, CVB3 infects multiple human organs, yet cardiac tissue is among the most pertinent targets for CVB3 infection [102].

CVB3 is a non-enveloped single-stranded (+) RNA enterovirus of approximately 7.5 Kb, which codes for capsid proteins VP1-VP4 and for 7 non-structural proteins: two viral proteases (2A, and 3C), an RNA-dependent-RNA-polymerase (3D), two proteins involved in RNA synthesis (2B, and 2C), a primer of initiation of RNA synthesis (3AB) and a small polypeptide (VPg) of 20–24 amino acids derived from gene 3B [103-105] (**Figure 7**).

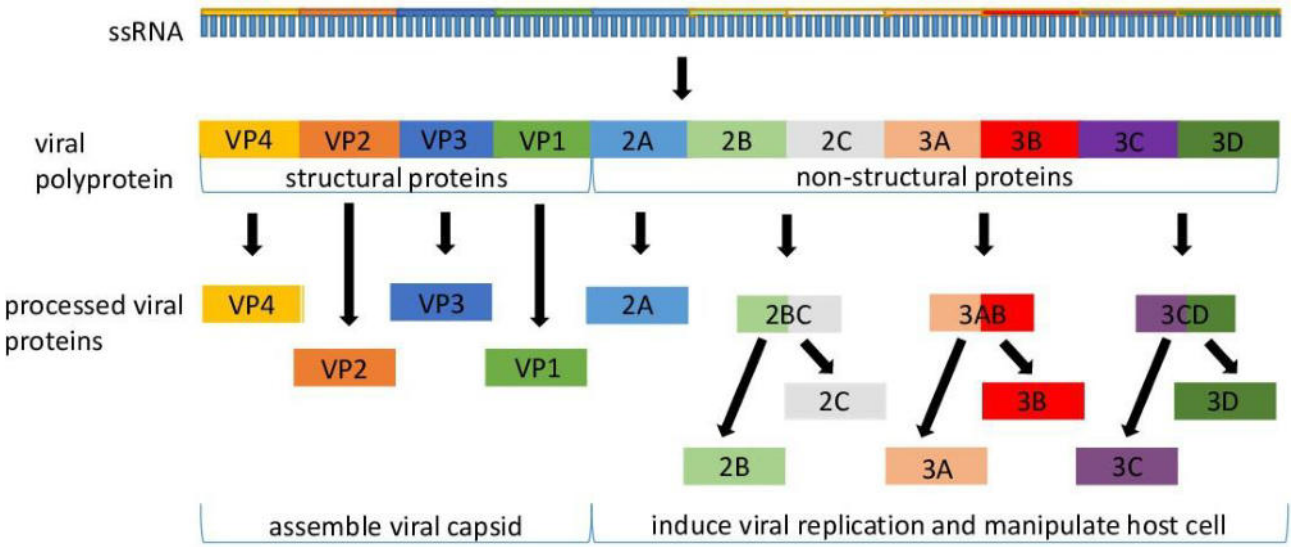


Figure 7. Organization of the Coxsackievirus B3 genome [105].

CVB3 infection begins by coupling the virus to host-cell receptors. CVB3 is dependent upon binding to both CAR (coxsackievirus, and adenovirus receptor), and DAF (decay-accelerating factor [DAF also known as CD55]) as the main receptor, and co-receptor, respectively [106]. Upon infection, the positive-strand RNA genome of CVB3 is translated into a large poly-protein precursor that is subject to successive proteolytic cleavage processes to generate functional and structural viral proteins [107]. In particular, $3C^{Pro}$ is a picornaviral cysteine protease required for the proteolytic processing, essential for viral replication [108]. Structural studies on $3C^{Pro}$ from hepatitis A virus (HAV), human rhinovirus (HRV), poliovirus (PV), FMDV and Coxsackievirus B (CVB) have shown that this family of cysteine protease shares an overall fold with many serine proteases, such as chymotrypsin. Picornaviral 3C proteases possess Cys-His-Glu/Asp catalytic triads similar in geometry with Ser-His-Asp triad found in serine proteases [109]. The sequence and structural alignment show that $3C^{Pro}$ contains a Gly-X-Cys-Gly motif, similar to that of the active site of chymotrypsin-like serine proteases, which is Gly-Asp-Ser-Gly [110]. This catalytic important motif Gly-X-Cys/Ser-Gly-Gly has the role of position Cys/Ser for nucleophilic attack and form the oxyanion hole, conserved in all picornaviral 3C proteases [109]. Another important structural element shared among picornaviral $3C^{Pro}$ is the flexible surface loop β-ribbon, which plays a crucial role in substrate recognition.

CVB3 $3C^{Pro}$ consists of 180 amino acid residues (~21KDa) and is a monomeric protease. The structure of CVB3 $3C^{Pro}$ adopts a chymotrypsin protein fold similar to those of RV $3C^{Pro}$ [111] and other viral 3C proteases. The root-mean-square deviation between CVB3 $3C^{Pro}$ and RV $3C^{Pro}$ is 0.73Å for 166 Cα atoms. The N-terminus starts with an α-helix of residues 1–14 and is followed by two topologically equivalent β-barrel domains comprising residues 15–77 and 99–173, which pack together to form an extended shallow groove for substrate binding (**Figure 8**). The catalytic triad of Cys147, His40, and Glu71 is located in the cleft between the two-barrel domains [112]. The structural elements that differentiate CVB3 $3C^{Pro}$ and RV $3C^{Pro}$ include the loop β-ribbon (143-146) and the different side-chain orientations of Cys147 and His40 [112]. As mentioned above, the loop β-ribbon has high flexibility and plays a crucial role in the substrate recognition and binding. Studies have shown that PV $3C^{Pro}$ preferentially cleaves polypeptides with P1-Gln/P1'-Gly junctions [109,112].

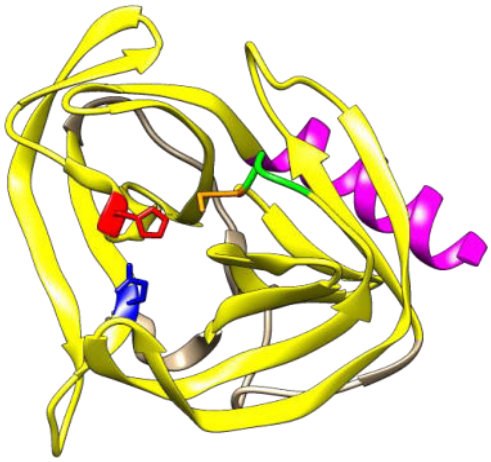


Figure 8. Structure of CVB3 $3C^{Pro}$ (PDB ID: 3ZYD) depicting the α-helix in magenta, the loop β-ribbon in green and the two equivalent β-barrel domains in yellow. The catalytic triad is shown with Cys147-His40-Glu71 residues in orange, red and blue, respectively.

1.4. - $3CL^{Pro}$ and $3C^{Pro}$ common inhibitors rationale

Both coronavirus $3CL^{Pro}$ and enterovirus $3C^{Pro}$ are cysteine proteases that utilize a cysteine nucleophile and share a preference for cleaving the P2–P1↓P1′–P2′ peptide sequence (hydrophobic residue – Gln – small residue – small residue), with the scissile bond located between P1 and P1′. Despite this shared specificity, subtle differences exist: coronavirus $3CL^{Pro}$ favors a medium-sized hydrophobic P2 residue (preferably Leu), whereas enterovirus $3C^{Pro}$ prefers a bulkier hydrophobic residue (preferably Phe). A major mechanistic distinction lies in their oligomeric state, as $3CL^{Pro}$ requires homodimerization for catalytic activity [87], whereas enterovirus $3C^{Pro}$ works as monomer. Structurally, both protease families adopt a two-barrel domain architecture with the catalytic site situated at the interdomain interface. Notably, coronavirus $3CL^{Pro}$ additionally contains a C-terminal helical domain, which contributes to the regulation of dimerization and thereby modulates enzymatic activity [87,113]. Finally, coronavirus $3CL^{Pro}$ features a catalytic dyad comprising a Cys ... His pair, whereas enterovirus $3C^{Pro}$ proteins have a catalytic triad consisting of Cys ... His ... Glu (or Asp) [114]. Despite of all these differences, Coronaviral and Enteroviral main proteases ($3C^{Pro}$ and $3CL^{Pro}$) share several common characteristics including a chymotrypsin-like fold, a Cys residue as a nucleophile in the catalytic triad ($3C^{Pro}$) or dyad ($3CL^{Pro}$), their almost absolute requirement for Gln in the P1 position of the substrate and space for only small amino-acid residues such as Gly, Ala, or Ser in the P1′ position on the natural substrate [97]. The related architecture of the substrate binding sites of $3CL^{Pro}$ and $3C^{Pro}$ encouraged the design and development of inhibitors targeting both families of enzymes with approximately equal potency.

In recent years, several studies have highlighted the antiviral potential of metal ions and metal-based compounds, including zinc, rhenium, gold, and copper derivatives, which have been shown to act as efficient inhibitors of viral proteases [115-117]. Among these, Auranofin, an FDA-approved oral gold(I)-based drug originally developed for the treatment of rheumatoid arthritis, has emerged as a particularly interesting candidate for its combined antiviral and anti-inflammatory effects. Ott and co-workers [116] reported the efficacy of Auranofin and related gold complexes against the Papain-Like protease (PL^{Pro}). Furthermore, the crystal structure of the adduct formed between SARS-CoV-2 $3CL^{Pro}$ and Auranofin demonstrated that gold can, at least partially, coordinate the catalytic cysteine residue [118]. A subsequent broader study [119] confirmed the ability of a larger set of gold compounds to inhibit the catalytic activity of both SARS-CoV-2 $3CL^{Pro}$ and PL^{Pro}, reinforcing the concept of gold-based compounds as protease inhibitors. Moreover, 3D structures of SARS-CoV-2 $3CL^{Pro}$ in complex with zinc ion [120-122] showed that the metal ion is bound to the catalytically relevant Cys145 and His 41 residues. The binding capability of SARS-CoV-2 $3CL^{Pro}$ towards the zinc ion is not completely unexpected as homologous proteases are reported to bind a zinc ion, not in its ionic form but as a metal conjugated ligand. Zinc-coordinating molecules were also tested and found active against the recombinant $3C^{Pro}$ of CVB3. The zinc-coordinating inhibitor is tetrahedrally coordinated to the His40-Cys147 catalytic triad of CVB3 $3C^{Pro}$ [112].

However, in the large landscape of protease inhibitors, peptidomimetics are the most suitable technology. Zhang et al., in 2020, designed and synthesized peptidomimetics of the general structure cinnamoyl—P2—γ-lactam— α-ketoamide—P1', compounds differing in the substituent at the P2 position some of which turned out to have broad-spectrum activity versus betacoronaviruses and enteroviruses [97]. Nucleophilic attack by cysteine residues on the α-ketone carbon results in sulfhydryl hemiketals through competitive inhibition with substrates [123]. After the SARS-CoV-2 outbreak in December 2019 starting from crystal structures of the $3CL^{Pro}$ in complexes with the Hepatitis C virus NS3/4A protease inhibitors boceprevir and telaprevir, Zhang et al. optimized the potency of the alpha ketoamide boceprevir against the $3CL^{Pro}$ by replacing its P1 cyclobutyl moiety by γ-lactam as a glutamine surrogate. They conclude that the modified boceprevir scaffold is suitable for obtaining high-potency inhibitors of the coronavirus $3CL^{Pros,}$ but further optimization would be needed to target enterovirus $3C^{Pro}$ efficiently. Indeed, only moderate inhibitory activity was observed against the 3C proteases of enteroviruses CVB3 and EV-A71, suggesting steric incompatibilities that likely diminish inhibitory potency [114].

Michael acceptors, characterized by olefinic or acetylenic bonds conjugated to electron-withdrawing groups, display strong electrophilicity and readily react with cysteine residues through nucleophilic addition to their sulfhydryl groups [124]. Rupintrivir or AG7088 is part of the class of Michel acceptor inhibitors, one of the first drug candidates and behaves as a covalent inhibitor. It was preliminary designed to target the 3C cysteine protease of human rhinovirus B [111] and subsequently also was shown to be a broad-spectrum inhibitor against viruses in the Picornaviridae family with a similar binding mode [125]. Despite their similarities however, rupintrivir only weakly inhibits the $3CL^{Pro}$ of SARS-CoV-1 and 2 with an IC50 of >100 μM and 68 ± 7 μM, respectively [126,127].

The mercaptan of the Cys145-His41 dimer plays an important role in the enzymatic reaction of coronavirus $3CL^{Pro}$. As electrophilic functional groups, aldehydes can undergo nucleophilic addition to mercaptan to form stable hemimercaptol [128]. Therefore, peptide aldehydes containing aldehyde warheads are potent inhibitors of $3CL^{Pro}$. Qiao et al. [129] as well as Xia et al. [130] replaced the P1 residues of both boceprevir and telaprevir with the canonical-lactam moiety. However, they used an aldehyde as the warhead of their inhibitors. Qiao et al. found that both boceprevir- and telaprevir-derived aldehydes were roughly equipotent inhibitors of the SARS-CoV-2 $3CL^{Pro}$, also previously shown by Zhu et al. [131]. GC376 is a dipeptidyl bisulfite adduct salt that has been reported to be a broad-spectrum inhibitor with excellent inhibitory effects against coronaviruses, picornaviruses, and noroviruses [109]. GC376 has been used in clinical trials to inhibit feline coronavirus (FIPV) [84], and showed to be a virtually non-cytoxic and safe antiviral candidate [132-134]. Feline coronavirus $3CL^{Pro}$ is very similar to SARS-CoV-2 $3CL^{Pro}$. In both, GC376 binds within the active site where the bisulfite salt of GC376 is easily desorbed under aqueous conditions and converted into dipeptidyl aldehyde (GC373). The aldehyde group is attacked by the nucleophilic reaction of Cys145 and covalently linked to Cys145, forming a dithioacetal [132]. Based on detailed structural analysis, similarities and differences of GC376 in interacting with different $3CL^{Pro}$s were revealed. The results reported in this study confirm the broad-spectrum inhibition of coronaviral $3CL^{Pro}$s by GC376 and provide valuable information for the development of broad-spectrum and potent inhibitors against SARS-CoV-2, SARS-CoV-2 variants, as well as other human coronaviruses. Moreover, GC376 and GC373 (**Figure 9**) compounds have broad-spectrum antiviral activity against multiple viruses in the picornavirus-like supercluster, including important classical and emerging animal and human pathogens [109].

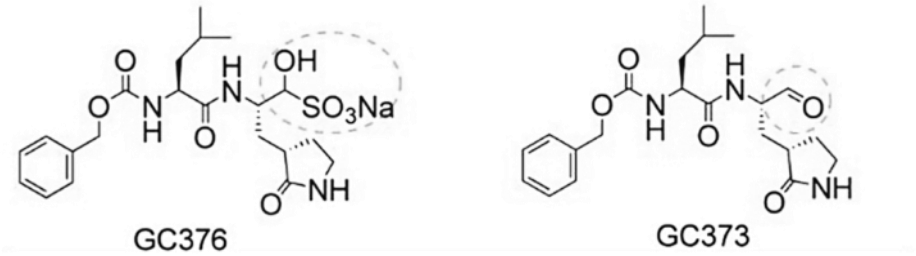


Figure 9. Chemical structures of GC376 and GC373 molecules (adapted from [135]). In grey dashed lines the functional groups of both molecules are represented: the bisulfite group in GC376 and the aldehyde group in GC373.

Owen et al. [136] also started from the boceprevir scaffold, but they replaced the ketoamide group with a nitrile as an electrophilic warhead. Their compound, PF-07321332 (nirmatrelvir), covalently reacts with Cys145 through its nitrile group, likely to give a complex in which a thioimidate electron pair occupies the $3CL^{Pro}$ oxyanion hole, which is formed by the main chain amide NHs of Cys145 and Gly143, mimicking binding by tetrahedral intermediates in catalysis. The use of a nitrile group as a warhead was apparently key to the successful development of nirmatrelvir, in part because the non-specific reactivity of nitriles with off-target nucleophiles may be lower than that of many other available electrophilic warheads and in part because the covalent reaction of the compound with the thiolate of Cys145 is reversible. Nirmatrelvir has been approved by both FDA and EMA and has been introduced into the market as a combination with ritonavir (as a booster) under the name Paxlovid. In the co-packaged medication nirmatrelvir/ritonavir, ritonavir serves to slow the metabolism of nirmatrelvir via cytochrome enzyme inhibition, thereby increasing the circulating concentration of the main drug. To overcome the metabolomic stability of nirmatrelvir, Pfizer developed a second-generation improvement, named Ibuzatrelvir (development code PF-07817883) with a similar chemical structure [137]. Ibuzatrelvir incorporates modifications to the chemical structure of nirmatrelvir that give it enhanced oral bioavailability, so it does not require coadministration with ritonavir [138].

1.5. - Targeted protein degradation (TPD)

Targeted protein degradation (TPD), via the proteasomal and lysosomal pathways, represents a promising therapeutic approach by inducing the depletion or reduction of a disease-causing protein via hijacking the endogenous protein degradation machineries. In addition to their therapeutic

potential, protein degraders are valuable chemical biology tools to validate targets and to gain a deeper understanding of protein function and cellular pathways [139]. In contrast to nucleotide-based methods like RNAi, antisense oligonucleotides or genome editing strategies, like CRISPR-Cas9 technology, which all suffer from limited therapeutic applications due to their low in vivo stability and bioavailability [140], TPD leverages more drug-like small molecules. Although nucleic acid-based tools have proved to be useful in research, the development of these agents as drug candidates has faced many challenges: unmodified nucleotides are highly unstable in serum [141], whereas modified nucleotides tend to accumulate in the kidney [142,143] and can be immunogenic [144]. Moreover, RNAi can engage off-target mRNA, which leads to undesired effects [145-147]. Finally, efficacy is ultimately dependent on the target protein half-life; thus, long-lived proteins are less affected by these approaches [148,149].

Small-molecule-induced protein degradation is emerging as a strategy that also has the potential to target a broader range of proteins than standard small-molecule strategies which focus on binding site occupancy. Classical approaches in the drug discovery process typically aim to identify high affinity small molecules modulating the activity of target proteins. This happens in an occupancy-driven manner with the inhibitor binding to its target and occupying the binding site, thereby inhibiting target function [150-152]. This strategy has been very successful for many targets harboring a tractable active or allosteric site like enzymes or receptors. Furthermore, small-molecule therapeutic strategies that focus on occupation of a binding site on a protein may require high systemic drug exposures to achieve sufficient site occupancy *in vivo* [153] and thus potentially increasing the risk of off-target adverse effects. However, the majority of proteins remains difficult to target, with an estimated 80–85% of the human proteome considered undruggable due to the lack of suitable binding pockets and appropriate chemical matter [154,155]. TPD has the potential to target the undruggable proteome that limits current drug discovery efforts, as only a binder is required to recruit the target protein for degradation rather than high affinity inhibitors. This transition from inhibiting a target protein to promoting its elimination enables the therapeutic modulation of proteins that were previously regarded as ‘undruggable’ [156,157]. Furthermore, eliminating ‘druggable’ targets through induced protein degradation can act synergistically with conventional inhibitor-based therapies. This strategy not only reduces the pool of active proteins requiring inhibition but also mitigates compensatory overexpression commonly triggered by loss of protein function [151].

Induced protein degradation is not a novel concept. As early as 1999, inhibitors of the molecular chaperone heat shock protein 90 (HSP90), a key factor in ensuring proper protein folding, entered

clinical trials. The initial enthusiasm for these inhibitors stemmed from the observation that cancer cells upregulate HSP90 to sustain proliferation and survival pathways [158], rendering them particularly vulnerable to HSP90 inhibition due to the protein's hyperactive conformation in malignant cells [159]. HSP90 inhibitors have been hindered by poor *in vivo* pharmacological properties and severe hepatotoxicity. Unlike the broad protein depletion induced by HSP90 inhibitors, selective estrogen receptor degraders (SERDs) and selective androgen receptor degraders (SARDs) display high specificity towards their respective targets. Notably, these compounds were initially designed as modulators of receptor function but were later unexpectedly found to promote targeted protein degradation [160]. Although the efficacy of these compounds has been impressive, this strategy is not generalizable to other targets: the compounds can only induce the degradation of the estrogen and androgen receptors.

More recently, platform technologies were developed including hetero-bifunctional molecules such as proteolysis-targeting chimeras (PROTACs) [161-164] and Hydrophobic tagging (HyT) molecules [165,166]. Based on the concept of 'event-driven' pharmacology, these techniques combine the modularity of nucleic acid-based strategies with the *in vivo* pharmacology of small-molecule therapeutics (**Figure 10**).

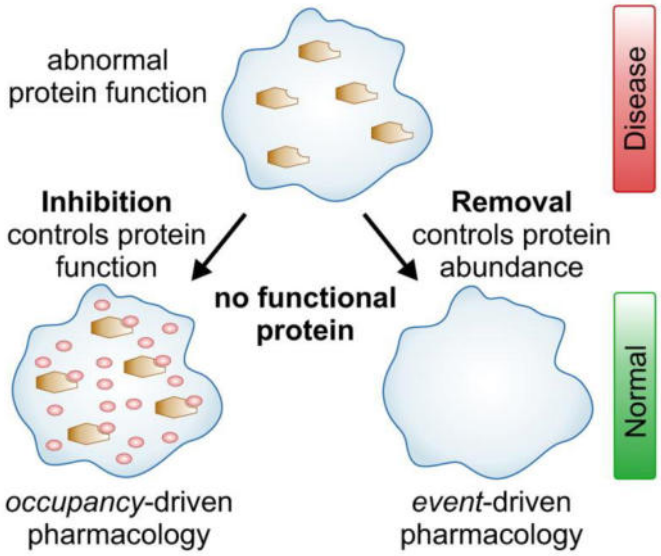


Figure 10. The *event*-driven pharmacology approach compared to the *occupancy*-driven pharmacology approach. The disease-implicated protein is displayed in dark yellow, the applied inhibitor in red [161].

1.5.1. - Insights into TPDs strategies

Native protein degradation occurs via one of two main mechanisms in the cell: the ubiquitin-proteasome system (UPS) and the autophagy-lysosome pathway. Proteins are marked for degradation by a covalent post-translational modification with the 76 amino acid protein ubiquitin [167,168]. Covalent attachment of ubiquitin to the substrate follows a three-step mechanism involving E1, E2 and E3 enzymes [169].

The ubiquitin-activating enzyme (E1) initiates the process by forming a high-energy thioester bond between the Glycine residue of the ubiquitin C-terminal domain and its own Cystine residue by using ATP. The activated ubiquitin is then transferred to a ubiquitin-conjugating enzyme E2. In the presence of a ubiquitin ligase E3, the ubiquitin molecule is transferred from E2 to the target ε-NH_2 -Lysine residue forming an isopeptide bond (**Figure 11**). Mono-ubiquitination can serve as a signal for various non-proteolytic cellular processes, including endocytosis, histone regulation, DNA repair, viral budding, and nuclear export. Subsequent sequential ubiquitination of the conjugated ubiquitin results in the formation of a polyubiquitin chain. Ubiquitin contains seven lysine (Lys) residues for polyubiquitin chain formation (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, or Lys63). While Lys48 and Lys11 linkages mediate proteasomal degradation, Lys63-linked ubiquitin chains are prone to degradation via autophagy [170-173]. While the 26S proteasome efficiently degrades short lived, soluble unfolded or misfolded proteins and polypeptides [174,175], the autophagy-lysosome system is responsible for elimination of long-lived proteins, insoluble protein aggregates and dysfunctional organelles such as degenerated mitochondria [176,177]. Lysosomes mediate the intracellular degradation of proteins and organelles in three different ways: endocytosis, phagocytosis, or autophagy. In contrast with proteasome-based TPD, which can only degrade certain intracellular proteins, lysosome-based TPDs have potential to remove proteins aggregates, damaged excess organelles, membrane and extracellular proteins. With the intensive research in the endosome-lysosome and autophagosome-lysosome degradation pathways, TPD strategies via the lysosomal pathway, such as LYsosome-TArgeting Chimeras (LYTAC), Antibody-based PROTAC (AbTAC), Autophagosome Tethering Compound (ATTEC), AUtophagy-TArgeting Chimera (AUTAC), Bispecific aptamer chimeras, and AUTOphagy-TArgeting Chimera (AUTOTAC) have emerged in recent years [178-180].

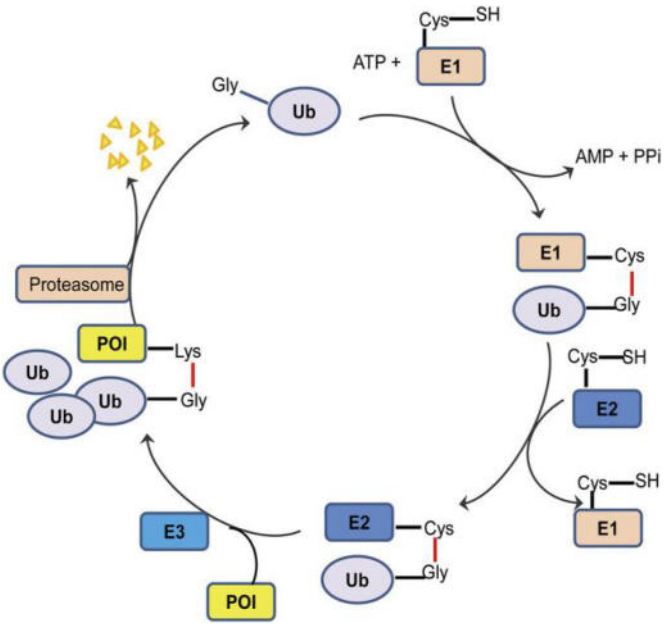


Figure 11. Schematic diagram of ubiquitination process [181].

Proteins, which are naturally degraded by the UPS, are displaying specific degradation signals, termed degrons, which represent E3 ligase binding motifs [175,182,183]. Degrons can range from a single amino acid to short peptide sequences or domains and are transferable between proteins. Therefore, attaching a degron to any other protein induces its degradation. Various constitutive and conditional degrons, activatable by small molecules, light, or temperature, have been successfully employed to trigger selective target protein degradation [183,184]. This kind of approach belongs to the Degradation of Tagged Target Proteins method, which allows target validation and target discovery.

A protein tag comprises a peptide sequence or protein genetically fused to the protein of interest (POI) to allow its selective detection and/or purification. The tag allows highly selective recognition of the POI without the need to identify a suitable binder like a small molecule, peptide or antibody. A variety of tags, ranging from short amino acid sequences to protein domains are available to be introduced at the N- or C-terminus of the POI [139]. Although it is not necessary to identify a binder for the target protein, the use of tagged proteins involves the prior modification of the genome of the target cell/organism. A very powerful tool to quickly achieve TPD for a wide range of target proteins and without the need to identify a specific binder comes from the use of tag specific degraders. In this case, the degrader molecules are optimized to engage the protein tag resulting in effective removal of the tag-POI fusion construct. The two main tags which are widely used as chemical biology tools for this technology are the HaloTag [185] and the dTag system [186].

Briefly, these bifunctional degraders bring the tagged protein as well as an E3 ubiquitin ligase in close spatial proximity inducing ubiquitination and subsequent proteasomal degradation of the tagged protein [161,162].

Differently, the untagged TPD strategies are based on specific binders and hold great potential for future therapeutic applications. Targeting untagged proteins offers the key advantage of avoiding any prior genetic modification of the system. This approach better reflects the natural pathological state, as the target protein remains unaltered in its function, expression, and availability. Moreover, it allows rapid switching between different cell lines, organisms or species without the limitations imposed by protein tagging [139]. All these approaches are functionally active *in vivo* (cell-based and, in some cases, animal) methods because they rely on the cell's endogenous degradation machinery.

Among untagged TPD strategies (**Figure 12**), a method termed Hydrophobic tagging (HyT) takes advantage of a process called unfolded protein response (UPR) by exposing hydrophobic patches on the surface of target proteins thereby triggering their degradation [187]. A hydrophobic group, such as an adamantyl group, is covalently attached to the protein ligand and presented on the target protein surface. The chemical group may mimic a partially unfolded protein domain, promoting UPR [161]. Another untagged TPD strategy involves molecular glues, which are monovalent degraders that induce the interaction between a ubiquitin ligase and a POI, leading to POI ubiquitination and subsequent degradation [188]. They can regulate a variety of biological processes, such as transcription, chromatin regulation, protein folding, localization, and degradation. Also, PROTACs employ the UPS for protein degradation, but in contrast to molecular glues, they are heterobifunctional degraders that simultaneously interact with the E3 ligase and the POI. Then, molecular glues do not have a linker, making them smaller molecular weight, increased oral bioavailability, and improved cellular permeability, compared to PROTACs. Examples of molecular glue degraders include thalidomide, lenalidomide, and pomalidomide [189], approved by the FDA for the treatment of various types of tumors long before their functional mechanisms were elucidated. Over the years, it was elucidated that this class of compounds mediates its antitumor effects through a molecular glue mechanism [190].

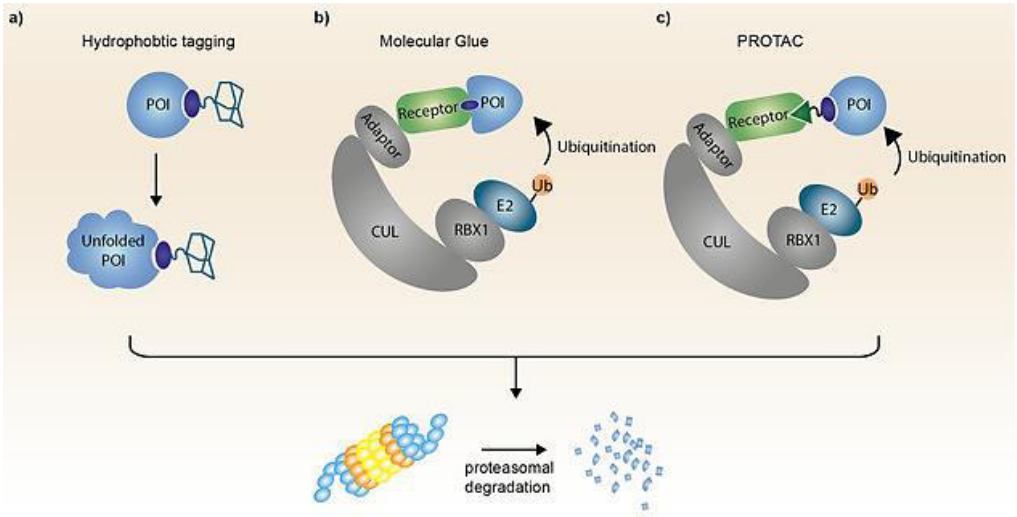


Figure 12. Schematic representation of Hydrophobic tagging (HyT) (a), Molecular glue (b) and PROTAC (c) mode of action [139]. The blue circle represents the molecule targeting the POI (Protein of Interest) in the three represented cases. Ub stands for ubiquitin.

1.6. - The PROTAC approach

The last few years have seen explosive growth in the field of TPD. Substantial progress has recently been made, in particular with the development of PROteolysis TArgeting Chimera (PROTAC) compounds [191,192].

In 2001, the concept of PROTAC was first introduced in the Crews lab at Yale University [193]. PROTACs consist of two ligands, one that recruits an E3 ligase, and the other for recruiting a target POI. The two ligands are joined by a linker that allows formation of a ternary complex between the ligase, the POI, and the PROTAC molecule (**Figure 13**). The ternary complex facilitates polyubiquitination of the POI, resulting in its ultimate degradation via the proteasome [194,195].

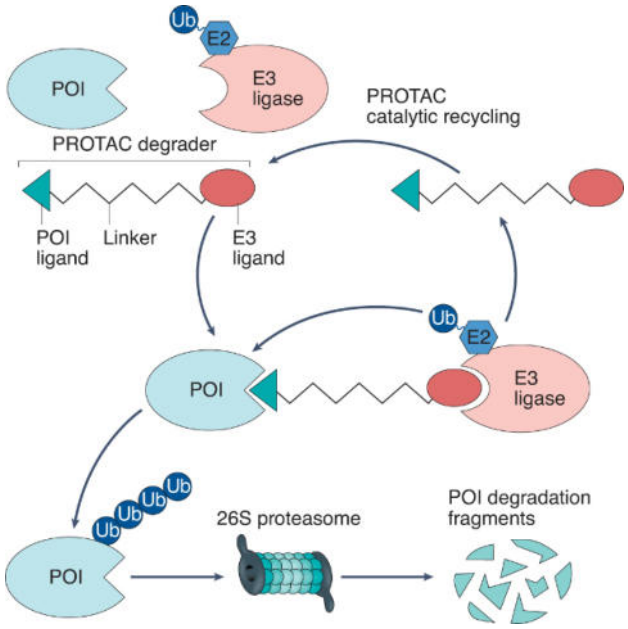


Figure 13. Schematic representation of PROTAC mechanism [196].

The first PROTAC was designed and synthesized to achieve the degradation of methionine aminopeptidase-2 (MetAP-2), consisting of a POI-binding warhead, the antiangiogenic drug Ovalicin, a linker, and an IκBα phosphorylated peptide that recruits the SCFβ-TRCP E3 ubiquitin ligase. This type of peptide-containing PROTAC is now called bioPROTAC, and its application is limited mainly due to the peptide's susceptibility to hydrolysis [197]. In 2012, Ciulli and Crews designed and synthesized a small-molecule analogue of the (hypoxia-inducible factor 1-alpha) HIF-1α recognition peptide [198], which makes the PROTAC based on small molecule structures a reality and opens the door for PROTAC to become an orally bioavailable and thus easy-to-use drug candidate. Since 2013, pharmaceutical giants such as Merck, Novartis, and Pfizer have laid out their PROTAC pipeline, and companies such as Arvinas and C4 Therapeutics have also been established and focused on PROTAC development. In particular, Arvinas, with ARV-110 for metastatic castration-resistant prostate cancer [199] and ARV-471 for ER+/HER2− breast cancer [200], first provided clinical proof-of-concept for PROTAC in mature tumor tar gets, demonstrating the ability to degrade target proteins in vivo and providing clinical benefit to patients.

Currently marketed degrader drugs, such as the IMiD class, along with the vast majority of newly developed PROTACs and molecular glue compounds in clinical trials, primarily target various cancers [192,201]. Nevertheless, due to its ability to induce degradation of virtually any protein of interest, the scope of TPD is expanding beyond oncology. [197]. In inflammation, immunity and immuno-oncology recent breakthroughs have been taking place. BTK degrader (NX-5948) and

IRAK4 degrader (KT-474), from Nurix Therapeutics and Kymera Therapeutics, respectively, could treat patients with various immuno-inflammatory diseases, such as rheumatoid arthritis. BTK is a well-established target in inflammation, as well as in oncology, and has been the subject of intense PROTAC molecule experimentation across academia and industry [202-204], particularly because of the emergence of the C481S resistance mutation in BTK that renders first-generation inhibitors less effective [205,206].

A key PROTAC property is the ability to degrade proteins that are not classically targetable by small-molecule inhibitors due to the absence of conventional active sites. This attribute makes targets for several neurodegenerative diseases, a group of disorders characterized by progressive motor or cognitive impairment [207] including Alzheimer's disease (AD), Parkinson's disease (PD), and Huntington disease (HD), are closely associated with insoluble aggregates formed by protein misfolding [208]. Tauopathies are an example of neurological diseases associated with toxic accumulation of aberrant tau species, leading tau to aggregate into paired helical filaments and ultimately into neurofibrillary tangles that result in neuronal cell death [191,209].

Interestingly, PROTACs are emerging as potential antiviral agents [191]. The feasibility of antiviral PROTACs was first established with the degradation of the NS3/4A protease of hepatitis C virus (HCV) in cell-based assays by a CRBN-linked HCV inhibitor but was brought to the forefront with the emergence of SARS-CoV-2 [210]. Potential target for the PROTAC approach against SARS-CoV-2 include the viral envelope protein [211], but also catalytic virulence factors, such as its two proteases, $3CL^{Pro}$ and PL^{pro} [212]. In particular, several PROTACs have been recently described to target coronaviral $3CL^{Pro}$ [135,213-215].

While PROTACs represent an important modality for certain target classes, some POIs lack the small-molecule binding sites needed for this approach. Even so, these targets are still susceptible to degradation by alternative PROTAC approaches, including bioPROTACs and hybrid PROTACs. Three PROTAC categories can be defined when considering small molecules and peptides [191]:

- PROTACs composed of traditional small-molecule ligands;

- bioPROTACs composed of peptide ligands;

- hybrid PROTACs containing both a peptide and a traditional small-molecule warhead.

PROTACs can be developed as oral therapeutics, but bioPROTACs and hybrid PROTACs are unlikely to be oral drugs owing to their peptide components.

1.6.1. - Features of PROTACs

PROTACs differ from small molecule inhibitors in that they act at sub-stoichiometric concentrations, de facto behaving as a catalyst rather than a true inhibitor and/or antagonist and degrade target proteins rather than just modulate their activity [216]. In particular, small-molecule PROTACs have the potential for oral delivery, and cellular levels can be controlled by dosing. Unlike conventional inhibitors that exert their activity through occupancy-driven pharmacology, PROTACs are examples of event-driven pharmacological agents. In this case, the event is ternary complex formation and ubiquitination of the POI, which results in proteasomal removal of the POI from within the cell and subsequent effects that may last well beyond the time of treatment [216]. Since only a transient binding event is required for activity, and thus, in contrast to the stoichiometric site occupancy of the occupancy-driven model, these event-driven molecules can cycle through multiple rounds of PROTAC activity, requiring relatively infrequent dosing regimens.

PROTACs are also distinct from traditional inhibitors because they do not rely on blocking enzyme function, which commonly limits drug surface area at the active site and compromises selectivity when targeting an enzyme that is a member of a large enzyme family, such as a protein kinase [216]. It is worth mentioning that a high affinity or covalent binding between the POI and the warhead is not required for a PROTAC molecule. Therefore, PROTACs have the potential to target proteins at any site on their surface, thus opening opportunities for expanding the range of targets to proteins previously thought to be undruggable due to the absence of well-defined small molecule binding pockets. In this regard, PROTACs can target proteins with common elusive characteristics, including a change from the natural state via mutation, aggregation, isoform expression or localization, that result in a disease-causing gain of function. If a binding surface that is approachable by an E3 ligase and, ideally, an unstructured region to thread into the proteasome are present the protein can be degraded [217,218]. Proteins that have evolved resistance mutations to targeted therapies, proteins that may not have an enzymatic function of their own but act as protein-protein interaction cores to recruit and direct signaling complexes (scaffolding), and proteins considered ‘undruggable’ with other modalities may be highly suitable as PROTAC targets [219].

Although ‘PROTACable’ POIs do not need an enzyme active site, they do need a small-molecule binding site that is approachable by an E3 ligase. Using these sites does not require a high-affinity ligand if coupled to the right E3 ligand, but moderate affinity (≥1–500 nM) is typically needed, and access to the POI surface near the binding site by a recruited E3 ligase is essential [220].

Multiple studies have shown that PROTACs maintain, and in some cases even enhance, the binding selectivity of their protein ligands. Consequently, binding to off-target proteins does not automatically result in their degradation. This provides an additional layer of specificity, which can be further tuned by the choice of recruited E3 ligase [221-223]. Additionally, it is expected that the resistance acquired to drugs will be lower than that associated with the use of small molecule drugs. The catalytic mode of action renders PROTACs particularly effective in countering drug resistance due to overexpression of the POI. PROTACs can also be effective in drug resistance cases caused by point mutations and by scaffolding functions of the target proteins [224].

In general, PROTAC physicochemical property space falls beyond the 'rule of 5' [225,226], owing to the bifunctional nature of the molecules. The rule of 5 was developed in the 1990s as a guideline for assembling high-throughput screening (HTS) libraries. PROTACs typically have molecular weights above 500 Da, tend to contain a greater number of hydrogen bond donors (HBDs) and acceptors (HBAs) [227]. Despite breaking the 'rule of 5', several PROTACs have been reported to exhibit good oral bioavailability [228,229]. This poses challenges to the physicochemical drug-like space of PROTACs, such as ADMET properties including solubility, cell permeability, metabolic stability, and absorption, thus calling into question oral bioavailability. An optimal PROTAC target is typically characterized by the following features: (a) it undergoes pathogenic gain-of-function changes, such as overexpression, mutation, or mislocalization; (b) it possesses a binding pocket for the warhead; (c) it has a surface site accessible for ubiquitination by an E3 ligase; and (d) it ideally contains a flexible region capable of entering the proteasome's barrel cavity [191,230].

A key consideration in linker design is the choice of attachment sites on both the warhead and the E3 ligase ligand. Ideally, the linker should extend from either the ligand-binding pocket or a solvent-exposed region, minimizing any impact on the interaction between the small molecule and its target protein. The linker can contribute to the PROTAC solution conformation [231] and degradation efficiency [232] and influence the E3 ligase–POI interactions and presentation of the POI within the zone of ubiquitylation. PEG and alkyl chains are the most common linker motifs.

The main challenge in linker design lies in the inability to predict, a priori, the optimal length or conformation needed for effective ubiquitination of the POI by the recruited E3 ligase. A linker that is too short may favor the formation of POI–linker or E3 ligase–linker binary complexes rather than the desired POI–linker–E3 ligase ternary complex. Conversely, excessively long linkers can compromise ternary complex stability and reduce the PROTAC's ability to efficiently penetrate cell membranes [197].

Finding a suitable E3 ubiquitin ligase ligand has become a major difficulty limiting the application of E3 ubiquitin ligases in PROTAC. There are more than 600 E3 ubiquitin ligases in the human body [162]. Researchers have focused on E3 ubiquitin ligases that are highly expressed in tumors, aiming to exploit their tissue specificity to enhance the antitumor efficacy of PROTACs while minimizing toxicity in healthy tissues [219]. In this regard, VHL and CRBN are the two most frequently redirected E3 ligases (**Figure 14**). Functionally, VHL is responsible for regulating intracellular levels of the hypoxia-inducible factor alpha subunits (HIFα), an important transcription factor that controls how cells sense oxygen and respond to low oxygen levels (hypoxia) [233]. VHL forms part of the Cullin2–EloBC–RBX1 E3 ligase complex, also referred to as CRL2VHL and functions as the substrate (HIFα) recognition subunit [234]. CRBN forms part of Cullin 4A–DDB1–RBX1 (CRL4CRBN) assembly [190] and plays an incompletely understood role in DNA damage response [235].

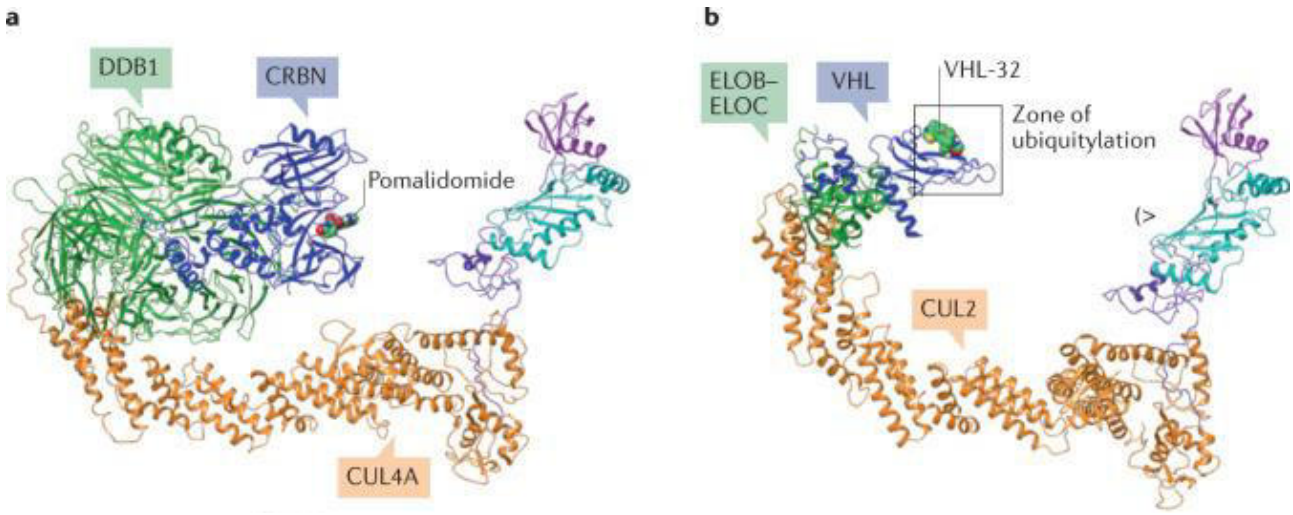


Figure 14. CRBN (a) and VHL (b) E3 ubiquitin ligases and their substrate adaptors shown with their corresponding colours (adapted from [191]). In violet and in blue the representation of RBX1 and the E2 conjugating enzyme.

Expanding the "toolbox" of E3 ligases that can be recruited by PROTACs is important for several reasons, including expanding the chemical space of bifunctional PROTAC degraders, overcoming resistance mechanisms, off-target activity, and limitations due to tissue- and cell-type dependence of E3 ligase expression and distribution [236]. Examples of other E3 ligases include DDB1-CUL4-associated factor (DCAF) family member such as DCAF15, which through interaction with sulfonamides is responsible for the degradation of the pre-mRNA splicing factor RNA-binding motif protein 39 (RBM39) in some solid tumours and leukaemias [237]. Other investigated E3

ligases comprise DCAF16, ring finger protein 4 (RNF4) [238], RNF114 [239], KEAP1 [240] and fem-1 homologue B (FEM1B) [241] able to degrade the prototypical TPD substrate bromodomain-containing protein 4 BRD4 in human breast cancer cells.

1.6.2. - Investigation of the ternary complex

The main function of the PROTACs is to bring a target protein into spatial proximity with an E3 ubiquitin ligase to achieve degradation through "hijacking" the cell's UPS by bringing together the target protein and the E3 ligase (**Figure 15**).

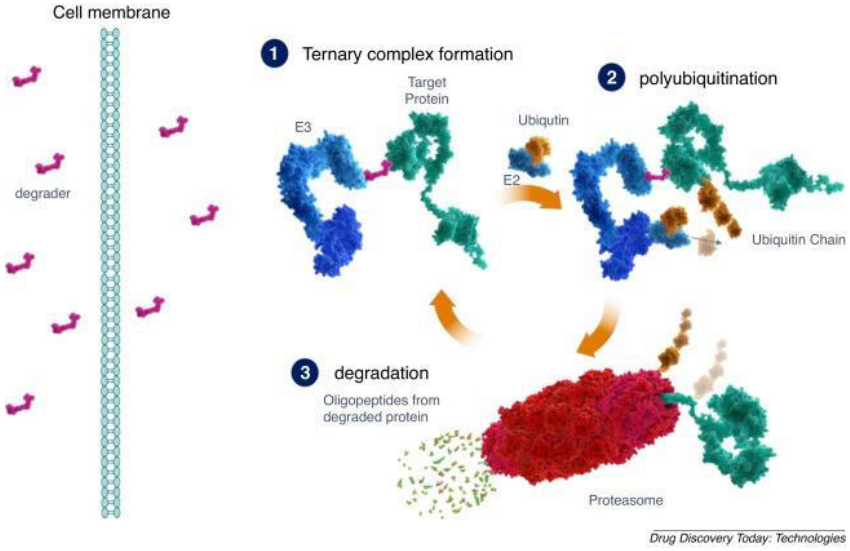


Figure 15. Catalytic cycle of the targeted protein degradation [242].

An important aspect of PROTAC validation and optimization is measuring their affinity for the E3 ligase, their affinity for the POI, and their ability to mediate ternary complex formation [216]. In general, there are numerous biophysical and cell biology methods to quantify a binding event between a small molecule and a biomolecule, and many of these techniques can be applied to measure the binary target engagement of PROTAC molecules or in other terms the independent binding of the PROTAC to either the E3 ligase or to the POI [243]. Label-free ligand binding assays such as fluorescence polarization (FP) and fluorescence energy transfer (FRET) are popular assays for measuring inhibition or dissociation constants IC_{50} and K_d of PROTACs. Isothermal titration calorimetry (ITC) is a well-established method for measurement of binding affinities that can also provide information about binding stoichiometry and is often used in PROTAC characterization.

Surface plasmon resonance (SPR) is another label-free technique that can be used to measure not only the binding affinity but kinetic information about interactions.

However, accumulated evidence suggests that binary target engagement alone is not sufficient to predict the outcome of protein degradation [244]. It is therefore essential to examine ternary complex formation that follows the binary binding step. Ternary complex formation is influenced by several key parameters, including the concentration of the target and E3 ligase, the intracellular concentration of the degrader, the affinity that the degrader has for each protein, and protein-protein interactions between the target protein and the recruited E3 ligase (defined as cooperativity). This parameter reflects the energy difference before and after ternary complex formation, is represented by a scaling factor called α, defined mathematically as K_d(binary)/K_d(ternary). If the degrader facilitates novel interactions between the E3 ligase and target protein, the affinity of the ternary complex will be higher than the binary interactions between the degrader and each protein, meaning $\alpha>1$. If the degrader does not stabilize interactions between the E3 ligase and the target protein, the probability of the ternary complex formation remains the same as in the binary complexes and, as a result, α is equal to 1. When ternary complex formation is unfavorable due to steric clashes or electrostatic repulsions induced by the degrader bringing the two proteins into close proximity, a weaker binding affinity will be observed for the ternary complex compared to the binary complex, resulting in an α value below 1 [242]. The value of α can be determined through techniques like ITC [245,246], TR-FRET [203,247] and SPR [248]. However, it is not clear whether there is a positive correlation between cooperativity and potency of the degrader-mediated protein degradation. Degraders with low cooperativity ($\alpha<1$) may still induce effective degradation. For example, a series of CRBN-BRD4 degraders that showed negative cooperativity ($\alpha<1$) led to effective degradation [247]. Therefore, high degradation potency can be achieved, even for degraders with low cooperativity, suggesting that other factors, such as permeability and ubiquitination, could also heavily contribute to the cellular potency [249].

1.6.3. - Cereblon E3 ubiquitin ligase

CRBN, first identified as a genetic cause of inherited autosomal recessive mental retardation, is one of the substrate receptors of the cullin 4 (Cul4)-RING E3 ubiquitin ligase (CRL4), which is the complex responsible for the poly-ubiquitination of the substrate [195]. CRL4 ubiquitin ligases are formed by a cullin protein (Cul4), which acts as an assembly factor that provides a scaffold for assembly of a RING box–domain protein (RBX1) and the adaptor protein damaged DNA–binding

protein 1 (DDB1) [250]. RBX1 is the docking site for the activated E2 protein, and DDB1 recruits substrate-specificity receptors or DCAFs (DDB1–cullin 4–associated factors) to form the substrate-presenting side of the CRL4 complex [251,252].

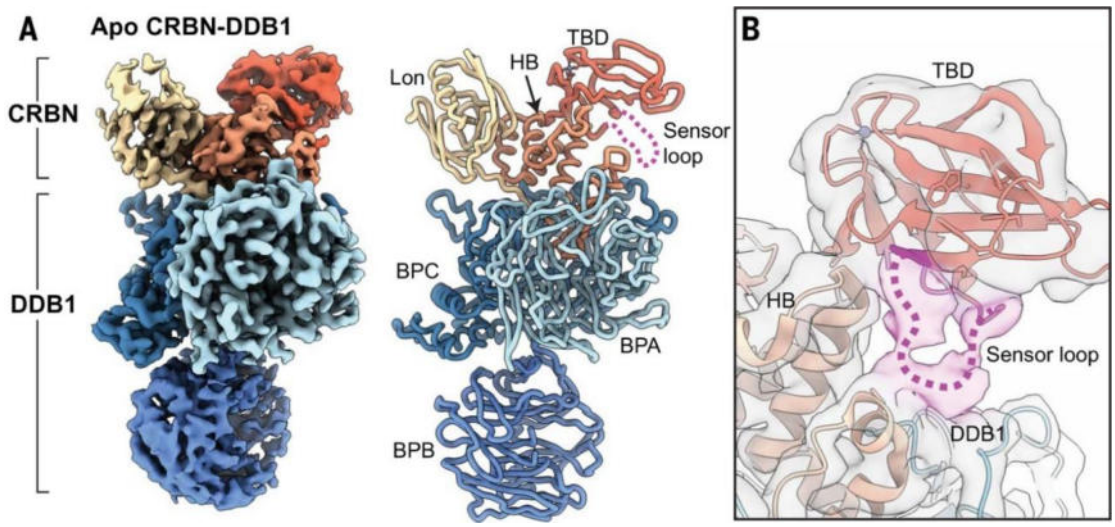


Figure 16. ~3.5-Å-resolution cryo-EM reconstruction of CRBN-DDB1 in the unliganded apo form (**A**) and cryo-EM map enlargement of the sensor loop, which interacts with the HB and BPA of DDB1 (**B**) [253].

CRBN is retained as the direct target of thalidomide and other IMiD drugs [190,254]. Thalidomide, α-(N-phthalimido) glutarimide, was initially used in pregnant women as a sedative drug. However, it was discovered to cause severe teratogenicity in many tissues and organs, including limb teratogenicity represented as amelia and phocomelia in infants [255,256]. Since it has been demonstrated that thalidomide and its derivatives, lenalidomide and pomalidomide, have immunomodulatory activity and anti-proliferative effects on several hematological cancers, they are widely applied for the therapy of multiple myeloma and other hematologic malignancies as IMiD drugs [257,258]. Biochemical and genetic evidence has shown that CRBN is a DCAF for CRL4 [190,250] and that binding of IMiD compounds to CRBN affects the ubiquitin E3 ligase activity of CRBN bound CRL4 (CRL4CRBN), thus mediating the antiproliferative effects on multiple myeloma (MM) cells and the immunomodulatory effects on T cells [190,254,259]. It was reported that the transcription factors Ikaros and Aiolos are substrates of the CRL4CRBN–IMiD drug complex, thus explaining many of the therapeutic effects of IMiD compounds on immune and tumor cells [260-262].

Cereblon proteins occur throughout eukaryotes, however not in fungi. They are typically 400–600 residues long (442 in the case of human CRBN) and their genes occur in single copy per genome [263]. Human CRBN is a 50-kDa protein containing three folded domains: an N-terminal Lon protease like domain (Lon domain), an intermedial helical bundle (HB), and a C-terminal thalidomide-binding domain (TBD) that harbors the drug-binding pocket. CRBN is recruited to the Cullin-4–ROC1 ligase module by the adaptor protein DDB1, which consists of three WD-40 beta propeller domains (BPA, BPB, and BPC) [253] (**Figure 16**, panel **A**). The N-terminal region including the Lon domain and HB resembles the N-terminal domain of bacterial Lon proteases, which led to its initial annotations as an ATP-dependent Lon protease [264]. The HB of CRBN docks into a central hydrophobic cleft formed at the interface of BPA and BPC of DDB1, while the mobile BPB domain interacts with Cullin-4, positioning CRBN-bound substrates for ubiquitination. The TBD is located on the face of CRBN that is oriented away from the DDB1-binding site. The TBD is composed of a six-stranded antiparallel β-sheet core, with a structural zinc ion coordinated by four cysteine residues (Cys323, Cys326, Cys391 and Cys394). The IMiD-compound pocket is formed by three tryptophan residues, Trp380, Trp386 and Trp400, with a phenyl alanine residue at the base (Phe402) [251]. These residues form a small hydrophobic pocket (tri-Trp pocket) in which the glutarimide portion of IMiD-compounds is accommodated. A combination of mutations at both Tyr384 and Trp386 has been reported to cause a loss of the IMiD-compound effects. A beta-insert hairpin within CRBN's TBD (residues ~341 to 361) has been shown to directly bind immunomodulatory drugs and neosubstrates in prior structures [251,265], at its canonical position adjacent to the thalidomide-binding pocket (**Figure 16**, panel **B**).

Although the CRBN-TBD shares its fold with proteins such as methionine sulfoxide reductase, dsRNA-binding proteins (RIG-I, LGP2, MDA5), and MSS4, the tri-Trp-pocket residues are not conserved, preventing IMiD binding. In contrast, CRBN orthologs across animals and plants retain complete conservation of this pocket [251].

1.7. - Nsp13 as another pharmacological target

As mentioned before, the genome of SARS-CoV-2 encodes structural proteins as well as 16 non-structural proteins (Nsp1-16) [266] that collectively form the machinery for viral replication and transcription. The RNA-dependent RNA polymerase (RdRp) consists of a catalytic subunit, Nsp12, and two auxiliary factors, Nsp7 and Nsp8, and is required to replicate the viral genome [267,268]. In addition to its central role in replication and transcription of the viral genome, the

RdRp contains an N-terminal nidovirus RdRp-associated nucleotidyl-transferase (NiRAN) domain required for the formation of the cap core structure, which consists of a guanosine residue linked to the first transcribed nucleotide by a 5′-5′ triphosphate bond (GpppA) [269]. The holo-RdRp is known to associate with multiple accessory factors during the viral life cycle. One of these factors is Nsp13, a 67-kDa member of the helicase superfamily 1B, where the Walker A (phosphate loop) and Walker B (Mg^{2+} binding site) motifs play a fundamental role in the adenosine triphosphate (ATP) binding and catalysis [270], which utilizes the energy of nucleotide triphosphate hydrolysis to catalyze the unwinding of double-stranded DNA or RNA in a 5′ to 3′ direction [271]. Although Nsp13 is believed to act on RNA, in vivo enzymatic characterization shows a significantly more robust activity on DNA in vitro assays with relatively weak non processive helicase activity when compared to other superfamily 1B enzymes [272,273]. However, binding of Nsp13 to the RdRp significantly augments its activity, possibly by means of mechanoregulation [272]. In addition to its helicase activity, Nsp13 also possesses RNA 5′ triphosphatase activity within the same active site [274], suggesting a further essential role for Nsp13 in the formation of the viral 5′ mRNA cap.

Furthermore, Nsp13 is involved in many other molecular mechanisms related to the evasion of the human immune system. Indeed, it antagonizes IFN-I signaling through blocking signal transducer and activator of transcription 1 (STAT1)/STAT2 phosphorylation or nuclear translocation [275], downregulates NF-κB promoter signaling by limiting TBK1 and IRF3 activation, as phospho-TBK1 and phospho-IRF3 protein levels are reduced with increasing levels of Nsp13 protein expression [276,277].

Structurally, Nsp13 contains 5 domains, a N-terminal Zinc binding domain (ZBD) (1 to 100), that coordinates 3 structural zinc ions, a helical “stalk” domain (101 to 150), a beta-barrel 1B domain (151 to 261) and two “RecA like” helicase subdomains 1A (262 to 442) and 2A (442 to 596) delimiting the central channel for RNA unwinding [278], characteristic of SF1 helicases [279,280]. Domains 1A, 2A, 1B and stalk are essential for helicase activity [281]. The ZBD packs against the three-helix bundle of the “stalk” domain which also makes contacts with 1A domain on its opposite side (**Figure 17**).

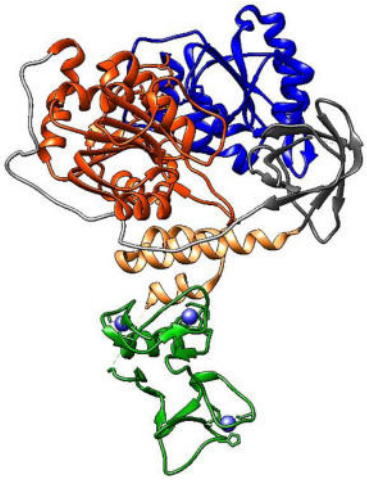


Figure 17. Crystal structure of SARS-CoV-2 NSP13 (PDB ID 7NIO), showing the zinc-binding domain (ZBD; green), stalk domain (wheat), 1B domain (grey), RecA1 domain (orange), RecA2 domain (blue).

This same basic 5-domain architecture is shared by other Nidovirus helicases such as the Nsp10 proteins from Equine arteritis virus [282] and Porcine reproductive and respiratory syndrome virus [283] and to a lesser extent the human nonsense-mediated mRNA decay factor UPF117.

Previous X-ray structures of Nsp13 have been solved for MERS-CoV and the highly related SARS-CoV to 3.0 Å and 2.8Å respectively [284,285]. More recently, cryo-electron microscopy studies have revealed the architecture of the Nsp13 containing replication and transcription complex, which contains two copies of Nsp13 that interact with Nsp8 via the N-terminal ZBD [269,286,287]. One of the Nsp13 protomers makes additional interactions with Nsp12 and is located with its RNA binding site in the path of downstream RNA [286]. However, the opposing polarities of helicase translocation (5′ to 3′) and polymerase activity suggest that Nsp13 may facilitate backtracking, template switching, or disruption of downstream secondary structures. [286].

The Nsp13 helicase exhibits a remarkable sequence conservation across the Coronaviridae family especially in the residues delimiting the central channel. 26 residues of the channel are conserved not only in SARS-CoV, SARS-CoV-2 and MERS human coronavirus, but also in many non-human beta as well as in alphacoronavirus with only two exceptions involving the point mutations R178K and A316S [278].

Notably, an X-ray crystallography fragment screen identified prominent druggable binding sites within the nucleotide and RNA-binding pockets, as well as in the cleft between the ZBD and the stalk domain. [278]. This study highlighted that the residues lining the nucleotide and RNA-binding pockets are highly conserved across α- and β-coronavirus helicases, with the 5′ RNA-binding site

in Nsp13 exhibiting a level of conservation comparable to the RdRp active site and the ADP-binding pocket of the NiRAN domain in Nsp12 [278].

The SARS-CoV-2 Nsp13 protein represents an ideal target for development of anti-CoV drugs because of its sequence conservation and indispensable role across all CoV species [273,288]. The primary amino acid sequences of SARS-CoV Nsp13 and SARS-CoV-2 Nsp13 share a sequence identity of 99.8% and a sequence similarity of 100.0% [289]. This implies that previously found therapeutics that target this enzyme in other CoVs species might be effective against the new COVID-19 disease.

Inhibition of Nsp13 can be achieved through various strategies, including ATP binding, direct nucleoside triphosphatase (NTPase) activity, nucleic acid binding to the helicase or blocking helicase translocation. Several classes of compounds have been reported as Nsp13 inhibitors, including benzotriazoles, imidazoles, imidazodiazepines, phenothiazines, quinolines, triphenylmethanes, pyrroles, acridones, small peptides, and bananin derivatives, all representing potential therapeutic agents against SARS-CoV-2 [290]. In response to the COVID-19 pandemic, natural products have gained increased attention as potential lead compounds against SARS-CoV-2 [291]. Among these, myricetin, a flavonoid that is abundant in fruits, vegetables and teas, has emerged as a particularly promising lead compound. The crystal structure of myricetin-bound Nsp13 was obtained revealing that the inhibitor binds to the surface of the RecA1 domain near its interface with the stalk domain [292] (**Figure 18**). Caffeic acid derivatives rosmarinic acid and chlorogenic acid, identified as potential inhibitors of Nsp13, along with myricetin, were tested in ATPase and unwinding assays. Myricetin inhibited both activities of Nsp13, consistent with previous studies [293-295]. In contrast, rosmarinic acid and chlorogenic acid inhibited only the unwinding activity of Nsp13.

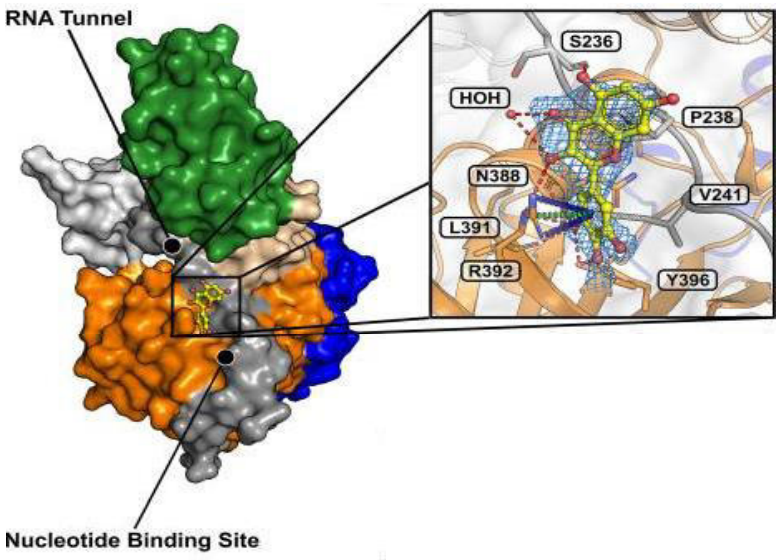


Figure 18. Surface representation of NSP13, colour-coded as in Figure 17, with myricetin displayed as a yellow ball-and-stick model. The positions of the nucleotide-binding site and RNA-binding channel are indicated. The inset shows the binding site with interacting residues (sticks), a water molecule (sphere), hydrogen bonds (red dashed lines), carbon–π interactions (white dashed lines), cation–π interactions (green dashed lines) and donor–π interactions (blue dashed lines) (adapted from [292]).

1.8. - Thesis objectives

Viral proteins retained as important drug target represent the object of research of my PhD period. In particular, my PhD project focuses on the structural characterization of protein complexes between selected pharmacological targets and small molecules. These investigations will serve as a starting point for validating specific proteins or domains as drug targets with the final goal to develop therapies with the desired results. The proteins under analysis are the main viral proteases SARS-CoV-2 $3CL^{Pro}$ and CVB3 $3C^{Pro}$ and the viral helicase SARS-CoV-2 Nsp13.

SARS-CoV-2 $3CL^{Pro}$ exhibits a high structural conservation of the main core among the various Coronaviruses [296]. Therefore, the discovery of new antiviral compounds that are active against a broad spectrum of Coronaviruses could be relevant [124]. The $3CL^{Pro}$ catalytic dyad suggests a good ability of binding transition metal ions. In section 3.1 gold metallodrugs, in particular Auranofin and five related gold(I) complexes, in complex with SARS CoV-2 $3CL^{Pro}$ are studied as potential inhibitors of the protease activity.

The structural similarity across 3CL proteases in all Coronaviruses genera is conserved as well as across 3C proteases belonging to the Enterovirus genus, in the Picornavirus family [96,97]. The latter proteases are essential for viral replication, and they possess distinctive cleavage site recognition features that are not found in closely related human homologs. With the intention of developing broad-spectrum antiviral drugs, we decided to exploit the PROTAC peptidomimetic approach able to target both the coronaviral $3CL^{Pro}$ and the enteroviral CVB3 $3C^{Pro}$ proteases. Indeed, the PROTAC molecules under investigation - KME16, ACH213, and GC-376 PROTACs - are bifunctional compounds designed to engage the active sites of the SARS-CoV-2 $3CL^{Pro}$ and CVB3 $3C^{Pro}$ proteases on one end (derived from a PROTAC precursor), while recruiting the E3 ubiquitin ligase CRBN on the other (**Figure 19**).

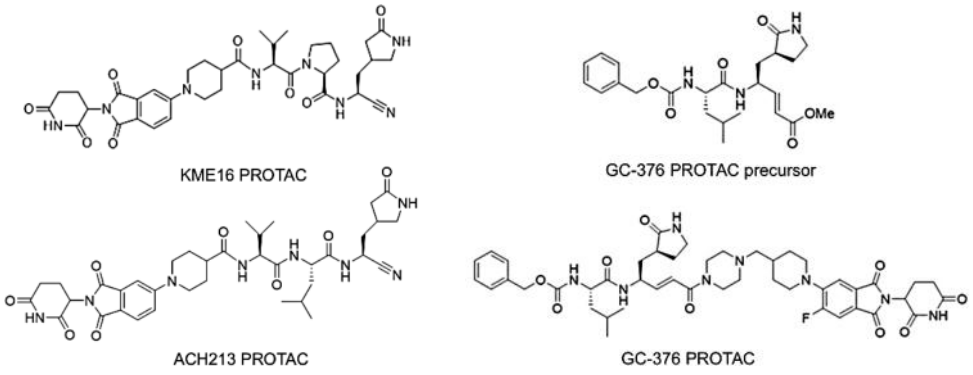


Figure 19. Chemical formula of investigated PROTAC molecules.

The structural biology platform allows us to set up a procedure for the validation of molecular targets for drug discovery. The research goals include the development of successful protocols for the expression of recombinant proteins SARS-CoV-2 $3CL^{Pro}$, CVB3 $3C^{Pro}$ and CRBN E3 ubiquitin ligase, in E. coli as isotopically and not isotopically enriched samples. X-ray crystallography and solution NMR spectroscopy are used to structurally characterize the interaction of PROTACs both with the protein targets and with the CRBN.

Finally, I set up and optimized a protocol for the production and purification of SARS-CoV-2 Nsp13 protein. Activities involving the characterization of Nsp13 entail extensive collaboration with G. Caminati and S. Menichetti laboratory group (DICUS, UNIFI), which are participants in the PNRR project, Spoke 4.

2. Chapter – Methodological aspects

2.1. - Protein production

2.1.1. - Expression systems

Recombinant protein expression has revolutionized all aspects of the biological sciences. In particular, it has greatly expanded the number of proteins that can be studied from both a biochemical and structural perspective. As the number of proteins produced by recombination increases, so does awareness of the difficulties and limitations inherent in this process [297].

