Supporting Information for:

Mapping Inhibitor Binding Modes on an Active Cysteine Protease via NMR Spectroscopy

Gregory M. Lee[1], Eaman Balouch[1], David H. Goetz[1], Ana Lazic[1],
James H. McKerrow[2], and Charles S. Craik[1]*

[1] Department of Pharmaceutical Chemistry
University of California, San Francisco, CA 94158-2280
USA

[2] Department of Pathology and
Center for Discovery and Innovation in Parasitic Diseases
University of California, San Francisco, CA, 94158-2250
USA

*Correspondence should be addressed to C.S.C.
email: craik@cgl.ucsf.edu
Tel: (415) 476-8146

Short title: Cruzain-inhibitor NMR assay

SUPPORTING INFORMATION TABLE OF CONTENTS

SUPPORTING METHODS

Modifications to auto-induction media recipes[1]

For the initial growth tests, unmodified auto-induction recipe ZYP-5052[1] was used to express unlabeled cruzain. In the case of the N-5052 (uniform ^{15}N-labeling) and PA-5052 (selective ^{15}N-Cys, ^{15}N-His, and ^{13}C-Met labeling) auto-induction media recipes,[1] the base buffer consisted of 6.8 g Na_2HPO_4, 3.0 g KH_2PO_4 (~ 70 mM phosphate, pH ~ 7.0) and 1.5 g NaCl (~ 25 mM) dissolved in 1 L water. No Na_2SO_4 was used in this base buffer. Following autoclaving and cooling to room temperature, the solution was supplemented with 2 mM $MgSO_4$, 30 μM $CaCl_2$, 1.0 mL vitamin solution, and 2 mL "O" solution containing trace metals.[2] Carbon sources in the form of sterile-filtered 0.5% glycerol (v/v), 0.05% glucose (w/v), and 0.2% lactose (w/v) were as recommended[1] and also added after the autoclaved buffer cooled. In the case of uniform ^{15}N-labeling (recipe N-5052), the sole nitrogen source was 2.7 g $^{15}NH_4Cl$ (50 mM). In the case of selective ^{15}N- or ^{13}C-labeling (recipe PA-5052), 26.6 mL of a sterile-filtered 0.75% (w/v) stock amino acid solution was added to the growth media, with a final concentration of 200 mg/L of each unlabeled amino acid. Due to limited solubility, 200 mg tyrosine was added directly to 1 L media. In addition, the stock solution did not contain cysteine, histidine, and methionine. Depending on the selective labeling scheme, 100 mg/L ^{15}N-Cys or ^{15}N-His, or 250 mg/L ^{13}C-Met, and 200 mg/L of the other remaining unlabeled amino acids were dissolved in a 30-40 mL aliquot of the auto-induction media, which was then reintroduced via sterile filtering. In the case of uniform ^{13}C/^{15}N/^{2}H-labeling (recipe C-750501), the labeled rich media mixture was supplemented with 0.05% (w/v) ^{13}C/^{2}H-labeled glucose (Cambridge Isotope Laboratories) and 0.75% (w/v) unlabeled glycerol and 0.01% (w/v) unlabeled lactose

NMR Spectral Parameters and Processing

Spectral offsets (widths) for the 3D HNCACB, HN(CO)CACB, CC(CO)NH experiments were as follows: ^{1}H, 4.703 ppm (16.02 ppm); ^{13}C, 39.0 ppm (75.0 ppm); ^{15}N, 117.0 ppm (40.0 ppm). 12 transients, with 32 dummy scans used to reach thermal equilibrium, and

1024 ^{1}H, 44 ^{13}C, and 128 ^{15}N complex data points were collected for each experiment. Data was collected in Echo-AntiEcho mode for the ^{13}C-dimension, States-TPPI mode for the ^{15}N-dimension and with ^{2}H-decoupling applied. General parameters for the 3D-HNCO experiment were the same as for the other triple-resonance experiments with the exception of ^{13}C: 173.0 ppm offset, 14.0 ppm spectral width, and 72 complex data points. The total acquisition time for all four of the 3D-triple resonance experiments was approximately 100 hours. The ^{1}H and ^{15}N spectral offsets (widths) for the 3D-^{15}N/^{1}H-NOESY-HSQC were as above. 32 transients, 32 dummy scans, and 1024 (t3) ^{1}H, 48 (t1) ^{1}H, and 128 ^{15}N complex data points were collected. The NOESY mixing time was set to 120 ms, and the total acquisition time was approximately 67 hours.

Spectral offsets and widths for the ^{15}N-HSQC and ^{13}C-HSQC titration experiments were as follows: ^{1}H, 4.70 ppm (16.0 ppm for ^{15}N, 12.0 ppm for ^{13}C); ^{15}N, 117.0 ppm (34.0 ppm); ^{13}C, 17.0 ppm (18.0 ppm). 32 to 128 transients, 16 dummy scans, and 1024 ^{1}H and 96 ^{15}N or ^{13}C complex data points were collected for each spectrum. The acquisition time for the 2D-HSQC experiments ranged from 1 to 4 hours each. Polynomial baseline correction to deconvolute residual ^{1}HO^{2}H signals, squared sine bell apodization functions and zero-filling were applied to both dimensions of the raw data using NMRpipe[3] prior to Fourier transformation. Linear prediction was also applied to the indirectly-detected ^{15}N or ^{13}C dimensions as needed.

NMR-based pH Titrations

NMR sample preparation was as described above. Initial sample concentrations were 0.050 – 0.075 mM selectively ^{15}N-His, ^{15}N-Cys, or ^{13}C-Met labeled cruzain in 0.5 mL 20 mM phosphate buffer. HSQC spectra were acquired from pH 3 – 10 at approximately 0.5 pH unit intervals. Prior to acquiring the first pH data point, cruzain samples were inhibited with either MMTS or K777 to prevent self-proteolysis. 1 mM DTT (final concentration) was added to the cruzain-K777 samples. Spectra of the MMTS-inhibited cruzain samples were acquired without additional reducing agents. Determination of apparent pK$_a$ values of the selectively labeled ^{15}N-His, ^{15}N-Cys, and ^{13}C-Met residues were performed using the Ekin module of PEAT_DB,[4] using previously described equations and curve fitting models.[5] Estimated experimental errors of ± 0.1 pH units and ± 0.05 ppm were used for the individual residue curve fittings, each performed in triplicate. Reported pK_a values represent the average and propagated errors calculated for the proton and either nitrogen or carbon titration curves.

Cruzain Residue Numbering

There are currently two main residue numbering systems associated with cruzain. The first, we have termed the "classical" system, is based upon papain residue numbers.[6] The second, termed the "sequential" system, is used in the published cruzain-K777 crystal structure (2OZ2)[7] and for many subsequent cruzain-inhibitor structures published since 2009. In the "classical" numbering system, the catalytic residues are identified as Gln19, Cys25, His159, and Asn175. In the "sequential" numbering system, these residues are Gln19, Cys25, His162, and Asn182. The cruzain residue numbering used herein uses the "sequential" system.

SUPPORTING RESULTS

K777 can limit *in vitro* cruzain self-activation.

To determine whether *in vitro* self-activation of the zymogen can be inhibited, five-fold stoichiometric excess of K777 was added to MMTS- and PMSF-inhibited procruzain prior to activation with DTT (**Figure 1c**). Examination of the SDS-PAGE gel indicates that the untreated procruzain undergoes proteolysis of the 14 kDa pro-region segment after 30 minutes incubation time. Conversely, the K777-treated procruzain remains relatively intact after 3 hours. Weak bands with approximate molecular weights larger than activated cruzain appear at the 30 – 60 minute mark. This result implies that excess DTT may also form thiol adducts to K777, thereby preventing inhibitor binding to the zymogen, and that the remaining uninhibited procruzain retains basal proteolytic activity. Therefore, pre-treatment of procruzain with K777 can impede, but not completely abolish self-activation under the current conditions. The higher molecular weight bands observed in the SDS-PAGE gel also suggest that sections of the pro-region are proteolyzed in a step-wise manner, eventually leading to the mature catalytic domain sequence.

Apparent cruzain pK_a values.

In addition to helping probe inhibitor binding modes to cruzain, the selective ^{15}N-Cys, ^{15}N-His, and ^{13}C-Met labeled cruzain samples were used to determine apparent pK_a values in an effort to gain further insight into general cysteine protease mechanisms (**Supplemental Figs. S8-S11**). Because the apo-form of cruzain was too unstable, the protease was inhibited with MMTS, and the HSQC data were acquired in the absence of any reducing agents. The resulting methanethiol adduct is a minimal modification of the catalytic Cys25 thiol group and serves as a proxy for the apo state. The averaged pK_a values determined from the proton and heteronucleus titration curves are reported in **Supplemental Table S2**. Several residues also displayed biphasic pH titration curves. In these cases, the “primary pK_a value” (pK_{a1}) and “secondary pK_a value” (pK_{a2})

described below would correspond to the acidic and basic pH arms of the titration curves, respectively.

Local inter-residue ionization effects are observed for the histidine residues (**Figure S8**). For example, His115 exhibits a biphasic titration curve for the MMTS-inhibited cruzain spectra, giving rise to two apparent pK_a values ($pK_{a1} = 4.36 \pm 0.46$; $pK_{a2} = 7.42 \pm 0.05$). A less dramatic biphasic titration curve is also observed for the catalytic His162 ($pK_{a1} = 3.98 \pm 0.51$; $pK_{a2} = 6.75 \pm 0.07$). In the case of the cruzain-K777 crystal structure,[7] the His115 side chain is in contact with those of Glu73 and Glu117, while His162 is adjacent to Asp161. The acidic pK_a value of His162 in the MMTS-inhibited cruzain sample agrees, within experimental error, with that reported for the catalytic His159 residue of MMTS-inhibited papain (3.45 ± 0.07),[8] but is significantly lower than active papain (8.34 ± 0.04).[9] These differences in apparent pK_a values may be attributed to the absence of the Cys25-His162 thiolate-imidazolium ion pair, which in the active state, is known to be critical for papain catalysis.[10] The remaining two histidine residues are remotely located from other ionizable residues and display monophasic titration curves with higher pK_a values (His43, $pK_a = 8.04 \pm 0.04$; His106, $pK_a = 7.88 \pm 0.02$).

Inhibition of cruzain with K777 induces significant perturbations in the pH titration curves and associated pK_a values for His162 ($pK_{a1} = 5.35 \pm 0.25$, $pK_{a2} = 9.13 \pm 0.74$). This result is likely a reflection of the K777 sulfonyl group being in close proximity (< 3.4 Å) to the histidine side chain imidazole. As expected, the pH titration curves for histidine residues distally located from the cruzain active site display no significant perturbations between the MMTS and K777-inhibited states.

Because the cysteine or methionine residues have no ionizable functional groups under the current conditions, their respective amide (Cys) or methyl (Met) pH titration curves (**Supplemental Figs. S9-S11**) reflect local inter-residue ionization effects imparted by nearby acidic or basic residues.[5] In particular, the ^{15}N-Cys and ^{13}C-Met pH titration curves exhibited small chemical shift perturbation ranges and higher degrees of uncertainty relative to the ^{15}N-His data. Notably, the apparent pKa value obtained for the catalytic Cys25 of MMTS-inhibited cruzain ($pK_a = 3.83 \pm 0.52$) is within experimental error with the reported value of uninhibited papain ($pK_a = 3.32 \pm 0.01$).[9] As expected, overall differences in the pH titration curves between the MMTS- and K777-inhibited cruzain samples are more significant for the residues located within or proximal to the active site.

With the exception of Cys101 (pK_a = 7.06 ± 0.13), all the cysteine residues exhibit primary pK_a values less than pH 6.0 in both the MMTS- and K777-inhibited cruzain samples. Although the majority of the cysteine residues are located in the predominantly negatively charged regions of the protease, close contacts between Cys101 and the side chain of Lys58 may increase its apparent pK_a value. Several of the cysteine residues also displayed secondary pK_a values greater than pH 7.0 (**Supplemental Figs. S9-S10**). Cys25 (pK_a = 3.83 ± 0.52) and Cys155 (pK_a = 3.82 ± 0.36) have the most acidic pK_a values of any of the cysteine groups in the cruzain-MMTS complex. Cysteine residues that had the most dramatic chemical shift perturbations between the MMTS- and K777-inhibited states (Cys22, Cys25, and Cys63) also exhibited the largest differences in their respective apparent pK_a values.

