

Article

**Peptidomimetic Vinyl Heterocycle Inhibitors  
of Cruzain Effect Anti-Trypanosomal Activity**

BALA CHANDRA CHENNA, Linfeng Li, Drake Mellott, Xiang Zhai, Jair L Siqueira-Neto, Claudia Calvet Alvarez, Jean Bernatchez, Emily Desormeaux, Elizabeth Alvarez Hernandez, Jana Gomez, James McKerrow, Jorge Cruz-Reyes, and Thomas D. Meek

*J. Med. Chem.*, **Just Accepted Manuscript** • DOI: 10.1021/acs.jmedchem.9b02078 • Publication Date (Web): 03 Mar 2020

Downloaded from pubs.acs.org on March 4, 2020

**Just Accepted**

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.

**Peptidomimetic Vinyl-Heterocyclic Inhibitors of Cruzain Effect Anti-Trypanosomal Activity**

Bala C. Chenna,<sup>†,‡</sup> Linfeng Li,<sup>†,‡</sup> Drake M. Mellott,<sup>†</sup> Xiang Zhai,<sup>†,‡</sup> Jair L. Siqueira-Neto,<sup>§</sup> Claudia Clavet,<sup>§</sup> Jean Bernatchez,<sup>§</sup> Emily Desormeaux,<sup>†,‡</sup> Elizabeth Alvarez Hernandez,<sup>†</sup> Jana Gomez,<sup>†</sup> James H. McKerrow,<sup>§</sup> Jorge Cruz-Reyes,<sup>†</sup> and Thomas D. Meek<sup>†,\*</sup>

#These authors contributed equally

\*Corresponding author

<sup>†</sup>Department of Biochemistry & Biophysics, Texas A&M University, 301 Old Main Drive, College Station, TX 77843, United States

<sup>§</sup>Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California-San Diego, 9500 Gilman Drive, La Jolla, CA 92093, United States

**ABSTRACT**

Cruzain, an essential cysteine protease of the parasitic protozoan, *Trypanosoma cruzi*, is an important drug target for Chagas disease. We describe here a new series of reversible, but time-dependent, inhibitors of cruzain, composed of a dipeptide scaffold appended to vinyl-heterocycles meant to provide replacements for the irreversible reactive “warheads” of vinyl sulfone inactivators of cruzain. Peptidomimetic vinyl-heterocyclic inhibitors (PVHIs) containing Cbz-Phe-Phe/homoPhe- scaffolds with vinyl-2-pyrimidine, vinyl-2-pyridine, and vinyl-2-(N-methyl)-pyridine groups conferred reversible, time-dependent inhibition of cruzain ( $K_i^* = 0.1\text{-}0.4\ \mu\text{M}$ ). These cruzain inhibitors exhibited moderate to excellent selectivity vs. human cathepsins B, L, and S, and showed no apparent toxicity to human cells, but were effective in cell cultures of *Trypanosoma brucei brucei* ( $\text{EC}_{50} = 1\text{-}15\ \mu\text{M}$ ), and eliminated *Trypanosoma cruzi* in infected murine cardiomyoblasts ( $\text{EC}_{50} = 5\text{-}8\ \mu\text{M}$ ). PVHIs represent a new class of cruzain inhibitors that could progress to viable candidate compounds to treat Chagas disease and human sleeping sickness.

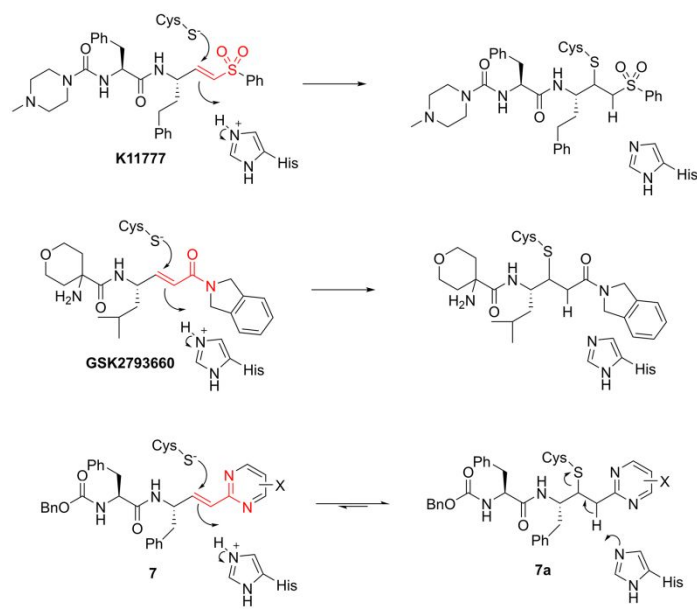
**INTRODUCTION**

Chagas disease (CD), caused by the American parasitic protozoan *Trypanosoma cruzi* (*T. cruzi*), affects nearly 8 million people in Mexico, Central and South America, resulting in 50,000 annual deaths.<sup>1-3</sup> It is estimated that more than 300,000 US residents are infected with *T. cruzi*, and cases have been reported in California, Texas and other parts of the Southwestern US in which the infections were acquired locally.<sup>4-6</sup> Approximately 20-30% of people infected with *T. cruzi* will develop debilitating and potentially fatal heart

dilation, arrhythmias, apical aneurysms,<sup>1</sup> and/or a dilated colon and esophagus.<sup>6</sup> The related African trypanosomes *Trypanosoma brucei rhodesiense* (*T. b. rhodesiense*) and *Trypanosoma brucei gambiense* (*T. b. gambiense*),<sup>7</sup> cause human sleeping sickness. Sleeping sickness is endemic in 36 African countries, threatening an estimated 60 million people, with more than 60,000 reported cases.<sup>8-10</sup> With the absence of vaccines, the standard treatment for Chagas disease is limited to nifurtimox and benznidazole which are so replete with side effects that abandonment of treatment is common.<sup>11-12</sup>

Cysteine proteases of the papain family are of similar structure<sup>13</sup> and substrate specificities,<sup>14-15</sup> and are implicated in numerous human diseases including cancer (cathepsin L),<sup>16</sup> chronic obstructive pulmonary disorder (cathepsins C<sup>17</sup> and S<sup>18</sup>) and osteoporosis (cathepsin K<sup>19</sup>). Likewise, the highly homologous, cathepsin L-like cysteine proteases, cruzain (E.C. 3.4.22.51; Uniprot ID: P25779),<sup>20-22</sup> rhodesain,<sup>23</sup> and brucipain (TbCatL),<sup>24</sup> found in the respective parasitic species *T. cruzi*, *T. b. rhodesiense*, and *T. b. brucei*, are essential to the establishment and maintenance of host infection by trypanosomes.<sup>20-24</sup> The roles of protozoal cysteine proteases in disease include the scavenging of iron from metalloproteins,<sup>24</sup> and the evasion of immune surveillance by proteolysis of NF- $\kappa$ B.<sup>25</sup> Gene deletion of cruzain demonstrated its role in undermining the prophylactic role of macrophages during infection.<sup>26</sup> The work of McKerrow and colleagues demonstrated the biochemical essentiality of cruzain in the pathology of *T. cruzi*.<sup>26-30</sup> They first expressed and characterized cruzain,<sup>21</sup> including its crystal structure,<sup>22</sup> and identified the drug candidate **K11777**, a dipeptide vinyl sulfone (**Figure 1**) which undergoes irreversible thia-Michael addition to active-site Cys<sub>25</sub>.<sup>20, 31-33</sup> The progression of **K11777** to human trials includes demonstration of a cure of *T. cruzi*-infected mammalian cells in culture,<sup>29</sup> and also elimination of parasites from acutely-infected mice.<sup>28</sup> However, **K11777** has stalled in pre-clinical evaluation due to liver toxicity, possibly arising from its irreversible mechanism.<sup>13, 34-35</sup>

Inactivators that form reversible covalent adducts with cysteine groups on enzymes have received recent attention.<sup>36-38</sup> Such reversible covalent inactivators demonstrate time-dependent inactivation, like irreversible inactivators, but may exert greater selectivity for the intended target rather than homologous “off-target” enzymes. This is because while the initial-collision complexes of irreversible inactivators with a panel of related enzymes may have variable affinities, over time the establishment of permanent covalent bonds may render this initial selectivity inconsequential. However, in the case of a reversible covalent inhibitor, their residence times on these enzymes are likely to be variable,<sup>38</sup> leading ultimately to “relief” from covalent inactivation for off-targets.



**Figure 1.** Structures of **K11777**, **GSK2793660**, and Cbz-Phe-Phe-vinyl-2Pyrm **7** with thia-Michael addition of Cys<sub>25</sub> to the vinyl groups in these compounds, including the putative adduct **7a**, which reverts to the fully conjugated **7** upon the reverse of adduct formation. The common bioisosteric atoms of **K11777**, **GSK2793660**, and **7** are highlighted in red. X is an electron-donating or electron-withdrawing substituent.

The design of dipeptide vinyl-heterocyclic inhibitors is predicated on existing irreversible covalent inactivators of cysteine proteases such as the vinyl sulfone of **K11777** and the acrylamide of **GSK27993660**, an inactivator of human cathepsin C that reached Phase I clinical trials.<sup>17</sup> Both compounds form irreversible covalent adducts with the active-site cysteines of the respective enzymes via a thia-Michael reaction (**Figure 1**), and both have encountered either toxicity issues or adverse events in, respectively, either animals<sup>39</sup> or humans.<sup>17</sup> Accordingly, we sought replacements of the vinyl sulfone and acrylamide “warheads” with less electrophilic moieties that would undergo *reversible* thia-Michael addition, in order to develop inactivators of high potency for cruzain, but also with suitable selectivity for trypanosomal over human cysteine proteases. One approach is the replacement of the vinyl sulfone and carboxamide group of the acrylamide with a bioisosteric heterocyclic group which is conjugated to the reactive vinyl group. A 2-vinyl pyrimidine is one such bioisosteric replacement (**7**, **Figure 1**), in which one of the ring nitrogens mimics the amide carbonyl and sulfone oxygen while the other substitutes for the amide nitrogen (red atoms in **Figure 1**). As the pyrimidine is conjugated to the vinyl group, the expected addition of the thiol group of the cysteine to the  $\beta$ -carbon of the vinyl group and attending protonation of the  $\alpha$ -carbon would eliminate this conjugation. Subsequent re-establishment of conjugation via reversal of the thia-Michael reaction would provide reversible covalent inactivation of the enzyme. Another feature afforded by this vinyl-heterocycle is the ability to modify the reactivity of the vinyl group by the substitution of the heterocycle with electron-donating or electron-withdrawing groups. In this study, we designed,

synthesized, and evaluated a panel of dipeptide compounds containing a vinyl group replacing the scissile amide group of the substrate, which is conjugated to a phenyl group or a collection of heterocycles. For some, we investigated the ability of glutathione to form covalent adducts with their vinyl groups, in order to explore the electronic nature of the heterocycle required for facile addition of thiols. Many of these compounds displayed potent, time-dependent inhibition of cruzain, as well as anti-trypansomal activity in cell culture.

## RESULTS AND DISCUSSION

**Evaluation of Dipeptide Substrates.** We previously conducted a detailed investigation of the catalytic mechanism of cruzain utilizing three dipeptide substrates, and by using solvent kinetic isotope effects in both pre-steady-state and steady-state kinetic modes.<sup>40</sup> As with other cysteine proteases, our kinetic studies supported a double-displacement chemical mechanism, in which a thioester intermediate of Cys<sub>25</sub> and substrate is first formed (enzyme acylation) concurrent with release of the amine product. This half-reaction is followed by enzyme-catalyzed hydrolysis of the thioester (enzyme de-acylation) to afford the carboxylate product. We determined that for Cbz-Phe-Arg-AMC (**S1**, **Table 1**), the most optimal substrate in terms of values of  $k_{\text{cat}}$  and  $k_{\text{cat}}/K_{\text{m}}$ , and Cbz-Arg-Arg-AMC (**S6**), enzyme-catalyzed de-acylation comprised the slow steps of the reaction ( $k_{\text{cat}} = k_{\text{dac}}$ ) while the rate of the acylation of active-site Cys<sub>25</sub> (described by  $k_{\text{ac}}$ ), was 10-fold and 4-fold, respectively, more rapid than  $k_{\text{cat}}$ .<sup>40</sup> The parameter  $k_{\text{cat}}/K_{\text{m}}$  contains the individual rate constants found in the expression for  $k_{\text{ac}}$ , and therefore reports on the rates of substrate binding and catalysis up to and including the acylation of active-site Cys<sub>25</sub>. Accordingly, this kinetic parameter comprises a useful guide for the selection of optimal dipeptide scaffolds to be incorporated into inhibitors that are meant to form reversible covalent complexes with Cys<sub>25</sub> upon binding, a reaction that chemically mimics acylation. If our dipeptidic vinyl-heterocyclic inhibitors do indeed inhibit cruzain by addition of the thiolate of Cys<sub>25</sub> to their olefinic replacements for the scissile amide bond, then the values of  $k_{\text{cat}}/K_{\text{m}}$  (the specificity constants) for the fluorogenic substrates should provide an approximate linear correlation with values of  $k_{\text{inact}}/K_{\text{I}}$  (which is the “specificity constant” of inactivation).<sup>41-42</sup>

We have kinetically characterized 12 dipeptide fluorogenic substrates of cruzain which are of the form: Cbz/NMePip-P<sub>2</sub>-P<sub>1</sub>-AMC, for which Cbz is benzyloxycarbonyl; NMePip is N-methyl-piperazinyl; P<sub>2</sub> = Phe, Leu, Arg or (4-Pyr)Ala (4-pyridyl-alanine); P<sub>1</sub> = Phe, hPhe (homoPhe), Arg, Ala and (4-Pyr)Ala; and AMC is 7-amino-4-methyl coumarin (**Table 1**). The largest specificity constant ( $k_{\text{cat}}/K_{\text{m}}$ ) measured was that of Cbz-Phe-Arg-AMC (11  $\mu\text{M}^{-1}\text{s}^{-1}$ , relative  $k_{\text{cat}}/K_{\text{m}} = 1.0$ ). Substrates **S2-S4** have values of  $k_{\text{cat}}/K_{\text{m}}$  which are 73%-35% of that of Cbz-Phe-Arg-AMC, indicating that Cbz-Phe-hPhe-AMC, Cbz-Leu-hPhe-AMC, and Cbz-Leu-Arg-AMC all comprise highly competent substrates likely effecting rapid acylation of

cruzain. The poorest substrate, Cbz-Phe-Ala-AMC (**S12**), demonstrated that a small sidechain in the P<sub>1</sub> position is less preferable than more bulky hydrophobic moieties. Conversely, the dipeptide Cbz-Phe-hPhe-AMC (**S2**) exhibited a value of  $k_{cat}/K_m = 8 \mu\text{M}^{-1}\text{s}^{-1}$ , indicating that a peptide substrate with the same P<sub>2</sub> and P<sub>1</sub> amino acids as **K11777** is an excellent substrate of cruzain, and likely involves rapid acylation. Interestingly, when the Cbz group of substrate **S2** is replaced with NMePip (**S7**), a substrate mimic of **K11777**,  $k_{cat}/K_m$  is 8-fold lower ( $1.3 \mu\text{M}^{-1}\text{s}^{-1}$ ). However, its value of  $k_{cat}$  exceeded that of Cbz-Phe-hPhe-AMC (**S2**), suggesting that the NMePip N-terminus of **S7** acts to retard cruzain acylation in comparison to the Cbz group of **S2**. While Cbz-Phe-Phe-AMC (**S9**) had a value of  $k_{cat}/K_m$  that was only 7% of Cbz-Phe-Arg-AMC, its  $K_m$  value of  $0.34 \mu\text{M}$  indicated potent binding, albeit, with slow turnover. This may indicate that a P<sub>1</sub> Phe substitution leads to favorable binding, but this sidechain impedes either the acylation or deacylation step in catalysis. This was observed for dipeptide-AMC substrates for the papain-like protease, human cathepsin C.<sup>43</sup>

Table 1. Kinetic Parameters of Peptide Substrates for Cruzain<sup>a</sup>

Substrate	Structure	$k_{cat} \text{ (s}^{-1}\text{)}$	$K_m \text{ (}\mu\text{M)}$	$k_{cat}/K_m \text{ (}\mu\text{M}^{-1}\text{s}^{-1}\text{)}$	$k_{cat}/K_m \text{ rel}$	$k_{cat} \text{ rel}$
<b>S1</b>	Cbz-Phe-Arg-AMC	$9.6 \pm 0.2$	$0.89 \pm 0.09$	$11 \pm 1$	1.00	1.00
<b>S2</b>	Cbz-Phe-hPhe-AMC	$2.0 \pm 0.2$	$0.26 \pm 0.06$	$8 \pm 2$	0.73	0.21
<b>S3</b>	Cbz-Leu-hPhe-AMC	$3.0 \pm 0.2$	$0.8 \pm 0.2$	$4 \pm 1$	0.36	0.21
<b>S4</b>	Cbz-Leu-Arg-AMC	$8.4 \pm 0.4$	$2.2 \pm 0.3$	$3.8 \pm 0.7$	0.35	0.88
<b>S5</b>	Cbz-Phe-(4-Pyr)Ala-AMC	$7.5 \pm 0.5$	$3.7 \pm 0.5$	$2.0 \pm 0.4$	0.18	0.76
<b>S6</b>	Cbz-Arg-Arg-AMC <sup>b</sup>	$7.2 \pm 0.1$	$3.7 \pm 0.4$	$1.9 \pm 0.2$	0.17	0.75
<b>S7</b>	NMePip-Phe-hPhe-AMC	$4.0 \pm 0.1$	$3.1 \pm 0.3$	$1.3 \pm 0.1$	0.12	0.42
<b>S8</b>	Cbz-(4-Pyr)Ala-hPhe-AMC	$5.3 \pm 0.2$	$4.6 \pm 0.5$	$1.1 \pm 0.1$	0.10	0.55
<b>S9</b>	Cbz-Phe-Phe-AMC	$0.26 \pm 0.01$	$0.34 \pm 0.06$	$0.8 \pm 0.2$	0.07	0.03
<b>S10</b>	Cbz-Arg-hPhe-AMC	$4.9 \pm 0.1$	$6.8 \pm 0.4$	$0.7 \pm 0.04$	0.06	0.51
<b>S11</b>	NMePip-Phe-Phe-AMC	$2.8 \pm 0.06$	$14 \pm 4$	$0.2 \pm 0.04$	0.02	0.29
<b>S12</b>	Cbz-Phe-Ala-AMC <sup>b</sup>	$0.89 \pm 0.06$	$38 \pm 2$	$0.023 \pm 0.003$	0.00	0.09

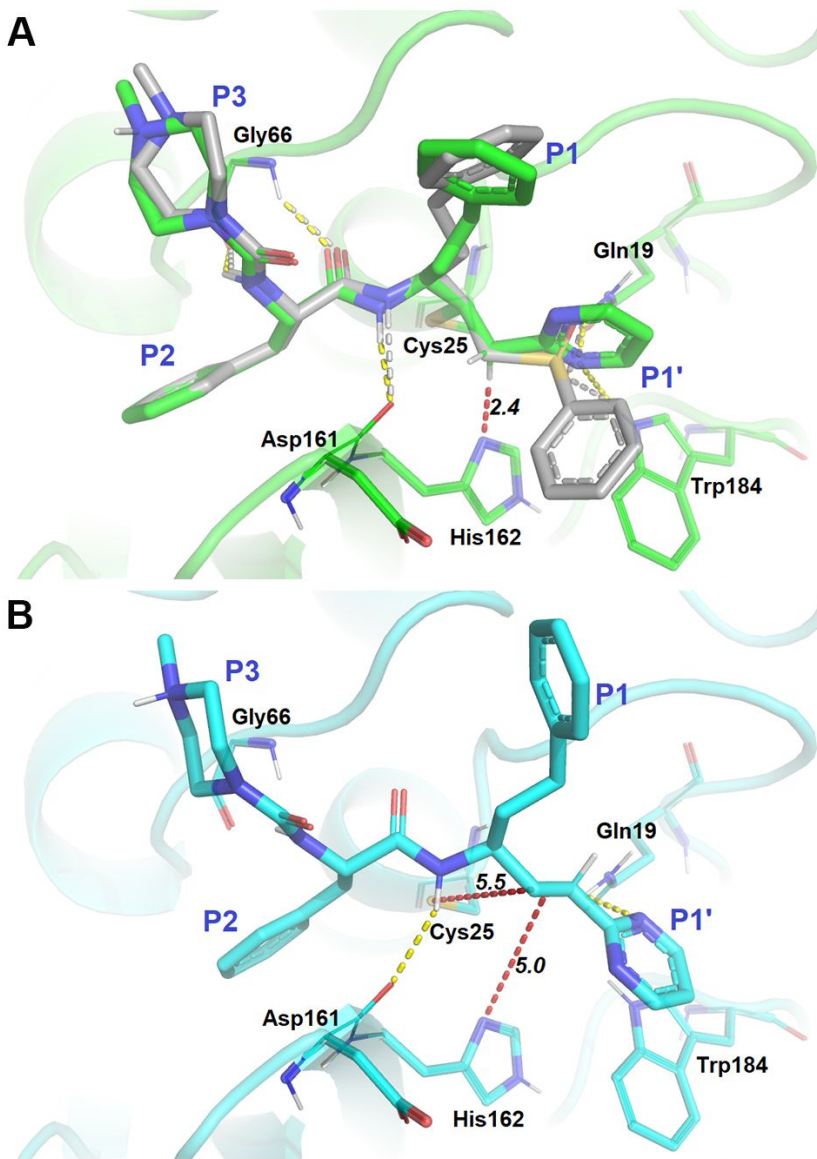
<sup>a</sup>Data obtained at pH 7.5, 25°C, 10% DMSO (v/v); <sup>b</sup>Data obtained at 2% DMSO (v/v) from Ref. 40.

We explored the viability of the (4-Pyr)Ala residue as a potential mimic of both Phe and Arg, for the latter residue when the pyridine is protonated. Cbz-Phe-(4-Pyr)Ala-AMC (**S5**) was found to be a good substrate (relative  $k_{cat}/K_m = 0.10$ ), but with a higher value of  $K_m$  ( $3.7 \mu\text{M}$ ) compared to substrates containing Phe and hPhe in the P<sub>1</sub> residue. To leverage the ability of cruzain to tolerate both charged basic and hydrophobic residues in the P<sub>2</sub> position, we prepared Cbz-(4-Pyr)Ala-hPhe-AMC and Cbz-Arg-hPhe-AMC (**S8** and **S10**). These substrates displayed efficient turnover numbers ( $k_{cat} = 5.3$  and  $4.9 \text{ s}^{-1}$ , respectively), but poor

values of  $K_m$  (4.6 and 6.8  $\mu\text{M}$ ) in comparison to Cbz-Phe-Phe/hPhe-AMC (0.26  $\mu\text{M}$ ). This demonstrates that the enzyme has a strong preference for substrates that contain a hydrophobic  $P_2$  residue. Together, these results indicated that dipeptide scaffolds in which  $P_2$  = Cbz-Phe or NMePip-Phe and  $P_1$  = Arg, hPhe, or (4-Pyr)Ala may be among the best to incorporate into the vinyl-heterocyclic framework. Accordingly, for our new peptidomimetic vinyl-heterocyclic inhibitors we have primarily utilized the Cbz-Phe-Phe and NMePip/Cbz-Phe-hPhe dipeptide scaffolds.