The choice of an expression system for the high-level production of recombinant proteins depends on many factors. These include cell growth characteristics, expression levels, intracellular and extracellular expression, post-translational modifications and biological activity of the protein of interest [298]. Many organisms can be used, such as bacteria, yeasts, plants, fungi, insects and mammalian cells, however, Escherichia coli is still the preferred host for recombinant protein expression [299]. The rationale is clear: E. coli is easy to genetically manipulate, it is inexpensive to culture, and expression is fast, with proteins routinely produced in one day. Moreover, protocols for isotope-labelling in E. coli for NMR spectroscopy are well established, making it highly suitable for structural studies [297].

Nevertheless, despite the extensive knowledge of the genetics and molecular biology of E. coli, not every gene can be expressed efficiently in this organism. This may be due to the unique and subtle structural features of the gene sequence, the stability and translational efficiency of mRNA, differences in codon usage between the foreign gene and native E. coli, the easy of protein folding and potential toxicity of the protein to the host. Moreover E. coli lacks the machinery required to perform certain eukaryotic post-translational modifications, such as glycosylation, which can be critical for the formation of folded, active protein and the reducing environment of the bacterial cytoplasm makes the efficient production of disulfide-containing proteins challenging [300,301]. Considerable efforts have been made in recent years to maximize the efficient production of soluble recombinant proteins in bacteria. A remarkable number of aspects such as expression vector design, host strains, promoter strength, co-expression and low temperature induction have been developed to successfully overcome the main limitations [297] (**Figure 20**) .

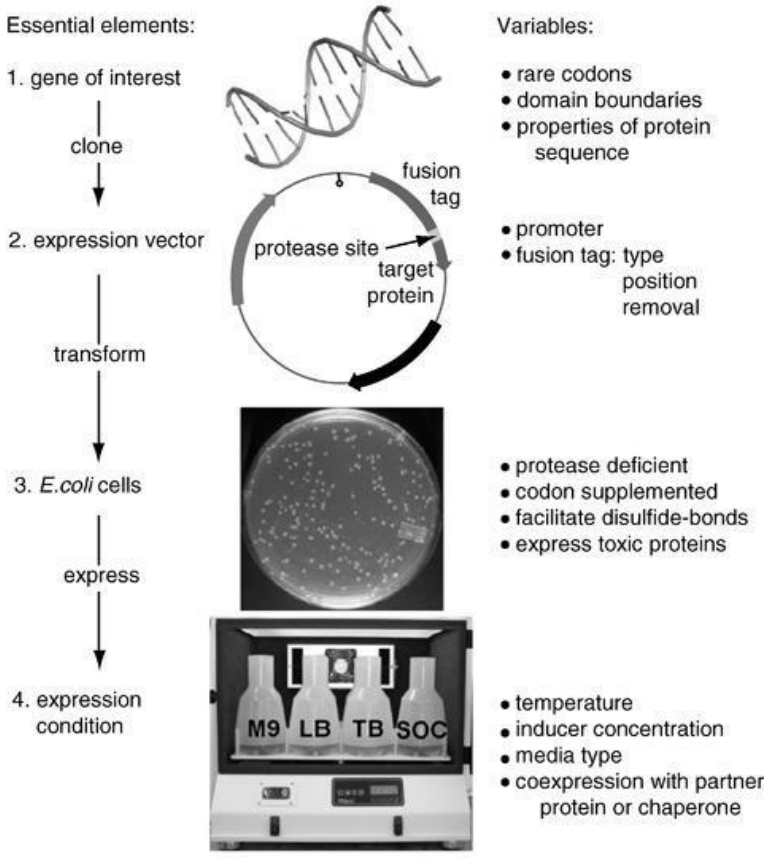


Figure 20. Overview of the protein expression strategies in E. coli [297].

2.1.2. - E. coli protein expression

The aim of recombinant protein expression in E. coli is to produce a protein product that is soluble, folded, and active. A fundamental element is the choice of the expression vector, which contains all DNA sequence elements that direct the transcription and translation of the target gene [302]. These elements include the origin of replication, promoters, and terminators. In addition, expression vectors contain a selection element, typically an antibiotic-resistance gene, to aid in plasmid selection within the host cell. The origin of replication of a vector is the site where replication is initiated. It also determines copy number of the vector in the host. The copy number for common E. coli expression plasmids ranges from low (2 to 20) to high (20 to 40). Typically, high-copy number plasmids are desired for protein expression in E. coli, as they result in the maximum protein yield for a given culture volume [303]. Promoters are another element of the vector that can have a profound effect on the strength and duration of transcription and, in turn, protein yield. An effective promoter for expressing heterologous proteins in E. coli has three characteristic features. Firstly, it is strong, resulting in robust expression of the target gene (typically 10% to 50% of total cellular proteins). Secondly, it exhibits low basal transcriptional

activity to prevent unwanted transcription prior to induction. Thirdly, induction is simple and inexpensive. For maximal protein yields, a strong promoter should be selected, such as T7 or tac promoters [297].

The T7 RNA polymerase system is the most commonly used promoter system in E. coli. Gene expression is driven by the T7 RNA polymerase (from the T7 bacteriophage), which transcribes DNA five times faster than the bacterial RNA polymerase [304]. Because E. coli lacks this enzyme, the polymerase must be delivered to the cell, via an inducible plasmid or, more often, by using an E. coli designated (DE3) strain, that contains a chromosomal copy of the T7 RNA polymerase gene allowing for simple and efficient expression of genes from T7-based expression vectors [302]. In the absence of an inducer, the polymerase, which itself is under the control of the lacUV5 promoter, is not produced, and correspondingly, the gene of interest is not transcribed [304]. Upon addition of the non-hydrolyzable lactose analog, Isopropyl β-D-1-thiogalactopyranoside (IPTG), the T7 RNA polymerase is transcribed and synthesized. The polymerase then initiates transcription of the target gene by binding to a T7 polymerase–specific promoter. Once induced, most of the cellular machinery is devoted to the production of the recombinant protein, comprising up to 50% of the total cellular protein [302,304].

Tac promoter is a hybrid of naturally occurring promoters, consisting of the −35 region of the trp promoter and the −10 region of the lacUV5 promoter [305,306]. The only difference between these two systems is the spacing between the −35 and −10 consensus sequences, with 16 bp separations in the tac promoter [307]. Expression is induced with IPTG. It can result in the accumulation of up to 15% to 30% of the total cellular protein [307].

Another critical feature of E. coli expression vectors is the presence of a fusion tag. The fusion tags are transcribed in-frame with the construct of interest. They are useful because they can improve protein expression, promote folding, increase protein solubility, and facilitate downstream processes such as purification and detection. The placement of the tag, either N-terminal or C-terminal, is also important as it can have a profound effect on soluble protein expression levels. Additionally, the presence of a fusion tag may interfere with the biological activity of the recombinantly expressed protein, and thus, in these cases, it may be important to enzymatically remove the tag after the fusion protein has been purified. The most common tags used in the experimental procedure in this work of thesis are hexahistidine (His6 or His-tag) which facilitate both solubility and purification through the immobilized metal affinity chromatography (IMAC) [308] and Glutathione-S-transferase (GST), a protein that catalyzes the nucleophilic attack of glutathione on electrophilic substrates in order to reduce their reactivity with other biomolecules

[309] and it binds tightly and specifically to glutathione agarose, allowing the target protein to be purified in a single step [310]. All fusion tags have the potential to influence the behavior of the expressed protein, and thus it is typically desirable to remove the tag following purification. This is achieved by engineering a protease-specific cleavage site (a sequence of ~7 amino acids that is specifically recognized by the protease) between the tag and target protein. Following expression and purification, the corresponding protease is used to cleave the tag from the protein in vitro.

Once the host organism and expression vector have been chosen, the first step is usually to perform expression tests on a small scale (5-100 mL of medium). The variables that can be screened are many: E. coli strains, expression temperatures, protein expression inducer concentrations, induction time, culture media [311]. E. coli BL21 and its derivatives are most frequently used for routine protein expression. Several BL21 (DE3) derivatives, known as pLysS strains, express T7 phage lysozyme, an enzyme that effectively inhibits T7 RNA polymerase activity. Inhibition of T7 RNA polymerase may decrease basal expression of the target protein, especially when the latter can result in toxicity to the host system. Moreover, numerous bacterial strains such as Rosetta and Rosetta2 contain plasmids that encode rare tRNAs to promote the efficient expression of genes that contain high frequencies of rare codons [312]. Finally, a derivative of the BL21 host strain, ArcticExpress, that coexpresses the cold-adapted chaperonins Cpn10 and Cpn60 from psychrophilic bacterium, Oleispira antarctica, has been developed and used for expression at low temperatures to improve protein solubility. Reducing the rate of transcription and/or translation facilitates folding by allowing the newly synthesized protein to fold before it aggregates and forms inclusion bodies. At lower temperatures, cell processes slow down leading to reduced protein aggregation [313,314]. Therefore, lowering the expression temperature routinely improves the solubility of recombinantly expressed proteins [315,316]. The bacterial culture should be cultivated at 37°C until mid-to-late log phase [317]. The culture is then induced and the recombinant protein synthesized between 15°C to 25°C [318]. Due to the reduced protein synthesis rate, longer induction times are necessary to obtain sufficient protein yield (typical induction time: 4 hr at 37◦C, 16 to 20 hr at 18°C). In addition to lowering the growth temperature, a reduction in transcription rate can also be achieved by lowering the concentration of the induction agent. The most common IPTG concentrations for protein induction range from 0.1 to 1.0 mM, but a decrease to even lower levels can affect solubility. Recombinant expression of target protein is usually performed in batch culture. Luria broth (LB) is the standard medium for the expression of proteins. This broth is composed of bacto tryptone, which provides peptides, peptones and essential amino acids, yeast extract, which provides vitamins and trace elements, and sodium chloride, which

provides sodium ions to maintain osmotic balance [314]. Terrific broth (TB), the second most widely used expression medium, is formulated to increase protein solubility and yield [314]. M9, or minimal medium, contains the minimum nutrients essential for bacterial species to grow. Minimal medium usually lacks amino acids, and thus is most widely used for selective labelling, as isotope labelling for proteins studied using NMR spectroscopy [319].

2.1.3. - Protein purification

The strategy of purification depends mainly upon the localization of the produced protein within the host; for example, the protein can be excreted into the growth media, transported in the periplasmic pace or produced like a soluble or insoluble (inclusion bodies) protein within the cytoplasm and finally directed to the cell membranes. In each of the cases the isolation has to be performed in different ways [320].

Cell lysis is the first step in purification. The latter can be performed mainly by sonication, mechanical methods (French Press) or the use of special reagents containing detergents. Once the cells have been lysed, centrifugation of the lysate at high speed (40000 rpm) allows the soluble fraction to be separated from the insoluble one. There are several chromatography methods for purification, based on the different physical chemical and biological features of the protein. The ones used in this study are presented below:

Immobilized Metal Ion Affinity Chromatography (IMAC)

IMAC is currently the most used affinity technique commonly used to purify recombinant proteins fused to a short peptide affinity tag constituted of 6 or more Histidines (His-tag). In IMAC, metal ions, typically nickel or zinc, are immobilized to resin via a metal chelator, such as nitrilotriacetic acid. The formation of weak coordination bonds between the metal ions immobilized on the column and the histidine residues allows the capture of the target protein. The latter is washed from the impurities and elution is usually carried out using an increasing concentration of a competitive agent such as imidazole (**Figure 21**).

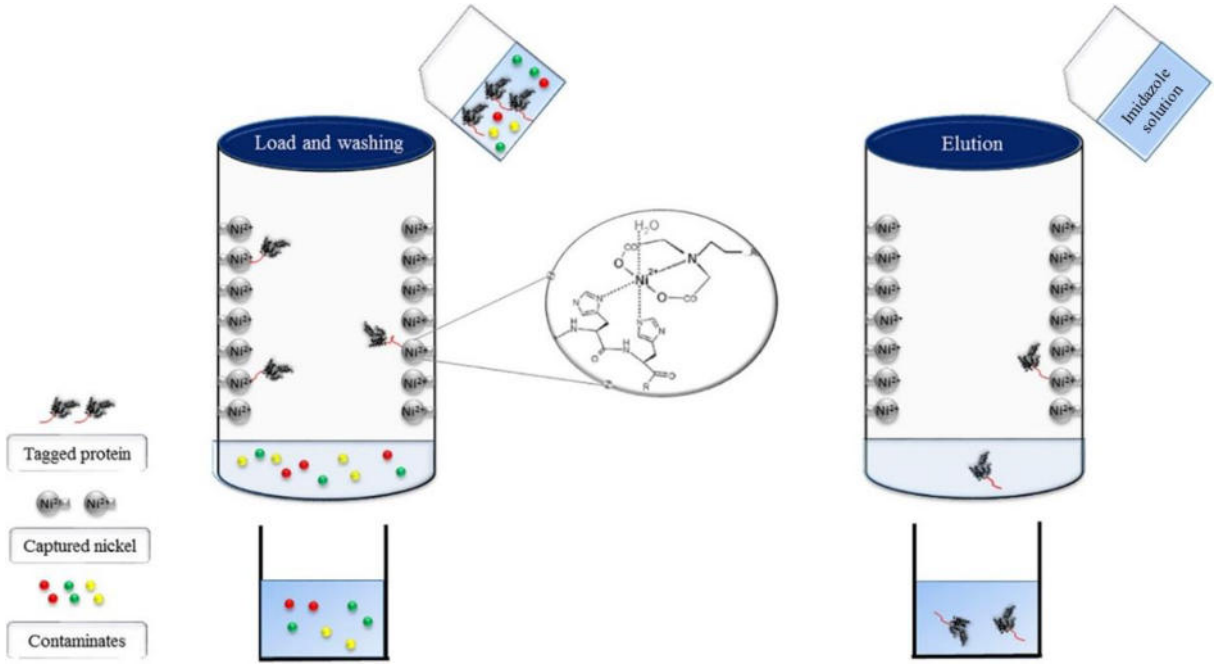


Figure 21. Schematic representation of IMAC purification (adapted from [321]).

Glutathione S-transferase (GST) Affinity Chromatography

GST-tagged proteins, as mentioned before, can be purified directly from pretreated bacterial lysates using affinity columns containing Glutathione Sepharose resin. GST-tagged proteins are eluted under mild, nondenaturing conditions using reduced glutathione (**Figure 22**).

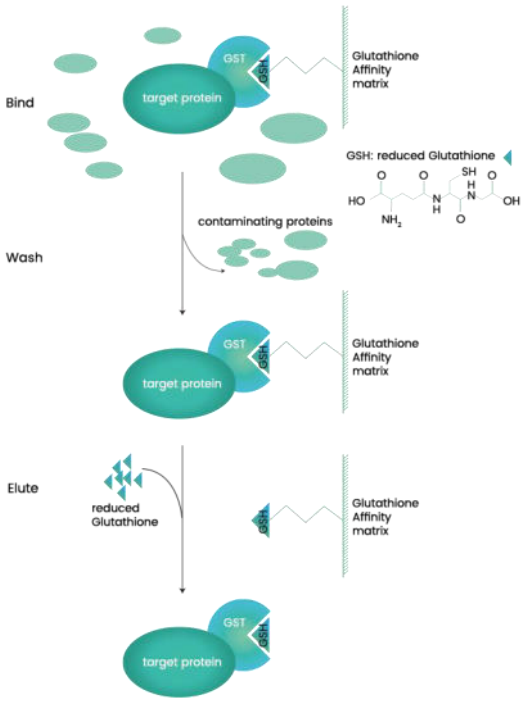


Figure 22. Graphical scheme of GST Affinity Chromatography
(https://www.sinobiological.com/resource/protein-review/gst-tag-protein-expression)

If desired, cleavage of the protein from Hist-tag or GST-tag can be achieved using a site-specific protease. One of the most widely used proteases is Tobacco Etch Virus (TEV) protease, which efficiently cleaves fusion tags without non-specific secondary cleavage [322]. Other examples of commonly used protease are PreScission, Thrombin and Factor Xa enzymes.

Size Exclusion Chromatography (SEC)

In SEC or gel filtration, molecules are separated based on their size or, more correctly, based on their hydrodynamic radius. The resin in the column consists of a series of gelatinous microspheres that have a specific pore size distribution. During the elution phase, smaller molecules will spend more time in the pores of the gel, delaying their elution, while larger molecules will pass quickly through the column (**Figure 23**). As preparative method SEC can be used to remove high or low molecular weight contaminants or to exchange buffer solutions, while analytical SEC allows to isolate one or more components from a protein mixture, to separate monomers from aggregates and to perform a molecular weight distribution analysis, provided that suitable standards are available.

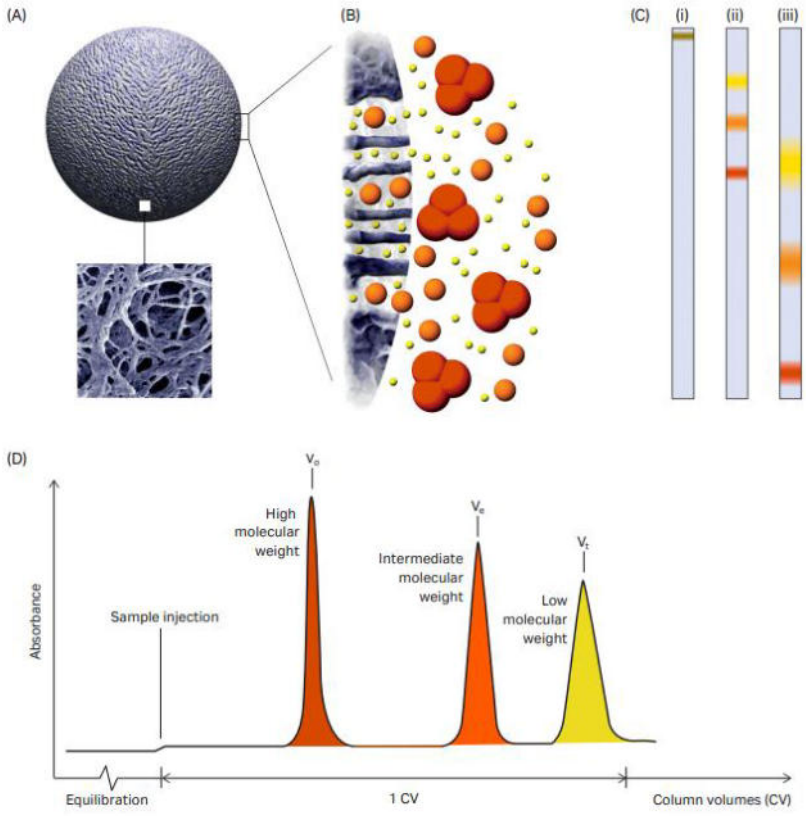


Figure 23. Process of SEC. (A) Schematic representation of a resin particle with an electron micrograph enlargement. (B) Illustration of sample molecules diffusing into the pores of the particle. (C) Graphical overview of the separation process: (i) sample application onto the column; (ii) smaller molecules (yellow) experience greater retention compared to larger molecules (dark orange); (iii) the largest species elutes first. Band broadening during chromatography leads to considerable dilution of protein zones. (D) Representative schematic chromatogram. (https://cdn.cytivalifesciences.com/api/public/content/digi-11639-pdf)

Modern gel filtration media allow separation across a molecular weight range from approximately 100 Da up to 80 million Da, making it possible to resolve biomolecules from small peptides to very large proteins and multiprotein complexes. The most critical parameter for achieving high resolution is the ratio of sample volume to column volume: larger relative sample volumes reduce resolution, while smaller ones enhance it. For preparative gel filtration, optimal resolution is typically achieved with loading volumes of 0.5–2% of the total column volume, although acceptable separation can often be maintained with volumes up to 5%. In cases where the difference in size between the target protein and impurities is substantial, even larger volumes may be tolerated. To increase effective loading capacity, samples are frequently concentrated prior to gel filtration.

Separation by SEC can be carried out over a wide range of pH, ionic strengths, and temperatures, and the resin is compatible with numerous additives such as cofactors, protein stabilizers, detergents, urea, and guanidine hydrochloride. In general, buffer composition does not markedly influence resolution; however, the inclusion of a low concentration of salt (typically 25–150 mM NaCl) is recommended to minimize weak electrostatic interactions between proteins and the gel-filtration matrix. Buffer conditions should be chosen according to the nature of the sample and the requirements for maintaining protein activity, since the target proteins are ultimately transferred into the equilibration buffer. This equilibration buffer can therefore be tailored to meet the needs of subsequent purification steps, analytical procedures, storage, or functional assays.

2.2. - SDS-PAGE analysis

Samples derived from chromatographic fractions can be tested for protein visualization using SDS-PAGE analysis. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is an

analytical technique for separating proteins based on their ability to migrate on a gel, depending on the length of their polypeptide chains and their molecular weight.

The principle behind this technique is that a protein molecule at a pH different from its isoelectric point has a total charge different from zero and moves in an electric field. This technique exploits the denaturing activity of SDS. SDS is a strongly ionic detergent that confers a negative electric charge on average for every 2 amino acid residues. Protein samples treated with buffers containing SDS are denatured and have a number of negative charges proportional to their molecular weight. In this way, proteins with the same charge have an electrophoretic mobility proportional to their size, using a three-dimensional sieve of polymerized acrylamide (gel) as a support. This type of electrophoresis is "discontinuous" because the gel consists of two parts: the lower part contains the running gel at pH 8.8, while the upper part contains the stacking gel at pH 6.8. The function of stacking gel is to align the protein molecules so that they enter the running gel simultaneously from the same starting point. The discontinuity of the gel allows for better resolution of the bands.

In detail, the preparation of samples for SDS-PAGE consists of adding 5 µL of protein sample and 10 µL of 3X FSB composed of Tris base, SDS, glycerol, 2-mercaptoethanol, and bromophenol blue. The samples are boiled in water for about 5–10 minutes, then placed on ice for 2 minutes and centrifuged for 1 minute at 13,000 × g. Subsequently, 7 µL of each sample are loaded into the wells of a 4-15% Mini-PROTEAN® TGX™ Precast Protein gel, 15-well, 15 µL (Bio-Rad), together with the molecular weight marker (Unstained Protein Standards 116-14.4 kDa, Bio-rad). The electrophoretic run is carried out for about 30 minutes at 220 Volts. Once the run is complete, the gel is immersed in Comassie Brilliant Blue G-250 dye used as a standard colorimetric method for detection of protein bands.

2.3. - Sample preparation for characterization analysis

The fractions collected at the end of the various purification steps can be subjected to buffer exchange to the desired condition using desalting columns. These columns contain Sephadex G-25 resin and are used for rapid buffer exchange, desalting and removal of small contaminants such as salts and dyes from samples. The chromatographic technique used is gel filtration, and molecules are separated based on differences in molecular size. Molecules larger than the pores in the Sephadex matrix are excluded from the matrix and are eluted first, while smaller molecules will penetrate the pores. Separation can be performed using two different protocols: gravity or centrifugation.

Then protein solution is concentrated using Amicon Ultra Filters with appropriate kDa cutoff. The concentrators have a cellulose membrane with a vertical design and a large surface area, which allows for easy sample processing, good sample recovery and a high concentration capacity of up to 80 times. Protein solution is concentrated by centrifugation at max 3500 rpm until the desired concentration is achieved. The concentration is determined by UV-visible spectroscopy. To calculate protein concentration Lambert-Beer's law is used:

$$A = \varepsilon \cdot l \cdot C$$

where A is the absorbance value measured at 280 nm, ε is the molar extinction coefficient in $M^{-1} \cdot cm^{-1}$, l is the side of the cuvette and C represents the concentration value in molarity. The molar extinction value is estimated by Protparam [www.expasy.org/tools/protparam/] or found in the literature.

In addition to UV absorbance, colorimetric methods are widely used for determining protein concentration, as they rely on the formation of colored complexes between proteins and specific reagents, with the resulting absorbance measured spectrophotometrically. The **Biuret assay** is one of the simplest, where peptide bonds react with copper (II) ions under alkaline conditions to yield a violet color, although it requires relatively high protein concentrations. The **Lowry method**, which combines the Biuret reaction with the Folin–Ciocalteu reagent, is more sensitive and detects aromatic amino acids such as tyrosine and tryptophan, but it can be affected by interfering compounds. The **Bradford assay** is one of the most employed approaches, based on the binding of Coomassie Brilliant Blue dye to proteins, resulting in a shift in absorbance measurable at 595 nm; it is rapid and convenient, even if sensitive to detergents. The **bicinchoninic acid (BCA) assay** improves on these limitations by using a copper reduction reaction coupled with BCA to form a stable purple complex, providing high sensitivity and greater tolerance to detergents, making it well suited for cell lysates. Together, these methods offer complementary options, with the choice depending on sensitivity requirements, sample purity, and buffer composition.

2.4. - Fluorescence enzymatic assay

Fluorescence is defined as the luminescence of a substance following the absorption of electromagnetic radiation. When a molecule emits fluorescence, the wavelength of the absorbed radiation is always lower (with higher energy) than the wavelength emitted (fluorescent). Fluorescence emission is often used in tests to determine the activity of an enzyme by monitoring the difference in fluorescence between the substrate and the product. This substrate, which is

generally synthetic in nature, is specially designed according to the type of enzyme being studied. In one of the studies in this thesis, the fluorescence enzymatic assay was used to evaluate the activity of the SARS-CoV-2 $3CL^{Pro}$ protein. The substrate is an 8-amino acid peptide that constitutes the recognition sequence of the viral protease. The N-terminal is linked to the fluorophore 7-methoxycoumarin-4-acetic acid (MCA), while the C-terminal is bound to the quencher 2,4-dinitrophenyl (Dnp). When the enzyme cleaves the MCA-Dnp amino acid substrate, the MCA part emits fluorescence. MCA/Dnp is a frequently used fluorophore/quencher combination that requires excitation and emission wavelengths of 320 and 405 nm, respectively, for fluorescence detection. To set the optimal conditions for fluorescence measurements, particular attention must be paid to the choice of reaction buffer, the concentration of the enzyme and fluorescent substrate, and the temperature.

2.5. - X-ray Crystallography

X-ray crystallography is currently the most favoured technique for structure determination of proteins and biological macromolecules. It enables the visualization of protein structures at an atomic level and the investigation about interactions between proteins and other molecules and how they undergo conformational changes. A crystal is an ordered 3D array of molecules held together by weak interactions. Once the crystal is exposed to X-rays, with a wavelength of around 0.1 nm or 1 Å, the latter are diffracted by the electrons in the molecule and consequently resulting in a diffraction pattern. The diffraction pattern contains all structural information regarding the electron density and position of each atom of the molecule [323]. To perform protein crystallography, a reliable source of proteins is needed, along with a purification/concentration protocol that produces soluble, homogeneous and high-quality material.

In the crystallization process the purified protein is slowly precipitated from an aqueous solution. which is brought into supersaturation. Crystallization proceeds in two phases, nucleation and growth. After nucleation, it is important to reach what is known as the ‘metastable zone’, in which the best conditions are found for the growth of large well-ordered crystals. As a result, the individual protein molecules align themselves in repeated series of “unit cells” and assume a consistent orientation. The growth of protein crystals of sufficient quality for structure determination represents the rate limiting phase of most protein crystallography work. The extent of the problem can be understood by considering different variables [324]: the choice of precipitant, its concentration, the buffer and its pH, the protein concentration, the temperature, the

crystallization technique and the possible inclusion of additives. The vapor diffusion method is the most frequently used technique to grow protein crystals. The two used common vapor diffusion techniques for protein crystallization are hanging drop and sitting drop methods.

Methodologies employed in the investigation of protein–ligand complex formation include soaking and co-crystallization. The soaking method consists of adding a drop of compound solution to the already-formed protein crystal and waiting for an indefinite time for the molecule to diffuse in the crystal through solvent channels. The solvent channels in protein crystals, a consequence of crystal packing, are primarily filled with crystallization buffer (up to 80% solvent), allowing the ligand's free diffusion into its functional binding site. The rate of diffusion is influenced by the ligand concentration and its affinity [325,326]. Using soaking is advantageous because it turns out to be faster and cost-effective. However, it may either alter the ordered structure of the crystal, cracking it, or its access to the protein site can be blocked by high packing of the protein's crystal lattice. By contrast, co-crystallization entails exposing the protein to the ligand in solution at the beginning of the experiment [325,327]. In comparison with soaking, co-crystallization is more accurate for determining the correct ligand-binding position [327] as crystal packing tends to favour the ligand position bound to the active site. However, co-crystallization is a more time-consuming and costly process.

Several crystallographic structures are proposed in this thesis. The structures of the SARS-CoV-2 $3CL^{Pro}$ protein complexes with Au(NHC)Cl, $Au(PEt_3)Br$ and GC-376 PROTAC precursor were obtained by soaking methods. Instead, crystal structures of the SARS-CoV-2 $3CL^{Pro}$ and CVB3 $3C^{Pro}$ protein complexes with GC-376 PROTAC were achieved by co-crystallization.

2.6. - NMR spectroscopy

NMR spectroscopy is an indispensable tool for investigation of biological molecules and their interactions [34,35]. The NMR method provides data that are in many ways complementary to those obtained from X-ray crystallography and thus promises to widen our view of protein molecules, giving a clearer insight into the relation between structure and function.

Whereas X-ray crystallography requires single crystals, NMR measurements are carried out in solution under conditions that can be as close as possible to the physiological state. Furthermore, the analysis through NMR spectroscopy easily allows the characterization under several different experimental conditions, such as such as different ionic strength and pH [328]. In addition to protein structure determination, NMR measurements can also include investigation of dynamic features

of the molecular structures, as well as studies of thermodynamic and kinetics aspects of interactions between proteins and other solution components [329].

NMR spectra are generated by placing a sample in a magnetic field and applying radiofrequency pulses, which perturb the equilibrium nuclear magnetization of these atoms with nuclei of non-zero spin. Transient time domain signals are detected as the system returns to equilibrium. Fourier transformation of the transient signal into a frequency domain yields a one-dimensional NMR spectrum, which is a series of resonances from the external fields. Therefore, the frequencies of different nuclei vary due to their different chemical environments. This effect is called 'chemical shift'. It is one of the major parameters of NMR spectroscopy since it causes the different positions of the signals in an NMR spectrum. The most prominent feature of NMR spectroscopy is its ability to differentiate chemically identical atoms, experiencing small variations in their magnetic environment influenced by neighboring atoms [35,330,331].

The fundamental step is the preparation of the protein solution, since a highly purified protein preparation is required. An inhomogeneous preparation and/or aggregation of the protein as well as low molecular weight protonated impurities may severely harm the structure determination. The first step in every protein NMR study therefore involves optimization of the measurement conditions as pH, ionic strength, and temperature that can often be adjusted to mimic physiological conditions [328]. The macromolecule under study should be stable in the chosen conditions for many weeks. Proteins with a molecular weight larger than 10 kDa must be isotope enriched in ^{15}N and ^{13}C for efficient NMR studies. ^{15}N and ^{13}C are used because the most abundant carbon isotope (^{12}C) does not give an NMR signal and the most abundant nitrogen isotope (^{14}N) has undesired NMR properties. Often, larger proteins require a degree of deuteration of non-exchangeable protons in order to reduce the efficiency of proton-proton relaxation mechanisms and thus obtain narrower and more observable signals [332]. For the NMR measurements the protein is dissolved in 0,25-0,5 mL of aqueous buffer that contains about 10 % of D_2O which is necessary for the stabilization of the NMR instrument during the measurement.

NMR spectra of biological macromolecules contain hundreds or even thousands of resonance lines which cannot be resolved in a conventional one-dimensional spectrum (1D). Further, the interpretation of NMR data requires correlations between different nuclei, which are implicitly contained in 1D spectra but often difficult to extract. Multidimensional NMR spectra provide both increased resolution and correlations which are easier to analyze [35].

The NMR multidimensional measurements almost always use protons (^{1}H) and, depending on the isotope labelling, ^{13}C and/or ^{15}N nuclei, exploit the transfer of the magnetization between these nuclei via through bond (scalar coupling) interactions. Such experiments allow the assignment of ^{15}N and/or ^{13}C backbone chemical shifts and the determination of protein structures up to 25 KDa range [333]. These experiments allow systematic information to be obtained on backbone resonances as they link the signals of the amide protons (HN) and those of the side chains (Cα,Cβ,Co). Through the analysis of 3D triple resonance experiments we can obtain the specific sequence assignment of the protein, which consists of linking the NMR signals to a specific nucleus of a specific amino acid in the protein sequence. Subsequently, the NMR method for protein structure determination relies on the measurement of through-space correlations via the Nuclear Overhauser Effect (NOE). A dense network of distance constraints are derived from NOEs between nearby hydrogen atoms in the protein [334]. NOEs connect pairs of hydrogen atoms separated by less than about 5 Å in space.

Sequence Resonance Assignment and Chemical Shift Perturbation Mapping

The application of multidimensional NMR spectroscopy allowed the development of general strategies for the assignment of signals in proteins. All procedures use the known protein sequence to connect nuclei of amino acid residues which are neighbours in the sequence. For unlabelled proteins smaller than 10 kDa the combination of the [^{1}H,^{1}H]-COSY or TOCSY, used for the sequential assignment, with the [^{1}H, ^{1}H]-NOESY spectrum allows the assignment of most proton NMR signals to individual protons [334]. For larger proteins extensive signal overlap prevents complete assignments of all ^{1}H signals in proton spectra. This barrier can be overcome with 3D NMR technique and uniformly ^{13}C and ^{15}N labelled proteins. With these methods, systems with molecular weights up to approximately 30 kDa can be studied [335]. The 3D triple resonance experiments exclusively correlate the resonances of the peptide backbone (HN(i), N(i), Cα(i), Hα(i), Cα(i-1), CO(i) and CO(i-1)). The 3D experiments used to identify the backbone resonances are, usually, HNCA or HNCACB, HN(CO)CA or HN(CO)CACB, HNCO, HN(CA)CO and HNHA [336].

One of the most important protein observed methods for the investigation of protein–ligand interactions is the chemical shift mapping (CSM) also known as the chemical shift perturbation (CSP) or complexation induced changes in chemical shifts (CIS) [337]. Thereby, a series of NMR spectra of the protein are recorded in the absence and presence of varying amounts of the binding ligand. Due to its superior signal dispersion, the most common experiment which is used for

chemical shift mapping is the ^{15}N-heteronuclear single quantum correlation (HSQC) experiment. If there is an interaction, the chemical shifts of the residues involved in the complex formation with ligand, seen as peaks in a ^{15}N-HSQC, are displaced from their original position. Two limiting cases can be described: low and fast chemical exchange between the free and protein-bound form of the ligand. In the fast exchange limit the two signals collapse into one, whose chemical shift represents the population averaged value of the free and ligand-saturated protein. In the slow exchange regime both signals of bound and free state are observed with signal integrals representing their relative amounts. When CSM is carried out on a protein with known resonance assignments, the residues which are involved in the interactions with the ligand are revealed. Chemical shift mapping is also particularly useful for the screening of ligands and is therefore often used in drug design. It not only provides information about binding and the binding strength, but also about the location on the protein where the interaction takes place. One essential requirement for chemical shift mapping is that both the protein and the ligand are dissolved in the exact same buffer and measured under the same conditions, since chemical shifts, especially those of amide protons are very sensitive to differences in pH value, temperature and buffer composition. Therefore, the preservation of constant solution conditions during a protein-ligand titration is important to obtain spectra with improved quality [338].

3. Chapter – Results and discussion

3.1. - Exploring Gold metallodrugs to inhibit the SARS-CoV-2 3CL^{Pro}

The emergence and swift proliferation of COVID-19 pandemic have created significant challenges for healthcare systems across the globe. In order to develop potent antiviral medications targeting SARS-CoV-2, broadening the range of chemical compounds tested, particularly by incorporating a diverse array of metal-based compounds, is highly beneficial. Recently, efforts have been made to monitor the metallodrugs' effectiveness against the COVID-19 pandemic. Cohen et al., in their work, have reported the description of metal complexes, containing bismuth, rhenium or gold, that have been shown to inhibit various SARS-CoV-2 enzymes and their proposed mechanism of action [119]. Numerous gold(I) and gold(III) complexes have been documented to effectively suppress the function of multiple proteins and enzymes that contain thiol groups, such as thioredoxin reductase and thioredoxin-glutathione reductase, complexes implicated in antiviral properties [339].

Messori et al. proposed the pharmaceutical compound Auranofin for its possible action against SARS-CoV-2, highlighting the excellent tolerance of the compound in the body, its low toxicity and its multi-targeting ability [340]. The prospect of using Auranofin as a therapeutic agent against SARS-CoV-2 has garnered significant attention. Several studies have shown that it can suppress viral replication by targeting and inhibiting particular enzymatic functions of the virus [341-343].

In the first presented paper I demonstrate, through in-vitro enzymatic activity assays, the capacity of Auranofin and some derivatives (the three halido analogs $Au(PEt_3)Br$, $Au(PEt_3)I$, $Au(PEt_3)Cl$ and the two gold carbene complexes Au(NHC)Cl and $[Au(NHC)_2]PF_6$) to suppress the activity of the SARS-CoV-2 main protease. Since all these gold compounds have previously exhibited notable cytotoxic and antimicrobial properties [344,345], a drug repurposing approach was employed by selecting this set of compounds to identify potential new therapeutics for COVID-19. With the only exception of $[Au(NHC)_2]PF_6$, all the compounds turned out to be potent inhibitors of the SARS-CoV-2 3CL^{Pro} catalytic activity. The measured inhibition constants fall within the low micromolar range.

Additionally, mass spectrometry analyses reveal that these gold(I) agents form strong, tight complexes with SARS-CoV-2, indicating the coordination of well-defined gold-containing fragments to the protease, again with the exclusion of $[Au(NHC)_2]PF_6$ in accordance with its low enzyme inhibitory activity. Furthermore, the crystal structures of two of these metal-protein

complexes, the protein adducts with $Au(PEt_3)Br$ and Au(NHC)Cl, were obtained at atomic resolution, providing detailed insights about the interaction between the gold(I) centers and the enzyme. In particular, the crystal complexes reveal that Cys 156 and the catalytic Cys 145 serve as primary sites for metal attachment within the protein. Collectively, the findings presented here strongly endorse the idea that the effective and promising suppression of $3CL^{Pro}$'s enzymatic activity by these gold compounds is probably due to the strong coordinate bond formed between a gold-containing fragment and the catalytic Cys 145.

Overall, these experiments highlight the strong affinity of gold(I) compounds for the active site thiols in SARS-CoV-2 $3CL^{Pro}$, suggesting that these compounds serve as promising starting points for the development of more potent gold-based inhibitors targeting SARS-CoV-2. A potential future direction for our work could involve investigating the capacity of these gold compounds to inhibit SARS-CoV-2 infection in vivo models, as well as comparing their mechanisms of action and therapeutic effectiveness to those of the established gold drug Auranofin.

My contribution to this part of the project consisted of the expression and purification of the SARS-CoV-2 $3CL^{Pro}$ target proteases as described previously [120]. In detail, SARS-CoV-2 $3CL^{Pro}$ protein was previously cloned in a pGEX plasmid with an N-terminal autocleavage sequence and a His-tag at the C-terminus. Its expression was performed in BL21 (DE3) bacterial cells in LB medium for natural abundance production. The induction was performed overnight at low temperature with IPTG 0,5 mM final concentration. An IMAC purification was carried out on a HisTrap FF column. The fractions containing target protein were collected and mixed with PreScission protease, resulting in the target protein with authentic N- and C-termini. The PreScission-treated $3CL^{Pro}$ was purified by the GST-tagged PreScission protease, the His-tag and protein with un-cleaved His-tag, using two connected columns: GSTtrap FF (GE Healthcare) and a nickel column (IMAC). Afterwards the His-tag-free $3CL^{Pro}$ was purified again by SEC (HiLoad Superdex 200 16/60) to remove unexpected impurities. The main trouble I dealt with was the high genomic instability of pGEX expression systems recorded by genome sequencing. This drawback could be related to RecA independent recombination events [346]. Although this drawback, I was able to optimize the expression conditions obtaining a final protein yield of about 10 mg/L of culture. Moreover, I carried out the fluorescence enzymatic assay for the determination of the inhibitory effects of gold compounds on the protein activity. SARS-CoV-2 $3CL^{Pro}$ samples were prepared by adding increasing concentrations of the metal-based compounds. Then the protein-drug mixtures were incubated at room temperature for 2 hours, after which a fluorogenic peptide substrate (Mca–AVLQSGFR-K(Dnp)K) was added for the fluorescence measurements (excitation

at 320 nm, emission at 405 nm). Furthermore, I was involved in the experimental trials for obtaining the native crystals of apo SARS-CoV-2 $3CL^{Pro}$ and crystal adducts of the protein in complex with the gold compounds $Au(PEt_3)Br$ and Au(NHC)Cl.

3.1.1. - Gold-Based Metal Drugs as Inhibitors of Coronavirus Proteins: The Inhibition of SARS-CoV-2 Main Protease by Auranofin and Its Analogs

Lara Massai [1,†], Deborah Grifagni [1,2,†], Alessia De Santis [1,2], Andrea Geri [1], Francesca Cantini[1,2,3,*], Vito Calderone [1,2,3,*], Lucia Banci [1,2,3] and Luigi Messori [1,*]

1 Department of Chemistry "Ugo Schiff", University of Florence

2 Magnetic Resonance Center (CERM), University of Florence

3 Consorzio Interuniversitario Risonanze Magnetiche Metallo Proteine (CIRMMP), University of Florence

* Correspondence: cantini@cerm.unifi.it (F.C.); calderone@cerm.unifi.it (V.C.); luigi.messori@unifi.it (L.M.)

† These authors contributed equally to this work.

Published: 10.3390/biom12111675

biomolecules

MDPI

Article

Gold-Based Metal Drugs as Inhibitors of Coronavirus Proteins: The Inhibition of SARS-CoV-2 Main Protease by Auranofin and Its Analogs

Lara Massai [1,†], Deborah Grifagni [1,2,†], Alessia De Santis [1,2], Andrea Geri [1], Francesca Cantini [1,2,3,*], Vito Calderone [1,2,3,*], Lucia Banci [1,2,3] and Luigi Messori [1,*]

[1] Department of Chemistry "Ugo Schiff", University of Florence, Via Della Lastruccia 3, 50019 Florence, Italy
[2] Magnetic Resonance Center (CERM), University of Florence, Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy
[3] Consorzio Interuniversitario Risonanze Magnetiche Metallo Proteine (CIRMMP), University of Florence, Via L. Sacconi 6, 50019 Florence, Italy
* Correspondence: cantini@cerm.unifi.it (F.C.); calderone@cerm.unifi.it (V.C.); luigi.messori@unifi.it (L.M.)
† These authors contributed equally to this work.

Abstract: Gold compounds have a long tradition in medicine and offer many opportunities for new therapeutic applications. Herein, we evaluated the lead compound Auranofin and five related gold(I) complexes as possible inhibitors of SARS-CoV-2 Main Protease (SARS-CoV-2 M^{pro}), a validated drug target for the COVID-19 disease. The investigational panel of gold compounds included Auranofin; three halido analogues, i.e., $Au(PEt_3)Cl$, $Au(PEt_3)Br$, and $Au(PEt_3)I$; and two gold carbene complexes, i.e., Au(NHC)Cl and $[Au(NHC)_2]PF_6$. Notably, all these gold compounds, with the only exception of $[Au(NHC)_2]PF_6$, turned out to be potent inhibitors of the catalytic activity of SARS-CoV-2 M^{pro}: the measured K_i values were in the range 2.1–0.4 μM. The reactions of the various gold compounds with SARS-CoV-2 M^{pro} were subsequently investigated through electrospray ionization (ESI) mass spectrometry (MS) upon a careful optimization of the experimental conditions; the ESI MS spectra provided clear evidence for the formation of tight metallodrug-protein adducts and for the coordination of well defined gold-containing fragments to the SARS-CoV-2 M^{pro}, again with the only exception of $[Au(NHC)_2]PF_6$. The metal-protein stoichiometry was unambiguously determined for the resulting species. The crystal structures of the metallodrug- M^{pro} adducts were solved in the case of $Au(PEt_3)Br$ and Au(NHC)Cl. These crystal structures show that gold coordination occurs at the level of catalytic Cys 145 in the case of Au(NHC)Cl and at the level of both Cys 145 and Cys 156 for $Au(PEt_3)Br$. Tight coordination of gold atoms to functionally relevant cysteine residues is believed to represent the true molecular basis of strong enzyme inhibition.

Keywords: COVID-19; nsp5; M^{pro}; SARS-CoV-2; Auranofin; gold compounds; ESI mass spectrometry

Citation: Massai, L.; Grifagni, D.; De Santis, A.; Geri, A.; Cantini, F.; Calderone, V.; Banci, L.; Messori, L. Gold-Based Metal Drugs as Inhibitors of Coronavirus Proteins: Thse Inhibition of SARS-CoV-2 Main Protease by Auranofin and Its Analogs. *Biomolecules* **2022**, *12*, 1675. https://doi.org/10.3390/biom12111675

Academic Editor: Georgi Momekov

Received: 3 October 2022
Accepted: 8 November 2022
Published: 11 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

The outbreak and the rapid spread of COVID-19 disease are posing dramatic problems to health systems worldwide [1]. The identification and rapid implementation of effective antiviral drugs against SARS-CoV-2 are urgently needed to fight this severe disease. To this end, expanding the chemical space of the tested compounds by including a large variety of metal compounds is highly desirable. It has to be stressed that the search for efficient inhibitors of SARS-CoV-2-related proteases has almost entirely focused on organic molecules [2–4]; much less attention has been deserved so far to inorganic moieties [5,6] that merit, in our opinion, a greater consideration.

Indeed, the known inorganic drugs contain a wide array of metals imparting peculiar chemical properties, that arise from the electronic structure of the metal center, its coordination sphere, the nature of the ligands, the redox properties, etc. It is evident that these chemical features cannot be completely reproduced by simple organic compounds. Accordingly, the unique chemical and biological properties of the various metal centers- in many cases non-physiological metals- should be taken into due account within new drug discovery programs. This approach might hopefully lead to successful pharmacological and therapeutic outcomes.

The main protease of the SARS-CoV-2 virus is nsp5, also known as $3CL^{pro}$ or M^{pro}. This is a homodimeric protein with each monomer having a molecular mass of 33796 Da. Each monomer contains three domains. More in detail, domain I (residues 8–101) and domain II (102–184) fold in an antiparallel β-barrel -where the active site with the Cys145-His41 catalytic dyad is located [7] whereas domain III (residues 201–306) is involved in the process of dimerization that is critical for the function of the enzyme [8]. The two subunits of M^{pro} bind to each other via the N-terminus (residues 1–7) where the Ser1 of each monomer completes and stabilizes the S1 substrate binding pocket of the adjacent monomer [9]. The M^{pro} function is indispensable for viral replication as it cleaves the polyproteins pp1a and pp1ab at eleven conserved sites, to allow the correct folding and function of essential viral proteins [10,11]. For this reason, M^{pro} is a widely accepted pharmacological target [12]. It is also important to consider that M^{pro} is an enzyme with a high structural conservation of the main core among the various coronaviruses [13,14]; therefore, the discovery of high-affinity inhibitors for this protease may be an excellent starting point for selecting new antiviral compounds that are active against a broad spectrum of coronaviruses [15]. The recently solved 3D structures of SARS-CoV-2 M^{pro} in complex with zinc ion [16–18] showed that the metal ion is bound to the catalytically relevant Cys145 and His 41 residues. Furthermore, a few recent works demonstrated that other metal ions or metal-based compounds, such as rhenium, gold and copper compounds, can be efficient inhibitors of viral proteases [19–21]. For instance, Auranofin is a representative gold drug that was claimed to exhibit remarkable antiviral properties; notably, the efficacy of Auranofin and some related gold compounds against Papain-like protease was nicely documented by Ingo Ott et al. [20]. In addition, the crystal structure of the adduct formed between SARS-CoV-2 M^{pro} and Auranofin was recently solved showing that gold is, at least partially, able to target the catalytic cysteine residue [22]. Another recent work [23] has analyzed an even larger set of gold compounds; this study proved once more the ability of such compounds to inhibit the catalytic activities of SARS-CoV-2 M^{pro} and papain-like protease (PL^{pro}).

Here we show, through in-vitro enzymatic activity assays, the ability of three gold(I) Auranofin analogs and two gold carbene complexes to inhibit the SARS-CoV-2 M^{pro} activity, and compare the results with those obtained with Auranofin. All of these gold complexes as already shown interesting cytotoxicity properties and antimicrobial activities [24–26], then, using the approach of the repurposing strategy we selected this panel of compounds to identify new possible drugs for the treatment of the SARS-CoV-2 disease. Moreover, we show, through mass spectrometry analysis, that these gold(I) compounds bind SARS-CoV-2 M^{pro} tightly forming well-defined metal-protein complexes. In addition, the crystallographic 3D structures for two of these metal-protein complexes have been solved providing atomic-level details about the binding mode of the gold(I) centers to the protein and disclosing the molecular basis for the observed enzyme inhibition. Summarizing, all the above experiments reveal a significant affinity of gold(I) compounds for active site thiols in SARS-CoV-2 M^{pro}. The resulting inhibition constants are estimated in the order of the low micromolar range, which is indeed a promising starting point for the design of new and potent gold-based inhibitors of SARS-CoV-2.

2. Materials and Methods

2.1. Sources of Compounds and Chemicals

The M^{pro} was produced through heterologous expression in E. Coli as previously described in [16]. The gold compounds were prepared and characterized according to the previously reported procedures [27,28].

The fluorescence-quenched peptide substrate used in the enzymatic assay has been purchased from GenScript Biotech, Leiden, Netherlands. If not otherwise specified, reagents and chemicals used have been purchased from Merck KGaA, Darmstadt, Germany.

2.2. Enzymatic Activity Assay

An initial stock of M^{pro} 10 μM in 20 mM Tris-HCl, 150 mM NaCl 1 mM EDTA, 5 mM DTT pH 7.8 was diluted in the measurement buffer (Tris-HCl 20 mM 20% glycerol pH 7.2) up to 0.2 μM. This solution was used to prepare different samples containing increasing concentrations of metal-based drugs as follows: Auranofin (0, 0.5, 1, 2, 4, 6, 8, 10, 12 and 15 μM,), Au(NHC)Cl (0, 0.3, 0.6, 1, 2, 3, 6, 8, 10, 15 and 20 μM), $Au(PEt_3)Br$ (0, 0.1, 0.2, 0.3, 0.4, 0.6, 0.8, 1, 1.5, 2, 3 and 6 μM), $Au(PEt_3)I$ (0, 0.5, 1, 1.5, 2, 3, 4, 5, 6, 8, 10 and 15 μM) and $Au(PEt_3)Cl$ (0, 0.1, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 2, 2.5, and 3 μM). The protein-molecule solutions were incubated at r.t. for 2 h, thereafter 4 μM of a fluorescence-quenched peptide substrate (Mca–AVLQ↓SGFR-K(Dnp)K) was added for each measurement. Fluorescence curves were acquired for 1 min at an excitation and emission wavelength of 320 and 405 nm, respectively. All experiments were performed in duplicate at least two times. The fluorescence curve slopes were calculated by linear fits at 18 s of the fluorescence signal increase as a function of time. For each molecule, all measurements were analyzed in a single fitting, which allowed the calculation of K_i values using the Hill equation where Vmax and Vmin were fixed at 100 and 0, respectively, while the Hill coefficient, n, was not fixed. Originpro 2018 was used to analyzed all the spectra.

3. Electrospray Ionization Mass Spectrometry Experimental Conditions

3.1. Sample Preparation

Stock solutions of SARS-CoV-2 M^{pro} were prepared by dissolving the protein in ammonium acetate solution 2×10^{-3} M (pH 6.8), mixed with aliquots of DTT in cysteine-to-reducing agent ratio 1:5. Stock solutions 10^{-2} M of the gold(I) compounds were prepared by dissolving the compounds in DMSO. For the experiments with SARS-CoV-2 M^{pro}, aliquots of the respective stock solutions were mixed with aliquots of gold compounds at a protein-to-metal ratio of 1:3 and diluted with ammonium acetate solution 2×10^{-3} M (pH 6.8) to 10^{-4} M final protein concentration. In the case of $Au(PEt_3)Br$, the protein-to-metal ratio was 1:1. The mixtures were incubated at 37 °C up to 24 h. After the incubation time, all solutions were sampled and diluted to a final protein concentration of 10^{-6} M using ammonium acetate solution 2×10^{-3} M, pH 6.8.

3.2. Instrumental Parameters

The ESI mass study was performed using a TripleTOF 5600 + high-resolution mass spectrometer (AB Sciex, Framingham, MA, USA), equipped with a DuoSpray® interface operating with an ESI probe. The general ESI source parameters optimized for protein analysis were as follows: SARS-CoV-2 M^{pro}: positive polarity, ion spray voltage floating 5500 V, temperature 0 °C, ion source Gas 1 (GS1) 25 L/min; ion source Gas 2 (GS2) 0; curtain gas (CUR) 45 L/min, collision energy (CE) 10 V; declustering potential (DP) 300 V, acquisition range 2400–5000 m/z. For acquisition, Analyst TF software 1.7.1 (AB Sciex, Framingham, MA, USA) was used, and deconvoluted spectra were obtained by using the Bio Tool Kit micro-application v.2.2 embedded in PeakViewTM software v.2.2 (AB Sciex, Framingham, MA, USA)).

3.3. Crystallization, Data Collection and Structure Solution

Crystals of apo SARS-CoV-2 M^{pro} were obtained in sitting drop by adding an aliquot of 2 µL of protein solution (20 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.8) to 2 µL of reservoir buffer (20 mM ammonium acetate, 20% PEG 3350 pH 7) and stored at 20 °C. The protein concentration in the sample was 5 mg/mL. The native crystals of apo SARS-CoV-2 M^{pro} were afterwards soaked in a 10% DMSO solution containing Au(PEt_3)Br Au(NHC)Cl respectively having a 2–3-fold ligand concentration with respect to the protein for about 20 h. Both adduct datasets were collected in-house, using a BRUKER D8 Venture diffractometer equipped with a PHOTON III detector, at 100 K; the crystals used for data collection were cryo-cooled using 25% ethylene glycol in the mother liquor. The crystals diffracted up to 2.2–2.3 Å resolution but structures have been refined at 2.4 Å: they belong to space group C2 with one molecule in the asymmetric unit, a solvent content of about 50%, and a mosaicity of 0.6°. The data were processed using the program XDS [29] reduced and scaled using XSCALE [29] and amplitudes were calculated using XDSCONV [29]. Both structures have been solved using the molecular replacement technique; in both cases, the model used was the 7NXH structure. The successful orientation hand translation of the molecule within the crystallographic unit cell was determined with MOLREP [30]. The refinement and water molecule fitting has been carried out using PHENIX [31], applying default TLS restraints. In between the refinement cycles, the model was subjected to manual rebuilding using COOT [32]. The quality of the refined structures was assessed using the program MOLPROBITY [33]. Data processing and refinement statistics for both adducts are shown in Table S1. Coordinates and structure factors have been deposited at the PDB under the accession code 8B0T for the Au(PEt_3)Br and 8B0S for Au(NHC)Cl.

4. Results and Discussion

4.1. Chemical Features of the Study Compounds

The screening of drugs already approved for a different disease (the so-called "drug repurposing") may result in an excellent drug discovery strategy, particularly in case a new therapeutical approach needs to be rapidly developed against an emerging disease. This strategy may allow to skip safety assessment, preclinical testing, and formulation development, i.e., the long and laborious phases that are typically required for the approval of a newly synthesized drug, thus drastically reducing the times for clinical implementation [34]. In this regard, Auranofin could be a very good candidate for drug repurposing. In fact, this drug was approved for the treatment of rheumatoid arthritis in 1988 and its safety profile is well-known and acceptable; Auranofin is currently under investigation for cancer therapy as a repurposed drug [35] and has also been proposed to fight a number of bacterial, parasitic and viral infections [36]. The potential of Auranofin as a drug for the treatment of SARS-CoV-2 recently emerged as a hot topic; a few studies demonstrated its ability to inhibit viral replication by contrasting specific enzymatic activities of the virus [5,6,37]. Auranofin could perform its antiviral activity in different ways: one of these is blocking the entry of the virus into the host cell; another is to inhibit the activity of selected viral enzymes that are crucial for virus replication and a last one is to counter host inflammatory response [38].

On the contrary, the other gold compounds presented herein, have not been approved yet for clinical use, but have been extensively tested for cellular cytotoxicity against a variety of cancer cells. Their IC_{50} values [39,40] typically fall in the low-micromolar range thus revealing the opportunity to be further investigated in cells and in vivo as prospective anticancer agents. In addition, some of these compounds may be exploited for different therapeutic uses, e.g., as antiviral or antimicrobial agents, if given at sub-cytotoxic concentrations.

The panel of gold compounds that we have used for the present investigation is shown in Figure 1. The panel contains six distinct gold(I) compounds, namely Auranofin,

three halido analogs, i.e., Au(PEt3)Cl, Au(PEt3)Br and Au(PEt3)I, and two gold carbene compounds, i.e., Au(NHC)Cl and [Au(NHC)2]PF6. All these compounds contain a linearly coordinated gold(I) center and were previously described and characterized [27,28]. In four cases, i.e., Au(PEt3)Cl, Au(PEt3)Br, Au(PEt3)I, and Au(NHC)Cl, the halido ligand behaves as the leaving group; the sugar ligand is the leaving group in the case of Auranofin. Activation of the gold(I) center in [Au(NHC)2]PF6 is far more difficult as it requires the removal of a strong carbene ligand.

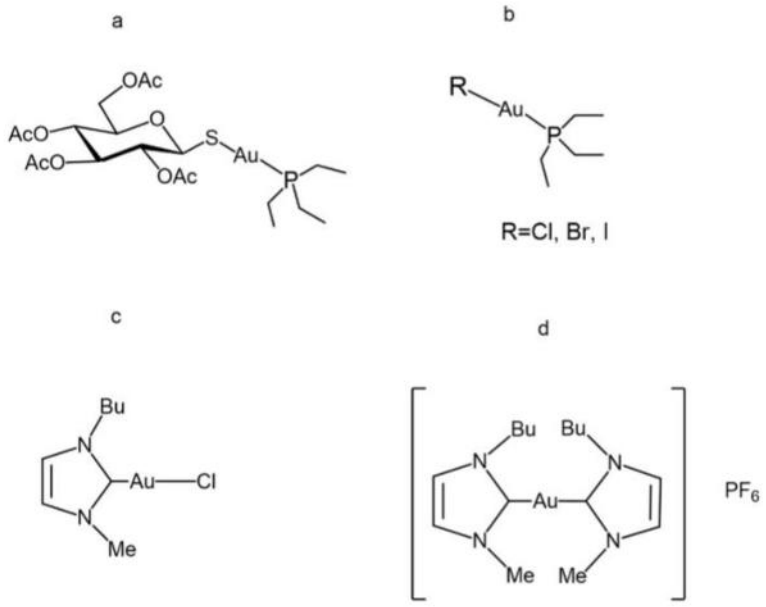


Figure 1. Gold(I) compounds chemical structure: (**a**) Auranofin, (**b**) Auranofin halido analogues, (**c**) Au(NHC)Cl and (**d**) [Au(NHC)2]PF6.

4.2. Enzyme Inhibition by the Study Compounds

The ability of the various gold compounds to inhibit the catalytic activity of SARS-CoV-2 M^{pro} was comparatively measured using a previously established procedure [16]. A dose-dependent M^{pro} inhibition assay was carried out by treating 0.2 μM solutions of M^{pro} with increasing concentrations of the various panel compounds (ranging from 0 to 20 μM; Figure S1). All the experiments were performed in duplicate and the initial rates of the reactions were measured through the linear fittings of the fluorescence curves. The resulting inhibition profiles are shown in Figure 2, whereas the values of the inhibition constants, K_i, are reported in Table 1. Notably, as all tested gold compounds, with the only exception of [Au(NHC)2]PF6, give a K_i value in the low μM range, they may be reputed as promising inhibitors of this enzyme.

Among the panel compounds, Au(PEt3)Br exhibited the maximum inhibitory potency for M^{pro}; the inhibitor potencies of Au(PEt3)Cl, Au(PEt3)I and Au(NHC)Cl were roughly similar, and close to the value of Auranofin. [Au(NHC)2]PF6 turned out to be a very weak inhibitor of M^{pro} activity causing no more than 10% loss of activity at the very high concentration of 60 μM (Figure S2).

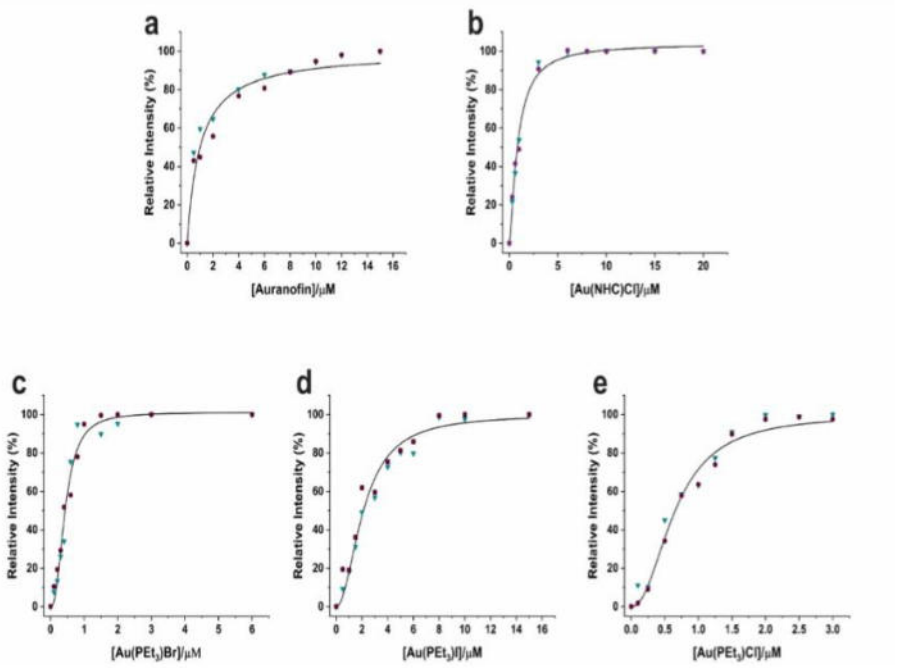


Figure 2. Inhibition M^{pro} profiles by (**a**) Auranofin (**b**) Au(NHC)Cl (**c**) Au(PEt_3)Br (**d**) Au(PEt_3)I (**e**) Au(PEt_3)Cl. Two independent measurements are shown as red dots and blue triangles, respectively. K_i values were determined by nonlinear regression.

Table 1. Inhibition constants, K_i values, measured for the interaction between gold compounds and M^{pro}.

Compound	K_i
Auranofin	0.99 ± 0.107 μM
Au(NHC)Cl	0.87± 0.04 μM
Au(PEt_3)Br	0.44± 0.022 μM
Au(PEt_3)I	2.12 ± 0.096 μM
Au(PEt_3)Cl	0.66 ± 0.026 μM

4.3. Metallodrug-Enzyme Interactions Probed by ESI MS

To better explain the above inhibition profiles further studies have been carried out to characterize in detail the metallodrug-M^{pro} interactions that form the molecular basis of enzyme inhibition. Thus, the interactions between all panel compounds and the enzyme were investigated through ESI MS experiments. In principle, ESI MS is an excellent tool to characterize protein metalation by metal-based drugs. Indeed, a number of successful studies were reported in the recent literature [41–46] However, in the case of M^{pro}, we encountered a lot of experimental difficulties in setting properly the ESI MS experiments to characterize the SARS-CoV-2 M^{pro}/ gold metallodrug adducts. A lot of efforts were done to determine the optimal conditions for recording the ESI MS spectra; the best conditions that were eventually identified and then applied to the various samples are described in the methods section. The resulting ESI MS spectra offer clear evidence that SARS-CoV-2 M^{pro} may bind metallic fragments originating from the various gold drugs upon activation (Figure 3) as described below.

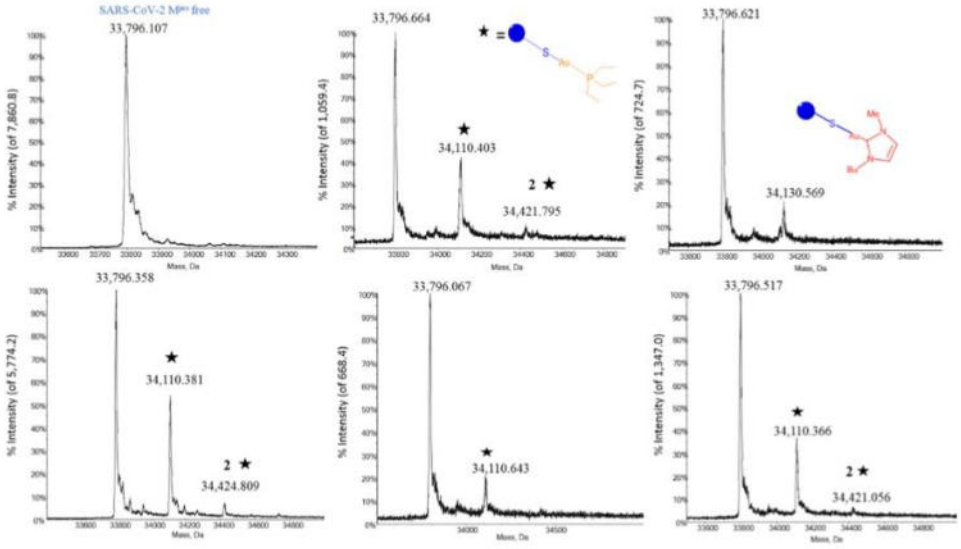


Figure 3. Deconvoluted ESI-QTOF mass spectra of M^{pro} incubated with different gold compounds for 4 h, (**a**) M^{pro} free (**b**) Auranofin (**c**) Au(NHC)Cl (**d**) Au(PEt_3)Br (**e**) Au(PEt_3)I (**f**) Au(PEt_3)Cl. The star indicates the compound moiety that is bound to the protein in the case of Auranofin and its derivates.

Notably, the deconvoluted mass spectrum of the free protein shows a well-resolved signal at 33796 Da that nicely matches the molecular mass of the native protein. Upon mixing, Auranofin and its halido analogues manifest a good reactivity with SARS-CoV/2-M^{pro} already after 4 h of incubation, giving rise to two different signals, respectively assigned to the mono- and the (generally less abundant) bis-adduct; in both cases, the metal-bound fragment nicely corresponds to the $[Au(PEt_3)]^+$ moiety (MW 315). Similarly, the interaction between the monocarbene gold complex Au(NHC)Cl and the protein results in the formation of a mono adduct bearing the 1-butyl-3-methylimidazole-2-ylidene-gold(I) moiety (MW 336) (Figure 3b). In all cases, the signal of unreacted SARS-CoV/2-M^{pro} is still present, indicating that the metalation of the protein is not complete under the applied experimental conditions. According to previous literature and previous experience of our laboratory on similar systems [43], it may be proposed that the resulting gold fragments will bind preferentially to solvent-exposed cysteines of the M^{pro}. Instead, in accord with the less reactive character of the bis-gold(I) carbene complex, no metalation of the protein was observed upon reaction with $[Au(NHC)_2]PF_6$ even after 24h of incubation, and just the free protein signal was invariably detected (Figure S2). This confirms that $[Au(NHC)_2]PF_6$ is poorly reactive with this protein in line with its scarce enzyme-inhibiting activity.

4.4. Crystallographic Results

Remarkably, for two of the above-mentioned gold drugs, i.e., Au(PEt_3)Br and Au(NHC)Cl, we succeeded in obtaining good quality crystals of their adducts with M^{pro} and in solving the respective 3D crystal structures. The polypeptide structure of both adducts is totally superimposable with the others already deposited on the PDB except for the last five C-terminal residues that have not been modeled due to their very weak electron density. The structure of both SARS-CoV-2 M^{pro} adducts clearly resembles that of the apo protein (PDBID 7NXH) with a negligible RMSD computed on backbone atoms. Only local loop regions showed backbone RMSD values higher than average (Figure 4).