The ^{13}C-Met pH titration curves displayed larger chemical shift perturbation ranges relative to the ^{15}N-Cys data (**Supplemental Fig. S11**). As with the cysteine residues, the methionines are located in the predominantly negatively-charged region of cruzain. However, the apparent pK_a values of the methionines are higher than those of the cysteine residues. Of the methionines, Met68 exhibits the most dramatic change in apparent pK_a values between the MMTS- and K777-inhibited states (MMTS, pK_{a1} = 6.53 ± 0.35; K777, pK_a = 7.83 ± 0.28). Importantly, the chemical shift perturbations observed for the resonance peak of the Met68 methyl group may reflect the ionization state of the Glu208 side chain. Both residues are in contact with each other, helping to form the critical S2 pocket in cruzain.**[11]**

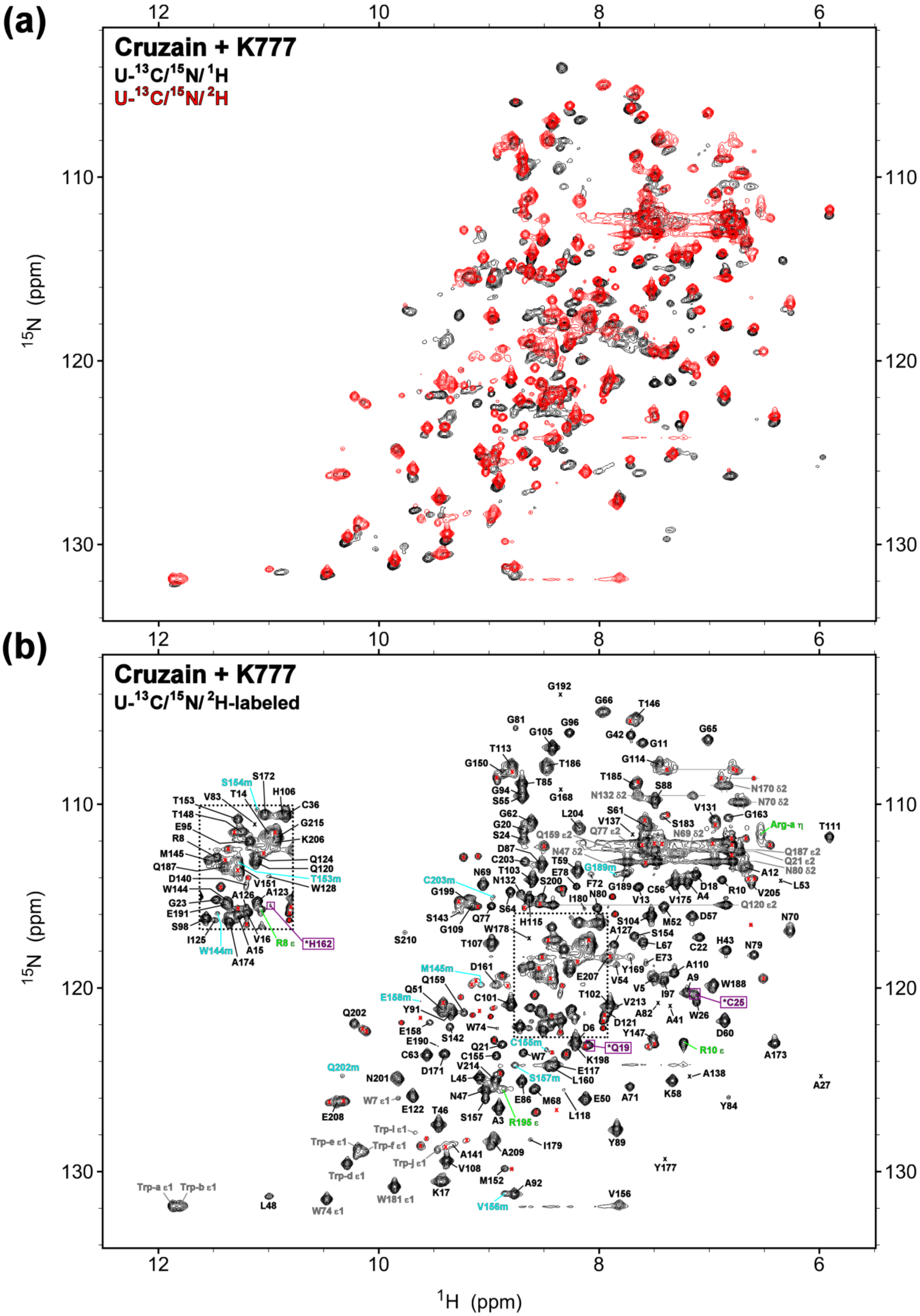
(a)
Cruzain + K777
U-^{13}C/^{15}N/^{1}H
U-^{13}C/^{15}N/^{2}H
^{15}N (ppm)
12
10
8
6
110
120
130
(b)
Cruzain + K777
U-^{13}C/^{15}N/^{2}H-labeled
^{1}H (ppm)
^{15}N (ppm)
G192
G66
T146
G81
G105
G96
G42
G11
G65
T113
T186
G114
G150
T85
G94
S55
G168
T185
S88
N132 δ2
N170 δ2
N70 δ2
V131
G163
G62
L204
S61
G20
V137
S183
Arg-a η
T111
S24
Q159 ε2
Q77 ε2
N69 δ2
D87
N47 δ2
Q187 ε2
C203
T59
G189m
Q21 ε2
N80 δ2
N69
T103
E78
A12
C203m
N132
S200
F72
G189
C56
D18
R10
L53
G199
S64
I180
N80
V13
V175
A4
V205
Q120 ε2
S143
Q77
H115
S104
M52
D57
N70
G109
W178
A127
S154
S210
T107
L67
C22
H43
N79
E73
D161
Y169
A110
M145m
E207
V54
A9
W188
Q159
V5
Q51
C101
T102
I97
E158m
V213
*C25
Y91
A82
W26
D121
A41
D60
Q202
E158
W74
D6
Y147
S142
E190
C155m
*Q19
R10 ε
A173
Q21
C63
C155
W7
K198
D171
E117
V214
S157m
A138
Q202m
N201
L45
L160
K58
A27
W7 ε1
E122
N47
E86
M68
E50
A71
Y84
E208
S157
A3
R195 ε
L118
T46
Trp-i ε1
I179
Y89
Trp-e ε1
Trp-f ε1
A141
Trp-j ε1
V108
A209
Trp-d ε1
M152
Y177
Trp-a ε1
Trp-b ε1
A92
V156
L48
W74 ε1
W181 ε1
K17
V156m
S154m
S172
T14
H106
T153
V83
C36
T148
G215
E95
K206
R8
M145
Q124
Q187
Q120
D140
T153m
W144
V151
W128
G23
A126
A123
E191
S98
I125
V16
R8 ε
*H162
W144m
A15
A174

Figure S1. ^{15}N-^{1}H HSQC spectra of protonated and deuterated cruzain in complex with K777. (**a**) The HSQC spectrum of uniformly ^{13}C/^{15}N/^{2}H-labeled cruzain (red) superimposed over that of uniformly ^{13}C/^{15}N/^{1}H-labeled cruzain (black), both in complex with K777. Missing resonance peaks in the deuterated cruzain-K777 sample indicate incomplete ^{2}H to ^{1}H back-exchange of the backbone amide groups during protein purification. (**b**) The annotated ^{15}N-^{1}H HSQC spectrum of uniformly ^{13}C/^{15}N/^{2}H-labeled cruzain in complex with unlabeled K777. Positions of the catalytic Gln19, Cys25, and His162 resonance peaks are indicated by purple boxes. The central region of the spectrum (dotted box) is annotated in the inset. Black X's indicate assigned resonances not observed in the deuterated cruzain spectra, but are present in the protonated cruzain. Red X's indicate unassigned resonance peaks. Gray annotations denote sidechain Trp, Gln, and Asn NH groups; cyan annotations denote minor conformers; and green annotations denote folded peaks corresponding to Arg sidechain NH groups.

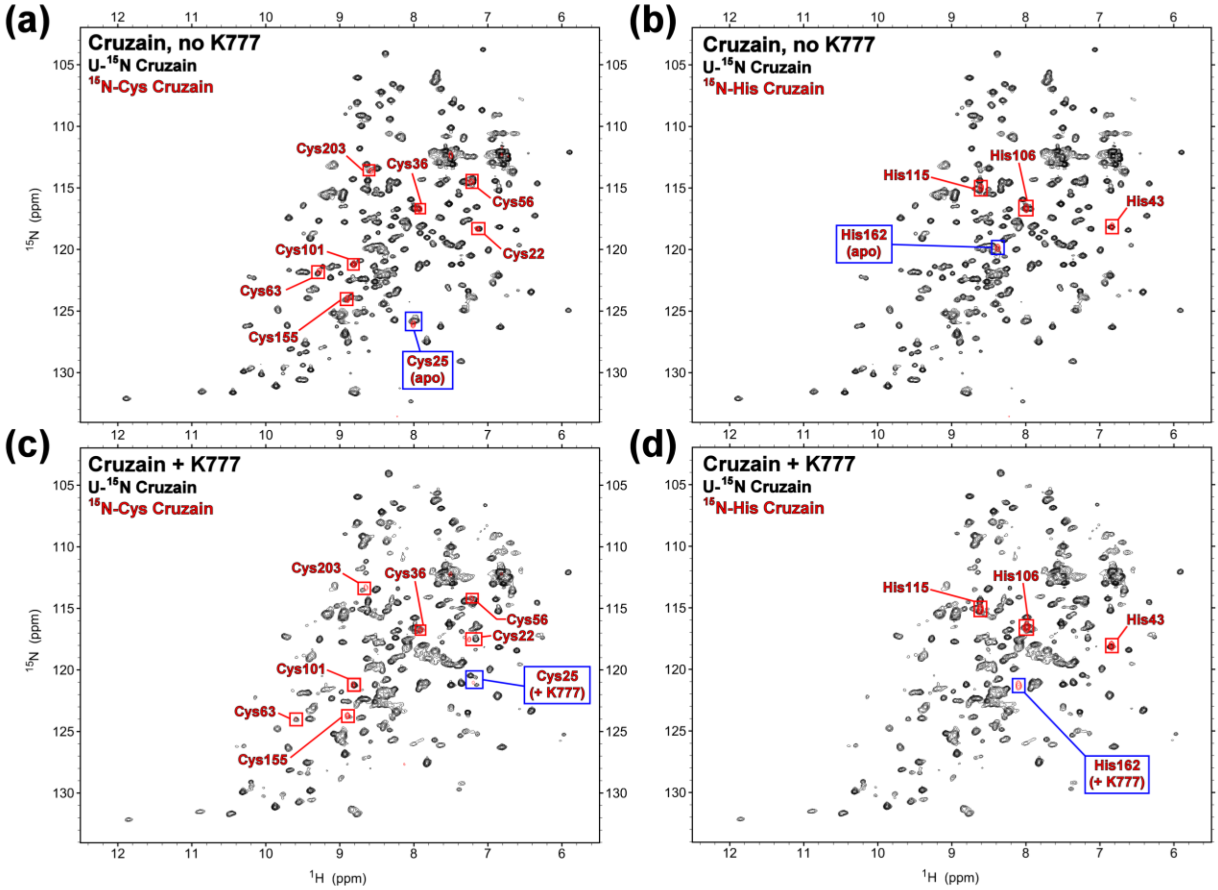

Figure S2: Selectively labeled Cys and His resonances in the ^{15}N/^{1}H-HSQC spectrum. Spectral overlays of the uniformly ^{15}N-labeled cruzain (black) and selectively **(a)** ^{15}N-Cys and **(b)** ^{15}N-His labeled cruzain (red) in their respective apo states. Spectral overlays of the uniformly ^{15}N-labeled cruzain (black) and selectively **(c)** ^{15}N-Cys and **(d)** ^{15}N-His labeled cruzain (red) in their respective K777-inhibited forms.