**Computer-Assisted Inhibitor Design.** To aid in the rational design of our vinyl-heterocyclic inhibitors we employed molecular docking of these compounds to a model constructed from the crystal structure of **K11777**-cruzain (PDB accession code: 2OZ2) which contains a covalent bond between the inactivator and Cys<sub>25</sub>.<sup>31</sup> Owing to our hypothesis that the vinyl-heterocyclic inhibitors have the ability to undergo a reversible thia-Michael addition with the active-site Cys<sub>25</sub> of cruzain, it is necessary to consider scenarios of both non-covalent and covalent binding. To this end, we first predicted the binding patterns for NMePip-Phe-hPhe-vinyl-2Pyrmid (**9**) which has the same scaffold as **K11777** using Glide<sup>44-46</sup> and CovDock<sup>47</sup> modules embedded in the Schrodinger software package. In the covalent model, the binding of **9** with cruzain was highly conserved when compared to that of **K11777** (**Figure 2A**). Pyrimidine N1 of **9** was within hydrogen bonding distance of Gln<sub>19</sub> and Trp<sub>184</sub>, allowing the stabilization of the vinyl-heterocycle in a nearly analogous fashion to the sulfone moiety in **K11777**. In addition, the  $\alpha$ -carbon of the inhibitor is positioned within 2.4 Å of His<sub>162</sub>, an interatomic distance that would easily allow facile proton transfer between this carbon and the imidazole nitrogen, supporting our hypothesis that a reversible adduct could be formed with cruzain. The non-covalent model (**Figure 2B**) shared similar shape complementarity with the covalent binding pose, except that it was slightly shifted away from the binding site as militated by the docking algorithm to avoid clashing with Cys<sub>25</sub>. This suggested that covalent bond formation would only slightly perturb the non-covalent binding conformation. Overall, these data suggested that the binding of our newly designed compounds containing a vinyl-heterocyclic warhead have the ability to interact with cruzain in a very similar fashion to the characterized, irreversible inactivators of the enzyme. Similarly, we carried out docking for five other PVH compounds (**7**, **11**, **12**, **13**, and **15** in **Figure S1**, and their covalent-docking affinity values (Cdock affinity) are summarized in **Table S1**. The corresponding inhibition constants (predicted  $K_i$ ) converted from these affinity values ranged from 0.79 to 6.1  $\mu\text{M}$ , and with the exception of compound **7**, were similar with the experimental values found in **Table 2**. In agreement with our dipeptide substrate kinetic data, we observed an increased Cdock affinity for **13** and **15** for which each contained a hPhe in the  $P_1$  position instead of Phe. In addition, the substitution of a pyridine ring at the  $P_1$  positions of **11**, **12**, **13**, and **15** may subtly improve the binding compared to the pyrimidine substituent of **7**. Further, the N-methylation of pyridine resulted in a fairly large shift in Cdock affinity, possibly resulting

from an additional ion-ion/dipole interaction. On the basis of these docking analyses, our inhibitor design focused on using a dipeptidic scaffold containing Phe in the P<sub>2</sub> position and either Phe or hPhe in the P<sub>1</sub> position. We varied the identity of the heterocycle in order to modify the electrophilicity of the olefin bond, but generally maintained functional groups that possibly could afford hydrogen bonding with Gln<sub>19</sub> to stabilize the binding of the compounds near Cys<sub>25</sub> of cruzain.



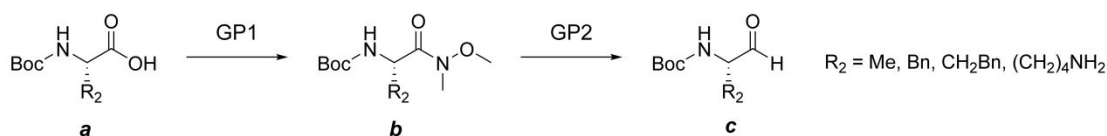
**Figure 2.** Molecular models of compound 9 bound to cruzain. (A) K11777 (gray) is superimposed with a binding pose (green) in which a covalent bond is formed between the  $\beta$ -carbon of the vinyl group of 9; (B) Binding pose (cyan) of 9 in which no covalent bond is formed with Cys<sub>25</sub>. Yellow and white dashed lines represent hydrogen bonds with surrounding residues for 9 and K11777, respectively. Red dashed lines are measurements between the catalytic dyad and vinyl moiety of 9.

**Synthesis of PVHIs.** The general synthetic routes employing either Wittig<sup>48</sup> or Horner-Wadsworth-Emmons<sup>49</sup> reactions shown in **Scheme 1** were used to synthesize peptidomimetic vinyl-heterocyclic compounds from aldehydes and halo-methyl heterocycles.

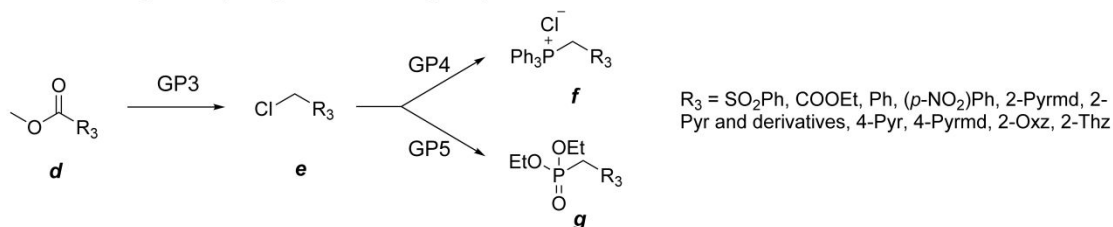
Commercially-available Boc-protected L-amino acids phenylalanine, homophenylalanine and alanine (**a**) were converted to Weinreb amides<sup>50</sup> by T3P-catalyzed coupling to *N,O*-dimethylhydroxylamine hydrochloride to afford **b** (GP1, General Procedure 1 in Experimental Section). Reduction of the Weinreb amide using LAH at -10 °C in anhydrous THF provided the Boc-amino acid aldehyde (**c**, GP2), generally in overall yields of ~80% (**a-c**).

### Scheme 1. General synthetic route to PVHIs<sup>a</sup>

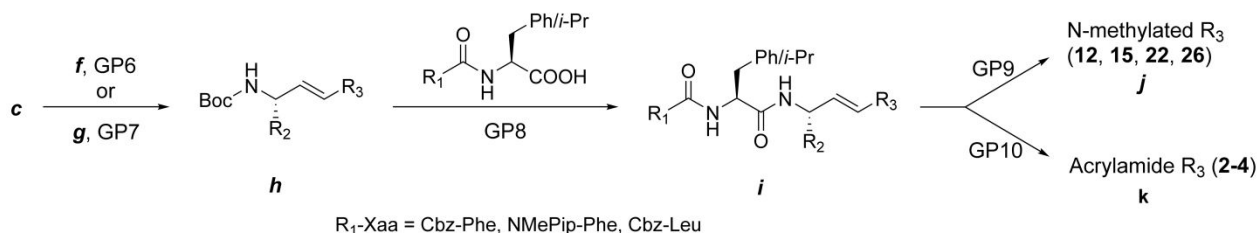
#### 1. P1 building block (Aldehyde)



#### 2. P1' building block (Wittig and HWE reagents)



#### 3. Coupling of building blocks



<sup>a</sup>Conditions and reagents (General Procedures are detailed in Experimental Section): GP1. *N,O*-dimethylhydroxylamine hydrochloride, T3P, DIPEA, DCM, 0 °C; GP2. LAH, THF, -10 °C; GP3. 1) NaBH<sub>4</sub>, EtOH, 0 °C, 2) SOCl<sub>2</sub>, DCM; GP4. PPh<sub>3</sub>, benzene, reflux; GP5. P(OEt)<sub>3</sub>, 150 °C; GP6. LHMDS, THF, -70 °C - 0 °C; GP7. LHMDS, THF, -70 °C - 0 °C; GP8. 1) TFA, DCM, 0 °C, 2) R<sub>1</sub>-Xaa-OH, T3P, DIPEA, DCM, 0 °C; GP9. MeI, MeCN, reflux; GP10. 1) LiOH, H<sub>2</sub>O, 2) ClCOOEt, NH<sub>4</sub>Cl.

Phosphonium salts of methyl-heterocycles were in general prepared by derivatization of either the 2-methylcarboxy- or 2-hydroxymethyl-heterocycle (**d** to **f**, **Scheme 1**). Methyl 2-carboxy-pyrimidine (or pyridine, oxazole, and thiazole) was reduced using sodium borohydride to the primary alcohol, followed by conversion of the alcohol to the 2-chloromethyl-pyrimidine (**e**) using SOCl<sub>2</sub> or POCl<sub>3</sub> in DCM or CHCl<sub>3</sub>.

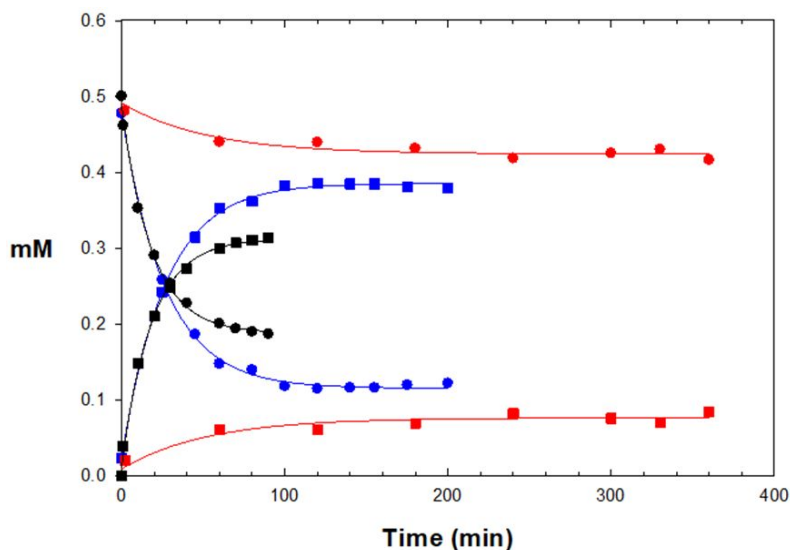
(GP3). Reaction of **e** with triphenylphosphine provided the Wittig reagent phosphonium salt (**f**) at overall yields of 28 - 80% (GP4). Wittig coupling of **f** with a peptide aldehyde (**c**) using LHMDS in anhydrous THF or sodium methoxide in benzene as base provided the peptide vinyl-heterocyclic product **h** (GP6), with general overall yields of 13 - 54%. Typically, the ratio of *E:Z* was 4:1, and separation of these regioisomers was readily achieved using silica gel column chromatography.

Alternatively, 2-chloromethyl-heterocycle **e** was converted to its phosphonate **g** by use of the Arbuzov reaction with triethyl-phosphite (~80% yields, GP5). The resulting phosphonate was de-protonated with LHMDS in THF, and then coupled with aldehyde **c** to provide the peptide vinyl-heterocycle **h** at 20% - 80% yield (GP7). The Boc group was removed quantitatively by treatment with TFA in DCM, then the free amine was coupled with the P<sub>3</sub>-P<sub>2</sub> fragment (R<sub>1</sub>-Xaa-OH) using T3P to give the inhibitor **i** (GP8).

In addition, some of the PVHIs underwent N-methylation of the heterocycle (**j**, GP9). Further, we also prepared several acrylamides (**k**) through hydrolysis of corresponding acrylate ester and subsequent treatment with ethyl chloroformate and NH<sub>4</sub>Cl (GP10). Final products were confirmed structurally by NMR and LCMS as described in Experimental Section and Supplementary Information. It is important to note that proton NMR analysis of the products (**i**, **j**, and **k**) indicated negligible epimerization at the  $\alpha$ -carbon in these products, as evidenced by the absence of diastereomers.

**Electrophilicity of Vinyl-Heterocycles.** To evaluate the chemical reactivity of the vinyl group in our PVHIs and **K11777** we treated selected compounds with glutathione (GSH) at pH 8.0 to determine their reactivity in a thia-Michael addition of the sulfhydryl group of GSH with the vinyl group of the inhibitors. Normally, the addition of glutathione to an enzyme inhibitor is to be avoided, but here this serves as a means to evaluate the electrophilicity of these inhibitors. **K11777** and compounds **7**, **11**, **12**, **15**, **17**, **25** and **26** (Table 2), which respectively contain a vinyl sulfone (**R**<sub>3</sub> = **I**), a vinyl-2-pyrimidine (**R**<sub>3</sub> = **IV**), a vinyl-2-pyridine (**R**<sub>3</sub> = **V**), a vinyl-2-N-methylpyridine (**R**<sub>3</sub> = **V**, **R**<sub>4</sub> = Me), a vinyl-2-(4-trifluoromethyl)-pyridine (**R**<sub>3</sub> = **V**, **R**<sub>5</sub> = CF<sub>3</sub>), a vinyl-2-thiazole (**R**<sub>3</sub> = **IX**), and a vinyl-2-N-methylthiazole (**R**<sub>3</sub> = **IX**, **R**<sub>4</sub> = Me). As previously reported,<sup>51</sup> the formation of a glutathione adduct with **K11777** was very slow ( $k = 0.00028\text{ s}^{-1}$ , Table S2, Figure 3), and we were unable to ascertain an equilibrium constant for the **K11777**-GSH adduct. For the PVHIs, the reaction between GSH and the vinyl-2-pyrimidine (**7**), the vinyl-pyridine (**11**), the vinyl-2-(4-trifluoromethyl)-pyridine (**17**), and the vinyl-thiazole (**25**), was negligible as no adduct was observed after 90 minutes incubation with either a 2:1 or 10:1 molar ratio of GSH:inhibitor. The electron-withdrawing 4-trifluoromethyl group on the pyridine of **17** had no effect on the electrophilicity of **11**. In contrast, addition of GSH to the vinyl group of vinyl-2-N-methylpyridine (**12**, **15**) and vinyl-2-N-methylthiazole (**26**) in a 2:1 molar ratio resulted in the rapid formation of GSH adducts at respective rates of  $0.037\text{ mM}^{-1}\text{s}^{-1}$ ,  $0.054\text{ mM}^{-1}\text{s}^{-1}$ , and  $0.015\text{ mM}^{-1}\text{s}^{-1}$ , and apparent equilibrium was achieved for these

compounds in 90 min ( $K_{eq} = 7400 \text{ M}^{-1}$ ,  $2400 \text{ M}^{-1}$ , and  $930 \text{ M}^{-1}$ , respectively). This demonstrated that the N-methylation of the PVHs afforded a significant increase in the electrophilicity of the vinylic position, enabling rapid addition to thiols, owing to the strong electron-withdrawing effect of the methylpyridinium moiety. For example, compounds **11** and **12** are identical except for the N-methylpyridine group of compound **12**; compound **12** readily forms an adduct with GSH (97% conversion of **12** to its GSH adduct in 20 min at a 10:1 molar ratio of GSH to compound), whereas compound **11** is unreactive towards GSH. Interestingly, the rate of thiolation of compound **12** is 50% that of **15**, while the values of  $K_{eq}$  indicated that the **12**-GSH adduct is three times more abundant than that of **15**-GSH. This suggested that the phenylalanyl sidechain of **12** may retard the addition of GSH to its vinyl group, and also slowed the presumed base-catalyzed elimination of GSH from its adduct with **12**. Overall, these results demonstrated that the reactivity of vinyl-heterocycles with GSH and presumably Cys<sub>25</sub> vary with the nature of the heterocycles. Hence, it is possible to tune the electronic properties of the vinyl bond in the PVHIs, thereby allowing for development of modifiable electrophilic inhibitors of other enzymes that have an active-site cysteine or other nucleophile.

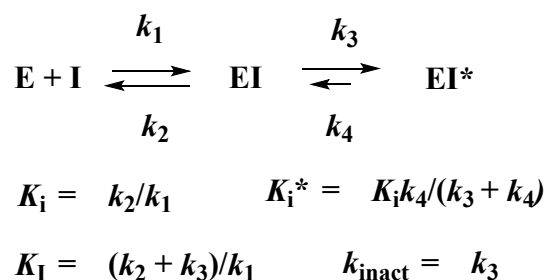


**Figure 3.** Time courses of depletion of 0.5 mM of **K11777** (red circles), compound **12** (blue circles), and compound **15** (black circles) upon formation of adducts (red, blue, and black squares, respectively) with 1 mM glutathione. Lines drawn through the curves for substrate depletion and adduct formation, were respectively,  $[\text{Substrate}] = (0.5 \text{ mM} - A)(1 - \exp(-k \cdot t)) + C$  and  $[\text{Adduct}] = A(1 - \exp(-k \cdot t)) + C$ , with resulting kinetic parameters found in **Table S3**.

**Kinetic Analysis of Cruzain Inhibitors and Inactivators.** Scheme 2 is a kinetic depiction of inhibition and inactivation of cruzain and the relevant kinetic parameters.<sup>52</sup> The initial, and usually rapid, formation of **EI** is characterized by the inhibition constant  $K_i$ . For time-dependent inhibitors **EI** progresses to a second, tighter complex **EI\***, generally over the course of minutes, characterized by  $K_i^*$ , for which  $K_i^* < K_i$  when

$k_4 < k_3$ . For irreversible covalent inactivators,  $k_4$  and  $K_i^* \sim 0$ , and the kinetic parameter  $k_{\text{inact}}/K_I$  is generally reported. For reversible time-dependent inhibitors, initiation of reaction by adding enzyme to substrate and inhibitor leads to concave-downward, curvilinear time courses of product formation in which reaction rates demonstrably decrease as the  $\text{EI}^*$  complex forms. Typical data, as exemplified for compound **15**, are shown in **Figure 4A**. Alternatively, extended pre-incubation of enzyme and inhibitor, followed by dilution of the inhibitor and initiation of reaction with high concentrations of substrate, leads to concave-upward curvilinear plots of product formation as  $\text{E}$  reforms from  $\text{EI}^*$  (**Figure 4B**). Results of this analysis for cruzain inhibitors and inactivators are collected in **Table 2**.

**Scheme 2. Kinetic depiction of inhibition and inactivation.**

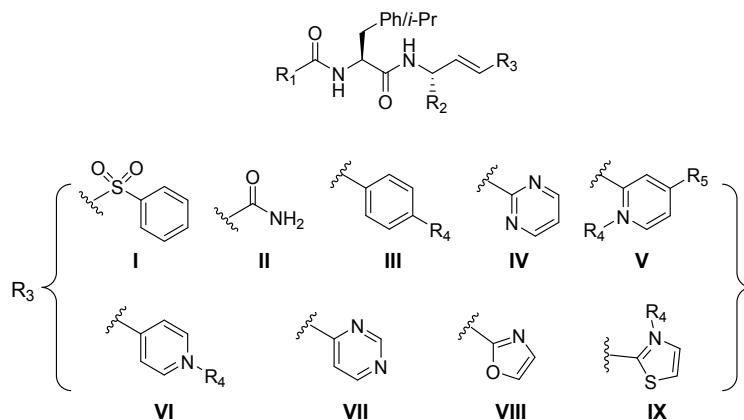


**K11777** comprises a useful benchmark compound despite the fact that it is an irreversible inactivator of cruzain (reported kinetic data: apparent  $\text{IC}_{50}$  of 2 nM,  $k_{\text{inact}}/K_I = 234,000 \text{ M}^{-1}\text{s}^{-1}$ ).<sup>32,53</sup> We replaced the  $\text{P}_1$  hPhe group of **K11777** with a Phe sidechain to provide vinyl sulfone **1**, which had apparently equivalent potency ( $K_i^* = 3.6 \text{ nM}$ ) to that of **K11777**, but which, interestingly, exhibited kinetically reversible inhibition of cruzain. However, a crystal structure we obtained for **1** bound to cruzain indicated the formation of a C-S bond between  $\text{Cys}_{25}$  and **1** (**Figure S4**). This may indicate that a phenylalanyl group at the  $\text{P}_1$  position partly impedes the ability of an adjacent vinyl electrophile to access  $\text{Cys}_{25}$ , as was observed with the solution phase GSH addition studies to our PVH compounds.

We next evaluated three C-terminal acrylamides ( $\text{R}_3 = \text{II}$ ) within the Cbz-Phe-Phe, Cbz-Phe-hPhe, and NMePip-Phe-hPhe scaffolds (**2-4**). The acrylamides within the Cbz-Phe-hPhe and NMePip-Phe-hPhe scaffolds afforded apparently irreversible covalent inactivation ( $k_{\text{inact}}/K_I = 1700\text{-}1900 \text{ M}^{-1}\text{s}^{-1}$ ), while Cbz-Phe-Phe-acrylamide (**2**) was less effective ( $k_{\text{inact}}/K_I = 22 \text{ M}^{-1}\text{s}^{-1}$ ). Comparing the values of  $k_{\text{inact}}/K_I$  for **K11777** and **4** indicated that the vinyl sulfone is overwhelmingly more effective as a covalent inactivator than its acrylamide counterpart, possibly owing to hydrogen bond contacts of the sulfone oxygen with  $\text{Gln}_{19}$ , which position the vinyl group proximal to  $\text{Cys}_{25}$  of cruzain. As with **1**, a Phe rather than a hPhe group at

the P<sub>1</sub> position, may retard covalent formation over the time course of kinetic analysis when one compares the rates of apparent inactivation of **2** vs. **3** and **4**, as was also seen with peptide substrates.

**Table 2. Kinetic Data of Peptidomimetic Vinyl-Heterocyclic Inhibitors of Cruzain<sup>a</sup>**



Compound	Structure	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	Cruzain Inhibition or Inactivation		
					K <sub>i</sub> (μM)	K <sub>i</sub> <sup>*</sup> (μM)	k <sub>inact</sub> /K <sub>inact</sub> (M <sup>-1</sup> s <sup>-1</sup> )
<b>K11777</b>	NMePip-Phe-hPhe-VSPH	NMePip	CH <sub>2</sub> Bn	<b>I</b>	NA	0.002 <sup>b</sup>	234,000 <sup>b,c</sup>
<b>1</b>	Cbz-Phe-Phe-VSPH	BnO	Bn	<b>I</b>	ND	0.0036 ± 0.0001	ND
<b>2</b>	Cbz-Phe-Phe-vinyl-CONH <sub>2</sub>	BnO	Bn	<b>II</b>	37 ± 2	NA	21.7 ± 0.8
<b>3</b>	Cbz-Phe-hPhe-vinyl-CONH <sub>2</sub>	BnO	CH <sub>2</sub> Bn	<b>II</b>	3 ± 1	NA	1700 ± 500
<b>4</b>	NMePip-Phe-hPhe-vinyl-CONH <sub>2</sub>	NMePip	CH <sub>2</sub> Bn	<b>II</b>	3.4 ± 0.4	NA	1900 ± 200
<b>5</b>	Cbz-Phe-Phe-vinyl-Ph	BnO	Bn	<b>III</b>	1.8 ± 0.1	0.87 ± 0.05	NA
<b>6</b>	Cbz-Phe-Phe-vinyl-(4-NO <sub>2</sub> )Ph	BnO	Bn	<b>III, R<sub>4</sub> = NO<sub>2</sub></b>	ND	0.37 ± 0.02	NA
<b>7</b>	Cbz-Phe-Phe-vinyl-2Pyrmd	BnO	Bn	<b>IV</b>	28 ± 1	0.364 ± 0.004	NA
<b>8</b>	Cbz-Phe-hPhe-vinyl-2Pyrmd	BnO	CH <sub>2</sub> Bn	<b>IV</b>	>35	NA	NA
<b>9</b>	NMePip-Phe-hPhe-vinyl-2Pyrmd	NMePip	CH <sub>2</sub> Bn	<b>IV</b>	>10	2.2 ± 0.1	NA
<b>10</b>	Cbz-Phe-Ala-vinyl-2Pyrmd	BnO	Me	<b>IV</b>	58 ± 6	25 ± 1	NA
<b>11</b>	Cbz-Phe-Phe-vinyl-2Pyr	BnO	Bn	<b>V</b>	5.5 ± 0.4	0.31 ± 0.01	NA
<b>12</b>	Cbz-Phe-Phe-vinyl-2PyrNMe	BnO	Bn	<b>V, R<sub>4</sub> = Me</b>	3.8 ± 0.4	0.28 ± 0.08	NA
<b>13</b>	Cbz-Phe-hPhe-vinyl-2Pyr	BnO	CH <sub>2</sub> Bn	<b>V</b>	1.06 ± 0.07	0.171 ± 0.004	NA
<b>14</b>	NMePip-Phe-hPhe-vinyl-2Pyr	NMePip	CH <sub>2</sub> Bn	<b>V</b>	ND	3.4 ± 0.1	NA
<b>15</b>	Cbz-Phe-hPhe-vinyl-2PyrNMe	BnO	CH <sub>2</sub> Bn	<b>V, R<sub>4</sub> = Me</b>	0.76 ± 0.04	0.126 ± 0.004	NA
<b>16</b>	Cbz-Phe-hPhe-vinyl-2-(4-OMe)-Pyr	BnO	CH <sub>2</sub> Bn	<b>V, R<sub>5</sub> = OMe</b>	>5	NA	NA
<b>17</b>	Cbz-Phe-hPhe-vinyl-2-(4-CF <sub>3</sub> )-Pyr	BnO	CH <sub>2</sub> Bn	<b>V, R<sub>5</sub> = CF<sub>3</sub></b>	NA	0.57 ± 0.05	NA
<b>18</b>	Cbz-Leu-hPhe-vinyl-2Pyr	BnO	CH <sub>2</sub> Bn	<b>V</b>	7.8 ± 0.6	1.42 ± 0.09	NA
<b>19</b>	Cbz-Phe-Ala-vinyl-2Pyr	BnO	Me	<b>V</b>	ND	4.8 ± 0.2	NA
<b>20</b>	Cbz-Phe-Lys-vinyl-2Pyr	BnO	(CH <sub>2</sub> ) <sub>4</sub> NH <sub>2</sub>	<b>V</b>	17.3 ± 0.3	0.87 ± 0.02	NA
<b>21</b>	Cbz-Phe-Phe-vinyl-4Pyr	BnO	Bn	<b>VI</b>	ND	5.5 ± 0.2	NA
<b>22</b>	Cbz-Phe-Phe-vinyl-4PyrNMe	BnO	Bn	<b>VI, R<sub>4</sub> = Me</b>	92 ± 5	4.0 ± 0.1	NA
<b>23</b>	Cbz-Phe-Phe-vinyl-4Pyrmd	BnO	Bn	<b>VII</b>	10.8 ± 1.4	1.14 ± 0.07	NA
<b>24</b>	Cbz-Phe-Phe-vinyl-2Oxz	BnO	Bn	<b>VIII</b>	10 ± 1	0.71 ± 0.01	NA
<b>25</b>	Cbz-Phe-Phe-vinyl-2Thz	BnO	Bn	<b>IX</b>	ND	1.71 ± 0.09	NA
<b>26</b>	Cbz-Phe-Phe-vinyl-2ThzNMe	BnO	Bn	<b>IX, R<sub>4</sub> = Me</b>	ND	0.94 ± 0.06	NA

<sup>a</sup> Data obtained at 25°C, pH 7.5; <sup>b</sup> Reported as apparent IC<sub>50</sub> in Ref. 32; <sup>c</sup> Reported as 32,500 M<sup>-1</sup> s<sup>-1</sup> (pH 8.0) in Ref. 53; NA, not applicable; ND, not determined; app *K*<sub>i</sub> and *K*<sub>i</sub>\* are respectively, the apparent initial and tight-binding inhibition constants.

