Organic & Biomolecular Chemistry

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d6ob00697c

2-Substituted oxazolyl and thiazolyl amino acids are suitable P1 surrogates for new SARS-CoV-2 main protease inhibitors

Zijie Liu, Tayla J. Van Oers, Conrad Fischer, Anna Nguyen, Johnson Chew and John C. Vederas *

Received 1st May 2026,
Accepted 26th May 2026

DOI: 10.1039/d6ob00697c

rsc.li/obc

The SARS-CoV-2 main protease is a cysteine protease that can bind in a reversible covalent fashion to the warhead of peptidic inhibitors. The most commonly used P1 amino acid subunit in SARS-CoV-2 therapeutics is a cyclic glutamine derivative, and this moiety has been shown to be relatively quickly metabolized, but contains heteroatoms that interact with key residues in the active site. Here, a structure–activity study has been explored with the synthesis of oxazolyl and thiazolyl unnatural amino acids as P1 moieties. This investigation involved different substitution patterns of these heteroaromatic rings, in an effort to maintain a scaffold that can form the important interactions with residues in the active site. IC_{50} values with the SARS-CoV-2 main protease reveal the 2-oxazolyl and 2-thiazolyl derivatives are similarly effective (290 nM and 270 nM, respectively) to compounds with the cyclic glutamine derivative, demonstrating alternative P1 moieties that can be used in the efficient design of novel coronavirus main protease inhibitors.

Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) main protease (M^{pro}) is a cysteine protease that is key to the replication of the SARS-CoV-2 virus. This protease is one of two proteases that cleaves the two polyproteins (pp1a and pp1ab) to release several functional proteins.[1] Due to the important role the M^{pro} plays in the replication cycle of the virus, this protease is a desirable therapeutic target. The main protease has a specific recognition sequence (Leu-Gln↓Ser-Ala-Gly), which allows it to be selective in the cleavage process.[1] This recognition sequence is unique, and not found among human proteases, which makes this enzyme a selective target, with likely fewer off-target effects. The protease works using a catalytic dyad (Cys145 and His41), where the cysteine thiol/thiolate acts as a nucleophile attacking the substrate amide carbonyl, with the histidine acting as a general base.[1,2]

Most of the current therapeutics that target the main protease are designed to mimic this recognition sequence, three literature examples are shown in Fig. 1. The protease can recognize the compound, and forms a reversible covalent bond with the warhead, which is an electrophilic motif that can react with a thiolate nucleophile. Three examples of commonly used warheads in this type of inhibitor are aldehydes, nitriles, and hydroxymethyl ketones.[1,3–6] In this work, we have decided to focus primarily on compounds with bisulfites and aldehydes. The bisulfites are pro-drugs of aldehydes, as bisulfites convert to aldehydes when under physiological conditions (neutral aqueous conditions), but are more soluble, and often result in lower and better IC_{50} values.[7] Aldehydes and bisul-

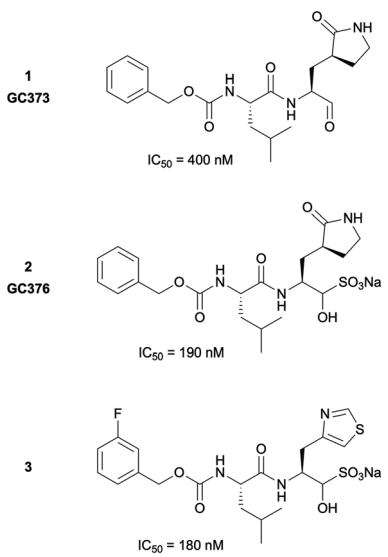


Fig. 1 Chemical structure of literature SARS-CoV-2 main protease inhibitors and their respective IC_{50} values.[1,3]

Department of Chemistry, University of Alberta, Edmonton, AB, T6G 2G2, Canada. E-mail: john.vederas@ualberta.ca

fites are easy to handle, and the synthetic route is shorter and yields two analogues in the same route (one step longer to the bisulfite).

Development of potent main protease inhibitors in the past has mostly revolved around improving P2 and P3 residues to fit tightly into the S2 and S3 subunits of the M^{pro} active site. It was a longstanding dogma that P1 has to be a glutamine analogue to warrant recognition of respective inhibitors by the protease.[8] This P1 is present in the FDA approved therapeutic nirmatrelvir (Paxlovid™), as well as in Pfizer's second-generation inhibitor in clinical trials, ibuzatrelvir.[5,9] However, this cyclic glutamine analogue has been shown to be rapidly metabolized by the P450 enzymes, as it is in nirmatrelvir (Fig. 2), which is the reason nirmatrelvir is provided in combination with ritonavir, an inhibitor of the P450 enzyme responsible for the metabolism.[5,10] This poses a problem, as many individuals cannot take ritonavir, and therefore cannot take Paxlovid™. Heteroaromatic rings may be more stable to metabolic enzymes and we initially believed they may be suitable replacements for the cyclic glutamine analogue.[3] In addition, mutants of the SARS-CoV-2 main protease have appeared in studies and patient isolates that have developed resistance to the commonly used drugs, such as Pfizer's Paxlovid™.[11,12] A different P1 scaffold may provide an alternative option that may be effective against mutants.

Thanks to the abundance of previously acquired crystal structures of various inhibitors with the main protease, it was previously established that there are key interactions present between the P1 subunit and the residues in the S1 pocket that aid in the binding and effectiveness of the inhibitor.[6,13–16] These key interactions are primarily hydrogen bonds to the carbonyl oxygen and the nitrogen of the side chain of the P1 amino acid. Only recently, alternative scaffolds including a 4-thiazole[3] and a 4-oxazole moiety[17] were challenged as P1 surrogates. Since both examples only included 4-substituted heteroazoles we were thus curious whether differently arranged heteroazoles (2- and 5-substituted rings) would potentially improve binding characteristics. We here report the synthesis and binding data of six novel aldehydes (**4a–4f**, Fig. 3) and their bisulfites (**5a–5f**, Fig. 3) that incorporate all three structural oxazole/thiazole isomers (permutations) in P1.

