Discovery, Synthesis, And Structure-Based Optimization of a Series of N-(tert-Butyl)-2-(N-arylamido)-2-(pyridin-3-yl) Acetamides (ML188) as Potent Noncovalent Small Molecule Inhibitors of the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) 3CL Protease

Jon Jacobs,†‡§ Valerie Grum-Tokars,⊥ Ya Zhou,†‡§ Mark Turlington,†‡§ S. Adrian Saldanha,∥ Peter Chase,∥ Aimee Eggler,▽ Eric S. Dawson,†‡§ Yahira M. Baez-Santos,▽ Sakshi Tomar,▽ Anna M. Mielech,¶ Susan C. Baker,¶ Craig W. Lindsley,†‡§ Peter Hodder,∥ Andrew Mesecar,*▽ and Shaun R. Staufer*†‡§

† Department of Pharmacology, Vanderbilt University Medical Center, Nashville, Tennessee 37232, United States
‡ Vanderbilt Center for Neuroscience Drug Discovery, Vanderbilt University Medical Center, Nashville, Tennessee 37232, United States
§ Vanderbilt Specialized Chemistry Center for Probe Development (MLPCN), Nashville, Tennessee 37232, United States
∥ Lead Identification Division, Translational Research Institute, Scripps Research Institute Molecular Screening Center, 130 Scripps Way, Jupiter, Florida 33458, United States
▽ Department of Biological Sciences, Purdue University, 915 West State Street, West Lafayette, Indiana 47907, United States
¶ Department of Microbiology and Immunology, Loyola University Medical Center, 2160 South First Avenue, Maywood, Illinois 60153, United States
# Department of Chemistry, Vanderbilt University, Nashville, Tennessee 37232, United States
* Supporting Information

ABSTRACT: A high-throughput screen of the NIH molecular libraries sample collection and subsequent optimization of a lead dipeptide-like series of severe acute respiratory syndrome (SARS) main protease (3CLpro) inhibitors led to the identification of probe compound ML188 (16-(R), (R)-N-(4-(tert-butyl)phenyl)-N-(2-(tert-butylamino)-2-oxo-1-(pyridin-3-yl)ethyl)furan-2-carboxamide, Pubchem CID: 46897844). Unlike the majority of reported coronavirus 3CLpro inhibitors that act via covalent modification of the enzyme, 16-(R) is a noncovalent SARS-CoV 3CLpro inhibitor with moderate MW and good enzyme and antiviral inhibitory activity. A multicomponent Ugi reaction was utilized to rapidly explore structure–activity relationships within S1’, S1, and S2 enzyme binding pockets. The X-ray structure of SARS-CoV 3CLpro bound with 16-(R) was instrumental in guiding subsequent rounds of chemistry optimization. 16-(R) provides an excellent starting point for the further design and refinement of 3CLpro inhibitors that act by a noncovalent mechanism of action.

INTRODUCTION

Coronaviruses (CoV) are enveloped, large plus-strand RNA viruses that cause medical disorders such as the common cold, lower respiratory tract infections, and diarrhea. The first characterized human CoV strains, 229E and OC43, were identified and studied extensively from 1965 to the mid-1980s. In 2003, the novel SARS-CoV was identified as the etiological agent of the global pandemic of severe acute respiratory syndrome (SARS), an atypical pneumonia that led to progressive respiratory failure in over 8000 individuals and 800 deaths by July of that year. The death rate of approximately 10% for the SARS virus and its ability to spread person-to-person via respiratory droplets make it particularly deadly among worldwide pandemic threats. With the cooperation of leading nations, a rigorous public healthcare campaign was fortunately successful in controlling the outbreak. However, a reemergence of the SARS-CoV is still considered a potential pandemic risk and potentially new strains of SARS could be more severe than that found from the 2003 outbreak. Since 2003, two additional human coronaviruses, NL63 and HKU1, have been identified in patients around the world, and

Received: October 25, 2012
the viruses have been characterized and found to be significantly less lethal than SARS-CoV.\textsuperscript{6--8} Most recently, a new SARS-like virus, called HCoV-EMC, has been identified in at least two individuals, one of whom died.\textsuperscript{9} Sequence analysis of HCoV-EMC indicates that this virus is more closely related to bat coronaviruses than to SARS-CoV. Therefore, the possibility of a future SARS-like pandemic remains possible, and to date, there are still no vaccines or antiviral agents available to prevent or treat SARS-like infections.

The SARS-CoV genome encodes a large polyprotein that is proteolytically processed by two cysteine proteases including the 3C-chymotrypsin-like protease (3CLpro) and the papain-like protease (PLpro). 3CLpro is essential for proteolytic processing at 11 different cleavage sites within the coronavirus polyprotein and is thus vital for viral replication.\textsuperscript{10} The 3CLpro enzyme exists primarily as a dimer in solution, and the dimer has been confirmed to be the active species for the enzyme reaction.\textsuperscript{11} The cloning and expression of recombinant SARS 3CLpro,\textsuperscript{12} along with studies showing that 3CLpro is essential for the viral life cycle,\textsuperscript{13} support a role for 3CLpro as an important pathogenic component of SARS-CoV and therefore a viable target for antiviral drug development.

The SARS-CoV 3CLpro has three domains: I (residues 8–101), II (residues 102–184), and III (residues 201–301). Domains I and II, which contain the active site region, are β-barrel domains, and domain III is an α-helical domain. The active site contains a catalytic dyad consisting of a cysteine residue (Cys-145) that acts as a nucleophile and a histidine residue (His-41) that acts as the general acid–base. Optimized octapeptide-based inhibitors using mutational and CoMFA models have been reported,\textsuperscript{14} and more recently, a systematic saturation mutagenesis study was conducted at the P5 through P3′ positions of the substrate.\textsuperscript{15} These results demonstrate a strong structure–activity relationship between 3CLpro and its substrate and have provided a basis for peptidomimetic inhibitor design. X-ray structures of the SARS-CoV 3CLpro
enzyme bound to hexapeptidyl chloromethyl ketone inhibitors were first reported, and numerous peptidic structures now exist in the context of targeted antiviral drug design. These first generation protease inhibitors maintain a peptidic nature, often five residues in length, and bear a reactive warhead group at the terminus that forms a covalent interaction with Cys-145 (Figure 1, 1–7). Reactive “warhead” groups for 3CLpro have included aldehydes, epoxo-ketones, halo-methyl ketones, trifluoromethyl ketones, and a number of examples of Michael acceptors. These inhibitors often first form a noncovalent interaction complex with the enzyme, positioning the warhead in close proximity to the catalytic cysteine. Attack of the thiolate anion of the catalytic cysteine onto the reactive atom of the warhead leads to formation of the covalent adduct, inactivating the enzyme. One of these compounds, TG-0205221 (9), reacts with SARS 3CLpro with a reported $K_i$ value of ∼60 nM.

