Novel cruzain inhibitors for the treatment of Chagas' disease

Kathleen Rogers[a,*], **Henrik Keränen**[b], **Jacob D. Durrant**[c], **Joseline Ratnam**[d], **Allison Doak**[d], **Michelle R. Arkin**[d], and **J. Andrew McCammon**[e,f,g]

[a]Biomedical Sciences Graduate Program, University of California San Diego, La Jolla, CA, 92093

[b]Department of Cell and Molecular Biology, Uppsala University, S-751 24 Uppsala, Sweden

[c]Department of Chemistry & Biochemistry, University of California San Diego, La Jolla, CA, 92093

[d]Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA, 94158

[e]Department of Chemistry & Biochemistry, NSF Center for Theoretical Biological Physics, National Biomedical Computation Resource, University of California San Diego, La Jolla, CA, 92093

[f]Department of Pharmacology, University of California San Diego, La Jolla, CA, 92093

[g]Howard Hughes Medical Institute, University of California San Diego, La Jolla, CA

Abstract

The protozoan parasite *Trypanosoma cruzi,* the etiological agent of Chagas' disease, affects millions of individuals and continues to be an important global health concern. The poor efficacy and unfavorable side effects of current treatments necessitate novel therapeutics. Cruzain, the major cysteine protease of *T. cruzi*, is one potential novel target. Recent advances in a class of vinyl-sulfone inhibitors are encouraging; however, as most potential therapeutics fail in clinical trials and both disease progression and resistance call for combination therapy with several drugs, the identification of additional classes of inhibitory molecules is essential. Using an exhaustive virtual-screening and experimental-validation approach, we identify several additional small-molecule cruzain inhibitors. Further optimization of these chemical scaffolds could lead to the development of novel drugs useful in the treatment of Chagas' disease.

Keywords

cruzain; cruzipain; Chagas' disease; *Trypanosoma cruzi*; computer-aided drug discovery; cysteine protease inhibitor

Introduction

The World Health Organization estimates that over 10 million people are infected by the protozoan parasite *Trypanosoma cruzi*, with another 25 million at risk (1). The associated illness, called Chagas' disease, is spread through a triatomine vector or through blood transfusion (2, 3). The acute phase lasts at most a few months and is characterized by mild symptoms such as fever, malaise, facial edema, generalized lymphadenopathy, and hepatosplenomegaly. In approximately 30% of infected patients, parasite multiplication *via* asynchronous cycles contributes to the chronic stage of the disease, with the associated cell destruction, reinfection within the reticuloendothelial system, and organ infection (4). Infection of the heart, digestive tract, and central nervous system can lead to fatal heart-rhythm abnormalities, megacolon, and dementia, respectively (5, 6).

T. cruzi is not susceptible to many of the drugs used to treat closely related parasites like *Trypanosoma brucei.* Benznidazole and nifurtimox are the only available therapies for acute-phase Chagas' disease. These nitroheterocyclics are highly toxic and have poor efficacy in long-lasting chronic infections (7–9). No extensive studies of the long-term sequellae of these therapeutics have been conducted in humans, but several reports of neuropathy and tumorigenic or carcinogenic effects have been described (7, 8). Efforts to develop a vaccine against *T. cruzi* have also failed thus far, likely because the disease pathology has an autoimmune component (10).

The major *T. cruzi* cysteine proteinase cruzain (also referred to as cruzipain, the full-length native enzyme) has been shown to be crucial for all stages of the parasite life cycle. This papain-like cysteine protease is thought to play an important role in differentiation, cell invasion, intracellular multiplication, and immune evasion (11, 12). Furthermore, studies have demonstrated that cysteine-proteinase inhibitors have trypanocidal activity with negligible mammalian toxicity (13).

Previous efforts have identified vinyl sulfones, sulfonates, and sulfonamides as high-affinity cruzain inhibitors (14, 15); one of these vinyl sulfones, K11777, is currently undergoing Investigational New Drug (IND) enabling studies (16, 17). α-ketoamide-, α-ketoacid, α-ketoester-, aldehyde-, and ketone-based inhibitors have also been described (18–20). While these successes are encouraging, many potential drugs, including those that enter clinical trails, ultimately fail to gain approval (21), and those that are approved are subject to growing parasitic resistance. Consequently, a diverse set of inhibitory scaffolds that can be optimized into distinct therapeutic candidates is urgently needed.

Hoping to contribute to this ever-growing diverse set of compounds, we here use an advanced virtual-screening methodology that accounts for receptor flexibility to identify three promising non-covalent inhibitors of *T. cruzi* cruzain.

Experimental Methods

Ligand Preparation

A small-molecule library was prepared from the ligands of the NCI Diversity Set II using the Schrödinger LigPrep program (22). Protonation states were assigned at pH 5.5 to mimic the natural acidic environment of the *T. cruzi* digestive vacuole. Multiple tautomers and stereoisomers were generated. One ligand could not be processed by LigPrep; instead, Discovery Studio (23) was used to add hydrogen atoms to this ligand and to optimize its geometry.