We wanted to keep the P2 and P3 residues the same throughout, and we selected cyclopropylalanine (P2) and 3-fluoro-benzyloxycarbonyl (3F-Cbz, P3) as those accommodate well the corresponding protease subunits and were generally found to give low IC_{50} values previously.[14]

Experimental

Materials and methods

IR spectroscopy. Infrared spectra were recorded on a Thermo Scientific Nicolet 8700 FT-IR spectrometer equipped with a Continuum FT-IR microscope (Thermo Fisher). Cast film refers to preparing a small amount of sample on an IR-transparent silica wafer by depositing a solution of the compound on the wafer and allowing it to dry. The spectrum of the wafer was collected as the background. Neat refers to placing the sample directly on the wafer for testing, rather than dissolving it in a solvent and evaporating it on the wafer.

NMR spectroscopy. Nuclear magnetic resonance (NMR) spectra were obtained on Agilent/Varian 400, 500, 600, and 700 MHz spectrometers and Bruker 400 and 500 MHz spectrometers. Deuterated chloroform ($CDCl_3$, Sigma-Aldrich), deuterated dichloromethane (CD_2Cl_2, Sigma-Aldrich), deute-

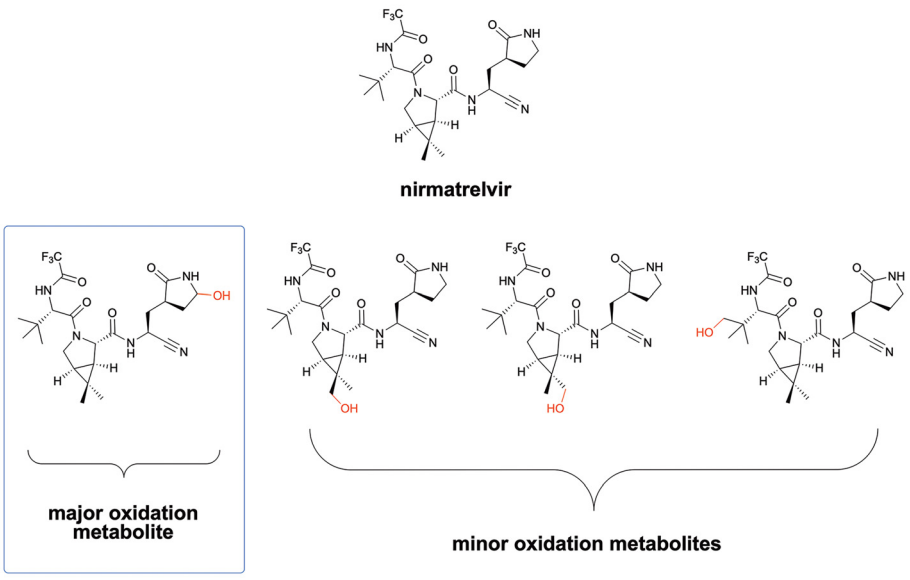


Fig. 2 Pfizer reported the major oxidation metabolite to be the one resulting from oxidation on the lactam ring of the cyclic glutamine amino acid, with the minor metabolites resulting from oxidation on the methyl groups.[5]

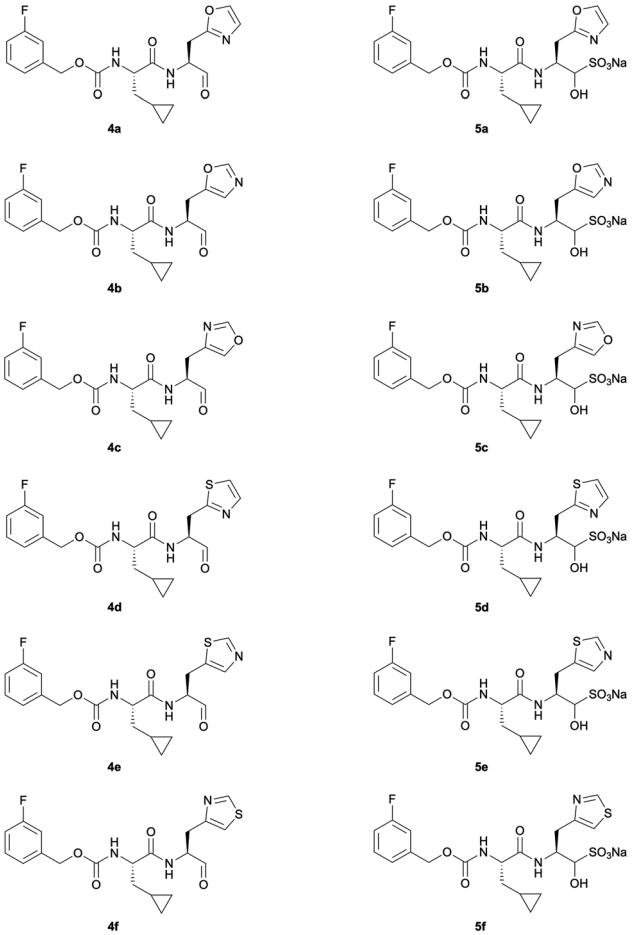


Fig. 3 Target compounds incorporating oxazolyl or thiazolyl amino acids with cyclopropyl alanine and 3F-Cbz.

rated methanol (CD_3OD, Sigma-Aldrich), dimethyl sulfoxide-d_6 (DMSO-d_6, Sigma-Aldrich) and deuterium oxide (D_2O, Sigma-Aldrich) were used as the solvent for NMR spectra. Deuterated chloroform was referenced to 7.26 ppm and 77.16 ppm for the ^{1}H NMR and ^{13}C NMR spectra. Deuterated dichloromethane was referenced to 5.32 ppm and 53.84 ppm for the ^{1}H NMR and ^{13}C NMR spectra. Deuterated methanol was referenced to 3.31 ppm and 49.00 ppm for the ^{1}H NMR and ^{13}C NMR spectra. Dimethyl sulfoxide-d_6 was referenced to 2.50 ppm and 39.52 ppm for the ^{1}H NMR and ^{13}C NMR spectra, respectively. D_2O was referenced to 4.79 ppm for the ^{1}H NMR spectra. When synthesizing peptide aldehydes/bisulfites, there is at least one aqueous work-up step, which allows the aldehyde to spontaneously convert to the hydrate form when exposed to water. Previous in-depth NMR studies by our group have shown that there is epimerization at the alpha-carbon of the C-terminal amino acid of the peptide aldehyde, resulting in a mixture of two aldehydes, and two hydrates (Fig. S10).[14] However, this does not cause issues with the enzyme, the protease has also been shown to bind only to the desired diastereomer.[14]

Optical rotation. PerkinElmer 241 polarimeter was used to measure the optical rotation. It was measured at ambient temperature (25 °C) and are reported in units of deg while all specific rotations were measured at sodium D line (589 nm) in units of deg mL g^{-1} dm^{-1}.