These first generation inhibitors achieved submicromolar activity and provided valuable insights into further structure-based inhibitor design, quickly leading to nonpeptidic, warhead-based small molecule inhibitors (Figure 2, 8–11).

Efforts utilizing virtual screening approaches also proved successful leading to nonpeptidic inhibitors (e.g., 12 and 13). One of the early compounds disclosed bearing a cinnamyl amide, cinanserin (Figure 2, 8), has a reported $IC_{50}$ value of ∼5 μM. Low molecular weight nonpeptide inhibitors bearing reactive esters (9–10b) and ketone moieties (11a,b) have demonstrated moderate to good micromolar inhibition and in the case of pyrdyl ester 10b achieved inhibition in a cell-based assay below 10 μM.

The utility and development of using covalently bound inhibitors can unfortunately be limited due to their potential for off-target side effects and toxicity. Despite the fact that there are currently 39 commercially available therapeutics that act through covalent-modification mechanisms, none of these target cysteine proteases. The fact that all of the protease classes (serine, threonine, aspartyl, and metallo), with the exception of cysteine, have been targeted by marketed therapeutics suggests that advancing covalent cysteine-protease inhibitors to market may have unique challenges. Therefore, researchers are now focusing more attention on the identification of noncovalent inhibitors. Legitimate noncovalent inhibitors described to date have been limited to high molecular weight peptidomimetics (MW > 800 amu) with low ligand efficiency and have included known protease drugs originally developed for HIV protease. In an effort to identify small molecule inhibitors of the SARS 3CLpro that inhibit via a noncovalent mechanism and possess low micromolar potency and good cell permeability, a screen was conducted in 2009 against the NIH molecular libraries sample collection (~293000 compounds) in collaboration with the Mesecar laboratory and the Scripps Research Institute Molecular Screening Center (SRIMSC) through the molecular libraries program center network (MLPCN). Herein we describe the results of this screening campaign and the prosecution of an optimization effort at the Vanderbilt Specialized Chemistry Center (VSCC) around a class of 2-(N-arylamido)-2-(pyridin-3-yl) acetamide inhibitors of moderate molecular weight. In addition, from these collaborative efforts and with the aid of X-ray crystallography, we report the details of the molecular interaction of 3CLpro inhibitor 16-(R) with the active site of 3CLpro and further describe the physicochemical and ancillary properties of 16-(R).

### RESULTS AND DISCUSSION

**SARS-CoV 3CL Protease Construct Used in Assays and Lack of Inhibition by Cinanserin.** The 3CLpro enzyme, also known as nonstructural protein 5 (nsp5), is liberated from coronavirus polyproteins by self-cleavage at its N-terminal (nsp4/S) and C-terminal (nsp5/6) cleavage sites. Although the autocatalytic mechanism is still being elucidated, current studies suggest that it involves an initial dimerization event between two unreleased nsp5s within the polyprotein followed by a series of catalytic steps that liberate a fully active, dimeric enzyme that can then proceed to processing the remaining 9 cleavage sites of the polyprotein. The SARS-CoV 3CLpro is highly active as a dimer and is essentially inactive as a monomer. The addition of affinity tags or additional amino acids to the N- and C-termini of 3CLpro is found to substantially increase the dimer dissociation constant ($K_d$) and dramatically decrease the enzymatic activity of the enzyme.

Therefore, we expressed and purified SARS-CoV 3CLpro from a newly designed expression construct that produces the authentic 3CLpro dimer that would be generated in a virus infected cell. The enzyme was purified by a multistep chromatographic procedure and was found to be highly active.

Because the compound cinanserin was reported to be a highly effective noncovalent inhibitor of SARS-CoV and HCoV-229E 3CLpros, with IC$_{50}$ values of 5 μM against each enzyme, we sought to use this compound as a control inhibitor to calculate $Z'$-factors during our high-throughput screening of compound libraries. We tested the ability of cinanserin to inhibit our authentic version of SARS-CoV 3CLpro at a concentration of 100 μM, and surprisingly, we found no inhibition of the enzyme. Next, we tested the ability of cinanserin to inhibit the 3CLpro enzymes from the human coronavirus HKU1 and the mouse coronavirus MHV. Once again, we found no inhibition of these 3CL proteases at a cinanserin concentration of 100 μM.

Since our kinetic inhibition results stand in contrast to those of Chen and co-workers, we sought an explanation for this large discrepancy in potency. Upon closer examination of the reported methods used to determine the IC$_{50}$ values for cinanserin, we found that the 3CLpro expression construct utilized by Chen and co-workers incorporated a (His)$_6$-affinity tag at the N-terminus that extended it by an additional 14 residues. The N-terminal affinity tag was never removed from the purified 3CLpro enzymes prior to their kinetic studies. As described above, the addition of N-terminal or C-terminal residues to SARS 3CLpro is known to have a dramatic effect in increasing the dissociation constant for dimerization of the enzyme, as well as decreasing the enzymatic activity of the enzyme. Therefore, cinanserin may be inhibiting a non-dimeric form of the enzyme. Another possible explanation is that Triton X-100 was not included in the assay buffer to remove potential promiscuous inhibition caused by nonspecific aggregation of the protein by the compounds. We tested cinanserin for inhibition of SARS-CoV, HKU1, and MHV 3CLpro at a concentration of 100 μM cinanserin in both the presence and absence of 0.01% Triton X-100 and found no inhibition of the enzymes in either case. Ultimately, we were unable to use cinanserin as a control inhibitor since no inhibition of SARS-CoV, HKU1, or MHV 3CL proteases was observed. Therefore we used a compound 10a (Figure 2), a covalent modifier, as a positive control compound.
Screening Campaign and Lead Identification at the SRIMSC. High-throughput screening was conducted by evaluating the inhibition of the 3CLpro-mediated peptide cleavage event using a novel FRET-based substrate (Pubchem: AID 1859). We utilized a number of key assay features in the screening campaign that have significant advantages over previous HTS campaigns against the 3CL enzyme. First, we utilized a SARS-CoV 3CLpro construct that consisted of the native sequence of the enzyme that represents the in vivo postproteolytic form of the enzyme. Second, we maintained a high concentration of fresh reducing agent in the primary assays, which helped to minimize but did not eliminate covalent compounds as primary leads. Finally, we utilized a highly fluorescent FRET-based 3CLpro peptide substrate composed of a HiLyte fluor 488 fluorescent label attached at the N-terminus and a QXL520 quencher dye attached to the C-terminus. This substrate allows for excitation in the spectral region of 488 nm that helps to eliminate spectral overlap with the compounds in the screen. Moreover, the high fluorescence quantum yield of the fluorescent probe and sufficiently low turnover number of the enzyme yields a highly sensitive assay that is amenable to screening in 1536 plate format with low ($$< K_m$$) substrate concentrations.