Initial Screen Against the Crystal Structure

The prepared ligand models of this small-molecule library were docked into a 1.20 Å crystal structure of cruzain (PDB ID: 1ME4)(19), with hydrogen atoms included using PDB2PQR (24, 25) at pH 5.5. Residues CYS25 and H159 (called H162 by some) formed the thiolate-imidazolium pair required for the catalytic mechanism (26) of the proteinase at this pH. This initial virtual screen was performed using the CDOCKER docking software (23) with a docking sphere 15 Å in diameter centered on the coordinates of the crystallographic ligand.

Rescoring Protocol

The CDOCKER-predicted pose of each ligand model was rescored using six additional scoring functions: LigScore1, LigScore2 (27), PLP1, PLP2 (28), PMF (29), and PMF04 (30). The best scoring models as evaluated using each of these seven scoring functions were compiled into a new small-molecule library of 302 models (182 unique ligands) enriched for predicted *cruzain* inhibitors.

Molecular Dynamics Simulations

The molecular dynamics simulations used in the current study have been described previously (31). In brief, the simulations were based on a 1.20 Å cruzain crystal structure (PDB ID: 1ME4) (19) protonated at pH 5.5 to mimic the natural acidic environment of the *T. cruzi* digestive vacuole. Following appropriate minimization and equilibration, five distinct 20-ns simulations of the cruzain protein bound to a hydroxymethyl ketone inhibitor, [1-(1-BENZYL-3-HYDROXY-2-OXO-PROPYLCARBAMOYL)-2-PHENYL-ETHYL]-CARBAMIC ACID BENZYL ESTER, were performed. The *gromos* clustering algorithm (32) was used to cluster 4,002 conformations extracted from the simulations every 50 fs. We found that decreasing the cutoff below 0.95 Å resulted in a precipitous rise in the number of clusters; consequently, we chose an RMSD cutoff of 0.95 Å, which yielded 24 clusters. The central member of each cluster, considered most representative, was selected for subsequent analysis; this set of central members is said to constitute an *ensemble*.

Relaxed-Complex Screen

The 302 compound models of the enriched small-molecule library were docked into the 24 clusters of the ensemble using CDOCKER (Accelrys). Each of these docked small-molecule models was rescored with the following scoring functions: LigScore2 (27), PLP1, PLP2 (28), PMF (29), and PMF04 (30). For each ligand/scoring-function pair, an ensemble-average score was calculated according to the following equation:

$$\bar{E} = \frac{\sum_{i=1}^{23} w_i E_i}{\sum_{i=1}^{23} w_i} \quad (1)$$

where   is the weighted ensemble-average score, w_i is the size of cluster i, and E_i is the best score of each unique ligand, independent of tautomeric or stereoisomeric form, docked into the centroid of cluster i.

Two methods were used to select compounds for subsequent experimental validation. First, for each of the five scoring functions, the compounds were ranked from best to worst by the ensemble-average score. The top seven compounds were selected from each of these five ordered lists and merged into a single list of potential binders. Second, the average rank of each compound across all five scoring functions was calculated. The compounds were then

reordered by this average rank, and the top thirty were likewise identified as potential binders. Any compound indicated by either of these two protocols was subsequently recommended for preliminary experimental validation.

Enzymatic Assays

Each compound was obtained from the National Cancer Institute's Development Therapeutics Program, which guaranteed 90% purity. Compounds were tested for cruzain enzymatic inhibition using a protocol that has been described previously (33). The eight compounds with the lowest IC_{50} values were assessed for aggregation by observing enzymatic activity under varying experimental conditions. As detergent is known to disrupt colloidal aggregation (34), inhibition in the absence of detergent was compared to inhibition in the presence of Triton X-100 (0.02%) and Tween (0.002%). The reducing agent was also varied (10mM DTT or 10mM Beta-Mercaptoethanol). Each experimental condition was tested in at least two separate experiments. Finally, a dynamic light scattering (DLS) technique, described in detail elsewhere (35), was applied to the top four compounds to further confirm the presence or absence of aggregation. Additional experimental details can be found in the Supporting Information.

Final Pose Predictions

All compounds submitted for experimental validation were subsequently docked a final time into the binding pocket of the crystal structure (PDB ID: 1ME4) using the Induced Fit Docking (IFD) module of the Maestro 9.2 (Schrödinger, LLC) computer package with Glide XP precision (36, 37). For each of the top three non-aggregating ligands, the best-scoring pose that positioned the inhibitor near the crucial catalytic triad was selected and visualized using PyMOL (38). Although the top poses using this IFD protocol were generally similar to those from the RCS CDOCKER work, we choose to show them here in the commonly represented crystal structure conformation for ease of recognition to the reader of subsites within the well-characterized active site of cruzain.

Results and Discussion

Discovered by Carlos Chagas in 1909 (39, 40), *T. cruzi* is one of only two known pathogenic *Trypanosoma* species. Current trypanocidal therapeutics like nifurtimox and benznidazole are inadequate because they are toxic (7–9), subject to growing resistance (41), and ineffective at eradicating the parasite and preventing cardiomyopathy over the long term (42). Given the dire need for novel therapies, we here use virtual-screening methods to identify three promising inhibitors of cruzain, a critical cysteine protease required for *T. cruzi* survival.