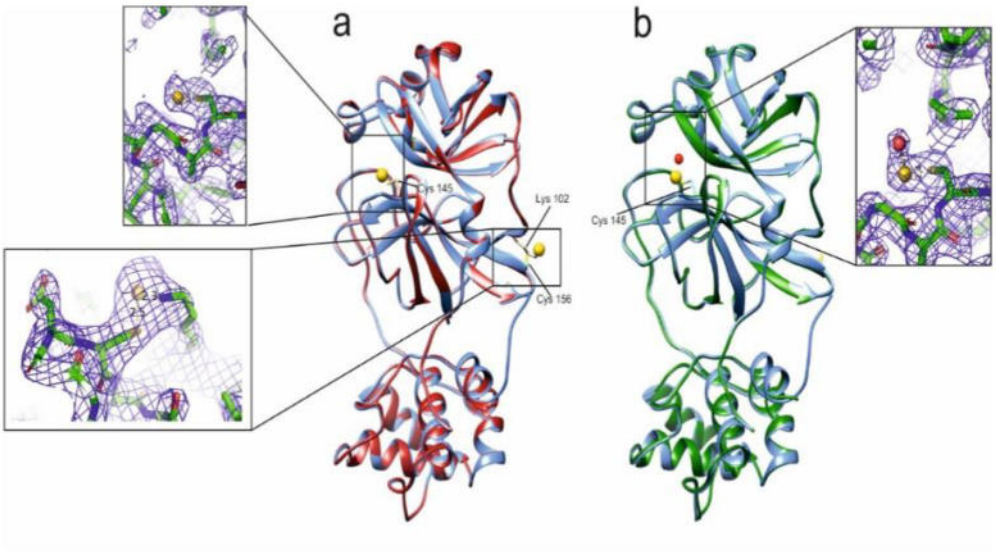


Figure 4. Overlays of (**a**) apo M^{pro} (cyan), (PDBID 7NXH), with M^{pro}-Au(PEt_3)Br adduct (red) and (**b**) apo M^{pro} (cyan) with M^{pro}- Au(NHC)Cl adduct (dark green). Gold ion is shown as yellow sphere and the water molecule as red sphere. Cys 145, Cys 156 and Lys 102 are shown as sticks. Insets show the 2Fo-Fc electron density contoured at 1s level of Au(PEt_3)Br Cys145, Au(PEt_3)Br Cys156 and Au(NHC)Cl Cys145 coordination environments.

To the best of our knowledge, the structures of only two M^{pro}/gold adducts are available on the PDB (7DAT and 7DAU [22]). In the 7DAT structure, M^{pro} has been soaked with Auranofin; the ionic gold targets the catalytic Cys145 with an occupancy of about 0.33 and Cys156 with an occupancy of 0.16 (Figure S3). In the 7DAU structure, ionic gold is bound to both Cys145 and Cys156 with occupancy values of 0.2 and 0.1, respectively. The distance between gold and sulfur in 7DAU is around 2.3 Å. The behavior of each of the two gold adducts analyzed in this work is slightly different, but they share a quite low gold occupancy, with the organic moiety originally attached to gold being no more visible in all the structures. In the case of Au(PEt_3)Br, both Cys145 and Cys156 are involved in the binding of gold (7DAT and 7DAU structures, Figure 4A and S3). The electron density of gold bound to Cys156 is quite well defined with an occupancy of about 0.4 and a B-factor in the range of the solvent atoms (Inset of Figure 4A). On the contrary, the gold ion bound to Cys145 has a weaker electron density due to a significantly lower occupancy that is as low as about 0.15. Concerning the coordination geometry, Cys145 binds gold in a monodentate fashion whereas, in the case of Cys156, gold shows a peculiar binding mode to the cysteine sulfur and to the terminal nitrogen of the Lys102 sidechain (insets of Figure 4A).

In the case of Au(NHC)Cl, only Cys145 is involved in the binding to gold but still with an occupancy of about 0.15, and a gold-sulfur distance of about 2.4 Å, which is in line with the expected values. At variance with Auranofin, a labile water molecule coordinates the gold ion in this complex (Figure 4B).

5. Conclusions

In conclusion, we have shown that Auranofin and its halido analogs behave as potent and effective inhibitors of the M^{pro} of SARS-CoV-2. Furthermore, while Au(NHC)Cl results to have inhibitory capabilities comparable to the Auranofin analogues, the gold dicarbene complex $[Au(NHC)_2]PF_6$ was found to inhibit the M^{pro} rather weakly in accordance with the mass data offering no evidence of adduct formation with the protein. The peculiar behavior of $[Au(NHC)_2]PF_6$ may be ascribed to the great strength of the two Au-

C bonds and to the intrinsic difficulties in breaking them. Notably, the measured inhibition constants for Auranofin, for its three halido derivatives, and for AuNHC all fall in the low micromolar range. As SARS-CoV-2 M^{pro} is a validated druggable target of the coronavirus, these strong inhibitory properties suggest that the gold compounds studied in this work may be prospective anti-COVID-19 agents. ESI MS measurements point out that relatively stable metallodrug-protein adducts are indeed formed when mixing the protein with the various metal compounds through the coordination of gold-containing fragments to accessible protein residues in accordance with previous studies on similar systems; clear evidence is offered for the formation of adducts bearing one or two protein-bound gold fragments. In two cases, i.e., the protein adducts with $Au(PEt_3)Br$ and Au(NHC)Cl, high-resolution crystal structures have been obtained allowing unambiguous localization of the gold binding sites. Crystal structures show that Cys 156 and the catalytic Cys 145 are the main sites of protein metalation. Overall, the here reported results coherently support the view that the potent and promising inhibition of the catalytic activity of M^{pro} produced by these gold compounds is the likely consequence of the tight coordinative binding of a gold-containing fragment to the catalytic Cys 145. As a final consideration, it has to be pointed out that there are currently no in vivo experimental data concerning the ability of the gold compounds described in this paper to inhibit the SARS-CoV-2 virus. Within this frame, a possible future development of our study would be to explore the ability of such gold compounds to contrast SARS-CoV-2 infection in in vivo models and to compare their mechanism of action and their pharmacological efficacy with that of the reference gold drug Auranofin.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/biom12111675/s1, Figure S1: Fluorescence curves of M^{pro} titration with (a) Auranofin (b) Au(NHC)Cl (c) $Au(PEt_3)Br$ (d) $Au(PEt_3)I$ (e) $Au(PEt_3)Cl$; Figure S2: Activity assay raw data showing the absence of inhibition by $[Au(NHC)_2]PF_6$ and ESI-MS spectrum of a 1:1 M^{pro}: $[Au(NHC)_2]PF_6$ solution showing no adduct formation; Figure S3: Superposition between $Au(PEt_3)Br$ red, 7DAT (Auranofin) magenta, Au(NHC)Cl green. Gold ions are represented as spheres and colored in yellow. Cys 145 and Cys156 are shown as sticks; Table S1: Data collection and refinement parameters.

Author Contributions: Data curation, L.M. (Lara Massai), D.G. and A.G.; funding acquisition, L.B.; investigation, L.M. (Lara Massai), D.G., A.D.S., A.G. and V.C.; Supervision, F.C., V.C. and L.M. (Luigi Messori); visualization, A.D.S.; writing—original draft, L.M. (Lara Massai), D.G., A.G. and L.M. (Luigi Messori); writing—review & editing, F.C., V.C. and L.B. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge the support by the Italian Ministry for University and Research (FOE funding) to the CERM/CIRMMP Italian Centre of Instruct-ERIC, a landmark ESFRI project. DG is a PhD student supported under the EOSC-Life Project (ID:824087) funded by the Horizon 2020 EC program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Coordinates and structure factors of the two crystallographic structures described in the paper are publicly available at the Protein Data Bank under the accession codes 8B0T for the $Au(PEt_3)Br$ and 8B0S for Au(NHC)Cl.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

DTT (dithiothreitol), EDTA (ethylenediaminetetraacetic acid), DMSO (Dimethylsulfoxide), PEG (polyethyleneglycole), ESI MS (Electrospray ionization mass spectrometry), PDB (protein data bank), RMSD (root mean square deviation)

References

1. Mullen, L.; Potter, C.; Gostin, L.O.; Cicero, A.; Nuzzo, J.B. An analysis of International Health Regulations Emergency Committees and Public Health Emergency of International Concern Designations. *BMJ Glob. Health* **2020**, *5*, e002502. https://doi.org/10.1136/bmjgh-2020-002502.
2. Cheudjeu, A. The SARS-CoV-2 Entry Inhibition Mechanisms of Serine Protease Inhibitors, OM-85, Heparin and Soluble HS Might Be Linked to HS Attachment Sites. *Molecules* **2022**, *27*, 1947. https://doi.org/10.3390/molecules27061947.
3. Xiang, R.; Yu, Z.; Wang, Y.; Wang, L.; Huo, S.; Li, Y.; Liang, R.; Hao, Q.; Ying, T.; Gao, Y.; et al. Recent advances in developing small-molecule inhibitors against SARS-CoV-2. *Acta Pharm. Sinica. B* **2022**, *12*, 1591–1623. https://doi.org/10.1016/j.apsb.2021.06.016.
4. Zhang, C.H.; Stone, E.A.; Deshmukh, M.; Ippolito, J.A.; Ghahremanpour, M.M.; Tirado-Rives, J.; Spasov, K.A.; Zhang, S.; Takeo, Y.; Kudalkar, S.N.; et al. Potent Noncovalent Inhibitors of the Main Protease of SARS-CoV-2 from Molecular Sculpting of the Drug Perampanel Guided by Free Energy Perturbation Calculations. *ACS Cent. Sci.* **2021**, *7*, 467–475. https://doi.org/10.1021/acscentsci.1c00039.
5. Rothan, H.A.; Stone, S.; Natekar, J.; Kumari, P.; Arora, K.; Kumar, M. The FDA-approved gold drug auranofin inhibits novel coronavirus (SARS-CoV-2) replication and attenuates inflammation in human cells. *Virology* **2020**, *547*, 7–11. https://doi.org/10.1016/j.virol.2020.05.002.
6. Sonzogni-Desautels, K.; Ndao, M. Will Auranofin Become a Golden New Treatment Against COVID-19? *Front. Immunol.* **2021**, *12*, 683694. https://doi.org/10.3389/fimmu.2021.683694.
7. Hu, Q.; Xiong, Y.; Zhu, G.H.; Zhang, Y.N.; Zhang, Y.W.; Huang, P.; Ge, G.B. The SARS-CoV-2 main protease (M(pro)): Structure, function, and emerging therapies for COVID-19. *MedComm* **2022**, *3*, e151. https://doi.org/10.1002/mco2.151.
8. Ho, B.L.; Cheng, S.C.; Shi, L.; Wang, T.Y.; Ho, K.I.; Chou, C.Y. Critical Assessment of the Important Residues Involved in the Dimerization and Catalysis of MERS Coronavirus Main Protease. *PLoS ONE* **2015**, *10*, e0144865. https://doi.org/10.1371/journal.pone.0144865.
9. Arutyunova, E.; Khan, M.B.; Fischer, C.; Lu, J.; Lamer, T.; Vuong, W.; van Belkum, M.J.; McKay, R.T.; Tyrrell, D.L.; Vederas, J.C.; et al. N-Terminal Finger Stabilizes the S1 Pocket for the Reversible Feline Drug GC376 in the SARS-CoV-2 M(pro) Dimer. *J. Mol. Biol.* **2021**, *433*, 167003. https://doi.org/10.1016/j.jmb.2021.167003.
10. Ziebuhr, J.; Snijder, E.J.; Gorbalenya, A.E. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J. Gen. Virol.* **2000**, *81*, 853–879. https://doi.org/10.1099/0022-1317-81-4-853.
11. Malone, B.; Urakova, N.; Snijder, E.J.; Campbell, E.A. Structures and functions of coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug design. *Nat. Rev. Mol. Cell Biol.* **2022**, *23*, 21–39. https://doi.org/10.1038/s41580-021-00432-z.
12. Ullrich, S.; Nitsche, C. The SARS-CoV-2 main protease as drug target. *Bioorg Med. Chem Lett* **2020**, *30*, 127377. https://doi.org/10.1016/j.bmcl.2020.127377.
13. Lee, J.T.; Yang, Q.; Gribenko, A.; Perrin, B.S., Jr.; Zhu, Y.; Cardin, R.; Liberator, P.A.; Anderson, A.S.; Hao, L. Genetic Surveillance of SARS-CoV-2 M(pro) Reveals High Sequence and Structural Conservation Prior to the Introduction of Protease Inhibitor Paxlovid. *mBio* **2022**, *13*, e0086922. https://doi.org/10.1128/mbio.00869-22.
14. Mengist, H.M.; Dilnessa, T.; Jin, T. Structural Basis of Potential Inhibitors Targeting SARS-CoV-2 Main Protease. *Front. Chem.* **2021**, *9*, 622898. https://doi.org/10.3389/fchem.2021.622898.
15. Yang, H.; Xie, W.; Xue, X.; Yang, K.; Ma, J.; Liang, W.; Zhao, Q.; Zhou, Z.; Pei, D.; Ziebuhr, J.; et al. Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol.* **2005**, *3*, e324. https://doi.org/10.1371/journal.pbio.0030324.
16. Grifagni, D.; Calderone, V.; Giuntini, S.; Cantini, F.; Fragai, M.; Banci, L. SARS-CoV-2 M(pro) inhibition by a zinc ion: Structural features and hints for drug design. *Chem. Commun.* **2021**, *57*, 7910–7913. https://doi.org/10.1039/d1cc02956h.
17. Panchariya, L.; Khan, W.A.; Kuila, S.; Sonkar, K.; Sahoo, S.; Ghoshal, A.; Kumar, A.; Verma, D.K.; Hasan, A.; Khan, M.A.; et al. Zinc(2+) ion inhibits SARS-CoV-2 main protease and viral replication in vitro. *Chem. Commun.* **2021**, *57*, 10083–10086. https://doi.org/10.1039/d1cc03563k.
18. Tao, X.; Zhang, L.; Du, L.; Lu, K.; Zhao, Z.; Xie, Y.; Li, X.; Huang, S.; Wang, P.H.; Pan, J.A.; et al. Inhibition of SARS-CoV-2 replication by zinc gluconate in combination with hinokitiol. *J. Inorg. Biochem.* **2022**, *231*, 111777. https://doi.org/10.1016/j.jinorgbio.2022.111777.
19. Rani, I.; Goyal, A.; Bhatnagar, M.; Manhas, S.; Goel, P.; Pal, A.; Prasad, R. Potential molecular mechanisms of zinc- and copper-mediated antiviral activity on COVID-19. *Nutr. Res.* **2021**, *92*, 109–128. https://doi.org/.
20. Gil-Moles, M.; Basu, U.; Büssing, R.; Hoffmeister, H.; Türck, S.; Varchmin, A.; Ott, I. Gold Metallodrugs to Target Coronavirus Proteins: Inhibitory Effects on the Spike-ACE2 Interaction and on PLpro Protease Activity by Auranofin and Gold Organometallics. *Chemistry* **2020**, *26*, 15140–15144. https://doi.org/10.1002/chem.202004112.
21. Abate, C.; Carnamucio, F.; Giuffre, O.; Foti, C. Metal-Based Compounds in Antiviral Therapy. *Biomolecules* **2022**, *12*, 933. https://doi.org/10.3390/biom12070933.
22. He, Z.; Ye, F.; Zhang, C.; Fan, J.; Du, Z.; Zhao, W.; Yuan, Q.; Niu, W.; Gao, F.; He, B.; et al. A comparison of Remdesivir versus gold cluster in COVID-19 animal model: A better therapeutic outcome of gold cluster. *Nano Today* **2022**, *44*, 101468. https://doi.org/10.1016/j.nantod.2022.101468.
23. Karges, J.; Cohen, S.M. Metal Complexes as Antiviral Agents for SARS-CoV-2. *ChemBioChem* **2021**, *22*, 2600–2607. https://doi.org/10.1002/cbic.202100186.

24. Marzo, T.; Cirri, D.; Pollini, S.; Prato, M.; Fallani, S.; Cassetta, M.I.; Novelli, A.; Rossolini, G.M.; Messori, L. Auranofin and its Analogues Show Potent Antimicrobial Activity against Multidrug-Resistant Pathogens: Structure-Activity Relationships. *ChemMedChem* **2018**, *13*, 2448–2454. https://doi.org/10.1002/cmdc.201800498.
25. Massai, L.; Messori, L.; Carpentieri, A.; Amoresano, A.; Melchiorre, C.; Fiaschi, T.; Modesti, A.; Gamberi, T.; Magherini, F. The effects of two gold-N-heterocyclic carbene (NHC) complexes in ovarian cancer cells: A redox proteomic study. *Cancer Chemother. Pharmacol.* **2022**, *89*, 809–823. https://doi.org/10.1007/s00280-022-04438-y.
26. Massai, L.; Cirri, D.; Marzo, T.; Messori, L. Auranofin and its analogs as prospective agents for the treatment of colorectal cancer. *Cancer Drug Resist.* **2022**, *5*, 1–14. https://doi.org/10.20517/cdr.2021.71.
27. Messori, L.; Marchetti, L.; Massai, L.; Scaletti, F.; Guerri, A.; Landini, I.; Nobili, S.; Perrone, G.; Mini, E.; Leoni, P.; et al. Chemistry and biology of two novel gold(I) carbene complexes as prospective anticancer agents. *Inorg. Chem.* **2014**, *53*, 2396–2403. https://doi.org/10.1021/ic401731a.
28. Marzo, T.; Cirri, D.; Gabbiani, C.; Gamberi, T.; Magherini, F.; Pratesi, A.; Guerri, A.; Biver, T.; Binacchi, F.; Stefanini, M.; et al. Auranofin, Et(3)PAuCl, and Et(3)PAuI Are Highly Cytotoxic on Colorectal Cancer Cells: A Chemical and Biological Study. *ACS Med. Chem. Lett.* **2017**, *8*, 997–1001. https://doi.org/10.1021/acsmedchemlett.7b00162.
29. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132.
30. Vagin, A.A.; Teplyakov, A. An approach to multi-copy search in molecular replacement. *Acta Crystallogr D Biol. Crystallogr.* **2000**, *56*, 1622–1624.
31. Adams, P.D.; Afonine, P.V.; Bunkòczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.-W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 213–221.
32. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486–501.
33. Chen, V.B.; Arendall, W.B., III; Headd, J.J.; Keedy, D.A.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 12–21.
34. Alegria-Arcos, M.; Barbosa, T.; Sepulveda, F.; Combariza, G.; Gonzalez, J.; Gil, C.; Martinez, A.; Ramirez, D. Network pharmacology reveals multitarget mechanism of action of drugs to be repurposed for COVID-19. *Front. Pharmacol.* **2022**, *13*, 952192. https://doi.org/10.3389/fphar.2022.952192.
35. Gamberi, T.; Chiappetta, G.; Fiaschi, T.; Modesti, A.; Sorbi, F.; Magherini, F. Upgrade of an old drug: Auranofin in innovative cancer therapies to overcome drug resistance and to increase drug effectiveness. *Med. Res. Rev.* **2022**, *42*, 1111–1146. https://doi.org/10.1002/med.21872.
36. Liu, Y.; Lu, Y.; Xu, Z.; Ma, X.; Chen, X.; Liu, W. Repurposing of the gold drug auranofin and a review of its derivatives as antibacterial therapeutics. *Drug Discov. Today* **2022**, *27*, 1961–1973. https://doi.org/10.1016/j.drudis.2022.02.010.
37. Tabassum, S.F.Z.; Ur-Rahman, M.S.; Hussain, M.A.; Etheshaam uz Zaman, U.S.M.; Koneru, A. Drug Information of Auranofin and Its Effectiveness in COVID-19. *IOSR J. Pharm.* **2021**, *11*, 29–37.
38. Laplantine, E.; Chable-Bessia, C.; Oudin, A.; Swain, J.; Soria, A.; Merida, P.; Gourdelier, M.; Mestiri, S.; Besseghe, I.; Bremaud, E.; et al. The FDA-approved drug Auranofin has a dual inhibitory effect on SARS-CoV-2 entry and NF-kappaB signaling. *iScience* **2022**, *25*, 105066. https://doi.org/10.1016/j.isci.2022.105066.
39. Cirri, D.; Schirmeister, T.; Seo, E.J.; Efferth, T.; Massai, L.; Messori, L.; Micale, N. Antiproliferative Properties of a Few Auranofin-Related Gold(I) and Silver(I) Complexes in Leukemia Cells and their Interferences with the Ubiquitin Proteasome System. *Molecules* **2020**, *25*, 4454. https://doi.org/10.3390/molecules25194454.
40. Tialiou, A.; Chin, J.; Keppler, B.K.; Reithofer, M.R. Current Developments of N-Heterocyclic Carbene Au(I)/Au(III) Complexes toward Cancer Treatment. *Biomedicines* **2022**, *10*, 1417. https://doi.org/10.3390/biomedicines10061417.
41. Wenzel, M.; Casini, A. Mass spectrometry as a powerful tool to study therapeutic metallodrugs speciation mechanisms: Current frontiers and perspectives. *Coord. Chem. Rev.* **2017**, *352*, 432–460. https://doi.org/10.1016/j.ccr.2017.02.012.
42. Zoppi, C.; Massai, L.; Cirri, D.; Gabbiani, C.; Pratesi, A.; Messori, L. Protein metalation by two structurally related gold(I) carbene complexes: An ESI MS study. *Inorg. Chim. Acta* **2021**, *520*, 120297. https://doi.org/10.1016/j.ica.2021.120297.
43. Zoppi, C.; Messori, L.; Pratesi, A. ESI MS studies highlight the selective interaction of Auranofin with protein free thiols. *Dalton Trans.* **2020**, *49*, 5906–5913. https://doi.org/10.1039/d0dt00283f.
44. Ugone, V.; Sanna, D.; Sciortino, G.; Crans, D.C.; Garribba, E. ESI-MS Study of the Interaction of Potential Oxidovanadium(IV) Drugs and Amavadin with Model Proteins. *Inorg. Chem.* **2020**, *59*, 9739–9755. https://doi.org/10.1021/acs.inorgchem.0c00969.
45. Merlino, A.; Marzo, T.; Messori, L. Protein Metalation by Anticancer Metallodrugs: A Joint ESI MS and XRD Investigative Strategy. *Chemistry* **2017**, *23*, 6942–6947. https://doi.org/10.1002/chem.201605801.
46. Hartinger, C.G.; Tsybin, Y.O.; Fuchser, J.; Dyson, P.J. Characterization of platinum anticancer drug protein-binding sites using a top-down mass spectrometric approach. *Inorg.Chem.* **2008**, *47*, 17–19. https://doi.org/10.1021/ic702236m.

Gold metallodrugs to target coronavirus proteins: the inhibition of SARS CoV-2 main protease by Auranofin and its analogs.

Lara Massai[*1], Deborah Grifagni[*1,2], Alessia De Santis[1,2], Andrea Geri[1], Francesca Cantini[1,2,3,±], Vito Calderone[1,2,3±], Lucia Banci[1,2,3], Luigi Messori[1±]

1 Department of Chemistry "Ugo Schiff", University of Florence, Via della Lastruccia 3 50019 Florence, Italy
2 Magnetic Resonance Center (CERM), University of Florence, Via L. Sacconi 6 50019 Sesto Fiorentino, Italy
3 Consorzio Interuniversitario Risonanze Magnetiche Metallo Proteine (CIRMMP), University of Florence, Via L. Sacconi 6 50019 Florence, Italy;

Correspondence to:

[±] Francesca Cantini: cantini@cerm.unifi.it

ORCID: 0000-0003-0526-6732

[±]Vito Calderone: calderone@cerm.unifi.it

ORCID: 0000-0002-7963-6241

[±]Luigi Messori: luigi.messori@unifi.it

ORCID: 0000-0002-9490-8014

This PDF file includes:

Figure S1 to S3

Table S1

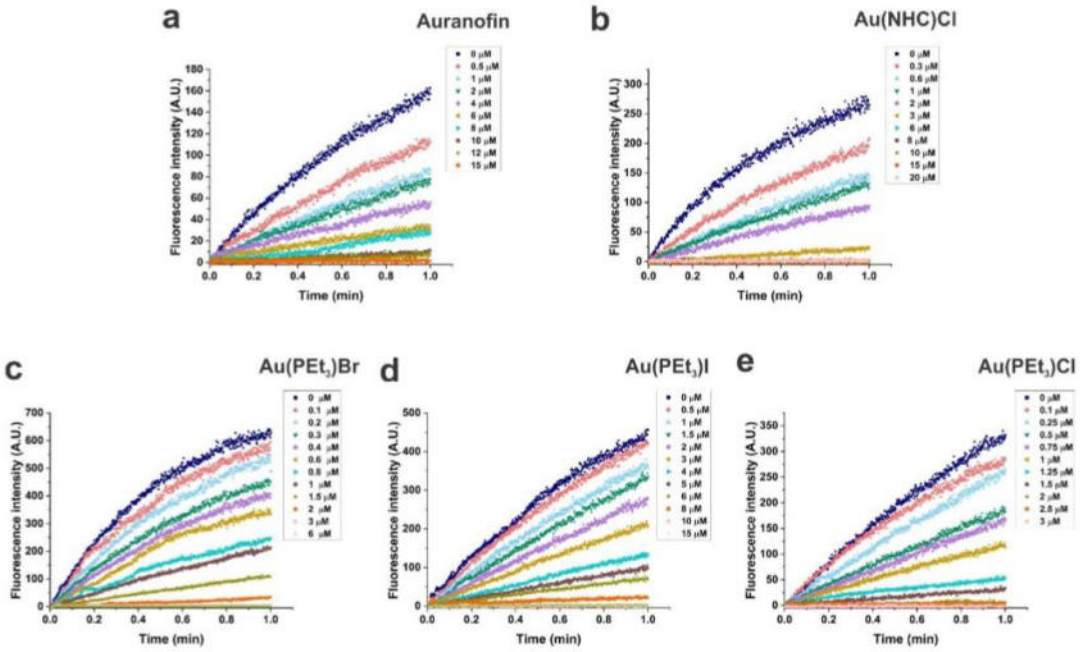


Figure S1. Fluorescence curves of M^{pro} titration with a) Auranofin b) Au(NHC)Cl c) $Au(Pet_3)Br$ d) $Au(PEt_3)I$ e) $Au(Pet_3)Cl$.

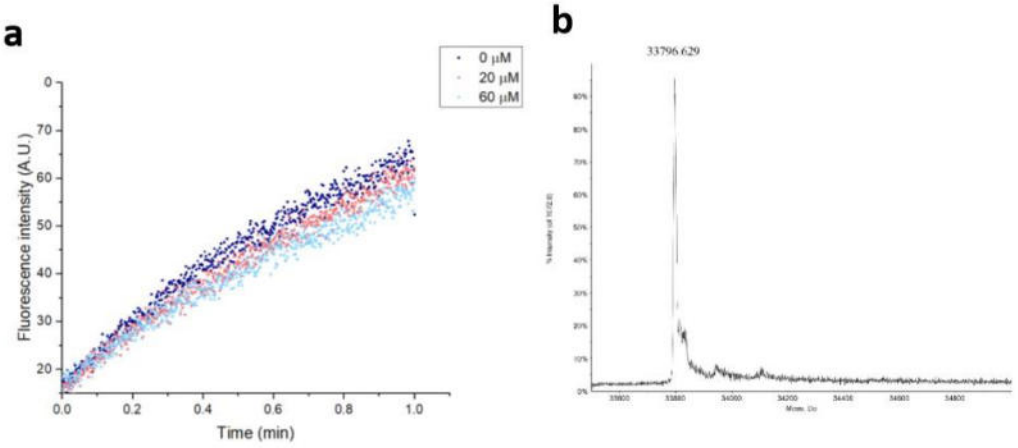


Figure S2. Activity assay raw data showing the absence of inhibition by $[Au(NHC)_2]PF_6$ and ESI-MS spectrum of a 1:1 M^{pro}: $[Au(NHC)_2]PF_6$ solution showing no adduct formation.

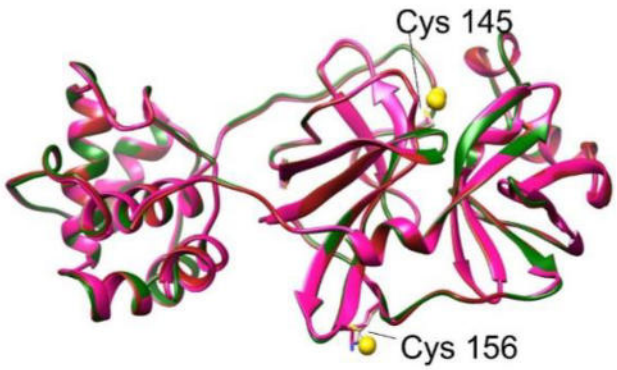


Figure S3. Superposition between Au(PEt3)Br red, 7DAT (Auranofin) magenta, Au(NHC)Cl green. Gold ions are represented as spheres and colored in yellow. Cys 145 and Cys156 are shown as sticks.

Table S1. Data collection and refinement parameters.

	Au(PEt3)Br	Au(NHC)Cl
Wavelength	1.541	1.541
Resolution range	26.19 - 2.41 (2.496 - 2.410)	38.4 - 2.42 (2.506 - 2.42)
Space group	C 1 2 1	C 1 2 1
Unit cell	114.34 53.65 44.84 90 102.1 90	111.92 53.27 44.55 90 102.85 90
Unique reflections	9801 (944)	9385 (741)
Multiplicity	3.3 (3.1)	3.0 (1.2)
Completeness (%)	94.18 (91.83)	94.52 (74.65)
Mean I/sigma(I)	5.95 (1.05)	5.32 (1.01)
Wilson B-factor	52.65	43.68
R-merge	0.1862 (0.765)	0.1671 (0.5823)
CC1/2	0.981 (0.433)	0.976 (0.657)
R-work	0.2471 (0.3621)	0.2538 (0.3575)
R-free	0.2751 (0.3519)	0.2726 (0.2996)
Number of non-hydrogen atoms	2346	2348
macromolecules	2329	2329
ligands	2	1
solvent	15	18
Protein residues	301	301
RMSD (bonds)	0.002	0.003
RMSD (angles)	0.51	0.57
Ramachandran outliers (%)	0	0
Average B-factors	56.03	44.78
macromolecule	56.07	44.83
ligands	58.96	45.09
solvent	47.57	37.87

3.2. - Characterization of SARS-CoV-2 3CL^{Pro} and CVB3 3C^{Pro} with KME16 and ACH213 PROTACs

KME16 and ACH213 PROTACs are bifunctional molecules able to recognize the active site of SARS-CoV-2 3CL^{Pro} and CVB3 3C^{Pro} proteases from one side (PROTAC precursor) and to select the E3 ubiquitin ligase Cereblon (CRBN) from the other side. The synthesis of PROTAC molecules has been conducted in the laboratory of Prof. Trabocchi (DICUS, UNIFI). In the synthesis of KME16 and ACH213 PROTACs, the warhead was chosen for its similarity to high-affinity SARS-CoV-2 3CL^{Pro} inhibitors such as Nirmatrelvir [136] and Ibuzatrelvir [137], that both contain a nitrile group (**Figure 24**). The warhead was conjugated to a pomalidomide CRBN ligand through a piperidine linker.

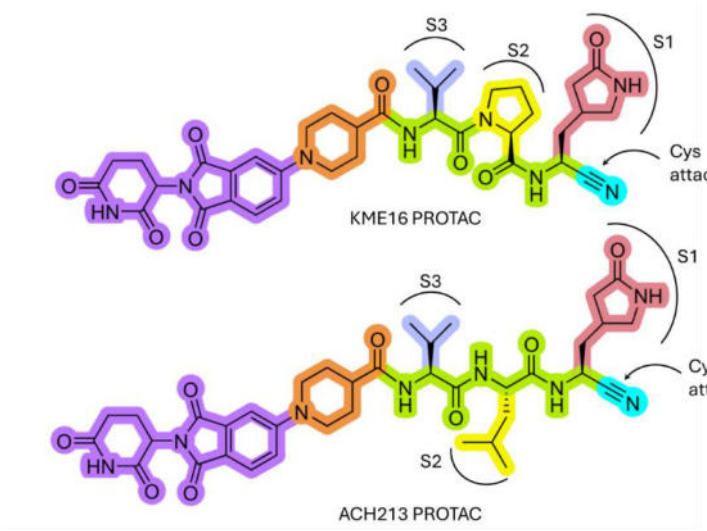


Figure 24. Molecular structure of KME16 and ACH213 PROTACs. The pomalidomide moiety, the piperidine linker and the warhead backbone of PROTACs are represented in violet, in orange and green, respectively. The functional chemical groups of PROTAC warhead are depicted in different colours: the isopropyl group in indigo, the pyrrolidine and leucine side chain groups in yellow, the γ-lactam ring in dark pink and the nitrile group in light blue. The protein subsites are marked for typical recognition of specific functional groups.

3.2.1. - SARS-CoV-2 3CL^{Pro} ^{1}H-^{15}N HSQC NMR Titrations experiments

^{15}N-isotopically labelled SARS-CoV-2 3CL^{Pro} was titrated with KME16 and ACH213 PROTACs and their interactions monitored through ^{1}H-^{15}N HSQC experiments. To perform the titration with

both molecules, multiple ^{1}H-^{15}N HSQC experiments using a 950 MHz spectrometer (Bruker) were acquired at 298 K. 500 µL of ^{15}N $3CL^{Pro}$ solution in the purification buffer (25 mM HEPES, 75 mM NaCl, 5 mM TCEP, 1 mM DTT pH 6.5) with 50 µL of D_2O was placed in a 5mm tube. At first a ^{1}H-^{15}N HSQC of the protein solution was acquired as a starting point. Subsequently, stock solutions of KME16 and ACH213, dissolved in acetonitrile-d_3, were used to set the titration. Starting with additions of 2 µL of the PROTACs stock solutions up to a total addition of 20 µL, we reached a 1:5 equivalent ratio between the protein and the PROTAC molecules. The NMR spectra processing was performed with TopSpin 3.6.1.

Upon subsequent additions of PROTACs, chemical shift changes were observed in the $^{1}H–^{15}N$ HSQC maps of SARS-CoV-2 $3CL^{Pro}$, corresponding to a slow/intermediate exchange regime on the NMR time scale. On reaching the 1:5 SARS-CoV-2 $3CL^{Pro}$-KME16 or ACH213 protein-molecule molar ratio, the signals of the free $3CL^{Pro}$ completely disappeared, while the signals of the protein-PROTAC complexes reached their maximal intensity, indicating that the latter were fully formed at the 1:5 molar ratio (**Figure 25**). However, during the titration of $3CL^{Pro}$ with ACH213 some protein precipitation was observed, which may have caused an overall decrease in quality of the NMR spectra.

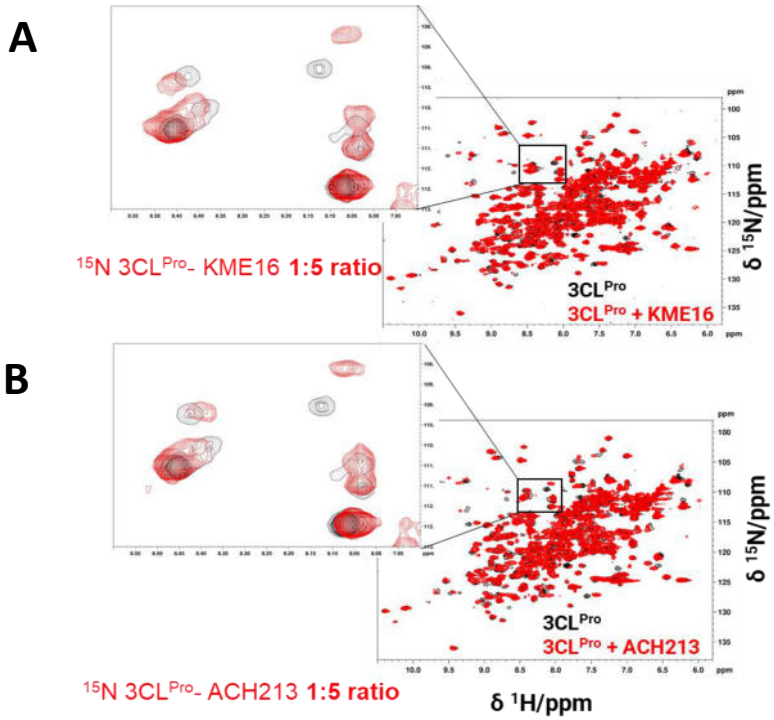


Figure 25. Overlay of the ^{1}H-^{15}N HSQC spectrum of SARS-CoV-2 $3CL^{Pro}$ (black) and after (red) the addition of 5 equivalents of the KME16 (**A**) or ACH213 (**B**) PROTACs, stock solutions dissolved in acetonitrile-d_3. The spectra were recorded at 298K, 950 MHz spectrometer (Bruker). In the insets of panel

A and B, NH signals in a slow exchange regime of the NMR time scale are shown at 0 (black) and 5 (red) equivalents of the KME16 and ACH213 PROTACs additions.

The observed chemical shift changes were calculated as backbone-weighted average chemical shift differences, i.e., $\Delta\delta_{avg}$(HN), i.e., $(((\Delta H)^2 + (\Delta N/5)^2)/2)^{1/2}$, where ΔH and ΔN are chemical shift differences for backbone amide ^{1}H and ^{15}N nuclei, respectively. The estimation of the chemical shift threshold value for defining meaningful chemical shift differences was obtained by averaging $\Delta\delta_{avg}$(HN) values plus one standard deviation (1δ). Chemical shift assignments of SARS CoV-2 3CL^{Pro} (under accession code BMRB: BMR50780) [347] and of the C145A variant of SARS-CoV-2 3CL^{Pro} (under accession code BMRB: BMR51456) [348] were used to map the chemical shift changes.

Most of the affected residues were found close to the Cys-His catalytic dyad of SARS-CoV-2 3CL^{Pro} thus revealing that KME16 and ACH213 PROTACs both interact with the enzyme through their warhead moiety (**Figure 26**).

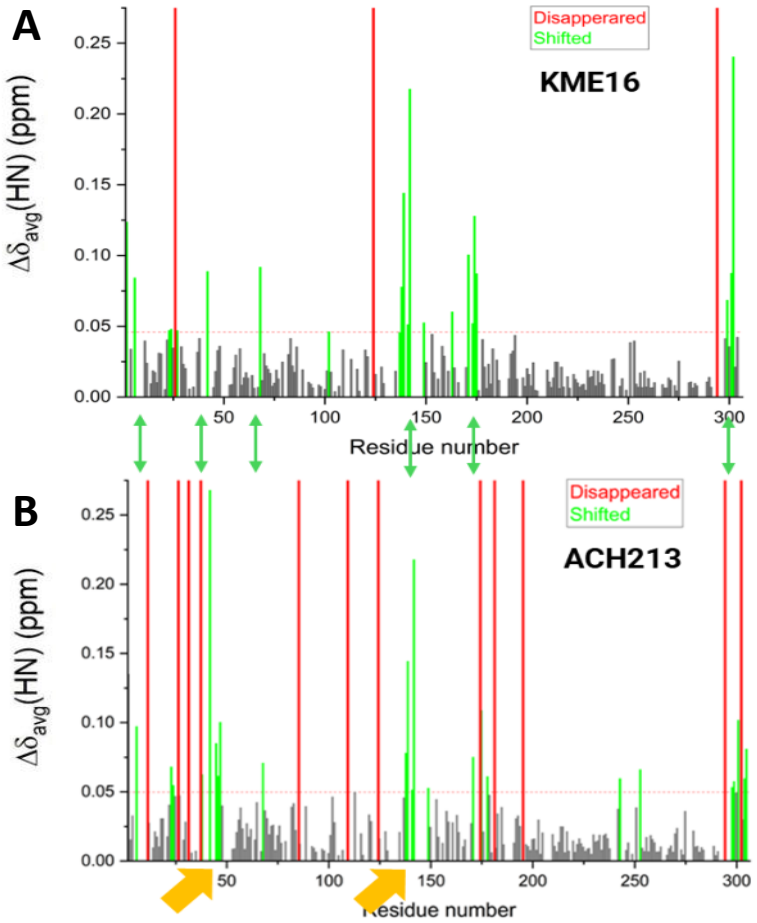


Figure 26. Backbone average chemical shift differences ($\Delta\delta_{avg}$(HN)) between ^{15}N-labelled SARS-CoV-2 3CL^{Pro} and its 1:5 mixture with KME16 (**A**) and ACH213 (**B**) PROTACs. A chemical shift threshold

value of 0.05 ppm, indicated as a dashed line in both panels, was estimated to define the most significant chemical shift differences. The red bars identify the disappeared residues, the green those with a $\Delta\delta_{avg}$(HN) higher than the threshold. The yellow arrows indicate the catalytic residues of the enzyme. The regions affected by significant chemical shift changes in the two complexes $3CL^{Pro}$ -KME16 and $3CL^{Pro}$-ACH213 are shown by green double-headed arrows. The plots were generated by OriginLab software (Origin 2022).

3.2.2. - CVB3 $3C^{Pro}$ ^{1}H-^{15}N HSQC NMR Titrations experiments

^{15}N-isotopically labelled CVB3 $3C^{Pro}$ was titrated with KME16 and ACH213 PROTACs and their interactions monitored through ^{1}H-^{15}N HSQC experiments. To perform the titration with both molecules, multiple ^{1}H-^{15}N HSQC experiments using a 950MHz spectrometer (Bruker) were acquired at 308 K. 500 µL of ^{15}N $3C^{Pro}$ solution in the purification buffer (10 mM HEPES, 300 mM NaCl, 1 mM EDTA and 1 mM DTT pH 7.5) with 50 µL of D_2O was placed in a 5 mm tube. At first a ^{1}H-^{15}N HSQC of the protein solution was acquired as a starting point. Subsequently, stock solutions dissolved in acetonitrile-d_3 of KME16 and ACH213 were used to set the titration, starting with additions of 2 µL of the PROTACs stock solutions up to a total addition of 20 µL to reach a 1:5 equivalent ratio between the protein and the PROTAC molecules. The NMR spectra processing was performed with TopSpin 3.6.1.

Upon subsequent additions of PROTACs, chemical shift changes were observed in the ^{1}H-^{15}N HSQC maps of ^{15}N CVB3 $3C^{Pro}$, corresponding to a slow exchange regime on the NMR time scale. However, the signals of the free $3C^{Pro}$ did not completely disappear when the signals of the protein-PROTAC complexes reached their maximal intensity, indicating that the latter were not fully formed at the 1:5 protein-molecule molar ratio (**Figure 27**).

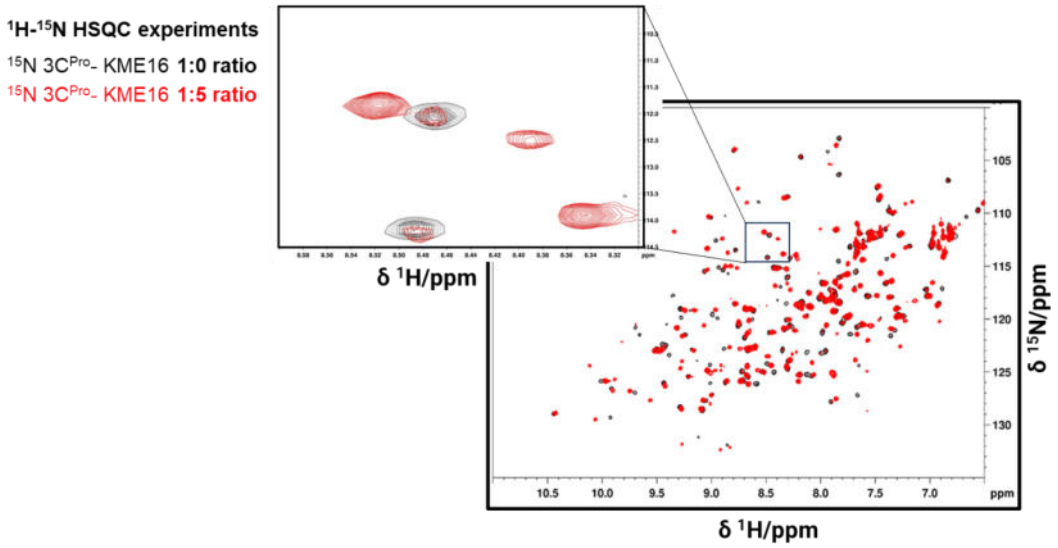


Figure 27. Overlay of the ^{1}H-^{15}N HSQC spectrum of CVB3 $3C^{Pro}$ (black) and after (red) the addition of 5 equivalents of KME16 PROTAC, stock solution dissolved in acetonitrile-d_3. The spectra were recorded at 298K, 950 MHz spectrometer (Bruker). In the insets of panel A and B, NH signals in a slow exchange regime of the NMR time scale are shown at 0 (black) and 5 (red) equivalents of the KME16 PROTACs additions.

Even at a 1:5 molar ratio, whereas complete complex formation was observed with SARS-CoV-2 $3CL^{Pro}$, the chemical shift data in this case suggest a reduced affinity of the KME16 and ACH213 PROTACs for CVB3 $3C^{Pro}$.

3.3. - Investigation of KME16 PROTAC ternary complex

The subsequent aim in this study was the investigation of Cereblon (CRBN) E3 ubiquitin ligase to characterize the ternary complex target protein:PROTAC:CRBN E3 ubiquitin ligase. The initial key steps involved the expression and further purification of the CRBN E3 ligase.

3.3.1. - CRBN-TBD E3 ubiquitin ligase production

Since the high molecular weight of full-length CRBN (~50 kDa) and the related limitations about its expression in bacteria [349], we decided to study the CRBN domain, called Cereblon-Thalidomide Binding Domain (CRBN-TBD), involved in the binding with the pomalidomide moiety of our PROTAC molecules. There are few protocols in literature that report the expression of this domain in E. coli and few available plasmids. We purchased an expression vector on Addgene from Knapp laboratory. This plasmid called pGTVL2-CRBN-TBD allows the expression

of the CRBN-TBD as a fusion protein with an N-terminal His-tag adjacent to a GST-tag and a TEV cleavage site. For the recombinant expression of CRBN-TBD protein, I followed and optimized the Adhikari et al. (2020) protocol [350]. In detail, for the expression I decided to use the Arctic Express Cells to overcome the hurdle of protein solubility. The protein expression was performed in TB medium supplied with zinc acetate solution for a final concentration of 50 μM. I observed that the expression of CRBN-TBD in absence of zinc acetate solution is lacking, therefore indicating that the zinc ion is mandatory for the production and the assembly of the protein. For the initial purification, I exploited the His-tag with an IMAC system (HisTrap FF column). Substantial protein precipitation was observed, which became particularly pronounced after the proteolytic cleavage step with TEV protease. Therefore, I moved on to the investigation of a different purification protocol, exploiting the GST-tag, and improved the proteolytic TEV cleavage with the addition of 5% glycerol to the buffer solution. Moreover, I replaced the counter-column step after the tag cleavage with a SEC purification (HiLoad Superdex 75pg 16 60 column (Cytiva)) (**Figure 28**). The final protein yield was about 5 mg/L culture.

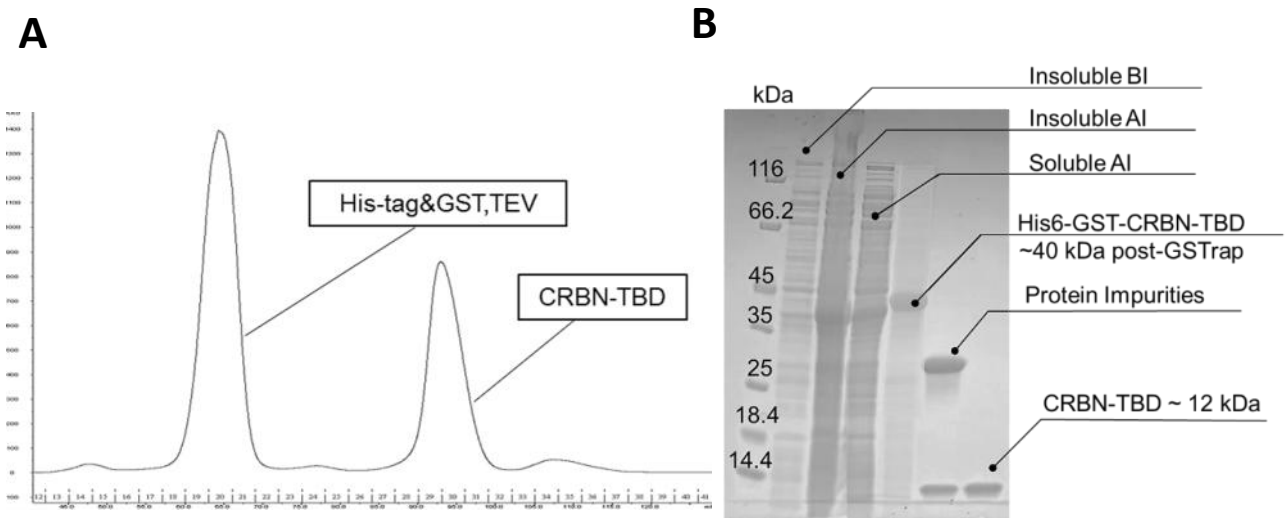


Figure 28. SEC profile of CRBN-TBD purification. Two peaks are visible, one eluted at 60-70 ml corresponding to His-tag, GST and TEV proteins, the other at 90-95 ml related to the CRBN-TBD (**A**). Isocratic elution was performed with 1.3 CV on a HiLoad Superdex 75pg 16 60 column (Cytiva). Elution buffer: 50 mM HEPES, 500 mM NaCl, 0,5 mM TCEP, pH 7.5. SDS-PAGE analysis of the main steps of purification on a 4–15% Mini-PROTEAN® TGX™ Precast Protein Gel with the molecular marker (Bio-rad) (**B**).

3.3.2. - Interaction of KME16 PROTAC with CRBN-TBD by NMR and fluorescence spectroscopy

^{1}H 1D NMR spectra were first recorded for CRBN-TBD, after which the protein was titrated with KME16 PROTAC from a stock solution in acetonitrile-d_3. Comparison of the titrated CRBN-TBD spectra with the free protein spectrum revealed noticeable differences, particularly in the amide and methyl regions (**Figure 29**). These observations are an indication of PROTAC-CRBN-TBD interaction.

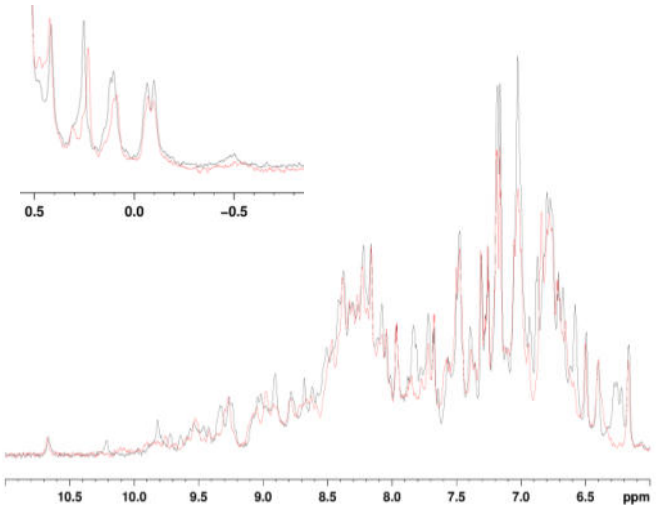


Figure 29. Methyl and amide regions of 1D ^{1}H NMR spectra of free CRBN-TBD (black) overlaid with CRBN-TBD-KME16 PROTAC mixture (red). The spectra were recorded at 298K, 950 MHz spectrometer (Bruker).

3.3.3. - Testing SARS-CoV-2 3CL^{Pro}:KME16 PROTAC:CRBN-TBD ternary complex formation

A mixture was prepared by incubating (4 hours) SARS-CoV-2 3CL^{Pro} (200 μM concentration) with the pre-formed CRBN-TBD–KME16 PROTAC complex, obtaining a 1:1:1 target protease:molecule:E3 ligase molar ratio. The mixture was subsequently concentrated and analyzed using a Superdex 200 Increase 10/300 GL column (Cytiva). Elution was performed with 25 mM HEPES, 150 mM NaCl, 1 mM TCEP, pH 6.5 buffer solution.

In the case of successful ternary complex assembly, the chromatographic profile would be expected to reveal three distinct peaks: an early-eluting species corresponding to the higher-molecular-weight ternary complex, followed by the peaks associated with SARS-CoV-2 3CL^{Pro}

(~66 kDa) and CRBN-TBD (~12 kDa), respectively. However, the chromatogram obtained for the mixture displayed only two peaks. Direct overlay of this chromatogram with those obtained for pure CRBN-TBD and pure SARS-CoV-2 $3CL^{Pro}$ under identical conditions demonstrated complete overlap of the peaks, strongly suggesting that no stable ternary complex was formed (**Figure 30**).

To corroborate this result, the elution fractions corresponding to the two peaks were analyzed by SDS-PAGE. Each fraction yielded a single, well-defined band, corresponding to the molecular weight of either CRBN-TBD or SARS-CoV-2 $3CL^{Pro}$ alone. Importantly, no additional bands or band shifts indicative of ternary complex formation were observed. This suggests either that the interaction between the two proteins mediated by KME16 PROTAC is too transient or weak to be detected by gel filtration, or that the chosen conditions are not conducive to ternary complex stabilization.

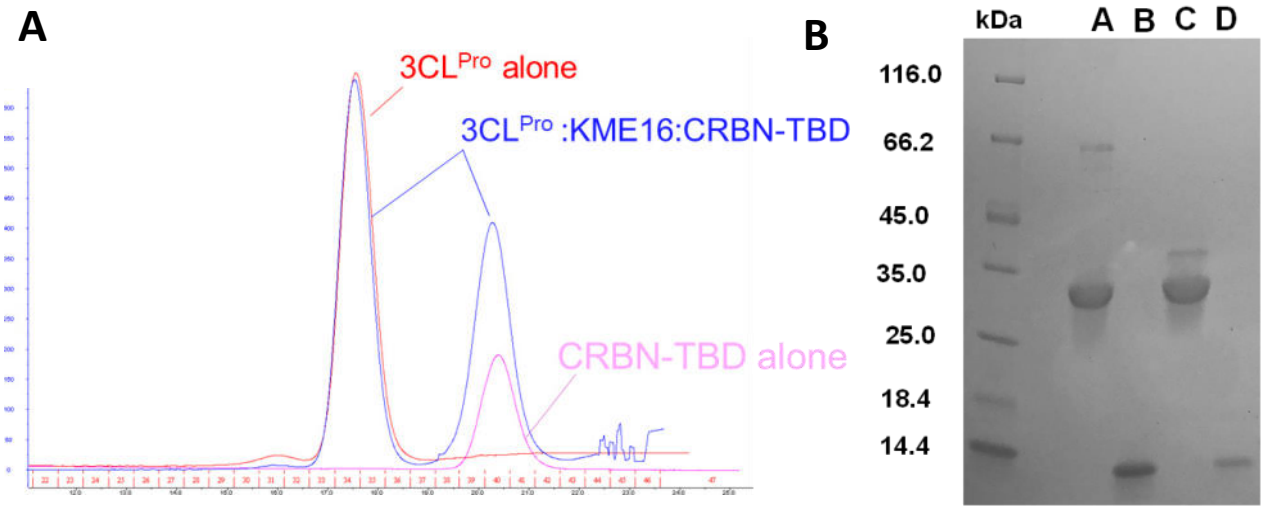


Figure 30. Analytical SEC of SARS-CoV-2 $3CL^{Pro}$:KME16 PROTAC:CRBN-TBD mixture (**A**). Comparative overlay of UV-chromatograms: in blue the curve corresponding to the complex mixture, in red and pink the runs related to the SARS-CoV-2 $3CL^{Pro}$ and CRBN-TBD alone, respectively. Isocratic elution was performed with 1.2 CV on a Superdex 200 Increase 10/300 GL column (Cytiva). Elution buffer: 25 mM HEPES, 150 mM NaCl, 1 mM TCEP, pH 6.5. SDS-PAGE analysis of the analytical SEC eluted fractions on a 4–15% Mini-PROTEAN® TGX™ Precast Protein Gel with the molecular marker (Bio-rad) (**B**). Label A indicates eluted fraction 35 of the mixture complex containing only $3CL^{Pro}$ and impurities at higher molecular weight. Label B refers to eluted fraction 40 of the mixture complex containing only CRBN-TBD. Labels C and D correspond to elution chromatograms of $3CL^{Pro}$ and CRBN-TBD alone, respectively.

My contribution to the project consisted of the development of successful protocols for the expression of the recombinant proteins SARS-CoV-2 $3CL^{Pro}$ and CVB3 $3C^{Pro}$ in E. coli as isotopically and not isotopically enriched samples. I produced SARS-CoV-2 $3CL^{Pro}$ protein as described previously (section 3.1.). To obtain the labelled protein, the production was performed in M9 medium supplied with ammonium sulfate enriched in ^{15}N. Concerning the CVB3 $3C^{Pro}$ production, I selected the most suitable expression construct, the pET-24a (+) plasmid encoding the 3C protease with a hexahistidine-coding sequence was used to express the protein into T7 Express pLysS competent BL21(DE3) E. coli cells. I induced the protein at 23°C overnight with IPTG (1 mM final concentration). Purification consisted of two steps: IMAC purification and SEC. Optimizing the protocol for CVB3 $3C^{Pro}$ production, I achieved a good yield protein expression and high level of purification (about 45 mg/L of culture). I also produced the CRBN-TBD E3 ligase and thanks to the modifications apported to the purification protocol, the stability of the protein appeared to be better after the TEV incubation.

Moreover, ^{1}H-^{15}N HSQC NMR titration experiments of the target proteases with the KME16 and ACH213 PROTAC molecules were performed. We analyzed the NMR bidimensional spectra to execute the chemical shift perturbation analysis of the SARS-CoV-2 $3CL^{Pro}$ in complex with KME16 or ACH213 PROTACs and identify the residues involved in the interactions. The interaction of KME16 with CRBN-TBD was characterized by performing 1D ^{1}H NMR titration experiments. Finally, we executed the analytical gel filtration of the complex mixture SARS-CoV-2 $3CL^{Pro}$-KME16 PROTAC-CRBN-TBD, which turned out to be unsuccessful.

In the next section (3.4.), results about the application of a promising PROTAC molecule, the GC-376 based PROTAC, against both viral proteases are displayed.

3.4. - Evaluation of a GC-376 based PROTAC targeting the main protease from SARS-CoV-2 and CVB3

The high degree of conservation of the main protease across viral genomes, coupled with the absence of a similar protease in humans, makes this enzyme family an excellent candidate for antiviral drug development with reduced host toxicity. These characteristics have allowed the creation of multiple classes of viral inhibitors capable of disrupting the function of this target. Among these, GC-376 stands out as one of the most potent inhibitors of the 3CL protease, effectively blocking the activity of coronavirus $3CL^{Pro}$ [351]. Recent research has demonstrated that GC-376 is a highly effective 3CL protease inhibitor against several coronaviral $3CL^{Pro}$ enzymes, including SARS-CoV-2 mutants [352]. Specifically, GC-376 has been shown to exhibit strong inhibitory activity against MERS-CoV, SARS-CoV and SARS-CoV-2 [352]. According to the structural information about SARS-CoV-2 $3CL^{Pro}$-GC-376 interaction, Prof. Trabocchi and his collaborators (DICUS, UNIFI) have designed and synthesized a GC-376 based PROTAC bifunctional molecule, able to recognize the active site of the dimeric $3CL^{Pro}$ of SARS-CoV-2 from one side and to select CRBN E3 ubiquitin ligase from the other side (section 3.3.1.). PROTAC molecule includes the key interacting elements of GC-376 and of its aldehyde derivative, potent inhibitor of SARS-CoV-2 $3CL^{Pro}$, named GC373 [353,354]. The precursor molecule, called GC-376 PROTAC precursor, was synthetized as a dipeptidyl ligand containing an (S)-γ-lactam ring, a leucine side chain and a α,β-unsaturated methyl ester group as electrophilic warhead. In the synthesis of GC-376 PROTAC, the C-terminus of the dipeptidyl GC-376 based ligand, was conjugated to a pomalidomide CRBN ligand through a piperazine-piperidine rigid linker (**Figure 31**), retained as an optimal strategy in CRBN-based PROTAC degraders [355]. NMR and crystallographic analyses (PDB IDs 8OKC and 8OKB for SARS-CoV-2 $3CL^{Pro}$ complexed with GC-376 PROTAC and its precursor, respectively), alongside enzymatic and cellular experiments, revealed that the dipeptidyl ligand of the PROTAC interacts with the active site of the dimeric form of SARS-CoV-2 $3CL^{Pro}$, forming a reversible covalent bond with the sulfur atom of the catalytic Cys145 residue. Moreover, we observed that the linker and the pomalidomide CRBN ligand of PROTAC exhibits significant flexibility and absence of interactions with other regions of the protein, suggesting that this molecular part of the molecule extends outward from the protein surface. To determine whether GC-376 PROTAC could interact with the monomeric form of SARS-CoV-2 $3CL^{Pro}$, NMR titration experiments were also performed by gradually adding the molecule to the monomeric enzyme. In this case we did not observe any notable chemical shift variations, showing no interaction between the protein and PROTAC molecule.

We have ultimately demonstrated that GC-376 PROTAC is effective within cells, significantly lowering the levels of SARS-CoV-2 $3CL^{Pro}$ protein while preserving cell health (section 3.4.1.).

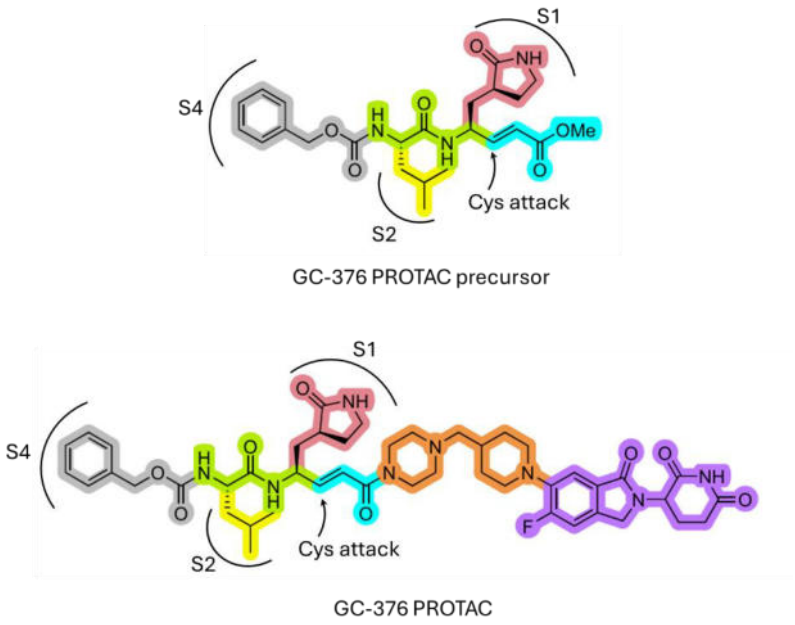


Figure 31. Molecular structure of GC-376 PROTAC precursor and GC-376 PROTAC. The pomalidomide moiety, the piperazine-piperidine linker and the warhead backbone of PROTACs are represented in violet, in orange and green, respectively. The functional chemical groups of PROTAC precursor and PROTAC warhead are depicted in different colours: the benzyloxycarbonyl group in grey, the leucine side chain in yellow, the γ-lactam ring in dark pink, the α,β-unsaturated methyl ester group of PROTAC precursor and the α,β-unsaturated PROTAC amide moiety in light blue. The protein subsites are marked for typical recognition of specific functional groups.

As mentioned in the Introduction (section 1.4.), the coronavirus $3CL^{Pro}$ shares a significant structural resemblance with the $3C^{Pro}$ enzymes found in the picornavirus family. The observation that all these proteases encoded by coronaviruses and enteroviruses play crucial roles in the viral replication process by cleaving both viral and host proteins, combined with the fact that no human protease is known to specifically target Gln at the cleavage site [356], has encouraged researchers to pursue the development of an effective antiviral strategy targeting both virus groups. These considerations prompted us to investigate the potential of the previously developed GC-376 PROTAC against SARS-CoV-2 $3CL^{Pro}$ for its effectiveness in also targeting enteroviral CVB3 $3C^{Pro}$. Initially, I selected the most suitable expression construct for expressing CVB3 $3C^{Pro}$ in E. coli bacteria cells to achieve high levels of protein both in natural abundance and in isotopically ^{15}N and ^{13}C enriched form. Subsequently, the interaction between PROTAC molecule and CVB3

$3C^{Pro}$ was characterized by solution NMR, X-ray crystallography and biochemical techniques. We obtained the crystal structure of CVB3 $3C^{Pro}$ bound to the GC-376 PROTAC precursor at 1.9 Å resolution (PDB ID 8S6F). The crystallographic data indicates that while there are some differences in the binding mode to the molecule between CVB3 $3C^{Pro}$ and SARS-CoV-2 $3CL^{Pro}$, their overall structures are highly similar. Moreover, NMR backbone assignments were performed for CVB3 $3C^{Pro}$ both in its free form and when bound to the GC-376 PROTAC precursor, achieving 80% and 97% coverage respectively. This information enabled us to determine that the GC-376 PROTAC interacts with CVB3 $3C^{Pro}$ in a manner very similar to that of the precursor molecule (section 3.4.2.). Comprehensively, structural analyses regarding the active site of both CVB3 $3C^{Pro}$ and SARS-CoV-2 $3CL^{Pro}$ demonstrate that the GC-376 PROTAC aligns effectively within the binding regions of the two proteases. This provides evidence of its potential as a versatile molecule for designing broad-spectrum antiviral PROTACs.

My contribution to this part of the project, presented in the next two papers (sections 3.4.1. and 3.4.2.), consisted of the expression and purification of the SARS-CoV-2 $3CL^{Pro}$ and CVB3 $3C^{Pro}$ target proteases as described previously (sections 3.1. and 3.3.3.). I also produced the SARS-CoV-2 $3CL^{Pro}$ protein in its monomeric form. The SARS-CoV-2 $3CL^{Pro}$ gene, optimized for E. coli expression, was cloned into the pET-28a(+) vector to produce a His-tagged monomeric protein. The construct included an N-terminal 6xHis tag separated from $3CL^{Pro}$ by a thrombin cleavage site. The plasmid was transformed into BL21(DE3) cells, which were grown in LB medium or M9 minimal medium containing ^{15}N-labelled ammonium sulfate. Protein expression was induced with 0.5 mM IPTG at an OD_{600} of 0.7–0.8 and continued at 35 °C for 6 hours. Regarding the protein purification, a HisTrap FF affinity column was used to elute the monomeric SARS-CoV-2 $3CL^{Pro}$ with a stepwise imidazole gradient (5–500 mM). After a buffer exchange step, the final protein yield turned out to be 25 mg/L of culture. Then I carried out the analytical gel filtration of monomeric CVB3 $3C^{Pro}$ and SARS-CoV-2 $3CL^{Pro}$ in its dimeric and monomeric form. I personally performed and optimized the crystallization trials to obtain the protein-molecule adducts SARS-CoV-2 $3CL^{Pro}$-GC-376 PROTAC precursor through soaking technique, while SARS-CoV-2 $3CL^{Pro}$-GC-376 PROTAC and CVB3 $3C^{Pro}$-GC-376 PROTAC precursor crystal complexes via co-crystallization. Bidimensional NMR titration experiments were conducted on target proteases in complex with GC-376 PROTAC, and analysis of 3D-triple resonance experiments on CVB3 $3C^{Pro}$ and CVB3 $3C^{Pro}$-GC 376 PROTAC precursor was achieved. I mapped the protein regions affected by the interaction after the chemical shift perturbation analysis for all the target protease-molecule complexes.

3.4.1. - Development of a GC-376 Based Peptidomimetic PROTAC as a Degrader of 3-Chymotrypsin-like Protease of SARS-CoV-2

Deborah Grifagni [1,2], Elena Lenci [2], Alessia De Santis [1,2], Andrea Orsetti [1,2], Carlo Giorgio Barracchia [3], Filomena Tedesco [2], Raffaele Bellini Puglielli [2], Francesca Lucarelli [1,2], Angela Lauriola [3], Michael Assfalg [3], Francesca Cantini [1,2], Vito Calderone [1,2], Daniele Guardavaccaro [3], Andrea Trabocchi [2,*], Mariapina D' Onofrio [3,*], Simone Ciofi-Baffoni [1,2,4,*]

[1]Magnetic Resonance Center CERM, University of Florence

[2] Department of Chemistry "Ugo Schiff", University of Florence

[3]Department of Biotechnology, University of Verona

[4]Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP)

* Correspondence: andrea.trabocchi@unifi.it; mariapina.donofrio@univr.it ; ciofi@cerm.unifi.it

Published: 10.1021/acsmedchemlett.3c00498

ACS Medicinal Chemistry Letters

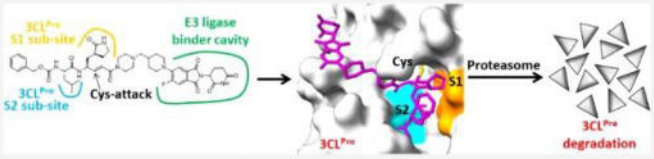
pubs.acs.org/acsmedchemlett Letter

Development of a GC-376 Based Peptidomimetic PROTAC as a Degrader of 3-Chymotrypsin-like Protease of SARS-CoV-2

Deborah Grifagni, Elena Lenci, Alessia De Santis, Andrea Orsetti, Carlo Giorgio Barracchia, Filomena Tedesco, Raffaele Bellini Puglielli, Francesca Lucarelli, Angela Lauriola, Michael Assfalg, Francesca Cantini, Vito Calderone, Daniele Guardavaccaro, Andrea Trabocchi,* Mariapina D'Onofrio,* and Simone Ciofi-Baffoni*

Cite This: *ACS Med. Chem. Lett.* 2024, 15, 250−257 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: We have applied a proteolysis targeting chimera (PROTAC) technology to obtain a peptidomimetic molecule able to trigger the degradation of SARS-CoV-2 3-chymotrypsin-like protease ($3CL^{Pro}$). The PROTAC molecule was designed by conjugating a GC-376 based dipeptidyl $3CL^{Pro}$ ligand to a pomalidomide moiety through a piperazine−piperidine linker. NMR and crystallographic data complemented with enzymatic and cellular studies showed that (i) the dipeptidyl moiety of PROTAC binds to the active site of the dimeric state of SARS-CoV-2 $3CL^{Pro}$ forming a reversible covalent bond with the sulfur atom of catalytic Cys145, (ii) the linker and the pomalidomide cereblon-ligand of PROTAC protrude from the protein, displaying a high degree of flexibility and no interactions with other regions of the protein, and (iii) PROTAC reduces the protein levels of SARS-CoV-2 $3CL^{Pro}$ in cultured cells. This study paves the way for the future applicability of peptidomimetic PROTACs to tackle $3CL^{Pro}$-dependent viral infections.