The primary screen of the molecular libraries small molecule repository (~293,000 compounds at the time of screen) was conducted using a 6 $$\mu$$M concentration of each compound. The primary screen produced 406 hits with greater than 12.25% inhibition. An excellent $$Z'$$ score of 0.92 and a relatively low hit-rate of 0.14% were observed. Among the 406 hits, 380 were available for reorder and were tested in a confirmation screen (singles at 10 $$\mu$$M). This second screen produced 136 active compounds with greater than 12.25% inhibition. A subsequent dose–response testing of 101 compounds provided 44 actives with IC50 values below 10 $$\mu$$M. A number of these actives were then counter-screened against the SARS-CoV papain-like protease.
protease, PLpro, which is also a potential target since it too is vital for SARS-CoV maturation.46,47 In general, the majority of IC₅₀ values were substantially greater than 10 μM against PLpro suggesting moderate to good selectivity for 3CLpro inhibition.

Prior to initiation of the chemistry phase of the project, SRIMSC performed structural clustering analysis of the 101 confirmed actives. Several interesting scaffold clusters and a handful of singletons were identified from this analysis. A dipeptide class, represented by hit 14 (Figure 3, SID49730186), was one of two analogues with an IC₅₀ below 10 μM. This furyl amide series was of particular interest due to its starting potency, modular nature, structural properties, and absence of reactive functionality found within typical covalent modifiers.

**Probe Optimization of Furyl Amide 14.** Within the furyl scaffold, an initial 80 compound library array 15 was designed based upon HTS data using an Ugi MCC (multicomponent condensation) strategy (Scheme 1).48,49

In the initial Ugi array, the isocyanide was held constant to a tert-butyl group at R₄. The α-aryl substituent R₃ was maintained either as a 3-pyridyl, as found in the original hit 14, or as a 2-thienyl, which was also identified as a tolerable group in another closely related dipeptide cluster from HTS. Five monomers were selected for the R₂ amine group to examine modifications of the aryl, and a cyclopropyl was included as a potential aryl replacement as a means to reduce MW and overall lipophilicity. For the R₁ carboxylic acid component, eight monomers were chosen focusing exclusively on substituted aromatics and heteroaromotics. As part of the R₁ group scan, a furan-3-carboxylic acid was included as a control for SAR comparisons relative to 14. All 80 compounds were successfully prepared and submitted for testing.

Disappointingly, only two Ugi library analogues displayed inhibition levels greater than 50% at 100 μM and subsequently were determined to have IC₅₀ values less than 50 μM. Notably, these active analogues were exclusively within the 3-pyridyl subseries (Figure 4, 16 and 17). Within the 3-pyridyl series, replacement of the lipophilic N-aryl R₂ group with either a cyclopropyl group or a 4-fluoro phenyl derivative (Figure 4, 18 and 19) were deleterious for inhibition, indicating the importance of steric bulk for optimal interaction within this putative subpocket of the enzyme. The 3-thienyl congener 20, a direct analogue of 16, resulted in a complete loss of inhibition activity.

**Concurrent with synthetic plans to prepare a second generation Ugi library, SARS-CoV 3CLpro crystals suitable for X-ray diffraction and analysis were obtained with inhibitor 16 bound to 3CLpro.50** The electron density associated with inhibitor 16 is clearly resolved and allowed for accurate refinement of its position (Figure 6). The binding orientation of 16 is overall similar to known covalent peptidomimetic inhibitors, with inhibitor 16 preferentially occupying the S₃−S₁′ subpockets of the 3CLpro enzyme as the R enantiomer (Figures 5−7).21 In this orientation, the tert-butyl amide occupies the S₃ pocket, the tert-butylanilido group occupies a deep S₂ pocket, the 3-pyridyl group occupies the S₁, and the furyl amide acts as a P₁′ group. For comparison, the

**Figure 4.** Selected SAR highlights from first round Ugi library.

**Figure 5.** Schematic of enzyme pockets occupied by 16 and 4.

**dx.doi.org/10.1021/jm301580n J. Med. Chem. XXXX, XXX, XXX−XXX**
mechanism-based covalent inhibitor 4 (PDB code 2ALV) and its occupied subpockets are represented schematically in Figure 5 relative to inhibitor 16.21 In addition to the S3–S1′ pockets occupied by 16, 4 occupies the S1 pocket, thus relative to 4 the MW of 16 is reduced by over 100 amu and furthermore 16 lacks a reactive warhead.