High-resolution mass spectrometry. ESI–TOF (Electrospray ionization, Time of Flight) mass spectra were obtained using Agilent Technologies 6220 oaTOF. DART–Orbitrap (Direct Analysis in Real Time ionization, Orbitrap) mass spectrum was obtained with a Thermo Scientific Exactive Orbitrap.

TLC. Thin-layer chromatography (TLC) was performed on normal phase silica TLC plates (Sigma Aldrich 60 F_{254}) and one or more methods were used for visualizing each compound on TLC plates: UV absorption by fluorescence quenching, potassium permanganate stain (1 L of H_2O, 7.5 g $KMnO_4$, 50 g K_2CO_3, 0.625 g NaOH), ninhydrin stain (0.75 g ninhydrin in 50 mL of *n*-butanol, 1.5 mL acetic acid), or 2,4-dinitrophenylhydrazine (DNP) stain (1.2 g of 2,4-dinitrophenylhydrazine, 6 mL of conc. sulfuric acid, and 8 mL of water in 20 mL of 95% ethanol).

Synthetic procedures

General procedure A (alcohol to mesylate). This procedure was based on a literature procedure.[18] Alcohol (1.0 equiv.) was dissolved in dry DCM (0.1 M) and cooled to 0 °C. After 5 min, triethylamine (1.5 equiv.) was added with a syringe. After that, methanesulfonyl chloride (1.4 equiv.) was added dropwise to the stirred solution. The reaction mixture was allowed to warm to room temperature and stirred for 1 hour. DCM was added to the mixture, and the organic layer was then washed with saturated $NaHCO_3$. The resulting organic layer was separated and dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. Purification was achieved by silica gel column chromatography with elution monitored by TLC with $KMnO_4$ stain.

General procedure B (mesylate to alkylated DEAM). This procedure was based on a literature procedure.[18] To a freshly prepared solution of sodium ethoxide in ethanol (from 1.05 equiv. of sodium in absolute ethanol, 0.50 M) was added diethyl acetamidomalonate (DEAM) (1.05 equiv.). The resulting mixture was heated to reflux for 30 min, then cooled to 55 °C. An ethanolic solution of mesylate (1.0 equiv., in ethanol, 1.0 M) was added dropwise to this solution and then the reaction mixture was stirred at 55 °C for 24 h. the solution was then cooled to 5 °C, then the reaction mixture was diluted with ethyl acetate and filtered through a small plug of Celite to remove the salt. The filtrate was evaporated under reduced pressure, and the residue was redissolved in 20 mL of ethyl acetate. After washing with water, the organic layer was dried over anhydrous Na_2SO_4 and filtered. The solvent was removed under reduced pressure. Purification was achieved by silica gel column chromatography with elution monitored by TLC with $KMnO_4$ stain.

General procedure C (alcohol to alkyl chloride). This procedure was based on a literature procedure.[19] Alcohol (1.0 equiv.) was dissolved in DCM (0.2 M) and then cooled down to 0 °C. Thionyl chloride (1.5 equiv.) was added dropwise to the solution to afford a turbid white mixture, and the resulting mixture was stirred overnight at room temperature. The

mixture was washed with water and the white precipitation disappeared to afford a clear organic layer. The layers were separated, and the organic layer was then washed with saturated aq. $NaHCO_3$ and dried over anhydrous Na_2SO_4. The solvent was removed under reduced pressure to afford a pure product without further purification.

General procedure D (alkyl chloride to alkylated DEAM). This procedure was based on a literature procedure.[20] To a flame-dried three-neck flask, absolute ethanol (0.50 M) and sodium (1.05 equiv.) were added. The solution was stirred until all the sodium was consumed. Then the diethyl acetamidomalonate (1.05 equiv.) was added, and the resulting mixture was heated to reflux for 30 min and then cooled down to 55 °C. Potassium iodide (0.1 equiv.) was added, followed by the slow addition of the ethanolic solution of alkyl chloride (1.0 equiv.). The reaction temperature was held at 50 °C for 24 h. The reaction mixture was then cooled to 5 °C, then the reaction mixture was diluted with ethyl acetate and filtered through Celite to remove the salt. The filtrate was concentrated under reduced pressure, and the residue was redissolved in ethyl acetate. After washing with water, the organic layer was dried over anhydrous Na_2SO_4, concentrated under reduced pressure. Purification was achieved by silica gel column chromatography with elution monitored by TLC with $KMnO_4$ stain.

General procedure E (alkylated DEAM to *rac N*-acetyl amino acid). This procedure was based on a literature procedure.[18] Alkylated diethyl acetamidomalonate (1.0 equiv.) was dissolved in H_2O, followed by the addition of 1 M NaOH (aq) (2.0 equiv.). The resulting solution was stirred at room temperature overnight. Then, to complete the hydrolysis of the two ethyl esters, 1 M NaOH (aq) (0.2 equiv.) was added, and the reaction mixture was stirred overnight at 50 °C. The reaction mixture was then cooled to room temperature, and the pH was adjusted to ~4 using HCl (aq) (6 M). The resulting reaction mixture was heated to reflux at 100–105 °C. Every 6 h, the pH was adjusted to ~4 until the decarboxylation was complete (~6 h). HCl and water were removed under reduced pressure to afford a mixture of racemic *N*-acetyl amino acid and inorganic salts, and were then moved forward to the next step without purification and full characterization. A small sample was taken and purified by HPLC.