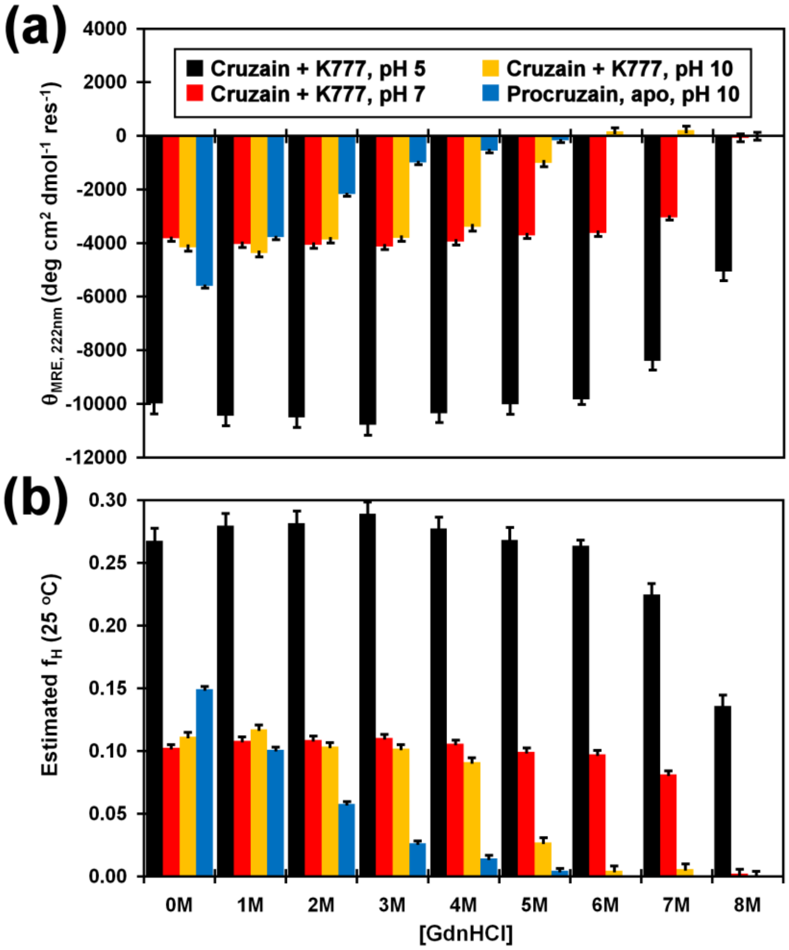

Figure S3. CD denaturation study of cruzain-K777 and procruzain. **(a)** Data corresponding to the θ_{222} bands displayed in **Figure 2** normalized and converted to units of mean residue ellipticity (MRE) as a function of final guanidinium hydrochloride concentration. **(b)** θ_{222} data converted to estimated fractional helicity (f_H) values. Both sets of data indicate that the cruzain-K777 exhibits enhanced helicity under acidic conditions, and is relatively stable against chemical denaturation. Conversely, the apo form of procruzain at pH 10 is structurally labile against chemical denaturation relative to its inhibited counterpart.

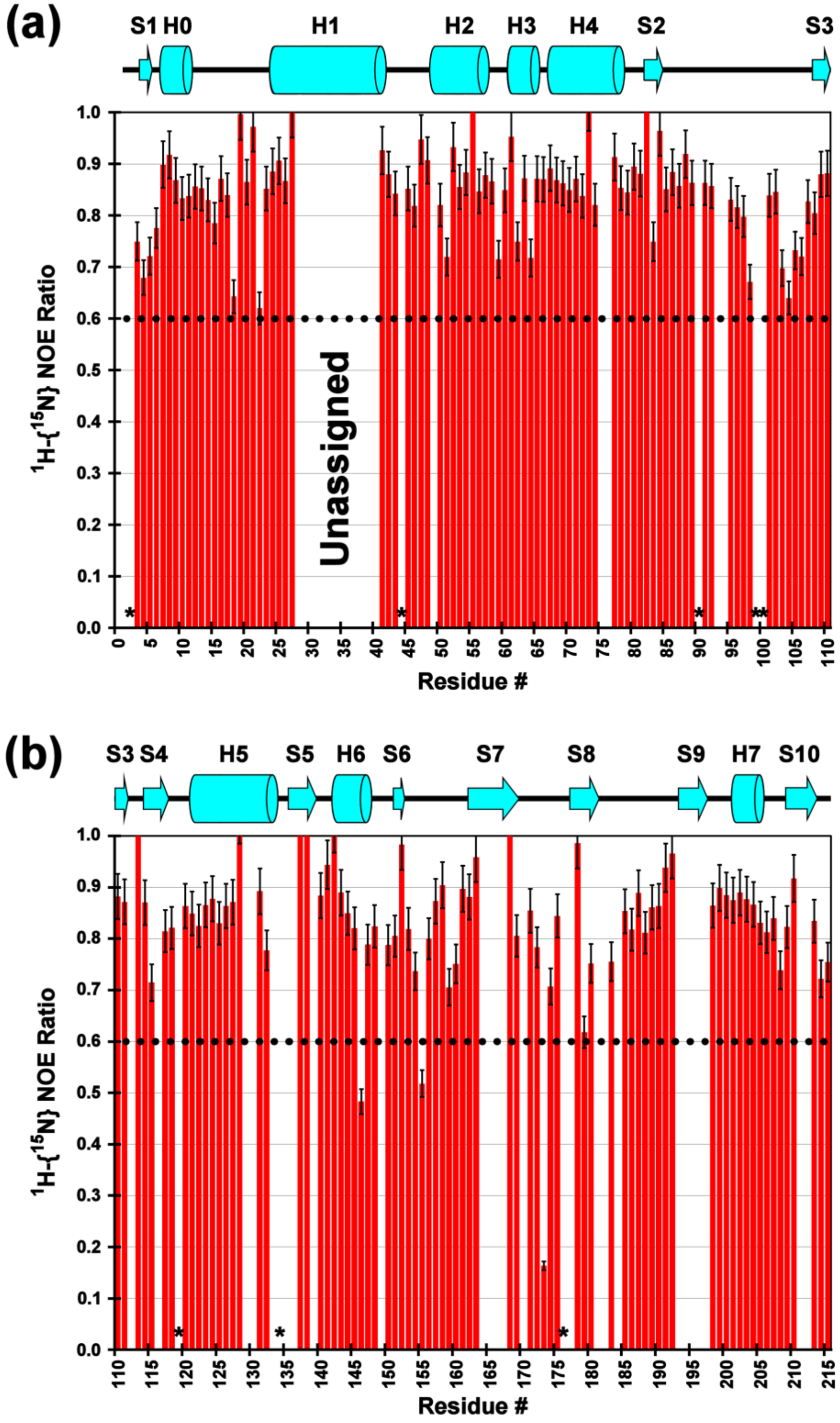
(a)
S1
H0
H1
H2
H3
H4
S2
S3
^{1}H-{^{15}N} NOE Ratio
Unassigned
Residue #
(b)
S3
S4
H5
S5
H6
S6
S7
S8
S9
H7
S10
^{1}H-{^{15}N} NOE Ratio
Residue #

Figure S4: Backbone dynamics data of K777-inhibited cruzain. Heteronuclear ^{1}H-{^{15}N} NOE ratios of the cruzain backbone $^{15}N/^{1}H$ amide resonances greater than 0.6 (dotted line) indicates a structurally stable protease-K777 complex. Residue numbers **(a)** 1-110 and **(b)** 110-215 and the secondary structure motifs correspond to those of the cruzain-K777 crystal structure, 2OZ2.[7] Asterisks denote the positions of proline residues. Other blank regions indicate unassigned residues of the uniformly $^{13}C/^{15}N/^{2}H$ labeled cruzain-K777 sample.

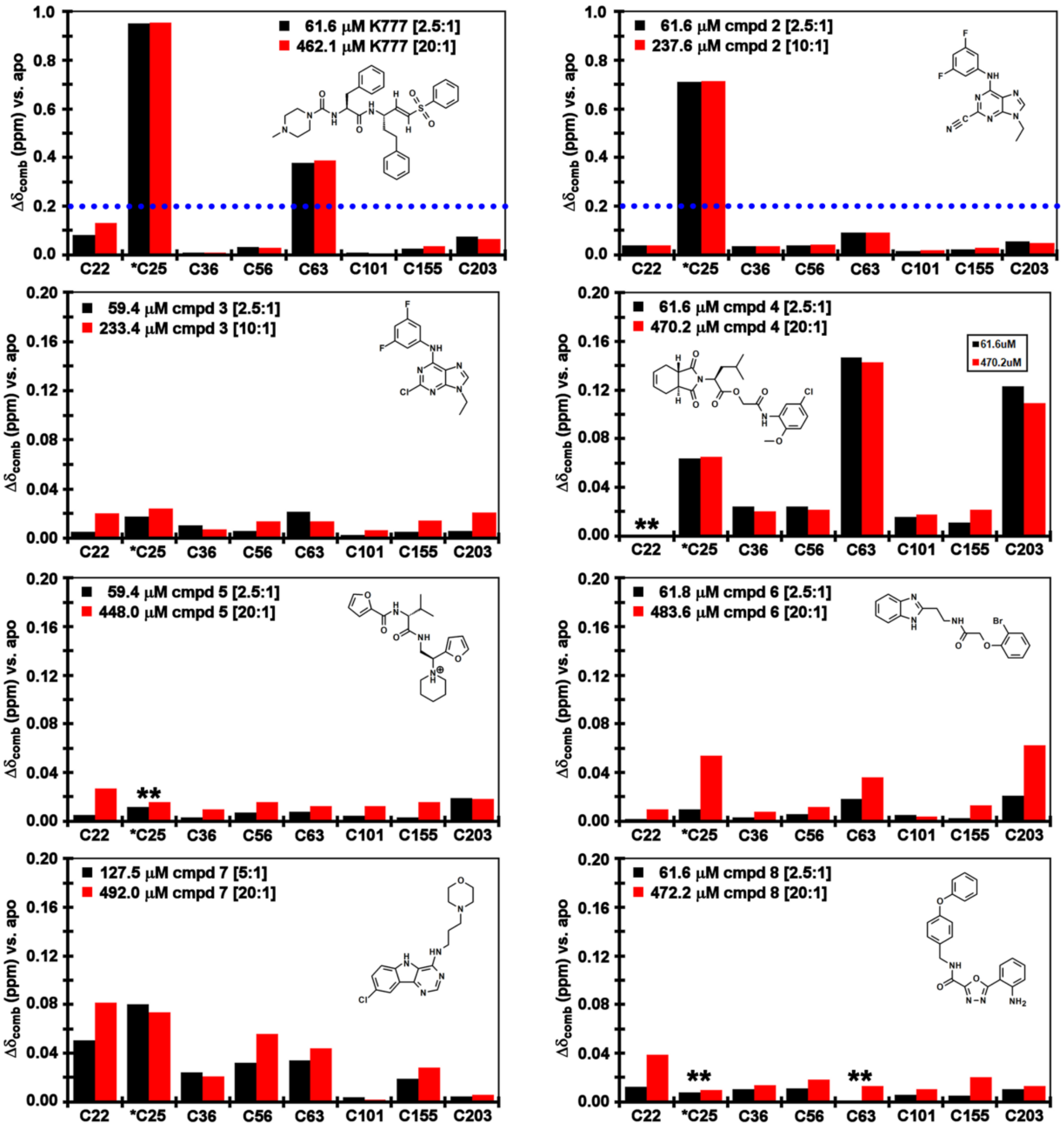

Figure S5: Summary of the cruzain-inhibitor ^{15}N-Cys shift perturbation data. Chemical shift perturbations of the backbone amide ^{15}N-Cys resonances upon addition of inhibitors listed in **Table 1** (as indicated) at 2.5 (black) and 10 or 20-fold (red) molar equivalents of cruzain. The chemical structures of K777 and compounds **2** – **8** are indicated as insets. Single asterisks indicate the catalytic Cys25. Double asterisks signify peaks that display extensive peak broadening upon addition of the inhibitors. Note that the scale of the y-axis ranges from 0 - 1.0 ppm in the K777 and compound **2** bar charts and 0 - 0.2 ppm for compounds **3** – **8**, indicating larger shift perturbations for the covalently-bound inhibitors. The blue dotted line in the bar charts of K777 and compound **2** indicate the upper limit of the y-axis for the bar charts of compounds **3 – 8**.