An electron density map and a solvent-accessible surface depiction of 16 bound to 3CLpro are shown in Figure 6. A stereoview of the inhibitor-bound structure and associated interactions is also shown in Figure 7. Interestingly, the 3-pyridyl ring nitrogen serves as a hydrogen bond acceptor within the S1-subpocket engaging His-163 with a N–N interatomic distance of 2.8 Å. A similar hydrogen-bond interaction with His-163 engaged is retained in other 3CLpro crystal structures.21 The furan ring oxygen and the amide carbonyl oxygen have a bifurcated interaction with the backbone NH of Gly-143. Although the distance to the amide carbonyl is 0.5 Å closer to Gly-143 and appears to be a hydrogen bond, based upon distance (C=O to NH = 2.9 Å), the geometry is not ideal. In fact, the NH of Gly-143 is nearly coplanar with both oxygen atoms. On the floor of the S1′ pocket, the catalytic Cys-145 is positioned beneath both the amide carbonyl carbon and the furan oxygen at a distance of 3.5 Å. With this rather unique furyl amide interaction within the context of the catalytic residues and backbone Gly-143 interaction, we set out to understand this interaction further by preparing a second chemical library focusing exclusively on the P1′ group holding the P1–P3 groups constant. Results from this campaign are summarized in Figure 8 (21–48).

Replacement with five-membered π-excessive heterocycles proved most successful with 22, 27, and 31 having IC50’s ~50 μM and below. A fully saturated tetrahydrofuran (32) and a select set of acyclic analogues were >10-fold less active with IC50’s >50 μM. Substitution on the furan ring negatively impacted activity (21, 26) and although replacement of the...
furan oxygen atom with a larger sulfur atom in the related thiophene (27) was weakly tolerated (10-fold loss), the related pyrrole (37) heterocycle was completely inactive. The nearly equipotent inhibition by the imidazole 31 (SID93616598) however was encouraging, and thus a smaller subset of analogues was prepared to complete the survey of five-membered heterocycles (49–52). This effort identified 2-oxazolyl analogue 49 (SID99350507) with similar potency to the 4-oxazolyl isomer 22. Interestingly, in contrast to the 2-methyl furyl derivative 21, the 2-chloro-substituted furyl congener 50 (SID99350505) was nearly equipotent to 16 with an IC\textsubscript{50} of 5.2 μM. SAR around imidazole 31 proved steep as the N-methyl C(4) substituted analogue 52 was >10-fold less active and addition of a methylene spacer (51) was deleterious for inhibition.

A subsequent survey of P\textsubscript{1} replacements was performed in order to identify alternate hydrogen bond acceptor groups that might engage His-163, while retaining the 2-furyl amide P\textsubscript{1} group (Table 1).

### Table 1. Focused P\textsubscript{1} Group Heterocyclic Library (53–61)

<table>
<thead>
<tr>
<th>Compound</th>
<th>SID</th>
<th>P\textsubscript{1}</th>
<th>IC\textsubscript{50} μM&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>53</td>
<td>117686897</td>
<td>=N=N</td>
<td>10</td>
</tr>
<tr>
<td>54</td>
<td>117686898</td>
<td>=N=N</td>
<td>5.5</td>
</tr>
<tr>
<td>55</td>
<td>117686909</td>
<td>-N=N</td>
<td>Inactive</td>
</tr>
<tr>
<td>56</td>
<td>117686901</td>
<td>=N=N</td>
<td>45</td>
</tr>
<tr>
<td>57</td>
<td>117686900</td>
<td>-N=N</td>
<td>Inactive</td>
</tr>
<tr>
<td>58</td>
<td>117686899</td>
<td>=N=N</td>
<td>16% (100 μM)</td>
</tr>
<tr>
<td>59</td>
<td>117686904</td>
<td>=N=N</td>
<td>Inactive</td>
</tr>
<tr>
<td>60</td>
<td>117686905</td>
<td>=N=N</td>
<td>23% (100 μM)</td>
</tr>
<tr>
<td>61</td>
<td>117686906</td>
<td>=N=N</td>
<td>23% (100 μM)</td>
</tr>
</tbody>
</table>

<sup>a</sup>IC\textsubscript{50} values are average of two independent runs using triplicate concentrations; “Inactive” defined as %inhibition <15% at 100 μM, CV <0.3.

Among the various P\textsubscript{1} replacements examined (53–61) pyrazidine (53) and pyrazine (54) were well tolerated, although no improvement was found over lead compound 16. More dramatic modifications around the pyridyl ring were clearly limited for reasons consistent with the inhibitor bound X-ray structure because both the 4- and 2-pyridyl analogues were either weak inhibitors (56) or inactive (57) and addition of a 2-methyl in pyridyl analogue 55 was not permitted. Attempts to replace the 3-pyridyl P\textsubscript{1} with 1,2-diazoles (58, 59), 1,2,3-triazoles (60), or 1,3-diazoles (61) were also unsuccessful in identifying compounds with improved inhibition.

We next turned to chiral stationary phase supercritical fluid chromatography (SFC) to separate and isolate the single enantiomers of 16 (Figure 9). Highly specific inhibition was found for single stereoisomer 16-\textsuperscript{R} with an IC\textsubscript{50} of 1.5 ± 0.3 μM, which was identified as the first eluting peak from chiral SFC. In addition, inhibitor 16-\textsuperscript{R} was found to specifically inhibit 3CL\textsubscript{pro} versus PL\textsubscript{pro}. Inhibitor 16-\textsuperscript{R} was subsequently declared MLPCN probe ML188.\textsuperscript{41} Inhibitor 16-\textsuperscript{R} was inferred as the \textsuperscript{R} stereoisomer based upon the absolute configuration observed in the electron density map of the X-ray structure of the 16-3CL\textsubscript{pro} complex, and it was found to rotate plane polarized light in the positive direction. Finally, the mechanism of inhibition of SARS 3CL\textsubscript{pro} by 16-\textsuperscript{R} was determined to be competitive based on a double-reciprocal plot of the initial rate of the SARS 3CL\textsubscript{pro} catalyzed reaction versus variable concentrations of substrate and fixed, variable concentrations of 16-\textsuperscript{R} (see Figure 2, Supporting Information). The resulting K\textsubscript{i} value was determined to be 1.6 ± 0.26 μM, which is equivalent to its IC\textsubscript{50} value and is expected since the IC\textsubscript{50} values for the inhibitors are determined at a substrate concentration that is significantly lower than the K\textsubscript{m} value.