General procedure F (*N*-acetyl amino acid ethyl ester to *rac N*-acetyl amino acid). This procedure was based on a literature procedure.[18] *N*-acetyl amino acid ethyl ester was dissolved in H_2O, followed by the addition of 1 M NaOH (aq) (1.0 equiv.). The resulting solution was stirred at room temperature overnight. Then, to complete the hydrolysis of the two ethyl esters, 1 M NaOH (aq) (0.1 equiv.) was added, and the reaction mixture was stirred overnight at 50 °C.

General procedure G (*rac N*-acetyl amino acid to L-amino acid). This procedure was based on a literature procedure.[21] Racemic *N*-acetyl amino acid was dissolved in Milli-Q water, and 10 mM $CoCl_2$ (aq.) was added (the concentration of *N*-acetyl amino acid is ~0.6 M). And then the pH of this solution was adjusted to 7.5–8.0 with 6 M KOH (aq.) Acylase I from *Aspergillus melleus* (8.3 U $mmol^{-1}$ *N*-acetyl amino acid) was dissolved in another vial, same volume of Milli-Q water, and 10 mM $CoCl_2$ (aq.) was added. These two solutions were incubated separately at 37 °C for 30 min, and then the solutions were mixed to afford a solution in which the concentration of *N*-acetyl amino acid is ~0.3 M. The reaction proceeded for 24 h in an incubator (37 °C). During this time, the pH of the reaction mixture was monitored, and 0.1 M KOH (aq.) was added to maintain the pH of the reaction at approximately 7.5. After 24 h, the reaction was adjusted to pH 5.0 with 6 M HCl (aq.), heated to 60 °C with charcoal (25% wt/wt substrate mass) for 1 h, and centrifuged. The supernatant was collected and acidified to pH ≈ 1.5 with 6 M HCl (aq.) and extracted with ethyl acetate. The aqueous layer was retained, and the HCl and water were removed under reduced pressure, and the residue was redissolved in Milli-Q water (~5 mL). The pH of the mixture was adjusted to ~2 before loading onto the ion exchange column. The mixture was applied to the DOWEX 50WX8 cation exchange column, and the column was successively eluted with 50 mL each of 0.1 M, 0.3 M, and 0.5 M HCl to desalt and remove the *N*-acetyl amino acid (for the mixture of thiazolylalanine and *N*-acetyl thiazolylalanine, *N*-acetyl amino acid cannot be removed in this stage). Then the column was eluted with Milli-Q water until the excess HCl was removed, and then with 2% ammonium hydroxide (~50 mL) to elute the column. The fractions that contained L-amino acid were combined, and the water was removed under reduced pressure. The residue was co-evaporated with methanol and diethyl ether sequentially to remove the remaining water.

General procedure H (amino acid to amino acid methyl ester). This procedure was based on a literature procedure.[3] To methanol was added thionyl chloride (20% v/v MeOH) at 0 °C (ice-bath) dropwise; the resulting mixture was stirred for 0.5 h, and amino acid (~11 mM) was added. The reaction mixture was allowed to warm to room temperature and was stirred overnight. The reaction progress was monitored by TLC (*n*-BuOH/AcOH/H_2O = 4 : 1 : 1, v/v) with ninhydrin stain. The solvent was removed under reduced pressure after all the amino acid had been consumed, and this material was used directly in the next step.

General procedure I (coupling of carboxylic acid and amino acid methyl ester). This procedure was based on a literature procedure.[1,3] In a flame-dried round-bottom flask under argon, 7 (1.0 equiv.) and HATU (1.0 equiv.) were dissolved in DMF (0.2 M). DIPEA (3.0 equiv.) was added with a syringe, and the reaction mixture was left to stir at room temperature for 12 min, during which time it turned yellow. Amino acid methyl ester (1.0 equiv.) was dissolved in DMF (0.2 M) and was added dropwise to the stirred mixture. The reaction mixture was stirred under an argon atmosphere at room temperature for 1.5 h. After the reaction was complete, the reaction mixture was diluted with water and ethyl acetate. Then the layers were separated, and the aqueous layer was extracted further with ethyl acetate (3×). All ethyl acetate layers were combined and washed with sat. aq. $NaHCO_3$ (2×), 1 M HCl (2×), and brine (1×). The organic layer was then dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure. Purification was achieved by silica gel column chromatography with elution monitored by TLC with $KMnO_4$ stain.

General procedure J (reduction of methyl ester to alcohol). This procedure was based on a literature procedure.[1,3] The methyl ester (1.0 equiv.) was added to a flame-dried round-bottom flask under argon and dissolved in dry THF (0.1 M). Lithium borohydride (2.0 M in THF) (3.0 equiv.) was added dropwise. The reaction was left to stir at room temperature for 1.5 h under argon. The reaction was quenched with 1 M HCl until it reached pH 1–2. The reaction mixture was concentrated under reduced pressure to obtain a solid, which was then dissolved in ethyl acetate and brine. The layers were separated, and the ethyl acetate layer was dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure. Purification was achieved by silica gel column chromatography with elution monitored by TLC with $KMnO_4$ stain.

General procedure K (oxidation of alcohol to aldehyde). This procedure was based on a literature procedure.[1,3] The alcohol (1.0 equiv.) was dissolved in dry DCM (0.5 M) in a flame-dried round-bottom flask under argon. This was cooled to 0 °C, and Dess–Martin periodinane (1.5 equiv.) was added. The reaction mixture was capped under argon and allowed to warm to room temperature and stirred for 3 h. The cloudy white solution was quenched with 10% aq. $Na_2S_2O_3$ and stirred for 15 min. The reaction mixture was diluted with ethyl acetate and an additional 10% aq. $Na_2S_2O_3$ until a clear solution was obtained, and the layers were separated. The organic layer was washed with 10% aq. $Na_2S_2O_3$ (1×), sat. aq. $NaHCO_3$ (2×), water (2×), and brine (2×). It was then dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure. Purification was achieved by silica gel column chromatography with elution monitored by TLC with $KMnO_4$ stain and DNP stain.

