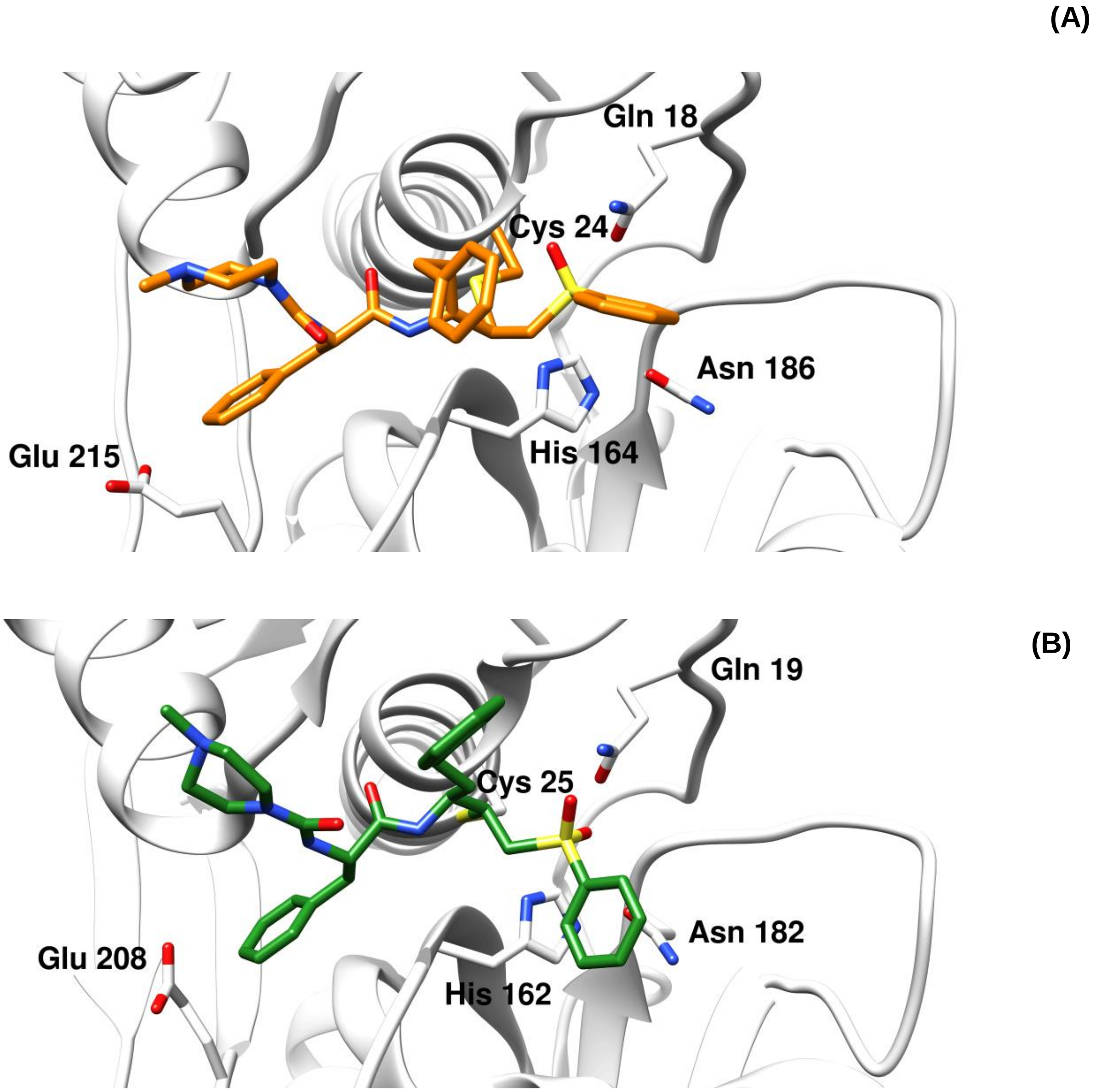

Supplementary Figure 1

Cryptopain 1 model ‘crypto1M’ with covalently docked K11777 structurally aligned to cruzain co-crystallized with K11777 (2OZ2). The images of the proteins are displayed separately here to show the individual residues more clearly. The active site residues of crypto1M **(A)** closely superimpose with those of **(B)** cruzain (PDB code 2OZ2) and the pose of the docked ligand is in good general agreement with the co-crystallized ligand. A glutamate at the bottom of the S2 pocket is present in both structures, and is seen in only a few parasite cathepsin L-like proteases, but is common in vertebrate cathepsin B-like proteases. Mammalian cathepsin L proteases typically have an Ala in this position.