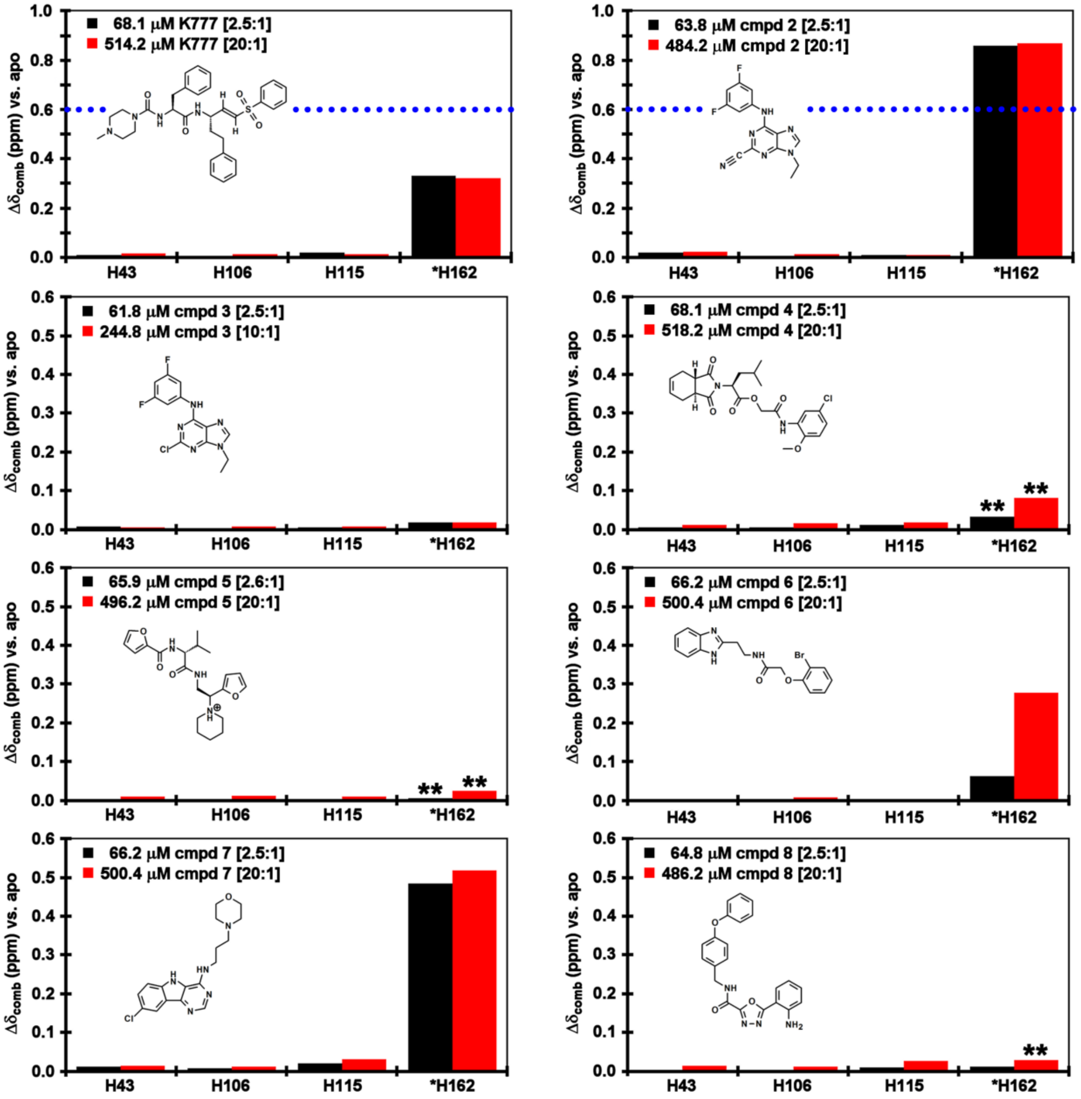

Figure S6: Summary of the cruzain-inhibitor ^{15}N-His shift perturbation data. Chemical shift perturbations of the backbone amide ^{15}N-His resonances upon addition of inhibitors listed in **Table 1** (as indicated) at 2.5 (black) and 10 or 20-fold (red) molar equivalents of cruzain. The chemical structures of K777 and compounds **2**-**8** are indicated as insets. Single asterisks indicate the catalytic His162. Double asterisks signify peaks that display extensive peak broadening upon addition of the inhibitors. Note that the scale of the y-axis ranges from 0 - 1.0 ppm in the K777 and compound **2** bar charts and 0 - 0.6 ppm for compounds **3** – **8**, indicating larger shift perturbations for the covalently-bound inhibitors. The blue dotted line in the bar charts of K777 and compound **2** indicate the upper limit of the y-axis for the bar charts of compounds **3 – 8**.

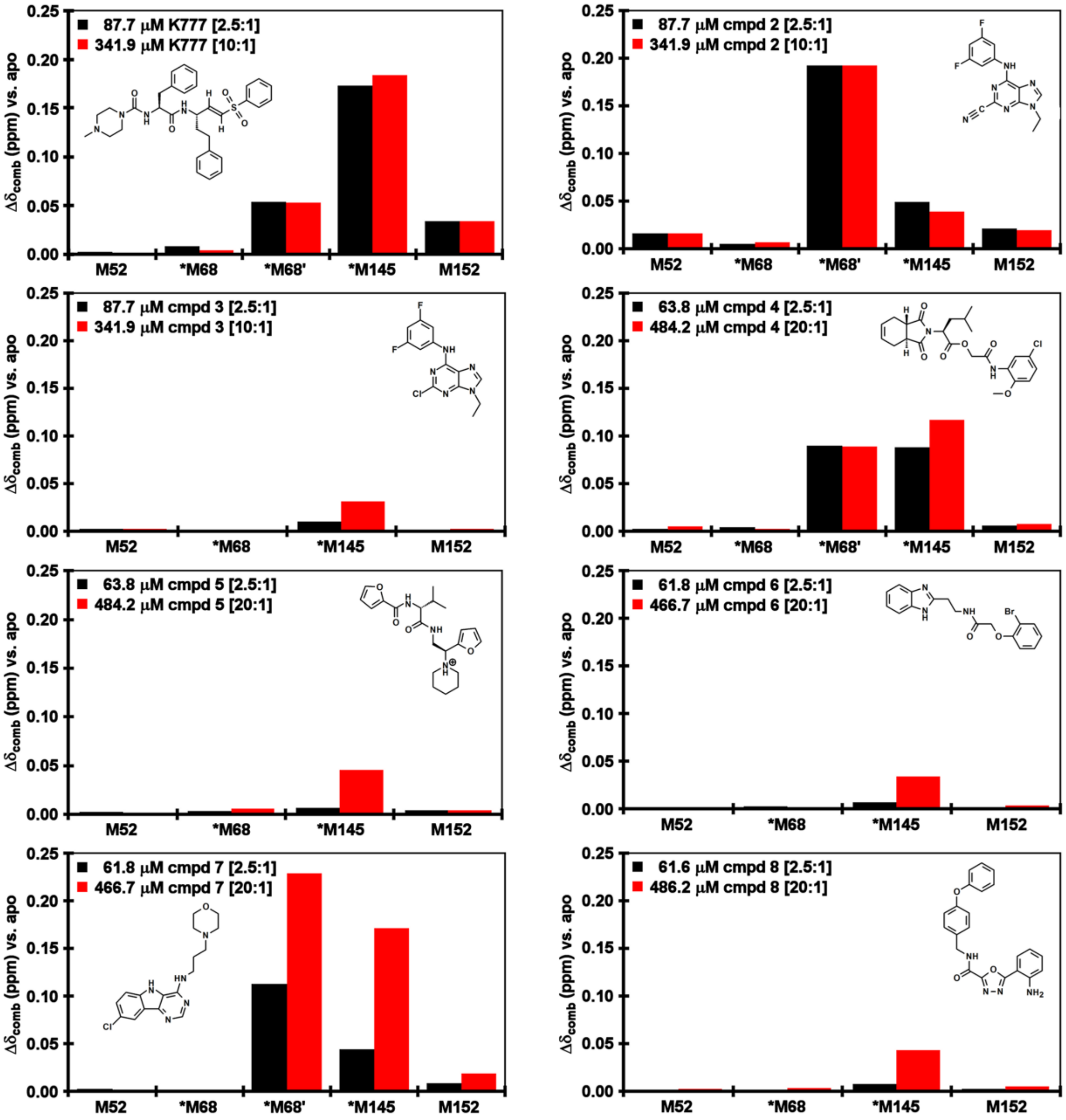

Figure S7: Summary of the cruzain-inhibitor ^{13}C-Met shift perturbation data. Chemical shift perturbations of the backbone amide ^{13}C-Met resonances upon addition of inhibitors listed in **Table 1** (as indicated) at 2.5 (black) and 10 or 20-fold (red) molar equivalents of cruzain. The chemical structures of K777 and compounds **2**-**8** are indicated as insets. Single asterisks indicate Met68, Met68', and Met145, which are positioned in the substrate binding pocket. Double asterisks signify peaks that display extensive peak broadening upon addition of the inhibitors.

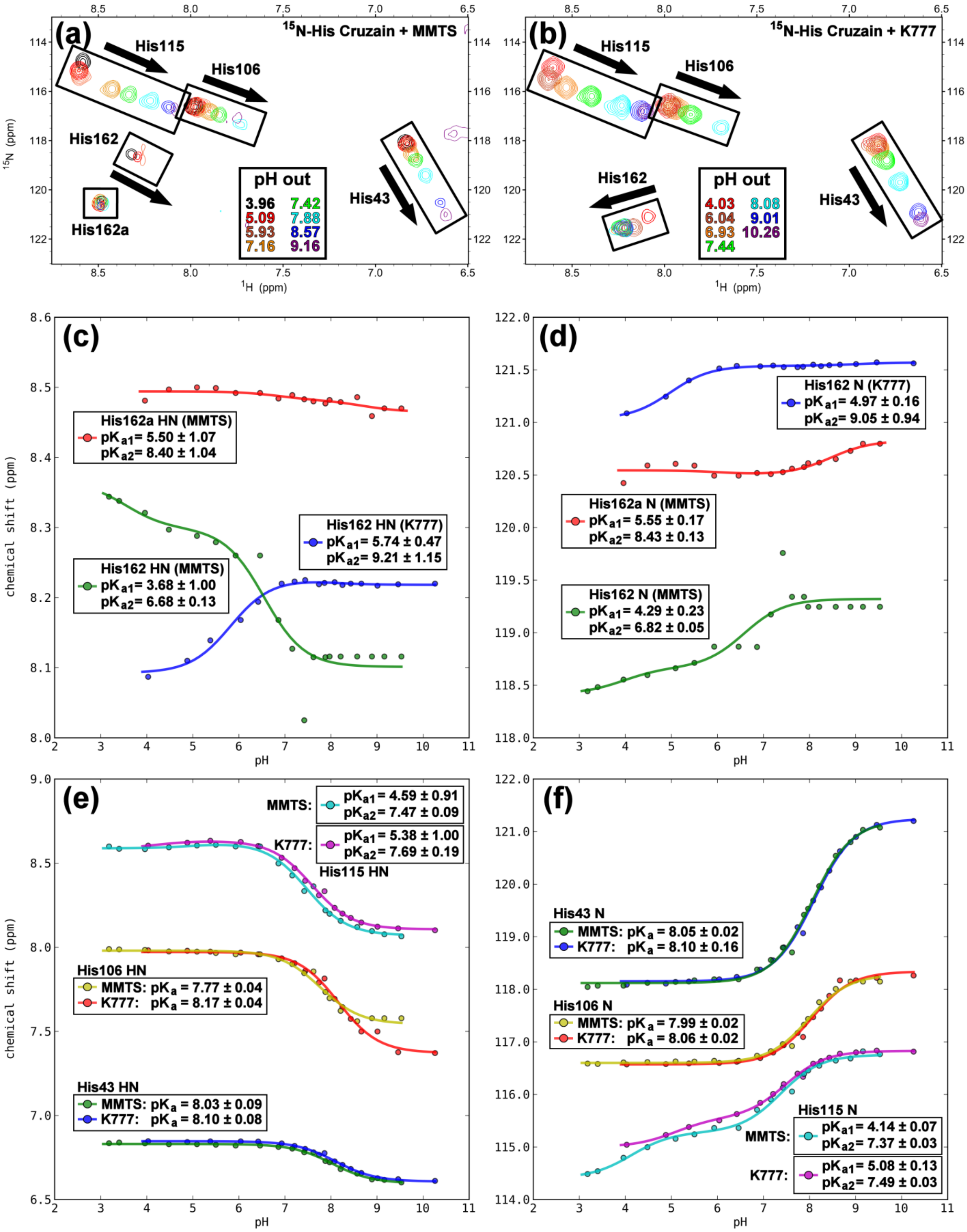
(a)
^{15}N-His Cruzain + MMTS
His115
His106
His162
His162a
His43
pH out
3.96 7.42
5.09 7.88
5.93 8.57
7.16 9.16
(b)
^{15}N-His Cruzain + K777
His115
His106
His162
His43
pH out
4.03 8.08
6.04 9.01
6.93 10.26
7.44
^{1}H (ppm)
^{15}N (ppm)
(c)
His162a HN (MMTS)
pK_{a1}= 5.50 ± 1.07
pK_{a2}= 8.40 ± 1.04
His162 HN (K777)
pK_{a1}= 5.74 ± 0.47
pK_{a2}= 9.21 ± 1.15
His162 HN (MMTS)
pK_{a1}= 3.68 ± 1.00
pK_{a2}= 6.68 ± 0.13
chemical shift (ppm)
pH
(d)
His162 N (K777)
pK_{a1}= 4.97 ± 0.16
pK_{a2}= 9.05 ± 0.94
His162a N (MMTS)
pK_{a1}= 5.55 ± 0.17
pK_{a2}= 8.43 ± 0.13
His162 N (MMTS)
pK_{a1}= 4.29 ± 0.23
pK_{a2}= 6.82 ± 0.05
pH
(e)
MMTS: pK_{a1}= 4.59 ± 0.91
pK_{a2}= 7.47 ± 0.09
K777: pK_{a1}= 5.38 ± 1.00
pK_{a2}= 7.69 ± 0.19
His115 HN
His106 HN
MMTS: pK_a = 7.77 ± 0.04
K777: pK_a = 8.17 ± 0.04
His43 HN
MMTS: pK_a = 8.03 ± 0.09
K777: pK_a = 8.10 ± 0.08
chemical shift (ppm)
pH
(f)
His43 N
MMTS: pK_a = 8.05 ± 0.02
K777: pK_a = 8.10 ± 0.16
His106 N
MMTS: pK_a = 7.99 ± 0.02
K777: pK_a = 8.06 ± 0.02
His115 N
MMTS: pK_{a1}= 4.14 ± 0.07
pK_{a2}= 7.37 ± 0.03
K777: pK_{a1}= 5.08 ± 0.13
pK_{a2}= 7.49 ± 0.03
pH