General procedure L (aldehyde to bisulfite adduct). This procedure was based on a literature procedure.[1,3] The aldehyde (1.0 equiv.) was dissolved in the solvent mixture (156 mM) of dry ethyl acetate and dry ethanol (v/v = 5 : 3). $NaHSO_3$ (aq) solution (1 M, 1.0 equiv.) was added to the solution of the aldehyde, the reaction vessel was then sealed, and the reaction mixture was stirred at 50 °C for 3 h. Afterwards, the reaction mixture was filtered to remove any solids, and then the solids were washed with dry ethanol. The filtrate was concentrated under reduced pressure, followed by co-evaporation with diethyl ether (3×) to remove the remaining solvent, furnishing a solid. Diethyl ether was then added, the mixture was sonicated to afford a dispersion, after which, the solid was separated by centrifugation after removing the supernatant.

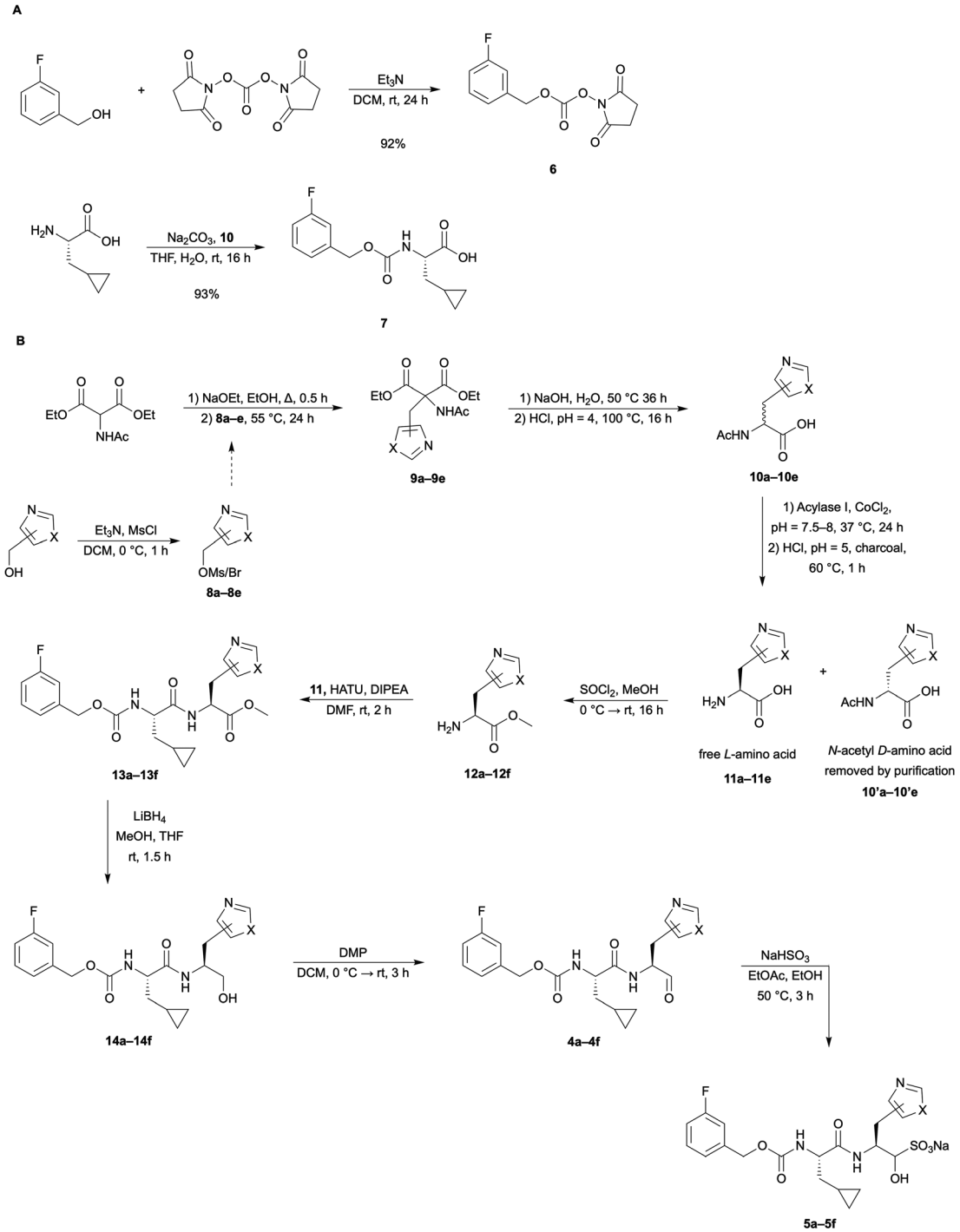
Results and discussion

The P1 surrogate unnatural amino acids were synthesized using robust methodology originally developed by Albertson and Archer in the 1940s, using diethyl acetamidomalonate (DEAM) as the glycine equivalent.[22] This method begins with the alkylation of DEAM with an electrophilic building block to form the side-chain of the target amino acid. For this, we primarily used the mesylates but discovered that in some cases it was preferable to use the alkyl chloride to provide better yields. We were able to purchase the oxazole and thiazole substituted methanol building blocks to start our synthetic route (Scheme 1). From the alcohols, we could easily convert them to a reactive electrophile using methanesulfonyl chloride or thionyl chloride to yield the mesylate or alkyl chloride in one step each. Alkylation of DEAM was done using sodium ethoxide in freshly distilled ethanol and the electrophilic oxazole or thiazole building block. The crude material was then purified and used as soon as possible, to obtain the best yields. In addition, the 5-oxazole and 5-thiazole mesylate derivatives were found to be the least stable of all. An alternative method was used, which was to convert the alcohols to alkyl chlorides, and use the *in situ* Finkelstein method in the alkylation of these electrophiles. This modified route improved the yields significantly (about 2-fold).