We therefore sought to explore the effects of replacement of both the vinyl-phenylsulfone and acrylamide groups with a phenyl and heterocyclic groups conjugated to the vinyl group. The Cbz-Phe-Phe-vinyl-benzene compound **5** is a time-dependent inhibitor of cruzain ( $K_i^* = 0.87 \mu\text{M}$ ), but substitution of the *para* position of the phenyl ring with an electron-withdrawing nitro group (compound **6**) led to a nearly 3-fold improvement in potency ( $K_i^* = 0.34 \mu\text{M}$ ), suggesting that the vinyl group of **6** is more capable of thiolation by the cruzain. As seen with **1**, these compounds also demonstrated reversible inhibition of cruzain, possibly due to the P<sub>1</sub> phenylalanine. Due to poor aqueous solubility (solubility of **5** and **6**  $\leq 2 \mu\text{M}$  in 10% DMSO), the inhibitors containing vinyl-benzene were not explored further.

Subsequently, six heterocyclic groups (**R**<sub>3</sub> = **IV-IX**) conjugated to the presumed electrophilic vinyl group were evaluated within several dipeptide scaffolds. The vinyl-2-pyrimidine (**R**<sub>3</sub> = **IV**), vinyl-2-pyridine (**R**<sub>3</sub> = **V**), vinyl-2-oxazole (**R**<sub>3</sub> = **VIII**), and vinyl-2-thiazole (**R**<sub>3</sub> = **IX**) groups, unlike the vinyl-4-pyridine (**R**<sub>3</sub> = **VI**) and vinyl-4-pyrimidine (**R**<sub>3</sub> = **VII**), maintain bioisosteric similarity to the reactive acrylamides and vinyl sulfones, which is reflected in their more potent inhibition of cruzain as detailed below. Most of these compounds induced time-dependent inhibition on cruzain and were found to be kinetically reversible with residence times ( $\tau$ ) of 6-20 minutes.

The vinyl-2-pyrimidine moiety (**R**<sub>3</sub> = **IV**) in the Cbz-Phe-Phe scaffold afforded compound **7**, which exerted time-dependent inhibition of cruzain with an initial value of  $K_i = 5 \mu\text{M}$  and subsequent tight-binding inhibition of  $K_i^* = 0.38 \mu\text{M}$ . Substitution of the phenyl group of **5** by a pyrimidine group greatly improved the solubility of **7** ( $\geq 100 \mu\text{M}$  in 10% DMSO). Extended pre-incubation with **7**, followed by dilution, and addition of an excess of substrate, resulted in slow recovery of cruzain activity, indicating that any covalent reaction between cruzain and **7** was kinetically reversible ( $k_4 = 0.0018 \pm 0.0003 \text{ s}^{-1}$ ;  $\tau = 9 \text{ min}$ ). Interestingly, when the 2-pyrimidinyl moiety is appended to Cbz-Phe-hPhe (**8**), the resulting compound is a poor inhibitor of cruzain ( $K_i > 35 \mu\text{M}$ ); however, when the 2-pyrimidinyl group is attached to afford the same scaffold as **K11777**, we obtained an inhibitor of low micro-molar potency (**9**,  $K_i^* = 2.2 \mu\text{M}$ ). Substitution of the P<sub>1</sub> Phe with Ala (**10**,  $K_i = 25 \mu\text{M}$ ) produced a poor inhibitor of cruzain, indicating the essentiality of a larger sidechain in the P<sub>1</sub> position, as was observed with dipeptide substrates. To probe the importance of the vinyl group for the inhibition of cruzain, we prepared an analogue in which the vinyl group of **7** was reduced (compound **27**). This inhibitor lacked time-dependent behavior ( $K_i = 22 \mu\text{M}$ ), and was 100-fold less potent than its vinyl analogue **7**, which demonstrated the importance of the vinyl group for the inhibition of cruzain. We prepared inhibitor **23** which contains a vinyl-4-pyrimidinyl (**R**<sub>3</sub> = **VII**) group that does not maintain bioisosteric similarity to the acrylamides. **23** exhibited three-fold less potency than the bioisosteric

vinyl-2-pyrimidine (**7**). Similarly, inhibitors containing the vinyl-4-pyridyl ( $R_3 = VI$ ) (**21**, **22**) lack bioisosteric equivalence to the acrylamides and were found to be only modest inhibitors of cruzain.

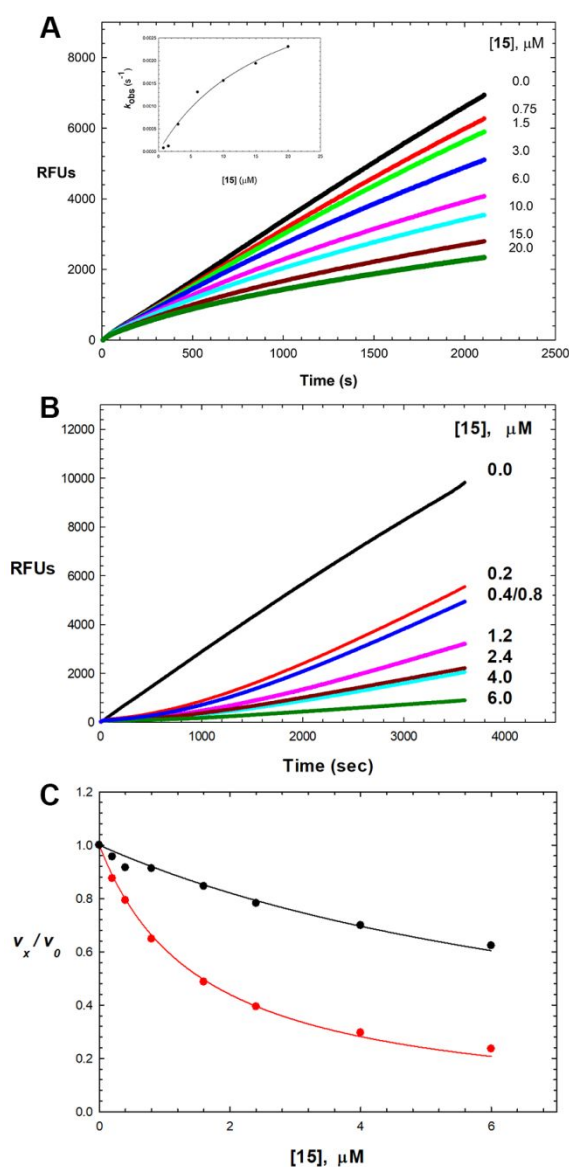
Inhibitors containing a vinyl-2-pyridinyl group ( $R_3 = V$ ) were explored more widely. Cbz-Phe-Phe-vinyl-2-pyridine **11** exhibited time-dependent inhibition of cruzain with an initial value of  $K_i = 5.5 \mu M$  and subsequent tight-binding inhibition of  $K_i^* = 0.31 \mu M$  ( $k_4 = 0.0012 \pm 0.0002 s^{-1}$ ;  $\tau = 13$  min), and solubility of **11** was  $\leq 30 \mu M$  in 10% DMSO. Unlike the vinyl-pyrimidinyl group of **7**, placement of the vinyl-2-pyridinyl group in the Cbz-Phe-hPhe scaffold improved inhibition by 3-fold (**13**,  $K_i^* = 0.17 \mu M$ ), while the vinyl-2-pyridinyl group was much less effective in the NMePip-Phe-Phe scaffold (**14**,  $K_i^* = 3.4 \mu M$ ). Substitution of an electron-donating methoxy group on the pyridine ring (**16**) of the Cbz-Phe-hPhe scaffold diminished the inhibitory activity of the vinyl-2-pyridinyl heterocycle compared to its unsubstituted counterpart **13** by  $>50$ -fold, suggesting that the methoxy group is large enough to create a steric barrier to inhibitor binding. In contrast, the substitution at C-4 of the pyridine with the electron-withdrawing trifluoromethyl group resulted in better inhibition (**17**,  $K_i^* = 0.57 \mu M$ ), but nonetheless was less potent than the unsubstituted pyridine **13**. Apparently, this result arises from steric crowding as  $OMe > CF_3 > H$ , implicating that substitution at the C-4 position of the pyridine heterocycles are not well tolerated.

We next investigated how the  $P_1$  and  $P_2$  sidechains of these vinyl-2-pyridinyl inhibitors effect inhibition. The replacement of the  $P_2$  Phe with Leu resulted in diminished potency (**18**,  $K_i^* = 1.42 \mu M$ ) compared to the Cbz-Phe-Phe and Cbz-Phe-hPhe scaffolds, overall demonstrating that inhibitors with bulky hydrophobic substituents in  $P_1$  and  $P_2$  enhanced binding to cruzain. To analyze how short alkyl and charged groups effected inhibition, we prepared Cbz-Phe-Ala-vinyl-2-pyridine (**19**) and Cbz-Phe-Lys-vinyl-2-pyridine (**20**). We found that the Cbz-Phe-Lys scaffold, which mimics our most optimal substrate, Cbz-Phe-Arg-AMC, exhibited good inhibition ( $K_i^* = 0.87 \mu M$ ), whereas **19** was a poor inhibitor ( $K_i^* = 4.8 \mu M$ ), in concert with the poor substrate activity of Cbz-Phe-Ala-AMC.

Seeking to improve the electrophilicity of the vinyl-2-pyridinyl group, we prepared N-methylated analogues **12** and **15**. This modification resulted in improved aqueous solubility ( $\geq 50 \mu M$  in 10% DMSO), and provided potent time-dependent inhibition of cruzain (**12**,  $K_i^* = 0.28 \mu M$ ; **15**,  $K_i^* = 0.126 \mu M$ ) comparable to, or exceeding, the inhibition exerted by their un-methylated counterparts (**11** and **13**).

Inhibition data for compound **15** were fitted by all methods outlined in Experimental Section, and as shown in **Figure 4**. We fitted each curve in **Figure 4A** to eq 3, and the resulting values of  $k_{obs}$  were re-plotted vs. [**15**] (**Figure 4A**, inset), which demonstrated a hyperbolic dependence of the inhibitor (fitting to eq 4:  $K_i = 2.0 \pm 0.9 \mu M$ ,  $k_3 = 0.004 \pm 0.001 s^{-1}$ , and  $k_4 \sim 0$ ). Alternatively, global fitting of these curves to eq 6 provided values of  $K_i = 4.3 \pm 0.1 \mu M$ ,  $k_3 = 0.0012 \pm 0.0004 s^{-1}$ , and  $k_4 = 0.00019 \pm 0.00005 s^{-1}$ , from which was

calculated a value of  $K_i^* = 0.6 \pm 0.2 \mu\text{M}$ . Fitting of the data globally for **15** and other potent PVHIs using Kintek Explorer is found in Supplementary Information. Pre-incubation of cruzain and variable concentrations of **15**, followed by initiation of reaction by the addition of substrate, produced time courses like that shown in **Figure 4B**. These data demonstrated a significant lag phase for recovery of cruzain activity, indicative of the slow desorption of the inhibitor, with or without the formation of a covalent bond with Cys<sub>25</sub>. Finally, analysis of inhibition of cruzain by **15** at early and late phases of the time courses in **Figure 4A** by fitting to eq 5 provided values of  $K_i = 0.76 \pm 0.04 \mu\text{M}$  and  $K_i^* = 0.126 \pm 0.004 \mu\text{M}$  (**Table 2**). Of note, in pre-incubation studies, all PVHIs which contain the vinyl-pyridinyl substituent displayed kinetic reversibility.



**Figure 4.** Time-dependent inhibition of cruzain by **15**. (A) Reaction initiated by addition of cruzain (100 pM) with Cbz-Phe-Arg-AMC (10  $\mu$ M) and 0-20  $\mu$ M **15** (pH 7.5). Lines drawn through the experimental data points were from fitting of each inhibitor concentration to eq 3, from which the replot of  $k_{\text{obs}}$  vs. [**15**] is shown in the inset (fitting to eq 4:  $K_i = 2.00 \pm 0.9 \mu\text{M}$ ,  $k_3 = 0.004 \pm 0.001 \text{ s}^{-1}$ , and  $k_4 \sim 0$ ); (B) Following 1 hr pre-incubation of cruzain (100 pM) with 0-6  $\mu$ M **15**, reaction was initiated by addition of Cbz-Phe-Arg-AMC (10  $\mu$ M). (C) Fitting of cruzain inhibition by compound **15** for  $v_i/v_0$  (black) and  $v_s/v_0$  (red) using eq 5 with results of this found in **Table 2**.

We investigated 5-membered ring heterocycles that are bioisosteric with acrylamide inactivators. The syntheses of the vinyl-2-oxazole (**24**), the vinyl-2-thiazole (**25**) and its N-methylated counterpart (**26**) into the Cbz-Phe-Phe scaffold proved facile, and provided useful inhibitors. Vinyl-2-oxazole **24** was a sub-micromolar inhibitor of cruzain ( $K_i^* = 0.71 \mu\text{M}$ ). Vinyl-2-thiazole inhibitors **25** and **26** were inhibitors of similar potency ( $K_i^* = 1.71$  and  $0.94 \mu\text{M}$ , respectively), for which N-methylation of the thiazole improved potency by nearly two-fold.

Cruzain inhibitors **5-26** allowed the evaluation of six heterocyclic groups ( $R_3 = \text{IV-IX}$ ) appended to the presumed electrophilic vinyl group within several dipeptide scaffolds. The vinyl-2-pyrimidine, vinyl-2-pyridine, vinyl-2-N-methylpyridinium, vinyl-2-oxazole, and vinyl-2-thiazole substituents, unlike the vinyl-4-pyrimidine and vinyl-4-pyridine heterocycles, maintain bioisosteric similarity to the reactive acrylamides, and provided potent, time-dependent inhibitors in accord with our hypothesis. Of these PVHIs, the 2-pyridine, the charged 2-N-methylpyridine, and the vinyl-2-pyrimidine presented the most interesting heterocycles for further exploration. The inhibition of cruzain displayed by these PVHIs may be due to the reversible formation of an adduct with active-site Cys<sub>25</sub>, as is supported by the loss of time dependent inhibition when the vinyl group is saturated. Importantly, we have no evidence that such a reversible covalent bond is formed, and ongoing studies are underway to address this point.

**Selectivity of PVHIs for Cruzain over Homologous Human Cathepsins.** Cruzain has, respectively, 25%, 15%, and 23% amino acid identity with human cathepsins L, B, and S.<sup>31, 55</sup> It is preferable to proceed with cruzain inhibitors that do not readily inhibit these human lysosomal cathepsins which might engender cellular toxicity. We evaluated selected cruzain inhibitors vs. the human cysteine proteases cathepsins L, B, and S (**Table 3**). For this selectivity comparison all inhibition data were obtained at pH 5.5, for which  $K_i^*$  values were invariant for all inhibitors except compound **15** ( $K_i^* = 88 \text{ nM}$ ). The cruzain inhibitors demonstrated moderate selectivity vs. cathepsins L and S (generally, 3-fold or greater), while all of these inhibitors displayed 40-fold or higher selectivity vs. cathepsin B. The vinyl-2-pyridine inhibitors **13** and **15** are particularly selective as their  $K_i$  values are over 10-fold lower than the corresponding values with the three human cathepsins. In contrast, **K11777** showed potent inactivation at nanomolar concentrations for all three human cathepsins; this apparent lack of selectivity possibly arising from its irreversible mode of

inactivation. These results suggest that suitable selectivity for reversible cruzain inhibitors may be more easily attained than for irreversible ones.

**Table 3. Enzymatic Selectivity of Cruzain Inhibitors<sup>a</sup>**

Compound	Structure	$K_i^*$ ( $\mu$ M)			
		Cruzain	Human Cathepsin L	Human Cathepsin B	Human Cathepsin S
<b>K11777</b>	NMePip-Phe-hPhe-VSPH	IC <sub>50</sub> = 0.2 nM <sup>b</sup>	IC <sub>50</sub> = 0.2 nM <sup>b</sup>	IC <sub>50</sub> = 5.7 nM <sup>b</sup>	IC <sub>50</sub> = 0.6 nM <sup>b</sup>
<b>7</b>	Cbz-Phe-Phe-vinyl-2Pyrmd	0.28 ± 0.01	1.1 ± 0.1	32 ± 3	0.37 ± 0.02
<b>11</b>	Cbz-Phe-Phe-vinyl-2Pyr	0.28 ± 0.02	4.3 ± 0.5	28 ± 4	1.8 ± 0.3
<b>12</b>	Cbz-Phe-Phe-vinyl-2PyrNMe	0.30 ± 0.02	0.70 ± 0.04	19 ± 4	0.87 ± 0.07
<b>13</b>	Cbz-Phe-hPhe-vinyl-2Pyr	0.123 ± 0.004	1.9 ± 0.2	6.5 ± 0.9	1.41 ± 0.06
<b>15</b>	Cbz-Phe-hPhe-vinyl-2PyrNMe	0.088 ± 0.002	0.88 ± 0.06	37 ± 6	0.32 ± 0.04

<sup>a</sup>Inhibition data obtained at pH 5.5, 25°C in 10% DMSO (v/v); <sup>b</sup>Reported as apparent IC<sub>50</sub> in Ref. 32.

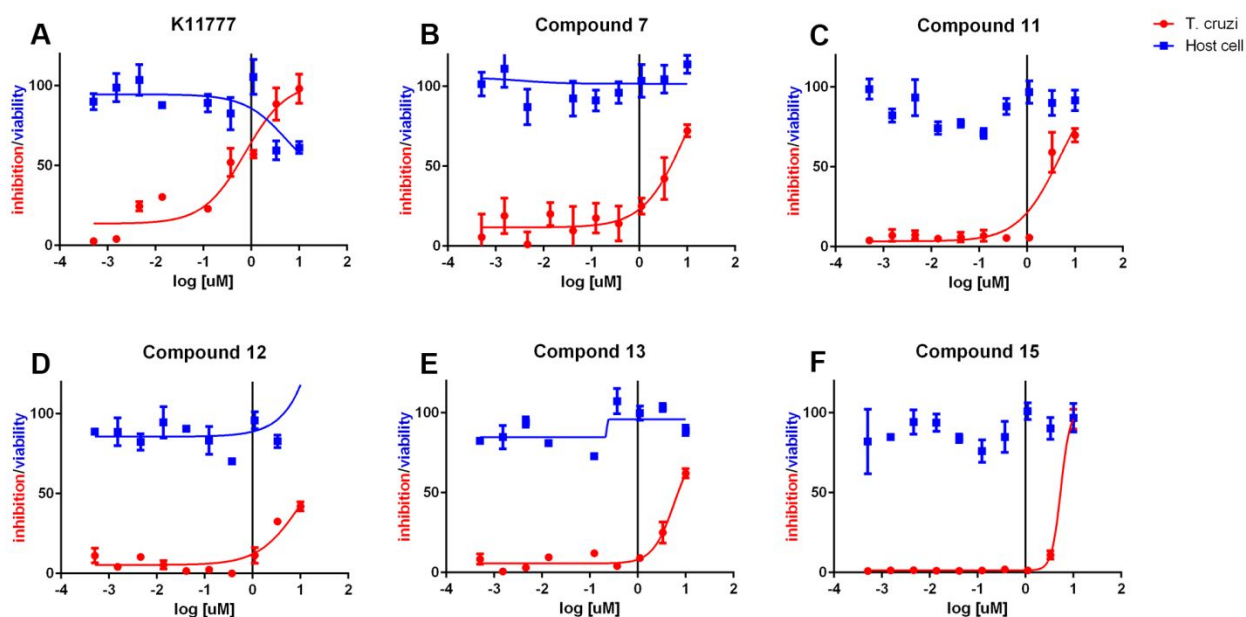
**Effects of PVHIs in axenic cultures of *T. cruzi* and in a cell model of *T. cruzi* infection.** Initially, we tested selected compounds against epimastigotes of *T. cruzi* (strain Y, ATCC 50832GFP) in axenic culture. As is observed here (**Table 4**), and has been shown previously, **K11777** weakly inhibited the growth of *T. cruzi* epimastigotes (EC<sub>50</sub> ~ 60  $\mu$ M).<sup>54</sup> PVHIs **7**, **12**, and **15** inhibited the growth of epimastigotes of *T. cruzi* (EC<sub>50</sub> = 2-20  $\mu$ M) while **11**, **13**, and **24** were poorly effective. Compounds **12** and **15** were comparably potent against cultures of *T. cruzi* (EC<sub>50</sub> = 8.6 and 2.1  $\mu$ M, respectively), and were, at a minimum, 10-fold more active than **K11777**.

**Table 4. Effects of Cruzain Inhibitors on Trypanosome and Human Cell Growth<sup>a</sup>**

Compound	Cruzain $K_i^*$ ( $\mu$ M)	<i>T. cruzi</i> axenic culture	<i>T. cruzi</i> -infected cardiomyoblasts (C2C12)	<i>T. brucei</i> <i>brucei</i> PCFs	<i>T. brucei</i> <i>brucei</i> BSFs	Human Cell Cytotoxicity	C2C12 Cytotoxicity	Selectivity Index
		EC <sub>50</sub> ( $\mu$ M)	EC <sub>50</sub> ( $\mu$ M)	EC <sub>50</sub> ( $\mu$ M)	EC <sub>50</sub> ( $\mu$ M)	CC <sub>50</sub> ( $\mu$ M)	CC <sub>50</sub> ( $\mu$ M)	CC <sub>50</sub> /EC <sub>50</sub>
<b>K11777</b>	IC <sub>50</sub> = 2 nM <sup>b</sup>	>20	0.7 ± 0.2	1.7 ± 0.5	0.09 ± 0.06	60-100	>10	140
<b>7</b>	0.364 ± 0.004	20	9.0 ± 0.5	7.1 ± 0.9	10.4 ± 0.2	>100	>10	>10
<b>9</b>	2.2 ± 0.1	20	ND	15 ± 2	>20	ND	ND	ND
<b>11</b>	0.31 ± 0.01	>20	4.9 ± 0.2	5 ± 1	5 ± 4	>100	>10	>20
<b>12</b>	0.28 ± 0.08	8.7 ± 0.1	9.9 ± 0.5	13 ± 3	6.6 ± 0.6	>100	>10	>10
<b>13</b>	0.171 ± 0.004	>20	5.9 ± 0.3	>10	4 ± 2	>100	>10	>20
<b>15</b>	0.126 ± 0.004	2.1 ± 0.1	5.4 ± 0.9	5.9 ± 0.2	2.8 ± 0.1	>100	>10	>20
<b>24</b>	0.71 ± 0.01	>20	ND	27 ± 5	ND	>100	ND	ND

<sup>a</sup>Effects of inhibitors were evaluated as EC<sub>50</sub> for axenic *T. cruzi*, PCFs/BSFs of *T. b. brucei* and *T. cruzi*-infected murine cardiomyoblasts. The Selectivity Index is the ratio of inhibitor cytotoxicity in human dermal fibroblasts (CC<sub>50</sub>)/ trypanosomacidal activity (EC<sub>50</sub>) in infected cardiomyoblasts. <sup>b</sup>Ref. 32.