Weaknesses of Virtual Screening

Virtual-screening techniques have been used to identify a number of inhibitors in recent years (see, for example, references (43–49)). Though widely used, these screens are often characterized by many false positives and negatives. Two principal weaknesses explain these inaccuracies. First, there are errors intrinsic to the scoring functions themselves. Because virtual-screening efforts often attempt to identify true binders from among the many thousands of molecules in a compound library, they are generally optimized for speed at the expense of accuracy. Second, current docking programs account for, at best, only limited receptor flexibility. When a small-molecule binder encounters its receptor *in vivo*, that receptor often undergoes conformational rearrangements or an "induced fit" to better accommodate the ligand. These *holo* conformations can differ significantly from those of x-ray crystallographic structures. Even if the perfect docking scoring function did exist, it

could not accurately predict binding affinity if receptor flexibility were ignored. In the current work, we use a scheme designed to minimize both these sources of error.

Overcoming Inaccuracies Inherent to the Scoring Functions Themselves

The ligands of the NCI Diversity Set II (50) were initially docked into a cruzain crystal structure using CDOCKER (23) because this program was able to capture the crystallographic poses of two positive controls. Docking with AutoDock (51) was initially performed, but docked and crystallographic poses differed significantly. The CDOCKER-docked poses were subsequently rescored using several different scoring functions, and potential binders were selected by consensus rather than by the score of any single function. Consensus scoring has two advantages. First, when multiple scoring functions are combined, the errors of each may in part cancel out (52). Second, different scoring functions likely have different intrinsic weaknesses and strengths. Some, for example, may better account for hydrophobic contacts, while others better capture electrostatic interactions. When combined, accuracy may be improved if the weaknesses and strengths of the constituent functions are complementary (53).

The scoring functions used in the current work come from different classes and thus provide independent assessments of ligand binding that may be complementary. Scoring functions fall into one of three categories: force field, empirical, and knowledge based. Force-field scoring functions like that used by CDOCKER, based on the CHARMm force field, evaluate ligand binding by accounting for bonded and non-bonded atomic interactions explicitly. Empirical scoring functions like LigScore1, LigScore2, PLP1, and PLP2 are based on counting the number of different types of receptor-ligand interactions (e.g. hydrogen-bond and hydrophobic interactions) as well as countable changes in molecular properties (e.g., the number of rotatable bonds immobilized upon ligand binding). Finally, knowledge-based scoring functions like PMF and PMF04 are based on statistical analyses of large structural databases like subsets of the Protein Data Bank (54). Intermolecular interactions that occur more often than expected by pure chance are assumed to be energetically favorable.

The top candidate binders from this initial crystal-structure screen, as judged by consensus scoring, were compiled into a single list of 302 small-molecule models that were subsequently subjected to further study.

Overcoming Inaccuracies Caused by Ignoring Full Receptor Flexibility

While consensus scoring may have helped overcome in part the errors intrinsic to each individual scoring function, those caused by ignoring receptor flexibility remained. In order to overcome this second challenge, we first studied receptor dynamics by performing five distinct 20-ns molecular dynamics simulations of cruzain, described elsewhere (31). The protein conformations sampled during these simulations were clustered into 24 groups using an RMSD-based clustering algorithm (32). Each of the centroid members of each cluster, considered to be the most representative, was identified, and the group of all centroids is said to constitute an *ensemble*.

Having identified multiple cruzain conformations, we re-docked the 302 small-molecule models identified in the initial crystal-structure screen into each of the 24 members of the ensemble, again using CDOCKER. The ligands were then re-ranked by an ensemble-average score that was not dependent on a single crystal structure but rather accounted for receptor flexibility. This multi-receptor docking protocol, called the relaxed complex scheme (RCS), has been used successfully in the past to identify inhibitors of FKBP (44), HIV integrase (43), and *T. brucei* RNA editing ligase 1 (45), among others (46–48, 55).

Final Rescoring

As each of the 302 small-molecule models was docked into 24 different *cruzain* conformations, there were 7,248 docked poses in all. Each of these 7,248 poses was rescored, again using multiple scoring functions. Ensemble-average scores were calculated for each scoring function, and the 302 small-molecule models were appropriately ranked.

Two criteria were used to identify likely inhibitors from among these ranked compounds. First, we selected the 7 best inhibitors as predicted by each of the individual scoring functions and merged them into a single list of 22 candidates. Second, the top 30 predicted ligands as judged by the ensemble-average rank were likewise identified. In all, 37 unique candidate inhibitors were selected; of these, the best 30 were tested experimentally.