The alkylated DEAM product then underwent hydrolysis and decarboxylation (Scheme 1) to yield the racemic *N*-acetyl amino acids, which could then be resolved by enzymatic enantioselective hydrolysis by Acylase I.[23] Acylase I from *Aspergillus melleus* is a metalloenzyme that selectively removes the acetyl from the L-enantiomer only.[23] After the reaction time is complete, the enzyme was adsorbed by activated charcoal addition, and then through a series of centrifugation, decantation, and extraction we were able to remove the enzyme and charcoal.

The resulting mixture of free L-amino acid and *N*-acetyl amino acid was then separated using an ion-exchange column loaded with Dowex 50WX8 (cation exchange resin). The resin was equilibrated with 0.01 M HCl, and the sample was acidified to pH = 2. The column was eluted first using MilliQ water to remove any uncharged species, which was followed by the elution of the *N*-acetyl amino acid and the inorganic salts with 0.5 M HCl. Then using MilliQ water only, the remaining HCl was washed from the column, and finally the free L-amino acid was eluted with an aqueous 2% ammonia solution. Solvent evaporation of the pure product fraction resulted in the pure L-amino acid as a solid. This purification procedure worked well for the oxazolyl derivatives but caused co-elution problems for the thiazolyl analogues. The free L-amino acids were subjected to Marfey's assay to determine if the presumed stereochemistry was correct, and we were pleased to see that the procedure worked as expected with only the pure L-amino acid present (Fig. S1–S5).

The *N*-acetyl thiazolyl amino acids co-eluted with the free L-thiazolyl amino acids. This is possibly because the nitrogen in the thiazole ring was basic enough to be protonated on the resin and caused it to be retained. It was likely protonated in both the *N*-acetyl and free forms, causing the separation to be unsuccessful. Reversed-phase high-performance liquid chromatography (HPLC) was then attempted, and while we were able to optimize the methods to obtain some separation, the scale was limited to very small amounts, and we determined that the separation at this step was not completely necessary. Because the *N*-acetyl amino acid amine was protected with the acetyl group (now an amide), it would not be nucleophilic



Scheme 1 (A) Synthetic steps to intermediate **7**, which is used to couple with the synthesized unnatural amino acid methyl esters. (B) Full general synthetic route from DEAM to the final inhibitor compound showing a general thiazole or oxazole side chain. While the aldehyde maintains the stereochemistry in organic solvents, it can epimerize in water at the stereocenter at the aldehyde. The bisulfite is synthesized as a pair of diastereomers, with the epimerization at the bisulfite carbon shown. Slight variations were made to this route for some analogues. See synthetic procedures and SI for specific procedures.

enough to participate in a HATU-mediated amide bond formation, and we opted to proceed to the next steps with the mixture, with the understanding that only the free L-amino acid would react in the key coupling step.

The mixture of free L-thiazolyl amino acid and the *N*-acetyl thiazolyl amino acid or the pure free L-oxazolyl amino acid was subjected to thionyl chloride and methanol to convert the carboxylic acid to a methyl ester. The next step was the HATU-

mediated coupling with 7, which proceeded as predicted, and at this point, the *N*-acetyl amino acid methyl ester impurity was able to be mostly extracted out in the work-up step for the thiazoles and fully separated out by flash column chromatography over silica gel. Following the coupling and purification, the remaining steps were reduction of the methyl ester with lithium borohydride and oxidation of the resulting alcohol to obtain the aldehyde. Column purification here yielded one of the target compounds as a pure peptide aldehyde. Upon addition of sodium bisulfite the adduct of each peptide aldehyde was obtained after lyophilisation. This route gave access to five target aldehydes, and their respective bisulfite adducts (**4a–4e** and **5a–5e**, Fig. 2). Target candidate **4f** and **5f** were synthesised from a commercially available P1 building block following a slightly adapted procedure (see SI).

We used pure enzyme and a 10-amino acid Förster Resonance Energy Transfer (FRET) substrate to determine the IC_{50} values for each inhibitor. The FRET substrate is a linear sequence of amino acids that matches the natural recognition sequence of the main protease and contains a fluorophore (2-aminobenzoic acid) and a quencher (3-nitrotyrosine) moiety on either end of the peptide (Abz-SVTLQ↓SG-Y$^{(NO_2)}$-R).[24] The main protease recognizes the peptide sequence and cleaves it, separating the fluorophore and quencher, which causes fluorescence at a certain wavelength that we can monitor. Concentration dependent monitoring of fluorescence for each inhibitor allows us to calculate the IC_{50} for each compound. The IC_{50} values we found (Table 1) reveal activity differences for the different substitution patterns of the side chain and the warhead of the inhibitor.

In nearly all cases, the bisulfite compound (**5a–5f**) has a lower IC_{50} value than its aldehyde analogue (**4a–4f**), indicating that the pro-drug versions of the compounds provide some benefit over the aldehydes. The substitution pattern also appears to play a large role in the efficiency of the inhibitor. Comparing compound sets **4a–4c** and **5a–5c**, all of these are oxazoles with the only difference being where the methylene is attached to the ring. Compounds **4b** and **5b** (with the substitution at the 5-position) appear to have a much higher IC_{50} value (about 3–4 fold higher) than the others. This same pattern appears with the thiazoles, though the 5-thiazole performs even worse than the 5-oxazole. The 2- and 4-substituted rings do not have as large of a difference, but the data does show that the 2-substituted rings may be the best performing examples of the bisulfite set, with compounds **5a** and **5d** exhibiting the lowest IC_{50} values of about 290 and 270 nM, respectively. This trend is supported by the Morrison K_i values determined for the bisulfite compounds (Table 2). The lowest K_i of the oxazoles is found for the 2-substituted heteroaromatic ring (**5a**), and the same is true of the thiazoles (**5d**). The lowest IC_{50} value overall was found for **4c** with 190 nM, the 4-substituted oxazole aldehyde analogue. Because of the benefit of increased stability with the bisulfite, **5a** and **5d** would likely be preferable compounds, as the IC_{50} values are still very similar. In terms of oxazoles *versus* thiazoles, there does not seem to be a large difference, they exhibit similar values, although the oxazoles were more straightforward to synthesize, as they were easily purified with the ion-exchange column at the free amino acid stage.