#### SARS-CoV Antiviral Activity of 16-\textsuperscript{R}.

The potency of 16-\textsuperscript{R} against SARS-CoV 3CL\textsubscript{pro} was deemed sufficient to determine the antiviral potency of the compound against SARS-CoV Urbani infected Vero E6 cells. Using our established antiviral assay and BSL3 protocols, we generated a dose–response curve of 16-\textsuperscript{R} against mock-infected and SARS-CoV infected cells, and the results are shown in Figure 10.\textsuperscript{47}

A fit of the data presented in Figure 10 yields an antiviral EC\textsubscript{50} value of 12.9 ± 0.7 μM. A second experiment, performed one month later, yielded an EC\textsubscript{50} value of 13.4 ± 1.0 μM. The two independent experiments demonstrate that the resulting noncovalent inhibitor, 16-\textsuperscript{R}, can effectively inhibit SARS-CoV replication in cell culture.

We searched the literature for other noncovalent inhibitors against SARS-CoV 3CL\textsubscript{pro} that also have antiviral data associated with the compound. The compound cinanserin is reported to have antiviral activity.\textsuperscript{26} In that report, SARS-CoV infected Vero cells were treated with 50 μg/mL (134 μM) cinanserin. The authors conducted an RNA quantitation assay of the cell culture media and reported a 3 log reduction and indicate that this reduction correlates with the titer of infectious particles. The authors also did a control cytotoxicity assay and reported that the cytotoxic concentration that gives 50% reduction (CC\textsubscript{50}) is 31 μM and the CC\textsubscript{50} is 66 μM. Given the fact that under these high cinanserin concentrations (134 μM) the majority of Vero cells (>90%) are likely killed by cinanserin, interpretation of the antiviral data in these studies is deemed ambiguous.

We next compared the EC\textsubscript{50} of 16-\textsuperscript{R} value to another noncovalent compound tested against SARS-CoV infected cells that targets the SARS-CoV papain-like protease (PL\textsubscript{pro}) known as GRL0617.\textsuperscript{37} Compound GRL0617 has an IC\textsubscript{50} against the PL\textsubscript{pro} enzyme of 600 nM and an EC\textsubscript{50} value of 14.5 μM against SARS infected Vero E6 cells.\textsuperscript{37} We next compared the EC\textsubscript{50}/IC\textsubscript{50} ratios of the noncovalent compounds, 16-\textsuperscript{R} and GRL0617, which are ∼8.7 and ∼24.1, respectively. This ratio can help guide compound design in terms of estimating what value of an IC\textsubscript{50} needs to be achieved in order to achieve a pharmacologically relevant effect. Therefore, we
compared these values to the most potent covalent 3CLpro inhibitor identified to date, TG-0205221 (S). This compound has an EC₅₀/Kᵢ value of 0.6 μM/0.052 μM or ~11.5. Thus, noncovalent and covalent inhibitors of 3CLpro can achieve similar antiviral effects.

Based upon the structural information, excellent 3CLpro protease inhibition activity, and SARS-CoV Ubani antiviral activity, 16-(R) was elected as a first in class probe candidate from the furyl amide series. Probe 16-(R) was found to be highly selective for 3CLpro versus PLpro, and in a Ricerca lead-profiling screen against 68 discrete GPCRs, ion channels and transporters, no significant activity was found (<50% at 10 μM). In PBS buffer at neutral pH, probe 16-(R) was found to have excellent aqueous solubility up to 95 μg/mL or 219 μM. Extended characterization of probe inhibitor 16-(R) and analogues against other cysteine proteases and in SARS-CoV infected cells as well as DMPK studies to assess metabolic stability and plasma protein binding are ongoing.

**CONCLUSIONS**

In summary, we have described the optimization and molecular interaction details for a potent, noncovalent inhibitor of SARS-CoV 3CLpro identified through the MLPCN initiative. A four-component Ugi reaction was utilized to rapidly generate SAR and optimize potency focusing on the S₁, S₂, and S₃ binding pockets. X-ray analysis of the 3CLpro−16-(R) complex reveals several key interactions crucial for activity; in particular, a hydrogen bonding interaction between the inhibitor 3-pyridyl ring nitrogen and the active site His-163 side-chain located within the S₁-subpocket. Probe 16-(R) (ML188) is a modest molecular weight SARS-CoV 3CLpro inhibitor with demonstrated antiviral activity and a noncovalent mechanism of action and thus provides the opportunity to facilitate further structure-based design and inhibitor refinement in the quest for potential novel anti-SARS CoV therapeutics. Collaborative efforts in these laboratories continue toward the identification of submicromolar 3CLpro inhibitors with optimal properties for antiviral activity and potential testing in animal models. ML188 is an MLPCN probe and is freely available upon request.

**EXPERIMENTAL SECTION**

**General.** All NMR spectra were recorded on a Bruker 400 and 600 MHz instrument. H chemical shifts are reported in δ values in ppm downfield from TMS as the internal standard in d₅-MeOH or CDCl₃. Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, br = broad, m = multiplet), integration, coupling constant (Hz). Low-resolution mass spectra were obtained on an Agilent 1200 series 6130 mass spectrometer. High-resolution mass spectra were recorded on a Waters Q-TOF API-MS. Analytical thin layer chromatography was performed on Analtech silica gel GF 250 μm plates. Analytical HPLC was performed on an HP1100 with UV detection at 214 and 254 nm along with ELSD detection, LC-MS (J-Sphere80-C18, 3.0 mm × 50 mm, using either a 1.1 or 3.1 min gradient, 5%[0.05%TFA/CH₃CN]/95%[0.05%TFA/H₂O] to 100%[0.05%TFA/CH₃CN]). Preparative RP-HPLC purification was performed on a custom HP1100 automated purification system with collection triggered by mass detection or using a Gilson Inc. preparative UV-based system using a Phenomenex Luna C18 column (50 mm × 30 mm I.D., 5 μm) with an acetonitrile (unmodified)−water (0.1% TFA) custom gradient. Normal-phase silica gel preparative purification was performed using an automated Combi-flash companion from ISCO. Solvents for extraction, washing, and chromatography were HPLC grade. All reagents were purchased from Aldrich Chemical Co. and were used without purification. All polymer-supported reagents were purchased from Argonaut Technologies and Biotage.