```
Conservation:
cgd4                  1   ---M---FKSL-------------GKTILL-----AFFYIVFVKCDLPIHVTTADIIGRW   36
cgd2                  1   TFLMG-GFKSR-FHI---SLRRKWALFCWF-----FSILIREISSDLPPNCIHGDVVGTW   50
cgd7                  1   ---MQIGFTEKDLQIPNKSCGRKWNKFIVFINFISFLSIITKINGNLL------------   45
cgd6                  1   ---MDIGN----------------------------------------------------    5
cgd3                      ------------------------------------------------------------
2oz2_chainA_p001          ------------------------------------------------------------
1fh0_chainA_p002          ------------------------------------------------------------
Consensus_aa:             ............................................................
Consensus_ss:

Conservation:
cgd4                 37   EFQYSEPANN------WIDGCGSGVPNKNTQNLQPSLADYNKWLTTKTNGKLEKLDLILT   90
cgd2                 51   RIHVGTYKPCPSDQKFEDPTCGFPSPDRDFA-----------------------------   81
cgd7                 46   --HNGN------------------------------------------------------   49
cgd6                      ------------------------------------------------------------
cgd3                  1   -------------------------------------------CKKEFGIKVSKVN----   13
2oz2_chainA_p001          ------------------------------------------------------------
1fh0_chainA_p002          ------------------------------------------------------------
Consensus_aa:             ............................................................
Consensus_ss:

Conservation:
cgd4                 91   DLLFSENDYSDLDRSMFPGRSDWTFLAVKDP----------------------------E  122
cgd2                 82   --------HSMLTPTERGLMNNNFKLSFTFDVKFEGSELKVISTQNLDGVDTGKFNDFEK  133
cgd7                 50   --------HSMNG-NESGVS---------RTLKQEKDDHSILI---------GKKNETN-   81
cgd6                      ------------------------------------------------------------
cgd3                 14   ----------------KKKKRKLIIILNLI------------------------------   26
2oz2_chainA_p001          ------------------------------------------------------------
1fh0_chainA_p002          ------------------------------------------------------------
Consensus_aa:             ............................................................
Consensus_ss:

Conservation:
cgd4                123   NGKVVGRWTMVYDEGFEIVVNNM-----------------------------------TI  147
cgd2                134   VGT-TGNWTVILDQAFTFWTKSY-----------------------------------RY  157
cgd7                 82   --------TEIFSP------NNYSSQSSTKLVDHRTENFIEKDFQKNNSLVEEDNLTSKF  127
cgd6                  6   ---------------------NV-------------------------------------    7
cgd3                 27   ----------LYSP----------------------------------------------   30
2oz2_chainA_p001          ------------------------------------------------------------
1fh0_chainA_p002          ------------------------------------------------------------
Consensus_aa:             ............................................................
Consensus_ss:

Conservation:
cgd4                148   FGIMKYNLLNNGNCSAKDGDNETTKGETLCYETDSSQIQIGWY-----------------  190
cgd2                158   TAFFKYINEHED----------INASYCYCHTT-----LLGWWDSYPESVISNQDNKEEN  202
cgd7                128   FDPFVFN-FSID----------VHQCISKYEKL-----NKYLQKI---SLIFYLFNKTRE  168
cgd6                  8   ---------------------EEHQE--------------YISGPYI-ALI----NGTNQ   27
cgd3                      ------------------------------------------------------------
2oz2_chainA_p001          ------------------------------------------------------------
1fh0_chainA_p002          ------------------------------------------------------------
Consensus_aa:             ............................................................
Consensus_ss:

Conservation:
cgd4                191   ------------------------------YLRDS-KKRGCVSGRKDESYEN--------  211
cgd2                203   -----------KILSLEERSKYNNYPWLVMSEDLK-LRRGCWYGMRIFDGD---------  241
cgd7                169   LENEGLNNNSTSILKAEN--------ERVVNTTIK-INIATNSG-KMDNHLDIIIQPEVF  218
cgd6                 28   --------------------------QREPNKKLKNIIIATL------------------   43
cgd3                      ------------------------------------------------------------
2oz2_chainA_p001          ------------------------------------------------------------
1fh0_chainA_p002          ------------------------------------------------------------
Consensus_aa:             ............................................................