Predicted Binding Modes

Of the 30 compounds tested, eight inhibited cruzain at 100 μM (Figure 1, all structures below the first arrow). Using standard conditions, the best-scoring compound had an initial IC_{50} value of 471 nM (see Figure 1, asterisk). Subsequent experiments suggested that this compound inhibited cruzain nonspecifically *via* aggregation (Supporting Information). Fortunately, three other compounds, although less potent, did in fact appear to inhibit cruzain specifically (Supporting Information). These three compounds, NSC 227186, NSC 67436, and NSC 260594, have IC_{50} values of 16 μM, 63 μM, and 66 μM, respectively (see Figure 1, bottom row). While these compounds lack the nanomolar affinity characteristic of approved drugs, they do possess the low micromolar affinity typical of lead compounds. With proper optimization, including fragment addition, moiety swapping, and similarity searching, these compounds could be transformed into viable drug candidates. We are hopeful that these new leads will be helpful to those in the drug-discovery community targeting Chagas' disease.

Figure 2 shows the predicted binding poses (a,c,e) and important interactions (b,d,f) associated with each of the three most promising compounds using the IFD protocol and visualized in PyMOL (described in *Methods*). Fig. 2b shows the predicted polar contacts of NSC 227186 with GLN19, THR185, and TRP184; TRP184 is also oriented for possible aromatic stacking with one of the ligand aromatic moieties. GLU208 (called GLU205 by some) is predicted to swing even further away from the cruzain S2 subsite than is seen in the 1ME4 crystal structure, allowing a ligand methylpyrrolidine moiety to occupy the site, reminiscent of the Phe and Tyr rings of several known inhibitors (14–19, 56, 57). An oxane moiety is also predicted to occupy the S3' subsite (14).

NSC 67436 (Fig.2d) is predicted to form even more interactions with the cruzain receptor, including various hydrogen-bond interactions with residues GLN19, GLY66, ALA141, ASP161, and ASN182, as well as possible aromatic stacking with HIS162. Two distinct cyclohexanamine moieties are predicted to occupy the S2 and S1' subsites, and a chlorocyclohexane (4-chlorocyclohexane-1,3-dicarbaldehyde) moiety is predicted to bind near S1. S4 and S3' are also occupied, both with distinct imidazoline rings (14).

Fig. 2f shows the hydrogen bonds predicted to form between NSC 260594 and MET68, ASN69, and ASP161. Again reminiscent of several known ligands, the S2 subsite is occupied by an aromatic ring system, and a slightly rotated GLU208 accommodates the larger aromatic moiety (1-methyl-6-nitro-decahydroquinolin-4-amine).

In an attempt to identify predicted binding characteristics that might aid the identification of future inhibitors, we used the Induced Fit Docking (IFD) module of the Schrödinger Maestro computer package to re-dock the 30 experimentally tested compounds into a crystallographic active-site model (PBD ID: 1ME4) at pH 5.5. Unlike CDOCKER docking,

the IFD protocol allows for local active-site conformational changes and so is judged to better predict ligand binding, albeit at the cost of speed. The most favorable IFD poses of the validated inhibitors (Figure 2) consistently placed ligand atoms near the catalytic triad in a position that could conceivably compromise cruzain enzymatic activity. Additionally, the IFD poses of these compounds also place at least one aromatic ring in the S2 subsite, a site known to be favorable for the binding of Phe and Tyr aromatic side chains (14–19, 56, 57).

These predicted binding poses are promising because they represent unique scaffolds that differ significantly from those of previously identified small-molecule inhibitors. While the poses included here are mere predictions, we are hopeful that they will guide future optimization of these experimentally validated ligands.

Conclusions

Chagas' disease is the leading cause of heart failure in Latin America, affecting over 10 million individuals worldwide. Although progress has been made in the treatment of the disease, especially given the recent success of K11777, multiple cruzain inhibitors are needed given the difficulty of obtaining FDA approval and ever-progressing drug resistance. The work presented here provides chemical scaffolds that, with further optimization, could be developed into promising therapeutics for Chagas' disease. The exhaustive virtual-screening approach and thorough experimental validation used to identify these leads also represents a promising and useful method for inhibitor identification.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. (WHO) WHO. Chagas Disease (American trypanosomiasis), Fact Sheet 340. WHO; Chagas Disease (American trypanosomiasis), Fact Sheet 340.

2. Rassi A, Rassi A, Marin-Neto JA. Chagas disease. Lancet. 2010; 375:1388–1402. [PubMed: 20399979]

3. Shulman IA, Appleman MD, Saxena S, Hiti AL, Kirchhoff LV. Specific antibodies to Trypanosoma cruzi among blood donors in Los Angeles, California. Transfusion. 1997; 37:727–731. [PubMed: 9225937]

4. Higuchi Mde L, Benvenuti LA, Martins Reis M, Metzger M. Pathophysiology of the heart in Chagas' disease: current status and new developments. Cardiovasc Res. 2003; 60:96–107. [PubMed: 14522411]

5. Tanowitz HB, Kirchhoff LV, Simon D, Morris SA, Weiss LM, Wittner M. Chagas' disease. Clin Microbiol Rev. 1992; 5:400–419. [PubMed: 1423218]

6. Prata A. Clinical and epidemiological aspects of Chagas disease. Lancet Infect Dis. 2001; 1:92–100. [PubMed: 11871482]

7. Coura JR, de Castro SL. A critical review on Chagas disease chemotherapy. Memorias Do Instituto Oswaldo Cruz. 2002; 97:3–24. [PubMed: 11992141]