We then used the best performing oxazole and thiazole bisulfite inhibitors, **5a** and **5d**, to test metabolic stability with cytochrome P450 3A4 (CYP3A4), with GC376 as a comparison. Each inhibitor was assayed in duplicate with controls to determine how much of the parent compound was remaining after 2 h of incubation with the oxidative enzyme (Table 3). GC376 was found to be the most stable with CYP3A4 of the three tested compounds. The oxazole, **5a** appeared to be more stable than the thiazole **5d**. Both **5a** and **5d** are similar to GC376, with more than 50% of the inhibitor still unoxidized after the 2 h incubation. In addition, we preformed liquid chromatography tandem mass spectrometry (LC-MS/MS) on the oxidized species detected by LC-MS (+16 *m*/*z* for the addition of an oxygen atom). Through analysis of the fragmentation patterns, we were able to determine that the oxidation was primarily taking place in the P1/warhead part of all three compounds. We have previously shown that oxidation of the aldehyde to the carboxylic acid occurs with the CYP3A4 enzyme, which may be occurring here as well.[10]

Table 1 IC_{50} values for each of the synthesized inhibitors against the SARS-CoV-2 main protease

Compound	IC_{50} (nM)	Compound	IC_{50} (nM)
4a	1220 ± 250	**5a**	290 ± 70
4b	20 200 ± 2400	**5b**	1300 ± 200
4c	190 ± 45	**5c**	420 ± 90
4d	590 ± 120	**5d**	270 ± 70
4e	12 200 ± 1800	**5e**	11 700 ± 2300
4f	500 ± 100	**5f**	380 ± 60

Table 2 Experimentally determined K_i values for each of the bisulfite inhibitors **5a–f**. See SI for procedure and graphical results (Fig. S9)

Compound	K_i (nM)
5a	250 ± 45
5b	970 ± 160
5c	350 ± 75
5d	185 ± 40
5e	4530 ± 850
5f	250 ± 70

Table 3 Comparison of remaining compound after incubation with purified CYP3A4 enzyme for 2 h. See SI for procedure

Compound	% of compound remaining after 2 h
GC376	79
5a	66
5d	53

Conclusions

Based on our data, 2-substituted oxazole and thiazole moieties outperform 4- and 5-substituted heteroazoles as chemical P1

surrogates in SARS-CoV-2 M^{pro} inhibitor design, as illustrated by compounds **5a** and **5d**. These inhibitors are comparable to other SARS-CoV-2 inhibitors in literature in terms of IC_{50} data (270–290 nM), metabolic stability, and present alternative chemical motifs from the commonly used cyclic glutamine derivatives (as in **1** and **2**). Because these new compounds stray from the commonly used scaffold of P1, to a novel P1 amino acid, these inhibitors may also have potential in the future, as an alternate therapeutic against mutants of the SARS-CoV-2 main protease that become immune to the current treatments. This approach adds to pandemic preparedness by expanding the chemical scope by usage of alternative scaffolds in the design of anti-coronaviral compounds.

Author contributions

Zijie Liu: investigation, visualization, writing (review and editing). Tayla J. Van Oers: investigation, writing (original draft, review, and editing). Conrad Fischer: investigation, supervision, writing (review and editing). Anna Nguyen: investigation. Johnson Chew: investigation. John C. Vederas: supervision, funding acquisition, conceptualization.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the supplementary information (SI). Supplementary information: Fig. S1–S13, NMR spectra and further experimental details. See DOI: https://doi.org/10.1039/d6ob00697c.

Ref. 25 is cited in the SI.

Acknowledgements

We would like to thank Nhat Pham for his assistance with IC_{50} testing of inhibitors. In addition, we thank Béla Reiz for his assistance with LC-MS/MS analysis. We gratefully acknowledge support by the Canadian Institutes of Health Research (CIHR) (J. C. V.: grant PS186126) and Alberta Ministry of Technology and Innovation through SPP-ARC (Striving for Pandemic Preparedness–The Alberta Research Consortium (J. C. V.)).