KEYWORDS: *Peptidomimetics, Bioconjugate chemistry, COVID-19, 3-Chymotrypsin-like protease, PROTAC, Infectious diseases*

Proteolysis targeting chimera (PROTAC) technology provides an attractive approach to modulate protein levels of therapeutic target by hijacking the cellular machinery responsible for the physiological elimination of endogenous proteins.[1−3] This strategy has been widely established to be successful in targeting cancer-related proteins,[4] with more than a dozen drug candidates entering clinical investigation.[5−7] On the contrary, the applicability of PROTAC technology in the field of antivirals remains marginal.[8−12] Only a few studies of PROTAC molecules targeting viral proteins have been reported in the case of hepatitis and influenza viruses,[13−15] and just a preliminary attempt investigating the applicability of PROTAC technology against coronaviruses (CoV) has been documented.[16] The latter study targets the viral spike protein receptor binding domain of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). PROTAC degraders acting against the major viral protease (3-chymotrypsin-like protease, $3CL^{Pro}$ hereafter), one of the best characterized viral drug targets to date,[17−19] have been only hypothesized[20,21] or computationally predicted.[22] $3CL^{Pro}$ turns out to be an ideal viral target for several reasons: (i) it is shared by all *Coronavirus* genera as well as by members of the large genus *Enterovirus* in the picornavirus family,[23] (ii) it is essential for viral replication,[24,25] (iii) it has unique features of its cleavage site recognition that are absent in closely related human homologues,[26] and (iv) it has a high degree of structural similarity among $3CL^{Pro}$ proteins of all *Coronaviruses* genera and 3C proteases of enteroviruses.[23]

In this study, we envisaged selecting $3CL^{Pro}$ of SARS-CoV-2 as a target for degradation activated via PROTAC technology, and we provide a PROTAC molecule specifically targeting and degrading the dimeric SARS-CoV-2 $3CL^{Pro}$ protein. This PROTAC molecule has been designed exploiting a structural biology approach to target the active site of SARS-CoV-2 $3CL^{Pro}$ by a peptidomimetic[27] ligand on one side, and to target cereblon (CRBN) by the immunomodulatory drug pomalidomide on the other side. CRBN is one of the substrate receptors of the CRL4 E3 ubiquitin ligase complex and is responsible for

Received: October 31, 2023
Revised: December 26, 2023
Accepted: January 3, 2024
Published: January 10, 2024

https://doi.org/10.1021/acsmedchemlett.3c00498
ACS Med. Chem. Lett. 2024, 15, 250−257

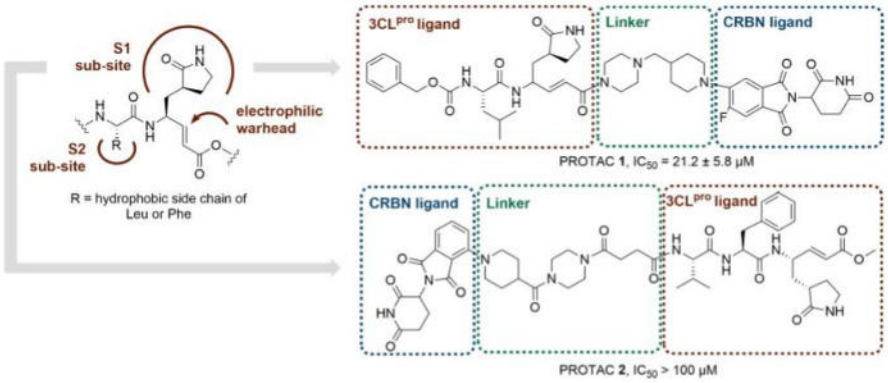


Figure 1. Schematic representation of synthesized PROTACs **1** and **2**. The synthesis is based on a core structure, shown on the left, containing common functional groups in specific positions, which are typically recognized by S1 and S2 subsites of $3CL^{Pro}$ proteins.

the polyubiquitination of the substrate (i.e., SARS-CoV-2 $3CL^{Pro}$ in this approach), which is the step required to trigger substrate degradation by the proteasome.[28,29] The use of CRBN in the design of PROTAC recruiters has been prompted by the discovery and the structural characterization of the interaction between CRBN and immunomodulatory imide drugs.[30−32] In the last years, many CRBN-based PROTACs have been designed to target a variety of substrate proteins for proteasomal degradation.[33] The immunomodulatory drug pomalidomide is a frequently used CRBN ligand to design PROTAC degraders.[34,35] In our designed PROTAC molecule, the pomalidomide and the peptidomimetic ligands were joined through a rigid linker previously shown to be optimal in CRBN-based PROTAC degraders.[36] X-ray crystallography and solution NMR were applied to structurally characterize the interaction between PROTAC and SARS-CoV-2 $3CL^{Pro}$, and cellular studies were used to investigate the degradation activity of the PROTAC molecule.

GC-376 is a typical $3CL^{Pro}$ inhibitor that has shown antiviral activity against feline infectious peritonitis CoV in experimentally infected cats and with broad-spectrum activity against coronavirus.[37,38] More recently, the antiviral activity of GC-376 against SARS-CoV-2 was demonstrated *in vitro*,[39,40] and *in vivo* studies supported that GC-376 could represent a promising lead candidate for therapeutic treatment of SARS-CoV-2 infections.[41] Accordingly, we designed two PROTAC molecules based on the structural information available for the SARS-CoV-2 $3CL^{Pro}$-GC-376 interaction. Both PROTAC molecules contain a SARS-CoV-2 $3CL^{Pro}$ peptidomimetic ligand that conserves the key interacting elements of GC-376 and of its aldehyde derivative GC-373.[42] This ligand contains an (*S*)-γ-lactam ring occupying the S1 subsite of SARS-CoV-2 $3CL^{Pro}$ enzyme and a hydrophobic side chain of an amino acid, such as leucine or phenylalanine, occupying the S2 subsite (Figure 1, left). As the electrophilic warhead of both SARS-CoV-2 $3CL^{Pro}$ ligands, we reasoned to replace the aldehyde functionality of GC-373 with a α,β-unsaturated methyl ester to achieve reversible ligands (Figure 1, left). On this structural basis, PROTACs **1** and **2** were designed and synthesized in a complementary approach, exploiting the *C*-terminus of the GC-376-derived dipeptidyl ligand for PROTAC **1** and its *N*-terminus for PROTAC **2**. For the synthesis of PROTAC **1**, the *C*-terminus of a dipeptidyl SARS-CoV-2 $3CL^{Pro}$ ligand (named ligand **1′** hereafter) was conjugated to the CRBN ligand pomalidomide through a piperazine−piperidine linker, the latter chosen following the success of Arvinas company in delivering the first two PROTACs as clinical candidates (Figure 1, right).[43] Briefly, PROTAC **1** was achieved following ester hydrolysis of **1′** and subsequent amide coupling with the amino group of piperazine−piperidine-functionalized pomalidomide in 65% yield. PROTAC **2** was achieved in 13% yield by installing the pomalidomide CRBN ligand at the *N*-terminus position of a tripeptidyl SARS-CoV-2 $3CL^{Pro}$ ligand through the piperazine−piperidine linker as for PROTAC **1**. The tripeptidyl ligand was synthesized from **1′** through Cbz removal and subsequent couplings at the *N*-terminus of valine and of succinic acid as a linker for the bioconjugation to the piperazine−piperidine-functionalized pomalidomide unit (Figure 1, right; details of the synthetic routes for PROTACs **1** and **2** are reported in Supporting Information (SI), Schemes S1 and S2, respectively).

We evaluated ligand **1′** and PROTACs **1** and **2** for their ability to inhibit SARS-CoV-2 $3CL^{Pro}$ (at 60 nM) through a fluorimetric enzyme inhibition kinetics assay using Hilyte Fluor-488-ESATLQSGLRKAK-(QXL-520)-NH_2 as substrate. Ligand **1′** and PROTAC **1** showed inhibition activity with a IC_{50} of 1.35 μM and 21.2 μM, respectively (Table 1, and SI,

Table 1. Inhibitory Activity of Ligand 1′, PROTACs 1 and 2,[a] and of Representative SARS-CoV-2 Inhibitors against $3CL^{Pro}$

compd	IC_{50} (μM)
Ligand **1′**	1.35 ± 0.34
PROTAC **1**	21.2 ± 5.8
PROTAC **2**	N/A
representative $3CL^{Pro}$ inhibitors	**reported IC_{50} (μM)**
GC-376[41,44,45]	0.19 ± 0.04
GC-373[45]	0.40 ± 0.05
Nirmatrelvir[46]	0.051 ± 0.01
Boceprevir[47,48]	5.40 ± 1.53
Dalcetrapib-thiol[49]	14.4 ± 3.3

[a]Mean from three different assays, errors were in the range of 5−10% of the reported values; IC_{50} values were retrieved from dose−response assays as the concentration of compound required for 50% inhibition, as estimated by nonlinear correlation using GraphPad Prism software.

https://doi.org/10.1021/acsmedchemlett.3c00498
ACS Med. Chem. Lett. 2024, 15, 250−257

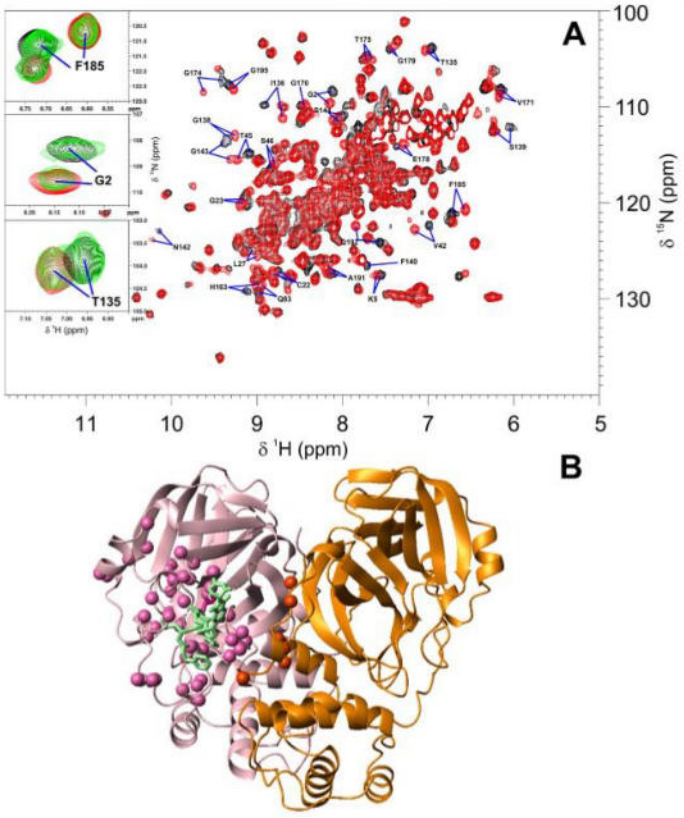


Figure 2. Mapping the interaction between PROTAC **1** and dimeric SARS-CoV-2 3CL^{Pro}. (A) Overlay of ^{1}H−^{15}N HSQC maps of ^{15}N-labeled SARS-CoV-2 3CL^{Pro} (red) and a 1:1 ^{15}N-labeled SARS-CoV-2 3CL^{Pro}-PROTAC **1** mixture (black). Some residues affected by the interaction are highlighted. In the insets, ^{1}H−^{15}N HSQC maps at 1:0 (red), 1:0.5 (green), and 1:1 (black) SARS-CoV-2 3CL^{Pro}-PROTAC **1** ratios are shown for some residues that are in slow exchange regime on the NMR time scale upon complex formation. (B) The backbone NH protons of the residues affected by the interaction between PROTAC **1** and dimeric SARS-CoV-2 3CL^{Pro} are shown as pink and orange spheres on the two subunits (depicted in orange and pink) of the dimeric structure of SARS-CoV-2 3CL^{Pro} (PDB 8OKC). The side chains of the Cys-His catalytic dyad are shown in yellow on the pink subunit.

Figure S1), whereas PROTAC **2** did not display any significant activity, showing only 64% inhibition at 100 μM and 0% inhibition at 10 μM. Therefore, it was not further considered in subsequent structural and cellular studies.

In order to elucidate the binding mode of PROTAC **1**, solution NMR was applied to screen the interaction[50] between PROTAC **1** and dimeric SARS-CoV-2 3CL^{Pro}. Specifically, NMR titration experiments were performed by stepwise addition of PROTAC **1** to ^{15}N-labeled dimeric SARS-CoV-2 3CL^{Pro}. Upon subsequent additions of PROTAC **1**, chemical shift changes were observed in the ^{1}H−^{15}N heteronuclear single quantum coherence (HSQC) maps of dimeric SARS-CoV-2 3CL^{Pro}, corresponding to a slow exchange regime on the NMR time scale (Figure 2A). This means that signals of dimeric SARS-CoV-2 3CL^{Pro} decreased in intensity, and those of a new species, assigned to the adduct formed between dimeric SARS-CoV-2 3CL^{Pro} and PROTAC **1**, appeared and increased in intensity during the titration (Figure 2A, insets). On reaching the 1:1 SARS-CoV-2 3CL^{Pro}-PROTAC **1** ratio, the signals of SARS-CoV-2 3CL^{Pro} completely disappeared and the signals of the PROTAC **1**−protein complex reached their maximal intensity, indicating that the complex was fully formed at the 1:1 molar ratio. These data strongly support that a tight binding of PROTAC **1** to the protein occurred. Most of the affected residues were found close to the Cys-His catalytic dyad of SARS-CoV-2 3CL^{Pro} (Figure 2B), thus revealing that PROTAC **1** specifically binds to the active site of the enzyme. The NMR data were confirmed by a fluorescence quench on the spectra recorded with or without PROTAC **1** addition after excitation of the tryptophan residues (SI, Figure S2A).

To further characterize the protein-PROTAC **1** interaction, NMR experiments were performed titrating ^{15}N-labeled dimeric SARS-CoV-2 3CL^{Pro} with ligand **1′**. We observed chemical shift changes in a slow exchange on the NMR time scale (SI, Figure S3), which reproduced those observed in the titration between PROTAC **1** and SARS-CoV-2 3CL^{Pro}. This result indicates that the dipeptidyl moiety of PROTAC **1** is actively involved in binding to the active site of the enzyme, while the linker and the pomalidomide moieties of PROTAC **1** do not significantly interact with SARS-CoV-2 3CL^{Pro}. To

https://doi.org/10.1021/acsmedchemlett.3c00498
ACS Med. Chem. Lett. 2024, 15, 250−257

provide information on the structural organization of the protein in the presence of PROTAC **1** and ligand **1′**, circular dichroism (CD) spectra were recorded on the dimeric SARS-Cov2 $3CL^{Pro}$ in the presence and absence of PROTAC **1** or ligand **1′** (SI, Figure S4). The recorded spectra are fully superimposable, thus showing that no significant changes of the secondary structure of the dimeric SARS-Cov2 $3CL^{Pro}$ occurred upon the PROTAC **1** interaction. Finally, to ascertain whether PROTAC **1** could recognize monomeric SARS-CoV-2 $3CL^{Pro}$, NMR titration experiment were carried out by adding PROTAC **1** to a monomeric form of SARS-CoV-2 $3CL^{Pro}$ (SI, Figures S5 and S6). No significant chemical shift changes were observed for the monomeric form, indicating no binding of PROTAC **1** to the protein (SI, Figure S6). Thus, we conclude that the dimeric state of SARS-CoV-2 $3CL^{Pro}$ is specifically targeted by PROTAC **1** and only the dimeric state of SARS-CoV-2 $3CL^{Pro}$ can be subjected to protein degradation.

To gain further insight into the interaction mode of PROTAC **1** with SARS-CoV-2 $3CL^{Pro}$, the crystal structures of SARS-CoV-2 $3CL^{Pro}$ in complex with ligand **1′** and PROTAC **1** were obtained (see SI, Table S1). The polypeptide structure of both adducts is well superimposable with previously reported structures of SARS-CoV-2 $3CL^{Pro}$.[51,52] By comparing the structure of the apo form of SARS-CoV-2 $3CL^{Pro}$ with those of the two adducts with ligand **1′** and PROTAC **1**, only loop regions showed local backbone RMSD values higher than average, indicating that the binding of either ligand **1′** or PROTAC **1** to the protein does not require significant structural rearrangement of the protein active site. Concerning the occupancy of PROTAC **1** and ligand **1′**, we found that the electron density is quite well-defined for most atoms of ligand **1′** (SI, Figure S7A) and, in the case of PROTAC **1**, the electron density is better defined for the chemical moiety corresponding to ligand **1′** (SI, Figure S7B). For the remaining PROTAC **1** structure, the electron density is virtually absent, indicating a high degree of mobility of the linker and of the pomalidomide moiety as well as the absence of interactions with other regions of the protein. This suggests that this molecular fragment protrudes out of the protein and is completely solvent exposed (Figure 3), in agreement with solution NMR data and in support of the fruitful involvement of PROTAC **1** in recruiting CRL4 E3 ubiquitin ligase. The binding of ligand **1′** and PROTAC **1** to the SARS-CoV-2 $3CL^{Pro}$ protein is quite similar for the structural part that is shared by the two compounds, highlighting their similar positioning in the protein active site (Figure 4A). The most relevant feature in common between the two compounds is their binding mode to the catalytic Cys145. Accordingly, the crystallographic structures of both adducts revealed that the distance between the sulfur atom of Cys145 and the Cβ carbon of the α,β-unsaturated amide moiety of ligand **1′** and PROTAC **1** spontaneously refines to a value around 1.7 Å, which is typical of a C−S covalent bond. From a chemical point of view, it is known that the formation of a covalent bond between the unsaturated β-carbon and sulfur atoms has been already observed to occur between analogous peptidomimetic inhibitors and SARS-CoV $3CL^{Pro}$.[53]

The formation of a covalent complex between SARS-CoV-2 $3CL^{Pro}$ and ligand **1′** or PROTAC **1** was confirmed by matrix-assisted laser desorption/ionization−time-of-flight (MALDI-TOF) mass spectrometry data. Incubation of SARS-CoV-2 $3CL^{Pro}$ with PROTAC **1** (exact mass 885 Da) or ligand **1′** (exact mass 459 Da) yielded a peak shift from the protein mass

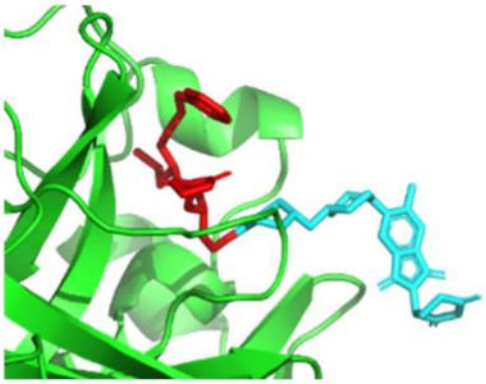


Figure 3. Close-up of the binding of PROTAC **1** to the SARS-CoV-2 $3CL^{Pro}$ active site. The dipeptidyl moiety of PROTAC **1** with a well-defined electron density is in red. The piperazine−piperidine linker and the pomalidomide moiety with absent electron density are shown in cyan.

by 881 and 458 Da, respectively (SI, Figure S2B−C). Increased mass corresponds to a single molecule of PROTAC **1** or ligand **1′** covalently captured by SARS-CoV-2 $3CL^{Pro}$. Besides the covalent bond, other polar and nonpolar interactions keep both ligand **1′** and PROTAC **1** in place.

In the case of PROTAC **1**, the side chain of His41, and the backbone of Phe140 and Glu166 were found as the protein residues involved in direct hydrogen bonding interactions (Figure 4B, left). In addition, a water molecule bridges the nitrogen atom of Gln189 via hydrogen bonding. Several residues are involved in hydrophobic interactions (Figure 4B, left), in agreement with most of the molecule atoms being nonpolar. In the case of the ligand **1′**-protein structure, protein−ligand interactions are similar but not identical to those found in the PROTAC **1**−protein structure (Figure 4B, right), as the hydrogen bonding network partially differs. In particular, the hydrogen bond involving the His41 side chain is absent, the one involving the Phe140 backbone is looser, and those involving Glu166 and the water molecule bridging Gln189 are strictly maintained (Figure 4B, right). Two additional hydrogen bonds are present involving the backbone amides of Cys145 and of Gly143 and the same oxygen of the ester group of the ligand. The absence of such hydrogen bonds in the PROTAC **1**−protein structure is due to the substitution of the methyl ester of ligand **1′** with the bulky piperazine−piperidine linker (Figure 4B, left). Although the occurrence of these slightly different hydrogen bond networks, all the residues involved in the molecular recognition of SARS-CoV-2 $3CL^{Pro}$ with both PROTAC **1** and ligand **1′** are conserved among human coronavirus main proteases, with the only exception of Asn 142 (SI, Figure S8). This finding supports the idea that the active sites of other viral proteases might also be targeted by PROTAC **1**.

Considering the covalent C−S bond formation detected in the crystal structure between Cys145 and the Cβ carbon of the α,β-unsaturated amide moiety of PROTAC **1**, we investigated whether PROTAC **1** is a reversible covalent ligand for SARS-CoV-2 $3CL^{Pro}$ by applying an assay protocol already available in the literature.[54] Thus, the enzyme was incubated with or without a large excess of PROTAC **1** for 30 min. Then, such

https://doi.org/10.1021/acsmedchemlett.3c00498
ACS Med. Chem. Lett. 2024, 15, 250−257

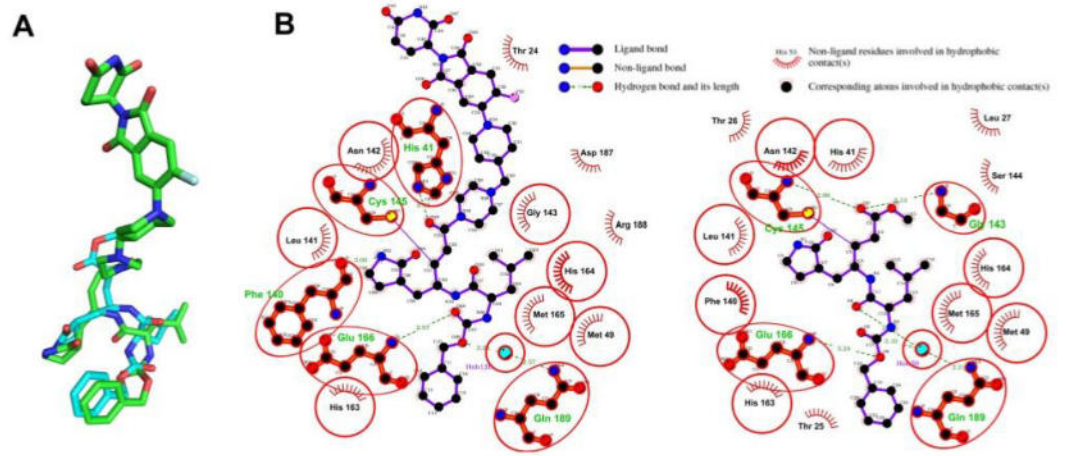


Figure 4. Binding mode of PROTAC **1** and ligand **1′** to SARS-CoV-2 3CL^{Pro} by X-ray crystallography. (A) Superimposition of PROTAC **1** (green) and ligand **1′** (cyan) in the respective structures showing a similar binding mode to that of the protein. (B) Ligplot schematic view of the binding modes of PROTAC **1** (left) and ligand **1′** (right). The residues of SARS-CoV-2 3CL^{Pro} involved in covalent, hydrophobic, and hydrogen-bonding contacts with PROTAC **1** and ligand **1′** are indicated according to the legend. The red circles and ellipses identify the residues that interact with the ligand (PROTAC **1** or ligand **1′**) on both structures upon superposition of the two structures.

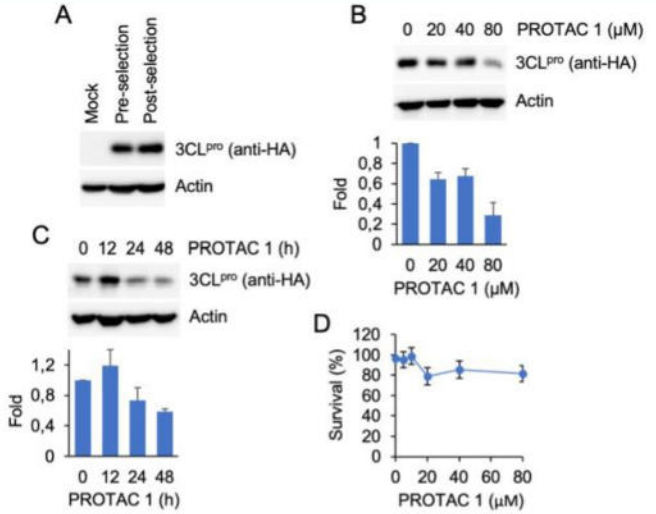


Figure 5. Effects of PROTAC **1** on SARS-CoV-2 3CL^{Pro} protein levels and cell viability in HeLa cells. (A) HeLa cells transduced with lentiviruses expressing C-terminally tagged SARS-CoV-2 3CL^{Pro} were analyzed by immunoblotting with antibodies specific for the indicated proteins before (preselection expressing HA-epitope) and after (postselection) incubation with puromycin. (B) HeLa cells expressing HA-epitope C-terminally tagged SARS-CoV-2 3CL^{Pro} were treated with PROTAC **1** at the indicated concentrations for 24 h. Cells were harvested and lysed. Whole cell extracts were analyzed by immunoblotting with antibodies specific to the indicated proteins. The graph shows the quantification of SARS-CoV-2 3CL^{Pro} levels normalized to actin levels. (C) HeLa cells expressing HA-epitope C-terminally tagged SARS-CoV-2 3CL^{Pro} were treated with 80 μM PROTAC **1** for the indicated times. Cells were analyzed by immunoblotting as in (B). The graph shows the quantification of SARS-CoV-2 3CL^{Pro} levels normalized to actin levels. (D) Viability of HeLa cells treated with PROTAC **1** for 72 h at the indicated concentration was assessed by 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay. Data are reported as mean ± SEM; $n = 3$.

mixtures were diluted 100-fold into a reaction buffer containing a fluorogenic substrate to initiate the enzymatic reaction of SARS-CoV-2 3CL^{Pro}. The progress curves for these samples were obtained, and the corresponding initial velocities were measured and compared to those of the enzyme incubated and diluted in the absence of PROTAC **1**. This experiment revealed the complete recovery of the enzyme activity (SI, Figure S9), thus showing the reversibility of the covalent C−S bond formed between SARS-CoV-2 3CL^{Pro} and PROTAC **1**. As a general rule for a PROTAC molecule,

https://doi.org/10.1021/acsmedchemlett.3c00498
ACS Med. Chem. Lett. 2024, 15, 250−257

combining covalent bond formation with its reversibility is expected to represent the best condition to efficiently degrade target proteins in a cellular context.[55] Indeed, the reversibility ensures dissociation and recycling of the PROTAC molecule, whereas the covalency confers higher selectivity.[56] Overall, these effects favor low doses usage of PROTAC molecules.

Finally, we investigated the degradation efficiency in the cellular context by assaying whether PROTAC **1** affects SARS-CoV-2 $3CL^{Pro}$ protein levels in cultured cells. We first generated HeLa cells stably expressing hemagglutinin (HA)-epitope C-terminally tagged SARS-CoV-2 $3CL^{Pro}$ by lentiviral transduction (Figure 5A). These cells were then treated with increasing concentrations of PROTAC **1** and the expression of SARS-CoV-2 $3CL^{Pro}$ was examined by immunoblotting. We observed a pronounced decrease in SARS-CoV-2 $3CL^{Pro}$ levels in the presence of PROTAC **1** at concentrations that did not affect the cell viability (Figure 5B–D).

In this study, we designed and synthesized a peptidomimetic PROTAC molecule specifically targeting the dimeric SARS-CoV-2 $3CL^{Pro}$ protein. We structurally characterized the interaction of the PROTAC molecule with SARS-CoV-2 $3CL^{Pro}$ and observed the formation of a covalent bond between the Cβ carbon of the α,β-unsaturated amide moiety of the PROTAC **1** molecule and the catalytic Cys145 sulfur atom of SARS-CoV-2 $3CL^{Pro}$. The covalent reversibility of the PROTAC **1** inhibition was demonstrated as a key requirement for the efficacy of the PROTAC molecule that needs to be recycled at the cellular level once the degradation of the viral target has occurred. We finally provided the first evidence that PROTAC **1** is active at the cellular level drastically reducing protein levels of SARS-CoV-2 $3CL^{Pro}$ without affecting cell viability, indicating that a peptidomimetic-based PROTAC strategy can be an efficient tool to block viral infections in *Coronavirus* genera also in the μM range.

Further studies will be aimed at improving the PROTAC efficacy in reaching the target in the cellular environment. We expect that the insights attained from this study can extend the applicability of peptidomimetic-based PROTACs to a more general setting to attack *Coronavirus* genera, as well as some members of the large genus Enterovirus.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications Web site. The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsmedchemlett.3c00498.

Experimental procedures and characterization data for the synthesis of PROTAC **1** and **2**; inhibition curves for PROTAC **1** and ligand **1′**; fluorescence and MALDI TOF mass spectrometry data monitoring complex formation of SARS-CoV-2 $3CL^{Pro}$ with PROTAC **1** and ligand **1′**; NMR data monitoring complex formation between ligand **1′** and dimeric SARS-CoV-2 $3CL^{Pro}$; far-UV circular dichroism spectroscopy monitoring structural changes upon complex formation of SARS-CoV-2 $3CL^{Pro}$ with PROTAC **1** and ligand **1′**; analytical gel filtration of monomeric and dimeric SARS-CoV-2 $3CL^{Pro}$; NMR data monitoring the interaction between PROTAC **1** and monomeric SARS-CoV-2 $3CL^{Pro}$; electron density of ligand **1′** and PROTAC **1**; sequence alignment of $3CL^{Pro}$ proteins from different coronaviruses and mapping of the residues interacting with both PROTAC **1** and ligand **1′** on the SARS-CoV-2 $3CL^{Pro}$ structure; recovery of the enzymatic activity in the SARS-CoV-2 $3CL^{Pro}$-PROTAC **1** adduct; HPLC chromatograms of PROTAC **1** and PROTAC **2**; data collection and refinement statistics of SARS-CoV-2 $3CL^{Pro}$ complexed with ligand **1′** and PROTAC **1**; NMR spectra of ligands **1′**, **2′**, PROTAC **1**, PROTAC **2**, **S7**, and **S12** (PDF)

Accession Codes

Coordinates and structure factors have been deposited at the PDB under the accession code 8OKB for SARS-CoV-2 $3CL^{Pro}$ complexed with ligand **1′** and 8OKC for SARS-CoV-2 $3CL^{Pro}$ complexed with PROTAC **1**.

■ AUTHOR INFORMATION

Corresponding Authors

Andrea Trabocchi − *Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Florence, Italy*; orcid.org/0000-0003-1774-9301; Email: andrea.trabocchi@unifi.it

Mariapina D'Onofrio − *Department of Biotechnology, University of Verona, 37134 Verona, Italy*; orcid.org/0000-0002-8699-0847; Email: mariapina.donofrio@univr.it

Simone Ciofi-Baffoni − *Magnetic Resonance Center CERM, University of Florence, 50019 Sesto Fiorentino, Florence, Italy; Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Florence, Italy; Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), 50019 Sesto Fiorentino, Italy*; orcid.org/0000-0002-2376-3321; Email: ciofi@cerm.unifi.it

Authors

Deborah Grifagni − *Magnetic Resonance Center CERM, University of Florence, 50019 Sesto Fiorentino, Florence, Italy; Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Florence, Italy*

Elena Lenci − *Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Florence, Italy*; orcid.org/0000-0001-7408-2828

Alessia De Santis − *Magnetic Resonance Center CERM, University of Florence, 50019 Sesto Fiorentino, Florence, Italy; Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Florence, Italy*

Andrea Orsetti − *Magnetic Resonance Center CERM, University of Florence, 50019 Sesto Fiorentino, Florence, Italy; Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Florence, Italy*

Carlo Giorgio Barracchia − *Department of Biotechnology, University of Verona, 37134 Verona, Italy*; orcid.org/0000-0002-0488-6350

Filomena Tedesco − *Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Florence, Italy*

Raffaele Bellini Puglielli − *Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Florence, Italy*

Francesca Lucarelli − *Magnetic Resonance Center CERM, University of Florence, 50019 Sesto Fiorentino, Florence, Italy; Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Florence, Italy*

Angela Lauriola − *Department of Biotechnology, University of Verona, 37134 Verona, Italy*

Michael Assfalg − *Department of Biotechnology, University of Verona, 37134 Verona, Italy;* orcid.org/0000-0001-9331-3169

Francesca Cantini − *Magnetic Resonance Center CERM, University of Florence, 50019 Sesto Fiorentino, Florence, Italy; Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Florence, Italy*

Vito Calderone − *Magnetic Resonance Center CERM, University of Florence, 50019 Sesto Fiorentino, Florence, Italy; Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Florence, Italy;* orcid.org/0000-0002-7963-6241

Daniele Guardavaccaro − *Department of Biotechnology, University of Verona, 37134 Verona, Italy*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsmedchemlett.3c00498

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by "Fondo di beneficienza ed opere di carattere sociale e culturale" of Intesa Sanpaolo (project no. B/2021/0212), by "the European Union-NextGenerationEU-National Recovery and Resilience Plan, Mission 4 Component 2-Investment 1.5-THE-Tuscany Health Ecosystem-ECS00000017-CUP B83C22003920001, and by Italian Ministry of Education and Research (MUR) through Dipartimenti di Eccellenza 2023-2027 (DICUS 2.0) to the Department of Chemistry "Ugo Schiff" of the University of Florence. Moreover, Centro Piattaforme Tecnologiche of the University of Verona is acknowledged for providing access to the mass spectrometer and CD instruments.

■ ABBREVIATIONS

PROTAC, proteolysis targeting chimera; CoV, coronaviruses; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; $3CL^{Pro}$, 3-chymotrypsin-like protease; CRBN, cereblon; NMR, nucler magnetic resonance; HSQC, heteronuclear single quantum coherence; CD, circular dichroism; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; HA, hemagglutinin; MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide

■ REFERENCES

(1) Li, K.; Crews, C. M. *Chem. Soc. Rev.* **2022**, *51*, 5214−5236.
(2) Cowan, A. D.; Ciulli, A. *Annu. Rev. Biochem.* **2022**, *91*, 295−319.
(3) Toure, M.; Crews, C. M. *Angew. Chem., Int. Ed. Engl.* **2016**, *55*, 1966−1973.
(4) Farnaby, W.; Koegl, M.; McConnell, D. B.; Ciulli, A. *Curr. Opin. Pharmacol.* **2021**, *57*, 175−183.
(5) Mullard, A. *Nat. Rev. Drug. Discovery* **2021**, *20*, 247−250.
(6) Garber, K. *Nat. Biotechnol.* **2022**, *40*, 12−16.
(7) Chirnomas, D.; Hornberger, K. R.; Crews, C. M. *Nat. Rev. Clin. Oncol.* **2023**, *20*, 265−278.
(8) Desantis, J.; Goracci, L. *Future Med. Chem.* **2022**, *14*, 459−462.
(9) Yao, T.; Xiao, H.; Wang, H.; Xu, X. *Int. J. Mol. Sci.* **2022**, *23*, 10328.
(10) Venkatesan, J.; Murugan, D.; Rangasamy, L. A. *Antibiotics* **2022**, *11*, 1717.
(11) Wang, C.; Zhang, Y.; Xing, D.; Zhang, R. *Bioorg. Chem.* **2021**, *114*, No. 105109.
(12) Espinoza-Chávez, R. M.; Salerno, A.; Liuzzi, A.; Ilari, A.; Milelli, A.; Uliassi, E.; Bolognesi, M. L. *ACS Bio & Med. Chem. Au* **2023**, *3*, 32−45.
(13) Montrose, K.; Krissansen, G. W. *Biochem. Biophys. Res. Commun.* **2014**, *453*, 735−740.
(14) de Wispelaere, M.; Du, G.; Donovan, K. A.; Zhang, T.; Eleuteri, N. A.; Yuan, J. C.; Kalabathula, J.; Nowak, R. P.; Fischer, E. S.; Gray, N. S.; Yang, P. L. *Nat. Commun.* **2019**, *10*, 3468.
(15) Li, H.; Wang, S.; Ma, W.; Cheng, B.; Yi, Y.; Ma, X.; Xiao, S.; Zhang, L.; Zhou, D. *J. Med. Chem.* **2022**, *65*, 7154−7169.
(16) Chatterjee, P.; Ponnapati, M.; Kramme, C.; Plesa, A. M.; Church, G. M.; Jacobson, J. M. *Commun. Biol.* **2020**, *3*, 715.
(17) Ziebuhr, J.; Snijder, E. J.; Gorbalenya, A. E. *J. Gen. Virol.* **2000**, *81*, 853−879.
(18) Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R. *Science* **2003**, *300*, 1763−1767.
(19) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. *Science* **2020**, *368*, 409−412.
(20) Liu, Y.; Liang, C.; Xin, L.; Ren, X.; Tian, L.; Ju, X.; Li, H.; Wang, Y.; Zhao, Q.; Liu, H.; Cao, W.; Xie, X.; Zhang, D.; Wang, Y.; Jian, Y. *Eur. J. Med. Chem.* **2020**, *206*, No. 112711.
(21) Martinez-Ortiz, W.; Zhou, M. M. *Drug Discovery Today* **2020**, *25*, 1894−1896.
(22) Shaheer, M.; Singh, R.; Sobhia, M. E. *J. Biomol. Struct. Dyn.* **2022**, *40*, 10905−10917.
(23) Zhang, L.; Lin, D.; Kusov, Y.; Nian, Y.; Ma, Q.; Wang, J.; von Brunn, A.; Leyssen, P.; Lanko, K.; Neyts, J.; de Wilde, A.; Snijder, E. J.; Liu, H.; Hilgenfeld, R. *J. Med. Chem.* **2020**, *63*, 4562−4578.
(24) Qiao, J.; Li, Y. S.; Zeng, R.; Liu, F. L.; Luo, R. H.; Huang, C.; Wang, Y. F.; Zhang, J.; Quan, B.; Shen, C.; Mao, X.; Liu, X.; Sun, W.; Yang, W.; Ni, X.; Wang, K.; Xu, L.; Duan, Z. L.; Zou, Q. C.; Zhang, H. L.; Qu, W.; Long, Y. H. P.; Li, M. H.; Yang, R. C.; Liu, X.; You, J.; Zhou, Y.; Yao, R.; Li, W. P.; Liu, J. M.; Chen, P.; Liu, Y.; Lin, G. F.; Yang, X.; Zou, J.; Li, L.; Hu, Y.; Lu, G. W.; Li, W. M.; Wei, Y. Q.; Zheng, Y. T.; Lei, J.; Yang, S. *Science* **2021**, *371*, 1374−1378.
(25) La Monica, G.; Bono, A.; Lauria, A.; Martorana, A. *J. Med. Chem.* **2022**, *65*, 12500−12534.
(26) Fan, K.; Wei, P.; Feng, Q.; Chen, S.; Huang, C.; Ma, L.; Lai, B.; Pei, J.; Liu, Y.; Chen, J.; Lai, L. *J. Biol. Chem.* **2004**, *279*, 1637−1642.
(27) Lenci, E.; Trabocchi, A. *Chem. Soc. Rev.* **2020**, *49*, 3262−3277.
(28) Yang, Q.; Zhao, J.; Chen, D.; Wang, Y. *Mol. Biomed.* **2021**, *2*, 23.
(29) Pickart, C. M.; Fushman, D. *Curr. Opin. Chem. Biol.* **2004**, *8*, 610−616.
(30) Fischer, E. S.; Böhm, K.; Lydeard, J. R.; Yang, H.; Stadler, M. B.; Cavadini, S.; Nagel, J.; Serluca, F.; Acker, V.; Lingaraju, G. M.; Tichkule, R. B.; Schebesta, M.; Forrester, W. C.; Schirle, M.; Hassiepen, U.; Ottl, J.; Hild, M.; Beckwith, R. E. J.; Harper, J. W.; Jenkins, J. L.; Thomä, N. H. *Nature* **2014**, *512*, 49−53.
(31) Ito, T.; Ando, H.; Suzuki, T.; Ogura, T.; Hotta, K.; Imamura, Y.; Yamaguchi, Y.; Handa, H. *Science* **2010**, *327*, 1345−1350.
(32) Petzold, G.; Fischer, E. S.; Thomä, N. H. *Nature* **2016**, *532*, 127−130.
(33) Cao, C.; He, M.; Wang, L.; He, Y.; Rao, Y. *Chem. Soc. Rev.* **2022**, *51*, 7066−7114.
(34) Ishida, T.; Ciulli, A. *SLAS Discovery Adv. Life Sci. R. D* **2021**, *26*, 484−502.
(35) Gadd, M. S.; Testa, A.; Lucas, X.; Chan, K. H.; Chen, W.; Lamont, D. J.; Zengerle, M.; Ciulli, A. *Nat. Chem. Biol.* **2017**, *13*, 514−521.
(36) Han, X.; Zhao, L.; Xiang, W.; Qin, C.; Miao, B.; McEachern, D.; Wang, Y.; Metwally, H.; Wang, L.; Matvekas, A.; Wen, B.; Sun, D.; Wang, S. *J. Med. Chem.* **2021**, *64*, 12831−12854.
(37) Pedersen, N. C.; Kim, Y.; Liu, H.; Galasiti Kankanamalage, A. C.; Eckstrand, C.; Groutas, W. C.; Bannasch, M.; Meadows, J. M.; Chang, K. O. *J. Feline Med. Surg.* **2018**, *20*, 378−392.

(38) Kim, Y.; Liu, H.; Galasiti Kankanamalage, A. C.; Weerasekara, S.; Hua, D. H.; Groutas, W. C.; Chang, K. O.; Pedersen, N. C. *PLoS Pathog.* **2016**, *12*, No. e1005531.
(39) Vuong, W.; Khan, M. B.; Fischer, C.; Arutyunova, E.; Lamer, T.; Shields, J.; Saffran, H. A.; McKay, R. T.; van Belkum, M. J.; Joyce, M. A.; Young, H. S.; Tyrrell, D. L.; Vederas, J. C.; Lemieux, M. J. *Nat. Commun.* **2020**, *11*, 4282.
(40) Fu, L.; Ye, F.; Feng, Y.; Yu, F.; Wang, Q.; Wu, Y.; Zhao, C.; Sun, H.; Huang, B.; Niu, P.; Song, H.; Shi, Y.; Li, X.; Tan, W.; Qi, J.; Gao, G. F. *Nat. Commun.* **2020**, *11*, 4417.
(41) Cáceres, C. J.; Cardenas-Garcia, S.; Carnaccini, S.; Seibert, B.; Rajao, D. S.; Wang, J.; Perez, D. R. *Sci. Rep.* **2021**, *11*, 9609.
(42) Kim, Y.; Shivanna, V.; Narayanan, S.; Prior, A. M.; Weerasekara, S.; Hua, D. H.; Kankanamalage, A. C.; Groutas, W. C.; Chang, K. O. *J. Virol.* **2015**, *89*, 4942−4950.
(43) Halford, B. *C&EN Global Enterprise* **2021**, *99*, 5.
(44) Gao, S.; Sylvester, K.; Song, L.; Claff, T.; Jing, L.; Woodson, M.; Weiße, R. H.; Cheng, Y.; Schäkel, L.; Petry, M.; Gütschow, M.; Schiedel, A. C.; Sträter, N.; Kang, D.; Xu, S.; Toth, K.; Tavis, J.; Tollefson, A. E.; Müller, C. E.; Liu, X.; Zhan, P. *J. Med. Chem.* **2022**, *65*, 13343−13364.
(45) Vuong, W.; Fischer, C.; Khan, M. B.; van Belkum, M. J.; Lamer, T.; Willoughby, K. D.; Lu, J.; Arutyunova, E.; Joyce, M. A.; Saffran, H. A.; Shields, J. A.; Young, H. S.; Nieman, J. A.; Tyrrell, D. L.; Lemieux, M. J.; Vederas, J. C. *Eur. J. Med. Chem.* **2021**, *222*, No. 113584.
(46) Iketani, S.; Mohri, H.; Culbertson, B.; Hong, S. J.; Duan, Y.; Luck, M. I.; Annavajhala, M. K.; Guo, Y.; Sheng, Z.; Uhlemann, A. C.; Goff, S. P.; Sabo, Y.; Yang, H.; Chavez, A.; Ho, D. D. *Nature* **2023**, *613*, 558−564.
(47) Ghahremanpour, M. M.; Tirado-Rives, J.; Deshmukh, M.; Ippolito, J. A.; Zhang, C. H.; Cabeza de Vaca, I.; Liosi, M. E.; Anderson, K. S.; Jorgensen, W. L. *ACS Med. Chem. Lett.* **2020**, *11*, 2526−2533.
(48) Mody, V.; Ho, J.; Wills, S.; Mawri, A.; Lawson, L.; Ebert, M. C. C. J. C.; Fortin, G. M.; Rayalam, S.; Taval, S. *Commun. Biol.* **2021**, *4*, 93.
(49) Niesor, E. J.; Boivin, G.; Rhéaume, E.; Shi, R.; Lavoie, V.; Goyette, N.; Picard, M. E.; Perez, A.; Laghrissi-Thode, F.; Tardif, J. C. *ACS Omega* **2021**, *6*, 16584−16591.
(50) Pellecchia, M.; Sem, D. S.; Wüthrich, K. *Nat. Rev. Drug. Discovery* **2002**, *1*, 211−219.
(51) Grifagni, D.; Calderone, V.; Giuntini, S.; Cantini, F.; Fragai, M.; Banci, L. *Chem. Commun.* **2021**, *57*, 7910−7913.
(52) Massai, L.; Grifagni, D.; De Santis, A.; Geri, A.; Cantini, F.; Calderone, V.; Banci, L.; Messori, L. *Biomolecules* **2022**, *12*, 1675.
(53) Lee, C. C.; Kuo, C. J.; Ko, T. P.; Hsu, M. F.; Tsui, Y. C.; Chang, S. C.; Yang, S.; Chen, S. J.; Chen, H. C.; Hsu, M. C.; Shih, S. R.; Liang, P. H.; Wang, A. H. J. *J. Biol. Chem.* **2009**, *284*, 7646−7655.
(54) Owen, D. R.; Allerton, C. M. N.; Anderson, A. S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R. D.; Carlo, A.; Coffman, K. J.; Dantonio, A.; Di, L.; Eng, H.; Ferre, R.; Gajiwala, K. S.; Gibson, S. A.; Greasley, S. E.; Hurst, B. L.; Kadar, E. P.; Kalgutkar, A. S.; Lee, J. C.; Lee, J.; Liu, W.; Mason, S. W.; Noell, S.; Novak, J. J.; Obach, R. S.; Ogilvie, K.; Patel, N. C.; Pettersson, M.; Rai, D. K.; Reese, M. R.; Sammons, M. F.; Sathish, J. G.; Singh, R. S. P.; Steppan, C. M.; Stewart, A. E.; Tuttle, J. B.; Updyke, L.; Verhoest, P. R.; Wei, L.; Yang, Q.; Zhu, Y. *Science* **2021**, *374*, 1586−1593.
(55) Ciulli, A. H.; Hsia, O. PROTAC Degraders Mechanism, Recent Advances, and Future Challenges. In *Protein Homeostasis in Drug Discovery: A Chemical Biology Perspective*; Jones, M. K. L. H., Ed.; John Wiley & Sons, 2022; pp 317−356.
(56) Gabizon, R.; Shraga, A.; Gehrtz, P.; Livnah, E.; Shorer, Y.; Gurwicz, N.; Avram, L.; Unger, T.; Aharoni, H.; Albeck, S.; Brandis, A.; Shulman, Z.; Katz, B. Z.; Herishanu, Y.; London, N. *J. Am. Chem. Soc.* **2020**, *142*, 11734−11742.

https://doi.org/10.1021/acsmedchemlett.3c00498
ACS Med. Chem. Lett. 2024, 15, 250−257

Development of a GC-376 based peptidomimetic PROTAC as a Degrader of 3-Chymotrypsin-like Protease of SARS-CoV-2

Deborah Grifagni[1,2], Elena Lenci[2], Alessia De Santis[1,2], Andrea Orsetti[1,2], Carlo Giorgio Barracchia[3], Filomena Tedesco[2], Raffaele Bellini Puglielli[2], Francesca Lucarelli[1,2], Angela Lauriola[3], Michael Assfalg[3], Francesca Cantini[1,2], Vito Calderone[1,2], Daniele Guardavaccaro[3], Andrea Trabocchi[2,*], Mariapina D'Onofrio[3,*], Simone Ciofi-Baffoni[1,2,4,*]

[1] *Magnetic Resonance Center CERM, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy. E-mail: ciofi@cerm.unifi.it*

[2] *Department of Chemistry "Ugo Schiff", University of Florence, Via della Lastruccia 3-13, 50019 Sesto Fiorentino, Florence, Italy. E-mail: andrea.trabocchi@unifi.it;*

[3] *Department of Biotechnology, University of Verona, Strada le Grazie 15, 37134 Verona, Italy. E-mail: mariapina.donofrio@univr.it*

[4] *Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (CIRMMP), Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy.*

Supporting Information

SARS-CoV-2 3CL^{Pro} expression and purification

Dimeric SARS-CoV-2 3CL^{Pro} was expressed and purified as previously described.[7] The sequence encoding SARS-CoV-2 3CL^{Pro}, once optimized for *E. coli* expression, was cloned into pET-28a(+) between the restriction sites NdeI/XhoI to produce a monomeric form of the protein. This construct results in an N-terminal 6xHis tagged protein with a thrombin recognition site that divides the His tag from the SARS-CoV-2 3CL^{Pro} sequence. BL21(DE3) cells were transformed by the plasmid. A single colony was picked to inoculate a 100 mL pre-culture in LB medium containing 50 mg/mL kanamycin. After overnight incubation at 37 °C, 10 mL of the pre-culture was used to inoculate 1 L of M9 minimal growth medium (1.2 g/L ^{15}N-labelled $(NH_4)_2SO_4$). When an optical density of 0.7-0.8 (at a wavelength of 600 nm) was reached, the protein overexpression was induced with 0.5 mM isopropyl β-d-1-thiogalactopyranoside. The culture was then shaken at 35 °C for 6 hours and centrifuged at 5000 rpm for 20 minutes. The resulting pellet was resuspended in 40 mL lysis buffer (of 50 mM KPi, pH 8.0, 300 mM NaCl, 10 mM DL-dithiothreitol (DTT) buffer). The cell suspension was lysed by sonication and the soluble cellular fraction was obtained by centrifugation at 4 °C at 35000 rpm for 35 minutes. The supernatant was collected, diluted up to 100 mL with the lysis buffer and then loaded onto a 5 mL HisTrap FF column, pre-equilibrated with the same buffer. An imidazole stepwise-gradient was applied to wash the column. The concentration of imidazole for each step (30 mL) was 5 mM, 20 mM, 40 mM, 300 mM and 500 mM. The pure His-tagged protein was detected in the 300 mM imidazole fraction by running an SDS-PAGE electrophoresis of the different washing-step fractions. The protein solution was buffer exchanged to in 20 mM HEPES buffer, 50 mM NaCl, 0.5 mM TCEP, pH 6.5 resulting in a final protein yield of 25 mg/L of culture.
The monomeric and dimeric states of SARS-CoV-2 3CL^{Pro} proteins here produced were verified by analytical gel filtration, as shown in **Fig. S5**.

Enzyme inhibition kinetics assay

The synthesized PROTACs were assayed through a fluorometric assay using the fluorogenic substrate Hilyte Fluor-488-ESATLQSGLRKAK-(QXL-520)-NH_2 (Anaspec). All of the measurements were performed in 96-well plates with a Fluostar Optima microplate reader (BMG Labtech, Ortenberg, Germany) analogously to what previously performed.[8] Excitation and emission wavelengths were 490 and 520 nm, respectively. All incubations were performed at 30 °C in 25 mmol of N-(2-hydroxyethyl)piperazine-N′-ethanesulfonic acid (HEPES), 0.2% Tween-20 at pH 7.0. The inhibitors were preincubated with SARS-CoV-2 3CL^{Pro} enzyme (60 nM) for 10 min at 30 °C before the reaction was started by the addition of the fluorogenic substrate (1 μM). The increase of fluorescence signal was monitored over 30 min (λex = 490 nm, λem = 520 nm) at 30 °C. The percentages of inhibition for the tested PROTAC molecules were determined through the equation $(1 - V_s/V_o) \times 100$, where V_s is

the initial velocity in the presence of the inhibitor and Vo is the initial velocity of the uninhibited reaction. The IC_{50} values were obtained by dose−response measurements using an inhibitor range of concentrations 1 nM to 1 mM. All the experiments were performed in triplicate, and data collected were analyzed using GraphPad 5.0 Software Package (GraphPad Prism, Inc., San Diego, CA).

Enzyme activity recovery assay

SARS-CoV-2 $3CL^{Pro}$ (600 nM) was incubated with or without 1 mM of PROTAC **1** for 30 minutes, before the reaction was started by the addition of the fluorogenic substrate (1 μM). These mixtures were diluted 100-fold into reaction buffer containing the fluorogenic substrates to initiate reaction. The progress curve for these samples was measured and the initial velocity was measured and compared to that of enzyme incubated and diluted in the absence of PROTAC **1**. All the experiments were performed in triplicate, and data collected were analyzed using GraphPad 5.0 Software Package (GraphPad Prism, Inc., San Diego, CA).

NMR spectroscopy

^{1}H-^{15}N HSQC experiments were performed at 298 K in 20 mM HEPES buffer, 50 mM NaCl, 0.5 mM TCEP, 5 mM DTT, pH 6.5, 10% (v/v) D_2O. These NMR spectra were recorded on Bruker AVANCE 950 MHz, processed using the standard Bruker software (Topspin) and analyzed with CARA program.[9] In order to monitor the interaction between SARS-CoV-2 $3CL^{Pro}$ and the small molecules, NMR titration experiments were performed adding aliquots (up to a maximum of 20 μL) of the small molecule dissolved in d_3-acetonitrile to the NMR tube containing ^{15}N-labeled SARS-CoV-2 $3CL^{Pro}$ (0.2-0.3 mM) up to a 1:1 ratio. All NMR titration data were analysed comparing the ^{1}H-^{15}N HSQC spectra recorded along the additions of the small molecule with that of the initial state as well as with ^{1}H-^{15}N HSQC blank spectra of ^{15}N-labeled SARS-CoV-2 $3CL^{Pro}$ recorded titrating the protein with the same aliquot of d_3-acetonitrile, in order to identify the chemical shift changes induced by the solvent. In such a way, following the chemical shift changes observed in the ^{1}H-^{15}N HSQC maps along each stepwise titration we were able to assign selectively the residues affected by protein-protein interaction. Signals showing both chemical shift changes on a slow exchange on the NMR time scale and broadening beyond detection effects were considered to map the interaction surface on SARS-CoV-2 $3CL^{Pro}$ upon the binding with the small molecule. Chemical shift assignments of SARS-CoV-2 $3CL^{Pro}$ (under accession code BMRB: BMR50780)[10] and of the C145A variant of SARS-CoV-2 $3CL^{Pro}$ (under accession code BMRB: BMR51456)[11] were used to map the chemical shift changes, obtained upon the addition of one equivalent of PROTAC **1**, on the protein structure.

Fluorescence spectroscopy

Fluorescence measurements were performed on a Jasco FP-8500 spectrofluorometer (Jasco, Easton, MD, USA) with a 1 cm path-length quartz cuvette, at room temperature. The excitation wavelength was 290 nm (slit width 2.5 nm), and emission spectra were collected in the range of 300-450 nm. All experiments were performed at room temperature in 20 mM

potassium phosphate buffer at pH 7.8. Three spectra accumulations were averaged for each sample and the spectrum of the buffer was considered as a blank and subtracted. Fluorescence spectra were recorded on sample of dimeric SARS-CoV-2 $3CL^{Pro}$ (2 μM), in the absence or presence of PROTAC **1** (previously re-suspended in acetonitrile) at a 1:2 ratio (protein:molecule).

Far-UV circular dichroism (CD) spectroscopy

CD measurements were carried out on a Jasco J-1500 spectropolarimeter equipped with a Peltier type thermostated cell holder (Jasco, Easton, MD, USA). Far-UV spectra (200–260 nm) were recorded at 25 °C at a scan rate of 50 nm min^{-1}, a bandwidth of 1 nm, and an integration time of 2 s, in 0.1 cm cuvettes. Spectra were recorded on samples of dimeric SARS-CoV-2 $3CL^{Pro}$ in the absence or presence of PROTAC **1** or ligand **1'** at a 1:2 ratio. Five spectra accumulations were averaged for each sample and the spectrum of the buffer was considered as a blank and subtracted. The dimeric SARS-CoV-2 $3CL^{Pro}$ concentration was 6 μM.

Mass spectrometry

MALDI-TOF mass spectra were acquired on a Bruker Ultraflextreme MALDI-TOF/TOF instrument (Bruker Daltonics, Billerica, MA, USA) at the Centro Piattaforme Tecnologiche of the University of Verona. Samples were resuspended and acidified with TFA 0.1% solution and incubated for 30 min at room temperature. The two-layer method was used, 0.5 μL of sinapinic acid saturated solution in 30:70 (v/v) Acetonitrile: TFA 0.1% in water was deposited onto the target as a seed layer and left for drying. The protein sample was mixed 1:1 with a saturated solution of sinapinic acid in 30% acetonitrile. 0.5 μL of the sample/matrix mixture was then deposited onto the seed layer and left for drying on Ground steel MALDI target plate (Bruker Daltonics).

Crystallographic data

Crystals of apo SARS-CoV-2 $3CL^{Pro}$ were obtained in sitting drop by adding an aliquot of 2 μL of protein solution (20 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.8) to 2 μL of reservoir buffer (200 mM ammonium acetate, 20% PEG 3350 pH 7.0) and stored at 20 °C. The protein concentration in the sample was 5 mg/mL. The native crystals of apo SARS-CoV-2 $3CL^{Pro}$ were afterwards soaked in a 10% DMSO solution containing ligand **1'** having a 2-3-fold concentration with respect to the protein for about 2-3 days. The adduct with PROTAC **1** was, on the contrary, obtained through co-crystallization in the same conditions as above with a small molecule-protein ratio of ~3. Both adduct datasets were collected in-house, using a BRUKER D8 Venture diffractometer equipped with a PHOTON III detector, at 100 K; the crystals used for data collection were cryo-cooled using 25% ethylene glycol in the mother liquor. The ligand **1'** adduct crystal has been refined at 2.3 Å resolution whereas the PROTAC **1** adduct crystal reaches 2.0 Å resolution: they both belong to space group C2 with one molecule in the asymmetric unit, a solvent content of about 50%, and a mosaicity of 0.4-0.6°. The data were processed using the program XDS,[12] reduced and scaled using XSCALE,[12] and amplitudes were calculated using XSCALE.[12] Both

structures have been solved using the molecular replacement technique using 7NXH structure as the template model. The successful orientation and translation of the molecule within the crystallographic unit cell was determined with MOLREP.[13, 14] The refinement and water molecule fitting has been carried out using PHENIX,[15] applying default TLS restraints. In between the refinement cycles, the model was subjected to manual rebuilding using COOT.[16] The quality of the refined structures was assessed using the program MOLPROBITY.[17] Data processing and refinement statistics for both adducts are shown in **Table S1**. Coordinates and structure factors have been deposited at the PDB under the accession code 8OKB for ligand **1'** and 8OKC for PROTAC **1**.

Cell culture and drug treatment

HeLa cells and HEK293T cells (from American Type Culture Collection) were maintained in Dulbecco's modified Eagle's medium (DMEM; Thermo Fisher Scientific) containing 10% fetal calf serum, 100 U/mL penicillin, and 100 µg/mL streptomycin. PROTAC **1** was added at the indicated concentration.

Biochemical methods

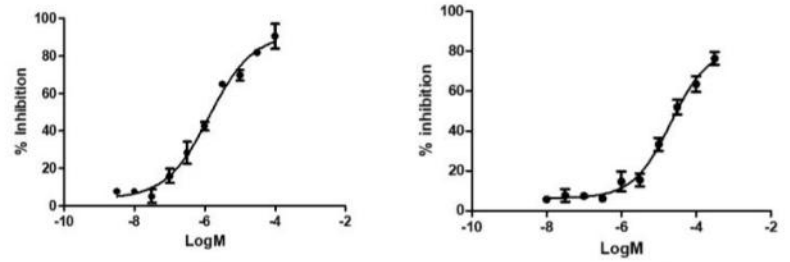
For preparation of cell extracts, cells were washed and collected in ice-cold PBS and lysed in lysis buffer (50 mM Tris pH 7.5, 250 mM NaCl, 0.1% Triton X-100, 1 mM EDTA, 50 mM NaF and protease and phosphatase inhibitors) for 30 min on ice, followed by 20 min centrifugation at 4 °C. Total cell extracts (30 µg per sample) were then analyzed by immunoblotting as described.[18] Mouse monoclonal antibodies were from Santa Cruz Biotechnology (Actin) and BioLegend (HA). Data are reported as mean ± SEM; n = 3.

Lentivirus-mediated gene transfer

Lentiviruses were produced in HEK293T cells by five-plasmid cotransfection.[18] Briefly, HEK293T cells were transfected with the pHAGE2 vector together with packaging vectors encoding Gag-Pol, Rev, Tat, and the G protein of the vesicular stomatitis virus (VSV) by the polyethylenimine method. The virus-containing supernatants were collected every 24 hours on two consecutive days starting 24 hours after transfection, filtered, and transferred onto HeLa cells in the presence of polybrene (4 µg/mL). The medium was replaced after 24 hours. Two days after transduction, Hela cells were selected with 2 µg/ml puromycin for 4 days.

Cell viability assay

Cells were seeded in 24-well plates at a density of 5×10^5 cells per well. After 24 hours, cells were exposed to increasing concentrations of PROTAC **1** for 72 h. 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide reagent (MTT) (Merck) was added to the cells at a final concentration of 0.5 mg/mL. Cells were incubated for 4 hours at 37 °C in a humidified incubator with 5% CO_2, until purple precipitates were visible. The medium was removed, and cells were rinsed once with PBS. DMSO was added to each well to dissolve the formazan precipitate, and the absorbance was read at OD = 535 nm. Data are reported as mean ± SEM; n = 3.



Ligand 1': IC_{50} = 1.35 ± 0.34 µM **PROTAC 1**: IC_{50} = 21.2 ± 5.8 µM

Figure S1. Inhibition curves for PROTAC 1 and ligand 1'. IC_{50} values of ligand 1' and PROTAC 1 are reported.

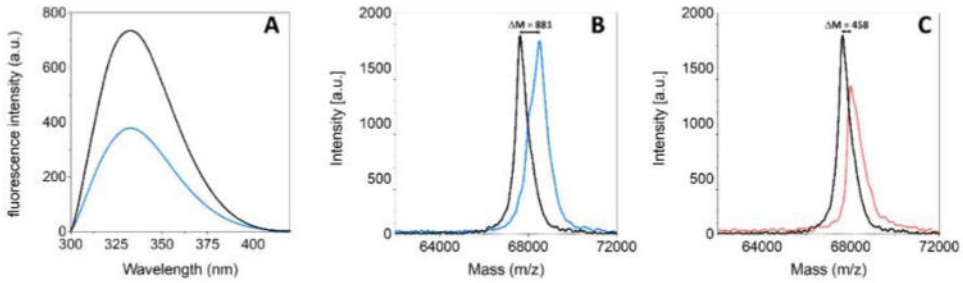


Figure S2. Investigating complex formation by fluorescence and binding mode of PROTAC 1 and ligand 1' to SARS-CoV-2 $3CL^{Pro}$ by mass spectrometry. (**A**) Fluorescence emission spectra (λ_{ex} 290 nm) measured on 2 µM SARS-CoV-2 $3CL^{Pro}$ in the absence (black) or in the presence of 4 µM of PROTAC **1** (blue). MALDI TOF mass spectrometry of SARS-CoV-2 $3CL^{Pro}$ samples in the absence (black, **B** and **C**) or in the presence of PROTAC **1** (**B**, blue) or ligand **1'** (**C**, salmon). ΔM: Mass increase in Da of SARS-CoV-2 $3CL^{Pro}$ upon incubation with PROTAC **1** or ligand **1'**.

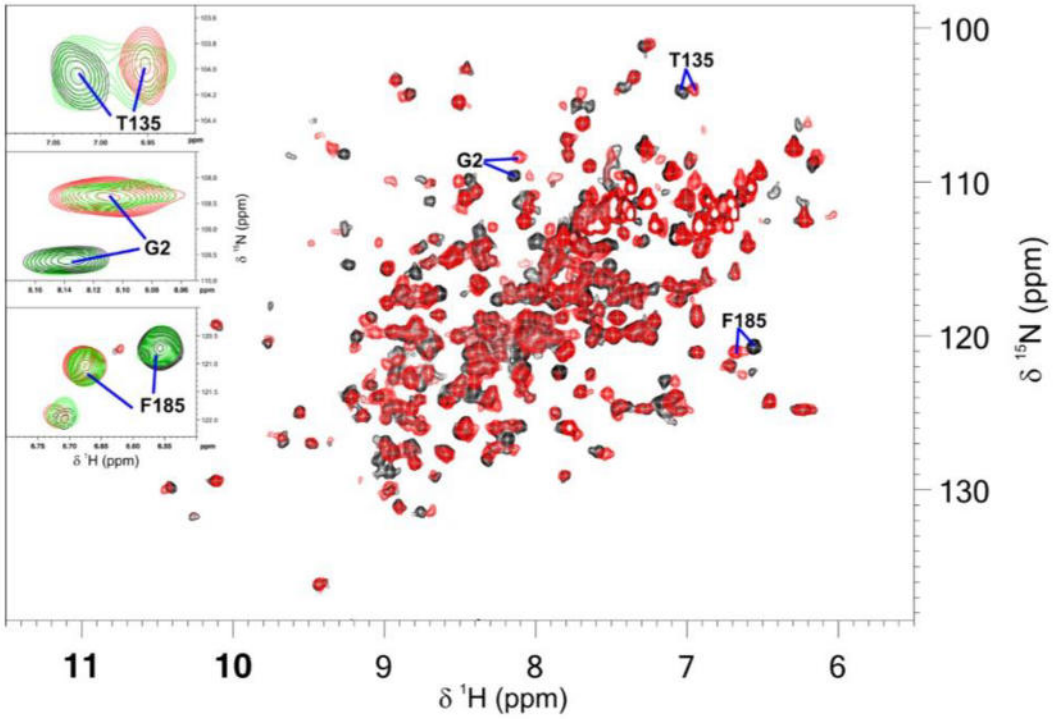


Figure S3. Ligand 1' interacts with dimeric SARS-CoV-2 3CL^{Pro}. Overlay of ^{1}H-^{15}N HSQC maps of ^{15}N-labelled SARS-CoV-2 3CL^{Pro} (black) and a 1:1 ^{15}N-labelled SARS-CoV-2 3CL^{Pro}-ligand **1'** mixture (red). In the insets, ^{1}H-^{15}N HSQC maps at 1:0 (black), 1:0.5 (green), 1:1 (red) SARS-CoV-2 3CL^{Pro}-ligand **1'** ratios are shown for some residues that are in slow exchange regime on the NMR time scale upon complex formation.