Consensus_ss:

Conservation:
cgd4                      ------------------------------------------------------------
cgd2                      ------------------------------------------------------------
cgd7                219   FNSIYRTNTTWMDRLLDTPNINSSPMNLLNLTNNQDNSQSIIDCICIFVSEVQSNLGLNR  278
cgd6                 44   -------------------------------------------IAIFIVLVVTV------   54
```

```
cgd3                     ------------------------------------------------------------
2oz2_chainA_p001         ------------------------------------------------------------
1fh0_chainA_p002         ------------------------------------------------------------
Consensus_aa:            ............................................................
Consensus_ss:

Conservation:
cgd4                212  ------SLDIIKNRVEIPIKKSFESSIRSLSNSLQEFAHRHNRKDS----------KNWR  255
cgd2                242  -------------------------GKMTDRTEWTLKLPRWRSSPLGLPPDHPVNAQNPT  276
cgd7                279  SSLEICNIKIQEIISAIKSNPEARFELTQLSLDRPPSKLIIPKNLG----------DKVT  328
cgd6                 55  -SLYITNNTSDKIDDFV-PGDYVDPATREYRKSFEEFKKKYHKVYS----------SMEE  102
cgd3                     ------------------------------------------------------------
2oz2_chainA_p001         ------------------------------------------------------------
1fh0_chainA_p002         ------------------------------------------------------------
Consensus_aa:            ............................................................
Consensus_ss:                                         hhhhhh  hhhh                hhh

Conservation:
cgd4                256  -ARSNFSLTD--VAFSHHSYYMN-AVNFRKIHP------------FRIMGSQHSSFIEMF  299
cgd2                277  -IAEYLNEVSAKYSEFESKNKLWSRSNKFEIGDIQNLDTLPKVRKALKISPHFVQRKKSN  335
cgd7                329  GYSWKESIAEGKEIQSENLLGRS-VQKLEPTHS------------SGNEILKELNTENSN  375
cgd6                103  -ENQRFEIYKQNMNFIKTTNSQG-FSYVLEMNE------------FGDLSKEEFMARFTG  148
cgd3                     ------------------------------------------------------------
2oz2_chainA_p001         ------------------------------------------------------------
1fh0_chainA_p002         ------------------------------------------------------------
Consensus_aa:            ............................................................
Consensus_ss:             hhhhhhhhh  hhhhhhh                               hhhhhhhh

Conservation:                                                   7 5
cgd4                300  K-------DGVGSVDNYDNIVDSIFACNVVRRKTDEVKLPSSFSIGDPFTD----EKYSD  348
cgd2                336  DEYITEISSVREEEQIS-GSKNNTDEPIWMRIRSFDWANPQHVYGRLGKKV-----RLIP  389
cgd7                376  A-------GNGVNETGE-ITN--------------KNIYFDSRDVNGG--------SCVY  405
cgd6                149  Y-------IKDSKDDER-VFKS---SRVSASESEEEFVPPNSINWVEA--------GCVN  189
cgd3                 31  ------------------------------------FRIPESFDLRSNEANRKNGGSCIT   54
2oz2_chainA_p001      1  --------------------------------------APAAVDWRAR--------GAVT   14
1fh0_chainA_p002      1  --------------------------------------LPKSVDWRKK--------GYVT   14
Consensus_aa:            .......................................Ppthsh............hls
Consensus_ss:                                                      ee

Conservation:               696 97 9 7655                          79 7    9       5
cgd4                349  FPFNQGNCGSCYAVSSVYILAKRAELKLNKLTNGASREEKILLSPQSVLSCSP----FNQ  404
cgd2                390  EVFNQGDCGDCFAVTAATIITSRLWIKYSNEPDILR---KIYASSIQMSECNV----YNQ  442
cgd7                406  IPFDQGKCGGCYAFVVSASISISNCIQKLVL--------PAPLSPQQIIDCSMS--FGNL  455
cgd6                190  PIRNQKNCGSCWAFSAVAALEGATCAQTNRG--------LPSLSEQQFVDCSKQ--NGNF  239
cgd3                 55  FPDNQGKCANCYNFSALSAIEGATCRQLGKR--------IPPLSQQHSLDCWLQSKGSVK  106
2oz2_chainA_p001     15  AVKDQGQCGSCWAFSAIGNVECQWFLAGHP---------LTNLSEQMLVSCDKT--DSGC   63
1fh0_chainA_p002     15  PVKNQKQCGSCWAFSATGALEGQMFRKTGK---------LVSLSEQNLVDCSRP--QGNQ   63