Selected cruzain inhibitors were further evaluated in a more relevant cellular model of Chagas disease: *T. cruzi*-infected murine cardiomyoblasts (C2C12 cells) (Table 4, Figure 5). Inhibitors 7, 11, 12, 13, and 15 exhibited antiparasitic efficacy at values of  $EC_{50} = 5\text{--}10\text{ }\mu\text{M}$ , while displaying no cytotoxicity against the host cardiomyoblasts ( $CC_{50} > 10\text{ }\mu\text{M}$ ). These  $EC_{50}$  values demonstrated that the anti-trypanosomal activities of the reversible PVHIs are within an order of magnitude of potency of the irreversible inactivator, **K11777** ( $EC_{50} = 0.7\text{ }\mu\text{M}$ ), despite the large difference in activity vs. cruzain. Accordingly, the PVHIs, while reversible in action, and with no apparent mammalian or human cytotoxicity, are nearly as effective as the potent, irreversible inactivator **K11777**. Further, the best of the PVHIs are less than 3-fold less potent than the currently-used anti-chagasic drug benznidazole ( $LD_{50} = 1.5\text{ }\mu\text{M}$ )<sup>55</sup> suggesting that a second generation of PVHIs may provide clinical candidates.

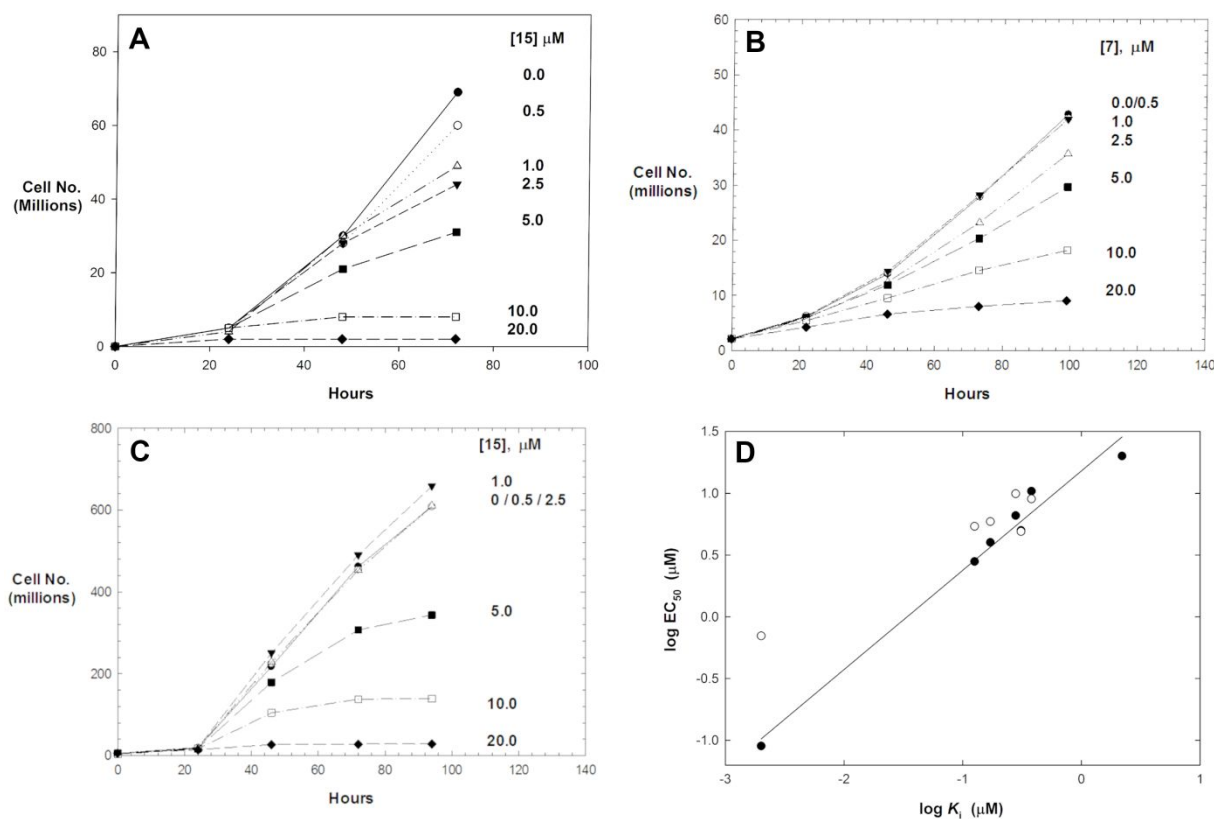


**Figure 5.** Effects of cruzain inhibitors on growth of *T. cruzi*-infected murine cardiomyoblasts in which growth of inhibition of *T. cruzi* (red) is superimposed with the viability of the cardiomyoblasts (blue).

**Effects of PVHIs in axenic cultures of *T. b. brucei*.** We additionally tested our cruzain inhibitors in axenic cultures of the related protozoan *T. b. brucei*, owing to the high structural similarity and reported essentiality of the cysteine proteases brucipain (TbCatL) in *T. b. brucei*.<sup>24, 34, 57</sup> It has been demonstrated that the cruzain inhibitor **K11777** is active in cellular cultures of both *T. b. brucei* and *T. cruzi*, supporting the notion that our PVHIs could be effective in growth inhibition of both species of parasite. For insect procyclic forms (PCFs) of *T. b. brucei* (ATCC PRA-381), compounds 7, 9, 11, 12, and 15 demonstrated growth inhibition at  $EC_{50}$  values of  $5\text{--}15\text{ }\mu\text{M}$  (Table 4, Figure 6). When compared to **K11777** ( $EC_{50} = 1.7\text{ }\mu\text{M}$ ), these PVHIs exhibited potent cell-growth inhibition. For example, compound 15 ( $EC_{50} = 5.9\text{ }\mu\text{M}$ ) was only 3-fold less

potent vs. *T. b. brucei* than **K11777**. Values of  $EC_{50}$  for these PVHIs roughly correlated with their values of  $K_i^*$ , with the exception of compound **13**.

We next evaluated these inhibitors in axenic cultures of human bloodstream forms (BSFs) of *T. b. brucei* (ATCC PRA-383). All PVHIs that were active vs. procyclic forms of *T. b. brucei* were also trypanocidal vs. the bloodstream forms, but with equal or lower  $EC_{50}$  values compared to the procyclic forms (**Table 4**). Compared to PVHIs that had similar potencies in both PCFs and BSFs, **K11777** was nearly 20-fold more potent in *T. b. brucei* BSFs than in PCFs. These results suggested that a cathepsin L-like cysteine protease in *T. b. brucei*, such as brucipain (or *TbCatL*),<sup>24</sup> is essential for growth of procyclic and bloodstream *T. b. brucei*, but perhaps an additional cysteine protease, such as *TbCatB*, is also essential in BSFs of *T. b. brucei*, as this enzyme is sensitive to **K11777** but not to the PVHIs. This is similar to the findings of Yang *et al*<sup>58</sup> who showed using an activity-based protein probe of **K11777** that *TbCatB* and brucipain (*TbCatL*) are both labeled in BSFs of *T. b. brucei* while only brucipain is labeled in PCFs. This could explain the exceptional trypanocidal activity of **K11777** in BSFs. This will be the focus of our future studies. Nonetheless, the activity of the PVHIs vs. *T. b. brucei* BSFs may hold promise for progression to their evaluation in models of African trypanosomiasis.



**Figure 6.** Cell-growth inhibition of *T. b. brucei*. (A) Inhibition of bloodstream forms by **15**; (B) Inhibition of bloodstream forms by **7**; (C) Inhibition of procyclic forms by **15**; (D) A correlation plot of values of EC<sub>50</sub> for trypanocidal activity vs. *T. b. brucei* BSFs (closed circles,  $r^2 = 0.979$ , slope = 0.80) and *T. cruzi* in murine cardiomyoblasts (open circles).

Interestingly, the values of EC<sub>50</sub> obtained for PCFs of *T. b. brucei* and amastigotes of *T. cruzi* were nearly identical for most PVHIs, despite their more modest inhibition of cruzain. Shown in **Figure 6D** is a correlation plot of log EC<sub>50</sub> for anti-trypanosomal activity for bloodstream forms of *T. b. brucei* and the amastigote forms of *T. cruzi* from the murine cardiomyoblast infection model. For the former, the correlation is excellent ( $r^2 = 0.979$ , slope = 0.80), and the activity against parasites is nearly a 1:1 correlation with logK<sub>i</sub> with these inhibitors. This result provided support that our PVHIs are targeting a cruzain-like protease in *T. b. brucei*. For *T. cruzi*, this correlation is not as strong, in part due to an absence of a sufficient range of data. We have also compared the cytotoxicity of selected inhibitors in human dermal fibroblasts vs. *T. cruzi*-infected cardiomyoblasts (Selectivity Index in **Table 4**), which demonstrates the PVHIs are more than 10-fold selective for trypanosomes vs. human cells.

## CONCLUSIONS

We have developed a novel class of reversible inhibitors for the essential cysteine protease of *Trypanosoma cruzi*, cruzain. These compounds, the peptidomimetic vinyl heterocycles, contain bioisosteric replacements for the acrylamide and vinyl sulfone warheads present in irreversible, covalent inactivators such as **K11777**. We also demonstrated that PVHIs containing vinyl-2-N-methylpyridine or vinyl-2-N-methylthiazole groups, unlike other inhibitors, readily form Michael adducts with glutathione. Our survey demonstrated that the most optimal cruzain inhibitors contained vinyl-2-pyrimidine, vinyl-2-pyridine, and vinyl-2-N-methylpyridinium groups. These PVHIs proved to be potent, time-dependent inhibitors of cruzain, albeit, fully reversible in terms of mode of action. These PVHIs are significantly active in both axenic cultures of *T. b. brucei* and in a cell infection model of *T. cruzi*, and further optimization may produce more potent anti-trypanosomal agents. Importantly, the concept of reversible covalent inactivation by vinyl-heterocycles is potentially expandable to other enzymes which contain active-site cysteines, such as EGFR, G12C K-ras and other protein kinases for which irreversible acrylamide inactivators comprise effective drugs.<sup>56</sup>

## EXPERIMENTAL SECTION

**General Synthetic Chemistry Methods and Compound Characterization.** All reagents and starting materials were obtained from commercial suppliers and used without further purification unless otherwise stated. Reactions were run under an atmosphere of nitrogen or argon and at ambient temperature unless otherwise noted. Reaction progress was monitored using thin layer chromatography and by analysis

employing an HPLC-MS (UltiMate 3000 equipped with a diode array coupled to a MSQ Plus Single Quadrupole Mass Spectrometer, ThermoFisher Scientific) using electrospray positive and negative ionization detectors. Reported liquid chromatography retention times (Rt) were established using the following conditions: column: Phenomenex Luna 5  $\mu$ m C18(2) 100 Å, 4.6 mm, 50 mm, Mobile phase A: water with 0.1% formic acid (v/v). Mobile phase B: MeCN with 0.1% formic acid (v/v). Temperature: 25 °C. Gradient: 0–100% B over 6 min, then a 2 min hold at 100% B. Flow: 1 mL/min. Detection: MS and UV at 254, 280, 214, and 350 nm.

Semi-preparative HPLC purification of compounds was performed on a Thermo Fisher Scientific UltiMate 3000 with a single wavelength detector coupled to a fraction collector. Purifications were conducted using the following conditions: column: Phenomenex Luna 5  $\mu$ m C18(2) 100 Å, 21.2 mm, 250 mm, Mobile phase A: water with 0.1% formic acid (v/v). Mobile phase B: MeCN with 0.1% formic acid (v/v). Temperature: Room temperature. Gradient: 0–100% B over 30 min, then a 5-min hold at 100% B. Flow: 20 mL/min. Detection: UV (254 nm).

$^1\text{H}/^{13}\text{C}$  NMR magnetic resonance spectra were obtained in  $\text{CDCl}_3$ ,  $\text{CD}_3\text{OD}$ , or  $\text{DMSO}-d_6$  at 400MHz/100MHz at 298 K on a Bruker Avance III NanoBay console with an Ascend magnet unless otherwise noted. The following abbreviations were utilized to describe peak patterns when appropriate: br = broad, s = singlet, d = doublet, q = quartet, t = triplet, and m = multiplet. All final compounds used for testing in assays and biological studies had purities that were determined to be >95% as evaluated by their proton NMR spectra and their HPLC/MS based on ultraviolet detection at 254 nm (see Supplementary Information). Similar RP-HPLC conditions were used for the experiments of GSH addition to vinyl-heterocycles. Masses detected were in the range 100 – 1000 Da and were detected in positive or negative mode depending on the ionization of the molecule.

General procedures (GP1-GP10 in **Scheme 1**) of synthesizing PVHIs were detailed below. Each GP described the synthesis of one representative compound. In addition, substrate synthesis and characterization were provided in Supplementary Information.

*GP1. Synthesis of Weinreb amides (a to b).* A solution of Boc-L-homophenylalanine (12.02 g, 43.03 mmol) in anh. DCM (200 mL) was cooled to 0 °C under an  $\text{N}_2$  atmosphere.  $\text{Et}_3\text{N}$  (18.1 mL, 129.09 mmol, 3 eq.) was added slowly, followed by addition of N,O-dimethylhydroxylamine hydrochloride (6.3 g, 64.5 mmol, 1.5 eq.) and dropwise addition of T3P (50% (w/v) in MeCN, 41.1 mL, 64.55 mmol, 1.5 eq.). The resulting mixture was stirred at 0°C for 30 min to 1h until TLC analysis (EtOAc/hexane=1:1, v/v) showed the disappearance of starting material. The reaction mixture was diluted with DCM and washed with  $\text{H}_2\text{O}$ . The organic layer was dried over anh.  $\text{Na}_2\text{SO}_4$  and filtered. The filtrate was concentrated *in vacuo* to afford the

crude product. Purification of the crude product by silica gel column chromatography using a gradient of 5% - 50% of EtOAc in hexane as eluent yielded the pure Weinreb amide *tert*-butyl (*S*)-(1-(methoxy(methyl)amino)-1-oxo-4-phenylbutan-2-yl)carbamate (**b**, 13.3 g, 41.31 mmol, 96% yield) as a colorless gum.

*GP2. LAH reduction of Weinreb amides (b to c).* To a solution of *tert*-butyl (*S*)-(1-(methoxy(methyl)amino)-1-oxo-4-phenylbutan-2-yl)carbamate (**b**, 6.7 g, 20.78 mmol) in anh. THF (120 mL) at -10°C under a N<sub>2</sub> atmosphere was added dropwise LAH (2.0 M in THF, 12.5 mL, 24.93 mmol, 1.2 eq.). The resulting mixture was stirred at -10°C for 30 min. Upon completion of reaction as shown by TLC analysis (EtOAc/Hexane=1:1, v/v), the reaction was quenched at the same temperature by adding dropwise 1N HCl, followed by removal of THF by rotary evaporation. Diethyl ether (500 mL) was added to the solid residue, and the solution was washed with aq. NaHCO<sub>3</sub> (1 X 50 mL) and brine (1 X 50 mL). The organic layer was dried over anh. Na<sub>2</sub>SO<sub>4</sub> and filtered. The filtrate was concentrated *in vacuo* to afford the crude product. Purification of the crude material by silica gel column chromatography using a gradient of 10% - 60% of EtOAc in hexane as eluent yielded the pure aldehyde *tert*-butyl (*S*)-(1-oxo-4-phenylbutan-2-yl)carbamate (**c**, 4.89 g, 18.57 mmol, 89% yield) as a white solid.

*GP3. Preparation of chloromethyl-heterocycles (d to e).* To a suspension of methyl pyrimidine-2-carboxylate (**d**, 1.156 g, 8.37 mmol) in anh. EtOH (20 mL) at 0°C under N<sub>2</sub> atmosphere, was added portion-wise NaBH<sub>4</sub> (0.443 g, 11.72 mmol, 5 eq.). The reaction mixture was stirred at 25°C for 2h. Upon the completion of the reaction as shown by TLC analysis (EtOAc/Hexane=1:1, v/v), the reaction solvents were removed by rotary evaporation. To the resultant colorless gummy residue was added ice cold H<sub>2</sub>O (20 mL) and extracted with DCM (5 X 50 mL). The organic layer was dried over anh. Na<sub>2</sub>SO<sub>4</sub> and filtered. The filtrate was concentrated *in vacuo* to afford the crude product pyrimidin-2-yl-methanol (0.900 g, 8.17 mmol). To this pyrimidin-2-yl-methanol in CHCl<sub>3</sub> (20 mL) at 0°C under N<sub>2</sub> atmosphere was added dropwise POCl<sub>3</sub> (1.95 mL, 3.21 g, 2.5 eq.). The reaction mixture was stirred at 25°C for 1h, followed by refluxing for an additional 3h under gentle heating until TLC analysis (EtOAc/Hexane=3:1, v/v) showed the completion of the reaction. The reaction was quenched by a careful addition of aq. NaHCO<sub>3</sub> and further addition of solid NaHCO<sub>3</sub> to afford a basic pH. The aqueous layer was extracted with CHCl<sub>3</sub> (3 X 50 mL), and the organic layer was dried over anh. Na<sub>2</sub>SO<sub>4</sub> and filtered. The filtrate was concentrated *in vacuo* to afford the pure product 2-(chloromethyl)pyrimidine (**e**, 0.948 g, 7.43 mmol, 63% yield) as a light yellow semi-solid, which was used further without any purification.

*GP4. Preparation of heterocyclic phosphonium ylides (e to f, Wittig reagents).* A mixture of 2-(chloromethyl)pyrimidine (**e**, 0.92 g, 7.22 mmol) and triphenyl phosphine (2.1 g, 7.94 mmol, 1.1 eq.) in anh. benzene (25 mL) was refluxed under N<sub>2</sub> atmosphere for 24h until TLC analysis (MeOH/DCM=1:19,

v/v) showed the completion of the reaction. The reaction mixture was concentrated via rotary evaporation, and the gummy residue was triturated with diethyl ether (3 X 10 mL). The solid obtained was purified by silica gel column chromatography using a gradient of 1% - 10% of MeOH in DCM as eluent to afford the pure product triphenyl(pyrimidin-2-ylmethyl)phosphonium chloride (**g**, 0.797 g, 2.039 mmol, 28% yield).

*GP5. Preparation of heterocyclic phosphonates (e to g, HWE reagents).* 2-(chloromethyl)pyridine hydrochloride (**e**, 16.5 g, 100.6 mmol) in DCM (100 mL) was treated with aq. NaHCO<sub>3</sub> (20 mL), and the DCM layer was dried over anh. Na<sub>2</sub>SO<sub>4</sub>. The filtrate was concentrated by rotary evaporation. The alkyl halide thus obtained along with triethyl phosphite (35 mL, 201.2 mmol, 2.0 eq.) were heated at 150°C under N<sub>2</sub> atmosphere for 5h until TLC analysis (MeOH/DCM=1:19, v/v) showed the completion of the reaction. The reaction mixture was purified by silica gel column chromatography using a gradient of 10% - 100% of EtOAc in hexane and later 1% - 10% of MeOH in DCM as eluent to yield the pure product 2-pyridyl methyl phosphonate (**g**, 18.26 g, 79.66 mmol, 79% yield).

*GP6. Wittig reaction (c + f to h).* To a suspension of the Wittig reagent triphenyl(pyrimidin-2-ylmethyl)phosphonium chloride (**f**, 0.719 g, 1.839 mmol) in anh. THF (40 mL) at -70°C under an N<sub>2</sub> atmosphere, was added dropwise LHMDS (1.0 M in THF, 2.03 mL, 2.024 mmol, 1.1 eq.), which was stirred at the same temperature for 15 min. To this mixture a solution of Boc-Phe-H (**c**, 0.321 g, 1.287 mmol, 0.7 eq.) in THF (10 mL) was added, and stirred over 2 h until the temperature reached -40°C. Upon completion of reaction as revealed by TLC analysis (EtOAc/hexane=1:1, v/v), the reaction was quenched by addition of 0.1 mL of glacial acetic acid, followed by aq. NaHCO<sub>3</sub>. Most of the THF was removed carefully using a rotary evaporator, and the residue was extracted with EtOAc (2X). The organic layer was dried over anh. Na<sub>2</sub>SO<sub>4</sub> and filtered. The filtrate was concentrated *in vacuo* to afford the crude material, which was purified by silica gel column chromatography using a gradient of 5% - 30% of EtOAc in hexane as eluent, yielding the pure olefin *tert*-butyl (*S,E*)-(1-phenyl-4-(pyrimidin-2-yl)but-3-en-2-yl)carbamate (**h**, *E*-isomer, 0.060 g, 14%). The other *Z*-isomer (0.014 g) was isolated as a side product and the ratio of *E* to *Z* isomers was typically 4:1.

*GP7. Horner–Wadsworth–Emmons reaction (c + g to h).* To a solution of the 2-pyridyl methyl phosphonate ester (**g**, 1.30 g, 5.65 mmol) in anh. THF (25 mL) at -70°C under N<sub>2</sub> atmosphere, was added dropwise LHMDS (1.0 M in THF, 6.22 mL, 6.22 mmol, 1.1 eq.). The reaction was stirred at the same temperature for 15 min, followed by dropwise addition of a solution of Boc-hPhe-H (**c**, 1.34 g in 10 mL THF, 5.09 mmol, 0.9 eq.). The reaction was stirred until it reached the temperature -20°C over 2 h. Upon completion of reaction revealed by TLC analysis (EtOAc/hexane=1:1, v/v), to the reaction mixture at 0°C was added glacial acetic acid (0.5 mL), followed by addition of 20 mL of saturated NaHCO<sub>3</sub>. The aqueous layer was extracted with EtOAc (3 X 100 mL). Extracts were washed with brine (1 X 50 mL), and the organic layer

was dried over anh.  $\text{Na}_2\text{SO}_4$  and filtered. The filtrate was concentrated *in vacuo* to afford the crude product, which was purified by silica gel column chromatography using a gradient of 10% - 50% of EtOAc in hexane as eluent to yield the pure product *tert*-butyl (*S,E*)-(5-phenyl-1-(pyridin-2-yl)pent-1-en-3-yl)carbamate (**h**, 0.344 g, 1.016 mmol, 20% yield).

**GP8. Amide coupling with  $P_3$ - $P_2$  fragment (**h** to **j**).** To a solution of *tert*-butyl (*S,E*)-(5-phenyl-1-(pyridin-2-yl)pent-1-en-3-yl)carbamate (**h**, 0.143 g, 0.423 mmol) in anh DCM (5 mL) at 0°C, was added dropwise TFA (1.5 mL in 1 mL DCM) with stirring at the same temperature for 1h. Upon the completion of the reaction as revealed by TLC analysis (EtOAc/hexane=1:1, v/v), the reaction solvent was removed by a rotary evaporator. The resulting oil was co-evaporated on a rotary evaporator with  $\text{CHCl}_3$  (3X) and ether (3X). The solid product was dried on high vacuum to yield the TFA salt (*S,E*)-5-phenyl-1-(pyridin-2-yl)pent-1-en-3-aminium trifluoroacetate (0.149 g, 0.423 mmol), which was used in subsequent synthetic steps without further purification. To a solution of above TFA salt in anh. DCM (5 mL) at -10°C under  $\text{N}_2$  atmosphere, was added dropwise DIPEA (0.6 mL, 0.3.44 mmol, 8 eq.), followed by addition of Cbz-Phe-OH (0.13 g, 0.43 mmol, 1 eq.) and T3P (50% in EtOAc, 0.41 mL, 1.5 eq.). The reaction was stirred at 0°C for an additional 1 h. Upon the completion of reaction revealed by TLC analysis (EtOAc/hexane=1:1, v/v), the reaction mixture was diluted with DCM (50 mL), and then washed with  $\text{H}_2\text{O}$  (3X) and brine (3X). The organic layer was dried over anh.  $\text{Na}_2\text{SO}_4$  and filtered. The filtrate was concentrated *in vacuo* to afford the crude product, which was purified by silica gel column chromatography using a gradient of 10% - 50% of EtOAc in hexane as eluent to yield the pure product benzyl ((*S*)-1-oxo-3-phenyl-1-(((*S,E*)-5-phenyl-1-(pyridin-2-yl)pent-1-en-3-yl)amino)propan-2-yl)carbamate (**i**, 0.113 g, 0.217 mmol, 51%).