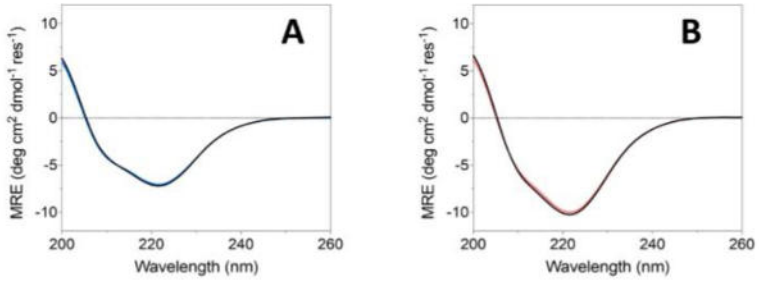


Figure S4. Monitoring structural changes by far-UV circular dichroism spectroscopy on SARS-CoV-2 3CL^{Pro} upon its interaction with PROTAC 1 and ligand 1'. Spectra acquired on SARS-CoV-2 3CL^{Pro} in the absence (black) (**A**, **B**) or in the presence of PROTAC **1** (blue) (**A**) or ligand **1'** (salmon) (**B**). MRE: Mean Residue Ellipticity.

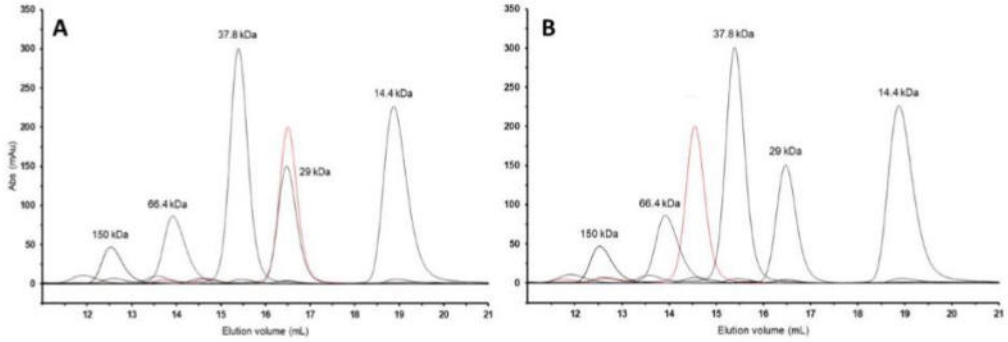


Figure S5. Analytical gel filtration of monomeric and dimeric SARS-CoV-2 3CL^{Pro}. Chromatograms of monomeric (**A**) and dimeric (**B**) SARS-CoV-2 3CL^{Pro} proteins (in red) compared with protein standard markers at 150, 66.4, 37.8, 29 and 14.4 kDa. The theoretical molecular weights of monomeric and dimeric SARS-CoV-2 3CL^{Pro} are 33.7 and 67.5 kDa. The samples were run on a Superdex 200 Increase 10/300 GL column using 25 mM MES, 150 mM NaCl, pH 6.5 running buffer.

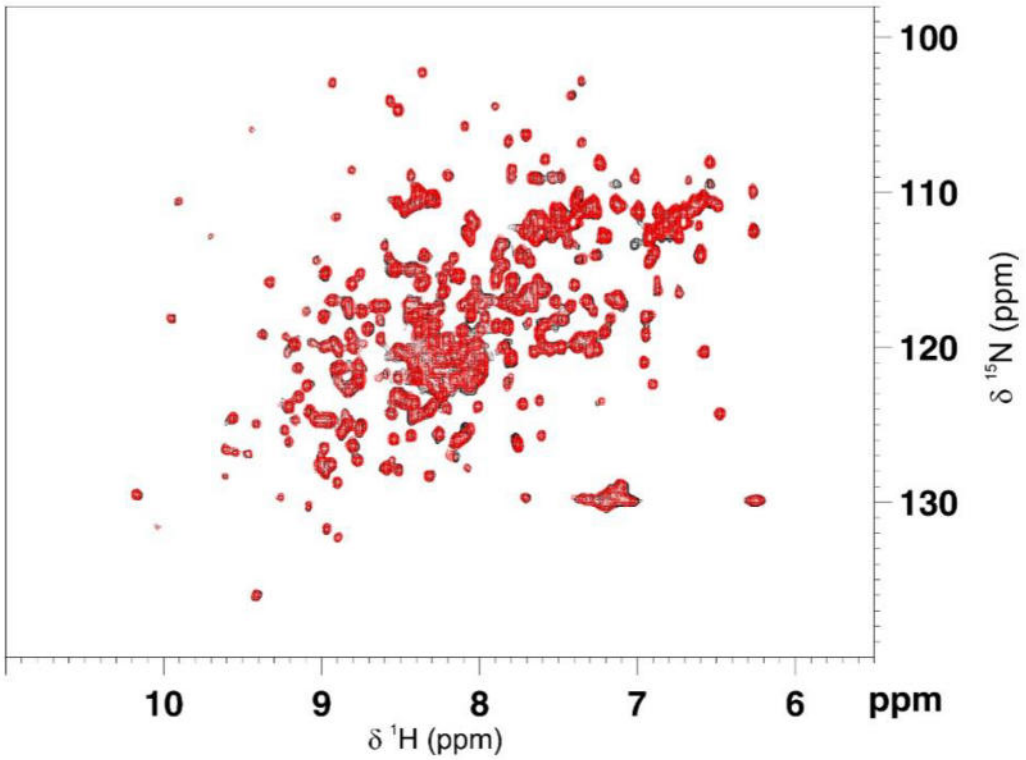


Figure S6. PROTAC 1 does not interact with monomeric SARS-CoV-2 3CL^{Pro}. Overlay of ^{1}H-^{15}N HSQC maps of ^{15}N-labelled SARS-CoV-2 3CL^{Pro} (red) and a 1:1 ^{15}N-labelled SARS-CoV-2 3CL^{Pro}-PROTAC **1** mixture (black).

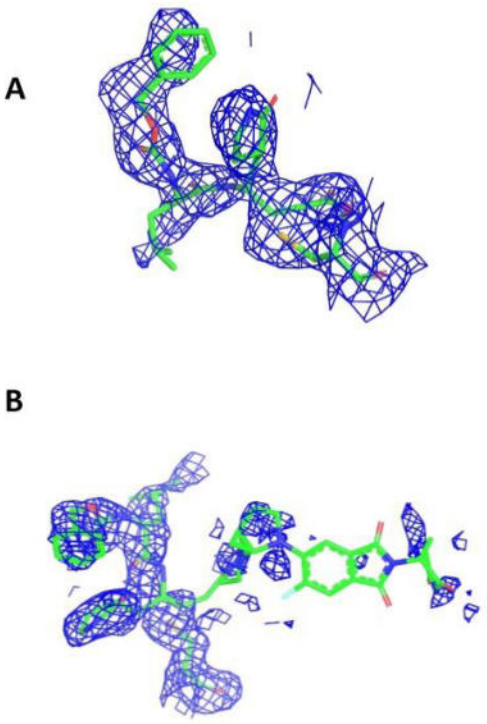


Figure S7. The electron density of ligand 1' and PROTAC 1. $2F_o$-F_c map of ligand **1'** and Cys145 (**A**) and PROTAC **1** and Cys145 (**B**), both contoured at 0.9s level.

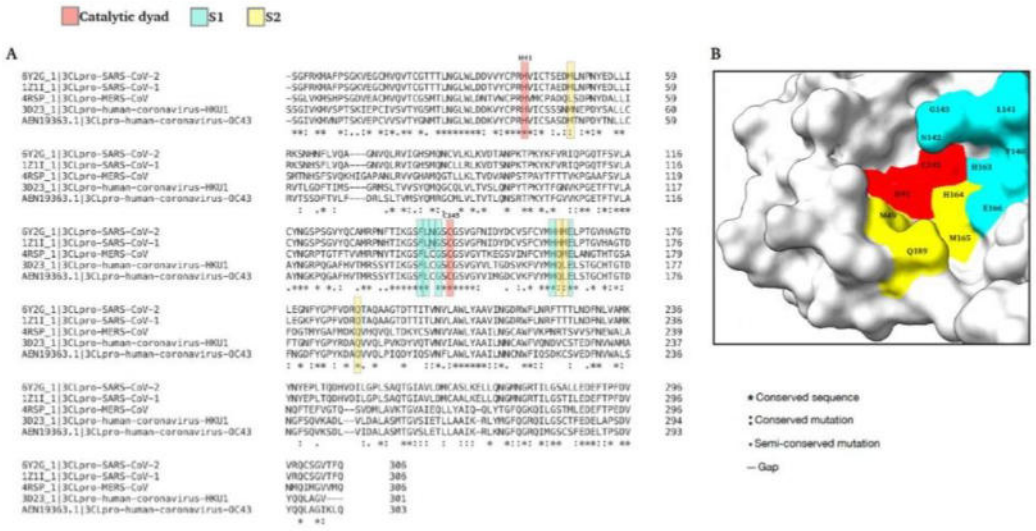


Figure S8. Sequence alignment of human coronavirus main proteases and mapping of the residues of the catalytic site of SARS-CoV-2 3CL^{Pro} interacting with both PROTAC 1 and ligand 1'. (**A**) Multiple sequence alignment by ClustalW between four human coronavirus main proteases. The residues, involved in the interaction with both PROTAC **1** and ligand **1'** and selected on the basis of the ligplot shown in Fig. 4B, have been highlighted according to the S1 and S2 sub-sites of the binding pocket. The catalytic dyad formed by His41 and Cys145 is also highlighted. (**B**) A view of the SARS-CoV-2 3CL^{Pro} binding pocket, using 8OKC as PDB template and Chimera-X as molecule visualization software, is shown. The catalytic dyad and the residues of S1 and S2 sites selected in panel A have been highlighted using the same color code.

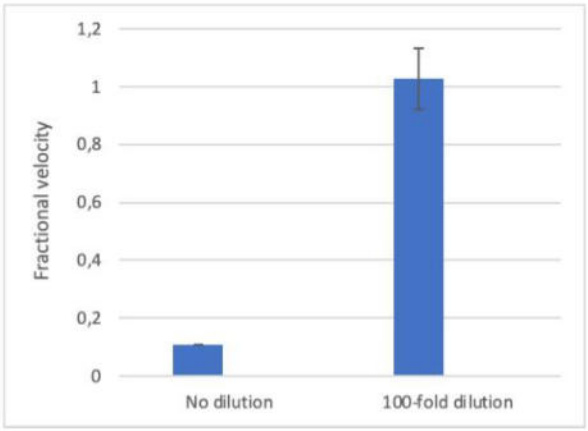
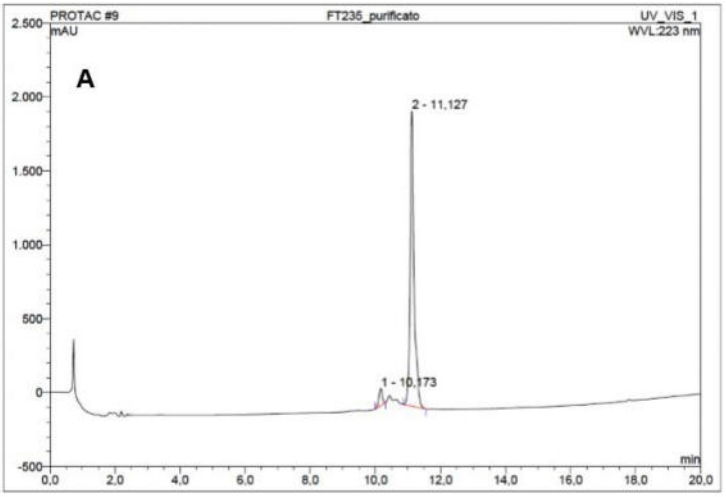
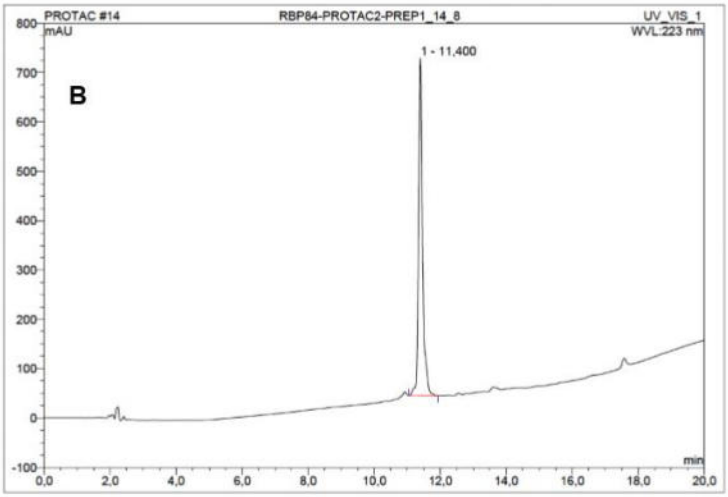


Figure S9. Recovery of the enzymatic activity after a 100-fold dilution of the SARS-CoV-2 3CL^{Pro}-PROTAC 1 adduct. Left (no dilution): Fractional velocity is calculated as the ratio between the slope of the curves obtained by the enzyme incubated in the presence and in the absence of the PROTAC **1** at initial situation (600 nM of the enzyme, 1 mM of PROTAC **1** and 1 μM of the fluorogenic substrate). Right (100-fold dilution): Fractional velocity is calculated as the ratio between the slope of the curves obtained by the enzyme incubated in the presence and in the absence of PROTAC **1** after 100-fold dilution (6 nM of the enzyme, 10 μM of PROTAC **1** and 1 μM of the fluorogenic substrate). Data are representative of three independent experiments.



No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	10,17	n.a.	114,455	13,592	4,61	n.a.	BMB
2	11,13	n.a.	1996,339	281,213	95,39	n.a.	BMB
Total:			2110,794	294,805	100,00	0,000	



No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	11,40	n.a.	685,173	96,865	100,00	n.a.	BMB
Total:			685,173	96,865	100,00	0,000	

Figure S10. HPLC chromatograms of PROTAC 1 (A) and PROTAC 2 (B).

Table S1. Data collection and refinement statistics of SARS-CoV-2 3CL^{Pro} complexed with ligand 1' and PROTAC 1.

	SARS-CoV-2 3CL^{Pro} complexed with ligand **1'**	SARS-CoV-2 3CL^{Pro} complexed with PROTAC **1**
Wavelength Å	1.541	1.541
Resolution range (Å)	44.14 - 2.31 (2.393 - 2.31)[1]	27.9 - 2.0 (2.072 - 2.0)
Space group	C 1 2 1	C 1 2 1
Unit cell (Å, °)	115.13 53.43 45.14 90 102.09 90	114.12 53.16 45.27 90 102.11 90
Total reflections	32842 (1859)	81051 (4064)
Unique reflections	11338 (929)	16697 (1211)
Multiplicity	2.9 (2.0)	4.9 (3.4)
Completeness (%)	96.11 (87.82)	92.40 (78.61)
Mean I/sigma(I)	4.80 (0.98)	7.77 (1.58)
Wilson B-factor	35.13	23.93
R-merge	0.2172 (0.9184)	0.1662 (0.828)
R-pim	0.1463 (0.7562)	0.08222 (0.4965)
CC1/2	0.969 (0.335)	0.99 (0.491)
Reflections used in refinement	11464 (926)	16689 (1211)
Reflections used for R-free	573 (46)	835 (60)
R-work	0.2195 (0.3519)	0.2156 (0.3396)
R-free	0.2729 (0.3931)	0.2509 (0.3607)
Number of non-hydrogen atoms	2457	2593
Macromolecules	2368	2368
Ligands	33	64
Solvent	56	161
Protein residues	306	306

RMS (bonds)	0.003	0.032
RMS (angles)	0.61	1.39
Ramachandran favoured (%)	96.05	95.72
Ramachandran allowed (%)	3.95	3.95
Ramachandran outliers (%)	0.00	0.33
Average B-factor	43.52	34.17
Macromolecules	43.41	31.62
Ligands	62.24	130.45
Solvent	37.42	33.31

[1]Statistics for the highest-resolution shell are shown in parentheses.

References

(1) McPherson, C. G.; Caldwell, N.; Jamieson, C.; Simpson, I.; Watson, A. J. B. Amidation of unactivated ester derivatives mediated by trifluoroethanol. *Organic & biomolecular chemistry* **2017**, *15*, 3507-3518.

(2) Shi, H.; Liu, K.; Leong, W. W.; Yao, S. Q. Expedient solid-phase synthesis of both symmetric and asymmetric diol libraries targeting aspartic proteases. *Bioorg Med Chem Lett* **2009**, *19*, 3945-3948.

(3) Vuong, W.; Khan, M. B.; Fischer, C.; Arutyunova, E.; Lamer, T.; Shields, J.; Saffran, H. A.; McKay, R. T.; van Belkum, M. J.; Joyce, M. A.; et al. Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nat Commun* **2020**, *11*, 4282.

(4) Gupta, M. K.; Jena, C. K.; Balachandra, C.; Sharma, N. K. Unusual Pseudopeptides: Syntheses and Structural Analyses of Ethylenediprolyl Peptides and Their Metal Complexes with Cu(II) Ion. *J Org Chem* **2021**, *86*, 16327-16336.

(5) Fuse, S.; Masuda, K.; Otake, Y.; Nakamura, H. Peptide-Chain Elongation Using Unprotected Amino Acids in a Micro-Flow Reactor. *Chemistry* **2019**, *25*, 15091-15097.

(6) Dai, W.; Jochmans, D.; Xie, H.; Yang, H.; Li, J.; Su, H.; Chang, D.; Wang, J.; Peng, J.; Zhu, L.; et al. Design, Synthesis, and Biological Evaluation of Peptidomimetic Aldehydes as Broad-Spectrum Inhibitors against Enterovirus and SARS-CoV-2. *J Med Chem* **2022**, *65*, 2794-2808.

(7) Grifagni, D.; Calderone, V.; Giuntini, S.; Cantini, F.; Fragai, M.; Banci, L. SARS-CoV-2 M(pro) inhibition by a zinc ion: structural features and hints for drug design. *Chemical communications (Cambridge, England)* **2021**, *57*, 7910-7913.

(8) Alcaraz, L. A.; Banci, L.; Bertini, I.; Cantini, F.; Donaire, A.; Gonnelli, L. MMP-inhibitor interaction: the solution structure of the catalytic domain of human matrix metalloproteinase-3 with three inhibitors. *J. Biol. Inorg. Chem* **2007**, *12*, 1197-1206.

(9) Keller, R. *The Computer Aided Resonance Assignment Tutorial*; CANTINA Verlag, 2004.

(10) Cantrelle, F. X.; Boll, E.; Brier, L.; Moschidi, D.; Belouzard, S.; Landry, V.; Leroux, F.; Dewitte, F.; Landrieu, I.; Dubuisson, J.; et al. NMR Spectroscopy of the Main Protease of SARS-CoV-2 and Fragment-Based Screening Identify Three Protein Hotspots and an Antiviral Fragment. *Angew Chem Int Ed Engl* **2021**, *60*, 25428-25435.

(11) Robertson, A. J.; Ying, J.; Bax, A. NMR Observation of Sulfhydryl Signals in SARS-CoV-2 Main Protease Aids Structural Studies. *Chembiochem* **2022**, *23*, e202200471.

(12) Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr* **2010**, *66*, 125-132.
(13) Vagin, A.; Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr* **1997**, *30*, 1022-1025.
(14) Vagin, A. A.; Teplyakov, A. An approach to multi-copy search in molecular replacement. *Acta Crystallogr D Biol Crystallogr* **2000**, *56*, 1622-1624.
(15) Adams, P. D.; Afonine, P. V.; Bunkòczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr* **2010**, *66*, 213-221.
(16) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 486-501.
(17) Chen, V. B.; Arendall, W. B., III; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr* **2010**, *66*, 12-21.
(18) Magliozzi, R.; Low, T. Y.; Weijts, B. G.; Cheng, T.; Spanjaard, E.; Mohammed, S.; van Veen, A.; Ovaa, H.; de Rooij, J.; Zwartkruis, F. J.; et al. Control of epithelial cell migration and invasion by the IKKβ- and CK1α-mediated degradation of RAPGEF2. *Dev Cell* **2013**, *27*, 574-585.

3.4.2. - A Structural Investigation of the Interaction between a GC-376-Based Peptidomimetic PROTAC and Its Precursor with the Viral Main Protease of Coxsackievirus B3

Alessia De Santis [1,2], Deborah Grifagni [1,2], Andrea Orsetti [1,2], Elena Lenci [2], Antonio Rosato[1,2], Mariapina D'Onofrio [3], Andrea Trabocchi [2], Simone Ciofi-Baffoni [1,2], Francesca Cantini[1,2,*], Vito Calderone [1,2,*]

[1]Magnetic Resonance Center CERM, University of Florence

[2]Department of Chemistry, University of Florence

[3]Department of Biotechnology, University of Verona

* Correspondence: cantini@cerm.unifi.it (F.C.); calderone@cerm.unifi.it (V.C.)

Published: 10.3390/biom14101260

biomolecules

MDPI

Article

A Structural Investigation of the Interaction between a GC-376-Based Peptidomimetic PROTAC and Its Precursor with the Viral Main Protease of Coxsackievirus B3

Alessia De Santis [1,2], Deborah Grifagni [1,2], Andrea Orsetti [1,2], Elena Lenci [2], Antonio Rosato [1,2], Mariapina D'Onofrio [3], Andrea Trabocchi [2], Simone Ciofi-Baffoni [1,2], Francesca Cantini [1,2,*] and Vito Calderone [1,2,*]

[1] Magnetic Resonance Center CERM, University of Florence, Via Luigi Sacconi 6, Sesto Fiorentino, 50019 Florence, Italy; desantis@cerm.unifi.it (A.D.S.); grifagni@cerm.unifi.it (D.G.); a.orsetti@uu.nl (A.O.); rosato@cerm.unifi.it (A.R.); ciofi@cerm.unifi.it (S.C.-B.)
[2] Department of Chemistry, University of Florence, Via della Lastruccia 3-13, Sesto Fiorentino, 50019 Florence, Italy; elena.lenci@unifi.it (E.L.); andrea.trabocchi@unifi.it (A.T.)
[3] Department of Biotechnology, University of Verona, Strada le Grazie 15, 37134 Verona, Italy; mariapina.donofrio@univr.it
* Correspondence: cantini@cerm.unifi.it (F.C.); calderone@cerm.unifi.it (V.C.)

Abstract: The conservation of the main protease in viral genomes, combined with the absence of a homologous protease in humans, makes this enzyme family an ideal target for developing broad-spectrum antiviral drugs with minimized host toxicity. GC-376, a peptidomimetic 3CL protease inhibitor, has shown significant efficacy against coronaviruses. Recently, a GC-376-based PROTAC was developed to target and induce the proteasome-mediated degradation of the dimeric SARS-CoV-2 $3CL^{Pro}$ protein. Extending this approach, the current study investigates the application of the GC-376 PROTAC to the $3C^{Pro}$ protease of enteroviruses, specifically characterizing its interaction with CVB3 $3C^{Pro}$ through X-ray crystallography, NMR (Nuclear Magnetic Resonance) and biochemical techniques. The crystal structure of CVB3 $3C^{Pro}$ bound to the GC-376 PROTAC precursor was obtained at 1.9 Å resolution. The crystallographic data show that there are some changes between the binding of CVB3 $3C^{Pro}$ and SARS-CoV-2 $3CL^{Pro}$, but the overall similarity is strong (RMSD on C-alpha 0.3 Å). The most notable variation is the orientation of the benzyloxycarbonyl group of GC-376 with the S4 subsite of the proteases. NMR backbone assignment of CVB3 $3C^{Pro}$ bound and unbound to the GC-376 PROTAC precursor (80% and 97%, respectively) was obtained. This information complemented the investigation, by NMR, of the interaction of CVB3 $3C^{Pro}$ with the GC-376 PROTAC, and its precursor allows us to define that the GC-376 PROTAC binds to CVB3 $3C^{Pro}$ in a mode very similar to that of the precursor. The NMR relaxation data indicate that a quench of dynamics of a large part of the protein backbone involving the substrate-binding site and surrounding regions occurs upon GC-376 PROTAC precursor binding. This suggests that the substrate cavity, by sampling different backbone conformations in the absence of the substrate, is able to select the suitable one necessary to covalently bind the substrate, this being the latter reaction, which is the fundamental step required to functionally activate the enzymatic reaction. The inhibition activity assay showed inhibition potency in the micromolar range for GC-376 PROTAC and its precursor. Overall, we can conclude that the GC-376 PROTAC fits well within the binding sites of both proteases, demonstrating its potential as a broad-spectrum antiviral agent.

Keywords: PROTAC; viral main protease; 3-chymotrypsin-like protease; Coxsackievirus B3; SARS-CoV-2; GC-376

Citation: De Santis, A.; Grifagni, D.; Orsetti, A.; Lenci, E.; Rosato, A.; D'Onofrio, M.; Trabocchi, A.; Ciofi-Baffoni, S.; Cantini, F.; Calderone, V. A Structural Investigation of the Interaction between a GC-376-Based Peptidomimetic PROTAC and Its Precursor with the Viral Main Protease of Coxsackievirus B3. *Biomolecules* **2024**, *14*, 1260. https://doi.org/10.3390/biom14101260

Academic Editor: Ajoy Basak

Received: 16 July 2024
Revised: 16 September 2024
Accepted: 2 October 2024
Published: 6 October 2024

1. Introduction

Cells use the ubiquitin–proteasome system to degrade unwanted or misfolded cellular proteins [1,2]. To perform this, the system targets proteins for proteasome-mediated

destruction by polyubiquitinating them. This process involves the sequential action of the E1 ubiquitin-activating enzyme, the E2 ubiquitin-conjugating enzyme and the E3 ubiquitin–protein ligase [3–6]. The E3 ligase is responsible for attaching multiple ubiquitin molecules to lysine residues at recognition sites in the protein to be degraded. This post-translational modification serves as a signal for proteolytic action by the 26S proteasome [7]. This cellular degradation system has been exploited by PROteolysis TArgeting Chimera (PROTAC) technology to modulate the cell levels of target proteins. PROTAC has proven highly effective in targeting proteins associated with cancer, leading to the initiation of clinical investigations for over a dozen drug candidates [8–10]. The utilization of the PROTAC technology as antiviral agent, although more limited, has recently grown. In the last year, it has indeed received a strong impetus with four works, including one from our group, that delineated the use of different PROTAC molecules against the 3-chymotrypsin-like protease of SARS-CoV-2 (SARS-CoV-2 $3CL^{Pro}$, hereafter) [11–14]. In our previous study [12], we have synthesized a PROTAC bifunctional molecule. This molecule conjugates a unit derived from the GC-376 inhibitor, which is able to recognize the active site of SARS-CoV-2 $3CL^{Pro}$ [15,16], with the pomalidomide ligand, which is able to select Cereblon (CRBN) E3 ligase [17–21]. The GC-376 inhibitor and the pomalidomide ligand are joined through a piperazine–piperidine linker. We showed that this PROTAC molecule is able to interact with SARS-CoV-2 $3CL^{Pro}$, forming a reversible covalent bond between the Cβ carbon of the α,β-unsaturated amide moiety and the catalytic Cys145 sulfur atom of the protein. Moreover, we observed that our PROTAC molecule reduces protein levels of SARS-CoV-2 $3CL^{Pro}$ in cultured cells without affecting cell viability. This demonstrates that a peptidomimetic-based PROTAC approach can attack viral infections within those coronaviruses displaying a high degree of structural similarity in the active site of their $3CL^{Pro}$ with that of SARS-CoV-2 $3CL^{Pro}$ [22].

Proteases are also important for the replication of viruses other than coronaviruses [23]. Among them, we have the proteases encoded by the 3C region of the enteroviruses genome ($3C^{Pro}$, hereafter), such as Coxsackievirus B3 (CVB3), EV-D68, EV-A71 and Coxsackievirus A16 [24–26]. All these enteroviral pathogens are part of the picornavirus family [27]. Infection with some of these viruses can lead to serious outcomes and cause clinical disease much more frequently than coronaviruses. As an example, infection with Coxsackievirus B3, a nonenveloped single-stranded (+) RNA enterovirus, is the most common reason for viral myocarditis and sudden cardiac death, within the enterovirus genus, with a very malicious disease pattern [28,29]. No effective therapeutic strategy for the prevention and treatment of these diseases is nowadays available, and enterovirus disease has a medical impact, with newborn infants and young children being at risk for septic-like disease [30]. The fact that all these virally encoded proteases from coronaviruses and enteroviruses perform essential functions in the viral replication cycle by cleaving both viral and host targets [31,32], as well as the fact that there is no known human protease with a specificity for Gln at the cleavage site of the substrate, has pushed the scientific community towards the development of a commercially viable antiviral drug targeting both of these virus families [33]. This approach is made viable because coronaviral $3CL^{Pro}$ exhibits a high degree of structural similarity across $3C^{Pro}$ of the picornavirus family [23]. Crystal structures of $3CL^{Pro}$ from both alpha and beta coronaviruses [34,35] revealed that two of the three domains of these enzymes together resemble the chymotrypsin-like fold of enteroviral $3C^{Pro}$, with the exception of an additional α-helical domain that is involved in the dimerization of coronaviral proteases. While this dimerization is essential for the catalytic activity of $3CL^{Pro}$, the enteroviral $3C^{pro}$ functions as a monomer [34]. In addition, the enteroviral $3C^{Pro}$ maintains a classical Cys···His···Glu/Asp catalytic triad, whereas the coronaviral $3CL^{Pro}$ only has a Cys···His dyad [34]. Nevertheless, these two types of proteases share common features, in particular their almost-absolute requirement for Gln in the P1 position of the substrate and space for only small amino-acid residues such as Gly, Ala, or Ser in the P1′ position of the substrate, indicating that the coronaviral $3CL^{Pro}$ and

the enteroviral $3C^{Pro}$ can be a common target for the design of broad-spectrum antiviral compounds [34].

$3CL^{Pro}$ has several subsites ("S") for substrate binding, which are identified by the Schechter and Berger (1967) nomenclature [36]. These include S1 (Phe140, Leu141, Asn142, His163, Glu166 and His172), S1′ (Thr24 and Thr25), S2 (His41, Met49, Tyr54, Met165 and Asp187), S4 (Leu167, Phe185, Gln189 and Gln192) and S5 (Pro168, Thr190 and Ala191), and they are partially maintained in $3C^{Pro}$ [37] (Figure 1A). The latter has a very similar backbone structural arrangement in the S1 and S1′ subsites although it has different residues (Thr142, Arg143, Ala144, Gln146, His161, Gly164 and Gly169 in S1 and Tyr22 and Gly23 in S1′), but it has an open and shallow S2 site with only His40 as conserved residue [38,39] (Figure 1A). The S4 and S5 subsites in $3C^{Pro}$ are also structurally different and more open compared to $3CL^{Pro}$. This difference arises because residues Gln189, Thr190, Ala191 and Gln192, which are located on the loop connecting domains I and II to domain III in $3CL^{Pro}$, are absent in $3C^{Pro}$. Notably, $3C^{Pro}$ lacks domain III, which is responsible for enzyme dimerization in $3CL^{Pro}$ (Figure 1A). Despite the structural differences present among the various subsites, broad-spectrum inhibitors of both $3CL^{Pro}$ and $3C^{Pro}$ have been largely documented in the literature [40–49], encouraging us to explore the capability of the PROTAC molecule previously designed by us against SARS-CoV-2 $3CL^{Pro}$ to also target enteroviral $3C^{Pro}$. To this aim, we have employed X-ray crystallography and solution NMR to characterize the interaction of $3C^{Pro}$ from Coxsackievirus B3 (CVB3 $3C^{Pro}$, hereafter) with the peptidomimetic PROTAC molecule previously selected on SARS-CoV-2 $3CL^{Pro}$ (GC-376 PROTAC) as well as with its precursor (GC-376 PROTAC precursor), i.e., a dipeptidyl protease ligand (Figure 1B) [12]. We have also compared the inhibitory activity of the GC-376 PROTAC precursor against CVB3 $3C^{Pro}$ with that obtained in the case of the SARS-CoV-2 $3CL^{Pro}$ enzyme.

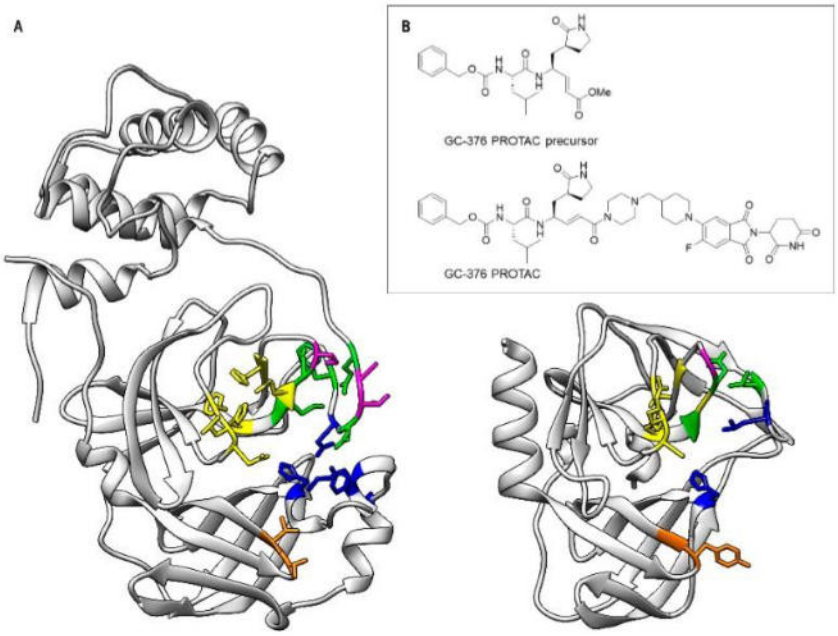


Figure 1. Comparison of the substrate-binding subsites (S1, S1′, S2, S4 and S5) of SARS-CoV-2 $3CL^{Pro}$ (left) and of CVB3 $3C^{Pro}$ (right) (**A**) and 2D structure of the GC-376 PROTAC and its precursor (**B**). In the two protein structures, the backbone and side chains of the residues of the subsites are color-coded (S1—yellow, S1′—orange, S2—blue, S4—green, S5—purple).

2. Materials and Methods

2.1. Expression and Purification of CVB3 3C^{Pro}

CVB3 3C^{Pro} was expressed and purified using a modification of Fili et al.'s protocol [50]. In detail, the pET-24a(+) plasmid encoding the CVB3 3C^{Pro} protein with an hexahistidine-coding sequence at the C-terminus was used to transform competent BL21(DE3)pLysS *E. coli* cells. A single colony from the transformation plate was picked to inoculate a 50 mL preculture in Luria–Bertani (LB) medium containing 50 μg mL^{-1} kanamycin and 35 μg mL^{-1} chloramphenicol. After overnight incubation at 37 °C, 20 mL of the preculture was used to inoculate 1 L LB medium containing 50 μg mL^{-1} kanamycin and 35 μg mL^{-1} chloramphenicol. When an optical absorption of 0.6 (at a wavelength of 600 nm) was reached, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to induce overexpression. The culture was then left overnight at 23 °C and then centrifuged at 5500 rpm for 20 min at 4 °C with a JA-10 Beckman rotor. The resulting pellet was resuspended in a 20 mL lysis buffer consisting of 50 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole, 5% glycerol and 0.1% Triton-X pH 7.7. The cell suspension was lysed by one cycle of freezing (−80 °C) and thawing (+10–15 °C) and sonication upon the addition of 20 mL lysis buffer, 2 mM DTT and lysozyme (0.15 mg mL^{-1}). The soluble cellular fraction was obtained by centrifugation at 4 °C at 40,000 rpm for 35 min with a 70Ti Beckman rotor. Then, the supernatant containing soluble CVB3 3C^{Pro} was loaded for Immobilized Metal Affinity Chromatography (IMAC) purification by using a HisTrap FF 5mL column, which was previously equilibrated with 50 mM Tris-HCl, 300 mM NaCl, 30 mM imidazole and 1 mM DTT pH 7.7 as a binding buffer. After lysate loading, the column was washed with 20 column volumes of the same buffer, and elution was executed with 10 column volumes of 50 mM Tris-HCl, 300 mM NaCl, 500 mM imidazole and 1 mM DTT pH 7.7 as a buffer solution. Fractions containing the target protein were pooled and concentrated by Amicon Ultra 10 kDa, reaching a volume of 3.5 mL. Subsequently, the protein solution was loaded onto a Hi Load Superdex 16/600 200 pg column for Size Exclusion Chromatography (SEC). Isocratic elution was performed with 1.2 column volumes of 10 mM HEPES, 300 mM NaCl, 1 mM EDTA and 1 mM DTT pH 7.5 as a buffer solution. The final protein yield resulted in 45 mg/L of culture. The purity of CVB3 3C^{Pro} was checked by SDS-PAGE analysis. The production of ^{15}N-labeled and ^{13}C,^{15}N-labeled CVB3 3C^{Pro} was performed by following the same protocol, just switching to a standard M9-medium containing $(^{15}NH_4)_2SO_4$ (1.2 g/L) and ^{13}C-glucose (4 g/L) as sources of nitrogen and carbon, respectively.

2.2. Analytical Gel Filtration

The protein size of CVB3 3C^{Pro} was analyzed using analytical gel filtration (Superdex 200 10/300 increase column; Cyativa, MA, USA) calibrated with gel filtration marker calibration kit, 6500–66,000 Da. Purified samples in 25 mM MES and 150 mM NaCl at pH 6.5 as a buffer was loaded on the column pre-equilibrated with 25 mM MES and 150 mM NaCl at pH 6.5 as a running buffer. Elution profiles were recorded at 280 nm with a flow rate of 0.75 mL/min. A standard curve of the logarithm of the molecular weight of the standards vs. $K_{av} = (V_e - V_0)/(V_t - V_0)$ (V_e, elution volume; V_0, dead volume; V_t, total volume) was used to calculate the apparent molecular mass of CVB3 3C^{Pro}.

2.3. Crystallization, Data Collection and Structure Solution

Crystals of CVB3 3C^{Pro} in complex with the GC-376 PROTAC precursor were obtained through co-crystallization in a sitting drop by adding a 2 μL aliquot of reservoir buffer (0.1 M Tris-HCl, 0.2 M $MgCl_2$ and 26% PEG4000, pH 8.5) to 2 μL of protein solution (20 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA and 1 mM DTT, pH 7.8) containing the ligand (2–3-fold concentration with respect to the protein and dissolved in 10% DMSO); trays were stored at 20 °C. The protein concentration in the sample was 5 mg/mL.

The dataset was collected in-house, using a BRUKER D8 Venture diffractometer equipped with a PHOTON III detector, at 100 K; the crystals used for data collection were cryo-cooled using 25% ethylene glycol in the mother liquor. The GC-376 PROTAC precursor–

protein adduct crystal was refined at 1.9 Å resolution: it belongs to space group C2 with one molecule in the asymmetric unit, a solvent content of about 50% and a mosaicity of 0.4–0.6°. The data were processed using the program XDS [51] and were reduced and scaled using XSCALE [51], and amplitudes were calculated using XSCALE [51]. The structure was solved with the molecular replacement technique using the 7QUW structure as the template model. The successful orientation and translation of the molecule within the crystallographic unit cell was determined with MOLREP [52]. The refinement and water molecule fitting was carried out using PHENIX [53], applying default TLS restraints. In between the refinement cycles, the model was subjected to manual rebuilding using COOT [54]. The quality of the refined structures was assessed using the program MOLPROBITY [55]. Data processing and refinement statistics for the GC-376 PROTAC precursor–protein adduct are shown in Table S1. The coordinates and structure factors have been deposited in the PDB under the accession code 8S6F. The software used for structural visualization is PyMOL (Schrodinger, LLC (2015), The PyMOL Molecular Graphics System, Version 1.8).

2.4. NMR Spectroscopy

Conventional multidimensional NMR techniques based on 3D triple-resonance experiments [56] were performed on an AVANCE 500 spectrometer equipped with a cryogenically cooled probe to obtain a backbone resonance assignment of CVB3 $3C^{Pro}$ with and without the GC-376 PROTAC precursor at 308 K. The used NMR sample condition was 50 mM phosphate buffer at pH 6.0 containing 100 mM NaCl, 50 mM arginine, 50 mM glutamate, 1 mM DTT and 1 mM EDTA and 10% (*v*/*v*) D_2O. All the NMR spectra were processed using the standard software Bruker: Topspin 4.2 and analyzed with the program CARA [57].

In order to monitor the interaction of CVB3 $3C^{Pro}$ with the GC-376 PROTAC and its precursor, we performed an NMR titration, based on the acquisition of ^{1}H-^{15}N HSQC experiments at 308 K on a Bruker AVANCE 950 MHz, by adding aliquots (up to a maximum of 10 µL of final added volume) of the small molecule (GC-376 PROTAC or its precursor) dissolved in D3-acetonitrile to the NMR tube containing ^{15}N-labeled CVB3 $3C^{Pro}$ (0.2–0.3 mM) up to a 1:1 ratio. All NMR titration data were analyzed comparing the ^{1}H-^{15}N HSQC spectra recorded along the additions of the small molecule with that of the initial state as well as with ^{1}H-^{15}N HSQC blank spectra of ^{15}N-labeled CVB3 $3C^{Pro}$ recorded from titrating the protein with the same aliquot of D3-acetonitrile, in order to identify the chemical shift changes induced by the organic solvent. In such a way, following the chemical shift changes observed in the ^{1}H-^{15}N HSQC maps along each stepwise titration, we were able to selectively assign the residues affected by protein–protein interactions. Signals showing chemical shift changes occurring on a slow exchange regime of the NMR time scale were considered when mapping the interaction surface on CVB3 $3C^{Pro}$ upon binding with the small molecule. The observed chemical shift changes were calculated as backbone-weighted average chemical shift differences, i.e., $\Delta\delta_{avg}$(HN), i.e., $(((\Delta H)^2 + (\Delta N/5)^2)/2)^{1/2}$, where ΔH and ΔN are chemical shift differences for backbone amide ^{1}H and ^{15}N nuclei, respectively. The estimation of the chemical shift threshold value for defining meaningful chemical shift differences was obtained by averaging $\Delta\delta_{avg}$(HN) values plus one standard deviation (1σ), following the standard procedure used in NMR protein–protein interaction studies [58].

The backbone dynamic properties of CVB3 $3C^{Pro}$ have been sampled through ^{15}N relaxation measurements. NMR experiments for measuring ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates [59] and $[^{1}H]^{15}N$ heteronuclear NOE values [60] were recorded at 298 K at 500 MHz, using a protein concentration of 300 µM. A temperature dependence of the backbone NH chemical shifts in the ^{1}H-^{15}N HSQC spectra allowed us to obtain the backbone resonance assignment of CVB3 $3C^{Pro}$ at 298 K. ^{15}N R_1, ^{15}N R_2 and steady-state $[^{1}H]^{15}N$ NOEs were obtained with previously described pulse sequences, which employ gradient selection and sensitivity enhancement, as well as minimal water suppression. ^{15}N R_2 were measured with a refocusing time (τ_{CPMG}) of 450 µs with the Carr–Purcell–Meiboom–Gill (CPMG) sequence. In all experiments, the water signal was

suppressed with the "water flipback" scheme. ^{15}N R_1 and ^{15}N R_2 relaxation rates were obtained by fitting the cross-peak volumes (I), measured as a function of the relaxation delay, to a single-exponential decay as described in the literature [61]. Heteronuclear $[^1H]^{15}N$ NOE values were calculated as the ratio of peak volumes in spectra recorded with and without saturation. The analysis of the uncertainties of ^{15}N R_1 and ^{15}N R_2 relaxation rates was carried out by comparing the peak intensity on duplicated spectra having the same relaxation delay. Estimates of the molecular tumbling value under the chosen experimental conditions of magnetic field and temperature were obtained using the program HydroNMR following the standard procedure [62].

2.5. Enzyme Inhibition Kinetics Assay

The CVB3 $3C^{Pro}$ enzyme inhibition activity of the synthesized GC-376 PROTAC and its precursor were analyzed through a fluorometric assay using the fluorogenic substrate Hilyte Fluor-488-ESATLQSGLRKAK-(QXL-520)-NH_2 (Anaspec). All the measurements were performed in 96-well plates with a Fluostar Optima microplate reader (BMG Labtech, Ortenberg, Germany) analogously to what has previously been performed [12]. Excitation and emission wavelengths were 490 and 520 nm, respectively. All incubations were performed at 30 °C in 10 mM HEPES buffer containing 150 mL NaCl, 1 mM EDTA and 1mM DTT at pH = 7.4. The GC-376 PROTAC or the GC-376 PROTAC precursor were preincubated with the CVB3 $3C^{Pro}$ enzyme (59 nM) for 10 min at 30 °C before the reaction was started by the addition of the fluorogenic substrate (1 μM). The increase in fluorescence signal was monitored over 30 min (λ_{ex} = 490 nm, λ_{em} = 520 nm) at 30 °C. The percentages of inhibition for the tested molecules were determined through the equation $(1 - V_s/V_o) \times 100$, where vs. is the initial velocity in the presence of the inhibitor and V_o is the initial velocity of the uninhibited reaction. The initial velocity was calculated, using the data analysis software MARS 2875A embedded in the Fluostar Optima instrument, as the slope of the linear regression of the curves in the initial range (0–5 min). The IC_{50} values were obtained by dose–response measurements using an inhibitor range of concentrations from 0.1 nM to 0.3 mM for the GC-376 PROTAC and 0.01 nM to 0.1 mM for the GC-376 PROTAC precursor. All the experiments were performed in triplicate, and the data collected were analyzed using GraphPad 5.0 Software Package (GraphPad Prism, Inc., San Diego, CA, USA). The percentages of inhibition at each inhibitor concentration were fitted using the non-linear fitting function 'log(inhibitor) vs. response' implemented in GraphPad.

3. Results

3.1. Crystal Structure of CVB3 $3C^{Pro}$ in Complex with GC-376 PROTAC Precursor

To characterize the interaction of CVB3 $3C^{Pro}$ with both the GC-376 PROTAC and its precursor, the same compounds previously used by us with SARS-CoV-2 $3CL^{Pro}$ (Figure 1B), we performed crystallization trials through both ligand-co-crystallization and soaking strategies. Well-diffracting crystals were obtained only in the case of the GC-376 PROTAC precursor. This result differs from what was obtained with SARS-CoV-2 $3CL^{Pro}$, whose well-diffracting crystals were obtained with both the GC-376 PROTAC precursor and the GC-376 PROTAC by soaking and co-crystallization, respectively. The crystal structure of CVB3 $3C^{Pro}$ in complex with the GC-376 PROTAC precursor was solved at 1.9 Å resolution, which provided atomic details of the CVB3 $3C^{Pro}$–ligand interactions. The overall 3D structure clearly resembles that of the free protein (PDB IDs 2VB0 and 3ZYD) with RMSD computed on C-alpha atoms of 0.55 and 0.33 Å, respectively. The structure of CVB3-$3C^{pro}$ adopts a chymotrypsin protein fold. The N-terminus starts with an α-helix of residues 1–14 and is followed by two topologically equivalent β-barrels comprising residues 15–77 and 100–173, which pack together to form a narrow groove for substrate binding. The catalytic triad of Cys147, His40 and Glu71 is located in the cleft between the two β-barrels.

Concerning the presence of the protease ligand, we observed that the electron density is well defined for all atoms (Figure 2A). The ligand has been refined at full occupancy and the B-factor values of the atoms that are visible in the electron density are in line with

those of the neighboring protein atoms. The most relevant feature is the binding mode of the ligand to the catalytic Cys147, which occurs via a covalent bond between C5 and the sulfur atom of Cys147. From a crystallographic point of view, this is supported by the fact that the bond length spontaneously refines to values around 1.8 Å, which is coherent with a C-S covalent bond. From a chemical point of view, it is known that the formation of a covalent (even reversible) bond between the unsaturated carbon and the sulfur atom is very likely [63,64]. Besides the covalent bond, other polar and non-polar interactions keep the protease ligand in place (Figure 2B). The protein atoms involved in direct hydrogen bond interactions with the ligand are, namely, the backbone oxygen of Thr142, one of the imidazole nitrogen atoms of His161, the backbone oxygen of Val162, the backbone nitrogen atom of Gly164 and the backbone nitrogen atom of Cys147. Several more residues are involved in hydrophobic interactions, and this is in agreement with the relatively non-polar nature of the ligand.

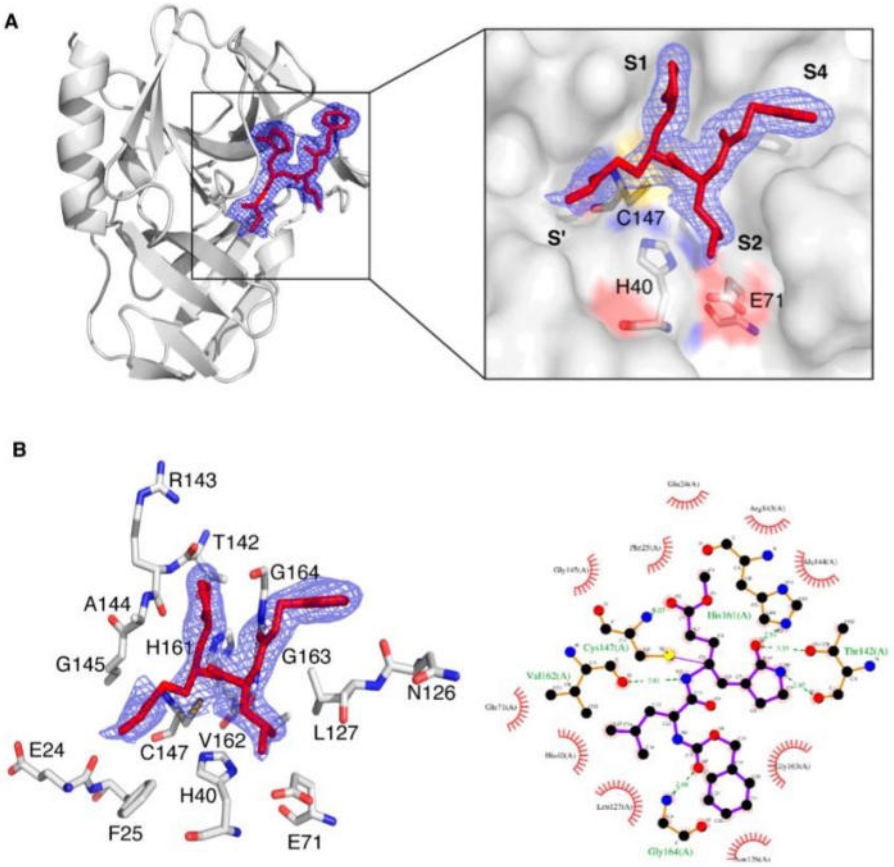


Figure 2. *Cont.*

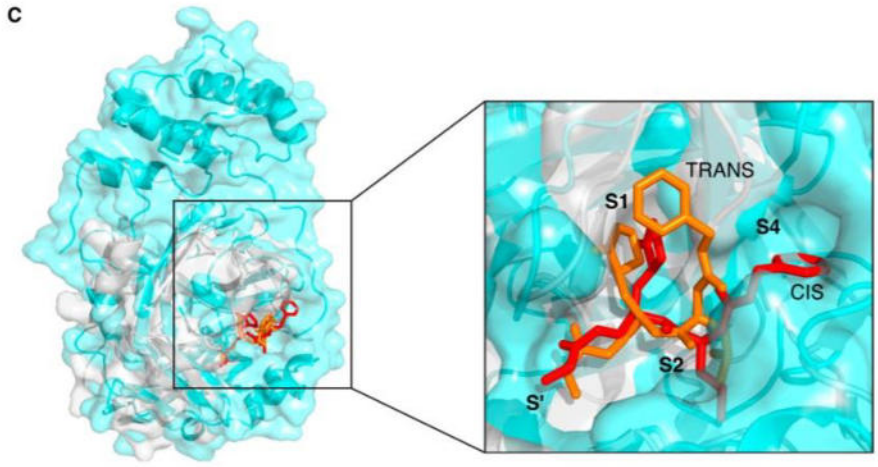


Figure 2. Crystal structure of CVB3 $3C^{Pro}$ in complex with the GC-376 PROTAC precursor. (**A**) Left: Ribbon diagram of CVB3 $3C^{Pro}$ bound to the GC-376 PROTAC precursor, with the 2Fo-Fc electron density map (blue mesh) contoured around the ligand at 1.0σ. Right: Close-up view of the binding site, highlighting the interaction of the GC-376 PROTAC precursor with key catalytic residues (Cys147, His40 and Glu71) in the active site. The subsites are also indicated. (**B**) Left: Representation of the binding interactions between the GC-376 PROTAC precursor and the surrounding residues of CVB3 $3C^{Pro}$, with the 2Fo-Fc electron density map as above. Right: Schematic 2D Ligplot representation showing the detailed binding interactions and hydrogen bonds between the ligand and active site residues, with hydrophobic interactions depicted as red arcs. (**C**) Left: Structural superimposition of CVB3 $3C^{Pro}$ (light gray, PDB ID 8S6F) and SARS-CoV-2 $3CL^{Pro}$ (cyan, PDB ID 8OKB) in complex with the GC-376 PROTAC precursor. The latter is shown as red sticks and orange sticks for CVB3 $3C^{Pro}$ and SARS-CoV-2 $3CL^{Pro}$, respectively. Right: Close-up of the binding pocket, showing the analogous orientation of the precursor in both active sites. The different conformations of the benzyloxycarbonyl group are also visible: "cis" for CVB3 $3C^{Pro}$ and "trans" for SARS-CoV-2 $3CL^{Pro}$.

Despite the structural difference in the subsites between SARS-CoV-2 $3CL^{Pro}$ and CVB3 $3C^{Pro}$ (shown in Figure 1A), the cavity of CVB3 $3C^{Pro}$ that hosts the ligand has a similar shape with respect to that of SARS-CoV-2 $3CL^{Pro}$ (PDB ID 8OKC) (Figure 2C); for this reason, the binding mode and interactions of the GC-376 PROTAC precursor in CVB3 $3C^{Pro}$ are similar to those in the 8OKB and 8OKC SARS-CoV-2 $3CL^{Pro}$ structures [12]. The only exception is the conformation of the phenyl ring of the benzyloxycarbonyl group, which is basically flipped by 90°. This effect is due to a different structural arrangement of the C-terminal segment of CVB3 $3C^{Pro}$ with respect to that of SARS-CoV-2 $3CL^{Pro}$, which makes the substrate cavity of CVB3 $3C^{Pro}$ more open. The larger available volume allows the phenyl ring to rotate in a position that is structurally more congested in the structure of SARS-CoV-2 $3CL^{Pro}$ (Figure 2C). This evidence also suggests that the ligand is versatile and is able to efficiently interact with different protein targets, provided, of course, that there is a certain degree of structural similarity among them. The alpha-helical region capping from the top the substrate cavity in SARS-CoV-2 $3CL^{Pro}$, which is absent in the CVB3 $3C^{Pro}$, does not determine a different orientation of the nearby Leu side chain of the ligand in the two structures. The ligand-binding site of CVB3 $3C^{Pro}$ has also been superposed to that of CVB3 $3C^{Pro}$ bound to a ligand similarly to the GC-376 PROTAC precursor (PDB ID 5NFS); in this case, the overall positions of the ligands are superposable (Figure S1).

All these results support that an antiviral PROTAC molecule acting on both coronaviral and enteroviral proteases can be constructed starting from our dipeptidyl protease ligand.

3.2. Structural Characterization of CVB3 $3C^{Pro}$ by Solution NMR

Because of the failure to obtain crystals of the GC-376 PROTAC-CVB3 $3C^{Pro}$ complex, we decided to apply a solution NMR strategy to characterize the interaction between the GC-376 PROTAC and CVB3 $3C^{Pro}$. Since the backbone resonance assignment of CVB3 $3C^{Pro}$ was not available, the first step was to find the experimental sample conditions suitable for obtaining high-quality 3D triple-resonance experiments required to extract the sequential information. The protein was produced in a ^{15}N- and ^{13}C-labeled medium and purified to a purity > 95% with two purification steps (Figure S2). Analytical gel filtration showed that the protein is homogenous with just one species present in solution corresponding to the monomeric form (Figure S2). The ^{1}H-^{15}N HSQC spectrum of CVB3 $3C^{Pro}$ at 298 K in 50 mM phosphate buffer with 100 mM NaCl at pH 6.0 shows well-dispersed amide signals with few peaks clustered in the random-coil region (Figure S3). Nevertheless, the high quality of the ^{1}H-^{15}N HSQC spectrum is not reflected in 3D triple-resonance NMR experiments needed to achieve a complete backbone resonance assignment (Figure S3). Indeed, several NHs in the 3D triple-resonance NMR experiments lack the sequential pattern of signals required to unambiguously perform backbone resonance assignment. Moreover, in these conditions, the protein has a strong tendency to precipitate over time in the NMR tube. Thus, to find suitable conditions to accomplish complete backbone resonance assignment, we performed a temperature dependence NMR analysis at different pHs and ionic strengths. As a result, we were able to improve the quality of the 3D triple-resonance experiments by using a 50 mM phosphate buffer at pH 6.0 containing 100 mM NaCl, 50 mM arginine, 50 mM glutamate, 1 mM DTT and 1 mM EDTA [65] and by acquiring the NMR spectra at 308 K. In these experimental conditions, the backbone amide NMR signals significantly narrowed, and the number of Cα and Cβ resonances observed in the 3D spectra increased significantly. We were thus able to assign about 80% of the backbone NHs of CVB3 $3C^{Pro}$. In particular, we assigned all the cross-peaks detected in the ^{1}H-^{15}N HSQC spectrum, indicating that the unassigned backbone NHs are missing since they are very weak or broadened beyond detection (Table S2). Mapping the latter NHs on the structure of CVB3 $3C^{Pro}$, we observed that they are located in a large area involving the Cys···His···Glu residues of the catalytic triad, the substrate-binding subsites, part of the β-strand in contact with subsite S1 and the proximal N-terminal helix (Figure 3). This finding suggests that conformational motions on the μs-ms time scale occur in this region. To further investigate this aspect, we have measured ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates and $[^{1}H]^{15}N$ heteronuclear NOE values of CVB3 $3C^{Pro}$ (Figure S4). These data were used to obtain insights both into the motions of the backbone NHs at different time scales and to calculate, from the R_2/R_1 ratio, the overall correlation time for molecular tumbling (τ_m). In agreement with the results of the analytical gel filtration, the experimentally calculated τ_m value is 13.8 ± 1.9 ns, as expected for a protein of this size in a monomeric state and perfectly matching the τ_m value estimated by HYDRONMR [62] (13.7 ns) based on the monomeric crystal structure of CVB3 $3C^{Pro}$ (PDB ID 3ZYD). The residues showing negative $[^{1}H]^{15}N$ NOE values (Lys42, Ser107, Phe109 and Thr132), indicating backbone motions on the ps-ns time scale, as well as those showing high R_2 over R_1 ratios (His40, Asn80, Asn105, Glu121, Leu127, Thr132 and Arg134), indicating backbone motions on the μs-ms time scale, were mapped on the structure of CVB3 $3C^{Pro}$ (Figure 3). These NHs are again located in the area involving the catalytic triad, the substrate-binding subsites and some residues of the loop close to subsite S1. We can thus conclude from both the broadening-beyond-detection effects as well as from high internal flexibility in the ps time scale, that the backbone of a large protein region, (i.e., the catalytic triad, the substrate-binding subsites, the β-strand in contact with subsite S1 and the proximal N-terminal helix), is highly dynamic, thus preventing a complete backbone resonance assignment. However, these findings suggested to us that the addition of a ligand could dampen the highly dynamic properties of the ligand-binding region. To test this hypothesis, we have added one equivalent of the GC-376 PROTAC precursor to CVB3 $3C^{Pro}$ and recorded a ^{1}H-^{15}N HSQC spectrum (Figure S3) and a full set of the 3D triple-resonance experiments. The results show that the quality of

the 3D triple-resonance experiments is significantly improved (Figure S3), in such a way allowing us to obtain a complete and accurate backbone resonance assignment at 308 K for the GC-376 PROTAC precursor-bound state of CVB3 $3C^{Pro}$ (97% of assigned backbone NHs), with only five missing NHs (Table S2). The backbone resonance assignment at 308 K has been deposited in the BioMagResBank (BMRB ID 52547).

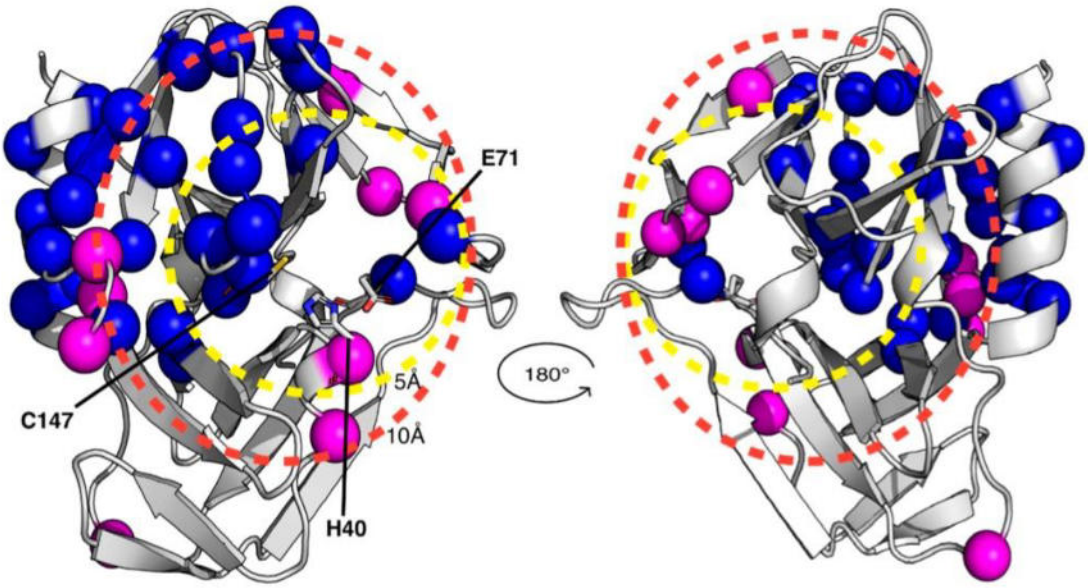


Figure 3. Structural and dynamic characterization of CVB3 $3C^{Pro}$ by solution NMR at 308 K. Backbone amide protons that are broadened beyond detection and those exhibiting elevated R_2/R_1 relaxation rates or negative ^1H-^{15}N NOE values are highlighted as blue and magenta spheres, respectively, on the crystal structure of CVB3 $3C^{Pro}$. Dashed circles indicate spatial distances of 5 Å and 10 Å from the active site, demarcating regions surrounding the binding pocket. The catalytic triad is also indicated as sticks.

The observed quench of dynamics of a large part of the protein backbone indicates that the cavity hosting the substrate samples different backbone conformations in the absence of the substrate. This structural feature allows a selection of the conformation necessary to covalently bind the substrate, which is the chemical step fundamentally required to functionally activate the enzymatic reaction.

3.3. Mapping the Interaction of the GC-376 PROTAC and Its Precursor with CVB3 $3C^{Pro}$ by Solution NMR

Thanks to having obtained the backbone resonance assignment of CVB3 $3C^{Pro}$ with and without the GC-376 PROTAC precursor, we can identify the residues of CVB3 $3C^{Pro}$ whose NH resonances are affected by the binding of the GC-376 PROTAC precursor as well as of the GC-376 PROTAC. An NMR titration was first performed by adding the GC-376 PROTAC precursor stepwise to ^{15}N-labeled CVB3 $3C^{Pro}$. Several NH signals are affected by the GC-376 PROTAC precursor additions, showing a slow exchange regime on the NMR time scale (Figures 4A and S5A, Table S3). Moreover, new NH signals appear with increasing intensities along the stepwise additions of the GC-376 PROTAC precursor with their chemical shifts corresponding to those of the GC-376 PROTAC precursor-bound species (Figure S5A, Table S4). Taken together, these findings indicate that the GC-376 PROTAC precursor is tightly bound to CVB3 $3C^{Pro}$. Mapping both these types of NH changes on the crystal structure of CVB3 $3C^{Pro}$ in complex with the GC-376 PROTAC

precursor (PDBID 8S6F) (Figure 5A), we can observe that the large majority of the backbone NHs surround the ligand but also include the β-strand in contact with subsite S1 and the proximal N-terminal helix. These results indicate the specific binding of the GC-376 PROTAC precursor to the cavity containing the Cys···His···Glu/Asp catalytic triad. We then used the same approach to investigate the binding of the GC-376 PROTAC: an NMR titration was thus performed by adding the GC-376 PROTAC stepwise to ^{15}N-labeled CVB3 3C^{Pro}. The observed chemical shift changes similarly reproduce those occurring upon the GC-376 PROTAC precursor additions, with signals displaying a slow exchange regime on the NMR time scale and new signals appearing with increasing intensities along the titration (Figures 4B and S5B, Tables S3 and S4). The residues whose chemical shifts change upon the addition of PROTAC were mapped on the crystal structure of 3C^{Pro} bound the PROTAC precursor (PDBID 8S6F) (Figure 5B). Upon interaction with 3C^{Pro}, PROTAC affects the same regions of its precursor. This is also confirmed by comparing the effects on the backbone chemical shifts of the binding of the GC-376 PROTAC to ^{15}N-labeled CVB3 3C^{Pro} with respect to those generated by the GC-376 PROTAC precursor (Figure S6). Indeed, the obtained $\Delta\Delta\delta_{avg}$ values are all below a threshold of ±0.05 ppm with the exception of four residues (Leu37, Asn69, Leu127 and Gly129) [58]. However, the latter four discrepancies do not indicate a meaningful difference in binding between PROTAC and its precursor. This is because the chemical shift differences in Leu37 and Asn69 are not meaningful in both NMR titrations with the GC-376 PROTAC and its precursor (i.e., below the threshold of 0.14 ppm; see Figure S5). Furthermore, the chemical shift differences in Leu127 and Gly129, although above the threshold in both titrations, show a difference as little as 0.054–0.062 ppm (Figure S5). Only two small differences over 175 backbone NHs indicate that the two molecules interact with CVB3 3C^{Pro} in a very similar mode. Thus, we can conclude that the linker and the pomalidomide ligand moiety are not tightly interacting with the protein surface, thus displaying a high degree of mobility. These data are coherent with our previous findings concerning the binding of the same GC-376 PROTAC molecule to SARS-CoV-2 3CL^{Pro} [12].

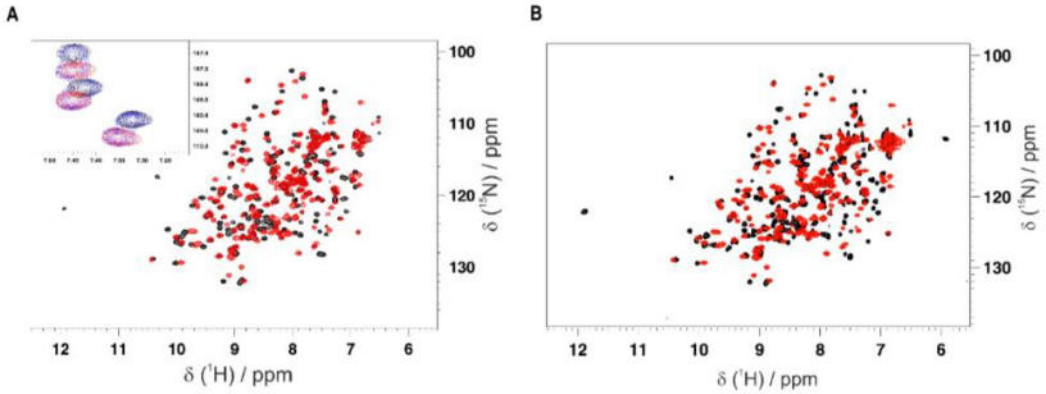


Figure 4. Monitoring the interaction of the GC-376 PROTAC and its precursor with CVB3 3C^{Pro} by solution NMR. Overlay of the ^{1}H-^{15}N HSQC spectra of CVB3 3C^{Pro} before (red) and after (black) the addition of 1 equivalent of the GC-376 PROTAC precursor (**A**) and GC-376 PROTAC (**B**) recorded at 308K. In the inset of panel A, three NH signals in a slow exchange regime of the NMR time scale are shown at 0 (red), 0.5 (blue) and 1 (black) equivalent of the GC-376 PROTAC precursor additions.