**General Procedure for 4CC-Ugi Reaction in Library Format.** To a series of 13 mm × 100 mm screw top glass tubes fitted with a magnetic stir bar equimolar amounts (0.08 mmol) of carbamaldehyde, amine, and carboxylic acid were combined in methanol (0.2 M, 1.5 mL) and subsequently treated with tert-butylisocyanide (0.08 mmol). The mixture was stirred for 16 h at ambient temperature and then concentrated under a stream of nitrogen in a well-ventilated hood. The crude mixtures were reconstituted in MeOH, treated with Argosresin MP-Trisamine (Biotage Inc.) scavenger for 2 h and applied to a Celite pad using a manifold fitted with polypropylene filter tubes capable of
filtering 24 samples in parallel. The filtrates were purified directly using an automated mass-guided RP-HPLC, and product containing fractions were concentrated to give final products >95% purity as judged by LC-MS (215 nm and ELSD) and $^1$H NMR (representative library members >10%). All final library array product characterization data, $^1$H NMR spectra, and SARS-3CLpro % inhibition at 100 µM can be found in Supporting Information.

Analogues 21–48 and 49–61 were prepared in an analogous manner and purified by Gilson Inc. preparative UV-based RP-HPLC. Chromatographic, LC-MS data, and representative $^1$H and $^{13}$C NMR spectra for these remaining analogues can be found in Supporting Information. $^1$H NMR data for 20% of these libraries have been presented above.

Preparation of (R)-(N-4-(tert-Butyl)phenyl)-(N-2-(tert-Butylamino)-2-oxo-1-(pyridin-3-yl)ethyl)-N-(4-(tert-Butyl)phenyl)furan-2-carboxamide (16-R, ML188). To a 20 mL scintillation vial, equimolar amounts (0.5 mmol) of pyridine-3-carboxaldehyde, 4-tert-butylaniline, and furan-2-carboxylic acid were combined in methanol (0.2M, 2.5 mL) and treated with tert-butylisocyanate (0.5 mmol). The mixture was stirred for 16 h at ambient temperature and then concentrated and applied as a DCM solution to a silica column (12 g) and purified over a gradient of 0 to 100% EtOAc in hexanes. The racemic product was obtained as a light yellow solid (198 mg, 91%): LC-MS (>98%) $m/z$ = 434 [M + H$^+$]. Separation of enantiomers was accomplished using supercritical fluid chromatography: the column was an IA UV 250 mm, 10 mm (± 25 mm, Chiral Technologies), eluent 6% MeOH in CD$_2$Cl$_2$. The desired peak was confirmed by the primary assay was identified as the first eluting peak, retention time = 291 min (CDI 46897844).

$^1$H NMR (400 MHz, CDCl$_3$) δ 8.46 (2H, m), 8.49 (1H, d, $J$ = 7.9 Hz), 7.38 (1H, s), 7.07 (3H, m), 6.98 (2H, bs), 6.15 (2H, m), 6.10 (1H, s), 5.38 (1H, d, $J$ = 3.6 Hz). 2.87 (1 H, p, $J$ = 6.9 Hz), 1.37 (9H, s), 2.10 (6H, d). $^{13}$C NMR (100 MHz, CDCl$_3$) δ 167.8, 159.5, 152.4, 151.3, 149.4, 146.1, 144.9, 138.1, 136.4, 130.4, 130.1, 126.0, 122.7, 117.0, 111.1, 63.5, 51.7, 34.6, 31.2, 28.6. HRMS (ES$^+$, M + H) calc for C$_{27}$H$_{30}$N$_4$O$_3$: 434.2440, found 434.2444.

N-(2-(tert-Butylamino)-2-oxo-1-(pyridin-3-yl)ethyl)-N-(4-(tert-Butyl)phenyl)pyridin-2-carboxamide, 17. $^1$H NMR (400 MHz, CDCl$_3$) δ 8.46 (2H, m), 8.49 (1H, d, $J$ = 7.9 Hz), 7.38 (1H, s), 7.07 (3H, m), 6.98 (2H, bs), 6.15 (2H, m), 6.10 (1H, s), 5.38 (1H, d, $J$ = 3.6 Hz). 2.87 (1 H, p, $J$ = 6.9 Hz), 1.37 (9H, s), 2.10 (6H, d). $^{13}$C NMR (100 MHz, CDCl$_3$) δ 167.8, 159.5, 151.4, 150.0, 149.4, 146.1, 144.8, 138.0, 136.6, 130.5, 127.0, 120.7, 117.0, 111.1, 63.5, 51.6, 33.6, 28.5, 23.8, 23.7. HRMS (ES$^+$, M + H) calc for C$_{27}$H$_{30}$N$_4$O$_3$: 434.2287, found 434.2290.

N-(2-(tert-Butylamino)-2-oxo-1-(pyridin-3-yl)ethyl)-N-(4-(tert-Butyl)phenyl)cyclopropylfuran-2-carboxamide, 18. $^1$H NMR (400 MHz, CDCl$_3$) δ 8.68 (1H, s), 8.57 (1H, d, $J$ = 4.0 Hz), 7.97 (1H, d, $J$ = 7.9), 7.53 (1H, s), 7.32 (1H, d, $J$ = 7.7, 4.8), 7.10 (1H, d, $J$ = 3.4 Hz), 6.57 (1H, s), 6.51 (1H, d, $J$ = 3.4, 1.6 Hz), 5.66 (1H, s), 2.90 (1H, m), 1.36 (9H, s), 0.99 (1H, m), 0.80 (1H, m), 0.68 (m, 2H). $^{13}$C NMR (100 MHz, CDCl$_3$) δ 168.3, 162.8, 150.5, 148.9, 147.2, 144.5, 137.1, 131.1, 122.8, 117.1, 111.2, 64.7, 51.1, 31.2, 28.4, 9.5. HRMS (ES$^+$, M + H) calc for C$_{27}$H$_{30}$N$_4$O$_3$: 434.1818, found 434.1815.