8. Castro JA, de Mecca MM, Bartel LC. Toxic side effects of drugs used to treat Chagas' disease (American trypanosomiasis). Human & Experimental Toxicology. 2006; 25:471–479. [PubMed: 16937919]

9. Jackson Y, Alirol E, Getaz L, Wolff H, Combescure C, Chappuis F. Tolerance and safety of nifurtimox in patients with chronic chagas disease. Clin Infect Dis. 2010; 51:e69–e75. [PubMed: 20932171]

10. Quijano-Hernandez I, Dumonteil E. Advances and challenges towards a vaccine against Chagas disease. Hum Vaccin. 2011; 7:1184–1191. [PubMed: 22048121]

11. Tomas AM, Miles MA, Kelly JM. Overexpression of cruzipain, the major cysteine proteinase of Trypanosoma cruzi, is associated with enhanced metacyclogenesis. Eur J Biochem. 1997; 244:596–603. [PubMed: 9119029]

12. Doyle PS, Zhou YM, Hsieh I, Greenbaum DC, McKerrow JH, Engel JC. The Trypanosoma cruzi protease cruzain mediates immune evasion. PLoS Pathog. 2011; 7 e1002139.

13. McKerrow JH, Rosenthal PJ, Swenerton R, Doyle P. Development of protease inhibitors for protozoan infections. Curr Opin Infect Dis. 2008; 21:668–672. [PubMed: 18978536]

14. Brinen LS, Hansell E, Cheng J, Roush WR, McKerrow JH, Fletterick RJ. A target within the target: probing cruzain's P1' site to define structural determinants for the Chagas' disease protease. Structure. 2000; 8:831–840. [PubMed: 10997902]

15. Kerr ID, Lee JH, Farady CJ, Marion R, Rickert M, Sajid M, et al. Vinyl sulfones as antiparasitic agents and a structural basis for drug design. J Biol Chem. 2009; 284:25697–25703. [PubMed: 19620707]

16. Bryant C, Kerr ID, Debnath M, Ang KK, Ratnam J, Ferreira RS, et al. Novel non-peptidic vinylsulfones targeting the S2 and S3 subsites of parasite cysteine proteases. Bioorg Med Chem Lett. 2009; 19:6218–6221. [PubMed: 19773167]

17. Sajid M, Robertson SA, Brinen LS, McKerrow JH. Cruzain : the path from target validation to the clinic. Adv Exp Med Biol. 2011; 712:100–115. [PubMed: 21660661]

18. Choe Y, Brinen LS, Price MS, Engel JC, Lange M, Grisostomi C, et al. Development of alpha-keto-based inhibitors of cruzain, a cysteine protease implicated in Chagas disease. Bioorg Med Chem. 2005; 13:2141–2156. [PubMed: 15727867]

19. Huang L, Brinen LS, Ellman JA. Crystal structures of reversible ketone-based inhibitors of the cysteine protease cruzain. Bioorg Med Chem. 2003; 11:21–29. [PubMed: 12467703]

20. Huang L, Lee A, Ellman JA. Identification of potent and selective mechanism-based inhibitors of the cysteine protease cruzain using solid-phase parallel synthesis. J Med Chem. 2002; 45:676–684. [PubMed: 11806719]

21. Dimasi JA. Risks in new drug development: approval success rates for investigational drugs. Clin Pharmacol Ther. 2001; 69:297–307. [PubMed: 11371997]

22. LigPrep, version 2.5. New York, NY: Schrödinger, LLC; 2011.

23. Discovery Studio, version 2.1. San Diego, CA, 92121, U.S.A.: Accelrys; 2008.

24. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, et al. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic Acids Res. 2007; 35:W522–W525. [PubMed: 17488841]

25. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res. 2004; 32:W665–W667. [PubMed: 15215472]

26. Duschak VG, Couto AS. Cruzipain, the major cysteine protease of Trypanosoma cruzi: a sulfated glycoprotein antigen as relevant candidate for vaccine development and drug target. A review. Curr Med Chem. 2009; 16:3174–3202. [PubMed: 19689291]

27. Krammer A, Kirchhoff PD, Jiang X, Venkatachalam CM, Waldman M. LigScore: a novel scoring function for predicting binding affinities. J Mol Graph Model. 2005; 23:395–407. [PubMed: 15781182]

28. Gehlhaar DK, Verkhivker GM, Rejto PA, Sherman CJ, Fogel DB, Fogel LJ, et al. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. Chem Biol. 1995; 2:317–324. [PubMed: 9383433]

29. Muegge I, Martin YC. A general and fast scoring function for protein-ligand interactions: a simplified potential approach. J Med Chem. 1999; 42:791–804. [PubMed: 10072678]

30. Muegge I. PMF scoring revisited. J Med Chem. 2006; 49:5895–5902. [PubMed: 17004705]

31. Durrant JD, Keranen H, Wilson BA, McCammon JA. Computational identification of uncharacterized cruzain binding sites. PLoS Negl Trop Dis. 2010; 4:e676. [PubMed: 20485483]