Consensus_aa:            .s.sQ.pCGsC@Ahoths.lp..hh.p............b..LS.QphlsCs......sb
Consensus_ss:                      hhhhhhhhhhhhhhhhhh                hhhhhh

Conservation:            57 77                    6   555
cgd4                405  GCEGGYPFLVGRQAEEI-----GISSEKCMGYYADSNQE----CNF-------SPFITPE  448
cgd2                443  GCGGGLITLAFKFAQEV-----GVRTQDCVDDYAKNLGL----KEIYPTPVYTPDEVGIA  493
cgd7                456  GCDGGFYSNGWSYLLEQNVPRNYICSWNEYPYIDSISSC----KAS-------ACNGCLT  504
cgd6                240  GCDGGTMGLAFQYAIKNK----YLCTNDDYPYFAEEKTC----MDS-----FCENYIEIP  286
cgd3                107  ACSGGQTFEVFNYAIKS-----KVCTRDSYPSTTHKTGKLGECKSN-CNECVGIKNFKWS  160
2oz2_chainA_p001     64  SGGLMNNAFEWIVQENNG----AVYTEDSYPYASGEGISP--PCTTS------GHTVGAT  111
1fh0_chainA_p002     64  GCNGGFMARAFQYVKENG----GLDSEESYPYVAVDEIC----KYR-------PENSVAQ  108
Consensus_aa:            tCsGG....s@phhbc.......l.opcshs.hs.p......................h.
Consensus_ss:                  hhhhhhhhhh                                         eee

Conservation:
cgd4                449  IEDRIYCE----------------------------------------------------  456
cgd2                494  DDGHSFLQTKKGKKFESKKSKIKEIKQQEGDNQEYDQEHDQEHDQEHEQEHEQEHEQEHE  553
cgd7                505  ISKYNVFTG---------------------------------------------------  513
cgd6                287  VKAYKYVF----------------------------------------------------  294
cgd3                161  YTGSSILY----------------------------------------------------  168
2oz2_chainA_p001    112  ITGHVELP----------------------------------------------------  119
1fh0_chainA_p002    109  DTGFTVVA----------------------------------------------------  116
Consensus_aa:            .ps..hh.....................................................
Consensus_ss:            eeeeeee
```

```
Conservation:
cgd4                   ------------------------------------------------------------
cgd2              554  QEHEREHDQEQGQKHDQKYQDDYKDMEEHLEGSDSAEHISNVHDYYGDSDDENSSDNEEL  613
cgd7                   ------------------------------------------------------------
cgd6                   ------------------------------------------------------------
cgd3                   ------------------------------------------------------------
2oz2_chainA_p001       ------------------------------------------------------------
1fh0_chainA_p002       ------------------------------------------------------------
Consensus_aa:          ............................................................
Consensus_ss:

Conservation:
cgd4              457  ------------------------------------------------EGERMYAEEYGY  468
cgd2              614  GHLSGAKYNIPKKFNVIQQLCWDLGGHMGAANTQCRKDIPITKNIPKSCSKIIKVKEYSY  673
cgd7                   ------------------------------------------------------------
cgd6                   ------------------------------------------------------------
cgd3                   ------------------------------------------------------------
2oz2_chainA_p001       ------------------------------------------------------------
1fh0_chainA_p002       ------------------------------------------------------------
Consensus_aa:          ............................................................
Consensus_ss:

Conservation:                              976 76       5  9  9
cgd4              469  VGGCYGCCDEDRMKEEIFKNGPIAVAMHIDT-SLLVYDNGVYDSIPNDHTKYCDLPNKQL  527
cgd2              674  VNNVYGKTTARDIMESLWNEGPVAVSLEPTL-EFSLYNSGVFKGFYDPITRQ---YPWSS  729
cgd7              514  -LALNGDDGWDFVTTILPKVGSISLSINSNLPGFSSYSDGIYKAPKCTT-----------  561
cgd6              295  ------PRNINALKTALAKYGPISVAIQADQTPFQFYKSGVFDAPCG-------------  335
cgd3              169  ------EDPWDVITDAIYNYGPVTVSVCSLMPGFNLYSGGYYEPPTCGSI---------W  213
2oz2_chainA_p001  120  -------QDEAQIAAWLAVNGPVAVAVDASS--WMTYTGGVMTSCVS-------------  157
1fh0_chainA_p002  117  ------PGKEKALMKAVATVGPISVAMDAGHSSFQFYKSGIYFEPDCS------------  158
Consensus_aa:          ........sbp.l.p.lhp.GPltVthpss..s@.hYpsGl@ps..s.............