N-(2-(tert-Butylamino)-2-oxo-1-(pyridin-3-yl)ethyl)-N-(4-(tert-Butyl)phenyl)fluoropyr-4-furan-2-carboxamide, 19. $^1$H NMR (400 MHz, CDCl$_3$) δ 8.41 (1H, d, $J$ = 4.0 Hz), 8.37 (1H, s), 7.43 (1H, d, $J$ = 8.0), 7.29 (1H, s), 7.05 (2H, dd, $J$ = 7.8, 4.8), 6.85 (2H, m), 6.29 (1H, s), 6.15 (1H, s), 6.13 (1H, dd, $J$ = 3.5, 1.6 Hz), 5.55 (1H, d, $J$ = 3.6 Hz), 5.24 (1H, s). $^{13}$C NMR (100 MHz, CDCl$_3$) δ 167.8, 162.5 (d, $J$ = 246), 159.3, 151.3, 149.5, 146.1, 144.9, 137.8, 134.9 (d, $J$ = 3.1 Hz). 132.9 (d, $J$ = 8.6 Hz), 130.3, 123.0, 117.2, 115 (d, $J$ = 22.4 Hz), 111.1, 62.8, 51.7, 28.5. HRMS (ES$^+$, M + H) calc for C$_{27}$H$_{30}$N$_4$O$_3$F: 396.1723, found 396.1725.

$^4$N-(2-(tert-Butylamino)-2-oxo-1-(thiophen-3-yl)ethyl)furan-2-carboxamide, 20. $^1$H NMR (400 MHz, CDCl$_3$) δ 7.37 (1H, s), 7.26 (2H, s), 7.24 (1H, s), 7.13 (1H, m), 6.99 (2H, bs), 6.89 (1H, d, $J$ = 5.0 Hz), 6.13 (1H, m), 6.10 (1H, s), 6.06 (1H, s) 5.33 (1H, d, $J$ = 3.6 Hz), 1.35 (9H, s), 1.29 (9H, s). $^{13}$C NMR (100 MHz, CDCl$_3$) δ 168.1, 159.3, 152.0, 146.4, 144.6, 137.2, 134.9, 129.6, 129.0, 126.7, 125.8, 125.1, 116.7, 111.1, 61.5, 51.4, 34.6, 31.3, 28.6. HRMS (ES$^+$, M + H) calc for C$_{26}$H$_{27}$N$_4$O$_3$: 397.1820, found 397.1821.
Expression and Purification of the SARS-CoV 3CLpro Enzyme. An expression construct for the SARS-CoV 3CLpro enzyme was designed to produce the exact coding region of the enzyme released in virus-infected cells. The expression construct contained a hexa-histidine affinity tag and a TEV protease site that is also recognized and cleaved by SARS-CoV 3CLpro. The coding region of SARS-CoV 3CLpro was codon-optimized and synthesized by BioBasic (Canada) and was inserted into a pET expression vector. The enzyme was expressed in Escherichia coli BL21(DE3) and was purified via multi-step purification protocol that employed nickel-charged metal-chelate HiTrap affinity column (GE Health sciences), DEAE anion-exchange chromatography, and size-exclusion chromatography. Purified SARS-CoV 3CLpro was then stored at −80 °C or used immediately for crystallization.

HTS Campaign Assays. 3CLpro HTS Assay Protocol. The primary assay began with the addition of 4 μL of 3CLpro enzyme (150 nM final concentration) in assay buffer (50 mM HEPES, 0.1 mg/mL BSA, 0.01% Triton-X 100, 2 mM DTT) at pH 7.5 into each well of a 1536 microtiter plate. Next, 30 nL of test compound in DMSO or DMSO alone (0.6% final concentration) was added to the appropriate wells. The plates were then incubated for 10 min at room temperature. After incubation, 1 μL of 3CLpro peptide substrate (2 μM final concentration) in 50 mM HEPES, pH 7.5 to each well. After 30 min of incubation at room temperature, 1 μL of 500 mM acetic acid was added to each well to terminate the assay and well fluorescence was read on a PerkinElmer Viewlux using fluorescein filters: excitation wavelength of 480 nm (with 20 nm bandwidth) and emission wavelength of 540 nm (with 20 nm bandwidth). All data was normalized to that of the positive control (3CLpro inhibitor) and wells containing DMSO only (negative control). This protocol was used for combinatorial dose response assays (PubChem AID 1890).

P3pro Counterscreen Assay Protocol. Prior to assay, P3pro peptide substrate and luciferase detection reagent were mixed in assay buffer (50 mM HEPES, 0.1 mg/mL BSA, 5 mM DTT, 0.5 mM EDTA and 1 mM magnesium sulfate) at pH 7.5 and incubated for 60 min. The assay was begun by dispensing 2.5 μL of P3pro enzyme (7.5 nM final concentration) in assay buffer or assay buffer alone into each well of a 1536 microtiter plate. Next, 30 nL of test compound in DMSO or DMSO alone (0.6% final concentration) was added to the appropriate wells. The plates were then incubated for 10 min at room temperature. Next, the enzyme reaction was initiated by dispensing 2.5 μL of the preincubuated mixture containing P3pro peptide substrate and luciferase detection reagent (1 μM final substrate concentration). Finally, well luminescence was read on a PerkinElmer Viewlux after 60 min of incubation at room temperature. All data was normalized to that of the positive control (no enzyme) and wells containing DMSO only (negative control). This protocol was used for catalytic dose response assays (PubChem AID 1944). Detailed assay protocols and data generated for HTS and probe development are found at PubChem AID 1859: pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=1859&loc=ea.

Crystallization and X-ray Structure Determination of SARS-CoV 3CLpro in Complex with Inhibitor 16-(R). Purified SARS-CoV 3CLpro was concentrated to 16 mg/mL in a buffer composed of 20 mM HEPES, pH 7.5, and 5 mM 2-mercaptoethanol. The enzyme was incubated on ice with a final concentration of 1 mM inhibitor. A 1:1 enzyme to crystallization solution ration was used, and the crystallization solution consisted of 14.5% PEG 20000, 50 mM MES, pH 6.0, 50 mM potassium chloride, and 1% MPD. Crystallization trials were set up at room temperature using the method of hanging drop vapor diffusion. Significantly better crystal formation was observed in the absence of sodium chloride, and crystals were sensitive to small fluctuations in pH. Single, well-formed crystals were used for soaking in mother-liquor supplemented with 2 mM of inhibitor at room temperature for 3 h. MPD (15%) was used as a cryoprotectant for freezing the crystals in liquid nitrogen where they were stored until synchrotron time was available. Crystals were transferred from liquid nitrogen into a stream of dry nitrogen gas at 100 K for X-ray data collection.