Figure S8 Comparison of the pH titration curves of MMTS- and K777-inhibited ^{15}N-His cruzain. Overlays of the ^{15}N-^{1}H HSQC spectra of (**a**) MMTS- and (**b**) K777-inhibited ^{15}N-His labeled cruzain. **(c)** Amide proton and **(d)** amide nitrogen pH titration curves corresponding to His162 of cruzain inhibited with MMTS (green) or K777 (blue). A minor conformer, His162a (red), is observed in the MMTS-inhibited cruzain spectra. **(e)** Amide proton and **(f)** amide nitrogen pH titration curves of the non-catalytic histidine residues (colored as indicated), display no significant differences between the MMTS- and K777-inhibited states. Individual amide proton and nitrogen pK_a values are reported in the boxes. Overall apparent pK_a values, representing the average values determined from the amide proton and nitrogen curve fittings are listed in **Supplemental Table S2**.

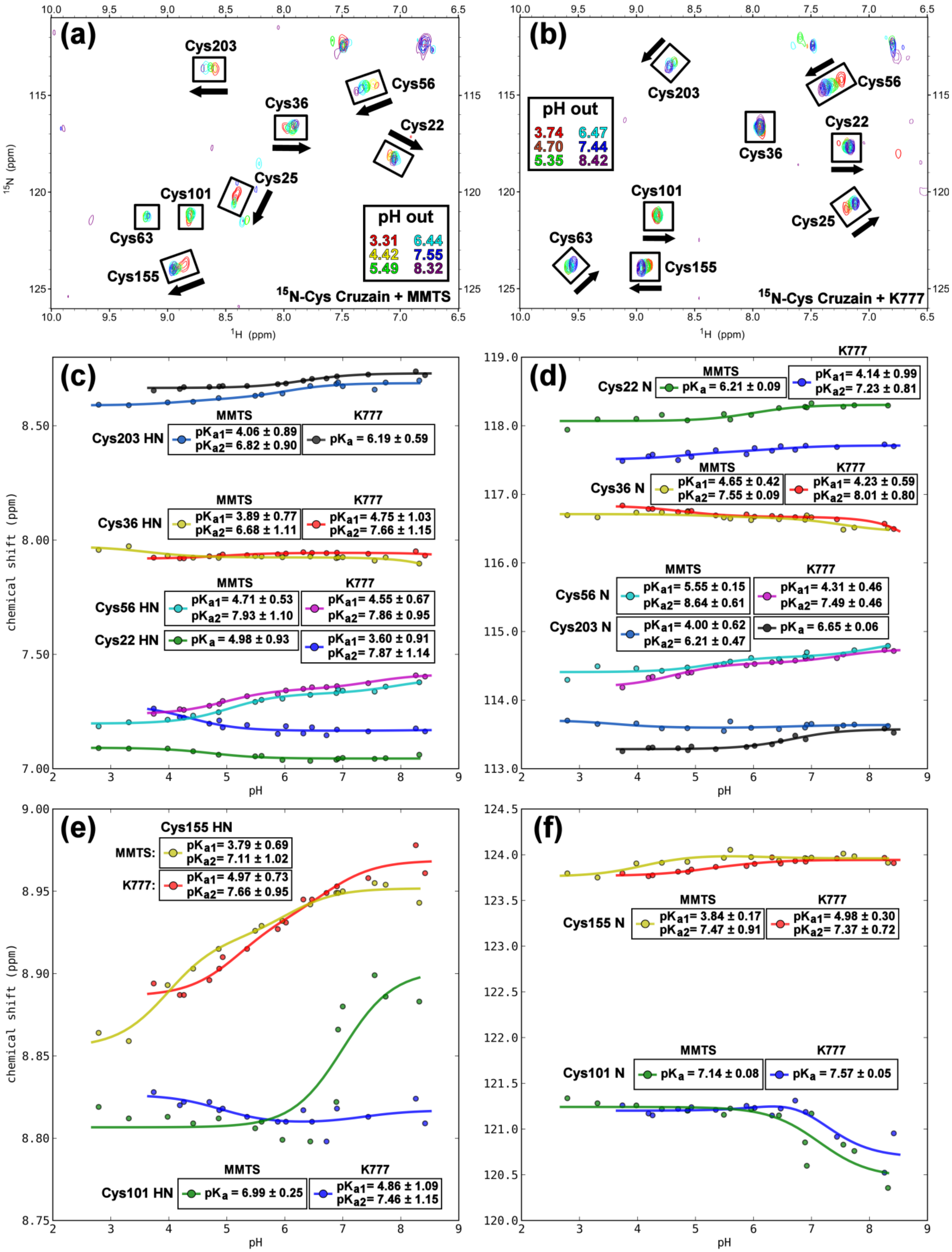
(a)
Cys203
Cys56
Cys36
Cys22
Cys25
Cys101
Cys63
Cys155
pH out
3.31 6.44
4.42 7.55
5.49 8.32
^{15}N-Cys Cruzain + MMTS
(b)
Cys203
Cys56
pH out
3.74 6.47
4.70 7.44
5.35 8.42
Cys36
Cys22
Cys101
Cys25
Cys63
Cys155
^{15}N-Cys Cruzain + K777
^{1}H (ppm)
^{15}N (ppm)
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5
115 120 125
(c)
Cys203 HN
MMTS
pK_{a1}= 4.06 ± 0.89
pK_{a2}= 6.82 ± 0.90
K777
pK_a = 6.19 ± 0.59
Cys36 HN
MMTS
pK_{a1}= 3.89 ± 0.77
pK_{a2}= 6.68 ± 1.11
K777
pK_{a1}= 4.75 ± 1.03
pK_{a2}= 7.66 ± 1.15
Cys56 HN
MMTS
pK_{a1}= 4.71 ± 0.53
pK_{a2}= 7.93 ± 1.10
K777
pK_{a1}= 4.55 ± 0.67
pK_{a2}= 7.86 ± 0.95
Cys22 HN
pK_a = 4.98 ± 0.93
pK_{a1}= 3.60 ± 0.91
pK_{a2}= 7.87 ± 1.14
chemical shift (ppm)
7.00 7.50 8.00 8.50
pH
2 3 4 5 6 7 8 9
(d)
Cys22 N
MMTS
pK_a = 6.21 ± 0.09
K777
pK_{a1}= 4.14 ± 0.99
pK_{a2}= 7.23 ± 0.81
Cys36 N
MMTS
pK_{a1}= 4.65 ± 0.42
pK_{a2}= 7.55 ± 0.09
K777
pK_{a1}= 4.23 ± 0.59
pK_{a2}= 8.01 ± 0.80
Cys56 N
MMTS
pK_{a1}= 5.55 ± 0.15
pK_{a2}= 8.64 ± 0.61
K777
pK_{a1}= 4.31 ± 0.46
pK_{a2}= 7.49 ± 0.46
Cys203 N
pK_{a1}= 4.00 ± 0.62
pK_{a2}= 6.21 ± 0.47
pK_a = 6.65 ± 0.06
113.0 114.0 115.0 116.0 117.0 118.0 119.0
pH
2 3 4 5 6 7 8 9
(e)
Cys155 HN
MMTS:
pK_{a1}= 3.79 ± 0.69
pK_{a2}= 7.11 ± 1.02
K777:
pK_{a1}= 4.97 ± 0.73
pK_{a2}= 7.66 ± 0.95
Cys101 HN
MMTS
pK_a = 6.99 ± 0.25
K777
pK_{a1}= 4.86 ± 1.09
pK_{a2}= 7.46 ± 1.15
chemical shift (ppm)
8.75 8.80 8.85 8.90 8.95 9.00
pH
2 3 4 5 6 7 8 9
(f)
Cys155 N
MMTS
pK_{a1}= 3.84 ± 0.17
pK_{a2}= 7.47 ± 0.91
K777
pK_{a1}= 4.98 ± 0.30
pK_{a2}= 7.37 ± 0.72
Cys101 N
MMTS
pK_a = 7.14 ± 0.08
K777
pK_a = 7.57 ± 0.05
120.0 120.5 121.0 121.5 122.0 122.5 123.0 123.5 124.0 124.5
pH
2 3 4 5 6 7 8 9

Figure S9: Comparison of NMR-based pH titration curves of MMTS- and K777-inhibited ^{15}N-Cys cruzain (part 1). Overlays of the ^{15}N-^{1}H HSQC spectra of (**a**) MMTS- and (**b**) K777-inhibited ^{15}N-Cys labeled cruzain. (**c**) Amide proton and (**d**) amide nitrogen pH titration curves of Cys22, Cys36, Cys56 and Cys203 (colored as indicated) from MMTS- and K777-inhibited cruzain. (**e**) Amide proton and (**f**) amide nitrogen pH titration curves of Cys101 and Cys155 (colored as indicated) from MMTS- and K777-inhibited cruzain. Individual “pK_a” values of the amide proton and amide nitrogen are indicated in the boxes. Average “pK_a” values with propagated errors are listed in **Supplementary Table S2**. pH titration curves for the Cys25 and Cys63 backbone amides are displayed in **Supplementary Figure S10**.