**GP9. *N*-methylation using methyl iodide (**i** to **j**).** To a suspension of benzyl ((*S*)-1-oxo-3-phenyl-1-(((*S,E*)-1-phenyl-4-(pyridin-2-yl)but-3-en-2-yl)amino)propan-2-yl)carbamate (**i**, 0.049 g, 0.098 mmol) in anh. MeCN (5 mL) and under  $\text{N}_2$  atmosphere, was added MeI (0.03 mL, 0.490 mmol, 5 eq.), and the reaction mixture was heated under reflux for 9h. Upon the completion of the reaction revealed by TLC analysis (EtOAc/hexane=1:1, v/v), the solvents were removed by rotary evaporation. The resulting gummy residue was dissolved in  $\text{CHCl}_3$  (1 mL), and precipitated with ether (5 mL). The solvents were decanted, and this procedure was repeated twice. The solid obtained was dried under high vacuum to give pure product 2-(((*S,E*)-3-((*S*)-2-(((benzyloxy)carbonyl)amino)-3-phenylpropanamido)-4-phenylbut-1-en-1-yl)-1-methylpyridin-1-ium iodide as a yellow solid (**j**, 0.039 g, 61%).

**GP10. Preparation of peptide acrylamide (**i** to **k**).** A solution of ethyl (*S,E*)-4-(((*S*)-2-(((benzyloxy)carbonyl)amino)-3-phenylpropanamido)-5-phenylpent-2-enoate (**i**, 0.346 g, 0.69 mmol) in THF (6mL) at 0°C was treated with LiOH (1N in  $\text{H}_2\text{O}$ , 0.83mL, 0.83mmol, 1.2 eq) and stirred overnight. The reaction was concentrated by rotary evaporation and the aqueous layer was added water and acidified

to pH 1-2, and extracted with EtOAc (3X). The combined organic layers were dried and concentrated to yield the crude acrylic acid. To a solution of this acrylic acid (0.124 g, 0.262 mmol) in THF (6mL) at -15°C was added Et<sub>3</sub>N (0.11 mL, 0.787 mmol, 3 eq.) and dropwise addition of ClCO<sub>2</sub>Et (0.035 mL, 0.367 mmol), which resulted in a white precipitate. The reaction mixture was stirred at the same temperature for an additional 30 min, then aq. 1M NH<sub>4</sub>Cl (0.4 mL) was added dropwise with continuous stirring over 3h until a temperature of 25°C was attained. Upon completion of reaction as revealed by TLC analysis (EtOAc/hexane=1:1, v/v), most of the reaction solvent was removed using a rotary evaporator, and the solid residue was extracted with EtOAc. The organic layer was washed with aq. NaHCO<sub>3</sub> (2X), H<sub>2</sub>O (1X) and brine (1X), and was dried over anh. Na<sub>2</sub>SO<sub>4</sub> and filtered. The filtrate was concentrated *in vacuo* to afford the crude product. Purification of the crude product by silica gel chromatography using a gradient of 20% - 100% of EtOAc in hexane as eluent, yielded the pure product benzyl ((S)-1-(((S,E)-5-amino-5-oxo-1-phenylpent-3-en-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (0.027 g, 0.057 mmol, 22% yield).

*4-Methyl-N-((S)-1-oxo-3-phenyl-1-(((S,E)-1-phenyl-4-(phenylsulfonyl)but-3-en-2-yl)amino)propan-2-yl)piperazine-1-carboxamide (1, Cbz-Phe-Phe-VSPh)*. White solid, 0.115 g, 0.202 mmol, 56% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 2.79 (d, 2H, *J* = 6.8 Hz), 2.85 - 3.12 (m, 2H), 4.26 (q, 1H, *J* = 7.3 Hz), 4.79 - 4.95 (m, 1H), 5.04 (s, 2H), 5.13 (s, 1H), 5.75 (s, 1H), 5.96 (dd, 1H, *J*<sub>1</sub> = 1.8 Hz, *J*<sub>2</sub> = 15.1 Hz), 6.78 (dd, 1H, *J*<sub>1</sub> = 4.8 Hz, *J*<sub>2</sub> = 15.1 Hz), 6.95 - 7.03 (m, 2H), 7.05 - 7.11 (m, 2H), 7.12 - 7.23 (m, 6H), 7.27 - 7.39 (m, 5H), 7.47 - 7.56 (m, 2H), 7.57 - 7.67 (m, 1H), 7.72 - 7.84 (m, 2H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 38.4, 40.3, 50.4, 56.7, 67.4, 127.3, 127.4, 127.8, 128.2, 128.5, 128.7 (2C), 128.8, 129.0, 129.3, 129.4, 131.1, 133.6, 135.5, 136.1, 136.2, 140.2, 144.6, 156.0, 170.5; LC-MS *t*<sub>R</sub> 7.27 min, *m/z* 569.31 [M+H]<sup>+</sup>, (C<sub>33</sub>H<sub>32</sub>N<sub>2</sub>O<sub>5</sub>S<sup>+</sup> Calcd 569.21).

*Benzyl ((S)-1-(((S,E)-5-amino-5-oxo-1-phenylpent-3-en-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (2, Cbz-Phe-Phe-vinyl-CONH<sub>2</sub>)*. White solid, 0.027 g, 0.057 mmol, 22% yield. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 2.64 - 2.77 (m, 1H), 2.84 (d, 2H, *J* = 7.2 Hz), 2.96 (dd, 1H, *J*<sub>1</sub> = 3.9 Hz, *J*<sub>2</sub> = 13.7 Hz), 4.19 - 4.31 (m, 1H), 4.62 (pentet, 1H, *J* = 6.8 Hz), 4.95 (s, 1H), 5.85 (d, 1H, *J* = 15.5 Hz), 6.56 (dd, 1H, *J*<sub>1</sub> = 5.9 Hz, *J*<sub>2</sub> = 15.5 Hz), 6.93 (s, 1H), 7.12 - 7.46 (m, 17H), 8.25 (d, 1H, *J* = 8.2 Hz); <sup>13</sup>C NMR (100 MHz, DMSO-*d*<sub>6</sub>) δ 28.7, 37.6, 50.9, 56.07, 65.1, 124.1, 126.2, 126.4, 127.4, 127.6, 127.9, 128.1, 128.2, 129.2, 129.5, 137.0, 137.8, 138.0, 142.0, 155.6, 166.2, 170.7; LC-MS *t*<sub>R</sub> 4.71 min, *m/z* 472.46 [M+H]<sup>+</sup>, (C<sub>28</sub>H<sub>29</sub>N<sub>3</sub>O<sub>4</sub><sup>+</sup> Calcd 472.22).

*Benzyl ((S)-1-(((S,E)-6-amino-6-oxo-1-phenylhex-4-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (3, Cbz-Phe-hPhe-vinyl-CONH<sub>2</sub>)*. White solid, 0.013 g, 0.027 mmol, 15% yield. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 1.64 - 1.92 (m, 2H), 2.54 - 2.72 (m, 2H), 2.81 (dd, 1H, *J*<sub>1</sub> = 10.6 Hz, *J*<sub>2</sub> = 13.6 Hz), 3.03 (dd, 1H, *J*<sub>1</sub> = 4.0 Hz, *J*<sub>2</sub> = 13.6 Hz), 4.25 - 4.33 (m, 1H), 4.35 - 4.43 (m, 1H), 4.85 - 5.04 (m,

2H), 5.90 (d, 1H,  $J = 15.5$  Hz), 6.54 (dd, 1H,  $J_1 = 5.7$  Hz,  $J_2 = 15.5$  Hz), 6.94 (s, 1H), 7.15 – 7.35 (m, 15H), 7.41 (s, 1H), 7.49 (d, 1H,  $J = 8.5$  Hz), 8.22 (d, 1H,  $J = 8.2$  Hz); LC-MS  $t_R$  4.89 min,  $m/z$  486.24  $[M+H]^+$ , ( $C_{29}H_{31}N_3O_4^+$  Calcd 486.24).

*N-((S)-1-(((S,E)-6-amino-6-oxo-1-phenylhex-4-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)-4-methylpiperazine-1-carboxamide (4, NMePip-Phe-hPhe-vinyl-CONH<sub>2</sub>)*. White solid, 0.022 g, 0.048 mmol, 18% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  1.66 – 1.91 (m, 2H), 2.23 (s, 3H), 2.25 – 2.34 (m, 4H), 2.50 – 2.64 (m, 2H), 2.99 – 3.15 (m, 2H), 3.25 – 3.41 (m, 4H), 4.44 – 4.57 (m, 1H), 4.62 (q, 1H,  $J = 7.4$  Hz), 5.37 (d, 1H,  $J = 7.6$  Hz), 5.60 (dd, 1H,  $J_1 = 1.3$  Hz,  $J_2 = 15.3$  Hz), 5.75 (s, 1H), 6.01 (s, 1H), 6.62 (dd, 1H,  $J_1 = 5.5$  Hz,  $J_2 = 15.3$  Hz), 6.93 (d, 1H,  $J = 8.2$  Hz), 7.04 – 7.31 (m, 11H); <sup>13</sup>C NMR (100 MHz, DMSO- $d_6$ )  $\delta$  32.0, 36.1, 38.5, 43.9, 46.1, 49.9, 54.6, 56.2, 122.9, 126.2, 126.9, 128.5, 128.6, 128.8, 129.7, 137.3, 141.1, 144.4, 157.3, 167.5, 172.0; LC-MS  $t_R$  2.54 min,  $m/z$  478.36  $[M+H]^+$ , ( $C_{27}H_{35}N_5O_3^+$  Calcd 478.28).

*Benzyl ((S)-1-(((S,E)-1,4-diphenylbut-3-en-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (5, Cbz-Phe-Phe-vinyl-Ph)*. Off-white solid, 0.054 g, 0.107 mmol, 35% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  2.87 (dt,  $J = 2.8, 6.3$  Hz, 2H), 2.96 – 3.16 (m, 2H), 4.35 (q,  $J = 7.6$  Hz, 1H), 4.80 – 4.96 (m, 1H), 5.29 (d,  $J = 9.6$  Hz, 1H), 5.69 (d,  $J = 9.9$  Hz, 1H), 5.94 (ddt,  $J = 3.0, 6.3, 15.9$  Hz, 1H), 6.27 (d,  $J = 15.9$  Hz, 1H), 7.03 – 7.14 (m, 2H), 7.14 – 7.28 (m, 10H), 7.28 – 7.40 (m, 9H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  29.7, 38.6, 41.3, 51.8, 67.1, 126.4, 126.6, 127.1, 127.6, 128.0, 128.2, 128.3, 128.4, 128.5, 128.7, 129.3, 129.4, 130.9, 136.5, 136.8, 169.8; LC-MS  $t_R$  6.20 min,  $m/z$  505.31  $[M+H]^+$ , ( $C_{33}H_{32}N_2O_3^+$  Calcd 505.25).

*Benzyl ((S)-1-(((S,E)-4-(4-nitrophenyl)-1-phenylbut-3-en-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (6, Cbz-Phe-Phe-vinyl-(4-NO<sub>2</sub>)Ph)*. White fluffy solid, 0.530 g, 0.964 mmol, 55% yield. <sup>1</sup>H NMR (400 MHz, DMSO- $d_6$ )  $\delta$  2.76 (dd, 1H,  $J_1 = 9.7$  Hz,  $J_2 = 13.4$  Hz), 2.85 – 2.97 (m, 3H), 4.18 – 4.35 (m, 1H), 4.68 (pentet, 1H,  $J = 6.7$  Hz), 4.99 (s, 2H), 6.41 (d, 1H,  $J = 16.1$  Hz), 6.50 (dd, 1H,  $J_1 = 5.4$  Hz,  $J_2 = 16.1$  Hz), 7.13 – 7.37 (m, 15H), 7.42 (d, 1H,  $J = 8.5$  Hz), 7.58 (d, 2H,  $J = 8.5$  Hz), 8.14 – 8.25 (m, 3H); <sup>13</sup>C NMR (100 MHz, DMSO- $d_6$ )  $\delta$  37.7, 39.0, 51.9, 56.2, 65.2, 123.9, 126.1, 126.2, 127.0, 127.4, 127.6, 127.9, 128.0, 128.1, 128.2, 129.2, 129.3, 135.6, 137.0, 137.8, 138.0, 143.3, 146.2, 155.6, 170.6; LC-MS  $t_R$  6.12 min,  $m/z$  550.28  $[M+H]^+$ , ( $C_{33}H_{31}N_3O_5^+$  Calcd 550.23).

*Benzyl ((S)-1-oxo-3-phenyl-1-(((S,E)-1-phenyl-4-(pyrimidin-2-yl)but-3-en-2-yl)amino)propan-2-yl)carbamate (7, Cbz-Phe-Phe-vinyl-2Pyrm)*. White solid, 0.026 g, 0.0513 mmol, 33% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  2.73 – 2.96 (m, 2H), 3.01 (d, 2H,  $J = 4.8$  Hz), 4.23 – 4.50 (m, 1H), 4.91 – 5.02 (m, 1H), 5.05 (s, 2H), 5.30 (s, 1H), 6.05 (s, 1H), 6.48 (d, 1H,  $J = 15.7$  Hz), 7.00 (dd, 1H,  $J_1 = 5.6$  Hz,  $J_2 = 15.7$  Hz), 7.05 – 7.37 (m, 16H), 8.63 (d, 2H,  $J = 4.0$  Hz); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  38.6, 41.0, 51.6,

56.5, 67.3, 119.0, 126.9, 127.2, 128.2, 128.3, 128.6, 128.7, 128.9, 129.4, 129.5, 130.3, 136.3, 136.5, 136.7, 139.2, 156.1, 157.1, 164.2, 170.3; LC-MS  $t_R$  5.14 min,  $m/z$  507.26  $[M+H]^+$ , ( $C_{31}H_{30}N_4O_3^+$  Calcd 507.24).

*Benzyl ((S)-1-oxo-3-phenyl-1-(((S,E)-5-phenyl-1-(pyrimidin-2-yl)pent-1-en-3-yl)amino)propan-2-yl)carbamate (8, Cbz-Phe-hPhe-vinyl-2PyrmD)*. Off-white solid, 0.280 g, 0.538 mmol, 29% yield.  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  1.76 – 2.03 (m, 2H), 2.63 (t,  $J$  = 7.9 Hz, 2H), 3.10 (t,  $J$  = 8.4 Hz, 2H), 4.47 (s, 1H), 4.74 (h,  $J$  = 7.3 Hz, 1H), 5.10 (d,  $J$  = 7.9 Hz, 2H), 5.53 (s, 1H), 6.27 (s, 1H), 6.56 – 6.72 (m, 1H), 6.96 – 7.05 (m, 1H), 7.06 – 7.14 (m, 3H), 7.19 (d,  $J$  = 7.3 Hz, 3H), 7.21 – 7.34 (m, 10H), 8.66 (dd,  $J$  = 4.9, 15.8 Hz, 2H);  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$  32.0, 36.2, 38.6, 50.5, 56.6, 67.1, 118.9, 126.0, 127.0, 128.0, 128.1, 128.3, 128.4, 128.5, 128.7, 128.8, 129.3, 129.4, 130.1, 136.4, 139.6, 141.2, 156.9, 164.1, 170.3; LC-MS  $t_R$  = 5.56 min,  $m/z$  521.24  $[M+H]^+$ , ( $C_{32}H_{32}N_4O_3^+$  Calcd 521.26).

*4-Methyl-N-((S)-1-oxo-3-phenyl-1-(((S,E)-5-phenyl-1-(pyrimidin-2-yl)pent-1-en-3-yl)amino)propan-2-yl)piperazine-1-carboxamide (9, NMePip-Phe-hPhe-vinyl-2PyrmD)*. Off-white gum, 0.054 g, 0.105 mmol, 54% yield.  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  1.82 – 1.99 (m, 2H), 2.29 (s, 3H), 2.34 – 2.46 (m, 4H), 2.61 (t, 1H,  $J$  = 7.5 Hz), 3.09 (d, 2H,  $J$  = 7.5 Hz), 3.39 (s, 4H), 4.58 – 4.74 (m, 2H), 5.47 (d, 1H,  $J$  = 6.3 Hz), 6.61 (d, 1H,  $J$  = 15.7 Hz), 6.87 (d, 1H,  $J$  = 8.2 Hz), 6.98 (dd, 1H,  $J_1$  = 6.3 Hz,  $J_2$  = 15.7 Hz), 7.06 – 7.31 (m, 11H), 8.66 (d, 2H,  $J$  = 4.8 Hz);  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$  32.1, 36.5, 38.8, 43.5, 45.6, 50.6, 54.3, 56.1, 118.9, 126.0, 126.9, 128.4, 128.5 (2C), 128.6, 129.6, 130.1, 137.2, 140.1, 141.3, 157.0, 164.3, 171.8; LC-MS  $t_R$  2.96 min,  $m/z$  513.17  $[M+H]^+$ , ( $C_{30}H_{36}N_6O_2^+$  Calcd 513.30).

*Benzyl ((S)-1-oxo-3-phenyl-1-(((S,E)-4-(pyrimidin-2-yl)but-3-en-2-yl)amino)propan-2-yl)carbamate (10, Cbz-Phe-Ala-vinyl-2PyrmD)*. White fluffy solid, 0.017 g, 0.039 mmol, 14% yield.  $^1H$  NMR (400 MHz,  $CDCl_3$  + MeOD)  $\delta$  1.32 (d, 3H,  $J$  = 6.6 Hz), 2.96 (dd, 1H,  $J_1$  = 8.0 Hz,  $J_2$  = 13.6 Hz), 3.10 (dd, 1H,  $J_1$  = 6.4 Hz,  $J_2$  = 13.6 Hz), 4.36 – 4.44 (m, 4H), 4.71 (pentet, 1H,  $J$  = 6.4 Hz), 4.97 – 5.12 (m, 2H), 6.52 (d, 1H,  $J$  = 15.7 Hz), 6.98 (dd, 1H,  $J_1$  = 5.8 Hz,  $J_2$  = 15.7 Hz), 7.13 – 7.36 (m, 11H), 8.70 (d, 2H,  $J$  = 4.9 Hz);  $^{13}C$  NMR (100 MHz,  $CDCl_3$  + MeOD)  $\delta$  19.6, 38.7, 46.1, 56.1, 66.7, 119.0, 126.7, 127.6, 127.9, 128.1, 128.3 (2C), 129.2, 136.3, 141.6, 156.9 (2C), 157.0, 163.9, 171.0; LC-MS  $t_R$  4.68 min,  $m/z$  430.91  $[M+H]^+$ , ( $C_{25}H_{26}N_4O_3^+$  Calcd 431.21).

*Benzyl ((S)-1-oxo-3-phenyl-1-(((S,E)-1-phenyl-4-(pyridin-2-yl)but-3-en-2-yl)amino)propan-2-yl)carbamate (11, Cbz-Phe-Phe-vinyl-2Pyr)*. White solid, 0.554 g, 1.096 mmol, 81% yield.  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  2.87 (dq, 1H,  $J_1$  = 6.8 Hz,  $J_2$  = 13.7 Hz), 3.01 (d, 2H,  $J$  = 7.1 Hz), 4.35 (q, 1H,  $J$  = 7.0 Hz), 4.91 (pentet, 1H,  $J$  = 6.8 Hz), 5.06 (s, 2H), 5.24 (s, 1H), 5.88 (d, 1H,  $J$  = 8.4 Hz), 6.32 (d, 1H,  $J$  = 15.7 Hz), 6.58 (dd, 1H,  $J_1$  = 6.1 Hz,  $J_2$  = 15.7 Hz), 7.08 – 7.36 (m, 17H), 7.59 (dt, 1H,  $J_1$  = 1.6 Hz,  $J_2$  =

7.7 Hz), 8.53 (d, 1H,  $J = 4.3$  Hz);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  38.4, 41.1, 51.7, 56.4, 67.1, 122.1, 122.2, 126.7, 127.0, 128.0, 128.2, 128.4, 128.5 (2C), 128.8, 129.4 (2C), 130.5, 133.0, 136.1, 136.4, 136.8, 149.5, 154.8, 155.9, 170.0; LC-MS  $t_R$  4.39 min,  $m/z$  506.24  $[\text{M}+\text{H}]^+$ , ( $\text{C}_{32}\text{H}_{31}\text{N}_3\text{O}_3^+$  Calcd 506.24).

2-((*S,E*)-3-((*S*)-2-(((benzyloxy)carbonyl)amino)-3-phenylpropanamido)-4-phenylbut-1-en-1-yl)-1-methylpyridin-1-ium iodide (**12**, *Cbz-Phe-Phe-vinyl-2PyrNMe*). Yellow solid, 0.039 g, 0.060 mmol, 61% yield.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  3.01 – 3.15 (m, 3H), 3.19 (dd, 1H,  $J_1 = 7.7$  Hz,  $J_2 = 13.6$  Hz), 4.19 (s, 3H), 4.51 (q, 1H,  $J = 7.0$  Hz), 4.90 – 5.02 (m, 2H), 5.06 (s, 1H), 5.81 (s, 1H), 6.69 (d, 1H,  $J = 15.7$  Hz), 6.86 (dd, 1H,  $J_1 = 4.6$  Hz,  $J_2 = 15.7$  Hz), 7.08 – 7.28 (m, 14H), 7.65 – 7.79 (m, 2H), 7.89 – 8.01 (m, 1H), 8.21 – 8.31 (m, 1H), 8.97 – 9.09 (m, 1H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  38.1, 39.8, 47.5, 52.6, 57.1, 66.8, 119.9, 126.1, 126.3, 126.9, 127.1, 127.5, 127.6, 127.9, 128.5, 128.6, 128.8, 129.6, 129.7, 136.7, 136.9, 144.9, 146.0, 147.6, 152.8, 156.2, 171.5; LC-MS  $t_R$  3.25 min,  $m/z$  520.32  $[\text{M}+\text{H}]^+$ , ( $\text{C}_{33}\text{H}_{34}\text{N}_3\text{O}_3^+$  Calcd 520.26).

Benzyl ((*S*)-1-oxo-3-phenyl-1-(((*S,E*)-5-phenyl-1-(pyridin-2-yl)pent-1-en-3-yl)amino)propan-2-yl)carbamate (**13**, *Cbz-Phe-hPhe-vinyl-2Pyr*). White solid, 0.113 g, 0.217 mmol, 51% yield.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  1.73 – 1.95 (m, 2H), 2.59 (t, 2H,  $J = 7.9$  Hz), 3.04 (d, 2H,  $J = 7.0$  Hz), 4.38 – 4.52 (m, 1H), 4.65 (pentet, 1H,  $J = 7.1$  Hz), 5.03 (s, 2H), 5.59 (d, 1H,  $J = 6.4$  Hz), 6.34 (d, 1H,  $J = 5.1$  Hz), 6.43 (d, 1H,  $J = 15.7$  Hz), 6.53 (dd, 1H,  $J_1 = 6.2$  Hz,  $J_2 = 15.7$  Hz), 7.02 – 7.32 (m, 17H), 7.57 (dt, 1H,  $J_1 = 1.4$  Hz,  $J_2 = 7.7$  Hz), 8.51 (d, 1H,  $J = 4.4$  Hz);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  32.1, 36.4, 38.7, 50.8, 56.6, 67.1, 122.1, 122.3, 126.0, 127.0, 128.0, 128.2, 128.4, 128.5 (2C), 128.7, 129.5, 130.5, 133.8, 136.2, 136.5 (2C), 141.4, 149.5, 155.0, 156.1, 170.4; LC-MS  $t_R$  4.68 min,  $m/z$  518.74, 520.41  $[\text{M}+\text{H}]^+$ , ( $\text{C}_{33}\text{H}_{33}\text{N}_3\text{O}_3^+$  Calcd 520.26).