32. Christen M, Hunenberger PH, Bakowies D, Baron R, Burgi R, Geerke DP, et al. The GROMOS software for biomolecular simulation: GROMOS05. J Comput Chem. 2005; 26:1719–1751. [PubMed: 16211540]

33. Ferreira RS, Bryant C, Ang KK, McKerrow JH, Shoichet BK, Renslo AR. Divergent Modes of Enzyme Inhibition in a Homologous Structure-Activity Series. J Med Chem. 2009; 52:5005–5008. [PubMed: 19637873]

34. Feng BY, Simeonov A, Jadhav A, Babaoglu K, Inglese J, Shoichet BK, et al. A high-throughput screen for aggregation-based inhibition in a large compound library. J Med Chem. 2007; 50:2385–2390. [PubMed: 17447748]

35. McGovern SL, Helfand BT, Feng B, Shoichet BK. A specific mechanism of nonspecific inhibition. J Med Chem. 2003; 46:4265–4272. [PubMed: 13678405]

36. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R. Novel procedure for modeling ligand/receptor induced fit effects. J Med Chem. 2006; 49:534–553. [PubMed: 16420040]

37. Sherman W, Beard HS, Farid R. Use of an induced fit receptor structure in virtual screening. Chem Biol Drug Des. 2006; 67:83–84. [PubMed: 16492153]

38. The PyMOL Molecular Graphics System, Version 1.2r3pre. Schrödinger, LLC;

39. Chagas C. New human trypanosomiasis. Morphology and life cycle of Schyzotrypanum cruzi, the cause of a new human disease. Mem Inst Oswaldo Cruz. 1909; 1:159–218.

40. Chagas C. A new human disease. Summary of etiological and clinical studies. Mem Inst Oswaldo Cruz. 1911; 3:219–275.

41. Wilkinson SR, Taylor MC, Horn D, Kelly JM, Cheeseman I. A mechanism for cross-resistance to nifurtimox and benznidazole in trypanosomes. Proc Natl Acad Sci U S A. 2008; 105:5022–5027. [PubMed: 18367671]

42. Lauria-Pires L, Braga MS, Vexenat AC, Nitz N, Simoes-Barbosa A, Tinoco DL, et al. Progressive chronic Chagas heart disease ten years after treatment with anti-Trypanosoma cruzi nitroderivatives. Am J Trop Med Hyg. 2000; 63:111–118. [PubMed: 11388500]

43. Schames JR, Henchman RH, Siegel JS, Sotriffer CA, Ni H, McCammon JA. Discovery of a novel binding trench in HIV integrase. J Med Chem. 2004; 47:1879–1881. [PubMed: 15055986]

44. Lin JH, Perryman AL, Schames JR, McCammon JA. Computational drug design accommodating receptor flexibility: the relaxed complex scheme. Journal of the American Chemical Society. 2002; 124:5632–5633. [PubMed: 12010024]

45. Amaro RE, Schnaufer A, Interthal H, Hol W, Stuart KD, McCammon JA. Discovery of drug-like inhibitors of an essential RNA-editing ligase in Trypanosoma brucei. Proceedings of the National Academy of Sciences. 2008; 105:17278–17283.

46. Durrant JD, Hall L, Swift RV, Landon M, Schnaufer A, Amaro RE. Novel Naphthalene-Based Inhibitors of Trypanosoma brucei RNA Editing Ligase 1. Plos Neglected Tropical Diseases. 2010; 4

47. Durrant JD, Cao R, Gorfe AA, Zhu W, Li J, Sankovsky A, et al. Non-bisphosphonate inhibitors of isoprenoid biosynthesis identified via computer-aided drug design. Chem Biol Drug Des. 2011; 78:323–332. [PubMed: 21696546]

48. Durrant JD, Urbaniak MD, Ferguson MA, McCammon JA. Computer-aided identification of Trypanosoma brucei uridine diphosphate galactose 4'-epimerase inhibitors: toward the development of novel therapies for African sleeping sickness. J Med Chem. 2010; 53:5025–5032. [PubMed: 20527952]

49. Wang Y, Hess TN, Jones V, Zhou JZ, McNeil MR, Andrew McCammon J. Novel inhibitors of Mycobacterium tuberculosis dTDP-6-deoxy-l-lyxo-4-hexulose reductase (RmlD) identified by virtual screening. Bioorg Med Chem Lett. 2011; 21:7064–7067. [PubMed: 22014548]

50. "The NCI/DTP Open Chemical Repository". http://dtp.cancer.gov ".