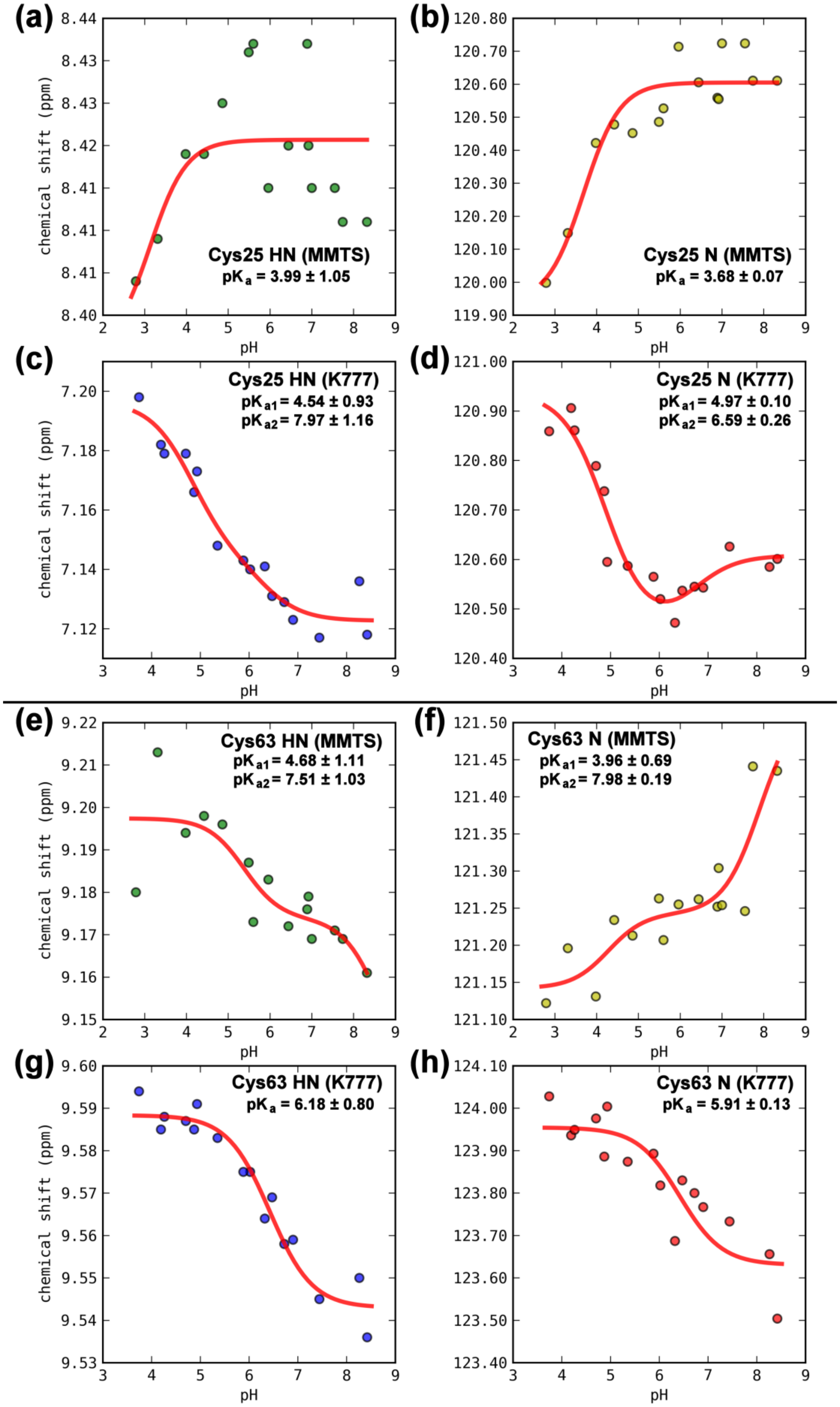
(a)
Cys25 HN (MMTS)
$pK_a = 3.99 \pm 1.05$
chemical shift (ppm)
8.44
8.43
8.43
8.42
8.41
8.41
8.41
8.40
pH
2 3 4 5 6 7 8 9
(b)
Cys25 N (MMTS)
$pK_a = 3.68 \pm 0.07$
120.80
120.70
120.60
120.50
120.40
120.30
120.20
120.10
120.00
119.90
pH
2 3 4 5 6 7 8 9
(c)
Cys25 HN (K777)
$pK_{a1} = 4.54 \pm 0.93$
$pK_{a2} = 7.97 \pm 1.16$
chemical shift (ppm)
7.20
7.18
7.16
7.14
7.12
pH
3 4 5 6 7 8 9
(d)
Cys25 N (K777)
$pK_{a1} = 4.97 \pm 0.10$
$pK_{a2} = 6.59 \pm 0.26$
121.00
120.90
120.80
120.70
120.60
120.50
120.40
pH
3 4 5 6 7 8 9
(e)
Cys63 HN (MMTS)
$pK_{a1} = 4.68 \pm 1.11$
$pK_{a2} = 7.51 \pm 1.03$
chemical shift (ppm)
9.22
9.21
9.20
9.19
9.18
9.17
9.16
9.15
pH
2 3 4 5 6 7 8 9
(f)
Cys63 N (MMTS)
$pK_{a1} = 3.96 \pm 0.69$
$pK_{a2} = 7.98 \pm 0.19$
121.50
121.45
121.40
121.35
121.30
121.25
121.20
121.15
121.10
pH
2 3 4 5 6 7 8 9
(g)
Cys63 HN (K777)
$pK_a = 6.18 \pm 0.80$
chemical shift (ppm)
9.60
9.59
9.58
9.57
9.56
9.55
9.54
9.53
pH
3 4 5 6 7 8 9
(h)
Cys63 N (K777)
$pK_a = 5.91 \pm 0.13$
124.10
124.00
123.90
123.80
123.70
123.60
123.50
123.40
pH
3 4 5 6 7 8 9

Figure S10: Comparison of NMR-based pH titration curves of MMTS- and K777-inhibited ^{15}N-Cys cruzain (part 2). pH titration curves of Cys25 amide proton and nitrogen from the (**a-b**) cruzain-MMTS complex and (**c-d**) cruzain-K777 complex. pH titration curves of Cys63 amide proton and nitrogen from the (**e-f**) cruzain-MMTS complex and (**g-h**) cruzain-K777 complex. Estimated pK_a values were calculated with either a 1 pK_a or 2 pK_a fitting model using the Ekin module of PEAT_DB.[4] Individual "pK_a" values of the amide proton and amide nitrogen are indicated. Average "pK_a" values with propagated errors are listed in **Supplementary Table S2**.

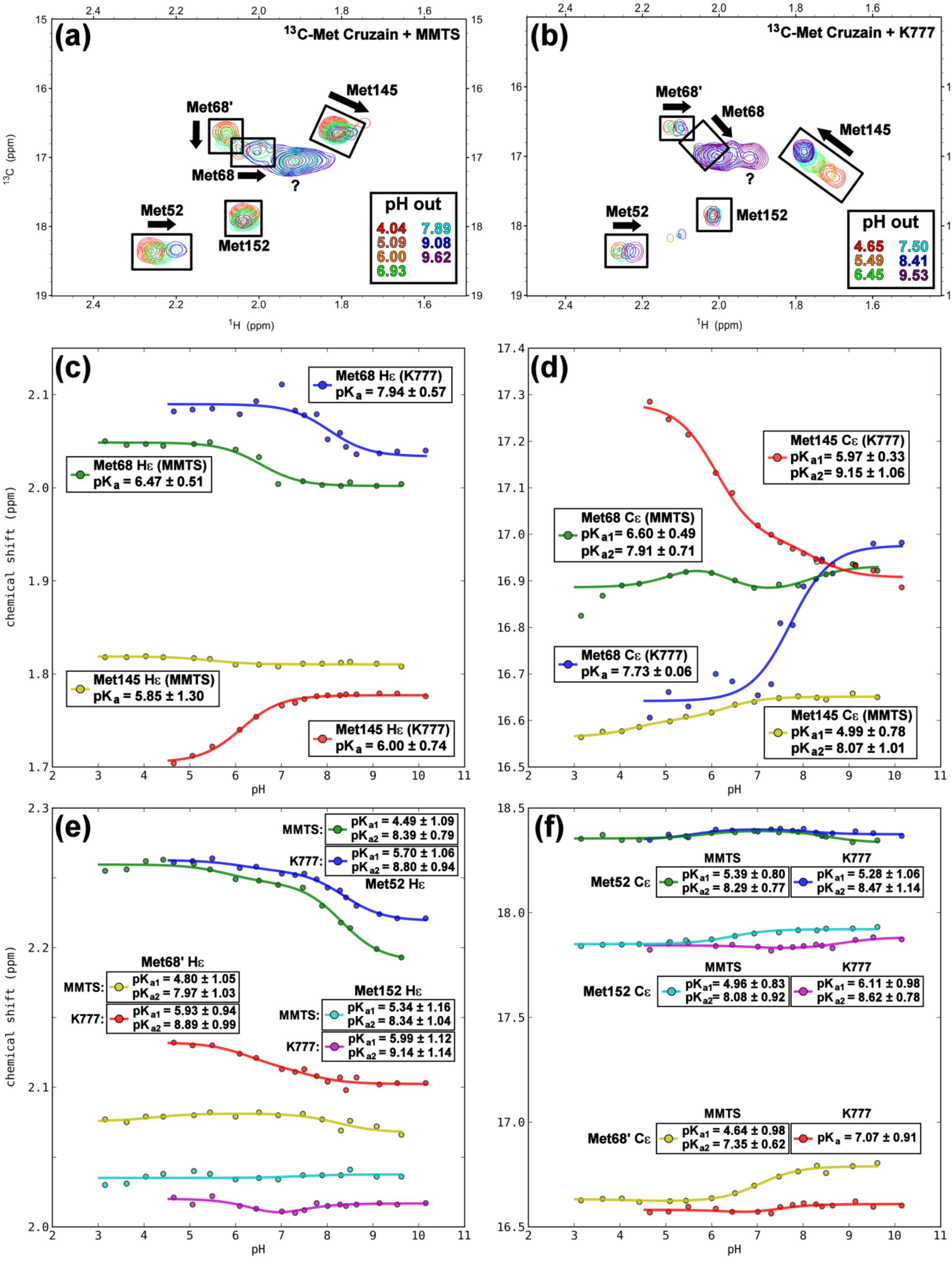
(a)
^{13}C-Met Cruzain + MMTS
Met145
Met68'
Met68
?
Met52
Met152
pH out
4.04
5.09
6.00
6.93
7.89
9.08
9.62
^{13}C (ppm)
^{1}H (ppm)
(b)
^{13}C-Met Cruzain + K777
Met68'
Met68
Met145
?
Met52
Met152
pH out
4.65
5.49
6.45
7.50
8.41
9.53
^{1}H (ppm)
(c)
Met68 Hε (K777)
pK$_a$ = 7.94 ± 0.57
Met68 Hε (MMTS)
pK$_a$ = 6.47 ± 0.51
Met145 Hε (MMTS)
pK$_a$ = 5.85 ± 1.30
Met145 Hε (K777)
pK$_a$ = 6.00 ± 0.74
chemical shift (ppm)
pH
(d)
Met145 Cε (K777)
pK$_{a1}$ = 5.97 ± 0.33
pK$_{a2}$ = 9.15 ± 1.06
Met68 Cε (MMTS)
pK$_{a1}$ = 6.60 ± 0.49
pK$_{a2}$ = 7.91 ± 0.71
Met68 Cε (K777)
pK$_a$ = 7.73 ± 0.06
Met145 Cε (MMTS)
pK$_{a1}$ = 4.99 ± 0.78
pK$_{a2}$ = 8.07 ± 1.01
pH
(e)
MMTS:
pK$_{a1}$ = 4.49 ± 1.09
pK$_{a2}$ = 8.39 ± 0.79
K777:
pK$_{a1}$ = 5.70 ± 1.06
pK$_{a2}$ = 8.80 ± 0.94
Met52 Hε
Met68' Hε
MMTS:
pK$_{a1}$ = 4.80 ± 1.05
pK$_{a2}$ = 7.97 ± 1.03
K777:
pK$_{a1}$ = 5.93 ± 0.94
pK$_{a2}$ = 8.89 ± 0.99
Met152 Hε
MMTS:
pK$_{a1}$ = 5.34 ± 1.16
pK$_{a2}$ = 8.34 ± 1.04
K777:
pK$_{a1}$ = 5.99 ± 1.12
pK$_{a2}$ = 9.14 ± 1.14
chemical shift (ppm)
pH
(f)
MMTS
K777
Met52 Cε
pK$_{a1}$ = 5.39 ± 0.80
pK$_{a2}$ = 8.29 ± 0.77
pK$_{a1}$ = 5.28 ± 1.06
pK$_{a2}$ = 8.47 ± 1.14
MMTS
K777
Met152 Cε
pK$_{a1}$ = 4.96 ± 0.83
pK$_{a2}$ = 8.08 ± 0.92
pK$_{a1}$ = 6.11 ± 0.98
pK$_{a2}$ = 8.62 ± 0.78
MMTS
K777
Met68' Cε
pK$_{a1}$ = 4.64 ± 0.98
pK$_{a2}$ = 7.35 ± 0.62
pK$_a$ = 7.07 ± 0.91
pH

Figure S11: Comparison of NMR-based pH titration curves of MMTS- and K777-inhibited ^{13}C-Met cruzain. Overlays of the ^{13}C-^{1}H HSQC spectra of (**a**) MMTS- and (**b**) K777-inhibited ^{13}C-Met labeled cruzain. Titration points are colored as indicated, and range from pH ~ 3 to ~ 10. The pH titration curves of the methionine ε-methyl (**c**) proton and (**d**) carbon of Met68 and Met145 in the MMTS- and K777-inhibited states (colored as indicated). Both Met68 and Met145 are surface exposed in the substrate binding pocket. For clarity, a minor conformer, Met68', is plotted in (**e**) and (**f**), below. The pH titration curves of the "non-catalytic" methionine residue ε-methyl (**e**) proton and (**f**) carbon in the MMTS- and K777-inhibited states (colored as indicated). Individual "pK_a" values of the methyl proton and methyl carbon are indicated in the boxes. Average "pK_a" values with propagated errors are listed in **Supplementary Table S2**.