4-Methyl-N-((*S*)-1-oxo-3-phenyl-1-(((*S,E*)-5-phenyl-1-(pyridin-2-yl)pent-1-en-3-yl)amino)propan-2-yl)piperazine-1-carboxamide (**14**, *NMePip-Phe-hPhe-vinyl-2Pyr*). Off-white solid, 0.090 g, 0.176 mmol, 44% yield.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  1.77 – 1.94 (m, 2H), 2.21 (s, 3H), 2.22 – 2.27 (m, 4H), 2.59 (t, 2H,  $J = 8.0$  Hz), 3.08 (d, 2H,  $J = 7.0$  Hz), 3.28 (m, 4H), 4.61 (pentet, 1H,  $J = 7.3$  Hz), 4.7 (q, 1H,  $J = 7.3$  Hz), 5.39 (d, 1H,  $J = 7.7$  Hz), 6.46 (d, 1H,  $J = 15.8$  Hz), 6.56 (dd, 1H,  $J_1 = 6.2$  Hz,  $J_2 = 15.8$  Hz), 6.98 (d, 1H,  $J = 8.4$  Hz), 7.06 – 7.24 (m, 12H), 7.59 (dt, 1H,  $J_1 = 1.8$  Hz,  $J_2 = 7.7$  Hz), 8.53 (d, 1H,  $J = 4.3$  Hz);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  32.1, 36.6, 39.0, 43.8, 46.1, 50.7, 54.6, 55.9, 122.0, 122.2, 125.9, 126.8, 128.4 (2C), 128.5, 129.6, 130.3, 134.2, 136.4, 137.2, 141.4, 149.5, 155.1, 157.0, 171.6; LC-MS  $t_R$  2.60 min,  $m/z$  512.28  $[\text{M}+\text{H}]^+$ , ( $\text{C}_{31}\text{H}_{37}\text{N}_5\text{O}_2^+$  Calcd 512.30).

2-((*S,E*)-3-((*S*)-2-(((benzyloxy)carbonyl)amino)-3-phenylpropanamido)-5-phenylpent-1-en-1-yl)-1-methylpyridin-1-ium (**15**, *Cbz-Phe-hPhe-vinyl-2PyrNMe*). Yellow solid, 0.029 g, 0.044 mmol, 89% yield.

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 1.98 – 2.10 (m, 1H), 2.11 – 2.25 (m, 1H), 2.56 – 2.85 (m, 2H), 3.04 – 3.35 (m, 2H), 4.22 (s, 3H), 4.62 (d, 1H, *J* = 5.5 Hz), 4.79 (s, 1H), 4.92 – 5.09 (m, 2H), 5.88 (s, 1H), 6.70 (d, 1H, *J* = 15.8 Hz), 6.77 (dd, 1H, *J*<sub>1</sub> = 3.6 Hz, *J*<sub>2</sub> = 15.8 Hz), 7.09 – 7.34 (m, 14H), 7.66 – 7.79 (m, 2H), 7.90 (d, 1H, *J* = 6.0 Hz), 8.17 – 8.30 (m, 1H), 8.98 (d, 1H, *J* = 4.5 Hz); LC-MS *t*<sub>R</sub> 3.48 min, *m/z* 534.25 [M+H]<sup>+</sup>, (C<sub>34</sub>H<sub>36</sub>N<sub>3</sub>O<sub>3</sub><sup>+</sup> Calcd 534.28).

*Benzyl ((S)-1-(((S,E)-1-(4-methoxypyridin-2-yl)-5-phenylpent-1-en-3-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (16, Cbz-Phe-hPhe-vinyl-2-(4-OMe)-Pyr)*. Pale-yellow solid, 0.072 g, 0.131 mmol, 33% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 1.73 – 2.01 (m, 2H), 2.59 (dt, *J* = 7.9, 37.9 Hz, 2H), 3.11 (dd, *J* = 7.0, 11.7 Hz, 2H), 3.87 (d, *J* = 9.2 Hz, 3H), 4.50 (dd, *J* = 7.6, 23.1 Hz, 1H), 4.68 (t, *J* = 7.1 Hz, 1H), 5.08 (d, *J* = 4.6 Hz, 2H), 5.74 (dd, *J* = 8.1, 65.8 Hz, 1H), 6.39 – 6.58 (m, 1H), 6.67 – 6.81 (m, 2H), 7.11 – 7.15 (m, 2H), 7.16 – 7.26 (m, 8H), 7.27 – 7.34 (m, 7H), 8.37 (dd, *J* = 5.8, 23.3 Hz, 1H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 32.0, 36.3, 38.7, 50.6, 55.2, 56.5, 67.1, 108.3, 108.5, 126.0, 126.6, 127.0, 127.9, 128.0, 128.1, 128.1, 128.3, 128.4, 128.4, 128.4, 128.5, 128.5, 128.7, 129.4, 129.4, 134.7, 135.1, 136.5, 141.3, 150.0, 156.1, 166.6, 170.3; LC-MS *t*<sub>R</sub> = 3.57 min, *m/z* 550.16 [M+H]<sup>+</sup>, (C<sub>34</sub>H<sub>35</sub>N<sub>3</sub>O<sub>4</sub><sup>+</sup> Calcd 549.26).

*Benzyl ((S)-1-oxo-3-phenyl-1-(((S,E)-5-phenyl-1-(4-(trifluoromethyl)pyridin-2-yl)pent-1-en-3-yl)amino)propan-2-yl)carbamate (17, Cbz-Phe-hPhe-vinyl-2-(4-CF<sub>3</sub>)-Pyr)*. White solid, 0.260 g, 0.442 mmol, 79% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 1.76 – 1.95 (m, 2H), 2.59 (t, 2H, *J* = 7.8 Hz), 3.06 (d, 2H, *J* = 7.1 Hz), 4.36 – 4.51 (m, 1H), 4.65 (pentet, 1H, *J* = 7.1 Hz), 5.05 (s, 2H), 5.49 (s, 1H), 6.13 (s, 1H), 6.35 (d, 1H, *J* = 15.7 Hz), 6.59 (dd, 1H, *J*<sub>1</sub> = 6.0 Hz, *J*<sub>2</sub> = 15.7 Hz), 7.06 – 7.11 (m, 2H), 7.14 – 7.33 (m, 15H), 8.67 (d, 1H, *J* = 4.9 Hz); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 32.1, 36.4, 38.7, 50.7, 56.8, 67.3, 117.5, 121.6, 124.3, 126.2, 127.0 (C-F), 127.2, 128.1, 128.3, 128.5, 128.6 (3C), 128.8, 129.2, 129.5, 136.2, 136.6, 138.7, 139.1, 141.2, 150.5, 156.2 (C-F), 156.4, 170.4; LC-MS *t*<sub>R</sub> 5.16 min, *m/z* 587.95 [M+H]<sup>+</sup>, (C<sub>34</sub>H<sub>32</sub>F<sub>3</sub>N<sub>3</sub>O<sub>3</sub><sup>+</sup> Calcd 588.25).

*Benzyl ((S)-4-methyl-1-oxo-1-(((S,E)-5-phenyl-1-(pyridin-2-yl)pent-1-en-3-yl)amino)pentan-2-yl)carbamate (18, Cbz-Leu-hPhe-vinyl-2-Pyr)*. White solid, 0.038 g, 0.078 mmol, 26% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 0.92 (t, 6H, *J* = 6.4 Hz), 1.45 – 1.57 (m, 1H), 1.59 – 1.73 (m, 2H), 1.79 (s, 1H), 1.86 – 2.05 (m, 2H), 2.68 (t, 2H, *J* = 7.9 Hz), 4.03 – 4.27 (m, 1H), 4.71 (pentet, 1H, *J* = 6.9 Hz), 5.10 (s, 2H), 5.14 (s, 1H), 6.26 (d, 1H, *J* = 6.2 Hz), 6.58 (d, 1H, *J* = 15.8 Hz), 6.69 (dd, 1H, *J*<sub>1</sub> = 6.0 Hz, *J*<sub>2</sub> = 15.8 Hz), 7.10 – 7.34 (m, 12H), 7.60 (dt, 1H, *J*<sub>1</sub> = 1.7 Hz, *J*<sub>2</sub> = 7.7 Hz), 8.54 (d, 1H, *J* = 4.6 Hz); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 23.1, 24.9, 32.3, 36.7, 41.2, 50.9, 54.0, 67.3, 122.3, 122.4, 126.1, 128.2, 128.4, 128.6 (2C), 128.7, 130.6, 134.0, 136.3, 136.7, 141.6, 149.7, 155.1, 156.5, 171.6; LC-MS *t*<sub>R</sub> 4.58 min, *m/z* 484.46, 485.49, 486.38 [M+H]<sup>+</sup>, (C<sub>30</sub>H<sub>35</sub>N<sub>3</sub>O<sub>3</sub><sup>+</sup> Calcd 486.28).

*Benzyl ((S)-1-oxo-3-phenyl-1-(((S,E)-4-(pyridin-2-yl)but-3-en-2-yl)amino)propan-2-yl)carbamate (19, Cbz-Phe-Ala-vinyl-2Pyr)*. White solid, 0.590 g, 1.374 mmol, 75% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 1.58 (s, 3H), 2.96 – 3.23 (m, 2H), 4.36 (q, *J* = 7.4 Hz, 1H), 4.63 – 4.80 (m, 1H), 5.10 (d, *J* = 1.4 Hz, 2H), 5.33 (s, 1H), 5.67 (d, *J* = 8.3 Hz, 1H), 6.40 (dd, *J* = 1.3, 15.8 Hz, 1H), 6.51 (dd, *J* = 5.5, 15.8 Hz, 1H), 7.13 (ddd, *J* = 1.1, 4.8, 7.4 Hz, 1H), 7.16 – 7.21 (m, 3H), 7.21 – 7.37 (m, 8H), 7.62 (td, *J* = 1.8, 7.7 Hz, 1H), 8.45 – 8.61 (m, 1H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 20.0, 38.8, 46.4, 56.2, 66.8, 121.8, 122.3, 126.7, 127.7, 128.0, 128.3, 128.7, 129.3, 135.6, 136.4, 137.0, 148.8, 155.0, 170.8; LC-MS *t<sub>R</sub>* = 3.49 min, *m/z* 430.35 [M+H]<sup>+</sup>, (C<sub>26</sub>H<sub>27</sub>F<sub>3</sub>N<sub>3</sub>O<sub>3</sub><sup>+</sup> Calcd 430.21).

*(S,E)-5-((S)-2-(((benzyloxy)carbonyl)amino)-3-phenylpropanamido)-7-(pyridin-2-yl)hept-6-en-1-aminium chloride (20, Cbz-Phe-Lys-vinyl-2Pyr)*. White solid, 0.027 g, 0.048 mmol, 51% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 1.44 – 1.59 (m, 2H), 1.65 – 1.84 (m, 4H), 2.89 – 3.02 (m, 3H), 3.09 – 3.19 (m, 1H), 4.43 (t, 1H, *J* = 7.5 Hz), 4.64 – 4.70 (m, 1H), 5.01 – 5.11 (m, 2H), 6.54 (d, 1H, *J* = 16.1 Hz), 6.92 (dd, 1H, *J<sub>1</sub>* = 5.7 Hz, *J<sub>2</sub>* = 16.1 Hz), 7.11 (t, 1H, *J* = 7.5 Hz), 7.21 (t, 2H, *J* = 7.5 Hz), 7.25 – 7.40 (m, 7H), 7.91 (t, 1H, *J* = 6.8 Hz), 8.10 (d, 1H, *J* = 8.0 Hz), 8.54 (t, 1H, *J* = 8.0 Hz), 8.69 (d, 1H, *J* = 5.2 Hz); <sup>13</sup>C NMR (100 MHz, MeOD) δ 23.8, 27.9, 34.0, 39.1, 40.6, 51.9, 58.2, 67.6, 121.6, 125.9, 126.5, 127.7, 128.5, 128.9, 129.5, 129.7, 130.5, 138.1, 138.4, 142.0, 146.3, 147.8, 151.2, 158.3, 174.1; LC-MS *t<sub>R</sub>* 3.16 min, *m/z* 486.98 [M+H]<sup>+</sup>, 508.94 [M+Na]<sup>+</sup> (C<sub>29</sub>H<sub>35</sub>N<sub>4</sub>O<sub>3</sub><sup>+</sup> Calcd 487.27, C<sub>29</sub>H<sub>35</sub>N<sub>4</sub>O<sub>3</sub>Na<sup>+</sup> Calcd 509.25).

*Benzyl ((S)-1-oxo-3-phenyl-1-(((S,E)-1-phenyl-4-(pyridin-4-yl)but-3-en-2-yl)amino)propan-2-yl)carbamate (21, Cbz-Phe-Phe-vinyl-4Pyr)*. Off-white solid, 0.051 g, 0.109 mmol, 50% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 2.83 (d, *J* = 6.8 Hz, 2H), 3.02 (d, *J* = 7.8 Hz, 2H), 4.36 (s, 1H), 4.86 (s, 1H), 5.08 (s, 2H), 5.90 (s, 1H), 6.13 (s, 1H), 7.10 (dd, *J* = 6.5, 31.0 Hz, 6H), 7.18 – 7.27 (m, 6H), 7.31 (dd, *J* = 3.6, 6.4 Hz, 5H), 8.50 (d, *J* = 6.0 Hz, 2H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 38.4, 41.0, 51.5, 52.1, 67.1, 105.0, 120.9, 126.9, 127.1, 127.9, 128.3, 128.4, 128.5, 128.6, 128.8, 129.3, 129.4, 133.4, 136.4, 150.0, 170.1; LC-MS *t<sub>R</sub>* = 3.49 min, *m/z* 506.29 [M+H]<sup>+</sup>, (C<sub>32</sub>H<sub>31</sub>N<sub>3</sub>O<sub>3</sub><sup>+</sup> Calcd 506.24).

*4-((S,E)-3-((S)-2-(((benzyloxy)carbonyl)amino)-3-phenylpropanamido)-4-phenylbut-1-en-1-yl)-1-methylpyridin-1-ium iodide (22, Cbz-Phe-Phe-vinyl-4PyrNMe)*. Yellow solid, 0.020 g, 0.030 mmol, 87% yield. <sup>1</sup>H NMR (400 MHz, MeOD + CDCl<sub>3</sub>) δ 2.75 – 3.09 (m, 4H), 4.28 – 4.43 (m, 5H), 4.87 (q, 1H, *J* = 6.3 Hz), 5.08 (s, 2H), 6.28 (d, 1H, *J* = 15.9 Hz), 6.83 (dd, 1H, *J<sub>1</sub>* = 5.0 Hz, *J<sub>2</sub>* = 15.9 Hz), 7.16 – 7.35 (m, 15H), 7.82 (d, 2H, *J* = 6.5 Hz), 8.71 (d, 2H, *J* = 6.5 Hz); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub> + MeOD) δ 38.2, 40.1, 47.8, 52.2, 56.3, 66.8, 124.5, 126.8 (2C), 127.6, 128.0, 128.3, 128.4 (3C), 128.5, 129.1, 129.2, 136.1, 136.4, 143.9, 144.7, 153.1, 156.2, 171.3; LC-MS *t<sub>R</sub>* 3.36 min, *m/z* 521.33 [M+H]<sup>+</sup>, (C<sub>33</sub>H<sub>34</sub>N<sub>3</sub>O<sub>3</sub><sup>+</sup> Calcd 521.27).

*Benzyl ((S)-1-oxo-3-phenyl-1-(((S,E)-1-phenyl-4-(pyrimidin-4-yl)but-3-en-2-yl)amino)propan-2-yl)carbamate (23, Cbz-Phe-Phe-vinyl-4Pyrm)*. Off-white solid, 0.800 g, 1.579 mmol, 64% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 2.79 – 2.93 (m, 2H), 3.01 (d, 2H, *J* = 6.9 Hz), 4.35 – 4.48 (m, 1H), 4.93 (pentet, 1H, *J* = 6.7 Hz), 5.04 (s, 2H), 5.49 (s, 1H), 6.16 (d, 1H, *J* = 15.6 Hz), 6.29 (s, 1H), 6.88 (dd, 1H, *J*<sub>1</sub> = 5.8 Hz, *J*<sub>2</sub> = 15.6 Hz), 6.97 (d, 1H, *J* = 4.6 Hz), 7.05 – 7.33 (m, 15H), 8.56 (d, 1H, *J* = 5.2 Hz), 9.06 (s, 1H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 38.5, 40.9, 51.6, 56.6, 67.2, 118.8, 126.9, 127.1, 128.0, 128.2, 128.3, 128.6 (3C), 128.8, 129.4 (2C), 136.2, 136.5, 138.6, 156.1, 157.4, 158.8, 161.5, 170.4; LC-MS *t*<sub>R</sub> 5.27 min, *m/z* 507.35 [M+H]<sup>+</sup>, (C<sub>31</sub>H<sub>30</sub>N<sub>4</sub>O<sub>3</sub><sup>+</sup> Calcd 507.24).

*Benzyl ((S)-1-(((S,E)-4-(oxazol-2-yl)-1-phenylbut-3-en-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (24, Cbz-Phe-Phe-vinyl-2Oxz)*. White solid, 0.020 g, 0.040 mmol, 29% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 2.72 – 2.88 (m, 2H), 2.93 – 3.11 (m, 2H), 4.35 (q, 1H, *J* = 7.0 Hz), 4.88 (pentet, 1H, *J* = 6.7 Hz), 5.05 (s, 2H), 5.30 (s, 1H), 5.99 (d, 1H, *J* = 5.8 Hz), 6.14 (d, 1H, *J* = 16.1 Hz), 6.48 (dd, 1H, *J*<sub>1</sub> = 5.9 Hz, *J*<sub>2</sub> = 16.1 Hz), 7.04 – 7.35 (m, 16H), 7.54 (s, 1H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 38.6, 40.9, 51.5, 56.7, 67.3, 117.3, 127.1, 127.4, 128.2, 128.4, 128.7 (3C), 129.0, 129.4 (2C), 136.2, 136.4 (2C), 137.1, 138.3, 156.1, 160.8, 170.3; LC-MS *t*<sub>R</sub> 5.41 min, *m/z* 496.25 [M+H]<sup>+</sup>, (C<sub>30</sub>H<sub>29</sub>N<sub>3</sub>O<sub>4</sub><sup>+</sup> Calcd 496.22).

*Benzyl ((S)-1-oxo-3-phenyl-1-(((S,E)-1-phenyl-4-(thiazol-2-yl)but-3-en-2-yl)amino)propan-2-yl)carbamate (25, Cbz-Phe-Phe-vinyl-2Thz)*. Light yellow solid, 0.206 g, 0.403 mmol, 54% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 2.75 – 2.89 (m, 2H), 2.99 (d, 2H, *J* = 7.1 Hz), 4.28 – 4.46 (m, 1H), 4.87 (d, 1H, *J* = 6.4 Hz), 4.96 – 5.11 (m, 2H), 5.44 (s, 1H), 6.19 (s, 1H), 6.39 (dd, 1H, *J*<sub>1</sub> = 5.1 Hz, *J*<sub>2</sub> = 16.0 Hz), 6.45 (d, 1H, *J* = 16.0 Hz), 7.01 – 7.39 (m, 16H), 7.71 (d, 1H, *J* = 3.2 Hz); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 38.7, 41.0, 51.5, 56.6, 67.2, 118.5, 123.9, 126.9, 127.2, 128.1, 128.3, 128.6 (2C), 128.8, 129.4 (2C), 135.0, 136.2, 136.5 (2C), 143.4, 156.0, 166.0, 170.4; LC-MS *t*<sub>R</sub> 5.64 min, *m/z* 512.18 [M+H]<sup>+</sup>, (C<sub>30</sub>H<sub>29</sub>N<sub>3</sub>O<sub>3</sub><sup>+</sup> Calcd 512.20).

*2-((S,E)-3-((S)-2-(((benzyloxy)carbonyl)amino)-3-phenylpropanamido)-4-phenylbut-1-en-1-yl)-3-methylthiazol-3-ium iodide (26, Cbz-Phe-Phe-vinyl-2ThzNMe)*. Yellow solid, 0.003 g, 0.005 mmol, 15% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 2.51 (s, 1H), 2.81 – 3.96 (m, 1H), 2.98 – 3.18 (m, 3H), 3.71 (s, 3H), 3.93 (s, 2H), 4.36 – 4.49 (m, 1H), 4.99 (s, 2H), 5.85 – 6.11 (m, 1H), 6.81 (d, 1H, *J* = 15.6 Hz), 6.91 (dd, 1H, *J*<sub>1</sub> = 3.4 Hz, *J*<sub>2</sub> = 15.6 Hz), 7.06 – 7.33 (m, 15H), 7.76 (s, 1H), 7.96 (m, 1H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 38.1, 39.8, 52.5, 54.7, 57.3, 66.7, 115.1, 121.6, 126.9, 127.1, 127.6, 128.0 (2C), 128.5, 128.7, 128.8, 129.5 (2C), 136.9, 137.0, 138.8, 150.6, 156.4, 168.7, 171.9; LC-MS *t*<sub>R</sub> 3.31 min, *m/z* 526.18 [M+H]<sup>+</sup>, (C<sub>31</sub>H<sub>32</sub>N<sub>3</sub>O<sub>3</sub>S<sup>+</sup> Calcd 526.22).

*Benzyl ((S)-1-oxo-3-phenyl-1-(((S)-1-phenyl-4-(pyrimidin-2-yl)butan-2-yl)amino)propan-2-yl)carbamate (27, Cbz-Phe-Phe-(CH<sub>2</sub>)<sub>2</sub>-2Pyrmid).* To a solution of Boc-Phe-vinyl-2Pyrmid (**h**, prepared following GP1-GP7, 0.05g, 0.15mmol) in anh. EtOAc (8mL) was added Pd/C (10%wt, 0.016mg) under H<sub>2</sub> atmosphere and stirred overnight. The reaction was filtered, and the filtrate was concentrated. The product was coupled with P<sub>3</sub>-P<sub>2</sub> fragment following GP8 to give compound **27**. White solid, 0.031 g, 0.060 mmol, 48% yield. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 1.63 - 1.82 (m, 1H), 1.85 - 2.01 (m, 1H), 2.63 - 2.77 (m, 2H), 2.78 - 2.94 (m, 2H), 2.99 (d, 2H, *J* = 6.8 Hz), 4.05 - 4.23 (m, 1H), 4.31 (pentet, 1H, *J* = 7.1 Hz), 5.07 (s, 2H), 5.34 (s, 1H), 6.43 (d, 1H, *J* = 6.4 Hz), 7.04 - 7.38 (m, 16H), 8.59 (d, 2H, *J* = 4.9 Hz); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 31.5, 35.8, 38.8, 41.1, 51.1, 56.8, 67.2, 118.6, 126.6, 127.0, 128.2, 128.3, 128.5, 128.7, 128.8, 129.4, 129.6, 136.4, 136.7, 137.9, 155.9, 157.0, 170.5, 170.9; LC-MS *t<sub>R</sub>* 5.10 min, *m/z* 509.15, 509.28 [M+H]<sup>+</sup>, (C<sub>31</sub>H<sub>32</sub>N<sub>4</sub>O<sub>3</sub>)<sup>+</sup> Calcd 509.26).

**Solubility of Cruzain Inhibitors.** Synthetic PVHIs such as **5-16** and **22-24** at final concentrations of 1-200 μM were added to 0.25 mL solutions of 10% DMSO (v/v) in clear 96-well Greiner plates, and absorbance at 620 nm was measured using a Biotek M2 Synergy plate reader at *t* = 0 and 120 min to evaluate increased light scattering, to evaluate precipitated inhibitor.