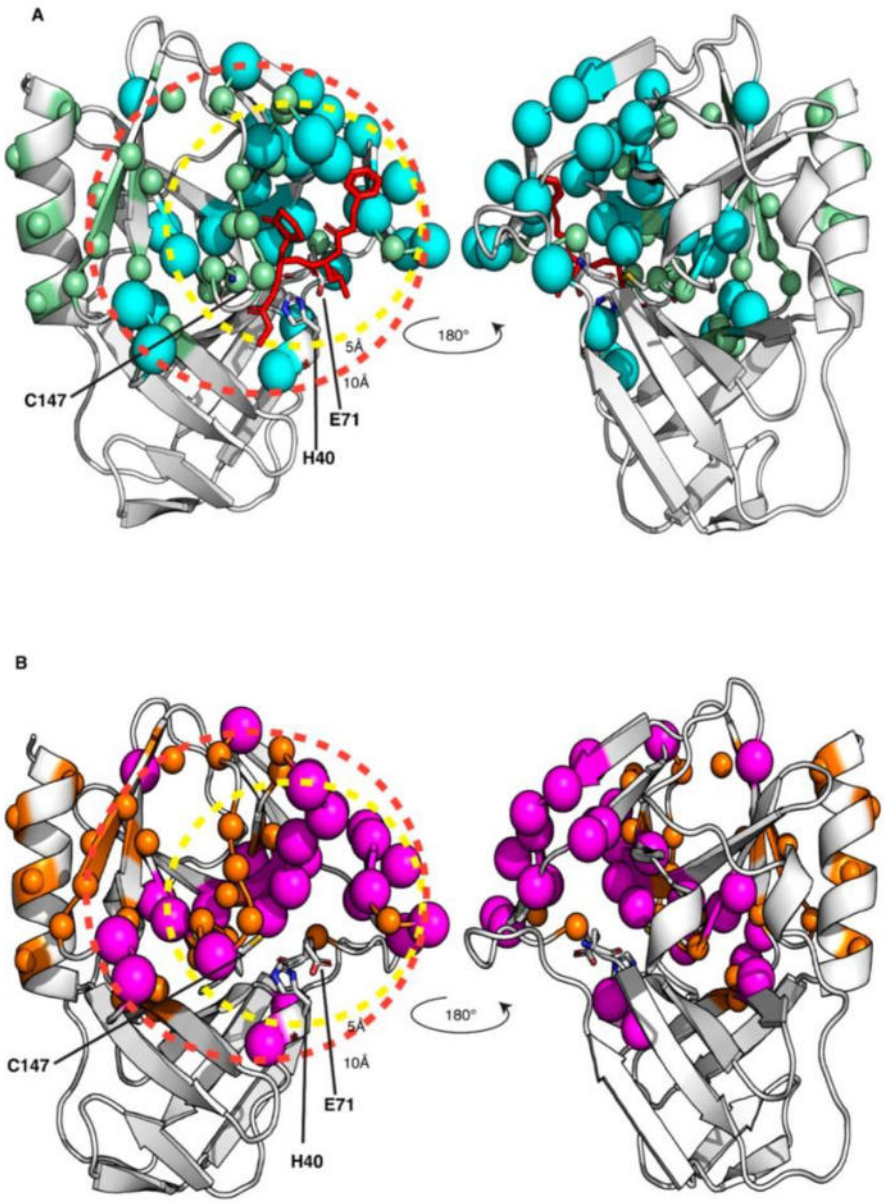


Figure 5. Mapping the interaction of the GC-376 PROTAC and its precursor with CVB3 $3C^{Pro}$ onto the crystal structure of the protein bound to the GC-376 PROTAC precursor (PDB ID 8S6F). The backbone NHs showing chemical shifts changes larger than the threshold value, as observed in Figure S5A,B, are mapped as cyan and magenta spheres for the GC-376 PROTAC precursor (**A**) and GC-376 PROTAC (**B**) titrations, respectively. The backbone NHs, unassigned in the CVB3 $3C^{Pro}$ and displaying increasing intensities along the stepwise additions of the GC-376 PROTAC precursor (light-green spheres) and of the GC-376 PROTAC (orange spheres) are also shown in panels (**A**) and (**B**) respectively. The dashed circles indicate spatial distances of 5 Å and 10 Å from the active site, demarcating regions surrounding the binding pocket. The catalytic triad is shown as sticks.

3.4. Enzyme Inhibition Kinetics of GC-376-PROTAC and Its Precursor

We evaluated GC-376-PROTAC and its precursor for their ability to inhibit CVB3 $3C^{Pro}$ through a fluorimetric enzyme inhibition kinetics assay using Hilyte Fluor-488-ESATLQSGLRKAK-(QXL-520)-NH_2 as the substrate. CVB3 $3C^{Pro}$ and EV71 3C (PDB ID 5BPE) are two viral proteases with a very high sequence identity (98.4%) and whose structures are fully superimposable. EV71 3C as well as SARS-CoV-2 $3CL^{Pro}$ are known to cleave several substrates in the dipeptide sequence consisting of Gln followed by a small amino acid (Ser or Gly), like in the substrate TSAVLQSGFRKM, used by several authors in the literature [48,49,66], and in the substrate used in this work containing the same LQSG epitope. These findings support that the peptide used here is also a suitable substrate for CVB3 $3C^{Pro}$. Both ligands showed an inhibition activity in the micromolar range. Specifically, the GC-376 PROTAC showed 32 ± 3% inhibition at 100 μM, and the GC-376 PROTAC precursor showed an IC_{50} of 4.6 μM (pIC_{50} = 5.28 ± 0.07) (Figure 6). Such values are in line with that displayed by SARS-CoV-2 $3CL^{Pro}$ (1.35 μM) [12], still supporting similar molecular recognition in the active site of both enzymes. The 4.6 μM value of the GC-376 PROTAC precursor is also comparable to those of compounds structurally similar to the GC-376 PROTAC precursor and inhibiting CVB3 $3C^{Pro}$ [48].

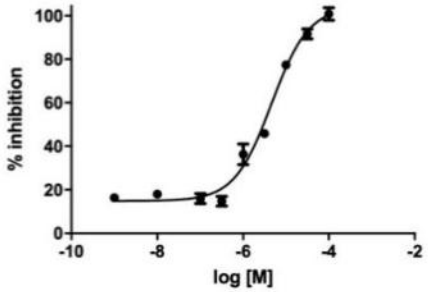


Figure 6. Inhibition curve of the GC-376 PROTAC precursor. M in the log refers to molar concentration.

4. Discussion

The conservation of the main protease in viral genomes and the fact that humans do not have a homologous protease make this enzyme family an ideal target for developing antiviral drugs with minimized toxicity against the host cell. These features have driven the development of several classes of viral inhibitors able to block the function of this target. Among them, GC-376 is one of the most efficient 3CL protease inhibitors found to inhibit the activity of coronavirus $3CL^{Pro}$ [67]. A recent study showed that GC-376 is a potent 3CL protease inhibitor that binds with different kinds of coronavirus $3CL^{Pro}$ and a variety of $3CL^{Pro}$ mutants of SARS-CoV-2 to exert inhibitory activity [68]. In particular, GC-376 was shown to have a strong inhibitory activity against three pathogenic coronaviruses, i.e., SARS-CoV, MERS-CoV and SARS-CoV-2 $3CL^{Pro}$. Recently, we have shown that a GC-376-based peptidomimetic PROTAC specifically targets and degrades the dimeric SARS-CoV-2 $3CL^{Pro}$ protein [12]. Several other PROTACs have been recently described to perform this function against both α- and β-groups of coronaviruses [11,13,14,69]. Here, we have extended the application of our GC-376-based peptidomimetic PROTAC to $3C^{Pro}$ protease, which is present in some members of the large genus enterovirus. Specifically, we structurally characterized the interaction of the GC-376 PROTAC with CVB3 $3C^{Pro}$, a member of the enteroviral proteases, and we found that the GC-376 PROTAC interacts similarly to what found with SARS-CoV-2 $3CL^{Pro}$. The GC-376 molecule penetrates in the substrate-binding site of CVB3 $3C^{Pro}$, while the linker and the pomalidomide molecule are not tightly bound to CVB3 $3C^{Pro}$ but remain exposed to the solvent. The binding of GC-376 to CVB3 $3C^{Pro}$ is similar to its binding in SARS-CoV-2 $3CL^{Pro}$, but there are some notable differences. The main difference lies in the orientation of the benzyloxycarbonyl

group of the GC-376 molecule, which is significantly altered between the two structures. This difference is also observed when the GC-376 inhibitor binds to $3CL^{Pro}$ of SARS-CoV and MERS-CoV [68]. In these crystal structures, the benzyloxycarbonyl group is indeed not immobilized, but it displays two conformations, "cis" and "trans", for each independent molecule and not a simultaneous presence of cis and trans in the same molecule. Namely, "cis" indicates that the benzyloxycarbonyl group shifted to Thr190, whereas "trans" indicates the benzyloxycarbonyl group shifted to the 2-pyrrolidone ring. Here, we have adopted the same criterion to identify the two possible conformers. Based on a more visual point of view, it is possible to say that in the "cis"-conformation, the phenyl ring lies on the opposite side with respect to the carbonyl oxygen and, conversely, that in the "trans"-conformation, the phenyl group lies on the same side with respect to the carbonyl oxygen. In Figure 7A,C, the trans-conformation of the benzyloxycarbonyl group of the GC-376 PROTAC precursor bound to $3CL^{Pro}$ of SARS-CoV-2 can be observed (PDB-ID 8OKB). The same trans-conformation of the benzyloxycarbonyl group is present in MERS-CoV (PDB-ID 8IG6). While a cis-conformation occurs for the benzyloxycarbonyl group of the GC-376 precursor bound to CVB3 $3C^{Pro}$ (Figure 7B,D), as occurs in SARS-CoV (PDB-ID 8IG5). This cis- or trans-conformation can be therefore present in $3C^{Pro}$/$3CL^{Pro}$ proteases, indicating that the S4 subsite of the substrate cavity is an open space where the benzyloxycarbonyl group of GC-376 can extend to the solvent, and it is relatively flexible due to limited stabilizing forces. Other minor structural differences include the ligand interactions with subsites S1 and S2. The isobutyl group, which inserts into the S2 subsite of both proteases, shows hydrophobic interactions with non-polar amino-acid residues (His41, Met49, His164 and Met165) in the $3CL^{Pro}$ structure of SARS-CoV-2, stabilizing its binding to the S2 site (Figure 7A). This hydrophobic patch is not fully maintained in the CVB3 $3C^{Pro}$ structure. Indeed, the catalytic Glu71, which is absent in coronaviral $3CL^{Pro}$s, is located with its negative charge next to the isobutyl group (Figure 7B). Finally, the γ-lactam ring of GC-376, which inserts into the S1 subsite of both proteases, forms similar interactions with the catalytic Cys and His161/163, residues conserved in both $3C^{Pro}$ and $3CL^{Pro}$ proteases. However, it also has other different surrounding residues in the two proteases. Specifically, there are several residues located in the same structural positions of the S1 subsite of both proteases that are exchanged: Phe140 → Thr142, Asn142 → Ala144, Ser144 → Gln146 and Glu166 → Gly163,164,166 (Figure 7C,D). Overall, we can conclude that, despite considerable sequence diversity in the substrate-binding sites of $3C^{Pro}$ and $3CL^{Pro}$, the GC-376 PROTAC can accommodate well in the binding sites of both proteases, suggesting this system as a valuable starting molecular platform for developing a broad-spectrum antiviral PROTAC. Further data concerning the degradation ability of our PROTAC against CVB3 $3C^{Pro}$ in viral cells will be crucial for verifying the applicability of our molecule.

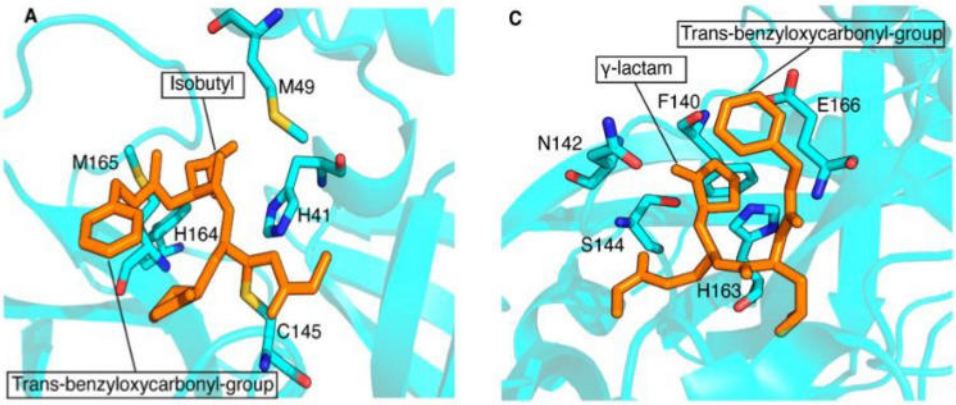


Figure 7. *Cont.*

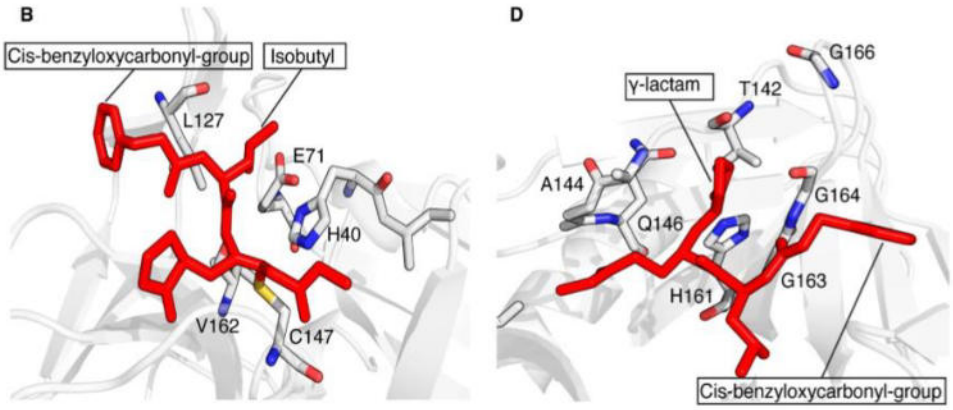


Figure 7. Structural comparison of the ligand–protein interactions in CVB3 $3C^{Pro}$ (PDB ID 8S6F) and SARS-CoV-2 $3CL^{Pro}$ (PDB ID 8OKB). The side chains of the residues surrounding the isobutyl group (**A**) and γ-lactam ring (**C**) of the GC-376 PROTAC precursor (in orange) are shown as sticks on the SARS-CoV-2 $3CL^{Pro}$ structure. The trans-conformation of benzyloxycarbonyl group is indicated. The side chains of the residues surrounding the isobutyl group (**B**) and the γ-lactam ring (**D**) of the GC-376 PROTAC precursor (in red) are shown as sticks on the CVB3 $3C^{Pro}$ structure. The cis-conformation of the benzyloxycarbonyl group is indicated. The key residues are shown as sticks.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/biom14101260/s1. Table S1, data collection and refinement statistics of the crystal structure of CVB3 $3C^{Pro}$ in complex with the GC-376 PROTAC precursor; Table S2, list of the unassigned backbone NHs of CVB3 $3C^{Pro}$ and CVB3 $3C^{Pro}$ in complex with the GC-376 PROTAC precursor; Table S3, list of the residues whose NH chemical shifts change above the threshold upon addition of the GC-376 PROTAC and of its precursor; Table S4, list of the residues unassigned in CVB3 $3C^{Pro}$ and displaying increasing intensities along the stepwise additions of the GC-376 PROTAC and of its precursor; Figure S1, superposition of CVB3 $3C^{Pro}$ bound to the GC-376 PROTAC precursor and a similar ligand; Figure S2, expression/purification steps and protein size of CVB3 $3C^{Pro}$; Figure S3, solution NMR spectra acquired at 500 MHz of CVB3 $3C^{Pro}$ and of CVB3 $3C^{Pro}$ bound to the GC-376 PROTAC precursor ; Figure S4, ^{15}N relaxation data of CVB3 $3C^{Pro}$; Figure S5, monitoring the interaction of CVB3 $3C^{Pro}$ with the GC-376 PROTAC precursor and GC-376 PROTAC by solution NMR ; Figure S6, comparing the interaction of CVB3 $3C^{Pro}$ with the GC-376 PROTAC with and its precursor CVB3 3CPro by solution NMR; Figure S7, representative kinetic curves of a control experiment in comparison with an inhibition experiment; Figures S8–S11, raw data of experimental and calculated inhibition curves of the GC-376 PROTAC and its precursor.

Author Contributions: Conceptualization, S.C.-B., V.C. and F.C.; Formal analysis, V.C., A.T. and D.G.; Funding acquisition, A.R., M.D. and S.C.-B.; Investigation, A.D.S., A.O. and E.L.; Project administration, A.R. and F.C.; Resources, D.G.; Supervision, F.C. and V.C.; Visualization, A.D.S. and D.G.; Writing—original draft, S.C.-B., F.C. and V.C.; Writing—review and editing, all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Intesa Sanpaolo Charity Fund (project number B/2021/0212); by "the European Union-NextGenerationEU-National Recovery and Resilience Plan", Mission 4 Component 2-Investment 1.5-THE-Tuscany Health Ecosystem-ECS00000017-CUP B83C22003920001; and by the National Recovery and Resilience Plan, Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2 February 2022 by the Italian Ministry of University and Research (MUR). This work was also supported by the European Union-Next Generation EU, Mission 4, Component 1, CUP B53D23015510006—Grant Assignment Decree No. 1064 adopted on 18 July 2023 by the Italian Ministry of Ministry of University and Research (MUR) and by PAN-HUB 2021-T4-AN-07-CUP B93C22000920001, "Hub multidisciplinare e interregionale di ricerca e sperimentazione clinica per il contrasto alle pandemie ed all'antibiotico resistenza".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Coordinates and structure factors have been deposited at the PDB under the accession code 8S6F for CVB3 $3C^{Pro}$ complexed with the GC-376 PROTAC precursor. The backbone resonance assignment of the GC-376 PROTAC precursor-bound state of CVB3 $3C^{Pro}$ at 308 K has been deposited in the BioMagResBank (BMRB ID 52547).

Acknowledgments: This publication was supported by the European Virus Archive Global (EVA-GLOBAL) project that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 871029. S.C.B. acknowledges support from COST Action, FeSImmChemNet. This article is based upon work from COST Action FeSImmChemNet, CA21115, supported by COST (European Cooperation in Science and Technology). F.C and V.C. acknowledge the support of the Italian Ministry of University and Research (MUR) through Dipartimenti di Eccellenza 2023–2027 (DICUS 2.0) to the Department of Chemistry "Ugo Schiff" of the University of Florence.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Komander, D.; Rape, M. The ubiquitin code. *Annu. Rev. Biochem.* **2012**, *81*, 203–229. [CrossRef] [PubMed]
2. Zheng, N.; Shabek, N. Ubiquitin Ligases: Structure, Function, and Regulation. *Annu. Rev. Biochem.* **2017**, *86*, 129–157. [CrossRef] [PubMed]
3. Morreale, F.E.; Walden, H. Types of Ubiquitin Ligases. *Cell* **2016**, *165*, 248–248.e241. [CrossRef] [PubMed]
4. Deshaies, R.J.; Joazeiro, C.A. RING domain E3 ubiquitin ligases. *Annu. Rev. Biochem.* **2009**, *78*, 399–434. [CrossRef] [PubMed]
5. Schulman, B.A.; Harper, J.W. Ubiquitin-like protein activation by E1 enzymes: The apex for downstream signalling pathways. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 319–331. [CrossRef] [PubMed]
6. Ye, Y.; Rape, M. Building ubiquitin chains: E2 enzymes at work. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 755–764. [CrossRef]
7. Finley, D. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev. Biochem.* **2009**, *78*, 477–513. [CrossRef]
8. Li, K.; Crews, C.M. PROTACs: Past, present and future. *Chem. Soc. Rev.* **2022**, *51*, 5214–5236. [CrossRef] [PubMed]
9. Ciulli, A.H.O. PROTAC Degraders Mechanism, Recent Advances, and Future Challenges. In *Protein Homeostasis in Drug Discovery: A Chemical Biology Perspective*; Jones, M.K.L.H., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2022; pp. 317–356.
10. Toure, M.; Crews, C.M. Small-Molecule PROTACS: New Approaches to Protein Degradation. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 1966–1973. [CrossRef] [PubMed]
11. Alugubelli, Y.R.; Xiao, J.; Khatua, K.; Kumar, S.; Sun, L.; Ma, Y.; Ma, X.R.; Vulupala, V.R.; Atla, S.; Blankenship, L.R.; et al. Discovery of First-in-Class PROTAC Degraders of SARS-CoV-2 Main Protease. *J. Med. Chem.* **2024**, *67*, 6495–6507. [CrossRef] [PubMed]
12. Grifagni, D.; Lenci, E.; De Santis, A.; Orsetti, A.; Barracchia, C.G.; Tedesco, F.; Bellini Puglielli, R.; Lucarelli, F.; Lauriola, A.; Assfalg, M.; et al. Development of a GC-376 Based Peptidomimetic PROTAC as a Degrader of 3-Chymotrypsin-like Protease of SARS-CoV-2. *ACS Med. Chem. Lett.* **2024**, *15*, 250–257. [CrossRef]
13. Desantis, J.; Bazzacco, A.; Eleuteri, M.; Tuci, S.; Bianconi, E.; Macchiarulo, A.; Mercorelli, B.; Loregian, A.; Goracci, L. Design, synthesis, and biological evaluation of first-in-class indomethacin-based PROTACs degrading SARS-CoV-2 main protease and with broad-spectrum antiviral activity. *Eur. J. Med. Chem.* **2024**, *268*, 116202. [CrossRef]
14. Sang, X.; Wang, J.; Zhou, J.; Xu, Y.; An, J.; Warshel, A.; Huang, Z. A Chemical Strategy for the Degradation of the Main Protease of SARS-CoV-2 in Cells. *J. Am. Chem. Soc.* **2023**, *145*, 27248–27253. [CrossRef] [PubMed]
15. Ma, C.; Sacco, M.D.; Hurst, B.; Townsend, J.A.; Hu, Y.; Szeto, T.; Zhang, X.; Tarbet, B.; Marty, M.T.; Chen, Y.; et al. Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res.* **2020**, *30*, 678–692. [CrossRef]
16. Kim, Y.; Lovell, S.; Tiew, K.C.; Mandadapu, S.R.; Alliston, K.R.; Battaile, K.P.; Groutas, W.C.; Chang, K.O. Broad-spectrum antivirals against 3C or 3C-like proteases of picornaviruses, noroviruses, and coronaviruses. *J. Virol.* **2012**, *86*, 11754–11762. [CrossRef]
17. Ito, T. Protein degraders—From thalidomide to new PROTACs. *J. Biochem.* **2024**, *175*, 507–519. [CrossRef]
18. Zhu, Y.X.; Braggio, E.; Shi, C.X.; Bruins, L.A.; Schmidt, J.E.; Van Wier, S.; Chang, X.B.; Bjorklund, C.C.; Fonseca, R.; Bergsagel, P.L.; et al. Cereblon expression is required for the antimyeloma activity of lenalidomide and pomalidomide. *Blood* **2011**, *118*, 4771–4779. [CrossRef] [PubMed]
19. Gandhi, A.K.; Kang, J.; Havens, C.G.; Conklin, T.; Ning, Y.; Wu, L.; Ito, T.; Ando, H.; Waldman, M.F.; Thakurta, A.; et al. Immunomodulatory agents lenalidomide and pomalidomide co-stimulate T cells by inducing degradation of T cell repressors Ikaros and Aiolos via modulation of the E3 ubiquitin ligase complex CRL4(CRBN.). *Br. J. Haematol.* **2014**, *164*, 811–821. [CrossRef] [PubMed]

20. Girardini, M.; Maniaci, C.; Hughes, S.J.; Testa, A.; Ciulli, A. Cereblon versus VHL: Hijacking E3 ligases against each other using PROTACs. *Bioorg. Med. Chem.* **2019**, *27*, 2466–2479. [CrossRef]
21. Diehl, C.J.; Ciulli, A. Discovery of small molecule ligands for the von Hippel-Lindau (VHL) E3 ligase and their use as inhibitors and PROTAC degraders. *Chem. Soc. Rev.* **2022**, *51*, 8216–8257. [CrossRef]
22. Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J.R.; Hilgenfeld, R. Coronavirus main proteinase (3CLpro) structure: Basis for design of anti-SARS drugs. *Science* **2003**, *300*, 1763–1767. [CrossRef] [PubMed]
23. Marjomäki, V.; Kalander, K.; Hellman, M.; Permi, P. Enteroviruses and coronaviruses: Similarities and therapeutic targets. *Expert Opin. Ther. Targets* **2021**, *25*, 479–489. [CrossRef] [PubMed]
24. Wen, W.; Qi, Z.; Wang, J. The Function and Mechanism of Enterovirus 71 (EV71) 3C Protease. *Curr. Microbiol.* **2020**, *77*, 1968–1975. [CrossRef] [PubMed]
25. Laitinen, O.H.; Svedin, E.; Kapell, S.; Nurminen, A.; Hytönen, V.P.; Flodström-Tullberg, M. Enteroviral proteases: Structure, host interactions and pathogenicity. *Rev. Med. Virol.* **2016**, *26*, 251–267. [CrossRef]
26. Tan, J.; George, S.; Kusov, Y.; Perbandt, M.; Anemüller, S.; Mesters, J.R.; Norder, H.; Coutard, B.; Lacroix, C.; Leyssen, P.; et al. 3C protease of enterovirus 68: Structure-based design of Michael acceptor inhibitors and their broad-spectrum antiviral effects against picornaviruses. *J. Virol.* **2013**, *87*, 4339–4351. [CrossRef]
27. Nikonov, O.S.; Chernykh, E.S.; Garber, M.B.; Nikonova, E.Y. Enteroviruses: Classification, Diseases They Cause, and Approaches to Development of Antiviral Drugs. *Biochem. Biokhimiia* **2017**, *82*, 1615–1631. [CrossRef] [PubMed]
28. Maier, R.; Krebs, P.; Ludewig, B. Immunopathological basis of virus-induced myocarditis. *Clin. Dev. Immunol.* **2004**, *11*, 1–5. [CrossRef]
29. Gauntt, C.; Huber, S. Coxsackievirus experimental heart diseases. *Front. Biosci.* **2003**, *8*, e23–e35. [CrossRef]
30. Lugo, D.; Krogstad, P. Enteroviruses in the early 21st century: New manifestations and challenges. *Curr. Opin. Pediatr.* **2016**, *28*, 107–113. [CrossRef]
31. Fan, W.; McDougal, M.B.; Schoggins, J.W. Enterovirus 3C Protease Cleaves TRIM7 To Dampen Its Antiviral Activity. *J. Virol.* **2022**, *96*, e0133222. [CrossRef]
32. Gorbalenya, A.E.; Donchenko, A.P.; Blinov, V.M.; Koonin, E.V. Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. A distinct protein superfamily with a common structural fold. *FEBS Lett.* **1989**, *243*, 103–114. [CrossRef]
33. Fan, K.; Wei, P.; Feng, Q.; Chen, S.; Huang, C.; Ma, L.; Lai, B.; Pei, J.; Liu, Y.; Chen, J.; et al. Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. *J. Biol. Chem.* **2004**, *279*, 1637–1642. [CrossRef] [PubMed]
34. Zhang, L.; Lin, D.; Kusov, Y.; Nian, Y.; Ma, Q.; Wang, J.; von Brunn, A.; Leyssen, P.; Lanko, K.; Neyts, J.; et al. α-Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J. Med. Chem.* **2020**, *63*, 4562–4578. [CrossRef]
35. Anand, K.; Palm, G.J.; Mesters, J.R.; Siddell, S.G.; Ziebuhr, J.; Hilgenfeld, R. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *Embo J.* **2002**, *21*, 3213–3224. [CrossRef] [PubMed]
36. Schechter, I.; Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162. [CrossRef]
37. Singh, E.; Khan, R.J.; Jha, R.K.; Amera, G.M.; Jain, M.; Singh, R.P.; Muthukumaran, J.; Singh, A.K. A comprehensive review on promising anti-viral therapeutic candidates identified against main protease from SARS-CoV-2 through various computational methods. *J. Genet. Eng. Biotechnol.* **2020**, *18*, 69. [CrossRef]
38. Azevedo, P.; Camargo, P.G.; Constant, L.E.C.; Costa, S.D.S.; Silva, C.S.; Rosa, A.S.; Souza, D.D.C.; Tucci, A.R.; Ferreira, V.N.S.; Oliveira, T.K.F.; et al. Statine-based peptidomimetic compounds as inhibitors for SARS-CoV-2 main protease (SARS-CoV-2 Mpro). *Sci. Rep.* **2024**, *14*, 8991. [CrossRef] [PubMed]
39. Hayek-Orduz, Y.; Vásquez, A.F.; Villegas-Torres, M.F.; Caicedo, P.A.; Achenie, L.E.K.; González Barrios, A.F. Novel covalent and non-covalent complex-based pharmacophore models of SARS-CoV-2 main protease (M(pro)) elucidated by microsecond MD simulations. *Sci. Rep.* **2022**, *12*, 14030. [CrossRef] [PubMed]
40. Fàbrega-Ferrer, M.; Herrera-Morandé, A.; Muriel-Goñi, S.; Pérez-Saavedra, J.; Bueno, P.; Castro, V.; Garaigorta, U.; Gastaminza, P.; Coll, M. Structure and inhibition of SARS-CoV-1 and SARS-CoV-2 main proteases by oral antiviral compound AG7404. *Antivir. Res.* **2022**, *208*, 105458. [CrossRef] [PubMed]
41. Göhl, M.; Zhang, L.; El Kilani, H.; Sun, X.; Zhang, K.; Brönstrup, M.; Hilgenfeld, R. From Repurposing to Redesign: Optimization of Boceprevir to Highly Potent Inhibitors of the SARS-CoV-2 Main Protease. *Molecules* **2022**, *27*, 4292. [CrossRef]
42. Lockbaum, G.J.; Henes, M.; Lee, J.M.; Timm, J.; Nalivaika, E.A.; Thompson, P.R.; Kurt Yilmaz, N.; Schiffer, C.A. Pan-3C Protease Inhibitor Rupintrivir Binds SARS-CoV-2 Main Protease in a Unique Binding Mode. *Biochemistry* **2021**, *60*, 2925–2931. [CrossRef] [PubMed]
43. Dai, W.; Jochmans, D.; Xie, H.; Yang, H.; Li, J.; Su, H.; Chang, D.; Wang, J.; Peng, J.; Zhu, L.; et al. Design, Synthesis, and Biological Evaluation of Peptidomimetic Aldehydes as Broad-Spectrum Inhibitors against Enterovirus and SARS-CoV-2. *J. Med. Chem.* **2022**, *65*, 2794–2808. [CrossRef]
44. Ramajayam, R.; Tan, K.P.; Liu, H.G.; Liang, P.H. Synthesis and evaluation of pyrazolone compounds as SARS-coronavirus 3C-like protease inhibitors. *Bioorg. Med. Chem.* **2010**, *18*, 7849–7854. [CrossRef] [PubMed]

45. Ramajayam, R.; Tan, K.P.; Liang, P.H. Recent development of 3C and 3CL protease inhibitors for anti-coronavirus and anti-picornavirus drug discovery. *Biochem. Soc. Trans.* **2011**, *39*, 1371–1375. [CrossRef] [PubMed]
46. Mandadapu, S.R.; Weerawarna, P.M.; Prior, A.M.; Uy, R.A.; Aravapalli, S.; Alliston, K.R.; Lushington, G.H.; Kim, Y.; Hua, D.H.; Chang, K.O.; et al. Macrocyclic inhibitors of 3C and 3C-like proteases of picornavirus, norovirus, and coronavirus. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 3709–3712. [CrossRef] [PubMed]
47. Kumar, V.; Shin, J.S.; Shie, J.J.; Ku, K.B.; Kim, C.; Go, Y.Y.; Huang, K.F.; Kim, M.; Liang, P.H. Identification and evaluation of potent Middle East respiratory syndrome coronavirus (MERS-CoV) 3CL(Pro) inhibitors. *Antivir. Res.* **2017**, *141*, 101–106. [CrossRef] [PubMed]
48. Lee, C.C.; Kuo, C.J.; Ko, T.P.; Hsu, M.F.; Tsui, Y.C.; Chang, S.C.; Yang, S.; Chen, S.J.; Chen, H.C.; Hsu, M.C.; et al. Structural basis of inhibition specificities of 3C and 3C-like proteases by zinc-coordinating and peptidomimetic compounds. *J. Biol. Chem.* **2009**, *284*, 7646–7655. [CrossRef]
49. Kuo, C.J.; Liu, H.G.; Lo, Y.K.; Seong, C.M.; Lee, K.I.; Jung, Y.S.; Liang, P.H. Individual and common inhibitors of coronavirus and picornavirus main proteases. *FEBS Lett.* **2009**, *583*, 549–555. [CrossRef]
50. Fili, S.; Valmas, A.; Christopoulou, M.; Spiliopoulou, M.; Nikolopoulos, N.; Lichière, J.; Logotheti, S.; Karavassili, F.; Rosmaraki, E.; Fitch, A.; et al. Coxsackievirus B3 protease 3C: Expression, purification, crystallization and preliminary structural insights. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2016**, *72*, 877–884. [CrossRef]
51. Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 125–132. [CrossRef]
52. Vagin, A.A.; Teplyakov, A. An approach to multi-copy search in molecular replacement. *Acta Cryst. D* **2000**, *56*, 1622–1624. [CrossRef] [PubMed]
53. Adams, P.D.; Afonine, P.V.; Bunkòczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.-W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 213–221. [CrossRef] [PubMed]
54. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. *Acta Cryst. D* **2010**, *66*, 486–501. [CrossRef] [PubMed]
55. Chen, V.B.; Arendall, W.B., III; Headd, J.J.; Keedy, D.A.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Cryst. D* **2010**, *66*, 12–21. [CrossRef] [PubMed]
56. Grzesiek, S.; Bax, A. Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins. *J. Biomol. NMR* **1993**, *3*, 185–204. [CrossRef]
57. Keller, R. *The Computer Aided Resonance Assignment Tutorial*; Cantina: Arth, Switzerland, 2004; pp. 1–81.
58. Williamson, M.P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16. [CrossRef] [PubMed]
59. Farrow, N.A.; Muhandiram, R.; Singer, A.U.; Pascal, S.M.; Kay, C.M.; Gish, G.; Shoelson, S.E.; Pawson, T.; Forman-Kay, J.D.; Kay, L.E. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* **1994**, *33*, 5984–6003. [CrossRef]
60. Grzesiek, S.; Bax, A. The Importance of Not Saturating H2o in Protein Nmr—Application to Sensitivity Enhancement and Noe Measurements. *J. Am. Chem. Soc.* **1993**, *115*, 12593–12594. [CrossRef]
61. Mandel, M.A.; Akke, M.; Palmer, A.G., III. Backbone dynamics of *Escherichia coli* ribonuclease HI: Correlations with structure and function in an active enzyme. *J. Mol. Biol.* **1995**, *246*, 144–163. [CrossRef]
62. Garcia de la Torre, J.; Huertas, M.L.; Carrasco, B. HYDRONMR: Prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* **2000**, *147*, 138–146. [CrossRef] [PubMed]
63. Matthews, D.A.; Dragovich, P.S.; Webber, S.E.; Fuhrman, S.A.; Patick, A.K.; Zalman, L.S.; Hendrickson, T.F.; Love, R.A.; Prins, T.J.; Marakovits, J.T.; et al. Structure-assisted design of mechanism-based irreversible inhibitors of human rhinovirus 3C protease with potent antiviral activity against multiple rhinovirus serotypes. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11000–11007. [CrossRef] [PubMed]
64. Johnson, T.O.; Hua, Y.; Luu, H.T.; Brown, E.L.; Chan, F.; Chu, S.S.; Dragovich, P.S.; Eastman, B.W.; Ferre, R.A.; Fuhrman, S.A.; et al. Structure-based design of a parallel synthetic array directed toward the discovery of irreversible inhibitors of human rhinovirus 3C protease. *J. Med. Chem.* **2002**, *45*, 2016–2023. [CrossRef] [PubMed]
65. Golovanov, A.P.; Hautbergue, G.M.; Wilson, S.A.; Lian, L.Y. A simple method for improving protein solubility and long-term stability. *J. Am. Chem. Soc.* **2004**, *126*, 8933–8939. [CrossRef]
66. Kuo, C.J.; Shie, J.J.; Fang, J.M.; Yen, G.R.; Hsu, J.T.; Liu, H.G.; Tseng, S.N.; Chang, S.C.; Lee, C.Y.; Shih, S.R.; et al. Design, synthesis, and evaluation of 3C protease inhibitors as anti-enterovirus 71 agents. *Bioorg. Med. Chem.* **2008**, *16*, 7388–7398. [CrossRef] [PubMed]
67. Fu, L.; Ye, F.; Feng, Y.; Yu, F.; Wang, Q.; Wu, Y.; Zhao, C.; Sun, H.; Huang, B.; Niu, P.; et al. Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. *Nat. Commun.* **2020**, *11*, 4417. [CrossRef] [PubMed]
68. Lin, C.; Zhu, Z.; Jiang, H.; Zou, X.; Zeng, X.; Wang, J.; Zeng, P.; Li, W.; Zhou, X.; Zhang, J.; et al. Structural Basis for Coronaviral Main Proteases Inhibition by the 3CLpro Inhibitor GC376. *J. Mol. Biol.* **2024**, *436*, 168474. [CrossRef] [PubMed]
69. Cheng, S.; Feng, Y.; Li, W.; Liu, T.; Lv, X.; Tong, X.; Xi, G.; Ye, X.; Li, X. Development of novel antivrial agents that induce the degradation of the main protease of human-infecting coronaviruses. *Eur. J. Med. Chem.* **2024**, *275*, 116629. [CrossRef]

SUPPLEMENTARY MATERIALS

Table S1. Data collection and refinement statistics of the crystal structure of CVB3 $3C^{Pro}$ in complex with GC-376 PROTAC precursor.

	8S6F
Wavelength	1.541
Resolution range	35.13 - 1.932 (2.13 - 1.93)*
Space group	C 1 2 1
Unit cell	76.92 64.35 38.99 90 115.72 90
Total reflections	48285 (10396)
Unique reflections	23725 (5260)
Multiplicity	2.0 (2.0)
Completeness (%)	94.73 (86.18)
Mean I/sigma(I)	3.56 (0.95)
Wilson B-factor	28.98
R-merge	0.1379 (0.7405)
R-meas	0.1834 (0.9709)
R-pim	0.12 (0.6225)
CC1/2	0.981 (0.494)
CC*	0.995 (0.813)
Reflections used in refinement	12206 (2726)
Reflections used for R-free	611 (136)
R-work	0.2384 (0.3192)
R-free	0.2634 (0.3220)
Number of non-hydrogen atoms	1496
macromolecules	1398
ligands	33
solvent	65
Protein residues	180
RMS(bonds)	0.003
RMS(angles)	0.69
Ramachandran favored (%)	94.94
Ramachandran allowed (%)	4.49
Ramachandran outliers (%)	0.56
Rotamer outliers (%)	2.03
Clashscore	5.99
Average B-factor	37.25
macromolecules	37.21
ligands	39.20
solvent	37.29

*Statistics for the highest-resolution shell are shown in parentheses.

Table S2. List of the unassigned backbone NHs for CVB3 $3C^{Pro}$ and CVB3 $3C^{Pro}$ in complex with GC-376 PROTAC precursor. Sample conditions: 50 mM phosphate buffer at pH 6.0 containing 100 mM NaCl, 50 mM arginine, 50 mM glutamate, 1 mM DTT and 1 mM EDTA.

	Unassigned backbone NHs
CVB3 $3C^{Pro}$	A3, E5, V8, K12, T26, M27, L70, L102, I104, T106, N111, M112, Y113, I114, V116, G128, F140, T142, R143, A144, Q146, C147, G148, H161, H167, Q168, F170
CVB3 $3C^{Pro}$ in complex with GC-376 PROTAC precursor	M11, K19, K76, G155, L175

Table S3. List of the residues whose NH chemical shifts change above the threshold upon addition of GC-376 PROTAC and of its precursor.

GC-376 PROTAC	L28, R39, A41, A100, K108, F109, Y122, G123, F124, L125, N126, L127, G129, T130, T132, R134, M135, N139, G149, V150, I160, V162, G163, G164, N165, G166, S171
GC-376 PROTAC precursor	L28, R39, A41, A100, K108, F109, Y122, G123, F124, N126, L127, G129, T130, T132, R134, M135, L136, N139, G149, V150, I160, V162, G163, N165, G166, G169, S171

Table S4. List of the residues unassigned in CVB3 $3C^{Pro}$ and displaying increasing intensities along the stepwise additions of GC-376 PROTAC and of its precursor.

GC-376 PROTAC	A3, E5, K12, T26, M27, L70, L102, I104, T106, N111, M112, Y113, I114, V116, G128, F140, T142, R143, A144, Q146, C147, G148, H161, H167, Q168, F170
GC-376 PROTAC precursor	A3, E5, K12, T26, M27, L70, L102, I104, T106, N111, M112, Y113, I114, V116, F140, T142, R143, A144, Q146, C147, G148, H161, H167, Q168, F170

Figure S1. Superposition of CVB3 $3C^{Pro}$ bound to PROTAC precursor (PDB ID 8S6F) (green) and to (S)-N-benzyl-3-((S)-2-cinnamamido-3-phenylpropanamido)-2-oxo-4-((S)-2-oxopyrrolidin-3-yl)butanamide (PDB-ID 5NFS) (cyan)..

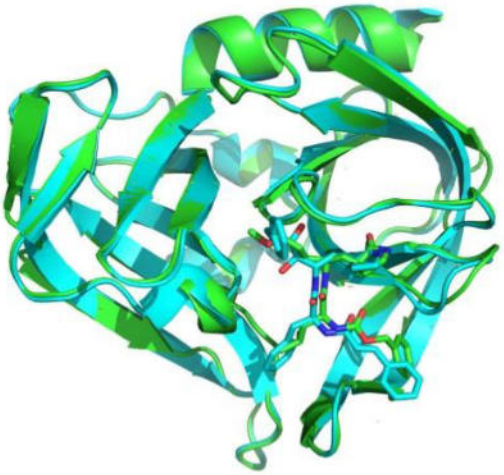


Figure S2. Expression/**purification steps and protein size of CVB3 $3C^{Pro}$**. (**A**) Analytical gel filtration of CVB3 $3C^{Pro}$ compared with protein standard markers (30 kDa in light blue Carbonic anhydrase (CA) and 8.5 kDa in orange Ubiquitin UB). The theoretical molecular weight of monomeric CVB3 $3C^{Pro}$ is 21 kDa. The protein eluted at a volume corresponding to a molecular mass of 22.9 kDa. (**B**) Analysis of CVB3 $3C^{Pro}$ expression and purification on a Coomassie Blue-stained SDS-PAGE gel. Lanes from the left: molecular marker (labelled in kDa); insoluble and supernatant CVB3 $3C^{Pro}$ expressing cells after overnight expression; Immobilized Metal Affinity Chromatography (IMAC) elution and flow-through; Size-Exclusion Chromatography (SEC) elution of the fractions containing the protein.

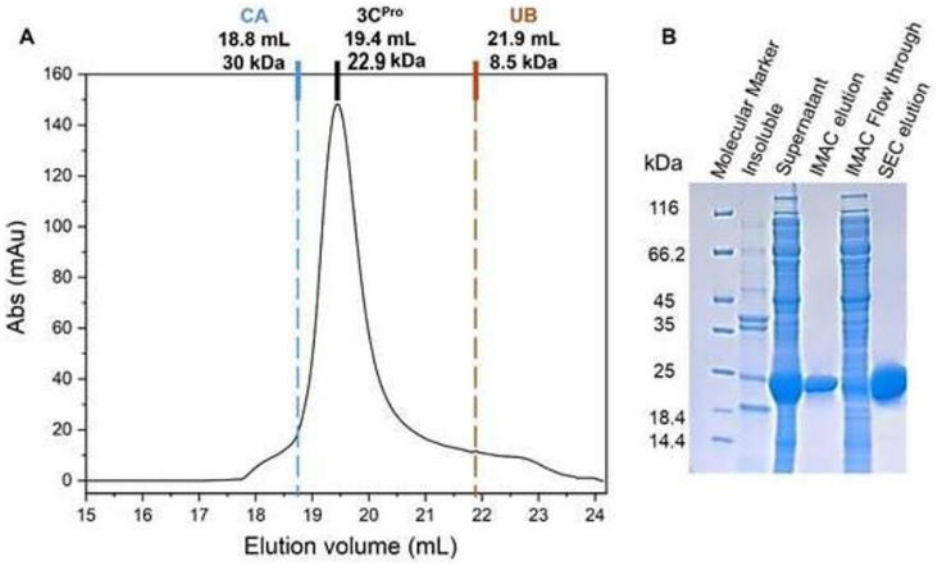


Figure S3. Solution NMR spectra acquired at 500 MHz of CVB3 3C^{Pro} and of CVB3 3C^{Pro} bound to GC-376 PROTAC precursor. ^{1}H-^{15}N HSQC and 3D CBCA(CO)NH spectra of CVB3 3C^{Pro} (**A** and **B**), acquired at 298 K in 50 mM phosphate buffer,100 mM NaCl at pH 6.0 and 10% (v/v) D_2O, and of CVB3 3C^{Pro} bound to GC-376 PROTAC precursor (**C** and **D**), acquired at 308 K in 50 mM phosphate buffer at pH 6.0 containing 100 mM NaCl, 50 mM arginine, 50 mM glutamate, 1 mM DTT and 1 mM EDTA and 10% (v/v) D_2O.

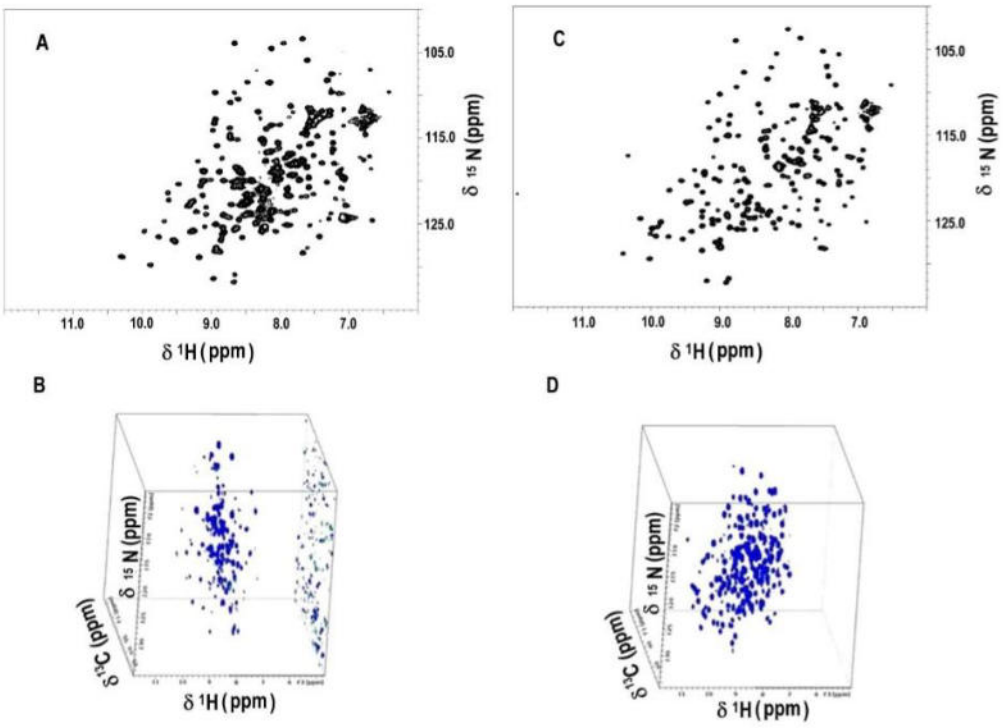


Figure S4. ^{15}N relaxation data of CVB3 $3C^{Pro}$. ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates and $[^1H]^{15}N$ heteronuclear NOE values determined at 500 MHz and 298 K in a 50 mM phosphate buffer at pH 6.0 containing 100 mM NaCl, 50 mM arginine, 50 mM glutamate, 1 mM DTT and 1 mM EDTA. The secondary structure elements are shown and the position of the catalytic triad is indicated with a star.

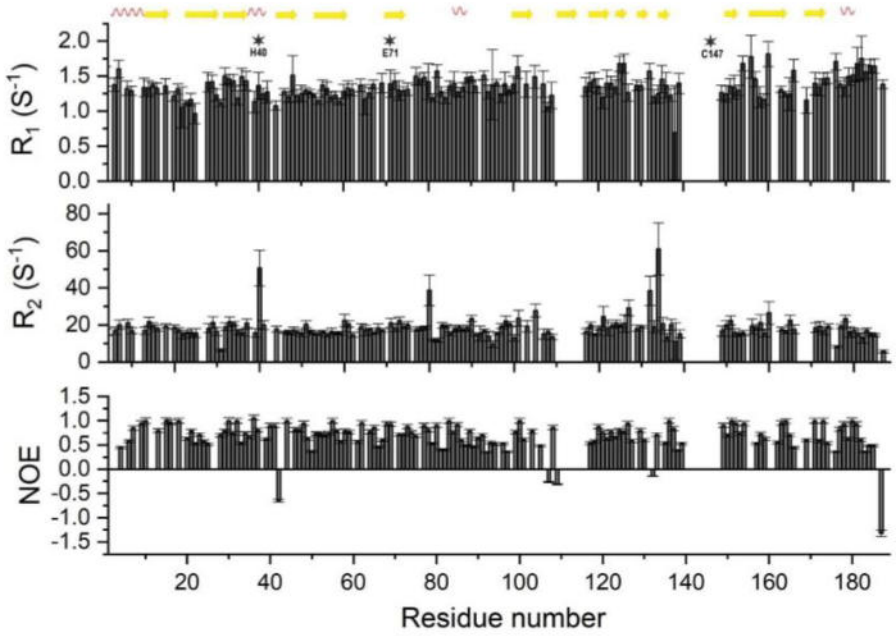


Figure S5. Monitoring the interaction of CVB3 3C^{Pro} with GC-376 PROTAC precursor and GC-376 PROTAC by solution NMR. Backbone weighted average chemical shift differences ($\Delta\delta_{avg}$(HN)) between ^{15}N-labelled CVB3 3C^{Pro} and its 1:1 mixture with GC-376 PROTAC precursor (**A**) and with GC-376 PROTAC (**B**). A chemical shift threshold value of 0.14 ppm, indicated as a dashed line in both panels, was estimated to define the most significant chemical shift differences (see the Materials and Methods section for details). The cyan (in A) and the magenta (in B) bars identify the residues having $\Delta\delta_{avg}$(HN) larger than the threshold value in the two titrations. The corresponding residues are shown as cyan and magenta spheres in Figures 5A and 5B) and are listed in Table S3. Light green and orange bars with $\Delta\delta_{avg}$(HN) = 1 indicate residues whose backbone NH signals are not assigned in the CVB3 3C^{Pro} and appear with increasing intensities along the stepwise additions of the GC-376 PROTAC precursor (light green) or GC-376 PROTAC (orange), respectively. At the end of the titrations their chemical shifts correspond to those of the GC-376 PROTAC precursor-bound or GC-376 PROTAC bound species. The backbone NHs of the latter residues are also identified as light green and orange spheres in Figure 5A and 5B, respectively. These residues are also listed in Table S4.

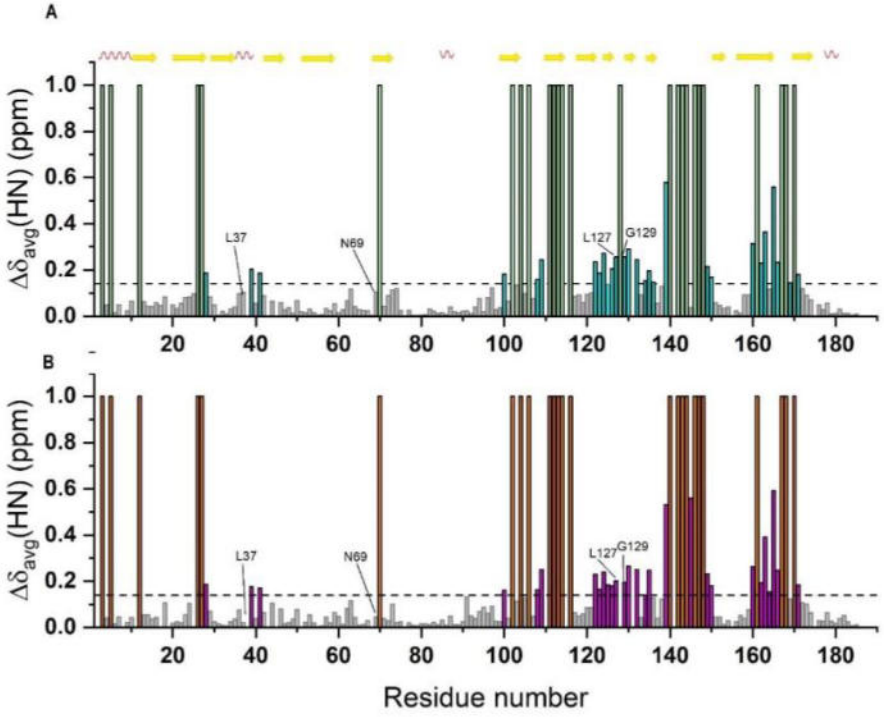


Figure S6. Comparing the interaction of CVB3 $3C^{Pro}$ with GC-376 PROTAC and its precursor by solution NMR. Subtraction of the backbone weighted average chemical shift differences ($\Delta\Delta\delta_{avg}$(HN)) obtained between ^{15}N-labelled CVB3 $3C^{Pro}$ and its 1:1 mixture with GC-376 PROTAC precursor and between ^{15}N-labelled CVB3 $3C^{Pro}$ and its 1:1 mixture with GC-376 PROTAC.

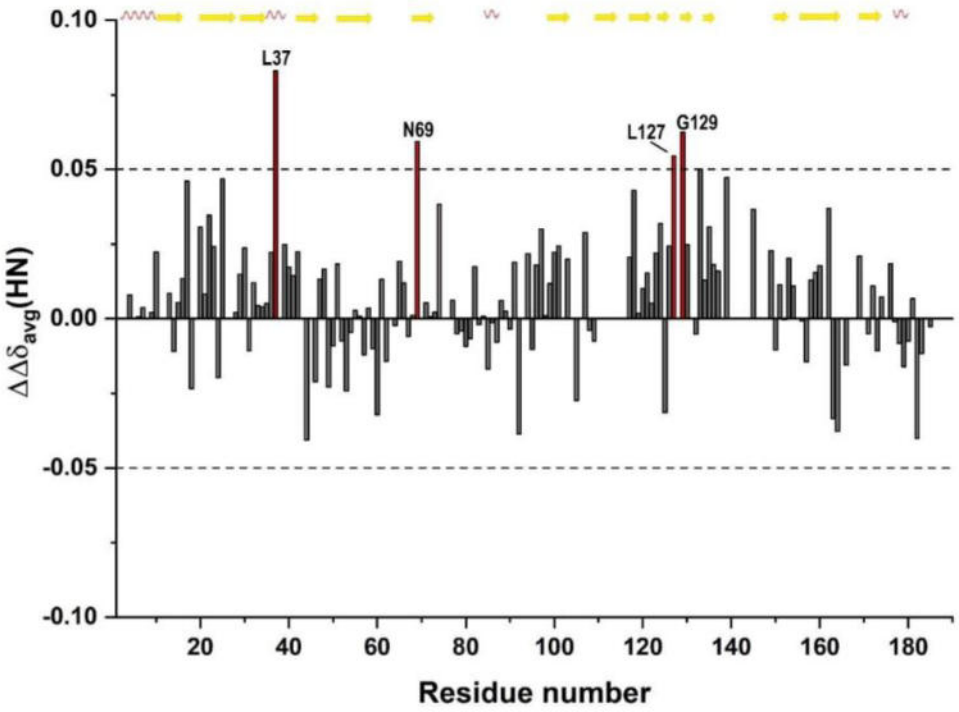


Figure S7. Representative kinetic curves of a control experiment in comparison with an inhibition experiment. Green points and line indicate the inhibition experiment, while the red ones the control experiment.

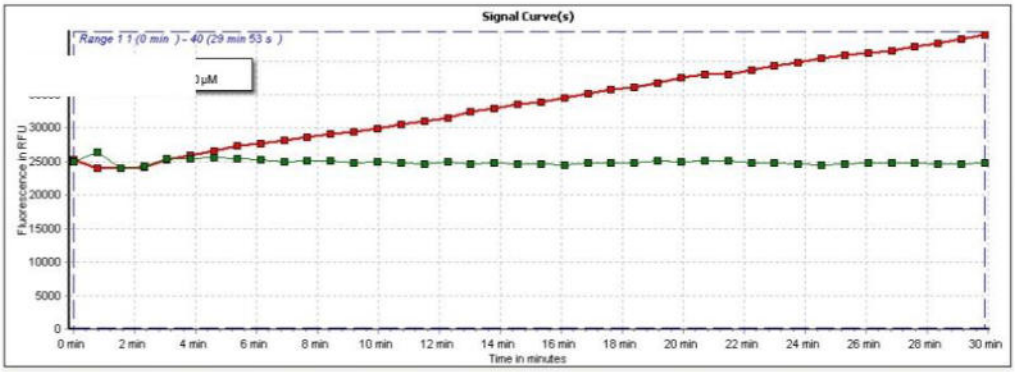


Figure S8. The raw data of enzyme inhibition kinetics of GC-376 PROTAC precursor.

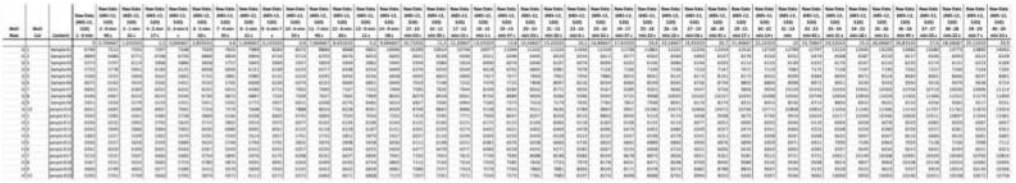


Figure S9. Raw data of calculated inhibition curve of GC-376 PROTAC precursor.

Well Row	Well Col	Content	Raw Data (485-12, 520) 1 - 0 min	Linear regression fit of Range 1 based on Raw Data (485-12, 520) 1 - 0 min	Slope of Linear regression fit (Range 1) based on Raw Data (485-12, 520)	r² of Linear regression fit (Range 1) based on Raw Data (485-12, 520)	% Inhibition (calculated using the formula 100-(v1/vo)*100	Inhibitor concentration			Log[M]	% Inhibition	% Inhibition
			0	0									
G	1	Sample X1	6740	6740	4,142022	0,969915	ctrl	0	3,733627	average ctrl	-4	103,732	97,9772
G	2	Sample X2	6899	6899	-0,13934	0,502185	103,732	100 µM			-4,5	93,90662	89,4247
G	3	Sample X3	5373	5373	0,227504	0,743477	93,90662	30 µM			-5	77,28203	77,5484
G	4	Sample X4	5755	5755	0,848205	0,893319	77,28203	10 µM			-5,5	45,77699	42,69942
G	5	Sample X5	5560	5560	2,024485	0,946434	45,77699	3 µM			-6	31,51275	40,99121
G	6	Sample X6	6675	6675	2,557059	0,969077	31,51275	1 µM			-6,5	16,97036	12,48713
G	7	Sample X7	6602	6602	3,100018	0,988537	16,97036	300 nM			-7	13,5861	18,27182
G	8	Sample X8	6546	6546	3,226373	0,994169	13,5861	100 nM			-8	17,36143	17,98093
G	9	Sample X9	5451	5451	3,085416	0,994116	17,36143	10 nM			-9	16,51646	16,20282
G	10	Sample X10	6022	6022	3,116964	0,993718	16,51646	1 nM					
H	1	Sample X11	5920	5920	3,325233	0,996018	ctrl	0					
H	2	Sample X12	6336	6336	0,075524	0,102954	97,9772	100 µM					
H	3	Sample X13	6543	6543	0,394842	0,875302	89,4247	30 µM					
H	4	Sample X14	5683	5683	0,838259	0,874699	77,5484	10 µM					
H	5	Sample X15	5936	5936	2,13939	0,943858	42,69942	3 µM					
H	6	Sample X16	6302	6302	0,203168	0,606018	94,55842	1 µM					
H	7	Sample X17	5510	5510	3,267405	0,995398	12,48713	300 nM					
H	8	Sample X18	5567	5567	3,051426	0,99541	18,27182	100 nM					
H	9	Sample X19	4660	4660	3,062287	0,98938	17,98093	10 nM					
H	10	Sample X20	5505	5505	3,128675	0,995616	16,20282	1 nM					

Figure S10. Raw data of enzyme inhibition kinetics of GC-376 PROTAC.

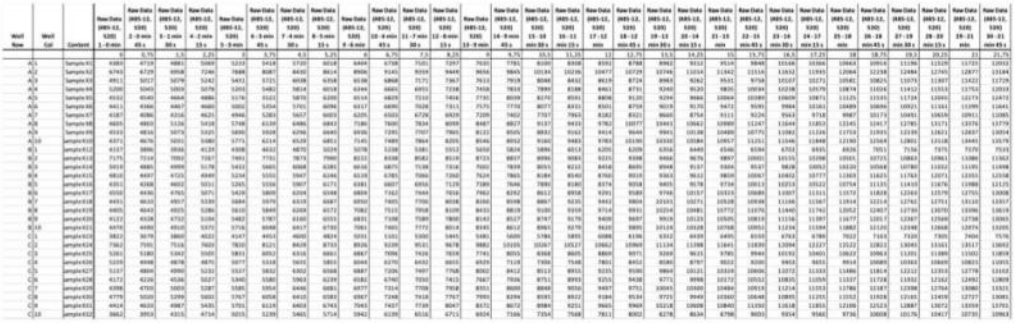


Figure S11. Raw data of calculated inhibition curve of GC-376 PROTAC.

Well Row	Well Col	Content	Raw Data (485-12, 520) 1 - 0 min	Linear regression fit of Range 1 based on Raw Data (485-12, 520) 1 - 0 min	Slope of Linear regression fit (Range 1) based on Raw Data (485-12, 520)	r^2 of Linear regression fit (Range 1) based on Raw Data (485-12, 520)	% Inhibition (calculated using the formula 100-(v1/vo)* 100	Inhibitor concentration				Log[M]	% Inhibition	% Inhibition	% Inhibition
			0	0											
A	1	Sample X1	4383	4383	8,189947	0,97291	ctrl	0	8,08531	average ctrl		-3,5	47,22549	55,49012	40,42337
A	2	Sample X2	6743	6743	4,267196	0,991048	47,22286	300 µM				-4	35,47965	29,8194	31,10195
A	3	Sample X3	4911	4911	5,216931	0,972472	35,47643	100 µM				-4,5	45,46198	*17,96231**	50,47114
A	4	Sample X4	5200	5200	4,409788	0,917966	45,46198	30 µM				-5	18,87188	16,19225	15,01112
A	5	Sample X5	4532	4532	6,559788	0,986794	18,87188	10 µM				-6	17,29813	4,233739	8,209004
A	6	Sample X6	4411	4411	6,687037	0,982113	17,29813	1 µM				-7	12,63905	7,521921	7,885094
A	7	Sample X7	4187	4187	7,063757	0,996297	12,63905	100 nM				-8	5,571915	6,76613	*25,52022**
A	8	Sample X8	4605	4605	7,635185	0,998848	5,571915	10 nM				-9	15,5019	1,48868	5,208742
A	9	Sample X9	4533	4533	6,832275	0,997156	15,5019	1 nM				-10	0,952101	0	*23,4557**
A	10	ample X10	4371	4371	8,00873	0,997394	0,952101	0.1 nM							
B	1	ample X12	4137	4137	8,308466	0,972267		0				*data excluded for the fitting			
B	2	ample X13	7175	7175	3,598942	0,949898	55,49012	300 µM							
B	3	ample X14	5019	5019	5,674603	0,981609	29,8194	100 µM							
B	4	ample X15	4810	4810	6,633333	0,993902	17,96231	30 µM							
B	5	ample X16	4351	4351	6,776455	0,994843	16,19225	10 µM							
B	6	ample X17	4550	4550	7,743386	0,998126	4,233739	1 µM							
B	7	ample X18	4431	4431	7,477513	0,998734	7,521921	100 nM							
B	8	ample X19	4405	4405	7,538624	0,993911	6,76613	10 nM							
B	9	ample X20	4122	4122	7,965344	0,998676	1,488679	1 nM							
B	10	ample X21	4470	4470	8,085714	0,996345	0	0.1 nM							
C	1	ample X23	3822	3822	7,757519	0,993549	4,058958	0							
C	2	ample X24	7362	7362	4,817196	0,932391	40,42337	300 µM							
C	3	ample X25	5261	5261	5,570899	0,991348	31,10195	100 µM							
C	4	ample X26	5229	5229	4,004762	0,901882	50,47114	30 µM							
C	5	ample X27	5137	5137	6,871958	0,990076	15,01112	10 µM							
C	6	ample X28	4172	4172	7,421958	0,995291	8,209004	1 µM							
C	7	ample X29	4398	4398	7,448148	0,994639	7,885094	100 nM							
C	8	ample X30	4779	4779	6,022222	0,994965	25,52022	10 nM							
C	9	ample X31	4414	4414	7,66455	0,997176	5,208742	1 nM							
C	10	ample X32	3662	3662	6,189153	0,986633	23,4557	0.1 nM							

3.5. - Investigation of GC-376 PROTAC ternary complex

3.5.1. - Testing the SARS-CoV-2 $3CL^{Pro}$ or CVB3 $3C^{Pro}$:GC-376 PROTAC: CRBN-TBD complex

As a first step, to confirm that the pomalidomide moiety of GC-376 PROTAC binds to CRBN-TBD, fluorescence experiments were carried out with a Cary Eclipse fluorescence spectrometer. The chosen excitation wavelength for the fluorescence experiment was 290 nm, which coincides with the maximum absorbance wavelength of the protein detected with a scanning excitation at a range of wavelengths between 300 and 500 nm. All experiments were carried out at 298 K. A 1 mM stock solution of GC-376 PROTAC was obtained in 20 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP buffer and added in small increments. At each subsequent addition of GC-376 PROTAC up to 4:1 the concentration of the protein (25 μM), a fluorescence spectrum was acquired and overlaid to the others (**Figure 32**).

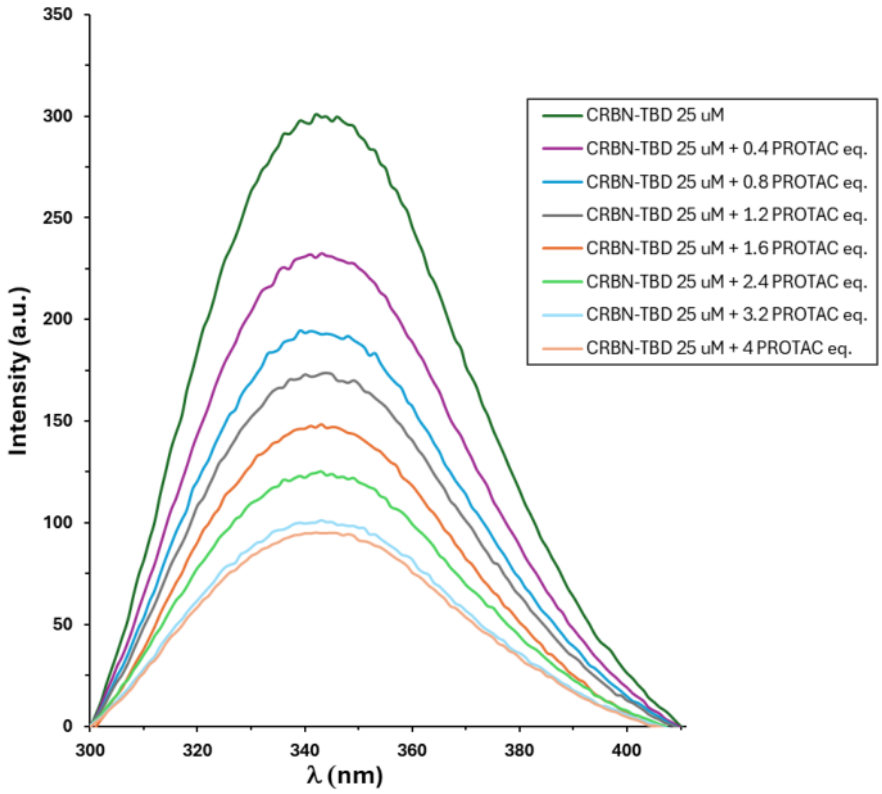


Figure 32. Overlaid fluorescence spectra of CRBN-TBD at increasing equivalents (eq.) of GC-376 PROTAC. Measurements were performed in 20 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP buffer solutions. Spectra were acquired at 298K with a Cary Eclipse fluorescence spectrometer. Excitation λ: 290 nm. Emission λ: 300-500 nm. The plot was generated with Excel (Microsoft Office) program.

We can observe that each addition of GC-376 PROTAC causes a decrease in fluorescence intensity and that, consequently, it monitors the interaction between the molecule and CRBN-TBD. Indeed, an excitation wavelength of 290 nm successfully targets Trp residues in the thalidomide binding site of CRBN-TBD which emits at maximal intensity at 348 nm [357]. The pomalidomide moiety of GC-376 PROTAC binds to CRBN-TBD since every additional aliquot of the molecule causes further quenching of fluorescent signal emitted by the protein binding pocket.