Table S1: U-^{13}C/^{15}N/^{2}H-cruzain + K777 resonance assignments
[a](Major conformer, 800 MHz, 27 ^{o}C)

Residue #			Chemical Shift (ppm)					
	[b]SN	[c]CN	N	H^N	C^{α}	C^{β}	C’	[d]Others
Ala	1	1						
Pro	2	2		---	61.84	31.11	177.2	Cγ: 26.69
Ala	3	3	126.5	8.91	54.06	17.73	175.2	Cγ1/γ2: 21.27, 19.80
Ala	4	4	114.2	7.19	51.17	19.99	175.4	
Val	5	5	119.5	7.50	60.79	35.25	173.2	
Asp	6	6	123.0	8.21	52.84	40.09	177.3	
Trp	7	7	123.6	8.68	60.50	28.23	178.4	Nε1: 126.0; Hε1: 9.82
Arg	8	8	118.8	8.62	58.11	28.65	180.5	Cγ: 24.30; Cδ: 42.26; Cζ: 159.2; Nε: 81.66; Hε: 8.21
Ala	9	9	120.2	7.20	53.25	17.16	178.1	
Arg	10	10	114.2	6.88	54.01	29.98	176.2	Cγ: 25.40; Cδ: 41.73; Cζ: 159.3 Nε: 82.92; Hε: 7.23
Gly	11	11	106.6	7.60	45.35	---	172.9	
Ala	12	12	113.5	6.68	51.25	21.16	173.9	
Val	13	13	114.5	7.64	60.87	33.57	176.3	Cγ1/γ2: 21.10, 20.22
Thr	14	14	117.5	8.14	61.51	70.48	175.5	Cγ2: 20.66
Ala	15	15	121.9	8.43	52.22	18.30	177.8	
Val	16	16	121.8	8.35	63.46	31.15	176.0	Cγ1/γ2: 20.65, 20.65
Lys	17	17	130.5	9.43	55.14	34.73	173.5	Cγ/Cδ: 25.27; Cε: 41.11
Asp	18	18	113.8	7.12	51.30	42.92	176.0	
Gln	**19**	**19**	**123.2**	**8.09**	**55.78**	**29.44**	**178.7**	**Cγ: 32.92**
Gly	20	20	111.4	8.68	45.84	---	174.8	
Gln	21	21	123.2	8.88	54.65	27.13	174.5	Cγ: 33.41; Cδ: 180.4 Nε2: 112.2; Hε21/ε22: 7.51, 6.73
Cys	22	22	117.2	7.07	54.68	45.55	174.9	
Gly	23	23	121.1	8.62	47.87	---	174.3	
Ser	24	24	111.9	8.69	55.87	62.12	176.0	
Cys	**25**	**25**	**120.4**	**7.14**	**59.26**	**35.42**	**172.4**	
Trp	26	26	120.8	7.11	56.46	27.12	174.6	
Ala	27	27	124.9	5.98	53.05	16.40		
Phe	28	28						
Ser	29	29						
Ala	30	30						
Ile	31	31						
Gly	32	32						
Asn	33	33						
Val	34	34						

Residue #			Chemical Shift (ppm)					
	[b]SN	[c]CN	N	H^N	C^{α}	C^{β}	C’	Others
Glu	35	35						
[e]Cys	36	36	116.9	7.91				
Gln	37	37						
Trp	38	38						
Phe	39	39						
Leu	40	40			56.28	39.82	177.3	
Ala	41	41	121.0	7.34	51.51	17.76	175.9	
Gly	42	42	106.2	7.71	44.70	---	172.9	
His	43	43	118.0	6.84	52.18	27.64		
Pro	44	44		---	62.55	31.09	177.2	
Leu	45	45	124.9	9.08	56.96	39.98	176.8	
Thr	46	46	127.4	9.45	61.50	71.67	172.5	Cγ2: 19.07
Asn	47	47	125.5	9.03	52.96	37.65	176.2	Cγ: 176.3; Nδ2: 112.80 Hδ21/δ22: 7.74, 6.95
Leu	48	48	131.3	10.98	53.95	42.17		
Ser	49	49			57.03	62.82	174.4	
Glu	50	50	126.0	8.10	61.03	27.65	180.3	Cγ: 34.67
Gln	51	51	121.2	9.44	56.90	28.27	175.8	Cγ: 33.48
Met	52	52	115.6	7.41	58.22	31.75	177.4	Cγ: 30.51; Cε: 18.47; Hε: 2.26
Leu	53	53	114.3	6.35	55.73	41.21	178.3	Cγ: 25.28
Val	54	54	118.7	7.84	66.90	30.63	177.8	Cγ1/γ2: 22.63, 20.68
Ser	55	55	109.6	8.69	62.37	64.05	175.3	
Cys	56	56	114.3	7.31	55.87	45.56	173.7	
Asp	57	57	116.1	7.15	51.11	37.31	176.4	
Lys	58	58	125.1	7.33	55.67	29.55	177.4	Cγ: 23.14; Cδ: 27.17
Thr	59	59	113.6	8.19	62.69	68.07	173.7	Cγ2: 21.81
Asp	60	60	121.8	6.86	53.25	42.19	173.3	
Ser	61	61	111.3	7.56	56.17	62.74	176.8	
Gly	62	62	111.0	8.61	48.10	---	176.2	
Cys	63	63	123.6	9.56	54.84	39.38	175.5	
Ser	64	64	114.7	8.79	58.05	62.14	172.7	
Gly	65	65	106.5	7.01	44.24	---	171.9	
Gly	66	66	105.0	7.96	44.48	---	170.5	
Leu	67	67	117.5	7.59	53.03	43.62	176.0	Cδ1/δ2: 25.27, 22.31
Met	68	68	125.5	8.57	63.40	26.95	176.1	Cγ: 33.55 Cε: 16.88; Hε: 2.13
Asn	69	69	114.4	9.03	55.84	36.57	177.8	Cγ: 175.2; Nδ2: 112.4 Hδ21/δ22: 7.32, 7.27
Asn	70	70	116.9	6.26	54.61	37.39	177.7	Cγ: 176.6; Nδ2: 109.6 Hδ21/δ22: 6.73, 6.69
Ala	71	71	125.4	7.71	54.51	17.08	179.8	

Residue #			**Chemical Shift (ppm)**					
	[b]SN	**[c]CN**	**N**	**H^N**	**C^α**	**C^β**	**C'**	**Others**
Phe	72	72	113.8	8.07	58.45	37.65	179.2	
Glu	73	73	118.6	7.58	58.69	27.89	177.3	Cγ: 34.65
Trp	74	74	122.2	8.92	62.74	25.32		Nε1: 131.5; Hε1: 10.47
Ile	75	75						
Val	76	76			66.67	31.83	176.8	Cγ: 21.88, 20.20
Gln	77	77	115.6	8.97	57.77	28.24	178.1	Cγ: 35.13, Cδ: 179.0 Nε2: 111.8; Hε21/ε22: 7.59, 6.85
Glu	78	78	114.4	8.20	55.27	28.48	176.6	Cγ: 35.13
Asn	79	78a	118.2	6.59	51.12	41.58	177.2	
Asn	80	78b	115.7	8.02	54.53	37.02	174.6	Cγ: 178.3; Nδ2: 113.0 Hδ21/δ22: 7.47, 6.78
Gly	81	78c	105.8	8.75	45.76	---		
Ala	82	79	120.9	7.48	52.56	19.56	176.3	
Val	83	80	117.2	8.27	60.40	32.23	176.3	
Tyr	84	81	126.0	6.82	57.22	38.15	174.7	
Thr	85	82	108.9	8.69	60.96	68.63	174.6	Cγ2: 21.65
Glu	86	83	125.0	8.69	58.55	28.75	178.5	Cγ: 34.86
Asp	87	84	113.3	8.52	56.21	39.95	177.4	
Ser	88	85	109.8	7.49	58.10	64.53	175.2	
Tyr	89	86	127.7	7.84	54.15	35.37		
Pro	90	87		---	63.05	32.35	176.4	Cγ: 26.03
Tyr	91	88	121.4	9.31	60.25	37.22	177.1	
Ala	92	89	131.3	8.76	51.04	20.11		
Ser	93	89a			57.91	63.67	175.6	
Gly	94	89b	109.5	8.69	47.09	---	175.7	
Glu	95	89c	117.9	8.31	55.76	27.85	176.8	Cγ: 35.80
Gly	96	90	106.1	8.27	45.01	---	172.7	
Ile	97	91	119.6	7.42	58.96	38.49	175.4	Cγ1: 25.86; Cδ1: 16.47
Ser	98	92	122.1	8.72	53.87	63.38		
Pro	99	93						
Pro	100	94		---	62.15	31.18	176.3	Cγ: 26.27; Cδ: 49.96
Cys	101	95	120.9	8.80	57.15	41.69	174.6	
Thr	102	96	121.0	7.90	58.66	70.38	175.5	Cγ2: 21.21
Thr	103	97	114.1	8.60	61.68	68.58	174.7	Cγ2: 21.23
Ser	104	98	116.0	7.52	56.90	64.03	173.7	
Gly	105	99	106.9	8.42	45.14	---	173.9	
His	106	100	116.5	7.99	51.83	27.75	172.6	
Thr	107	101	117.6	8.98	61.43	70.89	174.1	Cγ2: 20.73
Val	108	102	129.4	9.38	65.73	30.96	175.2	Cγ1/γ2: 21.85, 20.64
Gly	109	103	115.5	9.15	45.13	---	170.3	
Ala	110	105	119.2	7.32	49.94	21.79	173.8	

Residue #			Chemical Shift (ppm)					
	^{b}SN	^{c}CN	N	H^N	C$^\alpha$	C$^\beta$	C’	Others
Thr	111	106	111.8	5.90	59.41	71.01		
Ile	112	107			58.53	40.04	175.6	
Thr	113	108	107.9	8.79	60.39	69.31	175.0	Cγ2: 20.68
Gly	114	109	107.8	7.45	45.21	---	171.6	
His	115	110	115.1	8.64	55.10	30.61		
Val	116	111			58.55	34.05	173.9	Cγ1/γ2: 19.47, 18.25
Glu	117	112	124.2	8.40	54.83	28.24	175.7	Cγ: 35.13
Leu	118	113	125.5	8.32	52.77	39.43		
Pro	119	114		---	61.66	32.08	176.3	Cγ: 27.01; Cδ: 49.83
Gln	120	115	119.0	8.26	52.86	25.88	176.5	Cγ: 32.74; Cδ: 181.50 Nε2: 115.30 Hε21/ε22: 8.42, 6.85
Asp	121	116	121.3	7.92	53.39	43.19	176.2	
Glu	122	117	125.8	9.69	61.64	31.12	177.2	Cγ: 40.56
Ala	123	118	121.2	8.25	54.70	17.11	181.5	
Gln	124	119	118.7	8.25	58.61	28.25	180.0	Cγ: 33.90
Ile	125	120	122.1	8.63	65.80	37.58	177.4	
Ala	126	121	121.6	8.37	55.60	17.02	178.4	
Ala	127	122	117.7	7.86	54.58	17.35	180.4	
Trp	128	123	119.8	8.14	61.53			
Leu	129	124						
Ala	130	125			55.19	17.79	177.0	
Val	131	126	110.8	6.95	62.69	33.44	176.5	Cγ1/γ2: 20.07, 20.07
Asn	132	127	114.8	8.70	53.14	37.65		Cγ: 181.30; Nδ2: 109.20 Hδ21/δ22: 7.64, 7.15
Gly	133	128						
Pro	134	129						
Val	135	130						
Ala	136	131			56.35	17.63	175.5	
Val	137	132	111.7	7.71	58.61	34.09	173.5	
Ala	138	133	124.9	7.16	48.62	22.46		
Val	139	134			57.50	35.93	173.2	Cγ1/γ2: 20.77, 20.77
Asp	140	135	120.3	8.36	53.47	41.22	176.1	
Ala	141	136	128.5	9.32	49.36	20.02	178.7	
Ser	142	139	122.2	9.36	61.48	62.66	177.4	
Ser	143	140	115.9	9.24	59.14	64.24	175.9	
Trp	144	141	121.4	8.50	58.78	28.73	178.6	
Met	145	142	118.9	8.69	58.15	29.57	178.0	Cγ: 30.73, Cε: 17.27; Hε:1.65
Thr	146	143	105.3	7.66	60.45	68.57	174.2	Cγ2: 20.57
Tyr	147	144	123.1	7.50	58.60	38.08	175.7	
Thr	148	145	117.4	8.40	59.90	68.91		
Gly	149	146			44.07	---	172.9	