51. Morris GM, Huey R, Olson AJ. Using AutoDock for ligand-receptor docking. Curr Protoc Bioinformatics. 2008; Chapter 8(Unit 8):14. [PubMed: 19085980]

52. Wang R, Wang S. How does consensus scoring work for virtual library screening? An idealized computer experiment. J Chem Inf Comput Sci. 2001; 41:1422–1426. [PubMed: 11604043]

53. Bar-Haim S, Aharon A, Ben-Moshe T, Marantz Y, Senderowitz H. SeleX-CS: a new consensus scoring algorithm for hit discovery and lead optimization. J Chem Inf Model. 2009; 49:623–633. [PubMed: 19231809]

54. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucl Acids Res. 2000; 28:235–242. [PubMed: 10592235]

55. Durrant JD, Urbaniak MD, Ferguson MA, McCammon JA. Computer-Aided Identification of Trypanosoma brucei Uridine Diphosphate Galactose 4'-Epimerase Inhibitors: Toward the Development of Novel Therapies for African Sleeping Sickness. J Med Chem. 2010; 53:5025–5032. [PubMed: 20527952]

56. Gillmor SA, Craik CS, Fletterick RJ. Structural determinants of specificity in the cysteine protease cruzain. Protein Sci. 1997; 6:1603–1611. [PubMed: 9260273]

57. Mott BT, Ferreira RS, Simeonov A, Jadhav A, Ang KK, Leister W, et al. Identification and optimization of inhibitors of Trypanosomal cysteine proteases: cruzain, rhodesain, and TbCatB. J Med Chem. 2010; 53:52–60. [PubMed: 19908842]
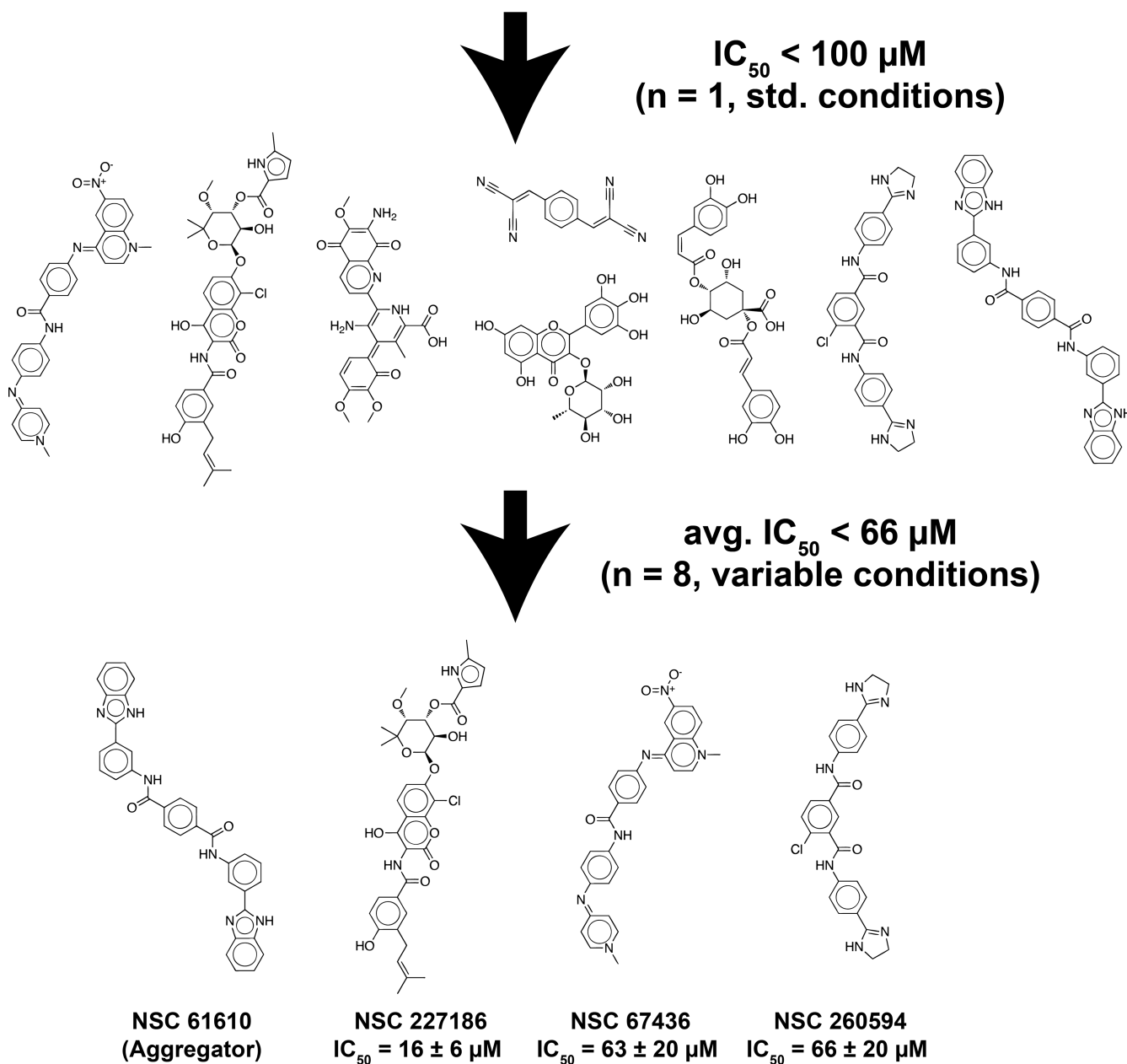
Thirty Predicted Inhibitors

$IC_{50} < 100$ μM
(n = 1, std. conditions)

avg. $IC_{50} < 66$ μM
(n = 8, variable conditions)