Consensus_ss:                   hhhhhhhhhh   eeeeeee   h       ee

Conservation:                6965   965          96   9997  99  96  5        5   5
cgd4              528  NGWEYTNHAIAIVGWGEE----NGIPYWIIRNSWGANWGKKGYAKIRRGK----NIGGIE  579
cgd2              730  IPWHKVDHAMVITGWGWETYGNERIPYWIVQNSWGKRWGEKGFCRIIRGV----NELSIE  785
cgd7              562  --HSELNHAVIMIGYGIND---NGDKYYVIQNSWGVSWGIGGFMNVSADSCDMFWYPGII  616
cgd6              336  ---TKVNHGVVLVGYDMDE--DTNKEYWLVRNSWGEAWGEKGYIKLALHSGK-KGTCGIL  389
cgd3              214  CGTRQVDHAVTLIGYGVSE---SGKRYYIMKNSWGLSWGNKGFMNISAD------MCSTF  264
2oz2_chainA_p001  158  ---EQLDHGVLLVGYNDS----AAVPYWIIKNSWTTQWGEEGYIRIAKGS----NQCLVK  206
1fh0_chainA_p002  159  --SKNLDHGVLVVGYGFEGANSDNSKYWLVKNSWGPEWGSNGYVKIAKDK---NNHCGIA  213
Consensus_aa:          ...pplsHtlhllG@s.p......ss..Y@llpNSWG.pWGppG@hpl..sp....s.stlb
Consensus_ss:                 eeeeeeeee         eeeeee           eeeee

Conservation:
cgd4              580  NQAVFIDPDFTRGMGLS-----LLNKYQNSTAYRTNLPDIKDSPQGSE--KVNVEIVNGG  632
cgd2              786  HAAVRASVKIYDSDRKY-----KAADLQKVHD-------------------ESV-FQYL-  819
cgd7              617  RQVSSESLPDKCTGNKLLLTGPGEINKPQKKSTEYNIFLKKTYYGICLIINILL-LVY--  673
cgd6              390  VEPVYPVINQSI------------------------------------------------  401
cgd3              265  FNPGWVTSVSMDGISDY---------------------CLNREPDKNE--IAH-DFVI--  298
2oz2_chainA_p001  207  EEASSAVVG---------------------------------------------------  215
1fh0_chainA_p002  214  TAASYPNVX---------------------------------------------------  222
Consensus_aa:          ..ss..s.s...................................................
Consensus_ss:             eeeeee

Conservation:
cgd4              633  VDQ  635
cgd2                   ---
cgd7                   ---
cgd6                   ---
cgd3                   ---
2oz2_chainA_p001       ---
1fh0_chainA_p002       ---
Consensus_aa:          ...
Consensus_ss:
```

Supplementary Figure 2.

Full multiple sequence alignment (MSA) of family C01 peptidases from *C. parvum* and the sequences for two structures (cruzain and human cathepsin L). Only the regions corresponding to active site residues are shown in Figure 7 in the main article. All five C01 C. parvum sequences as well as PDBs for cruzain (2OZ2) and human cathepsin L (1FH0) were used as input into PROMALS3D (see Methods) to generate the MSA. The active site residues (Q, C, H, and N in blue boxes) are strictly conserved but the S2 pocket Glu in cryptopain 1 is in a variable region and appears to be conserved only in cruzain (pink box). Cathepsin L-like sequences (“cyrptopains”): cgd6 = cryptopain 1, cgd3 = cyrptopain 2, cgd7 = cryptopain 3; cathepsin C-like sequences: cgd2 and cgd4. Higher conservation is shown with higher numbers in “Conservation.” Strictly conserved amino acids are shown as bold, upper case letters in “Consensus_aa.” Consensus_ss denotes where conserved secondary structure is predicted: e = beta strand (blue letters), h = alpha helix (red letters).