Subsequently, we tried to investigate the formation of the ternary complex target protease (SARS-CoV-2 $3CL^{Pro}$ or CVB3 $3C^{Pro}$):GC-376 PROTAC:CRBN-TBD by performing an analytical SEC. CRBN-TBD and target protease were first mixed at a molar ratio of 1:4, respectively. Then GC-376 PROTAC was added 4-fold-times the concentration of target protein, which was about 200 µM. The complex mixture was incubated for 4 hours at room temperature. After incubation the solution was loaded onto a Superdex 200 Increase 10/300 GL column (Cytiva). The elution was performed with a 20 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP buffer solution.

For both complex mixtures, the results turned out to be unsuccessful. **Figure 32** shows the case of SARS-CoV-2 $3CL^{Pro}$:GC-376 PROTAC:CRBN-TBD complex mixture. The chromatogram of the mixture exhibits only two distinct peaks. When overlaid with the chromatograms of pure CRBN-TBD (~12 kDa) and pure SARS-CoV-2 $3CL^{Pro}$ (~66 kDa), acquired under the same chromatographic conditions (identical column and buffer system), the retention times and peak profiles of the mixture coincide precisely with those of the individual components. This correspondence confirms that no additional species are present under the experimental conditions, thereby demonstrating the absence of ternary complex formation (**Figure 33**).

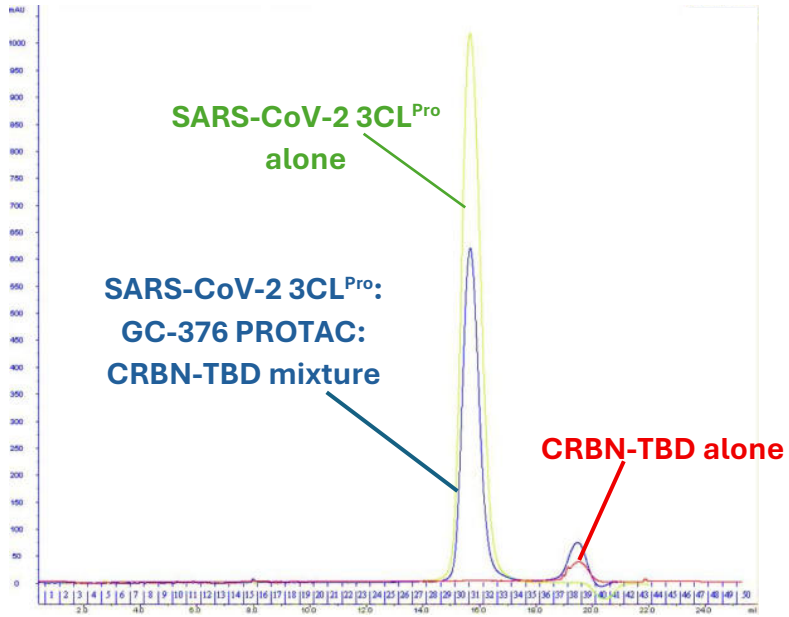


Figure 33. Analytical SEC of SARS-CoV-2 $3CL^{Pro}$:GC-376 PROTAC:CRBN-TBD mixture. Comparative overlay of UV-chromatograms: in blue the curve corresponding to the complex mixture, in green and red the runs related to the SARS-CoV-2 $3CL^{Pro}$ and CRBN-TBD alone, respectively. Isocratic elution was performed with 1.2 CV on a Superdex 200 Increase 10/300 GL column (Cytiva). Elution buffer: 20 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP.

Following incubation, the appearance of a yellowish precipitate was observed, most likely resulting from PROTAC precipitation. Such precipitation could significantly reduce the availability of soluble PROTAC, thereby impairing the ability to mediate the formation of a ternary complex. In addition, the chromatographic procedure itself may have contributed to the lack of ternary complex detection, as dilution of the complex mixture during the separation process could lower the effective concentrations of the components required for stable complex formation. Taken together, these factors provide explanations for the absence of a detectable ternary complex under the experimental conditions employed.

3.5.2. - CRBN-midi E3 ubiquitin ligase production

For studies of ternary complexes, which often involve interactions with the Lon domain of CRBN, the TBD alone may not be representative, since it lacks a significant portion of the full-length CRBN protein. To address these limitations, we chose to express and use a new truncated CRBN construct, containing both TBD and Lon domains, developed by A. Ciulli et al., (2024) [358]. It would conveniently express soluble and stable protein from E. coli with comparable functionality to the full-length wild-type CRBN. The plasmid, called His6-TEV-CRBN-midi, allows the expression of CRBN-midi as a fusion protein with an N terminal His-tag adjacent to a TEV cleavage site. The protocol for the expression and purification was adapted from [358]. Briefly, the protein was expressed in LB medium at 37°C and induced overnight at 18°C with IPTG 0,5 mM and $ZnCl_2$ 50 µM final concentration. The purification was performed by using a HisTrap column, exploiting the His-tag in the protein sequence. Then the protein solution was incubated overnight with TEV protease to remove the His-tag. After a counter-column step to elute the free His-tag CRBN-midi, the protein solution was subjected to SEC purification (Superdex 200 16 60 column (Cytiva)) (**Figure 34**). The final protein yield was about 2 mg/L of culture.

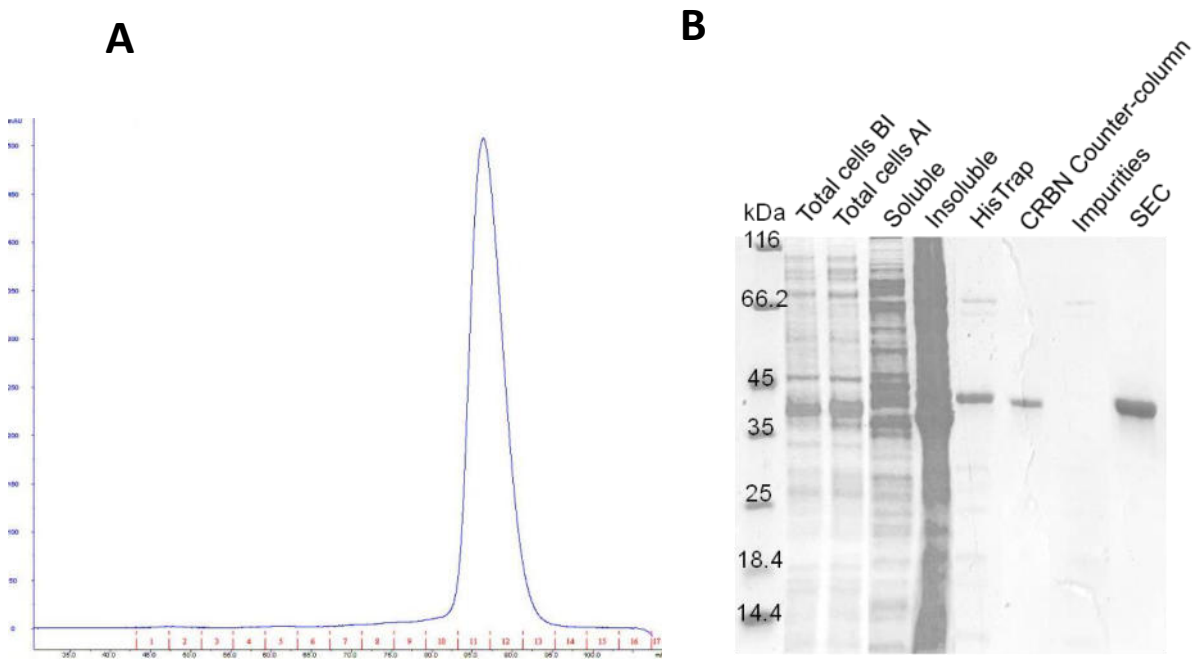


Figure 34. SEC profile of CRBN-midi purification (**A**). Two peaks are visible, one eluted at 60-70 ml corresponding to the protein impurities, the other at 90-95 ml related to the CRBN-midi. Isocratic elution was performed with 1.2 CV on a Superdex 200 16 60 column (Cytiva). Elution buffer: 20 mM HEPES pH 7.5, 500 mM NaCl, and 0.5 mM TCEP. SDS-PAGE analysis of protein samples in the main steps of purification on a 4–15% Mini-PROTEAN® TGX™ Precast Protein Gel with the molecular marker (Bio-rad) (**B**).

3.5.3. - Analysis of the SARS-CoV-2 3CL^{Pro} or CVB3 3C^{Pro}:GC-376 PROTAC:CRBN-midi complex

Since the unsuccessful results with CRBN-TBD I moved on to the investigation of the formation of the ternary complex with CRBN-midi.

To confirm that the pomalidomide moiety of GC-376 PROTAC binds to CRBN-midi, as described above, fluorescence experiments, were carried out with a Cary Eclipse fluorescence spectrometer. A 1 mM stock solution of GC-376 PROTAC was obtained in 20 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP buffer and added in small increments. At each subsequent addition of GC-376 PROTAC up to 4:1 the concentration of the protein (25 μM), a fluorescence spectrum was acquired and overlaid to the others (**Figure 35**). The chosen excitation wavelength for the fluorescence experiment was 290 nm, which coincides with the maximum absorbance wavelength

of the protein detected with a scanning excitation at a range of wavelengths between 300 and 500 nm. All experiments were carried out at 298 K.

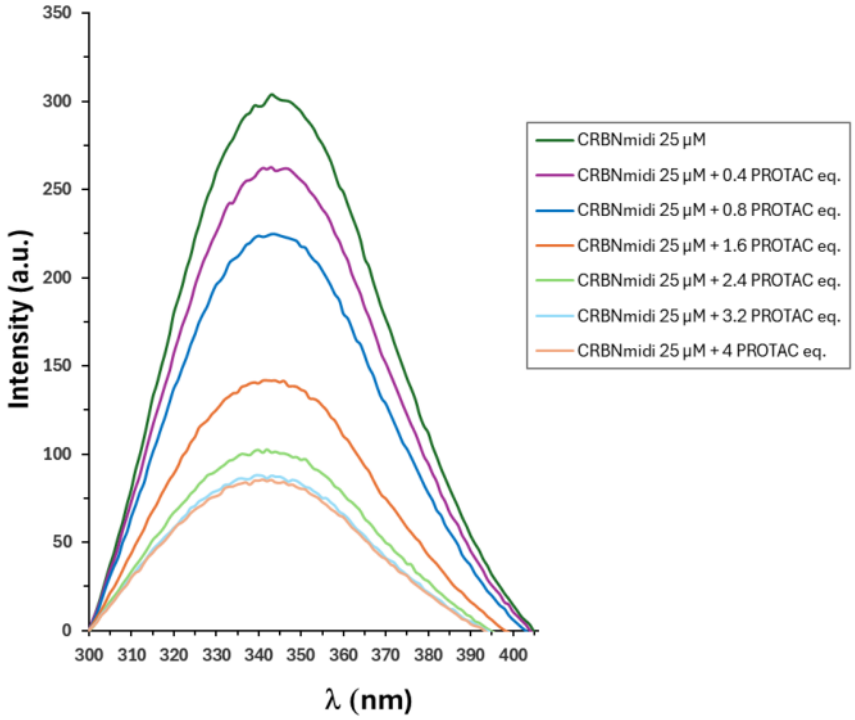


Figure 35. Overlaid fluorescence spectra of CRBN-midi at increasing equivalents (eq.) of GC-376 PROTAC. Measurements were performed in 20 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP buffer solution. Spectra were acquired at 298K with a Cary Eclipse fluorescence spectrometer. Excitation λ: 290 nm. Emission λ: 300-500 nm. The plot was generated with Excel (Microsoft Office) program.

Next, we investigated the potential formation of the ternary complex by NMR, focusing on the CVB3 $3C^{Pro}$: GC-376 PROTAC: CRBN-midi system. The ^{15}N-labelled CVB3 $3C^{Pro}$ protein (~70 µM in 20 mM HEPES, 200 mM NaCl, 0.5 mM TCEP, 2 mM DTT, pH 7.5) was first titrated with GC-376 PROTAC by adding small aliquots until reaching saturation (protein-to-ligand ratio of up to 1:5). After each addition, an HSQC spectrum was acquired. Subsequently, ^{15}N-labelled CRBN-midi (~40 µM in the same buffer conditions) was added to the sample, yielding a final NMR volume of 550 µL and a stoichiometric ratio of 1:1:1 for CVB3 $3C^{Pro}$:GC-376 PROTAC:CRBN-midi. Bidimensional NMR experiments were then recorded.

CSPs were calculated as backbone-weighted average differences ($\Delta\delta_{avg}$(HN)), as described previously (see Sections 3.2.1. and 3.4.2.). For the analysis, chemical shift assignment of CVB3 $3C^{Pro}$ in the GC-376 precursor–bound state (BMRB ID: 52547) was used as reference. The CSP

analysis revealed only minimal differences (**Figure 36**), with some residues showing $\Delta\delta_{avg}$ values above the 0.022 ppm threshold located near or within the catalytic region. However, these shifts were too small to provide clear evidence of ternary complex formation. This suggests that further optimization of experimental conditions will likely be required.

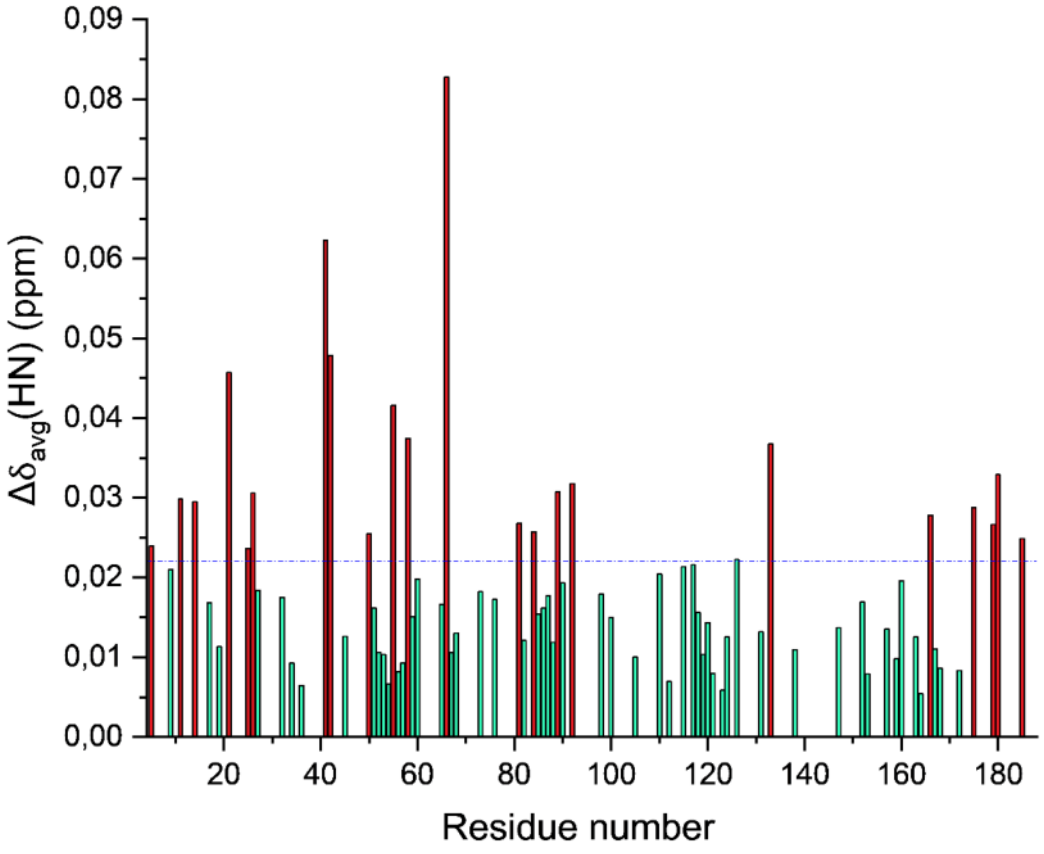


Figure 36. Backbone average chemical shift differences ($\Delta\delta_{avg}$(HN)) between ^{15}N-labelled CVB3 3C^{Pro} and its 1:1:1 mixture with GC-376 PROTAC and ^{15}N-labelled CRBN-midi. A chemical shift threshold value of 0.022 ppm, indicated as a blue dashed line, was estimated to define the most significant chemical shift differences. The red bars indicate the residues with a $\Delta\delta_{avg}$(HN) higher than the threshold. The plot was generated by OriginLab software (Origin 2022).

Since the NMR results were not satisfactory, we tried to investigate the formation of the ternary complex target protease CVB3 3C^{Pro}: GC-376 PROTAC:CRBN-midi by performing an analytical SEC. CRBN-midi and target protease were first mixed at a molar ratio of 1:4, respectively. Then GC-376 PROTAC was added 4-fold-times the concentration of target protein, which was about 200 μM. The complex mixture was incubated for 4 hours at room temperature. After incubation

the solution was loaded onto a Superdex 200 Increase 10/300 GL (Cytiva) column. The elution was performed with a 20 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP buffer solution. **Figure 36** shows the overlay of the mixture chromatogram with CRBN-midi and CVB3 $3C^{Pro}$ alone curves, acquired under the same chromatographic conditions (identical column and buffer solution, section 3.5.1.). In the chromatogram of the complex mixture, three distinct peaks are observed, consistent with the expected elution profile and providing clear evidence of ternary complex formation. The first peak corresponds to the assembled ternary complex (~60 kDa), confirming the simultaneous association of CRBN-midi and CVB3 $3C^{Pro}$ mediated by the PROTAC. The second peak corresponds to unbound CRBN-midi (~37 kDa), while the third represents free CVB3 $3C^{Pro}$ (~21 kDa). This data was further verified through SDS-PAGE of the fractions corresponding to the peaks and the presence of the ternary complex was confirmed by the bands belonging to 29-30 elution fractions of the complex mixture chromatogram (**Figure 37**).

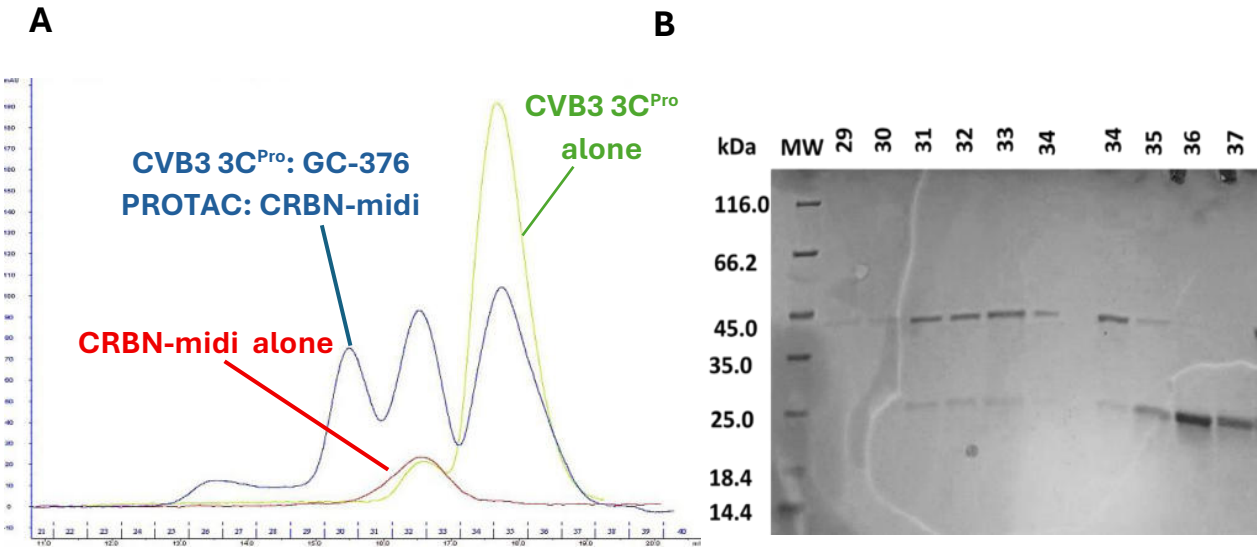


Figure 37. Analytical SEC of CVB3 $3C^{Pro}$:GC-376 PROTAC:CRBN-midi mixture (**A**). Comparative overlay of UV-chromatograms: in blue the curve corresponding to the complex mixture, in green and red the runs related to the CVB3 $3C^{Pro}$ and CRBN-midi alone, respectively. Isocratic elution was performed with 1.2 CV on a Superdex 200 Increase 10/300 GL column (Cytiva). Elution buffer: 20 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP. SDS-PAGE analysis of the analytical SEC elution fractions on a 4–15% Mini-PROTEAN® TGX™ Precast Protein Gel with the molecular marker (Bio-rad) (**B**).

We performed the same analytical SEC experiment onto the SARS-CoV-2 $3CL^{Pro}$:GC-376 PROTAC:CRBN-midi under the experimental conditions described above. In this experiment, evidence for ternary complex formation is inconclusive. The chromatogram displays two closely

eluting peaks that align with those of the individual species, SARS-CoV-2 3CL^{Pro} and CRBN-midi. To assess this, the elution fractions corresponding to the peaks were analyzed by SDS–PAGE. Fractions 30–31–32 show a band with a standard molecular weight below ~35 kDa, consistent with monomeric SARS-CoV-2 3CL^{Pro} under denaturing conditions (~33 kDa), along with a faint band below ~45 kDa corresponding to CRBN-midi (~37 kDa). These observations could indicate the presence of ternary complex, at least in fractions 30–31, given the chromatographic peak intensity. Alternatively, the results may simply reflect co-elution of the individual proteins, with analytical gel filtration lacking sufficient resolving power to separate the species. Accordingly, the data does not unambiguously demonstrate ternary complex formation (**Figure 38**).

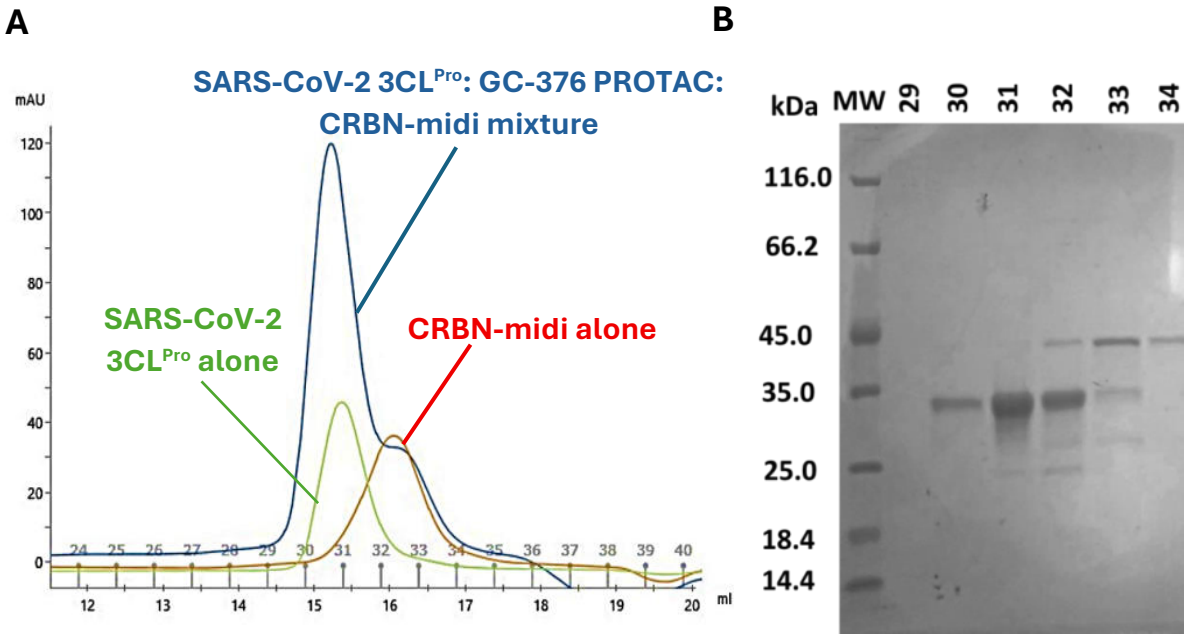


Figure 38. Analytical SEC of SARS-CoV-2 3CL^{Pro}:GC-376 PROTAC:CRBN-midi mixture (**A**). Comparative overlay of UV-chromatograms: in blue the curve corresponding to the complex mixture, in green and red the runs related to the SARS-CoV-2 3CL^{Pro} and CRBN-midi alone, respectively. Isocratic elution was performed with 1.2 CV on a Superdex 200 Increase 10/300 GL column (Cytiva). Elution buffer: 20 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP. SDS-PAGE analysis of the analytical SEC elution fractions on a 4–15% Mini-PROTEAN® TGX™ Precast Protein Gel with the molecular marker (Bio-rad) (**B**).

3.6. - Achievements in SARS-CoV- 2 Nsp13 production

Concerning the Nsp13 project, I was involved in the setting up and optimization of the protocol for protein expression and purification. SARS-CoV-2 Nsp13, encoded by pGEX-SARS CoV2 NSP12 NSP13 (DU67899 MRCPPU) plasmid, was expressed in Rosetta 2 (DE3) pLysS cells in HTMC medium supplemented with Zn acetate solution and induction was performed overnight at 16°C with IPTG 1 mM final concentration. The protein was purified using a GSTrap-affinity column exploiting its GST-tag at the N-terminus. In particular, zinc administration to the Rosetta pLysS strain during the protein production process and other minor adjustments to the purification steps, such as 5% glycerol in lysis buffer and use of protease inhibitors in buffer solutions, resulted to be essential to increase the protein output. The final protein yield was about 0,2 mg/L of culture (**Figure 39 A**).

Nsp13 contains an N- terminal domain reported to coordinate three zinc ions (zinc-binding domain or ZBD). The work by Rouault et al., 2023 [359], demonstrated that Nsp13 ligates a [Fe_4S_4] cluster rather than zinc in one of the three zinc sites in its ZBD. According to this study, the Fe–S cluster is crucial for the optimal activity of Nsp13, highlighting its potential as a therapeutic target for COVID-19 [359]. Therefore, in order to increase the yield, I tried to add a solution of $FeCl_3$ as an iron source in the protein expression. The obtained protein output was about 0,5 mg/L of culture (**Figure 39 B**).

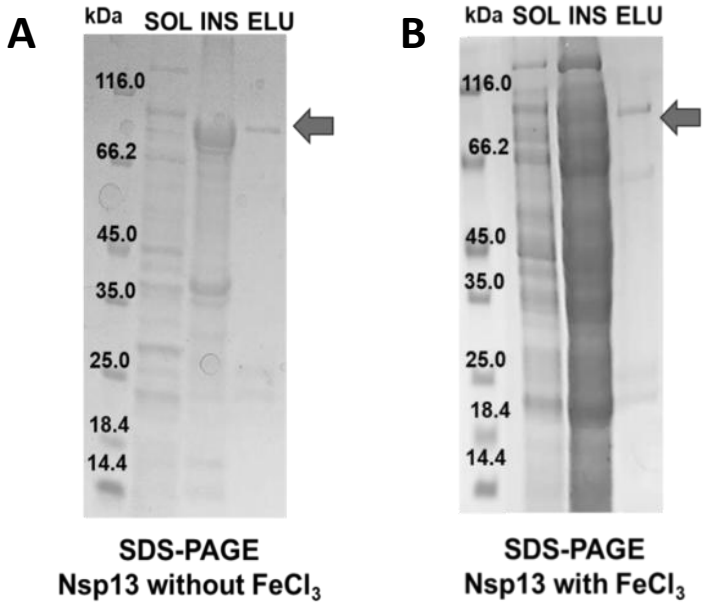


Figure 39. SDS-PAGE analysis of Nsp13 expression without (**A**) and in presence (**B**) of $FeCl_3$ solution (500 μM final concentration in the medium) on a 4–15% Mini-PROTEAN® TGX™ Precast Protein Gel (Bio-rad). Soluble, insoluble and elution after GSTrap Hp column (Cytiva) protein samples are shown after the molecular marker (Bio-rad).

4. Chapter – Conclusions and perspectives

My PhD project focuses on viral proteins that result key therapeutic targets, with particular emphasis on structural biology approaches to investigate their interactions with small molecules. These studies aim to provide a framework for validating specific proteins or domains as druggable entities, ultimately contributing to the development of more effective antiviral therapies. The proteins under investigation include the main viral proteases SARS-CoV-2 $3CL^{Pro}$ and CVB3 $3C^{Pro}$, as well as the SARS-CoV-2 helicase Nsp13.

Below I outline the main achievements and perspectives of my three-year PhD research on the viral proteins. SARS-CoV-2 $3CL^{Pro}$, a key enzyme in the viral replication cycle, was investigated with particular focus on the inhibitory activity of Auranofin and selected derivatives. The determined Ki values ranged from 2.1 to 0.4 μM. ESI-MS spectra confirmed the formation of adducts between the inhibitors and the protein, while X-ray crystallography identified the specific gold-binding sites. This work not only provided insights into the inhibitory mechanism of gold-based compounds but also led to the establishment of robust protocols for screening diverse chemical libraries to identify high-affinity inhibitors. Moreover, a broader exploration of gold-based and other metal-centered inhibitors could significantly enhance the understanding of structure–activity relationships underlying SARS-CoV-2 $3CL^{Pro}$ inhibition. Future studies should also focus on assessing the antiviral efficacy of the most promising compounds in relevant cellular and animal models, thereby bridging the gap between in vitro inhibition and potential therapeutic application.

The high conservation of coronavirus $3CL^{Pro}$, combined with its absence in humans, makes it an attractive antiviral target. GC-376 is among the most potent inhibitors of $3CL^{Pro}$, showing activity against important pathogenic coronaviruses. Building on structural data, we applied a PROTAC based approach derived from GC-376 inhibitor triggering the proteolysis of the SARS-CoV-2 $3CL^{Pro}$. Structural studies (NMR, X-ray) revealed that the ligand binds reversibly to Cys145 in the active site of SARS-CoV-2 $3CL^{Pro}$, while the linker-CRBN module remains flexible. Functional assays demonstrated that this GC-376 PROTAC reduces $3CL^{Pro}$ protein levels in cells without impairing viability.

Given the structural similarity between coronavirus $3CL^{Pro}$ and enterovirus $3C^{Pro}$, the PROTAC was also tested against $3C^{Pro}$ from the cardiotropic enterovirus CVB3. Expression of isotopically labelled CVB3 $3C^{Pro}$ enabled characterization by NMR, X-ray crystallography, and biochemical

assays. The crystal structure of CVB3 $3C^{Pro}$ bound to the PROTAC precursor (1.9 Å, PDB 8S6F) and backbone assignments confirmed a binding mode highly consistent with SARS-CoV-2 $3CL^{Pro}$.

In addition to GC-376 PROTAC, we evaluated the activity of the two related molecules KME16 and ACH213. Both compounds displayed markedly lower efficiency, particularly against CVB3 $3C^{Pro}$. A plausible explanation for their reduced activity lies in the relative orientation of the linker and the pomalidomide–CRBN ligand. In the case of GC-376 PROTAC, these elements are positioned on the right side of the γ-lactam ring, precisely where the covalent bond between the catalytic Cys145 and the β-unsaturated carbon is established (a structural model of the protein-GC-376 PROTAC complex is reported in **Figure 40**). Conversely, in KME16 and ACH213, the linker and pomalidomide–CRBN ligand are functionalized on the left side of the γ-lactam ring and, in other terms, on the opposite side of the expected covalent interaction site with the catalytic cysteine mediated by the nitrile group. This structural arrangement could explain their reduced capacity to promote efficient ternary complex formation.

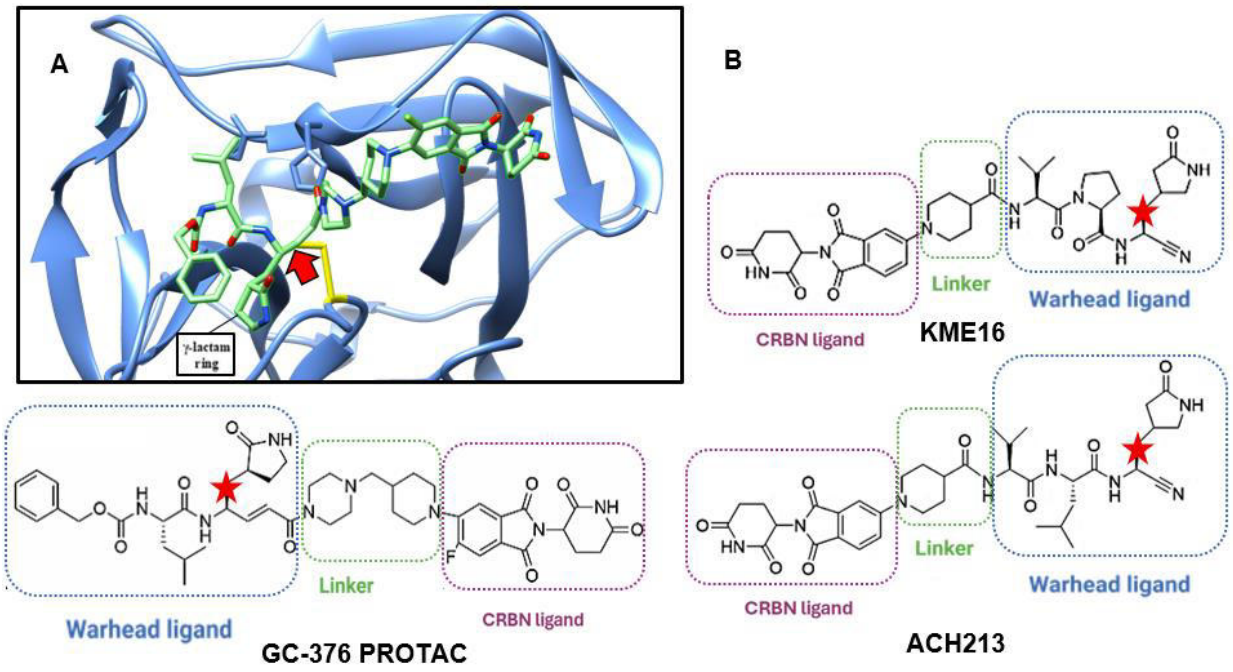


Figure 40. Model structure of CVB3 $3C^{Pro}$ (PDB ID 8S6F) in complex with GC-376 PROTAC (obtained by UCSF Chimera software) (**A**). The red arrow indicates the β-unsaturated carbon where the covalent bond would occur with the catalytic Cys147, highlighted in yellow. 2D structure of the PROTAC molecules (**B**). The red stars indicate the γ-lactam ring.

Finally, we investigated the ternary complex formation of both target proteases with CRBN-TBD or CRBN-midi and PROTAC molecules by analytical gel filtration and solution NMR.

Specifically, we obtained evidence of the ternary complex formation CVB3 $3C^{Pro}$:GC-376 PROTAC:CRBN-midi by analytical SEC.

In the case of SARS-CoV-2 $3CL^{Pro}$, we were unable to demonstrate ternary complex formation. A plausible explanation for this outcome may be related to the larger size of the protein and structural arrangement. As shown in **Figure 41**, a comparison between the monomeric CVB3 $3C^{Pro}$ in complex with GC-376 PROTAC and the crystallographic structure (PDB ID: 8OKC) of the dimeric SARS-CoV-2 $3CL^{Pro}$ bound to the same PROTAC molecule highlights a critical difference. In the latter, the linker and the pomalidomide–CRBN ligand are positioned near the C-terminal domain of the adjacent protomer within the dimer. This steric hindrance is probably enhanced by the presence of CRBN-midi, thus not encouraging the formation of the ternary complex.

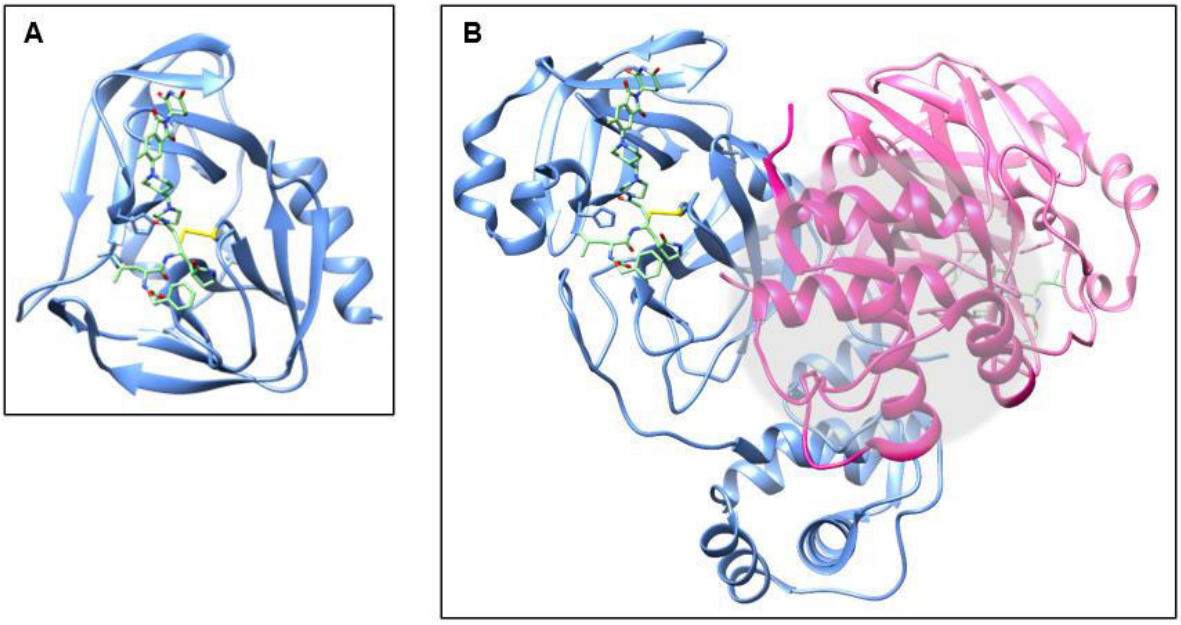


Figure 41. Structural model of monomeric CVB3 $3C^{Pro}$ in complex with GC-376 PROTAC (**A**) and crystallographic structure of the dimeric SARS-CoV-2 $3CL^{Pro}$ (PDB ID: 8OKC) (**B**). The PROTAC molecule is shown in green as sticks. The two protomers of SARS-CoV-2 $3CL^{Pro}$ are represented in light blue and pink, respectively. The C-terminal domain of the pink protomer, located in proximity to the PROTAC binding site of the blue protomer, is highlighted.

Overall, the achieved findings highlight PROTAC technology as a promising strategy, opening new avenues for the development of advanced antiviral therapies. In particular, the GC-376

PROTAC constitutes an auspicious scaffold for developing broad-spectrum antiviral degraders targeting both coronaviruses and enteroviruses. Future investigations should aim to address the design of alternative PROTAC architectures with optimized structural properties to better accommodate the dimeric arrangement of SARS-CoV-2 $3CL^{Pro}$. The structural optimization of the linker is widely recognized as a central determinant in the successful design of PROTAC molecules.

The design of PROTAC molecules typically begins with the selection of two key components: a ligand that binds the target protein and an E3 ligase recognition element. While these binding motifs can often be predetermined based on available structural and biochemical data, the optimal spatial separation between the two moieties, necessary to promote efficient ubiquitination and subsequent degradation, usually requires empirical determination on a case-by-case basis. Both the length and chemical composition of the linker have been shown to significantly affect critical physicochemical and structural properties of PROTACs, including molecular rigidity, hydrophobicity, and solubility [232]. As demonstrated in seminal works of Ciulli and co-workers [246,360,361], linker length, composition, and conformational flexibility profoundly influence ternary complex formation, cooperativity, and degradation efficiency. High-resolution structural studies from the Ciulli laboratory revealed that productive ternary complexes often result from specific protein–protein interfaces stabilized through geometrical complementarity conferred by the linker architecture, rather than from the intrinsic binary affinities of the individual ligand [246].

Further contributions from other leading groups, including those of Crews [193], Bradner [223], and Fischer [362], reinforced the notion that maintaining an optimal balance between linker rigidity and flexibility is essential for ternary complex stability. Excessive rigidity may hinder the conformational adaptability required to accommodate diverse spatial orientations of the E3 ligase and target protein, whereas excessive flexibility can introduce entropic penalties that weaken inter-protein interactions and reduce overall complex stability. Consequently, the rational design of PROTAC linkers must carefully calibrate these parameters to achieve a conformation conducive to productive ternary complex formation and efficient target degradation.

Combined with experimental validation through X-ray crystallography, solution NMR, and biophysical assays, computational modelling has become instrumental in elucidating linker-mediated structural dynamics and predicting productive geometries [363]. In particular, the use of the High Ambiguity Driven DOCKing (HADDOCK) platform could be a valuable tool for generating and refining structural models of multi-component assemblies [364]. HADDOCK proposes a hybrid method that can include a large variety of NMR and non-NMR experimental

information to guide the docking process [365], thereby allowing the generation of conformational ensembles that are both physically realistic and consistent with experimental observations.

In the context of this work, HADDOCK-based modelling could enable the exploration of possible orientations and contact interfaces between the viral protease (e.g., SARS-CoV-2 $3CL^{Pro}$) and the E3 ligase CRBN in the presence of the PROTAC molecule. The resulting models will provide valuable structural hypotheses to rationalize experimental findings, such as the absence of ternary complex formation observed in vitro, and to guide subsequent optimization of the PROTAC linker or attachment sites. In the context of viral proteases such as SARS-CoV-2 $3CL^{Pro}$, these structure-guided principles can inform the rational design of PROTACs capable of mitigating steric hindrance associated with dimeric assemblies, thereby enabling the formation of stable and degradation-competent ternary complexes.

Looking ahead, crystallization trials are currently underway to obtain high-resolution structural information on both the binary PROTAC–CRBN-midi complex and the corresponding protein:PROTAC:E3 ligase ternary assemblies. The crystallization setups are being designed by exploiting the molar ratios between target protein, PROTAC, and E3 ligase as reported in the literature and optimized through our experimental observations from analytical SEC and NMR studies. The eluted fractions corresponding to the ternary complexes, isolated via analytical SEC, will be directly employed in crystallization experiments using both soaking and co-crystallization strategies. Parallel optimization of crystallization conditions, such as buffer composition, pH, ionic strength, and additive screening, will be performed to enhance crystal quality and diffraction properties. These experiments aim to elucidate the precise molecular determinants governing ternary complex formation and stabilization, thereby providing direct structural evidence to complement the results described in this work.

Finally, cellular assays that evaluate target protein degradation will be essential to validate the applicability of our molecule. In this context, I can rely on the expertise in cellular and molecular biology acquired during my three-and-a-half-month internship experience in the laboratory of Dr. Golinelli at the Institut de Chimie des Substances Naturelles (ICSN), France. Specifically, I gained proficiency in cell culture techniques, including the growth and maintenance of human cell lines in flasks, as well as in biochemical assays, including protein extraction and quantification procedures from cellular samples, Western Blot, Immunofluorescence, MTT and Flow cytometry (FACS) experiments. These techniques have been employed to identify target proteins and evaluate the effects of known drugs or measure particular conditions in specific human cell lines.

The learned competencies will complement PROTAC in vitro studies by indirectly confirming ternary complex formation.

From a translational perspective, expanding the PROTAC concept beyond classical cellular targets to viral proteins represents a pioneering direction with potential implications for future next-generation antiviral degraders. The establishment of interdisciplinary networks, combining expertise in structural biology, medicinal chemistry, and computational modelling, will be essential to realize this vision. In a long-term prospective the refinement of computational and structural approaches will further enable the visualization and prediction of molecular assemblies, bridging the gap between *in silico* modelling and *in cellulo* activity.

Concerning the SARS-CoV-2 Nsp13, I optimized expression and purification conditions, obtaining ~0.2 mg/L of culture yield with zinc supplementation. Based on recent evidence that the ZBD coordinates a [Fe_4S_4] cluster [359], I supplemented $FeCl_3$ during expression, which increased the protein yield to ~0.5 mg/L of culture. As future prospects, thanks to the collaborative framework established within PNRR Spoke 4 initiative, which enables the integration of complementary expertise across different research groups (Prof. G. Caminati), ongoing and planned efforts will focus on elucidating the molecular basis of Nsp13 interactions with newly synthesized inhibitors. These studies aim to define the binding determinants and inhibitory mechanisms of potential drug candidates, thereby contributing to the broader understanding of SARS-CoV-2 replication machinery and informing the rational design of antiviral agents targeting viral helicases. Such an endeavour would not only deepen our understanding of host–pathogen molecular interplay but also contribute to the development of adaptable therapeutic strategies for future pandemics.

References

1. Gashaw, I.; Ellinghaus, P.; Sommer, A.; Asadullah, K. What makes a good drug target? *Drug Discov Today* **2012**, *17 Suppl*, S24-30.
2. Imming, P.; Sinning, C.; Meyer, A. Drugs, their targets and the nature and number of drug targets. *Nat Rev Drug Discov* **2006**, *5*, 821-834.
3. Hopkins, A.L.; Groom, C.R. The druggable genome. *Nat Rev Drug Discov* **2002**, *1*, 727-730.
4. Egner, U.; Hillig, R.C. A Structural View on Druggability. In *Structural Biology in Drug Discovery*; 2020; pp. 23-52.
5. Egner, U.; Krätzschmar, J.; Kreft, B.; Pohlenz, H.D.; Schneider, M. The target discovery process. *Chembiochem* **2005**, *6*, 468-479.
6. Egner, U.; Hillig, R.C. A structural biology view of target drugability. *Expert opinion on drug discovery* **2008**, *3*, 391-401.
7. Gao, Z.; Li, H.; Zhang, H.; Liu, X.; Kang, L.; Luo, X.; Zhu, W.; Chen, K.; Wang, X.; Jiang, H. PDTD: a web-accessible protein database for drug target identification. *BMC Bioinf.* **2008**, *9*, 104.
8. Henrich, S.; Salo-Ahen, O.M.; Huang, B.; Rippmann, F.F.; Cruciani, G.; Wade, R.C. Computational approaches to identifying and characterizing protein binding sites for ligand design. *J Mol Recognit* **2010**, *23*, 209-219.
9. Ferreira, L.G.; Dos Santos, R.N.; Oliva, G.; Andricopulo, A.D. Molecular docking and structure-based drug design strategies. *Molecules* **2015**, *20*, 13384-13421.
10. Anderson, A.C. The process of structure-based drug design. *Chem Biol* **2003**, *10*, 787-797.
11. Tuttle, K.R. Impact of the COVID-19 pandemic on clinical research. *Nat Rev Nephrol* **2020**, *16*, 562-564.
12. Cevik, M.; Bamford, C.G.G.; Ho, A. COVID-19 pandemic-a focused review for clinicians. *Clin Microbiol Infect* **2020**, *26*, 842-847.
13. Sohrabi, C.; Alsafi, Z.; O'Neill, N.; Khan, M.; Kerwan, A.; Al-Jabir, A.; Iosifidis, C.; Agha, R. World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *Int J Surg* **2020**, *76*, 71-76.
14. McCoy, J.; Wambier, C.G.; Vano-Galvan, S.; Shapiro, J.; Sinclair, R.; Ramos, P.M.; Washenik, K.; Andrade, M.; Herrera, S.; Goren, A. Racial variations in COVID-19 deaths may be due to androgen receptor genetic variants associated with prostate cancer and androgenetic alopecia. Are anti-androgens a potential treatment for COVID-19? *J Cosmet Dermatol* **2020**, *19*, 1542-1543.
15. Murray, M.F.; Kenny, E.E.; Ritchie, M.D.; Rader, D.J.; Bale, A.E.; Giovanni, M.A.; Abul-Husn, N.S. COVID-19 outcomes and the human genome. *Genet Med* **2020**, *22*, 1175-1177.
16. Ovsyannikova, I.G.; Haralambieva, I.H.; Crooke, S.N.; Poland, G.A.; Kennedy, R.B. The role of host genetics in the immune response to SARS-CoV-2 and COVID-19 susceptibility and severity. *Immunol Rev* **2020**, *296*, 205-219.
17. Agnello, S.; Brand, M.; Chellat, M.F.; Gazzola, S.; Riedl, R. A Structural View on Medicinal Chemistry Strategies against Drug Resistance. *Angew Chem Int Ed Engl* **2019**, *58*, 3300-3345.
18. Lionta, E.; Spyrou, G.; Vassilatis, D.K.; Cournia, Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr Top Med Chem* **2014**, *14*, 1923-1938.
19. Kalyaanamoorthy, S.; Chen, Y.P. Structure-based drug design to augment hit discovery. *Drug Discov Today* **2011**, *16*, 831-839.
20. Batool, M.; Ahmad, B.; Choi, S. A Structure-Based Drug Discovery Paradigm. *Int J Mol Sci* **2019**, *20*.
21. McIntyre, P.J.; Collins, P.M.; Vrzal, L.; Birchall, K.; Arnold, L.H.; Mpamhanga, C.; Coombs, P.J.; Burgess, S.G.; Richards, M.W.; Winter, A.; et al. Characterization of Three Druggable Hot-Spots in the Aurora-A/TPX2 Interaction Using Biochemical, Biophysical, and Fragment-Based Approaches. *ACS Chem Biol* **2017**, *12*, 2906-2914.
22. Kirsch, P.; Hartman, A.M.; Hirsch, A.K.H.; Empting, M. Concepts and Core Principles of Fragment-Based Drug Design. *Molecules* **2019**, *24*.
23. Smyth, M.S.; Martin, J.H. x ray crystallography. *Mol Pathol* **2000**, *53*, 8-14.
24. Feng, W.; Pan, L.; Zhang, M. Combination of NMR spectroscopy and X-ray crystallography offers unique advantages for elucidation of the structural basis of protein complex assembly. *Sci China Life Sci* **2011**, *54*, 101-111.
25. Müller, I. Guidelines for the successful generation of protein-ligand complex crystals. *Acta Crystallogr D Struct Biol* **2017**, *73*, 79-92.
26. Ehrmann, F.R.; Stojko, J.; Metz, A.; Debaene, F.; Barandun, L.J.; Heine, A.; Diederich, F.; Cianférani, S.; Reuter, K.; Klebe, G. Soaking suggests "alternative facts": Only co-crystallization discloses major ligand-induced interface rearrangements of a homodimeric tRNA-binding protein indicating a novel mode-of-inhibition. *PLoS One* **2017**, *12*, e0175723.
27. Wagner, G. An account of NMR in structural biology. *Nat Struct Biol* **1997**, *4 Suppl*, 841-844.
28. Fu, R.; Cross, T.A. Solid-state nuclear magnetic resonance investigation of protein and polypeptide structure. *Annu Rev Biophys Biomol Struct* **1999**, *28*, 253-268.
29. Wuthrich, K. Protein structure determination in solution by NMR spectroscopy. *J. Biol. Chem* **1990**, *265*, 22059-22062.
30. Cavanagh, J.; Fairbrother, W.J.; Palmer, A.G., III; Skelton, N.J. *Protein NMR Spectroscopy. Principles and practice*; Academic Press: San Diego, 1996.
31. Mittag, T.; Forman-Kay, J. Atomic-level characterization of disordered protein ensembles. *Curr. Opin. Struct. Biol* **2007**, *17*, 3-14.
32. Eliezer, D. Biophysical characterization of intrinsically disordered proteins. *Curr. Opin. Struct. Biol* **2009**, *19*, 23-30.
33. Betz, M.; Saxena, K.; Schwalbe, H. Biomolecular NMR: a chaperone to drug discovery. *Curr. Op. Chem. Biol.* **2006**, *10*, 219-225.
34. Schwalbe, H.; Stilz, H.U.; Kessler, H. NMR spectroscopy of biomacromolecules in drug discovery and beyond. *ChemBioChem* **2005**, *6*, 1475-1478.
35. Bonvin, A.M.; Boelens, R.; Kaptein, R. NMR analysis of protein interactions. *Curr. Opin. Chem. Biol* **2005**, *9*, 501-508.
36. Schmitz, C.; Bonvin, A.M. Protein-protein HADDocking using exclusively pseudocontact shifts. *J Biomol NMR* **2011**, *50*, 263-266.
37. Ludwig, C.; Guenther, U.L. Ligand based NMR methods for drug discovery. *Front Biosci (Landmark Ed)* **2009**, *14*, 4565-4574.
38. Kay, L.E. Protein dynamics from NMR. *Nat. Struct. Biol.* **1998**, *5*, 513-517.
39. Palmer, A.G., 3rd. Nmr probes of molecular dynamics: overview and comparison with other techniques. *Annu Rev Biophys Biomol Struct* **2001**, *30*, 129-155.
40. Case, D.A. Molecular dynamics and NMR spin relaxation in proteins. *Acc. Chem. Res.* **2002**, *35*, 325-331.
41. Korzhnev, D.M.; Kay, L.E. Probing invisible, low-populated States of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding. *Acc. Chem. Res* **2008**, *41*, 442-451.
42. Hiroaki, H. Recent applications of isotopic labeling for protein NMR in drug discovery. *Expert opinion on drug discovery* **2013**, *8*, 523-536.
43. Barile, E.; Pellecchia, M. NMR-based approaches for the identification and optimization of inhibitors of protein-protein interactions. *Chem. Rev.* **2014**, *114*, 4749-4763.
44. Orts, J.; Walti, M.A.; Marsh, M.; Vera, L.; Gossert, A.D.; Guntert, P.; Riek, R. NMR-Based Determination of the 3D Structure of the Ligand-Protein Interaction Site without Protein Resonance Assignment. *Journal of the American Chemical Society* **2016**, *138*, 4393-4400.

45. Walti, M.A.; Riek, R.; Orts, J. Fast NMR-Based Determination of the 3D Structure of the Binding Site of Protein-Ligand Complexes with Weak Affinity Binders. *Angewandte Chemie-International Edition* **2017**, *56*, 5208-5211.
46. Pellecchia, M.; Sem, D.S.; Wüthrich, K. NMR in drug discovery. *Nat Rev Drug Discov* **2002**, *1*, 211-219.
47. Sugiki, T.; Fujiwara, T.; Kojima, C. Latest approaches for efficient protein production in drug discovery. *Expert opinion on drug discovery* **2014**, *9*, 1189-1204.
48. Sugiki, T.; Furuita, K.; Fujiwara, T.; Kojima, C. Current NMR Techniques for Structure-Based Drug Discovery. *Molecules* **2018**, *23*.
49. Kühlbrandt, W. Biochemistry. The resolution revolution. *Science* **2014**, *343*, 1443-1444.
50. Ceska, T.; Chung, C.W.; Cooke, R.; Phillips, C.; Williams, P.A. Cryo-EM in drug discovery. *Biochem Soc Trans* **2019**, *47*, 281-293.
51. Renaud, J.P.; Chari, A.; Ciferri, C.; Liu, W.T.; Rémigy, H.W.; Stark, H.; Wiesmann, C. Cryo-EM in drug discovery: achievements, limitations and prospects. *Nat Rev Drug Discov* **2018**, *17*, 471-492.
52. Scheres, S.H. A Bayesian view on cryo-EM structure determination. *J Mol Biol* **2012**, *415*, 406-418.
53. Khoshouei, M.; Radjainia, M.; Baumeister, W.; Danev, R. Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate. *Nat Commun* **2017**, *8*, 16099.
54. Merk, A.; Bartesaghi, A.; Banerjee, S.; Falconieri, V.; Rao, P.; Davis, M.I.; Pragani, R.; Boxer, M.B.; Earl, L.A.; Milne, J.L.S.; et al. Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell* **2016**, *165*, 1698-1707.
55. Bajad, N.G.; Rayala, S.; Gutti, G.; Sharma, A.; Singh, M.; Kumar, A.; Singh, S.K. Systematic review on role of structure based drug design (SBDD) in the identification of anti-viral leads against SARS-Cov-2. *Curr Res Pharmacol Drug Discov* **2021**, *2*, 100026.
56. Lu, L.; Su, S.; Yang, H.; Jiang, S. Antivirals with common targets against highly pathogenic viruses. *Cell* **2021**, *184*, 1604-1620.
57. Ksiazek, T.G.; Erdman, D.; Goldsmith, C.S.; Zaki, S.R.; Peret, T.; Emery, S.; Tong, S.; Urbani, C.; Comer, J.A.; Lim, W.; et al. A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* **2003**, *348*, 1953-1966.
58. Drosten, C.; Günther, S.; Preiser, W.; van der Werf, S.; Brodt, H.R.; Becker, S.; Rabenau, H.; Panning, M.; Kolesnikova, L.; Fouchier, R.A.; et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* **2003**, *348*, 1967-1976.
59. Kuiken, T.; Fouchier, R.A.; Schutten, M.; Rimmelzwaan, G.F.; van Amerongen, G.; van Riel, D.; Laman, J.D.; de Jong, T.; van Doornum, G.; Lim, W.; et al. Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. *Lancet* **2003**, *362*, 263-270.
60. Zaki, A.M.; van Boheemen, S.; Bestebroer, T.M.; Osterhaus, A.D.; Fouchier, R.A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* **2012**, *367*, 1814-1820.
61. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **2020**, *382*, 727-733.
62. Zou, L.; Ruan, F.; Huang, M.; Liang, L.; Huang, H.; Hong, Z.; Yu, J.; Kang, M.; Song, Y.; Xia, J.; et al. SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. *N Engl J Med* **2020**, *382*, 1177-1179.
63. Wang, D.; Hu, B.; Hu, C.; Zhu, F.; Liu, X.; Zhang, J.; Wang, B.; Xiang, H.; Cheng, Z.; Xiong, Y.; et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA* **2020**, *323*, 1061-1069.
64. Xu, Z.; Shi, L.; Wang, Y.; Zhang, J.; Huang, L.; Zhang, C.; Liu, S.; Zhao, P.; Liu, H.; Zhu, L.; et al. Pathological findings of COVID-19 associated with acute respiratory distress syndrome. *Lancet Respir Med* **2020**, *8*, 420-422.
65. Wu, A.; Peng, Y.; Huang, B.; Ding, X.; Wang, X.; Niu, P.; Meng, J.; Zhu, Z.; Zhang, Z.; Wang, J.; et al. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host Microbe* **2020**, *27*, 325-328.
66. Naqvi, A.A.T.; Fatima, K.; Mohammad, T.; Fatima, U.; Singh, I.K.; Singh, A.; Atif, S.M.; Hariprasad, G.; Hasan, G.M.; Hassan, M.I. Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. *Biochim Biophys Acta Mol Basis Dis* **2020**, *1866*, 165878.
67. Brant, A.C.; Tian, W.; Majerciak, V.; Yang, W.; Zheng, Z.M. SARS-CoV-2: from its discovery to genome structure, transcription, and replication. *Cell & bioscience* **2021**, *11*, 136.
68. Mandala, V.S.; McKay, M.J.; Shcherbakov, A.A.; Dregni, A.J.; Kolocouris, A.; Hong, M. Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nat Struct Mol Biol* **2020**, *27*, 1202-1208.
69. Padhan, K.; Parvez, M.K.; Al-Dosari, M.S. Comparative sequence analysis of SARS-CoV-2 suggests its high transmissibility and pathogenicity. *Future Virol* **2021**.
70. Moustaqil, M.; Ollivier, E.; Chiu, H.P.; Van Tol, S.; Rudolffi-Soto, P.; Stevens, C.; Bhumkar, A.; Hunter, D.J.B.; Freiberg, A.N.; Jacques, D.; et al. SARS-CoV-2 proteases PLpro and 3CLpro cleave IRF3 and critical modulators of inflammatory pathways (NLRP12 and TAB1): implications for disease presentation across species. *Emerg Microbes Infect* **2021**, *10*, 178-195.
71. Zhu, J.; Zhang, H.; Lin, Q.; Lyu, J.; Lu, L.; Chen, H.; Zhang, X.; Zhang, Y.; Chen, K. Progress on SARS-CoV-2 3CLpro Inhibitors: Inspiration from SARS-CoV 3CLpro Peptidomimetics and Small-Molecule Anti-Inflammatory Compounds. *Drug Des Devel Ther* **2022**, *16*, 1067-1082.
72. Roe, M.K.; Junod, N.A.; Young, A.R.; Beachboard, D.C.; Stobart, C.C. Targeting novel structural and functional features of coronavirus protease nsp5 (3CL(pro), M(pro)) in the age of COVID-19. *J Gen Virol* **2021**, *102*.
73. Gahlawat, A.; Kumar, N.; Kumar, R.; Sandhu, H.; Singh, I.P.; Singh, S.; Sjöstedt, A.; Garg, P. Structure-Based Virtual Screening to Discover Potential Lead Molecules for the SARS-CoV-2 Main Protease. *J Chem Inf Model* **2020**, *60*, 5781-5793.
74. Graziano, V.; McGrath, W.J.; Yang, L.; Mangel, W.F. SARS CoV main proteinase: The monomer-dimer equilibrium dissociation constant. *Biochemistry* **2006**, *45*, 14632-14641.
75. Hsu, M.F.; Kuo, C.J.; Chang, K.T.; Chang, H.C.; Chou, C.C.; Ko, T.P.; Shr, H.L.; Chang, G.G.; Wang, A.H.; Liang, P.H. Mechanism of the maturation process of SARS-CoV 3CL protease. *J Biol Chem* **2005**, *280*, 31257-31266.
76. Chen, S.; Chen, L.; Tan, J.; Chen, J.; Du, L.; Sun, T.; Shen, J.; Chen, K.; Jiang, H.; Shen, X. Severe acute respiratory syndrome coronavirus 3C-like proteinase N terminus is indispensable for proteolytic activity but not for enzyme dimerization. Biochemical and thermodynamic investigation in conjunction with molecular dynamics simulations. *J Biol Chem* **2005**, *280*, 164-173.
77. Cheng, S.C.; Chang, G.G.; Chou, C.Y. Mutation of Glu-166 blocks the substrate-induced dimerization of SARS coronavirus main protease. *Biophys J* **2010**, *98*, 1327-1336.
78. Grottesi, A.; Bešker, N.; Emerson, A.; Manelfi, C.; Beccari, A.R.; Frigerio, F.; Lindahl, E.; Cerchia, C.; Talarico, C. Computational Studies of SARS-CoV-2 3CLpro: Insights from MD Simulations. *Int J Mol Sci* **2020**, *21*.
79. Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* **2020**, *368*, 409-412.
80. Goyal, B.; Goyal, D. Targeting the Dimerization of the Main Protease of Coronaviruses: A Potential Broad-Spectrum Therapeutic Strategy. *ACS combinatorial science* **2020**, *22*, 297-305.
81. Su, H.; Zhou, F.; Huang, Z.; Ma, X.; Natarajan, K.; Zhang, M.; Huang, Y.; Su, H. Molecular Insights into Small-Molecule Drug Discovery for SARS-CoV-2. *Angew Chem Int Ed Engl* **2021**, *60*, 9789-9802.
82. Razali, R.; Asis, H.; Budiman, C. Structure-Function Characteristics of SARS-CoV-2 Proteases and Their Potential Inhibitors from Microbial Sources. *Microorganisms* **2021**, *9*.

83. Schechter, I.; Berger, A. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **1967**, *27*, 157-162.
84. Dai, W.; Zhang, B.; Jiang, X.M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F.; et al. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **2020**, *368*, 1331-1335.
85. Yang, H.; Yang, M.; Ding, Y.; Liu, Y.; Lou, Z.; Zhou, Z.; Sun, L.; Mo, L.; Ye, S.; Pang, H.; et al. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc Natl Acad Sci U S A* **2003**, *100*, 13190-13195.
86. Ziebuhr, J.; Snijder, E.J.; Gorbalenya, A.E. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J Gen Virol* **2000**, *81*, 853-879.
87. Anand, K.; Palm, G.J.; Mesters, J.R.; Siddell, S.G.; Ziebuhr, J.; Hilgenfeld, R. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *Embo j* **2002**, *21*, 3213-3224.
88. Xiong, M.; Su, H.; Zhao, W.; Xie, H.; Shao, Q.; Xu, Y. What coronavirus 3C-like protease tells us: From structure, substrate selectivity, to inhibitor design. *Medicinal research reviews* **2021**, *41*, 1965-1998.
89. Xue, X.; Yu, H.; Yang, H.; Xue, F.; Wu, Z.; Shen, W.; Li, J.; Zhou, Z.; Ding, Y.; Zhao, Q.; et al. Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design. *J Virol* **2008**, *82*, 2515-2527.
90. Chen, C.C.; Yu, X.; Kuo, C.J.; Min, J.; Chen, S.; Ma, L.; Liu, K.; Guo, R.T. Overview of antiviral drug candidates targeting coronaviral 3C-like main proteases. *FEBS J* **2021**, *288*, 5089-5121.
91. Ferreira, J.C.; Fadl, S.; Villanueva, A.J.; Rabeh, W.M. Catalytic Dyad Residues His41 and Cys145 Impact the Catalytic Activity and Overall Conformational Fold of the Main SARS-CoV-2 Protease 3-Chymotrypsin-Like Protease. *Frontiers in chemistry* **2021**, *9*, 692168.
92. Kneller, D.W.; Phillips, G.; Weiss, K.L.; Pant, S.; Zhang, Q.; O'Neill, H.M.; Coates, L.; Kovalevsky, A. Unusual zwitterionic catalytic site of SARS-CoV-2 main protease revealed by neutron crystallography. *J Biol Chem* **2020**, *295*, 17365-17373.
93. Akaji, K.; Konno, H. Design and Evaluation of Anti-SARS-Coronavirus Agents Based on Molecular Interactions with the Viral Protease. *Molecules* **2020**, *25*.
94. Zhang, S.; Krumberger, M.; Morris, M.A.; Parrocha, C.M.T.; Kreutzer, A.G.; Nowick, J.S. Structure-based drug design of an inhibitor of the SARS-CoV-2 (COVID-19) main protease using free software: A tutorial for students and scientists. *European journal of medicinal chemistry* **2021**, *218*, 113390.
95. Wang, H.; He, S.; Deng, W.; Zhang, Y.; Li, G.; Sun, J.; Zhao, W.; Guo, Y.; Yin, Z.; Li, D.; et al. Comprehensive Insights into the Catalytic Mechanism of Middle East Respiratory Syndrome 3C-Like Protease and Severe Acute Respiratory Syndrome 3C-Like Protease. *ACS Catal* **2020**, *10*, 5871-5890.
96. Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J.R.; Hilgenfeld, R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **2003**, *300*, 1763-1767.
97. Zhang, L.; Lin, D.; Kusov, Y.; Nian, Y.; Ma, Q.; Wang, J.; von Brunn, A.; Leyssen, P.; Lanko, K.; Neyts, J.; et al. α-Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J Med Chem* **2020**, *63*, 4562-4578.
98. Martino, T.A.; Liu, P.; Sole, M.J. Viral infection and the pathogenesis of dilated cardiomyopathy. *Circ Res* **1994**, *74*, 182-188.
99. Maier, R.; Krebs, P.; Ludewig, B. Immunopathological basis of virus-induced myocarditis. *Clinical & developmental immunology* **2004**, *11*, 1-5.
100. Gauntt, C.; Huber, S. Coxsackievirus experimental heart diseases. *Front Biosci* **2003**, *8*, e23-35.
101. Lugo, D.; Krogstad, P. Enteroviruses in the early 21st century: new manifestations and challenges. *Current opinion in pediatrics* **2016**, *28*, 107-113.
102. Cheung, P.K.; Yuan, J.; Zhang, H.M.; Chau, D.; Yanagawa, B.; Suarez, A.; McManus, B.; Yang, D. Specific interactions of mouse organ proteins with the 5'untranslated region of coxsackievirus B3: potential determinants of viral tissue tropism. *J Med Virol* **2005**, *77*, 414-424.
103. Huber, S.A.; Gauntt, C.J.; Sakkinen, P. Enteroviruses and myocarditis: viral pathogenesis through replication, cytokine induction, and immunopathogenicity. *Adv Virus Res* **1998**, *51*, 35-80.
104. Garmaroudi, F.S.; Marchant, D.; Hendry, R.; Luo, H.; Yang, D.; Ye, X.; Shi, J.; McManus, B.M. Coxsackievirus B3 replication and pathogenesis. *Future Microbiol* **2015**, *10*, 629-653.
105. Peischard, S.; Ho, H.T.; Theiss, C.; Strutz-Seebohm, N.; Seebohm, G. A Kidnapping Story: How Coxsackievirus B3 and Its Host Cell Interact. *Cell Physiol Biochem* **2019**, *53*, 121-140.
106. Selinka, H.C.; Wolde, A.; Sauter, M.; Kandolf, R.; Klingel, K. Virus-receptor interactions of coxsackie B viruses and their putative influence on cardiotropism. *Med Microbiol Immunol* **2004**, *193*, 127-131.
107. Lee, E.S.; Lee, W.G.; Yun, S.H.; Rho, S.H.; Im, I.; Yang, S.T.; Sellamuthu, S.; Lee, Y.J.; Kwon, S.J.; Park, O.K.; et al. Development of potent inhibitors of the coxsackievirus 3C protease. *Biochem Biophys Res Commun* **2007**, *358*, 7-11.
108. Binford, S.L.; Maldonado, F.; Brothers, M.A.; Weady, P.T.; Zalman, L.S.; Meador, J.W., 3rd; Matthews, D.A.; Patick, A.K. Conservation of amino acids in human rhinovirus 3C protease correlates with broad-spectrum antiviral activity of rupintrivir, a novel human rhinovirus 3C protease inhibitor. *Antimicrob Agents Chemother* **2005**, *49*, 619-626.
109. Kim, Y.; Lovell, S.; Tiew, K.C.; Mandadapu, S.R.; Alliston, K.R.; Battaile, K.P.; Groutas, W.C.; Chang, K.O. Broad-spectrum antivirals against 3C or 3C-like proteases of picornaviruses, noroviruses, and coronaviruses. *J Virol* **2012**, *86*, 11754-11762.
110. Bazan, J.F.; Fletterick, R.J. Comparative analysis of viral cysteine protease structural models. *FEBS Lett* **1989**, *249*, 5-7.
111. Matthews, D.A.; Dragovich, P.S.; Webber, S.E.; Fuhrman, S.A.; Patick, A.K.; Zalman, L.S.; Hendrickson, T.F.; Love, R.A.; Prins, T.J.; Marakovits, J.T.; et al. Structure-assisted design of mechanism-based irreversible inhibitors of human rhinovirus 3C protease with potent antiviral activity against multiple rhinovirus serotypes. *Proc Natl Acad Sci U S A* **1999**, *96*, 11000-11007.
112. Lee, C.C.; Kuo, C.J.; Ko, T.P.; Hsu, M.F.; Tsui, Y.C.; Chang, S.C.; Yang, S.; Chen, S.J.; Chen, H.C.; Hsu, M.C.; et al. Structural basis of inhibition specificities of 3C and 3C-like proteases by zinc-coordinating and peptidomimetic compounds. *J Biol Chem* **2009**, *284*, 7646-7655.
113. Shi, J.; Song, J. The catalysis of the SARS 3C-like protease is under extensive regulation by its extra domain. *FEBS J* **2006**, *273*, 1035-1045.
114. Göhl, M.; Zhang, L.; El Kilani, H.; Sun, X.; Zhang, K.; Brönstrup, M.; Hilgenfeld, R. From Repurposing to Redesign: Optimization of Boceprevir to Highly Potent Inhibitors of the SARS-CoV-2 Main Protease. *Molecules* **2022**, *27*.
115. Rani, I.; Goyal, A.; Bhatnagar, M.; Manhas, S.; Goel, P.; Pal, A.; Prasad, R. Potential molecular mechanisms of zinc- and copper-mediated antiviral activity on COVID-19. *Nutrition Research* **2021**, *92*, 109-128.
116. Gil-Moles, M.; Basu, U.; Büssing, R.; Hoffmeister, H.; Türck, S.; Varchmin, A.; Ott, I. Gold Metallodrugs to Target Coronavirus Proteins: Inhibitory Effects on the Spike-ACE2 Interaction and on PLpro Protease Activity by Auranofin and Gold Organometallics*. *Chemistry* **2020**, *26*, 15140-15144.
117. Abate, C.; Carnamucio, F.; Giuffrè, O.; Foti, C. Metal-Based Compounds in Antiviral Therapy. *Biomolecules* **2022**, *12*.
118. He, Z.; Ye, F.; Zhang, C.; Fan, J.; Du, Z.; Zhao, W.; Yuan, Q.; Niu, W.; Gao, F.; He, B.; et al. A comparison of Remdesivir versus gold cluster in COVID-19 animal model: A better therapeutic outcome of gold cluster. *Nano today* **2022**, *44*, 101468.