Residue #			Chemical Shift (ppm)					
	[b]SN	[c]CN	N	H^N	C^α	C^β	C’	Others
Gly	150	147	108.3	8.86	42.89	---	173.4	
Val	151	148	119.3	8.41	61.30	30.53	175.8	Cγ1/γ2: 20.76, 20.76
Met	152	149	129.8	8.85	56.92	33.46	175.8	Cγ: 31.51 Cε: 18.02; Hε: 2.02
Thr	153	151	116.7	8.43	61.21	69.16	174.5	Cγ2: 21.24
Ser	154	152	117.2	7.67	55.77	62.38	173.6	
Cys	155	153	123.7	8.93	54.66	43.02	175.9	
Val	156	154	131.9	7.81	65.15	30.61	175.3	Cγ1/γ2: 21.15, 18.89
Ser	157	155	126.1	9.03	55.59	63.28	172.5	
Glu	158	156	121.9	9.54	56.81	31.82	176.3	Cγ: 34.64
Gln	159	156a	121.3	9.21	54.44	30.61	173.6	Cγ: 32.69; Cδ: 180.7 Nε2: 112.0 Hε21/ε22: 7.58, 6.79
Leu	160	157	124.2	8.45	54.94	42.23	177.9	
Asp	161	158	119.7	8.91	53.14	40.28	175.0	
His	**162**	**159**	**121.4**	**8.13**	**54.60**	**34.10**	**172.8**	
Gly	163	160	110.8	6.83	43.97	---		
Val	164	161						
Leu	165	162						
Leu	166	163						
Val	167	164			60.33	32.27	178.1	
Gly	168	165	109.2	8.35	46.62	---	173.6	
Tyr	169	166	118.3	7.72	56.05	41.14		
Asn	170	167			52.30	38.95	174.0	Cγ: 176.30; Nδ2: 108.7 Hδ21/δ22: 6.89, 6.83
Asp	171	167a	123.6	9.40	54.69	40.58	177.2	
Ser	172	167b	116.4	8.18	57.53	62.80	173.5	
Ala	173	167c	123.0	6.40	50.51	20.17	175.8	
Ala	174	167d	122.2	8.50	54.74	17.21	179.2	
Val	175	168	114.3	7.31	58.79	32.34		
Pro	176	169		---	56.12	29.43	175.2	Cδ: 41.77
Tyr	177	170	129.3	7.40	56.94	39.95	173.3	
Trp	178	171	117.2	8.62	58.07	28.77	176.3	
Ile	179	172	128.0	8.61	61.01	37.60	180.0	
Ile	180	173	115.8	8.13	63.97	42.91		
Lys	181	174						
Asn	**182**	**175**			**48.86**	**39.88**	**173.1**	
Ser	183	176	110.6	7.41	55.87	61.00		
Trp	184	177			52.23	28.16	177.8	
Thr	185	178	108.9	7.65	64.62	71.50	175.7	Cγ2: 20.71
Thr	186	179	107.9	8.48	62.79	68.53	174.8	Cγ2: 22.32

Residue #			Chemical Shift (ppm)					
	[b]SN	[c]CN	N	H^N	C^α	C^β	C’	Others
Gln	187	180	119.3	8.55	55.86	27.87	175.3	Cγ: 33.99; Cδ: 180.60 Nε2: 112.2 Hε21/ε22: 7.49, 6.80
Trp	188	181	119.9	6.96	56.80	30.05	176.0	Cε2: 140.1; Nε1: 130.8; Hε1: 9.85
Gly	189	182	113.7	7.76	46.56	---	173.8	
Glu	190	183	123.2	9.49	53.94	25.92	176.9	Cγ: 35.80
Glu	191	184	121.3	8.55	56.72	27.03	176.5	Cγ: 35.91
Gly	192	185	103.9	8.34	45.88	---		
Tyr	193	186						
Ile	194	187						
Arg	195	188						Cγ: 25.36; Cδ: 41.94 Cζ: 160.0; Nε: 85.53; Hε: 8.87
Ile	196	189						
Ala	197	190			52.13	18.21	175.3	
Lys	198	191	123.2	8.22	54.88	35.24	176.1	
Gly	199	192	115.3	9.19	45.21	---	174.1	
Ser	200	193	115.0	8.52	55.94	61.68	173.7	
Asn	201	198	125.0	9.81	52.27	36.87	176.7	
Gln	202	199	121.9	10.22	56.70	27.63	178.0	Cγ: 33.55
Cys	203	200	113.1	8.67	52.72	35.88	175.0	
Leu	204	201	111.2	8.18	55.77	37.02	176.2	Cδ1/δ2: 21.56, 21.56
Val	205	202	114.4	6.61	63.41	31.07	174.9	Cγ1/γ2: 21.12, 20.25
Lys	206	203	117.9	8.09	54.61	33.51	177.2	
Glu	207	204	118.3	7.91	58.66	29.35	177.6	Cγ: 35.86
Glu	208	205	126.2	10.38	55.34	28.81	172.6	Cγ: 33.95
Ala	209	206	128.4	8.97	49.29	20.62	177.9	
Ser	210	207	117.0	9.76	56.45	66.81		
Ser	211	208						
Ala	212	209			51.84	19.80	176.5	
Val	213	210	121.1	7.84	61.13	33.03	174.9	Cγ1/γ2: 20.00, 19.27
Val	214	211	125.0	8.94	62.18	32.28	175.5	Cγ1/γ2: 21.23, 20.23
Gly	215	212	117.6	8.07	45.80	---		

[a] Chemical shifts not corrected for ^{2}H-isotope effects; Catalytic residues in boldface.

[b] SN = sequential numbering (residues 1-215), used for cruzain-inhibitor crystal structures published after 2009, starting with 2OZ2.[7]

[c] CN = “classical” numbering (residues 1-212, includes insertions and deletions) corresponding to papain,[6] used for cruzain-inhibitor crystal structures published prior to 2009.

[d] Methionine ε-methyl ^{13}C/^{1}H resonances assigned from selective ^{13}C-Met labeled sample.

[e] Cys36 backbone amide resonances assigned from selective ^{15}N-Cys labeled sample.

Table S2: Selectively ^{15}N-His, ^{15}N-Cys, ^{13}C-Met labeled cruzain pK_a values (Major conformer, 27 ^{o}C)

Residue #			**[a]Fitted pK_a values**			
			[b]MMTS		**K777**	
	[c]SN	**[d]CN**	**pK_a 1**	**pK_a 2**	**pK_a 1**	**pK_a 2**
Cys amide	22	22	5.59 ± 0.47	----	3.87 ± 0.67	7.55 ± 0.70
	25	**25**	**3.83 ± 0.52**	**----**	**4.76 ± 0.47**	**7.28 ± 0.59**
	36	36	4.27 ± 0.44	7.12 ± 0.56	4.49 ± 0.60	7.83 ± 0.70
	56	56	5.13 ± 0.27	8.29 ± 0.63	4.43 ± 0.40	7.68 ± 0.53
	63	63	4.32 ± 0.66	7.75 ± 0.52	6.04 ± 0.41	----
	101	95	----	7.06 ± 0.13	4.85 ± 1.09	7.51 ± 0.57
	155	153	3.82 ± 0.36	7.29 ± 0.68	4.97 ± 0.40	7.52 ± 0.59
	203	200	4.03 ± 0.55	6.52 ± 0.51	----	6.42 ± 0.30
His amide	43	43	-----	8.04 ± 0.04	-----	8.10 ± 0.04
	106	100	-----	7.88 ± 0.02	-----	8.11 ± 0.02
	115	110	4.36 ± 0.46	7.42 ± 0.05	5.23 ± 0.50	7.59 ± 0.10
	162	**159**	**3.98 ± 0.51**	**6.75 ± 0.07**	**5.35 ± 0.25**	**9.13 ± 0.74**
	162a	**159a**	**5.52 ± 0.54**	**8.41 ± 0.53**	**----**	**----**
Met ε-CH_3	52	52	4.94 ± 0.67	8.34 ± 0.55	5.49 ± 0.75	8.63 ± 0.74
	68	**68**	**6.53 ± 0.35**	**(7.91 ± 0.71)**	**----**	**7.83 ± 0.28**
	68’	**68’**	**4.72 ± 0.72**	**7.66 ± 0.60**	**6.50 ± 0.66**	**(8.89 ± 0.99)**
	145	**142**	**5.42 ± 0.76**	**(8.07 ± 1.01)**	**5.98 ± 0.41**	**(9.21 ± 0.80)**
	152	149	5.15 ± 0.71	8.21 ± 0.69	6.05 ± 0.74	8.88 ± 0.69

Boldface indicates residues positioned in substrate binding pocket.

[a] Fitted pKa values calculated using PEAT/Ekin,[4] and represent averages and propagated errors of the ^{1}H and ^{15}N (Cys, His) or ^{13}C (Met) resonances. Estimated pH (σ = 0.1) and chemical shift (σ = 0.05) errors were used for the curve fitting. Curve fitting calculations were performed 3 times.

Parentheses indicate single atom curve fitting (either ^{1}H, or ^{13}C/^{15}N; no average).

[b] MMTS = methyl methylthiomethyl sulfoxide.

[c] SN = sequential numbering (residues 1-215), used for cruzain-inhibitor crystal structures published after 2009, starting with 2OZ2.[7]

[d] CN = “classical” numbering (residues 1-212, includes insertions and deletions) corresponding to Papain,[6] used for cruzain-inhibitor crystal structures published prior to 2009.

REFERENCES

1. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures, *Protein Expr Purif 41*, 207-234.
2. Weber, D. J., Gittis, A. G., Mullen, G. P., Abeygunawardana, C., Lattman, E. E., and Mildvan, A. S. (1992) NMR docking of a substrate into the X-ray structure of staphylococcal nuclease, *Proteins 13*, 275-287.
3. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J Biomol NMR 6*, 277-293.
4. Farrell, D., Miranda, E. S., Webb, H., Georgi, N., Crowley, P. B., McIntosh, L. P., and Nielsen, J. E. (2010) Titration_DB: storage and analysis of NMR-monitored protein pH titration curves, *Proteins 78*, 843-857.
5. Webb, H., Tynan-Connolly, B. M., Lee, G. M., Farrell, D., O'Meara, F., Sondergaard, C. R., Teilum, K., Hewage, C., McIntosh, L. P., and Nielsen, J. E. (2011) Remeasuring HEWL pK(a) values by NMR spectroscopy: methods, analysis, accuracy, and implications for theoretical pK(a) calculations, *Proteins 79*, 685-702.
6. Kamphuis, I. G., Kalk, K. H., Swarte, M. B., and Drenth, J. (1984) Structure of papain refined at 1.65 A resolution, *J Mol Biol 179*, 233-256.
7. 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., and Brinen, L. S. (2009) Vinyl sulfones as antiparasitic agents and a structural basis for drug design, *J Biol Chem 284*, 25697-25703.
8. Johnson, F. A., Lewis, S. D., and Shafer, J. A. (1981) Determination of a low pK for histidine-159 in the S-methylthio derivative of papain by proton nuclear magnetic resonance spectroscopy, *Biochemistry 20*, 44-48.
9. Pinitglang, S., Watts, A. B., Patel, M., Reid, J. D., Noble, M. A., Gul, S., Bokth, A., Naeem, A., Patel, H., Thomas, E. W., Sreedharan, S. K., Verma, C., and Brocklehurst, K. (1997) A Classical Enzyme Active Center Motif Lacks Catalytic Competence until Modulated Electrostatically, *Biochemistry 36*, 9968-9982.
10. Polgar, L. (1974) Mercaptide-imidazolium ion-pair: the reactive nucleophile in papain catalysis, *FEBS Lett 47*, 15-18.
11. Gillmor, S. A., Craik, C. S., and Fletterick, R. J. (1997) Structural determinants of specificity in the cysteine protease cruzain, *Protein Sci 6*, 1603-1611.