References

1 W. Vuong, M. B. Khan, C. Fischer, E. Arutyunova, T. Lamer, J. Shields, H. A. Saffran, R. T. McKay, M. J. van Belkum, M. A. Joyce, H. S. Young, D. L. Tyrrell, J. C. Vederas and M. J. Lemieux, *Nat. Commun.*, 2020, **11**, 4282.
2 Q. Hu, Y. Xiong, G. H. Zhu, Y. N. Zhang, Y. W. Zhang, P. Huang and G. B. Ge, *Med. Commun.*, 2022, **3**, 151.
3 J. R. Feys, K. Edwards, M. A. Joyce, H. A. Saffran, J. A. Shields, K. Garcia, D. L. Tyrrell and C. Fischer, *ACS Med. Chem. Lett.*, 2024, **15**, 2046–2052.
4 R. L. Hoffman, R. S. Kania, M. A. Brothers, J. F. Davies, R. A. Ferre, K. S. Gajiwala, M. He, R. J. Hogan, K. Kozminski, L. Y. Li, J. W. Lockner, J. Lou, M. T. Marra, L. J. Mitchell, B. W. Murray, J. A. Nieman, S. Noell, S. P. Planken, T. Rowe, K. Ryan, G. J. Smith, J. E. Solowiej, C. M. Steppan and B. Taggart, *J. Med. Chem.*, 2020, **63**, 12725–12747.
5 D. R. Owen, C. M. N. Allerton, A. S. Anderson, L. Aschenbrenner, M. Avery, S. Berritt, B. Boras, R. D. Cardin, A. Carlo, K. J. Coffman, A. Dantonio, L. Di, H. Eng, R. A. Ferre, K. S. Gajiwala, S. A. Gibson, S. E. Greasley, B. L. Hurst, E. P. Kadar, A. S. Kalgutkar, J. C. Lee, J. Lee, W. Liu, S. W. Mason, S. Noell, J. J. Novak, R. S. Obach, K. Ogilvie, N. C. Patel, M. Pettersson, D. K. Rai, M. R. Reese, M. F. Sammons, J. G. Sathish, R. S. P. Singh, C. M. Steppan, A. E. Stewart, J. B. Tuttle, L. Updyke, P. R. Verhoest, L. Wei, Q. Yang and Y. Zhu, *Science*, 2021, **374**, 1586–1593.
6 P. Chen, T. J. Van Oers, E. Arutyunova, C. Fischer, C. Wang, T. Lamer, M. J. van Belkum, H. S. Young, J. C. Vederas and M. J. Lemieux, *JACS Au*, 2024, **4**, 3217–3227.
7 A. C. Galasiti Kankanamalage, Y. Kim, A. D. Rathnayake, K. R. Alliston, M. M. Butler, S. C. Cardinale, T. L. Bowlin, W. C. Groutas and K. O. Chang, *J. Med. Chem.*, 2017, **60**, 6239–6248.
8 Y. L. Janin, *RSC Med. Chem.*, 2023, **15**, 81–118.
9 C. M. N. Allerton, J. T. Arcari, L. M. Aschenbrenner, M. Avery, B. M. Bechle, M. A. Behzadi, B. Boras, L. M. Buzon, R. D. Cardin, N. R. Catlin, A. A. Carlo, K. J. Coffman, A. Dantonio, L. Di, H. Eng, K. A. Farley, R. A. Ferre, S. S. Gernhardt, S. A. Gibson, S. E. Greasley, S. R. Greenfield, B. L. Hurst, A. S. Kalgutkar, E. Kimoto, L. F. Lanyon, G. H. Lovett, Y. Lian, W. Liu, L. A. Martínez Alsina, S. Noell, R. S. Obach, D. R. Owen, N. C. Patel, D. K. Rai, M. R. Reese, H. A. Rothan, S. Sakata, M. F. Sammons, J. G. Sathish, R. Sharma, C. M. Steppan, J. B. Tuttle, P. R. Verhoest, L. Wei, Q. Yang, I. Yurgelonis and Y. Zhu, *J. Med. Chem.*, 2024, **67**, 13550–13571.
10 T. J. Van Oers, A. Piercey, A. Belovodskiy, B. Reiz, B. L. Donnelly, W. Vuong, M. J. Lemieux, J. A. Nieman, K. Auclair and J. C. Vederas, *Org. Lett.*, 2023, **25**, 5885–5889.
11 D. Jochmans, C. Liu, K. Donckers, A. Stoycheva, S. Boland, S. K. Stevens, C. De Vita, B. Vanmechelen, P. Maes, B. Trüeb, N. Ebert, V. Thiel, S. De Jonghe, L. Vangeel, D. Bardiot, A. Jekle, L. M. Blatt, L. Beigelman, J. A. Symons, P. Raboisson, P. Chaltin, A. Marchand, J. Neyts, J. Deval and K. Vandyck, *mBio*, 2023, **14**, e02815–e02822.
12 C. Fischer, J. Lu, M. J. Van Belkum, S. Demmon, P. Chen, C. Wang, T. J. Van Oers, T. Lamer, M. J. Lemieux and J. C. Vederas, *RSC Med. Chem.*, 2025, **16**, 5032–5040.
13 E. Arutyunova, M. B. Khan, C. Fischer, J. Lu, T. Lamer, W. Vuong, M. J. van Belkum, R. T. McKay, D. L. Tyrrell,

J. C. Vederas, H. S. Young and M. J. Lemieux, *J. Mol. Biol.*, 2021, **433**, 167003.
14 W. Vuong, C. Fischer, M. B. Khan, M. J. van Belkum, T. Lamer, K. D. Willoughby, J. Lu, E. Arutyunova, M. A. Joyce, H. A. Saffran, J. A. Shields, H. S. Young, J. A. Nieman, D. L. Tyrrell, M. J. Lemieux and J. C. Vederas, *Eur. J. Med. Chem.*, 2021, **222**, 113584.
15 Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao and H. Yang, *Nature*, 2020, **582**, 289–293.
16 L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox and R. Hilgenfeld, *Science*, 2020, **368**, 409–412.
17 V. Kumar, J. Zhu, B. C. Chenna, Z. A. Hoffpauir, A. Rademacher, A. M. Rogers, C.-T. Tseng, A. Drelich, S. Farzandh, A. L. Lamb and T. D. Meek, *J. Am. Chem. Soc.*, 2025, **147**, 1631–1648.
18 Y. J. Lee, Y. Kim, H. Kim, J. Choi, G. H. Noh, K. S. Lee, J. Lee, C. H. Choi, S. H. Kim and J. Seo, *Inorg. Chem.*, 2023, **62**, 10279–10290.
19 D. Migulin, Y. Vysochinskaya, M. Buzin, A. Bakirov, G. Cherkaev and O. Shchegolikhina, *J. Photochem. Photobiol., A*, 2021, **407**, 113003.
20 C. N. Hsiao, M. R. Leanna, L. Bhagavatula, E. De Lara, T. M. Zydowsky, B. W. Horrom and H. E. Morton, *Synth. Commun.*, 1990, **20**, 3507–3517.
21 H. K. Chenault, J. Dahmer and G. M. Whitesides, *J. Am. Chem. Soc.*, 1989, **111**, 6354–6364.
22 N. F. Albertson and S. Archer, *J. Am. Chem. Soc.*, 1945, **67**, 308.
23 S.-C. J. Fu and S. M. Birnbaum, *J. Am. Chem. Soc.*, 1953, **75**, 918–920.
24 J. E. Blanchard, N. H. Elowe, C. Huitema, P. D. Fortin, J. D. Cechetto, L. D. Eltis and E. D. Brown, *Chem. Biol.*, 2004, **11**, 1445–1453.
25 P. Marfey, *Carlsberg Res. Commun.*, 1984, **49**, 591.