119. Karges, J.; Cohen, S.M. Metal Complexes as Antiviral Agents for SARS-CoV-2. *Chembiochem* **2021**, *22*, 2600-2607.
120. Grifagni, D.; Calderone, V.; Giuntini, S.; Cantini, F.; Fragai, M.; Banci, L. SARS-CoV-2 M(pro) inhibition by a zinc ion: structural features and hints for drug design. *Chemical communications (Cambridge, England)* **2021**, *57*, 7910-7913.
121. Panchariya, L.; Khan, W.A.; Kuila, S.; Sonkar, K.; Sahoo, S.; Ghoshal, A.; Kumar, A.; Verma, D.K.; Hasan, A.; Khan, M.A.; et al. Zinc(2+) ion inhibits SARS-CoV-2 main protease and viral replication in vitro. *Chemical communications (Cambridge, England)* **2021**, *57*, 10083-10086.
122. Tao, X.; Zhang, L.; Du, L.; Lu, K.; Zhao, Z.; Xie, Y.; Li, X.; Huang, S.; Wang, P.H.; Pan, J.A.; et al. Inhibition of SARS-CoV-2 replication by zinc gluconate in combination with hinokitiol. *J Inorg Biochem* **2022**, *231*, 111777.
123. Mondal, D.; Warshel, A. Exploring the Mechanism of Covalent Inhibition: Simulating the Binding Free Energy of α-Ketoamide Inhibitors of the Main Protease of SARS-CoV-2. *Biochemistry* **2020**, *59*, 4601-4608.
124. Yang, H.; Xie, W.; Xue, X.; Yang, K.; Ma, J.; Liang, W.; Zhao, Q.; Zhou, Z.; Pei, D.; Ziebuhr, J.; et al. Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol* **2005**, *3*, e324.
125. Patick, A.K.; Binford, S.L.; Brothers, M.A.; Jackson, R.L.; Ford, C.E.; Diem, M.D.; Maldonado, F.; Dragovich, P.S.; Zhou, R.; Prins, T.J.; et al. In vitro antiviral activity of AG7088, a potent inhibitor of human rhinovirus 3C protease. *Antimicrob Agents Chemother* **1999**, *43*, 2444-2450.
126. Shie, J.J.; Fang, J.M.; Kuo, T.H.; Kuo, C.J.; Liang, P.H.; Huang, H.J.; Wu, Y.T.; Jan, J.T.; Cheng, Y.S.; Wong, C.H. Inhibition of the severe acute respiratory syndrome 3CL protease by peptidomimetic alpha,beta-unsaturated esters. *Bioorg Med Chem* **2005**, *13*, 5240-5252.
127. Vatansever, E.C.; Yang, K.; Kratch, K.C.; Drelich, A.; Cho, C.-C.; Mellot, D.M.; Xu, S.; Tseng, C.-T.K.; Liu, W.R. Targeting the SARS-CoV-2 Main Protease to Repurpose Drugs for COVID-19. *bioRxiv : the preprint server for biology* **2020**, 2020.2005.2023.112235.
128. Ramos-Guzmán, C.A.; Ruiz-Pernía, J.J.; Tuñón, I. Multiscale Simulations of SARS-CoV-2 3CL Protease Inhibition with Aldehyde Derivatives. Role of Protein and Inhibitor Conformational Changes in the Reaction Mechanism. *ACS Catal* **2021**, *11*, 4157-4168.
129. Qiao, J.; Li, Y.S.; Zeng, R.; Liu, F.L.; Luo, R.H.; Huang, C.; Wang, Y.F.; Zhang, J.; Quan, B.; Shen, C.; et al. SARS-CoV-2 M(pro) inhibitors with antiviral activity in a transgenic mouse model. *Science* **2021**, *371*, 1374-1378.
130. Xia, Z.; Sacco, M.; Hu, Y.; Ma, C.; Meng, X.; Zhang, F.; Szeto, T.; Xiang, Y.; Chen, Y.; Wang, J. Rational Design of Hybrid SARS-CoV-2 Main Protease Inhibitors Guided by the Superimposed Cocrystal Structures with the Peptidomimetic Inhibitors GC-376, Telaprevir, and Boceprevir. *ACS Pharmacol Transl Sci* **2021**, *4*, 1408-1421.
131. Zhu, L.; George, S.; Schmidt, M.F.; Al-Gharabli, S.I.; Rademann, J.; Hilgenfeld, R. Peptide aldehyde inhibitors challenge the substrate specificity of the SARS-coronavirus main protease. *Antiviral research* **2011**, *92*, 204-212.
132. Lu, J.; Chen, S.A.; Khan, M.B.; Brassard, R.; Arutyunova, E.; Lamer, T.; Vuong, W.; Fischer, C.; Young, H.S.; Vederas, J.C.; et al. Crystallization of Feline Coronavirus M(pro) With GC376 Reveals Mechanism of Inhibition. *Frontiers in chemistry* **2022**, *10*, 852210.
133. Dampalla, C.S.; Zheng, J.; Perera, K.D.; Wong, L.R.; Meyerholz, D.K.; Nguyen, H.N.; Kashipathy, M.M.; Battaile, K.P.; Lovell, S.; Kim, Y.; et al. Postinfection treatment with a protease inhibitor increases survival of mice with a fatal SARS-CoV-2 infection. *Proc Natl Acad Sci U S A* **2021**, *118*.
134. Shi, Y.; Shuai, L.; Wen, Z.; Wang, C.; Yan, Y.; Jiao, Z.; Guo, F.; Fu, Z.F.; Chen, H.; Bu, Z.; et al. The preclinical inhibitor GS441524 in combination with GC376 efficaciously inhibited the proliferation of SARS-CoV-2 in the mouse respiratory tract. *Emerg Microbes Infect* **2021**, *10*, 481-492.
135. Cheng, S.; Feng, Y.; Li, W.; Liu, T.; Lv, X.; Tong, X.; Xi, G.; Ye, X.; Li, X. Development of novel antivrial agents that induce the degradation of the main protease of human-infecting coronaviruses. *European journal of medicinal chemistry* **2024**, *275*, 116629.
136. Owen, D.R.; Allerton, C.M.N.; Anderson, A.S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R.D.; Carlo, A.; Coffman, K.J.; et al. An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19. *Science* **2021**, *374*, 1586-1593.
137. Chen, P.; Van Oers, T.J.; Arutyunova, E.; Fischer, C.; Wang, C.; Lamer, T.; van Belkum, M.J.; Young, H.S.; Vederas, J.C.; Lemieux, M.J. A Structural Comparison of Oral SARS-CoV-2 Drug Candidate Ibuzatrelvir Complexed with the Main Protease (M(pro)) of SARS-CoV-2 and MERS-CoV. *JACS Au* **2024**, *4*, 3217-3227.
138. Brewitz, L.; Dumjahn, L.; Zhao, Y.; Owen, C.D.; Laidlaw, S.M.; Malla, T.R.; Nguyen, D.; Lukacik, P.; Salah, E.; Crawshaw, A.D.; et al. Alkyne Derivatives of SARS-CoV-2 Main Protease Inhibitors Including Nirmatrelvir Inhibit by Reacting Covalently with the Nucleophilic Cysteine. *J Med Chem* **2023**, *66*, 2663-2680.
139. Luh, L.M.; Scheib, U.; Juenemann, K.; Wortmann, L.; Brands, M.; Cromm, P.M. Prey for the Proteasome: Targeted Protein Degradation-A Medicinal Chemist's Perspective. *Angew Chem Int Ed Engl* **2020**, *59*, 15448-15466.
140. Tong, S.; Moyo, B.; Lee, C.M.; Leong, K.; Bao, G. Engineered materials for in vivo delivery of genome-editing machinery. *Nat Rev Mater* **2019**, *4*, 726-737.
141. de Smidt, P.C.; Le Doan, T.; de Falco, S.; van Berkel, T.J. Association of antisense oligonucleotides with lipoproteins prolongs the plasma half-life and modifies the tissue distribution. *Nucleic Acids Res.* **1991**, *19*, 4695-4700.
142. Geary, R.S.; Watanabe, T.A.; Truong, L.; Freier, S.; Lesnik, E.A.; Sioufi, N.B.; Sasmor, H.; Manoharan, M.; Levin, A.A. Pharmacokinetic properties of 2'-O-(2-methoxyethyl)-modified oligonucleotide analogs in rats. *J Pharmacol Exp Ther* **2001**, *296*, 890-897.
143. McMahon, B.M.; Mays, D.; Lipsky, J.; Stewart, J.A.; Fauq, A.; Richelson, E. Pharmacokinetics and tissue distribution of a peptide nucleic acid after intravenous administration. *Antisense Nucleic Acid Drug Dev* **2002**, *12*, 65-70.
144. Krieg, A.M. CpG motifs in bacterial DNA and their immune effects. *Annu Rev Immunol* **2002**, *20*, 709-760.
145. Fedorov, Y.; Anderson, E.M.; Birmingham, A.; Reynolds, A.; Karpilow, J.; Robinson, K.; Leake, D.; Marshall, W.S.; Khvorova, A. Off-target effects by siRNA can induce toxic phenotype. *RNA* **2006**, *12*, 1188-1196.
146. Jackson, A.L.; Bartz, S.R.; Schelter, J.; Kobayashi, S.V.; Burchard, J.; Mao, M.; Li, B.; Cavet, G.; Linsley, P.S. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* **2003**, *21*, 635-637.
147. Qiu, S.; Adema, C.M.; Lane, T. A computational study of off-target effects of RNA interference. *Nucleic Acids Res.* **2005**, *33*, 1834-1847.
148. Burnett, J.C.; Rossi, J.J. RNA-based therapeutics: current progress and future prospects. *Chem Biol* **2012**, *19*, 60-71.
149. Bennett, C.F.; Swayze, E.E. RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu Rev Pharmacol Toxicol* **2010**, *50*, 259-293.
150. Burslem, G.M.; Crews, C.M. Small-Molecule Modulation of Protein Homeostasis. *Chem. Rev.* **2017**, *117*, 11269-11301.
151. Lai, A.C.; Crews, C.M. Induced protein degradation: an emerging drug discovery paradigm. *Nat Rev Drug Discov* **2017**, *16*, 101-114.
152. Gong, L.; Cui, D.; Xiong, X.; Zhao, Y. Targeting Cullin-RING Ubiquitin Ligases and the Applications in PROTACs. *Adv Exp Med Biol* **2020**, *1217*, 317-347.
153. Adjei, A.A. What is the right dose? The elusive optimal biologic dose in phase I clinical trials. *J Clin Oncol* **2006**, *24*, 4054-4055.
154. Dang, C.V.; Reddy, E.P.; Shokat, K.M.; Soucek, L. Drugging the 'undruggable' cancer targets. *Nat Rev Cancer* **2017**, *17*, 502-508.

155. Verdine, G.L.; Walensky, L.D. The challenge of drugging undruggable targets in cancer: lessons learned from targeting BCL-2 family members. *Clin Cancer Res* **2007**, *13*, 7264-7270.
156. Crews, C.M. Targeting the undruggable proteome: the small molecules of my dreams. *Chem Biol* **2010**, *17*, 551-555.
157. Overington, J.P.; Al-Lazikani, B.; Hopkins, A.L. How many drug targets are there? *Nat Rev Drug Discov* **2006**, *5*, 993-996.
158. Moulick, K.; Ahn, J.H.; Zong, H.; Rodina, A.; Cerchietti, L.; Gomes DaGama, E.M.; Caldas-Lopes, E.; Beebe, K.; Perna, F.; Hatzi, K.; et al. Affinity-based proteomics reveal cancer-specific networks coordinated by Hsp90. *Nat Chem Biol* **2011**, *7*, 818-826.
159. Kamal, A.; Thao, L.; Sensintaffar, J.; Zhang, L.; Boehm, M.F.; Fritz, L.C.; Burrows, F.J. A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors. *Nature* **2003**, *425*, 407-410.
160. Dauvois, S.; Danielian, P.S.; White, R.; Parker, M.G. Antiestrogen ICI 164,384 reduces cellular estrogen receptor content by increasing its turnover. *Proc Natl Acad Sci U S A* **1992**, *89*, 4037-4041.
161. Cromm, P.M.; Crews, C.M. Targeted Protein Degradation: from Chemical Biology to Drug Discovery. *Cell Chem Biol* **2017**, *24*, 1181-1190.
162. Burslem, G.M.; Crews, C.M. Proteolysis-Targeting Chimeras as Therapeutics and Tools for Biological Discovery. *Cell* **2020**, *181*, 102-114.
163. Paiva, S.L.; Crews, C.M. Targeted protein degradation: elements of PROTAC design. *Curr Opin Chem Biol* **2019**, *50*, 111-119.
164. Maniaci, C.; Ciulli, A. Bifunctional chemical probes inducing protein-protein interactions. *Curr Opin Chem Biol* **2019**, *52*, 145-156.
165. Gustafson, J.L.; Neklesa, T.K.; Cox, C.S.; Roth, A.G.; Buckley, D.L.; Tae, H.S.; Sundberg, T.B.; Stagg, D.B.; Hines, J.; McDonnell, D.P.; et al. Small-Molecule-Mediated Degradation of the Androgen Receptor through Hydrophobic Tagging. *Angew Chem Int Ed Engl* **2015**, *54*, 9659-9662.
166. Xie, T.; Lim, S.M.; Westover, K.D.; Dodge, M.E.; Ercan, D.; Ficarro, S.B.; Udayakumar, D.; Gurbani, D.; Tae, H.S.; Riddle, S.M.; et al. Pharmacological targeting of the pseudokinase Her3. *Nat Chem Biol* **2014**, *10*, 1006-1012.
167. Hershko, A. Ubiquitin: roles in protein modification and breakdown. *Cell* **1983**, *34*, 11-12.
168. Finley, D. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu Rev Biochem* **2009**, *78*, 477-513.
169. Hershko, A.; Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem* **1998**, *67*, 425-479.
170. Kwon, Y.T.; Ciechanover, A. The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. *Trends Biochem Sci* **2017**, *42*, 873-886.
171. Tan, J.M.; Wong, E.S.; Kirkpatrick, D.S.; Pletnikova, O.; Ko, H.S.; Tay, S.P.; Ho, M.W.; Troncoso, J.; Gygi, S.P.; Lee, M.K.; et al. Lysine 63-linked ubiquitination promotes the formation and autophagic clearance of protein inclusions associated with neurodegenerative diseases. *Hum Mol Genet* **2008**, *17*, 431-439.
172. Glickman, M.H.; Ciechanover, A. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol Rev* **2002**, *82*, 373-428.
173. Olzmann, J.A.; Chin, L.S. Parkin-mediated K63-linked polyubiquitination: a signal for targeting misfolded proteins to the aggresome-autophagy pathway. *Autophagy* **2008**, *4*, 85-87.
174. Groll, M.; Huber, R. Substrate access and processing by the 20S proteasome core particle. *Int J Biochem Cell Biol* **2003**, *35*, 606-616.
175. Ravid, T.; Hochstrasser, M. Diversity of degradation signals in the ubiquitin-proteasome system. *Nat Rev Mol Cell Biol* **2008**, *9*, 679-690.
176. Klionsky, D.J.; Codogno, P. The mechanism and physiological function of macroautophagy. *J Innate Immun* **2013**, *5*, 427-433.
177. Anding, A.L.; Baehrecke, E.H. Cleaning House: Selective Autophagy of Organelles. *Dev Cell* **2017**, *41*, 10-22.
178. Lin, J.; Jin, J.; Shen, Y.; Zhang, L.; Gong, G.; Bian, H.; Chen, H.; Nagle, D.G.; Wu, Y.; Zhang, W.; et al. Emerging protein degradation strategies: expanding the scope to extracellular and membrane proteins. *Theranostics* **2021**, *11*, 8337-8349.
179. Miao, Y.; Gao, Q.; Mao, M.; Zhang, C.; Yang, L.; Yang, Y.; Han, D. Bispecific Aptamer Chimeras Enable Targeted Protein Degradation on Cell Membranes. *Angew Chem Int Ed Engl* **2021**, *60*, 11267-11271.
180. Ji, C.H.; Kim, H.Y.; Lee, M.J.; Heo, A.J.; Park, D.Y.; Lim, S.; Shin, S.; Ganipisetti, S.; Yang, W.S.; Jung, C.A.; et al. The AUTOTAC chemical biology platform for targeted protein degradation via the autophagy-lysosome system. *Nat Commun* **2022**, *13*, 904.
181. Qi, S.M.; Dong, J.; Xu, Z.Y.; Cheng, X.D.; Zhang, W.D.; Qin, J.J. PROTAC: An Effective Targeted Protein Degradation Strategy for Cancer Therapy. *Front Pharmacol* **2021**, *12*, 692574.
182. Lucas, X.; Ciulli, A. Recognition of substrate degrons by E3 ubiquitin ligases and modulation by small-molecule mimicry strategies. *Curr Opin Struct Biol* **2017**, *44*, 101-110.
183. Varshavsky, A. N-degron and C-degron pathways of protein degradation. *Proc Natl Acad Sci U S A* **2019**, *116*, 358-366.
184. Natsume, T.; Kanemaki, M.T. Conditional Degrons for Controlling Protein Expression at the Protein Level. *Annu Rev Genet* **2017**, *51*, 83-102.
185. Tovell, H.; Testa, A.; Maniaci, C.; Zhou, H.; Prescott, A.R.; Macartney, T.; Ciulli, A.; Alessi, D.R. Rapid and Reversible Knockdown of Endogenously Tagged Endosomal Proteins via an Optimized HaloPROTAC Degrader. *ACS Chem Biol* **2019**, *14*, 882-892.
186. Nabet, B.; Roberts, J.M.; Buckley, D.L.; Paulk, J.; Dastjerdi, S.; Yang, A.; Leggett, A.L.; Erb, M.A.; Lawlor, M.A.; Souza, A.; et al. The dTAG system for immediate and target-specific protein degradation. *Nat Chem Biol* **2018**, *14*, 431-441.
187. Neklesa, T.K.; Crews, C.M. Chemical biology: Greasy tags for protein removal. *Nature* **2012**, *487*, 308-309.
188. Stoeckli, E.T. Protocadherins: not just neuron glue, more too! *Dev Cell* **2014**, *30*, 643-644.
189. Asatsuma-Okumura, T.; Ito, T.; Handa, H. Molecular mechanisms of cereblon-based drugs. *Pharmacol Ther* **2019**, *202*, 132-139.
190. Ito, T.; Ando, H.; Suzuki, T.; Ogura, T.; Hotta, K.; Imamura, Y.; Yamaguchi, Y.; Handa, H. Identification of a primary target of thalidomide teratogenicity. *Science* **2010**, *327*, 1345-1350.
191. Békés, M.; Langley, D.R.; Crews, C.M. PROTAC targeted protein degraders: the past is prologue. *Nat Rev Drug Discov* **2022**, *21*, 181-200.
192. Mullard, A. Targeted protein degraders crowd into the clinic. *Nat Rev Drug Discov* **2021**, *20*, 247-250.
193. Sakamoto, K.M.; Kim, K.B.; Kumagai, A.; Mercurio, F.; Crews, C.M.; Deshaies, R.J. Protacs: chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. *Proc Natl Acad Sci U S A* **2001**, *98*, 8554-8559.
194. Sun, X.; Gao, H.; Yang, Y.; He, M.; Wu, Y.; Song, Y.; Tong, Y.; Rao, Y. PROTACs: great opportunities for academia and industry. *Signal transduction and targeted therapy* **2019**, *4*, 64.
195. Yang, Q.; Zhao, J.; Chen, D.; Wang, Y. E3 ubiquitin ligases: styles, structures and functions. *Molecular biomedicine* **2021**, *2*, 23.
196. Garber, K. The PROTAC gold rush. *Nat Biotechnol* **2022**, *40*, 12-16.
197. Wang, S.; He, F.; Tian, C.; Sun, A. From PROTAC to TPD: Advances and Opportunities in Targeted Protein Degradation. *Pharmaceuticals (Basel)* **2024**, *17*.
198. Buckley, D.L.; Van Molle, I.; Gareiss, P.C.; Tae, H.S.; Michel, J.; Noblin, D.J.; Jorgensen, W.L.; Ciulli, A.; Crews, C.M. Targeting the von Hippel-Lindau E3 ubiquitin ligase using small molecules to disrupt the VHL/HIF-1α interaction. *J. Am. Chem. Soc.* **2012**, *134*, 4465-4468.

199. Gao, X.; Iii, H.; Vuky, J.; Dreicer, R.; Sartor, A.; Sternberg, C.; Percent, I.; Hussain, M.; Kalebasty, A.; Shen, J.; et al. Phase 1/2 study of ARV-110, an androgen receptor (AR) PROTAC degrader, in metastatic castration-resistant prostate cancer (mCRPC). *Journal of Clinical Oncology* **2022**, *40*, 17-17.
200. Hamilton, E.; Schott, A.; Nanda, R.; Lu, H.; Keung, C.; Gedrich, R.; Parameswaran, J.; Han, H.; Hurvitz, S. ARV-471, an estrogen receptor (ER) PROTACdegrader, combined with palbociclib in advanced ER+/human epidermal growth factor receptor 2–negative (HER2-) breast cancer: Phase 1b cohort (part C) of a phase 1/2 study. *Journal of Clinical Oncology* **2022**, *40*, TPS1120-TPS1120.
201. Jan, M.; Sperling, A.S.; Ebert, B.L. Cancer therapies based on targeted protein degradation - lessons learned with lenalidomide. *Nat Rev Clin Oncol* **2021**, *18*, 401-417.
202. Schiemer, J.; Horst, R.; Meng, Y.; Montgomery, J.I.; Xu, Y.; Feng, X.; Borzilleri, K.; Uccello, D.P.; Leverett, C.; Brown, S.; et al. Snapshots and ensembles of BTK and cIAP1 protein degrader ternary complexes. *Nat Chem Biol* **2021**, *17*, 152-160.
203. Zorba, A.; Nguyen, C.; Xu, Y.; Starr, J.; Borzilleri, K.; Smith, J.; Zhu, H.; Farley, K.A.; Ding, W.; Schiemer, J.; et al. Delineating the role of cooperativity in the design of potent PROTACs for BTK. *Proc Natl Acad Sci U S A* **2018**, *115*, E7285-E7292.
204. Tinworth, C.P.; Lithgow, H.; Dittus, L.; Bassi, Z.I.; Hughes, S.E.; Muelbaier, M.; Dai, H.; Smith, I.E.D.; Kerr, W.J.; Burley, G.A.; et al. PROTAC-Mediated Degradation of Bruton's Tyrosine Kinase Is Inhibited by Covalent Binding. *ACS Chem Biol* **2019**, *14*, 342-347.
205. Buhimschi, A.D.; Armstrong, H.A.; Toure, M.; Jaime-Figueroa, S.; Chen, T.L.; Lehman, A.M.; Woyach, J.A.; Johnson, A.J.; Byrd, J.C.; Crews, C.M. Targeting the C481S Ibrutinib-Resistance Mutation in Bruton's Tyrosine Kinase Using PROTAC-Mediated Degradation. *Biochemistry* **2018**, *57*, 3564-3575.
206. Sun, Y.; Zhao, X.; Ding, N.; Gao, H.; Wu, Y.; Yang, Y.; Zhao, M.; Hwang, J.; Song, Y.; Liu, W.; et al. PROTAC-induced BTK degradation as a novel therapy for mutated BTK C481S induced ibrutinib-resistant B-cell malignancies. *Cell Res* **2018**, *28*, 779-781.
207. Vaquer-Alicea, J.; Diamond, M.I. Propagation of Protein Aggregation in Neurodegenerative Diseases. *Annu Rev Biochem* **2019**, *88*, 785-810.
208. Wan, Q.; Song, D.; Li, H.; He, M.L. Stress proteins: the biological functions in virus infection, present and challenges for target-based antiviral drug development. *Signal transduction and targeted therapy* **2020**, *5*, 125.
209. Goedert, M. Tau protein and neurodegeneration. *Semin Cell Dev Biol* **2004**, *15*, 45-49.
210. de Wispelaere, M.; Du, G.; Donovan, K.A.; Zhang, T.; Eleuteri, N.A.; Yuan, J.C.; Kalabathula, J.; Nowak, R.P.; Fischer, E.S.; Gray, N.S.; et al. Small molecule degraders of the hepatitis C virus protease reduce susceptibility to resistance mutations. *Nat Commun* **2019**, *10*, 3468.
211. Martinez-Ortiz, W.; Zhou, M.M. Could PROTACs Protect Us From COVID-19? *Drug Discov Today* **2020**, *25*, 1894-1896.
212. Ghosh, A.K.; Brindisi, M.; Shahabi, D.; Chapman, M.E.; Mesecar, A.D. Drug Development and Medicinal Chemistry Efforts toward SARS-Coronavirus and Covid-19 Therapeutics. *ChemMedChem* **2020**, *15*, 907-932.
213. Alugubelli, Y.R.; Xiao, J.; Khatua, K.; Kumar, S.; Sun, L.; Ma, Y.; Ma, X.R.; Vulupala, V.R.; Atla, S.; Blankenship, L.R.; et al. Discovery of First-in-Class PROTAC Degraders of SARS-CoV-2 Main Protease. *J Med Chem* **2024**, *67*, 6495-6507.
214. Desantis, J.; Bazzacco, A.; Eleuteri, M.; Tuci, S.; Bianconi, E.; Macchiarulo, A.; Mercorelli, B.; Loregian, A.; Goracci, L. Design, synthesis, and biological evaluation of first-in-class indomethacin-based PROTACs degrading SARS-CoV-2 main protease and with broad-spectrum antiviral activity. *European journal of medicinal chemistry* **2024**, *268*, 116202.
215. Sang, X.; Wang, J.; Zhou, J.; Xu, Y.; An, J.; Warshel, A.; Huang, Z. A Chemical Strategy for the Degradation of the Main Protease of SARS-CoV-2 in Cells. *J. Am. Chem. Soc.* **2023**, *145*, 27248-27253.
216. Ciulli, A.; Hsia, O. PROTAC Degraders. In *Protein Homeostasis in Drug Discovery*; 2022; pp. 317-356.
217. Guharoy, M.; Bhowmick, P.; Sallam, M.; Tompa, P. Tripartite degrons confer diversity and specificity on regulated protein degradation in the ubiquitin-proteasome system. *Nat Commun* **2016**, *7*, 10239.
218. Davis, C.; Spaller, B.L.; Matouschek, A. Mechanisms of substrate recognition by the 26S proteasome. *Curr Opin Struct Biol* **2021**, *67*, 161-169.
219. Liu, L.; Damerell, D.R.; Koukouflis, L.; Tong, Y.; Marsden, B.D.; Schapira, M. UbiHub: a data hub for the explorers of ubiquitination pathways. *Bioinformatics* **2019**, *35*, 2882-2884.
220. Farnaby, W.; Koegl, M.; Roy, M.J.; Whitworth, C.; Diers, E.; Trainor, N.; Zollman, D.; Steurer, S.; Karolyi-Oezguer, J.; Riedmueller, C.; et al. BAF complex vulnerabilities in cancer demonstrated via structure-based PROTAC design. *Nat Chem Biol* **2019**, *15*, 672-680.
221. Bondeson, D.P.; Mares, A.; Smith, I.E.; Ko, E.; Campos, S.; Miah, A.H.; Mulholland, K.E.; Routly, N.; Buckley, D.L.; Gustafson, J.L.; et al. Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nat Chem Biol* **2015**, *11*, 611-617.
222. Lai, A.C.; Toure, M.; Hellerschmied, D.; Salami, J.; Jaime-Figueroa, S.; Ko, E.; Hines, J.; Crews, C.M. Modular PROTAC Design for the Degradation of Oncogenic BCR-ABL. *Angew Chem Int Ed Engl* **2016**, *55*, 807-810.
223. Winter, G.E.; Buckley, D.L.; Paulk, J.; Roberts, J.M.; Souza, A.; Dhe-Paganon, S.; Bradner, J.E. DRUG DEVELOPMENT. Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science* **2015**, *348*, 1376-1381.
224. Burke, M.R.; Smith, A.R.; Zheng, G. Overcoming Cancer Drug Resistance Utilizing PROTAC Technology. *Front Cell Dev Biol* **2022**, *10*, 872729.
225. Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* **2001**, *46*, 3-26.
226. Doak, B.C.; Over, B.; Giordanetto, F.; Kihlberg, J. Oral druggable space beyond the rule of 5: insights from drugs and clinical candidates. *Chem Biol* **2014**, *21*, 1115-1142.
227. Edmondson, S.D.; Yang, B.; Fallan, C. Proteolysis targeting chimeras (PROTACs) in 'beyond rule-of-five' chemical space: Recent progress and future challenges. *Bioorg Med Chem Lett* **2019**, *29*, 1555-1564.
228. Wei, M.; Zhao, R.; Cao, Y.; Wei, Y.; Li, M.; Dong, Z.; Liu, Y.; Ruan, H.; Li, Y.; Cao, S.; et al. First orally bioavailable prodrug of proteolysis targeting chimera (PROTAC) degrades cyclin-dependent kinases 2/4/6 in vivo. *European journal of medicinal chemistry* **2021**, *209*, 112903.
229. Li, H.; Dong, J.; Cai, M.; Xu, Z.; Cheng, X.D.; Qin, J.J. Protein degradation technology: a strategic paradigm shift in drug discovery. *J Hematol Oncol* **2021**, *14*, 138.
230. Li, X.; Pu, W.; Zheng, Q.; Ai, M.; Chen, S.; Peng, Y. Proteolysis-targeting chimeras (PROTACs) in cancer therapy. *Molecular cancer* **2022**, *21*, 99.
231. Atilaw, Y.; Poongavanam, V.; Svensson Nilsson, C.; Nguyen, D.; Giese, A.; Meibom, D.; Erdelyi, M.; Kihlberg, J. Solution Conformations Shed Light on PROTAC Cell Permeability. *ACS Med Chem Lett* **2021**, *12*, 107-114.
232. Cyrus, K.; Wehenkel, M.; Choi, E.Y.; Han, H.J.; Lee, H.; Swanson, H.; Kim, K.B. Impact of linker length on the activity of PROTACs. *Mol Biosyst* **2011**, *7*, 359-364.
233. Maxwell, P.H.; Wiesener, M.S.; Chang, G.W.; Clifford, S.C.; Vaux, E.C.; Cockman, M.E.; Wykoff, C.C.; Pugh, C.W.; Maher, E.R.; Ratcliffe, P.J. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* **1999**, *399*, 271-275.
234. Cardote, T.A.F.; Gadd, M.S.; Ciulli, A. Crystal Structure of the Cul2-Rbx1-EloBC-VHL Ubiquitin Ligase Complex. *Structure* **2017**, *25*, 901-911 e903.

235. Zhou, L.; Xu, G. Cereblon attenuates DNA damage-induced apoptosis by regulating the transcription-independent function of p53. *Cell Death Dis* **2019**, *10*, 69.
236. Schapira, M.; Calabrese, M.F.; Bullock, A.N.; Crews, C.M. Targeted protein degradation: expanding the toolbox. *Nat Rev Drug Discov* **2019**, *18*, 949-963.
237. Ting, T.C.; Goralski, M.; Klein, K.; Wang, B.; Kim, J.; Xie, Y.; Nijhawan, D. Aryl Sulfonamides Degrade RBM39 and RBM23 by Recruitment to CRL4-DCAF15. *Cell Rep* **2019**, *29*, 1499-1510 e1496.
238. Ward, C.C.; Kleinman, J.I.; Brittain, S.M.; Lee, P.S.; Chung, C.Y.S.; Kim, K.; Petri, Y.; Thomas, J.R.; Tallarico, J.A.; McKenna, J.M.; et al. Covalent Ligand Screening Uncovers a RNF4 E3 Ligase Recruiter for Targeted Protein Degradation Applications. *ACS Chem Biol* **2019**, *14*, 2430-2440.
239. Spradlin, J.N.; Hu, X.; Ward, C.C.; Brittain, S.M.; Jones, M.D.; Ou, L.; To, M.; Proudfoot, A.; Ornelas, E.; Woldegiorgis, M.; et al. Harnessing the anti-cancer natural product nimbolide for targeted protein degradation. *Nat Chem Biol* **2019**, *15*, 747-755.
240. Wei, J.; Meng, F.; Park, K.S.; Yim, H.; Velez, J.; Kumar, P.; Wang, L.; Xie, L.; Chen, H.; Shen, Y.; et al. Harnessing the E3 Ligase KEAP1 for Targeted Protein Degradation. *J. Am. Chem. Soc.* **2021**, *143*, 15073-15083.
241. Henning, N.J.; Manford, A.G.; Spradlin, J.N.; Brittain, S.M.; Zhang, E.; McKenna, J.M.; Tallarico, J.A.; Schirle, M.; Rape, M.; Nomura, D.K. Discovery of a Covalent FEM1B Recruiter for Targeted Protein Degradation Applications. *J. Am. Chem. Soc.* **2022**, *144*, 701-708.
242. Zhang, Y.; Loh, C.; Chen, J.; Mainolfi, N. Targeted protein degradation mechanisms. *Drug Discov Today Technol* **2019**, *31*, 53-60.
243. Liu, X.; Zhang, X.; Lv, D.; Yuan, Y.; Zheng, G.; Zhou, D. Assays and technologies for developing proteolysis targeting chimera degraders. *Future medicinal chemistry* **2020**, *12*, 1155-1179.
244. Bondeson, D.P.; Smith, B.E.; Burslem, G.M.; Buhimschi, A.D.; Hines, J.; Jaime-Figueroa, S.; Wang, J.; Hamman, B.D.; Ishchenko, A.; Crews, C.M. Lessons in PROTAC Design from Selective Degradation with a Promiscuous Warhead. *Cell Chem Biol* **2018**, *25*, 78-87 e75.
245. Maniaci, C.; Hughes, S.J.; Testa, A.; Chen, W.; Lamont, D.J.; Rocha, S.; Alessi, D.R.; Romeo, R.; Ciulli, A. Homo-PROTACs: bivalent small-molecule dimerizers of the VHL E3 ubiquitin ligase to induce self-degradation. *Nat Commun* **2017**, *8*, 830.
246. Gadd, M.S.; Testa, A.; Lucas, X.; Chan, K.H.; Chen, W.; Lamont, D.J.; Zengerle, M.; Ciulli, A. Structural basis of PROTAC cooperative recognition for selective protein degradation. *Nat Chem Biol* **2017**, *13*, 514-521.
247. Nowak, R.P.; DeAngelo, S.L.; Buckley, D.; He, Z.; Donovan, K.A.; An, J.; Safaee, N.; Jedrychowski, M.P.; Ponthier, C.M.; Ishoey, M.; et al. Plasticity in binding confers selectivity in ligand-induced protein degradation. *Nat Chem Biol* **2018**, *14*, 706-714.
248. Roy, M.J.; Winkler, S.; Hughes, S.J.; Whitworth, C.; Galant, M.; Farnaby, W.; Rumpel, K.; Ciulli, A. SPR-Measured Dissociation Kinetics of PROTAC Ternary Complexes Influence Target Degradation Rate. *ACS Chem Biol* **2019**, *14*, 361-368.
249. Fisher, S.L.; Phillips, A.J. Targeted protein degradation and the enzymology of degraders. *Curr Opin Chem Biol* **2018**, *44*, 47-55.
250. Angers, S.; Li, T.; Yi, X.; MacCoss, M.J.; Moon, R.T.; Zheng, N. Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature* **2006**, *443*, 590-593.
251. Chamberlain, P.P.; Lopez-Girona, A.; Miller, K.; Carmel, G.; Pagarigan, B.; Chie-Leon, B.; Rychak, E.; Corral, L.G.; Ren, Y.J.; Wang, M.; et al. Structure of the human Cereblon-DDB1-lenalidomide complex reveals basis for responsiveness to thalidomide analogs. *Nat Struct Mol Biol* **2014**, *21*, 803-809.
252. Fischer, E.S.; Böhm, K.; Lydeard, J.R.; Yang, H.; Stadler, M.B.; Cavadini, S.; Nagel, J.; Serluca, F.; Acker, V.; Lingaraju, G.M.; et al. Structure of the DDB1-CRBN E3 ubiquitin ligase in complex with thalidomide. *Nature* **2014**, *512*, 49-53.
253. Watson, E.R.; Novick, S.; Matyskiela, M.E.; Chamberlain, P.P.; A, H.d.l.P.; Zhu, J.; Tran, E.; Griffin, P.R.; Wertz, I.E.; Lander, G.C. Molecular glue CELMoD compounds are regulators of cereblon conformation. *Science* **2022**, *378*, 549-553.
254. Lopez-Girona, A.; Mendy, D.; Ito, T.; Miller, K.; Gandhi, A.K.; Kang, J.; Karasawa, S.; Carmel, G.; Jackson, P.; Abbasian, M.; et al. Cereblon is a direct protein target for immunomodulatory and antiproliferative activities of lenalidomide and pomalidomide. *Leukemia* **2012**, *26*, 2326-2335.
255. Melchert, M.; List, A. The thalidomide saga. *Int J Biochem Cell Biol* **2007**, *39*, 1489-1499.
256. Lenz, W. Malformations caused by drugs in pregnancy. *Am J Dis Child* **1966**, *112*, 99-106.
257. Ramsay, A.G.; Evans, R.; Kiaii, S.; Svensson, L.; Hogg, N.; Gribben, J.G. Chronic lymphocytic leukemia cells induce defective LFA-1-directed T-cell motility by altering Rho GTPase signaling that is reversible with lenalidomide. *Blood* **2013**, *121*, 2704-2714.
258. Quach, H.; Ritchie, D.; Stewart, A.K.; Neeson, P.; Harrison, S.; Smyth, M.J.; Prince, H.M. Mechanism of action of immunomodulatory drugs (IMiDS) in multiple myeloma. *Leukemia* **2010**, *24*, 22-32.
259. Zhu, Y.X.; Braggio, E.; Shi, C.X.; Bruins, L.A.; Schmidt, J.E.; Van Wier, S.; Chang, X.B.; Bjorklund, C.C.; Fonseca, R.; Bergsagel, P.L.; et al. Cereblon expression is required for the antimyeloma activity of lenalidomide and pomalidomide. *Blood* **2011**, *118*, 4771-4779.
260. Gandhi, A.K.; Kang, J.; Havens, C.G.; Conklin, T.; Ning, Y.; Wu, L.; Ito, T.; Ando, H.; Waldman, M.F.; Thakurta, A.; et al. Immunomodulatory agents lenalidomide and pomalidomide co-stimulate T cells by inducing degradation of T cell repressors Ikaros and Aiolos via modulation of the E3 ubiquitin ligase complex CRL4(CRBN.). *Br J Haematol* **2014**, *164*, 811-821.
261. Krönke, J.; Udeshi, N.D.; Narla, A.; Grauman, P.; Hurst, S.N.; McConkey, M.; Svinkina, T.; Heckl, D.; Comer, E.; Li, X.; et al. Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science* **2014**, *343*, 301-305.
262. Lu, G.; Middleton, R.E.; Sun, H.; Naniong, M.; Ott, C.J.; Mitsiades, C.S.; Wong, K.K.; Bradner, J.E.; Kaelin, W.G., Jr. The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. *Science* **2014**, *343*, 305-309.
263. Lupas, A.N.; Zhu, H.; Korycinski, M. The thalidomide-binding domain of cereblon defines the CULT domain family and is a new member of the β-tent fold. *PLoS Comput Biol* **2015**, *11*, e1004023.
264. Higgins, J.J.; Pucilowska, J.; Lombardi, R.Q.; Rooney, J.P. A mutation in a novel ATP-dependent Lon protease gene in a kindred with mild mental retardation. *Neurology* **2004**, *63*, 1927-1931.
265. Matyskiela, M.E.; Lu, G.; Ito, T.; Pagarigan, B.; Lu, C.C.; Miller, K.; Fang, W.; Wang, N.Y.; Nguyen, D.; Houston, J.; et al. A novel cereblon modulator recruits GSPT1 to the CRL4(CRBN) ubiquitin ligase. *Nature* **2016**, *535*, 252-257.
266. Chan, J.F.; Kok, K.H.; Zhu, Z.; Chu, H.; To, K.K.; Yuan, S.; Yuen, K.Y. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg Microbes Infect* **2020**, *9*, 221-236.
267. Kirchdoerfer, R.N.; Ward, A.B. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat Commun* **2019**, *10*, 2342.
268. Gao, Y.; Yan, L.; Huang, Y.; Liu, F.; Zhao, Y.; Cao, L.; Wang, T.; Sun, Q.; Ming, Z.; Zhang, L.; et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **2020**, *368*, 779-782.
269. Yan, L.; Ge, J.; Zheng, L.; Zhang, Y.; Gao, Y.; Wang, T.; Huang, Y.; Yang, Y.; Gao, S.; Li, M.; et al. Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription Complex Reveals an Intermediate State in Cap Synthesis. *Cell* **2021**, *184*, 184-193 e110.
270. Caruthers, J.M.; McKay, D.B. Helicase structure and mechanism. *Curr Opin Struct Biol* **2002**, *12*, 123-133.
271. Tanner, J.A.; Watt, R.M.; Chai, Y.B.; Lu, L.Y.; Lin, M.C.; Peiris, J.S.; Poon, L.L.; Kung, H.F.; Huang, J.D. The severe acute respiratory syndrome (SARS) coronavirus NTPase/helicase belongs to a distinct class of 5' to 3' viral helicases. *J Biol Chem* **2003**, *278*, 39578-39582.

272. Mickolajczyk, K.J.; Shelton, P.M.M.; Grasso, M.; Cao, X.; Warrington, S.E.; Aher, A.; Liu, S.; Kapoor, T.M. Force-dependent stimulation of RNA unwinding by SARS-CoV-2 nsp13 helicase. *Biophys J* **2021**, *120*, 1020-1030.
273. Jang, K.J.; Jeong, S.; Kang, D.Y.; Sp, N.; Yang, Y.M.; Kim, D.E. A high ATP concentration enhances the cooperative translocation of the SARS coronavirus helicase nsP13 in the unwinding of duplex RNA. *Sci Rep* **2020**, *10*, 4481.
274. Ivanov, K.A.; Thiel, V.; Dobbe, J.C.; van der Meer, Y.; Snijder, E.J.; Ziebuhr, J. Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. *J Virol* **2004**, *78*, 5619-5632.
275. Xia, H.; Cao, Z.; Xie, X.; Zhang, X.; Chen, J.Y.; Wang, H.; Menachery, V.D.; Rajsbaum, R.; Shi, P.Y. Evasion of Type I Interferon by SARS-CoV-2. *Cell Rep* **2020**, *33*, 108234.
276. Vazquez, C.; Swanson, S.E.; Negatu, S.G.; Dittmar, M.; Miller, J.; Ramage, H.R.; Cherry, S.; Jurado, K.A. SARS-CoV-2 viral proteins NSP1 and NSP13 inhibit interferon activation through distinct mechanisms. *PLoS One* **2021**, *16*, e0253089.
277. Yuen, C.K.; Lam, J.Y.; Wong, W.M.; Mak, L.F.; Wang, X.; Chu, H.; Cai, J.P.; Jin, D.Y.; To, K.K.; Chan, J.F.; et al. SARS-CoV-2 nsp13, nsp14, nsp15 and orf6 function as potent interferon antagonists. *Emerg Microbes Infect* **2020**, *9*, 1418-1428.
278. Newman, J.A.; Douangamath, A.; Yadzani, S.; Yosaatmadja, Y.; Aimon, A.; Brandão-Neto, J.; Dunnett, L.; Gorrie-Stone, T.; Skyner, R.; Fearon, D.; et al. Structure, mechanism and crystallographic fragment screening of the SARS-CoV-2 NSP13 helicase. *Nat Commun* **2021**, *12*, 4848.
279. Singleton, M.R.; Dillingham, M.S.; Wigley, D.B. Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* **2007**, *76*, 23-50.
280. Saikrishnan, K.; Powell, B.; Cook, N.J.; Webb, M.R.; Wigley, D.B. Mechanistic basis of 5'-3' translocation in SF1B helicases. *Cell* **2009**, *137*, 849-859.
281. Yue, K.; Yao, B.; Shi, Y.; Yang, Y.; Qian, Z.; Ci, Y.; Shi, L. The stalk domain of SARS-CoV-2 NSP13 is essential for its helicase activity. *Biochem Biophys Res Commun* **2022**, *601*, 129-136.
282. Deng, Z.; Lehmann, K.C.; Li, X.; Feng, C.; Wang, G.; Zhang, Q.; Qi, X.; Yu, L.; Zhang, X.; Feng, W.; et al. Structural basis for the regulatory function of a complex zinc-binding domain in a replicative arterivirus helicase resembling a nonsense-mediated mRNA decay helicase. *Nucleic Acids Res.* **2014**, *42*, 3464-3477.
283. Tang, C.; Deng, Z.; Li, X.; Yang, M.; Tian, Z.; Chen, Z.; Wang, G.; Wu, W.; Feng, W.H.; Zhang, G.; et al. Helicase of Type 2 Porcine Reproductive and Respiratory Syndrome Virus Strain HV Reveals a Unique Structure. *Viruses* **2020**, *12*.
284. Jia, Z.; Yan, L.; Ren, Z.; Wu, L.; Wang, J.; Guo, J.; Zheng, L.; Ming, Z.; Zhang, L.; Lou, Z.; et al. Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis. *Nucleic Acids Res.* **2019**, *47*, 6538-6550.
285. Hao, W.; Wojdyla, J.A.; Zhao, R.; Han, R.; Das, R.; Zlatev, I.; Manoharan, M.; Wang, M.; Cui, S. Crystal structure of Middle East respiratory syndrome coronavirus helicase. *PLoS Pathog* **2017**, *13*, e1006474.
286. Chen, J.; Malone, B.; Llewellyn, E.; Grasso, M.; Shelton, P.M.M.; Olinares, P.D.B.; Maruthi, K.; Eng, E.T.; Vatandaslar, H.; Chait, B.T.; et al. Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. *Cell* **2020**, *182*, 1560-1573 e1513.
287. Yan, L.; Zhang, Y.; Ge, J.; Zheng, L.; Gao, Y.; Wang, T.; Jia, Z.; Wang, H.; Huang, Y.; Li, M.; et al. Architecture of a SARS-CoV-2 mini replication and transcription complex. *Nat Commun* **2020**, *11*, 5874.
288. Adedeji, A.O.; Marchand, B.; Te Velthuis, A.J.; Snijder, E.J.; Weiss, S.; Eoff, R.L.; Singh, K.; Sarafianos, S.G. Mechanism of nucleic acid unwinding by SARS-CoV helicase. *PLoS One* **2012**, *7*, e36521.
289. Yoshimoto, F.K. The Proteins of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS CoV-2 or n-COV19), the Cause of COVID-19. *Protein J* **2020**, *39*, 198-216.
290. Briguglio, I.; Piras, S.; Corona, P.; Carta, A. Inhibition of RNA Helicases of ssRNA(+) Virus Belonging to Flaviviridae, Coronaviridae and Picornaviridae Families. *Int J Med Chem* **2011**, *2011*, 213135.
291. Ebob, O.T.; Babiaka, S.B.; Ntie-Kang, F. Natural Products as Potential Lead Compounds for Drug Discovery Against SARS-CoV-2. *Nat Prod Bioprospect* **2021**, *11*, 611-628.
292. Kloskowski, P.; Neumann, P.; Kumar, P.; Berndt, A.; Dobbelstein, M.; Ficner, R. Myricetin-bound crystal structure of the SARS-CoV-2 helicase NSP13 facilitates the discovery of novel natural inhibitors. *Acta Crystallogr D Struct Biol* **2025**, *81*, 310-326.
293. Corona, A.; Wycisk, K.; Talarico, C.; Manelfi, C.; Milia, J.; Cannalire, R.; Esposito, F.; Gribbon, P.; Zaliani, A.; Iaconis, D.; et al. Natural Compounds Inhibit SARS-CoV-2 nsp13 Unwinding and ATPase Enzyme Activities. *ACS Pharmacol Transl Sci* **2022**, *5*, 226-239.
294. Kuzikov, M.; Reinshagen, J.; Wycisk, K.; Corona, A.; Esposito, F.; Malune, P.; Manelfi, C.; Iaconis, D.; Beccari, A.; Tramontano, E.; et al. Drug repurposing screen to identify inhibitors of the RNA polymerase (nsp12) and helicase (nsp13) from SARS-CoV-2 replication and transcription complex. *Virus Res* **2024**, *343*, 199356.
295. Inniss, N.L.; Rzhetskaya, M.; Ling-Hu, T.; Lorenzo-Redondo, R.; Bachta, K.E.; Satchell, K.J.F.; Hultquist, J.F. Activity and inhibition of the SARS-CoV-2 Omicron nsp13 R392C variant using RNA duplex unwinding assays. *SLAS discovery : advancing life sciences R & D* **2024**, *29*, 100145.
296. Lee, J.T.; Yang, Q.; Gribenko, A.; Perrin, B.S., Jr.; Zhu, Y.; Cardin, R.; Liberator, P.A.; Anderson, A.S.; Hao, L. Genetic Surveillance of SARS-CoV-2 M(pro) Reveals High Sequence and Structural Conservation Prior to the Introduction of Protease Inhibitor Paxlovid. *mBio* **2022**, *13*, e0086922.
297. Francis, D.M.; Page, R. Strategies to optimize protein expression in E. coli. *Curr Protoc Protein Sci* **2010**, *Chapter 5*, 5 24 21-25 24 29.
298. Demain, A.L.; Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv* **2009**, *27*, 297-306.
299. Yin, J.; Li, G.; Ren, X.; Herrler, G. Select what you need: a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. *J Biotechnol* **2007**, *127*, 335-347.
300. Stewart, E.J.; Aslund, F.; Beckwith, J. Disulfide bond formation in the Escherichia coli cytoplasm: an in vivo role reversal for the thioredoxins. *EMBO J* **1998**, *17*, 5543-5550.
301. Ritz, D.; Beckwith, J. Roles of thiol-redox pathways in bacteria. *Annu Rev Microbiol* **2001**, *55*, 21-48.
302. Studier, F.W.; Moffatt, B.A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **1986**, *189*, 113-130.
303. Huang, B.; Shi, Z.; Tsai, M.D. A small, high-copy-number vector suitable for both in vitro and in vivo gene expression. *Gene* **1994**, *151*, 143-145.
304. Studier, F.W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods In Enzymology* **1990**, *185*, 60-89.
305. Amann, E.; Brosius, J.; Ptashne, M. Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in Escherichia coli. *Gene* **1983**, *25*, 167-178.
306. de Boer, H.A.; Comstock, L.J.; Vasser, M. The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proc Natl Acad Sci U S A* **1983**, *80*, 21-25.
307. Brosius, J.; Erfle, M.; Storella, J. Spacing of the -10 and -35 regions in the tac promoter. Effect on its in vivo activity. *J Biol Chem* **1985**, *260*, 3539-3541.

308. Porath, J.; Carlsson, J.; Olsson, I.; Belfrage, G. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **1975**, *258*, 598-599.
309. Armstrong, R.N. Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chem Res Toxicol* **1997**, *10*, 2-18.
310. Smith, D.B.; Johnson, K.S. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene* **1988**, *67*, 31-40.
311. Andersen, K.R.; Leksa, N.C.; Schwartz, T.U. Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins* **2013**, *81*, 1857-1861.
312. Burgess-Brown, N.A.; Sharma, S.; Sobott, F.; Loenarz, C.; Oppermann, U.; Gileadi, O. Codon optimization can improve expression of human genes in Escherichia coli: A multi-gene study. *Protein Expr Purif* **2008**, *59*, 94-102.
313. Vasina, J.A.; Baneyx, F. Expression of aggregation-prone recombinant proteins at low temperatures: a comparative study of the Escherichia coli cspA and tac promoter systems. *Protein Expr Purif* **1997**, *9*, 211-218.
314. Sahdev, S.; Khattar, S.K.; Saini, K.S. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol Cell Biochem* **2008**, *307*, 249-264.
315. Shirano, Y.; Shibata, D. Low temperature cultivation of Escherichia coli carrying a rice lipoxygenase L-2 cDNA produces a soluble and active enzyme at a high level. *FEBS Lett* **1990**, *271*, 128-130.
316. Piserchio, A.; Ghose, R.; Cowburn, D. Optimized bacterial expression and purification of the c-Src catalytic domain for solution NMR studies. *J Biomol NMR* **2009**, *44*, 87-93.
317. Chou, C.P. Engineering cell physiology to enhance recombinant protein production in Escherichia coli. *Appl Microbiol Biotechnol* **2007**, *76*, 521-532.
318. Peti, W.; Page, R. Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost. *Protein Expr Purif* **2007**, *51*, 1-10.
319. Reilly, D.; Fairbrother, W.J. A novel isotope labeling protocol for bacterially expressed proteins. *J Biomol NMR* **1994**, *4*, 459-462.
320. Linn, S. Strategies and considerations for protein purifications. *Methods Enzymol* **2009**, *463*, 9-19.
321. Mahmoudi Gomari, M.; Saraygord-Afshari, N.; Farsimadan, M.; Rostami, N.; Aghamiri, S.; Farajollahi, M.M. Opportunities and challenges of the tag-assisted protein purification techniques: Applications in the pharmaceutical industry. *Biotechnology Advances* **2020**, *45*, 107653.
322. Phan, J.; Zdanov, A.; Evdokimov, A.G.; Tropea, J.E.; Peters, H.K., 3rd; Kapust, R.B.; Li, M.; Wlodawer, A.; Waugh, D.S. Structural basis for the substrate specificity of tobacco etch virus protease. *J Biol Chem* **2002**, *277*, 50564-50572.
323. Cerofolini, L.; Fragai, M.; Ravera, E.; Diebolder, C.A.; Renault, L.; Calderone, V. Integrative Approaches in Structural Biology: A More Complete Picture from the Combination of Individual Techniques. *Biomolecules* **2019**, *9*.
324. Carter, C.W., Jr.; Carter, C.W. Protein crystallization using incomplete factorial experiments. *J Biol Chem* **1979**, *254*, 12219-12223.
325. Danley, D.E. Crystallization to obtain protein-ligand complexes for structure-aided drug design. *Acta Cryst. D* **2006**, *62*, 569-575.
326. Kaščáková, B.; Koutská, A.; Burdová, M.; Havlíčková, P.; Kutá Smatanová, I. Revealing protein structures: crystallization of protein-ligand complexes - co-crystallization and crystal soaking. *Febs Open Bio* **2025**, *15*, 542-550.
327. Wienen-Schmidt, B.; Oebbeke, M.; Ngo, K.; Heine, A.; Klebe, G. Two Methods, One Goal: Structural Differences between Cocrystallization and Crystal Soaking to Discover Ligand Binding Poses. *ChemMedChem* **2021**, *16*, 292-300.
328. Wüthrich, K. NMR studies of structure and function of biological macromolecules. *Bioscience reports* **2003**, *23*, 119-168.
329. Wider, G. Structure determination of biological macromolecules in solution using nuclear magnetic resonance spectroscopy. *Biotechniques* **2000**, *29*, 1278-1282, 1284-1290, 1292 passim.
330. Stratmann, D.; Boelens, R.; Bonvin, A.M.J.J. Quantitative use of chemical shifts for the modeling of protein complexes. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 2662-2670.
331. Mielke, S.P.; Krishnan, V.V. Characterization of protein secondary structure from NMR chemical shifts. *Prog Nucl Magn Reson Spectrosc* **2009**, *54*, 141-165.
332. Vogl, D.P.; Conibear, A.C.; Becker, C.F.W. Segmental and site-specific isotope labelling strategies for structural analysis of posttranslationally modified proteins. *RSC Chem Biol* **2021**, *2*, 1441-1461.
333. Bax, A. Triple resonance three-dimensional protein NMR: before it became a black box. *J Magn Reson* **2011**, *213*, 442-445.
334. Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
335. Novakovic, M.; Kupče, Ē.; Oxenfarth, A.; Battistel, M.D.; Freedberg, D.I.; Schwalbe, H.; Frydman, L. Sensitivity enhancement of homonuclear multidimensional NMR correlations for labile sites in proteins, polysaccharides, and nucleic acids. *Nat. Commun.* **2020**, *11*, 5317.
336. Kay, L.E.; Ikura, M.; Tschudin, R.; Bax, A. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson* **1990**, *89*, 496-514.
337. Williamson, M.P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn Reson. Spectrosc* **2013**, *73*, 1-16.
338. Becker, W.; Bhattiprolu, K.C.; Gubensäk, N.; Zangger, K. Investigating Protein-Ligand Interactions by Solution Nuclear Magnetic Resonance Spectroscopy. *Chemphyschem* **2018**, *19*, 895-906.
339. Ioannou, K.; Vlasiou, M.C. Metal-based complexes against SARS-CoV-2. *Biometals* **2022**, *35*, 639-652.
340. Marzo, T.; Messori, L. A Role for Metal-Based Drugs in Fighting COVID-19 Infection? The Case of Auranofin. *ACS Med Chem Lett* **2020**, *11*, 1067-1068.
341. Rothan, H.A.; Stone, S.; Natekar, J.; Kumari, P.; Arora, K.; Kumar, M. The FDA-approved gold drug auranofin inhibits novel coronavirus (SARS-COV-2) replication and attenuates inflammation in human cells. *Virology* **2020**, *547*, 7-11.
342. Sonzogni-Desautels, K.; Ndao, M. Will Auranofin Become a Golden New Treatment Against COVID-19? *Front Immunol* **2021**, *12*, 683694.
343. Tabassum, S.F., Z.; Ur-Rahman, M.S.; Hussain, M.A.; Etheshaam uz Zaman, U.S.M; Koneru, A. Drug Information of Auranofin and Its Effectiveness in Covid-19. *IOSR Journal Of Pharmacy* **2021**, *11*, 29-37.
344. Massai, L.; Cirri, D.; Marzo, T.; Messori, L. Auranofin and its analogs as prospective agents for the treatment of colorectal cancer. *Cancer drug resistance (Alhambra, Calif.)* **2022**, *5*, 1-14.
345. Massai, L.; Messori, L.; Carpentieri, A.; Amoresano, A.; Melchiorre, C.; Fiaschi, T.; Modesti, A.; Gamberi, T.; Magherini, F. The effects of two gold-N-heterocyclic carbene (NHC) complexes in ovarian cancer cells: a redox proteomic study. *Cancer Chemother Pharmacol* **2022**, *89*, 809-823.
346. Borgi, I.; Gargouri, A. A spontaneous direct repeat deletion in the pGEX fusion vector decreases the expression level of recombinant proteins in Escherichia coli. *Protein Expr Purif* **2008**, *60*, 15-19.
347. Cantrelle, F.X.; Boll, E.; Brier, L.; Moschidi, D.; Belouzard, S.; Landry, V.; Leroux, F.; Dewitte, F.; Landrieu, I.; Dubuisson, J.; et al. NMR Spectroscopy of the Main Protease of SARS-CoV-2 and Fragment-Based Screening Identify Three Protein Hotspots and an Antiviral Fragment. *Angew Chem Int Ed Engl* **2021**, *60*, 25428-25435.
348. Robertson, A.J.; Ying, J.; Bax, A. NMR Observation of Sulfhydryl Signals in SARS-CoV-2 Main Protease Aids Structural Studies. *Chembiochem* **2022**, *23*, e202200471.

349. Lloyd, H.C.; Li, Y.; Payne, N.C.; Zhao, Z.; Xu, W.; Kroupova, A.; Zollman, D.; Long, T.; Kabir, F.; Chen, M.; et al. A method for the detection and enrichment of endogenous cereblon substrates. *Cell Chem Biol* **2025**, *32*, 1028-1041 e1013.
350. Adhikari, B.; Bozilovic, J.; Diebold, M.; Schwarz, J.D.; Hofstetter, J.; Schröder, M.; Wanior, M.; Narain, A.; Vogt, M.; Dudvarski Stankovic, N.; et al. PROTAC-mediated degradation reveals a non-catalytic function of AURORA-A kinase. *Nat Chem Biol* **2020**, *16*, 1179-1188.
351. Fu, L.; Ye, F.; Feng, Y.; Yu, F.; Wang, Q.; Wu, Y.; Zhao, C.; Sun, H.; Huang, B.; Niu, P.; et al. Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. *Nat Commun* **2020**, *11*, 4417.
352. Lin, C.; Zhu, Z.; Jiang, H.; Zou, X.; Zeng, X.; Wang, J.; Zeng, P.; Li, W.; Zhou, X.; Zhang, J.; et al. Structural Basis for Coronaviral Main Proteases Inhibition by the 3CLpro Inhibitor GC376. *J Mol Biol* **2024**, *436*, 168474.
353. Vuong, W.; Khan, M.B.; Fischer, C.; Arutyunova, E.; Lamer, T.; Shields, J.; Saffran, H.A.; McKay, R.T.; van Belkum, M.J.; Joyce, M.A.; et al. Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nat Commun* **2020**, *11*, 4282.
354. Kim, Y.; Shivanna, V.; Narayanan, S.; Prior, A.M.; Weerasekara, S.; Hua, D.H.; Kankanamalage, A.C.; Groutas, W.C.; Chang, K.O. Broad-spectrum inhibitors against 3C-like proteases of feline coronaviruses and feline caliciviruses. *J Virol* **2015**, *89*, 4942-4950.
355. Han, X.; Zhao, L.; Xiang, W.; Qin, C.; Miao, B.; McEachern, D.; Wang, Y.; Metwally, H.; Wang, L.; Matvekas, A.; et al. Strategies toward Discovery of Potent and Orally Bioavailable Proteolysis Targeting Chimera Degraders of Androgen Receptor for the Treatment of Prostate Cancer. *J Med Chem* **2021**, *64*, 12831-12854.
356. Fan, K.; Wei, P.; Feng, Q.; Chen, S.; Huang, C.; Ma, L.; Lai, B.; Pei, J.; Liu, Y.; Chen, J.; et al. Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. *J Biol Chem* **2004**, *279*, 1637-1642.
357. Teale, F.W.; Weber, G. Ultraviolet fluorescence of the aromatic amino acids. *Biochem J* **1957**, *65*, 476-482.
358. Kroupova, A.; Spiteri, V.A.; Rutter, Z.J.; Furihata, H.; Darren, D.; Ramachandran, S.; Chakraborti, S.; Haubrich, K.; Pethe, J.; Gonzales, D.; et al. Design of a Cereblon construct for crystallographic and biophysical studies of protein degraders. *Nat Commun* **2024**, *15*, 8885.
359. Maio, N.; Raza, M.K.; Li, Y.; Zhang, D.L.; Bollinger, J.M., Jr.; Krebs, C.; Rouault, T.A. An iron-sulfur cluster in the zinc-binding domain of the SARS-CoV-2 helicase modulates its RNA-binding and -unwinding activities. *Proc Natl Acad Sci U S A* **2023**, *120*, e2303860120.
360. Zengerle, M.; Chan, K.H.; Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem Biol* **2015**, *10*, 1770-1777.
361. Ramachandran, S.; Ciulli, A. Building ubiquitination machineries: E3 ligase multi-subunit assembly and substrate targeting by PROTACs and molecular glues. *Curr Opin Struct Biol* **2021**, *67*, 110-119.
362. Farnaby, W.; Koegl, M.; McConnell, D.B.; Ciulli, A. Transforming targeted cancer therapy with PROTACs: A forward-looking perspective. *Curr Opin Pharmacol* **2021**, *57*, 175-183.
363. Bemis, T.A.; La Clair, J.J.; Burkart, M.D. Unraveling the Role of Linker Design in Proteolysis Targeting Chimeras. *J Med Chem* **2021**, *64*, 8042-8052.
364. Sarnow, A.C.; Nassar, H.; Alfayomy, A.M.; Robaa, D.; Sippl, W. HADDOCK-Guided modeling and molecular simulations of cereblon-based ternary complexes: Development of novel PROTACs for Ataxia telangiectasia and RAD3-Related (ATR) kinase. *Comput Biol Med* **2025**, *194*, 110570.
365. Schmitz, C.; Melquiond, A.S.J.; de Vries, S.J.; Karaca, E.; van Dijk, M.; Kastritis, P.L.; Bonvin, A.M.J.J. Protein–Protein Docking with HADDOCK. In *NMR of Biomolecules*; **2012**; pp. 520-535.

Acknowledgements

Completing this PhD has been one of the most challenging and transformative experiences of my life, a journey filled with moments of struggle, growth, and profound satisfaction. Reaching this milestone has been possible not only through my own perseverance but also thanks to the support, encouragement, and kindness of so many wonderful people who have walked this path with me.

I am deeply grateful to my supervisor, Prof. Francesca Cantini, who gave me the opportunity to attend this pathway. Her guidance, patience, and support throughout this PhD journey shaped the direction of my work. My thanks also go to my co-supervisors, Prof. Simone Ciofi-Baffoni and Prof. Mario Piccioli, whose valuable feedback, constructive criticism, and thoughtful suggestions have greatly enriched my research experience.

This research was made possible thanks to the funding from the European Union – Next Generation EU – National Recovery and Resilience Plan, Mission 4 Component 2 – Investment 1.5, THE – Tuscany Health Ecosystem – ECS00000017 – CUP B83C22003920001. I am grateful to CERM and the Department of Chemistry "Ugo Schiff" at the University of Florence for providing the resources, facilities, and collaborative environment that enabled this work to take shape.

Special thanks go to my colleagues and friends in my research group for the stimulating discussions, for the moral support and for the sense of companionship that made every challenge more endurable. Of course, I am thankful to the entire CERM community: PhD students, postdoctoral researchers, thesis students, professors, as well as the administrative and technical staff. Your help, kindness, and shared passion for science have made these years truly special.

On a more personal note, I owe my deepest gratitude to my family. To my parents, for their unconditional love, endless patience, and unwavering belief in me, and to my sister, for always being there with words of encouragement, understanding, and warmth, especially in the hardest moments.

I also want to thank all those who, in one way or another, have accompanied me along this journey, through advice, friendship, or simple gestures of support.

Lastly, I want to pause to recognize my own effort. This PhD journey has challenged me in ways I could never have imagined, demanding resilience, patience, and courage. Looking back, I am proud of the strength and perseverance that have guided me through each difficulty and carried me to the completion of this incredible journey.

Alessia