**Evaluation of Covalent Adducts of Glutathione and Cruzain Inhibitors.** The compounds **7**, **11**, **12**, **15**, **25**, **26**, and **K1777** (0.5 mM) were added to 100 mM Tris (pH 8.0), 10% (v/v) DMSO, and 1 mM or 5 mM reduced glutathione to a final volume of 0.2 mL at room temperature. Samples were analyzed by HPLC-MS (as described above) by injecting 0.01-mL aliquots onto a Luna 5 mm C18(2) 100 Å, 4.6 mm, 50 mm column (Phenomenex) using the HPLC method prescribed in the chemistry section at 0-6 hr time points. The chromatographic peaks for each cruzain inhibitor and its covalent adduct with glutathione were characterized by their values of *m/z* using electrospray positive-ionization detection and UV absorbance at 254 nm: **K1777**, retention time: 4.75 min, *m/z*: 575.05; **K1777**-GSH, retention time: 4.64 min, *m/z*: 882.30; **7**, retention time: 6.71 min, *m/z*: 507.11; **11**, retention time: 5.48 min, *m/z*: 506.06; **12**, retention time: 4.77 min, *m/z*: 520.10; **15**, retention time: 4.81 min, *m/z*: 535.23; **15**-GSH, retention time: 4.72 min, *m/z*: 842.42; **17**, retention time: 5.03 min, *m/z*: 587.95; **25**, retention time: 4.43 min, *m/z*: 511.87; **26**, retention time: 4.42 min, *m/z*: 525.90; **26**-GSH, retention time: 3.61 min, *m/z*: 833.97. Integration of the chromatographic peaks of the inhibitors and their GSH-adducts at each time point was used to determine the rate of GSH-adduct formation. Integration of the chromatographic peaks of the inhibitors and their GSH-adducts was used at each time point to calculate the concentration of remaining inhibitor and its GSH-adduct.

**Enzyme Expression and Purification.** Recombinant human cathepsins B, L, and S were purchased from Millipore Sigma and used without further treatment. General procedures for cruzain expression, purification

and activation were performed according to published protocols<sup>59</sup> with modifications as described.<sup>40</sup> Activated cruzain at >95% purity was stored either as MMTS-conjugated samples, as described, at -80 °C in a buffer containing 50 mM Tris (pH 8.0) and 20% glycerol, at protein concentrations of 5 mg/mL. Prior to use, MMTS was removed by successive dialysis in the presence of 5 mM DTT. Activated samples of cruzain were used immediately or stored after use at 4 °C for a period of 1-2 months.

**Enzyme Assays and Evaluation of Inhibitors.** All enzyme assays were performed at 25 °C. Initial rates of the peptidolytic reaction catalyzed by cruzain were measured by monitoring the fluorescence generated by cleavage of the dipeptide-AMC bond. Assays were conducted in 96-well plates (Greiner, flat-bottom, clear black plates) in a total volume of 250  $\mu$ L, containing either 50 mM MES (pH 7.5), 50 mM TAPSO, 100 mM DEA, 1 mM CHAPS, 1 mM Na<sub>2</sub>EDTA, 5 mM DTT and 10% DMSO (v/v) or 50 mM sodium acetate (pH 5.5), 50 mM MES, 100 mM TEA, 1 mM CHAPS, 1 mM Na<sub>2</sub>EDTA, 5 mM DTT and 10% DMSO (v/v). Substrates were dissolved in 100% DMSO, and were then diluted 10-fold such that when added to reaction mixtures, final DMSO concentration were 10% (v/v). Reactions were initiated with addition of 1-10  $\mu$ L of cruzain (final concentrations: 0.1 – 3.0 nM (pre-incubation studies)). Fluorescence was measured on either a SpectraMax M5 (Molecular Devices) or a Synergy HTX (Biotek, Wisnooki, VT) microplate reader ( $\lambda_{\text{ex}}$  = 360 nm,  $\lambda_{\text{em}}$  = 460 nm). Initial rates were determined from continuous kinetic time courses, and calculated from the earliest time points, typically at less than 10 min.

Compounds were evaluated as inhibitors or inactivators of cruzain in two ways: (1) enzyme was added to reaction mixtures containing substrate (typically, 10  $\mu$ M Cbz-Phe-Arg-AMC) and inhibitor, and reaction time courses were measured for 0-40 min. In addition to other methods, the effects of all inhibitors on reaction rates were determined at  $t$  = 0-200 s;  $v_i$ ) and at longer incubation times ( $t$  > 1000 s;  $v_s$ ), to ascertain the respective inhibition constants  $K_i$  and  $K_i^*$ . (2) Enzyme and compound were pre-incubated over extended periods of time, and then aliquots were removed and diluted 50- to 100-fold into reaction mixtures containing 10  $\mu$ M of Cbz-Phe-Arg-AMC, followed by the assessment of the resulting time courses.

For assays of cathepsin B, L, and S, cruzain inhibitors were evaluated in reaction mixtures containing a buffer of sodium acetate (pH 5.5), 1 mM CHAPS, 1 mM Na<sub>2</sub>EDTA, and 5 mM DTT at 25 °C. The substrate Cbz-Phe-Arg-AMC was dissolved in 100% DMSO, as were all inhibitors, and aliquots of both substrates and inhibitors were added to 0.25 mL reaction mixtures to final concentrations of 10% DMSO (v/v). Michaelis constants for Cbz-Phe-Arg-AMC were determined for all three human cathepsins as: cathepsin L (2.9  $\mu$ M), cathepsin S (60  $\mu$ M), and cathepsin B (150  $\mu$ M), and fixed concentrations of Cbz-Phe-Arg-AMC of 1 or 2  $K_m$  were used to evaluate inhibitors. Cruzain inhibitors were added at seven concentrations

and one fixed concentration of Cbz-Phe-Arg-AMC, and time courses of AMC formation were analyzed as with cruzain.

**Evaluation of Cruzain Inhibitors in Axenic Cell Cultures of *T. b. brucei* and *T. cruzi*.** Selected cruzain inhibitors were evaluated in axenic cell cultures of *T. b. brucei* and *T. cruzi*. Procyclic trypomastigotes of *T. b. brucei* (ATCC PRA-381) were grown in SDM-79 medium at 26 °C, and bloodstream forms (ATCC PRA-383) were grown at 37 °C in HMI-9 medium at 5% CO<sub>2</sub>. *T. cruzi* (epimastigote forms, strain Y, ATCC 50832GFP) was grown in ATCC medium (1029 LIT medium). Both forms of *T. b. brucei* and *T. cruzi* were grown in media containing 10% fetal bovine serum (FBS) and penicillin/streptomycin (50 U/mL). Test compounds, including **K11777**, were dissolved in 100% DMSO, and added to cell cultures at final concentrations of 0.5-20 μM (maximum DMSO = 1% (v/v)). Control samples contained equal amounts of DMSO. *T. b. brucei* and *T. cruzi* (5 mL in flask cultures) were seeded at ~3 x 10<sup>6</sup> cells, and diluted daily maintaining a mid-log growth phase for up to 120 hrs. Treated cells were typically grown for 4 days (*T. b. brucei*) or 5 days (*T. cruzi*). After each cell dilution, fresh compound or an equal volume of DMSO (control samples), was supplemented into the cultures, while maintaining a constant concentration of each inhibitor. Cell counts were scored using a Z2 Coulter Counter.

**Evaluation of Cruzain Inhibitors in *T. cruzi*-Infected Murine Cardiomyoblasts.** For the evaluation of the anti-trypomastigote activity of cruzain inhibitors, we infected a C2C12 mouse cardiomyoblast cell line (ATCC CRL-1772) with *T. cruzi* strain Ca-I/72 (a gift from James Dvorak, National Institutes of Health) in 1536-microwell plates. In each well was added 10<sup>3</sup> cells and 10<sup>4</sup> parasites in a total volume of 10 μl including the test compounds in 10-point dose-response dilutions starting at 10 μM (3-fold dilutions). The plates were incubated at 37°C for 48h, and the wells were fixed with 2% paraformaldehyde in PBS, and stained with 5 μg/ml of 4',6-diamidino-2-phenylindole. After at least 30 minutes of incubation at room temperature in the dark, the plates were read in an automated microscope, ImageXpress MicroXL (Molecular Devices), and the images were analyzed by custom-built software to quantify and assess viability of the parasites, as well as the host cells independently. The compilation of data was used to calculate the antiparasitic activity (EC<sub>50</sub>) and host cytotoxicity (CC<sub>50</sub>).

**Evaluation of Human Cell Toxicity.** Primary human dermal fibroblast (HDF) cells were used to evaluate human cell toxicity of cruzain inhibitors. HDF cells were plated in a 384-well plate at 2400 cells/well (62,000 cells/mL). Inhibitors in 100% DMSO were added in duplicate to final concentrations of 0.001 – 0.1 mM and 1% DMSO (v/v) with 1% DMSO as a control sample), and cells were cultured at 37°C for 48 hrs, followed by addition of resazurin. Cell viability was then assessed by reading of fluorescence (excitation/ emission: 544 nm/590 nm) after an additional 24 hr of incubation.

**Kinetic Data Analysis.** Initial velocity data for cruzain-catalyzed reactions of fluorogenic peptide substrates were determined by fitting to eq. 1 using GraphPad Prism 6.0 or SigmaPlot 12.0. For eq. 1,  $k_{\text{cat}}$  is the turnover number,  $[E]_t$  is the concentration of active sites of cruzain, and  $K_a$  is the Michaelis constant for the substrate A. Cruzain concentrations were determined by spectrophotometric analysis of purified sample solutions as described.<sup>44</sup>

$$\frac{v}{[E]_t} = \frac{k_{\text{cat}}[A]}{K_a + [A]} \quad (1)$$

Competitive inhibition was fitted to eq 2, in which A and I are concentrations of substrate and inhibitor, respectively,  $V_{\text{max}}$  is the maximal velocity, and  $K_{\text{is}}$  is the slope inhibition constant.

$$v = \frac{V_{\text{max}}[A]}{K_a(1 + [I]/K_{\text{is}}) + [A]} \quad (2)$$

Data for time-dependent inhibition were fitted by several methods. All time-course data were fitted to eq (3) for studies in which reaction was initiated by the addition of enzyme, wherein P is the fluorescence generated by AMC formation, C is a non-zero constant,  $v_s$  and  $v_i$  are respectively the steady-state and initial enzymatic rates,  $t$  is time, and  $k_{\text{obs}}$  is the observed rate of conversion of the initial inhibited rate to the final inhibited rate.<sup>52</sup> In cases for which reaction was initiated with an excess of substrate, following pre-incubation of enzyme and inhibitor, for eq 3,  $v_i = 0$ .

$$P = v_s t + \left[ \frac{v_i - v_s}{k_{\text{obs}}} \right] [1 - e^{(-k_{\text{obs}} t)}] + C \quad (3)$$

Values of  $k_{\text{obs}}$  vs. [inhibitor] were then re-plotted and fitted to eq 4, for which  $k_3$  and  $k_4$  represent the respective rates of formation and dissolution of the EI\* complex as depicted in Scheme 1.

$$k_{\text{obs}} = k_4 + \frac{k_3[I]}{K_i \left( 1 + \frac{[A]}{K_a} \right) + [I]} \quad (4)$$

Inhibition constants were also obtained by fitting  $v_i$  and  $v_s$  data to eq 5, in which  $v_x$  is the rate in the presence of inhibitor I for either early ( $v_i$ ) or late ( $v_s$ ) phases of each time course,  $v_0$  is the rate in the absence of inhibitor,  $K_a$  is the Michaelis constant of the substrate A, and  $K_{\text{ix}}$  is the apparent inhibition constant,  $K_i$  or  $K_i^*$ , obtained from fitting of  $v_i$  or  $v_s$ , respectively.

$$\frac{v_x}{v_0} = \frac{1}{[1 + [I]/K_{\text{ix}} \left( 1 + \frac{[A]}{K_a} \right)]} \quad (5)$$

$$P = \left[ \frac{V_{\max}[A]t}{K_a(1 + \frac{[I]}{K_i(k_4/(k_3 + k_4))}) + [A]} \right] + \left[ \frac{\frac{V_{\max}[A]}{K_a(1 + [I]/K_i) + [A]}}{k_4 + \frac{k_3[I]}{K_i(1 + \frac{[A]}{K_a}) + [I]}} \right] \left[ 1 - \exp\left( \frac{-tk_3[I]}{K_i(1 + \frac{[A]}{K_a}) + [I]} \right) \right] + C \quad (6)$$

In addition, time-course data were fitted globally by two methods. First, expressions for  $v_s$ ,  $v_i$ , and  $k_{\text{obs}}$  were substituted into eq 3 to generate eq 6, which was used to fit time-course data at all concentrations of inhibitors simultaneously. Second, time course data were fitted using Kintek Explorer® as either a single-binding step model with the two kinetic parameters  $k_3 / K_i$  and  $k_4$ , or a two-step binding mechanism as in Scheme 1 from which rate constants  $k_1$ - $k_4$  are determined (Supplemental Information).

## ASSOCIATED CONTENT

### Supporting Information

Molecular modeling information, fitting of time-course inhibition data using Kintek Explorer®, thiolation of cruzain inhibitors by glutathione, structures of PVHIs, HPLC traces and NMR spectra of synthesized substrates and inhibitors, and unpublished crystallographic data. (PDF)

Model coordinates for cruzain bound to **7**, **9**, **11-13** and **15**. (PDB)

Molecular formula strings. (CSV)

## AUTHOR INFORMATION

### Corresponding Author

\*For T. D. M.: Phone: 979-458-9787, E-mail: [tdmeek@tamu.edu](mailto:tdmeek@tamu.edu)

### Present Addresses

‡AstraZeneca Pharmaceuticals, 35 Gatehouse Drive, Waltham, MA 02451, United States.

⊥Department of Chemistry, University of Illinois, 161 RAL, 600 South Matthews Avenue, Urbana, IL 61801, United States.

### ORCID

Thomas D. Meek: 0000-0002-1931-8073

Bala C. Chenna: 0000-0002-2660-5302

Linfeng Li: 0000-0001-6185-4999

Drake Mellott: 0000-0002-8043-8170

## Notes

The authors declare no competing financial interest

## Author Contributions

#Research contributions from L. L. and B. C. C. were equal. The manuscript was written by T.M., L. L., D.M.M., B.C.C., with contributions from X. Z., J. C. R. and J. L. S. N. Synthetic chemistry was performed by B. C. C., L.L., and D. M. M., and enzymatic studies were conducted by L. L., X. Z., D. M. M., and T. M.

## ACKNOWLEDGMENTS

We wish to thank Jasmin Ackermann, Ivani Pauli, for assisting with testing of cruzain inhibitors in *T. cruzi*-infected cardiomyoblasts, Thomas Snavelly, Dr. Deeann Wallis, and Kim Loesch in the laboratory of Prof. James C. Sacchettini for testing cruzain inhibitors for human cell toxicity, Prof. James Sacchettini and Dr. Su Tang for crystallography analysis, Jiyun Zhu, Charis Fernandez and Nora McGuffey for assistance with analysis of inhibitors vs. cruzain and human cathepsins, and Mireya Luna for support with synthetic chemistry. The authors thank Professor Charles S. Craik for providing a construct for the expression of cruzain and Dr. Larry Dangott for providing protein sequencing. The authors also thank Professor Ken Johnson for assistance with the use of Kintek Explorer. Financial support for the research was provided by NIH grant R21AI127634 and Texas A&M AgriLife Research.

## ABBREVIATIONS USED

AMC, 7-amino-4-methylcoumarin; BSF: bloodstream from; CHAPS, 3-[3-(cholamidopropyl)dimethylammonio]-1-propanesulfonate; DIPEA, *N,N*-Diisopropylethylamine; hPhe, homophenylalanine; MES, 2-(*N*-morpholino)ethanesulfonic acid; Oxz, oxazolyl; NMePip, *N*-methylpiperazinyl; PCF, Procyelic form; PVHI, peptidomimetic vinyl-heterocyclic inhibitor; Pyr, pyridinyl; Pyrmd, pyrimidinyl; Thz, thiazolyl; T3P, propylphosphonic anhydride; VS, vinyl sulfone.

## REFERENCES

1. Muñoz-Saravia, S. G.; Haberland, A.; Wallukat, G.; Schimke, I. Chronic Chagas' Heart Disease: a Disease on its Way to Becoming a Worldwide Health Problem: Epidemiology, Etiopathology, Treatment, Pathogenesis and Laboratory Medicine. *Heart Fail. Rev.* **2012**, *17*, 45-64.
2. Centers for Disease Control and Prevention. <https://www.cdc.gov/parasites/chagas> (Accessed October 2019).
3. Coura, J. R.; Viñas, P. A. Chagas Disease: a New Worldwide Challenge. *Nature* **2010**, *465*, S6-S7.
4. Castro, Y.; Marcus, R. .Epidemiology of Chagas Disease in the USA: High-Risk Patient Populations for Screening. *Curr. Trop. Med. Rep.* **2019**, *6*, 8-12.
5. Montgomery, S. P.; Parise, M. E.; Dotson, E. M.; Bialek, S. R. What Do We Know About Chagas Disease in the United States? *Am. J. Trop. Med. Hyg.* **2016**, *95*, 1225-1227.
6. Rassi, A.; Rassi, A.; Marcondes de Rezende, J. American Trypanosomiasis (Chagas Disease). *Infect. Dis. Clin. North. Am.* **2012**, *26*, 275-291.
7. Lejon, V.; Bentivoglio, M.; Franco, J. R. Human African Trypanosomiasis. In *Handbook of Clinical Neurology*; Garcia, H. H.; Tanowitz, H. B.; Del Brutto, O. H., Eds; Elsevier: Cambridge, MA, **2013**; pp 169-181.
8. World Health Organization. [https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-\(sleeping-sickness\)](https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness)) (Accessed October 2019).
9. Drugs for Neglected Diseases initiative. [https://www.dndi.org/wp-content/uploads/2019/09/Factsheet2019\\_SleepingSickness.pdf](https://www.dndi.org/wp-content/uploads/2019/09/Factsheet2019_SleepingSickness.pdf) (Accessed October 2019).
10. Simarro, P. P.; Jannin, J.; Cattand, P. Eliminating Human African Trypanosomiasis: Where Do We Stand and What Comes Next? *PLOS Medicine* **2008**, *5*, e55.
11. Kennedy, P. G. E. Clinical Features, Diagnosis, and Treatment of Human African Trypanosomiasis (Sleeping Sickness). *Lancet Neurol.* **2013**, *12*, 186-194.
12. Salomon, C. J. First Century of Chagas' Disease: An Overview on Novel Approaches to Nifurtimox and Benznidazole Delivery Systems. *J. Pharma. Sci.* **2012**, *101*, 888-894.
13. Cianni, L.; Feldmann, C. W; Gilberg, E.; Gütschow, M., Juliano, L.; Leitão, A.; Bajorath, J.; Carlos A. Montanari, C. A. Can Cysteine Protease Cross-Class Inhibitors Achieve Selectivity? *J. Med. Chem.* **2019**, *62*, 10497-10525.

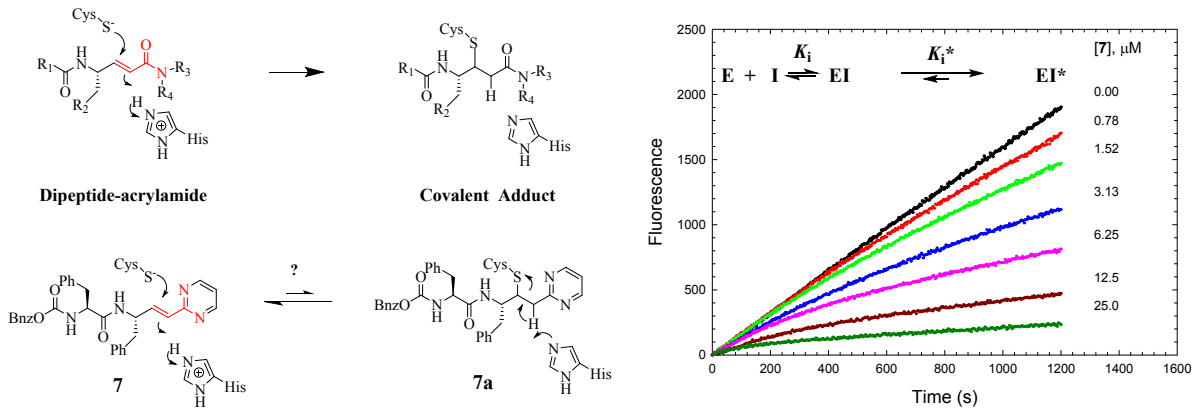
14. Choe, Y.; Leonetti, F.; Greenbaum, D. C.; Lecaille, F.; Bogyo, M.; Bromme, D.; Ellman, J. A.; Craik, C. S. Substrate Profiling of Cysteine Proteases Using a Combinatorial Peptide Library Identifies Functionally Unique Specificities. *J Biol. Chem.* **2006**, *281*, 12824-12832.
15. Lecaille, F.; Authié, E.; Moreau, T.; Serveau, C.; Gauthier, F.; Lalmanach, G. Subsite Specificity of Trypanosomal Cathepsin L-Like Cysteine Proteases. *Eur. J. Biochem.* **2001**, *268*, 2733-2741.
16. Sudhan, D. R.; Siemann, D. W. Cathepsin L Targeting in Cancer Treatment. *Pharmacol. Ther.* **2015**, *155*, 105-116.
17. Miller, B. E.; Mayer, R. J.; Goyal, N.; Bal, J.; Dallow, N.; Boyce, M.; Carpenter, D.; Churchill, A.; Heslop, T.; Lazaar, A. L. Epithelial Desquamation Observed in a Phase I Study of an Oral Cathepsin C Inhibitor (GSK2793660). *Br. J. Clin. Pharmacol.* **2017**, *83*, 2813-2820.
18. Wilkinson R. D.; Williams, R.; Scott C. J.; Burden R. E. Cathepsin S: Therapeutic, Diagnostic, and Prognostic Potential. *Biol. Chem.* **2015**, *396*, 867-882.
19. Drake, M. T.; Clarke, B. L.; Oursler, M. J.; Khosla, S. Cathepsin K Inhibitors for Osteoporosis: Biology, Potential Clinical Utility, and Lessons Learned. *Endocr. Rev.* **2017**, *38*, 325-350.
20. Sajid, M.; Robertson, S. A.; Brinen, L. S.; McKerrow, J. H. Cruzain. in *Cysteine Proteases of Pathogenic Organisms*; Robinson, M. W.; Dalton, J. P., Eds; Springer: Boston, MA, **2011**; pp 100-115.
21. Eakin, A. E.; Mills, A. A.; Harth, G.; McKerrow, J. H.; Craik, C. S. The Sequence, Organization, and Expression of the Major Cysteine Protease (Cruzain) from *Trypanosoma cruzi*. *J. Biol. Chem.* **1992**, *267*, 7411-7420.
22. McGrath, M. E.; Eakin, A. E.; Engel, J. C.; McKerrow, J. H.; Craik, C. S.; Fletterick, R. J. The Crystal Structure of Cruzain: A Therapeutic Target for Chagas' Disease. *J. Mol. Biol.* **1995**, *247*, 251-259.
23. Caffrey, C. R.; Hansell, E.; Lucas, K. D.; Brinen, L. S.; Alvarez Hernandez, A.; Cheng, J.; Gwaltney, S. L.; Roush, W. R.; Stierhof, Y. D.; Bogyo, M.; Steverding, D.; McKerrow, J. H. Active Site Mapping, Biochemical Properties and Subcellular Localization of Rhodesain, the Major Cysteine Protease of *Trypanosoma brucei rhodesiense*. *Mol. Biochem. Parasit.* **2001**, *118*, 61-73.
24. Steverding, D.; Sexton, D. W.; Wang, X.; Gehrke, S. S.; Wagner, G. K.; Caffrey, C. R., *Trypanosoma brucei*: Chemical Evidence that Cathepsin L is Essential for Survival and a Relevant Drug Target. *Int. J. Parasitol.* **2012**, *42*, 481-488.