NSC 61610
(Aggregator)

NSC 227186
$IC_{50} = 16 \pm 6$ μM

NSC 67436
$IC_{50} = 63 \pm 20$ μM

NSC 260594
$IC_{50} = 66 \pm 20$ μM

Figure 1.
The structures of all experimentally validated compounds. Associated IC_{50} values are given for the top three non-aggregating inhibitors. Means and standard deviations were calculated by considering all IC_{50} values associated with each compound, measured under varying experimental conditions as described in the Experimental Methods (*n* = 8). *NSC 61610 appears to inhibit cruzain nonspecifically via aggregation (see the Supporting Information).
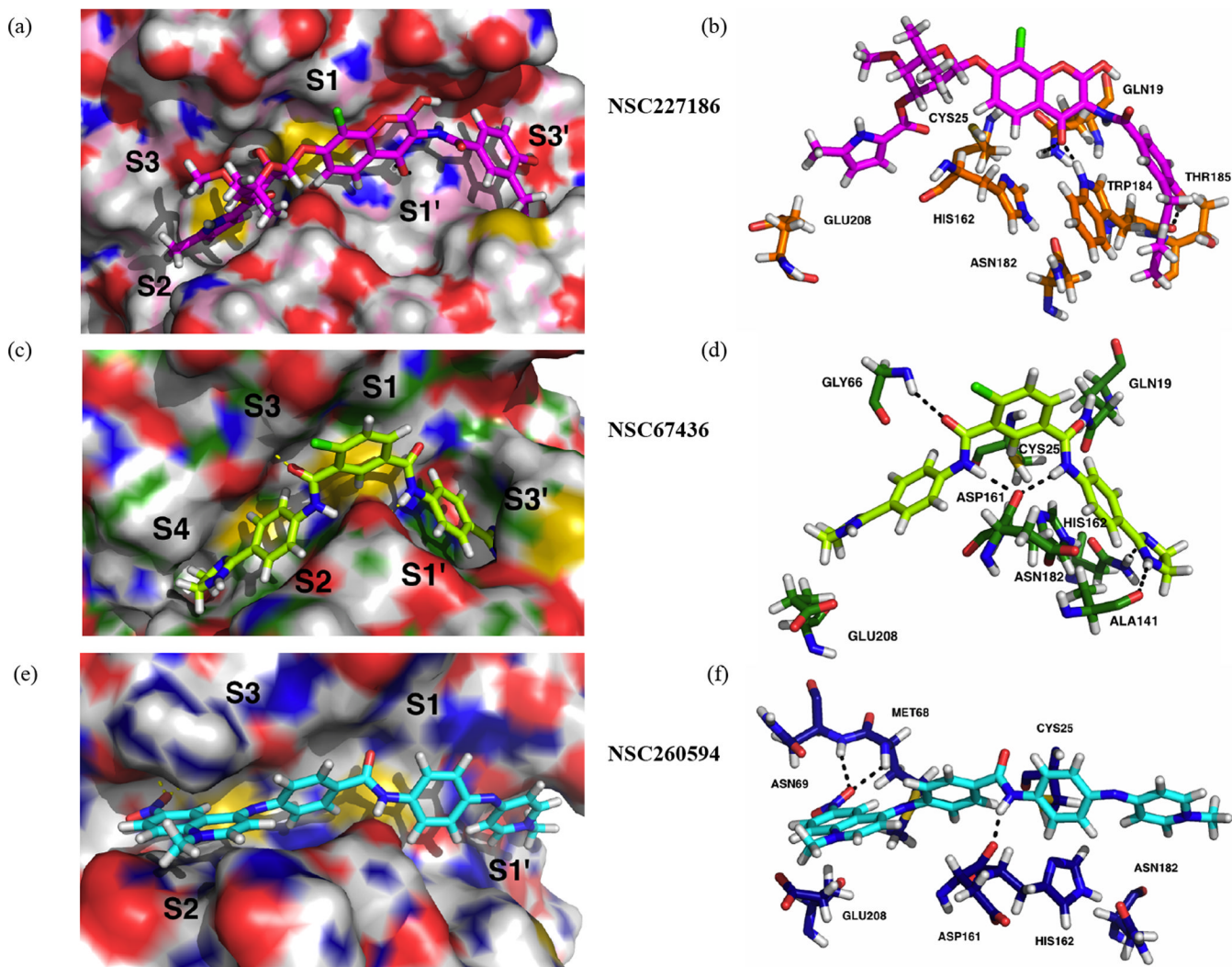
Figure 2.
The predicted binding poses and receptor-ligand interactions of the experimentally validated inhibitors. A, C, and E: Compounds and cruzain active sites are shown in stick and surface representation, respectively. Active site domains are also labeled (S1–S3, S1′, S3′). The poses shown were obtained using the Induced Fit Docking module of the Schrödinger Maestro computer package. B, D, and F: The predicted receptor-ligand interactions, with hydrogen bonds represented as black dotted lines. Residues predicted to participate in receptor-ligand interactions, as well as the residues of the catalytic triad, are visualized.