## **Material and Methods**

## Varying Conditions in Enzymatic Assays to Validate the Mechanism of Inhibition

The cysteine protease cruzain was expressed and purified according to the protocol described by Ferreira et al.<sup>26</sup> NSC61610 (Figure S1) was obtained from the National Cancer Institute's Developmental Therapeutics Program and serially diluted in DMSO to create final assay concentrations (FAC) of 1,000  $\mu$ M to 0.1  $\mu$ M. Enzyme inhibition assays, performed using a modified protocol from Ferreira et al.,<sup>26</sup> were used to determine the IC<sub>50</sub> value and the mode of NSC61610 inhibition. Initial assay conditions included 100 mM sodium acetate buffer at pH 5.5, 5 mM DTT as the reducing agent, and Triton X-100 (0.02%) as the detergent. Cruzain (0.4 nM FAC) enzymatic activity was initiated with the addition of the fluorogenic substrate Z-Phe-Argaminomethylcoumarin (Z-FR-AMC, 2.5  $\mu$ M FAC) following a five minute incubation with NSC61610.

To investigate detergent-dependent sensitivity, the IC<sub>50</sub> values of NSC61610 with and without Triton X-100 (0.02% and 0%) were compared. To investigate reducing-agent dependency and check for compound redox cycling that might lead to false positives,<sup>27</sup> the IC<sub>50</sub> value of NSC61610 with  $\beta$ -mercaptoethanol (BME) in place of DTT was likewise calculated.

Cruzain inhibition by NSC61610 was evaluated by measuring the increase in fluorescence (excitation wavelength = 355 nm, emission wavelength = 460 nm) for five minutes in a microtiter plate spectrofluorimeter (Molecular Devices, FlexStation). Percent inhibition was determined from initial velocity with SoftMax Pro 5.4, and dose response curves were created in Prism 4 (GraphPad, San Diego, USA) for IC<sub>50</sub>

calculations. Each curve included eight different concentrations of NSC61610. The assay was performed twice for each detergent/reducing-agent combination.

## **Results/Discussion**

## Biases in Docking Protocols: An Illustrative Example

To see if the performance of the Vina-Vina, Vina-NN1, and Vina-NN2 scoring functions could be further improved, we sought to identify possible biases in these scoring functions that could potentially be subjected to systematic correction. In harmony with what others have found, our own research does suggest significant bias.<sup>8, 14-17, 36</sup> For example, in our own virtual-screening projects, we have consistently seen certain compounds appear near the top of our ranked lists, independent of the target being studied.

Indeed, it was the frequent appearance in several of our own virtual screens of NSC61610, a compound provided by the National Cancer Institute, that prompted us to explore docking bias in the first place. When the NCI compounds and known DUD inhibitors were ranked by their average Vina-Vina ranking over all forty DUD receptors, NSC61610 was 3<sup>rd</sup> out of 4,795 compounds. A PubChem search suggests that NSC61610 modulates targets ranging from Shiga toxin to SUMO1-mediated protein-protein interactions. In one of our own recent projects, NSC61610 also turned up as a predicted inhibitor of cruzain, a therapeutic target in the fight against Chagas' disease.<sup>37</sup> The predicted inhibitors identified in this project were found by consensus scoring; compounds were docked with CDOCKER (Accelrys) and reevaluated using LigScore1, LigScore2,<sup>38</sup> PLP1, PLP2,<sup>39</sup> PMF,<sup>40</sup> and PMF04.<sup>41</sup>

When NSC61610 was experimentally tested in the presence of Triton X-100 and DTT to verify cruzain inhibition, it was found to be potent (Table S1). This was initially surprising given its observed promiscuity, leading us to suspect inhibition by non-specific self-association into colloidal aggregates. A detergent-based assay<sup>26, 42</sup> demonstrated attenuated inhibition in the presence of detergent (Triton X-100, 0.02%) (Table S1), typical of inhibition by aggregation.<sup>43</sup> A shift in the IC<sub>50</sub> value was also noted when the reducing agent was switched from DTT to BME (see Table S1). This dependency on experimental conditions suggests not only that NSC61610 may form aggregates by self-association and/or behave as a redox-cycling compound in the presence of strong reducing agents, but also that the mechanism of inhibition predicted by the virtual screen, i.e. one compound per active site, is incorrect at the micromolar concentrations assayed (1,000  $\mu$ M to 0.1  $\mu$ M).

**Table S1**: Cruzain inhibition assays. The measured  $IC_{50}$  values of NSC61610 vary with the use of detergent (Triton X-100) and according to the reducing agent (DTT vs. BME), suggesting that inhibition occurs via a non-specific perhaps aggregation-based mechanism and/or that NSC61610 is a redox cycling compound.

Triton X-100 (0.02%)	+	-	+	+
No Detergent	-	+	-	-
DTT (10mM)	+	+	+	-
BME (10mM)	-	-	-	+
IC <sub>50</sub> [µM]	9.56	1.39	12.7	37.5

Figure S1: The molecular structure of NSC61610.

`N´ H