25. O'Brien, T. C.; Mackey, Z. B.; Fetter, R. D.; Choe, Y.; O'Donoghue, A. J.; Zhou, M.; Craik, C. S.; Caffrey, C. R., A Parasite Cysteine Protease is Key to Host Protein Degradation and Iron Acquisition. *J. Biol. Chem.* **2008**, *283*, 28934-28943.
26. Doyle, P. S.; Zhou, Y. M.; Hsieh, I.; Greenbaum, D. C.; McKerrow, J. H.; Engel, J. C. The *Trypanosoma cruzi* Protease Cruzain Mediates Immune Evasion. *PLOS Pathogens* **2011**, *7*, e1002139.
27. Harth, G.; Andrews, N.; Mills, A. A.; Engel, J. C.; Smith, R.; McKerrow, J. H. Peptide-fluoromethyl Ketones Arrest Intracellular Replication and Intercellular Transmission of *Trypanosoma cruzi*. *Mol. Biochem. Parasit.* **1993**, *58*, 17-24.
28. Doyle, P. S.; Zhou, Y. M.; Engel, J. C.; McKerrow, J. H. A Cysteine Protease Inhibitor Cures Chagas Disease in an Immunodeficient-Mouse Model of Infection. *Antimicrob. Agents Chemother.* **2007**, *51*, 3932-3939.
29. Engel, J. C.; Doyle, P. S.; McKerrow, J. H. Trypanocidal Effect of Cysteine Protease Inhibitors *in vitro* and *in vivo* in Experimental Chagas Disease. *Medicina* **1999**, *59*, 171-175.
30. Engel, J. C.; Doyle, P. S.; Hsieh, I.; McKerrow, J. H. Cysteine Protease Inhibitors Cure an Experimental *Trypanosoma cruzi* Infection. *J. Exp. Med.* **1998**, *188*, 725-734.
31. Kerr, I. D.; Lee, J. H.; Farady, C. J.; Marion, R.; Rickert, M.; Sajid, M.; Pandey, K. C.; Caffrey, C. R.; Legac, J.; Hansell, E.; McKerrow, J. H.; Craik, C. S.; Rosenthal, P. J.; Brinen, L. S. Vinyl Sulfones as Antiparasitic Agents and a Structural Basis for Drug Design. *J. Biol. Chem.* **2009**, *284*, 25697-25703.
32. Chen, Y. T.; Brinen, L. S.; Kerr, I. D.; Hansell, E.; Doyle, P. S.; McKerrow, J. H.; Roush, W. R. *In vitro* and *in vivo* Studies of the Trypanocidal Properties of WRR-483 against *Trypanosoma cruzi*. *PLoS Negl. Trop. Dis.* **2010**, *4*, e825.
33. Mott, B. T.; Ferreira, R. S.; Simeonov, A.; Jadhav, A.; Ang, K. K. H.; Leister, W.; Shen, M.; Silveira, J. T.; Doyle, P. S.; Arkin, M. R.; McKerrow, J. H.; Inglese, J.; Austin, C. P.; Thomas, C. J.; Shoichet, B. K.; Maloney, D. J. Identification and Optimization of Inhibitors of Trypanosomal Cysteine Proteases: Cruzain, Rhodesain, and TbCatB. *J. Med. Chem.* **2010**, *53*, 52-60.
34. McKerrow, J. H.; Doyle, P. S.; Engel, J. C.; Podust, L. M.; Robertson, S. A.; Ferreira, R.; Saxton, T.; Arkin, M.; Kerr, I. D.; Brinen, L. S.; Craik, C. S. Two Approaches to Discovering and Developing New Drugs for Chagas Disease. *Mem. Inst. Oswaldo Cruz* **2009**, *104*, 263-269.
35. McKerrow, J. H. Update on Drug Development Targeting Parasite Cysteine Proteases. *PLoS Negl. Trop. Dis.* **2018**, *12*, e0005850.

36. Bradshaw, J. M.; McFarland, J. M.; Paavilainen, V. O.; Bisconte, A.; Tam, D.; Phan, V. T.; Romanov, S.; Finkle, D.; Shu, J.; Patel, V.; Ton, T.; Li, X.; Loughhead, D. G.; Nunn, P. A.; Karr, D. E.; Gerritsen, M. E.; Funk, J. O.; Owens, T. D.; Verner, E.; Brameld, K. A.; Hill, R. J.; Goldstein, D. M.; Taunton, J. Prolonged and Tunable Residence Time Using Reversible Covalent Kinase Inhibitors. *Nat. Chem. Biol.* **2015**, *11*, 525-531.
37. Krishnan, S.; Miller, R. M.; Tian, B.; Mullins, R. D.; Jacobson, M. P.; Taunton, J. Design of Reversible, Cysteine-Targeted Michael Acceptors Guided by Kinetic and Computational Analysis. *J. Amer. Chem. Soc.* **2014**, *136*, 12624-12630.
38. Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug-Target Residence Time and its Implications for Lead Optimization. *Nature Reviews Drug Discovery* **2006**, *5*, 730-739.
39. Barr, S. C.; Warner, K. L.; Kornreic, B. G.; Piscitelli, J.; Wolfe, A.; Benet, L.; McKerrow, J. H. A Cysteine Protease Inhibitor Protects Dogs from Cardiac Damage during Infection by *Trypanosoma cruzi*. *Antimicrob. Agents Chemother.* **2005**, *49*, 5160-5161.
40. Zhai, X.; Meek, T. D. Catalytic Mechanism of Cruzain from *Trypanosoma cruzi* As Determined from Solvent Kinetic Isotope Effects of Steady-State and Pre-Steady-State Kinetics. *Biochemistry* **2018**, *57*, 3176-3190.
41. Silverman, R. B. Mechanism-Based Enzyme Inactivators. In *Methods in Enzymology*; Johnson, M. L.; Brand, L., Eds; Academic Press: Cambridge, MA, **1995**; pp 240-283.
42. Brady, K.; Abeles, R. H. Inhibition of Chymotrypsin by Peptidyl Trifluoromethyl Ketones: Determinants of Slow-Binding Kinetics. *Biochemistry* **1990**, *29*, 7608-7617.
43. Rubach, J. K.; Cui, G.; Schneck, J. L.; Taylor, A. N.; Zhao, B.; Smallwood, A.; Nevins, N.; Wisnoski, D.; Thrall, S. H.; Meek, T. D. The Amino-Acid Substituents of Dipeptide Substrates of Cathepsin C Can Determine the Rate-Limiting Steps of Catalysis. *Biochemistry* **2012**, *51*, 7551-7568.
44. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739-1749.
45. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47*, 1750-1759.

46. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein–Ligand Complexes. *J. Med. Chem.* **2006**, *49*, 6177-6196.
47. Zhu, K.; Borrelli, K. W.; Greenwood, J. R.; Day, T.; Abel, R.; Farid, R. S.; Harder, E. Docking Covalent Inhibitors: A Parameter Free Approach To Pose Prediction and Scoring. *J. Chem. Inf. Model.* **2014**, *54*, 1932-1940.
48. Wittig, G.; Schöllkopf, U. Über Triphenyl-phosphin-methylene als Olefinbildende Reagenzien (I. Mitteil. *Chem. Ber.* **1954**, *87*, 1318-1330.
49. Wadsworth, W. S.; Emmons, W. D. The Utility of Phosphonate Carbanions in Olefin Synthesis. *J. Am. Chem. Soc.* **1961**, *83*, 1733-1738.
50. Nahm, S.; Weinreb, S. M. N-methoxy-N-Methylamides as Effective Acylating Agents. *Tetrahedron* **1981**, *22*, 3815-3818.
51. Klaus, J. L.; Palmer, J. T.; Rasnick, D. Irreversible Cysteine Protease Inhibitors Containing Vinyl Groups Conjugated to Electron Withdrawing Groups. U.S. Patent 6,287,840, **2001**.
52. Morrison, J. F.; Walsh, C. T. The Behavior and Significance of Slow-Binding Enzyme Inhibitors. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1988**, *61*, 201-300.
53. Jones, B. D.; Tochowicz, A.; Tang, Y.; Cameron, M. D.; McCall, L. I.; Hirata, K.; Siqueira-Neto, J. L.; Reed, S. L.; McKerrow, J. H.; Roush, W. R. Synthesis and Evaluation of Oxyguanidine Analogues of the Cysteine Protease Inhibitor WRR-483 against Cruzain. *ACS Med. Chem. Lett.* **2015**, *7*, 77-82.
54. Ndao, M.; Beaulieu, C.; Black, W. C.; Isabel, E.; Vasquez-Camargo, F.; Nath-Chowdhury, M.; Massé, F.; Mellon, C.; Methot, N.; Nicoll-Griffith, D. A. Reversible Cysteine Protease Inhibitors Show Promise for a Chagas Disease Cure. *Antimicrob. Agents Chemother.* **2014**, *58*, 1167-1178.
55. Boudreau, P. D.; Miller, B. W.; McCall, L. I.; Almaliti, J.; Reher, R.; Hirata, K.; Le, T.; Siqueira-Neto, J. L.; Hook, V.; Gerwick, W. H. Design of Gallinamide A Analogs as Potent Inhibitors of the Cysteine Proteases Human Cathepsin L and *Trypanosoma cruzi* Cruzain. *J. Med. Chem.* **2019**, *62*, 9026-9044.
56. Zhao, Z.; Bourne, P. E. Progress with Covalent Small-Molecule Kinase Inhibitors. *Drug Disc. Today* **2018**, *23*, 727-735.
57. Nkemgu, N. J.; Grande, R.; Hansell, E.; McKerrow, J. H.; Caffrey, C. R.; Steverding, D. Improved Trypanocidal Activities of Cathepsin L Inhibitors. *Int. J. Antimicrob. Agents* **2003**, *22*, 155-159.

- 1  
2  
3 58. Yang, P. Y.; Wang, M.; He, C. Y.; Yao, S. Q. Proteomic Profiling and Potential Cellular Target  
4 Identification of K11777, a Clinical Cysteine Protease Inhibitor, in *Trypanosoma brucei*. *Chem. Comm.*  
5 **2012**, 48, 835-837.  
6  
7  
8  
9 59. Lee, G. M.; Balouch, E.; Goetz, D. H.; Lazic, A.; McKerrow, J. H.; Craik, C. S. Mapping Inhibitor  
10 Binding Modes on an Active Cysteine Protease via Nuclear Magnetic Resonance Spectroscopy.  
11 *Biochemistry*, **2012**, 51, 10087-10098.  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



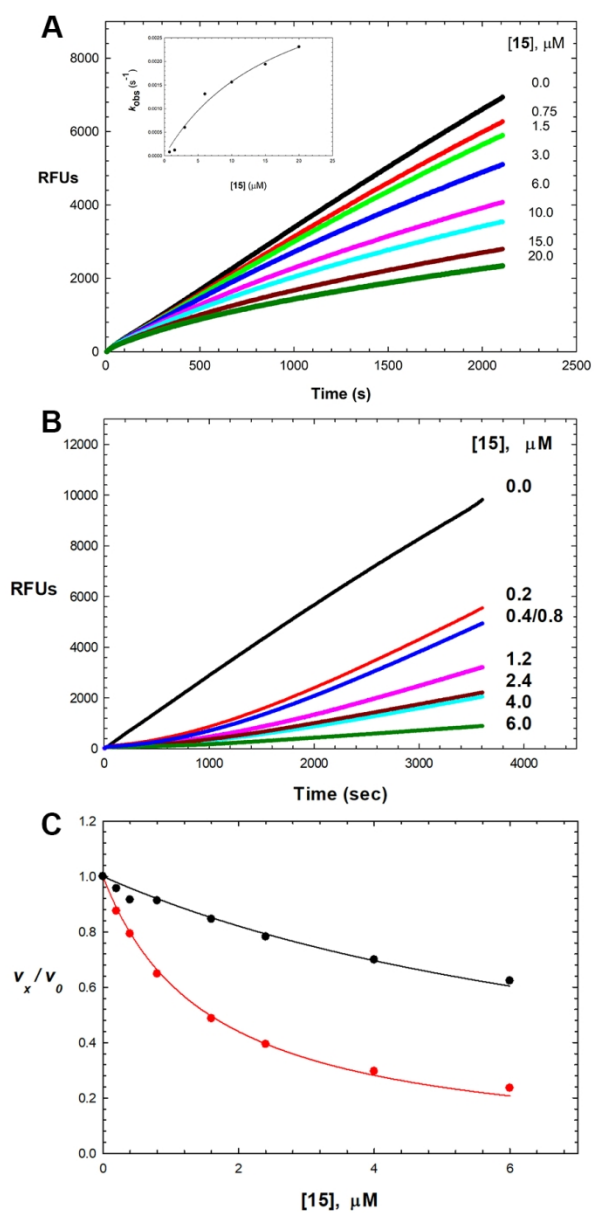


Figure 4

76x152mm (300 x 300 DPI)

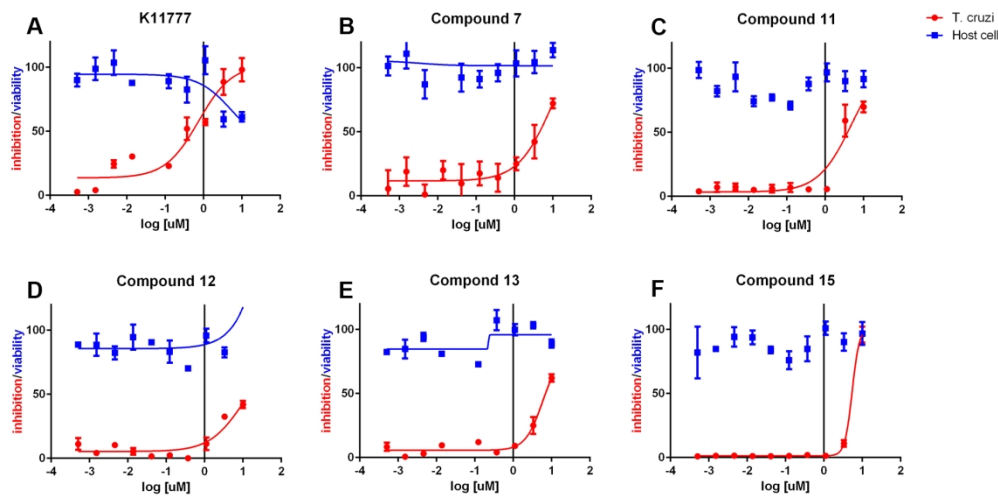


Figure 5

168x83mm (300 x 300 DPI)

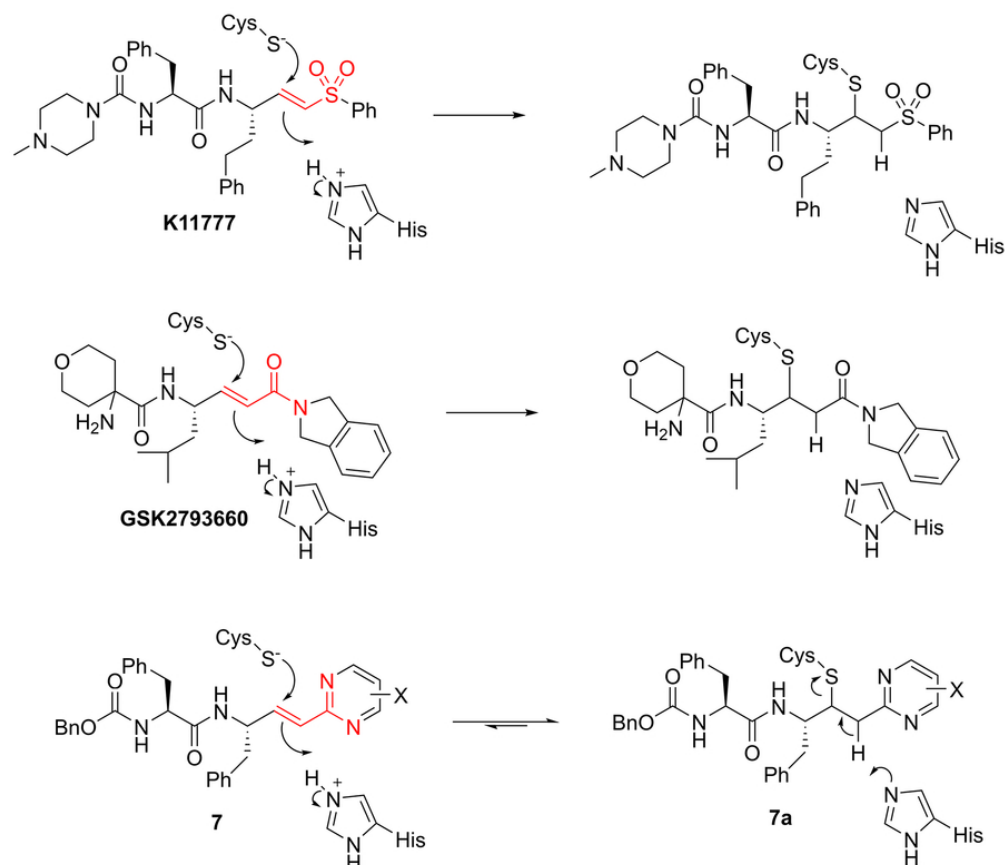
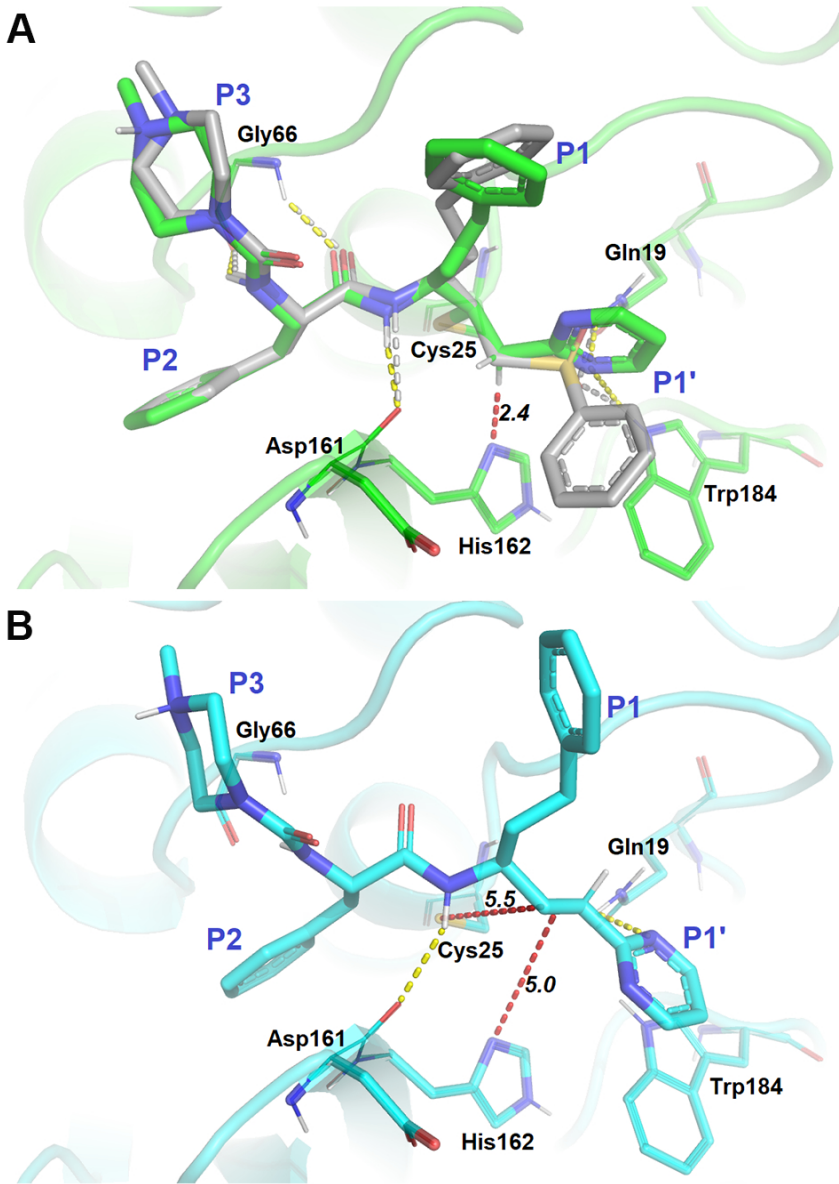


Figure 1

76x65mm (300 x 300 DPI)



81x114mm (300 x 300 DPI)

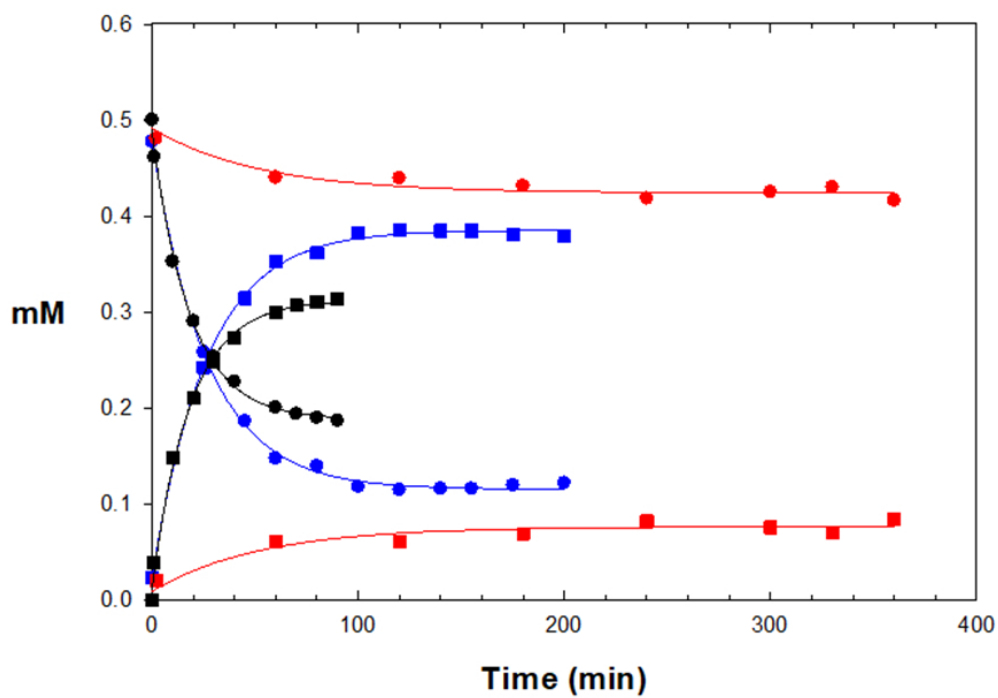


Figure 3

73x51mm (300 x 300 DPI)

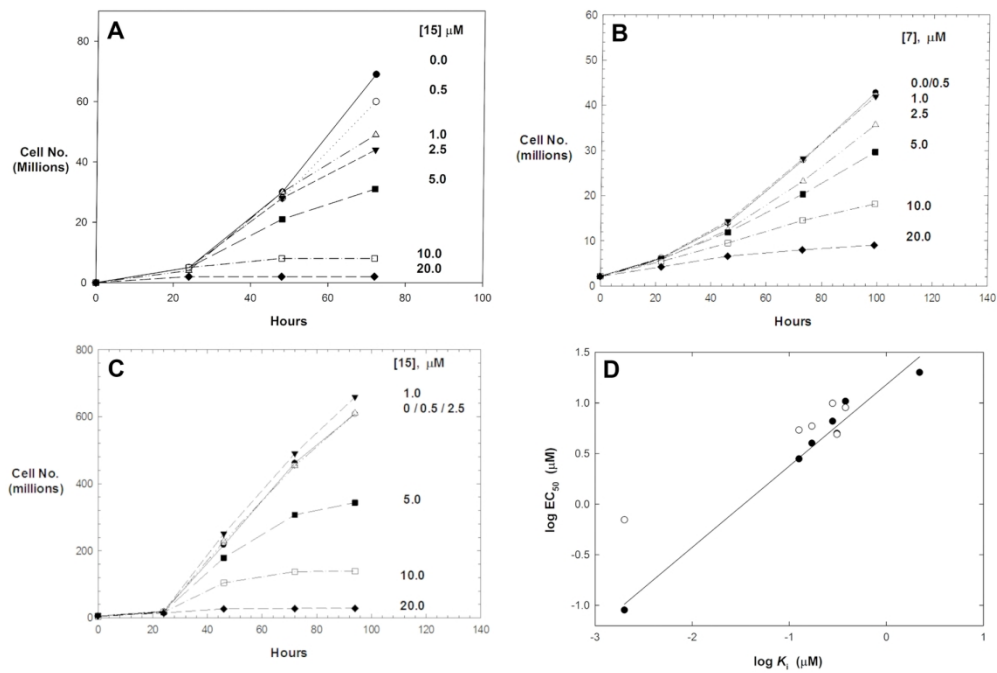


Figure 6

161x108mm (300 x 300 DPI)