X-ray data were collected on a Rayonix-300 CCD detector at the Advanced Photon Source, Argonne National Laboratory, on beamline 21 ID-F at the Life Sciences-Collaborative Access Team (LS-CAT). X-ray data were processed and scaled using HKL2000. The SARS-CoV 3CLpro–ML188 complex crystallized as a single monomer in the asymmetric unit and in space group C2 with unit cell dimensions of a = 106.7 Å, b = 82.67 Å, c = 53.12 Å, β = 106.0°. The crystal diffracted to a resolution 1.9 Å. The X-ray intensity data had a final Rmerge of 6.4%, and the data were 97.8% complete overall. Additional data processing statistics are provided in Table 4 in Supporting Information.

X-ray intensities were converted to structure-factor amplitudes by the method of French and Wilson using the program TRUNCATE in the CCP4 program suite.5 The initial phases for the model were determined by the method of molecular replacement using the program MOLREP in the CCP4 program suite. The search model used for molecular replacement consisted of a monomer of SARS-CoV 3CLpro from PDB code 2ALV with all side chains intact, and all waters and ligands removed. The final and optimal molecular replacement solution contained a single monomer in the asymmetric unit.

An initial round of combined positional and B-factor refinement was performed with the program REFMAC in the CCP4 program suite using a maximum-likelihood target function and no σ cutoff on structure factor amplitudes. Initial difference Fourier maps were calculated and visualized using the program COOT.5 The initial Fobs − Fcalc difference maps revealed strong (i.e., σ > 3) residual electron density peaks in the active site for the inhibitor 16-(R). A molecular model for the inhibitor was built using the Monomer Library Sketcher program in the CCP4 program suite, and a monomer library description was created for refinement. Iterative rounds of refinement were performed, and water molecules were added manually into strong (i.e., σ > 3) difference density peaks in the initial refinement stages and into peaks of (i.e., σ > 3) in the final stages of refinement. During these iterative refinements, residual electron density consistent with two DMSO molecules from the buffer solution were identified, and these molecules were built into residual density and included in all subsequent refinements.

Iterative refinement using REFMAC was continued until the Rwork and Rfree values plateaued at their lowest values, which were 21.5% and 27.1%, respectively. At this point, the coordinates for the resulting model were submitted to the TLS (translation/libration/screw) server to generate a multigroup TLS model.5 The resulting TLS groups were visualized using the molecular viewer on the TLS Web site, and 20 TLS groups were chosen.58 Two rounds of TLS and restrained refinement7 were performed in REFMAC with the weighting term set at 0.1. The resulting and final Rwork and Rfree values were 18.7% and 23.5%, respectively, justifying the inclusion of TLS groups in the standard refinement protocol.69 The final model coordinates have been deposited in the PDB under accession code 5YVM.59 A summary of the final X-ray data refinement statistics are given in Table 4 in Supporting Information.

SARS-CoV Antiviral Activity Assays. The SARS-CoV Urbani strain was used in these experiments and was provided by the Centers for Disease Control and Prevention. Maintenance of the Vero E6 cells used in these studies was achieved using Dulbecco’s minimal essential media (DMEM) (Gibco) supplemented with 100 units/mL penicillin, 100 μg/mL streptomycin (Gibco), and 10% FCS (Altanta Biologicals).
All antiviral experiments on SARS-CoV Urbani were carried out at Purdue University using a Biosafety Level 3 suite and approved biosafety protocols established in collaboration with the Purdue University Institutional Biosafety Committee.

Growth of Vero E6 cells was performed by seeding cells onto flat-bottom, 96-well plates at a density of approximately 9 × 10^3 cells per well. Vero E6 cells were either mock-infected with serum-free DMEM or infected with 100-fold the median tissue culture infective dose of SARS-CoV Urbani per well in 100 μL of serum-free MEM and incubated for 1 h at 37 °C with 5% CO₂. The viral inoculum was removed after the 1 h incubation period and then 100 μL of MEM, supplemented with 2% FCS and the 16-(R) inhibitor at concentrations ranging from 30 to 0.1 μM, was added. Cells were then incubated for 48 h at 37 °C with 5% CO₂. All controls and each inhibitor concentration were set up in triplicate, and the antiviral assays were performed independently on at least two separate occasions. Cell viability was determined approximately 48 h after infection using the CellTiter-Glo luminescent cell viability assay (Promega).

**ASSOCIATED CONTENT**

Supporting Information
Additional experimental procedures, analytical data for library compounds, in vitro assays, Lineweaver–Burk plot analysis, X-ray crystallography statistics, and parameters. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes
The final model coordinates have been deposited in the PDB under accession code 3V3M.

**AUTHOR INFORMATION**

Corresponding Author
*Tel 616-936-8407, e-mail shaun.stauffer@vanderbilt.edu; tel 765-494-1924, e-mail amesecar@purdue.edu.

Notes
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

The authors thank the MLPCN (1U54 MH084659 and MH084512) and NIH via P01 AI060915 (A.D.M.), 1R01AI085089 (A.D.M.), 2R01AI026603 (A.D.M.), and 1R03MH84162 (V.L.T.) for support, Nathan Kett and Chris Montgomery (Abzena), Pierre Baillargeon and Lina DeLuca (Scripps) for compound management assistance, and Katharine Emery, Jill Ferguson, and Becky Mercer (Scripps) for administrative assistance in reporting data and project details to the MLPCN. We gratefully acknowledge the synchrotron beamline (LS-CAT) personnel at the Advanced Photon Source at Argonne National Lab. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817).

**ABBREVIATIONS USED**

SARS, severe acute respiratory syndrome; CoV, coronavirus; 3CLpro, chymotrypsin-like protease; PLpro, papain-like protease; MCC, multicomponent condensation; 4CC, four-component condensation reaction; BSL3, biosafety level 3 suite; MLPCN, molecular libraries probe production centers network; SID, substance identifier; AID, assay identifier

**REFERENCES**


(28) For information on the MLPCN and information on how to request probe compounds, such ML188, see http://mli.nih.gov/ml/ mlpcn/.


(32) For information on the MLPCN and information on how to request probe compounds, such ML188, see http://mli.nih.gov/ml/ mlpcn/.


