

# De Novo Ligand Design to Partially Flexible Active Sites: Application of the ReFlex Algorithm to Carboxypeptidase A, Acetylcholinesterase, and the Estrogen Receptor

Stuart Firth-Clark,\* Stewart B. Kirton,<sup>†</sup> Henriëtte M. G. Willems,<sup>‡</sup> and Anthony Williams

De Novo Pharmaceuticals Ltd., Compass House, Vision Park, Chivers Way, Histon, Cambridge, United Kingdom CB24 9ZR

Received August 1, 2007

Reflex is a recent algorithm in the de novo ligand design software, SkelGen, that allows the flexibility of amino acid side chains in a protein to be taken into account during the drug-design process. In this paper the impact of flexibility on the solutions generated by the de novo design algorithm, when applied to carboxypeptidase A, acetylcholinesterase, and the estrogen receptor (ER), is investigated. The results for each of the targets indicate that when allowing side-chain movement in the active site, solutions are generated that were not accessible from the multiple static protein conformations available for these targets. Furthermore, an analysis of structures generated in a flexible versus a static ER active site suggests that these additional solutions are not merely noise but contain many interesting chemotypes.

## INTRODUCTION

De novo and fragment-based drug design set out to exploit knowledge of the architecture and chemistry of a target protein in order to optimize protein–ligand affinity. Knowledge of the differences in architecture and chemistry between the target and homologous proteins can be used to promote selectivity. The vast majority of these investigations utilize experimentally determined protein structures and/or ligands that display a significant level of affinity for the intended target.<sup>1–4</sup>

Several different methods have been used to elucidate protein structures. These include X-ray crystallography,<sup>5</sup> nuclear magnetic resonance (NMR) spectroscopy,<sup>6</sup> and electron microscopy.<sup>7</sup> Of these techniques, X-ray crystallography is by far the most popular due to its applicability to the widest range of proteins. In February 2006, there were 32 207 structures of proteins in the Brookhaven Protein DataBank (PDB),<sup>8</sup> of which 27 616 (86%) were solved using X-ray crystallography. Hence, it is X-ray crystal structures that are most often exploited in fragment-based and de novo drug design.

The major disadvantage of X-ray crystallography is that it provides a static, low-energy ‘snapshot’ representation of the protein structure,<sup>9</sup> which is in direct contrast to the dynamic nature of the protein in vivo. As such, many proteins have multiple entries in the PDB often including their apo (ligand-free) and holo (ligand-bound) forms. The changes in protein conformation between these static snapshots provide an insight into the inherently flexible nature of the protein, i.e., how conformations shift, architectures change, and active site environments alter depending upon the presence or absence of a ligand.

As proteins are dynamic by nature, protein models used for drug design would ideally incorporate protein flexibility. However, few experiments appear in the literature where the issue of drug design to flexible sites has been addressed.<sup>10–12</sup> This is in direct contrast to docking of known ligands into flexible protein sites (flexible docking), which is a well represented area of research.<sup>13–18</sup> Advantages of incorporating protein flexibility in the design process are numerous and include enhanced ligand diversity, improved sampling of potential ligand binding modes, and the exploration of novel chemical entities across a diverse range of protein conformations.

Reflex is an algorithm that enables partial flexibility of the active site during the drug-design process and is used as an addition to the de novo ligand design software, SkelGen.<sup>19</sup> Reflex has previously been validated using human collagenase MMP-1 and acetylcholinesterase.<sup>20,21</sup> It has been shown that neglecting to model partial protein flexibility in both docking and de novo ligand design to these targets has a detrimental effect on diversity of structures generated and accuracy of docking, due to an inherent exclusion of important regions of chemical space.

A detailed description of how the Reflex algorithm works is complex and beyond the scope of this article. However, a rigorous analysis of the algorithm is presented elsewhere.<sup>20,21</sup> In brief, Reflex functions as follows. Amino acid side-chain conformations, with the exception of alanine and glycine, which have no flexible side chains, can be represented by either rotamers from rotamer libraries or calculated  $\chi$  torsion angles. These angles, or rotamers, are then randomly modified to obtain the best fit with the protein and the ligand. Two recent and extensively validated rotamer libraries can be used to direct the algorithm.<sup>22,23</sup> Both experiments driven using the rotamer libraries, and those dependent upon random modification of the  $\chi$  torsion angle, only retain rotamers that do not clash with the ligand, backbone atoms, or nonflexible amino acid side chains as the algorithm progresses. It is important to emphasize that Reflex is currently only able to

\* Corresponding author phone: +44 (0)1223 238000; e-mail stuart.firth-clark@denovopharma.com.

<sup>†</sup> Present address: NCE Discovery, 418 Cambridge Science Park, Milton Road, Cambridge CB4 0PZ, U.K.

<sup>‡</sup> Present address: GlaxoSmithKline, Gunnels Wood Road, Stevenage, Herts SG1 2NY, U.K.

model side-chain flexibility (although we are aware that small C $\alpha$  shifts can have a major effect on the positions that side chains can explore via rotations around chi angles). Therefore, any test sets and targets should be chosen with these caveats in mind.

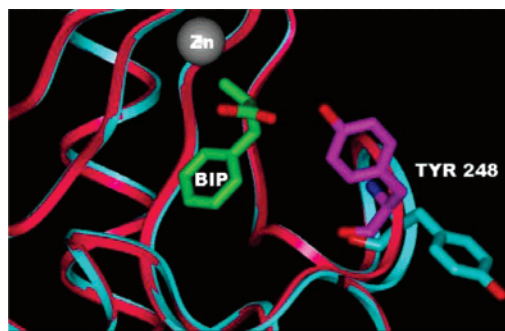
In this investigation we aim to broaden the scope of the initial Reflex validation, showing its applicability to a wider range of targets, and to analyze the impact of incorporating treatments of protein flexibility on de novo design by applying the algorithm to three different proteins: carboxypeptidase A, acetylcholinesterase, and the estrogen receptor. Analyzing the output of a de novo design experiment to a flexible site is more complex than analyzing output from a static site because in the first case the protein conformation has to be taken into account when scoring the solutions. We investigate for three examples whether this increased complexity is offset by extra diversity in the output. Another potential drawback of allowing active site flexibility in de novo design is that it may generate a much larger proportion of inactives than additional actives. Using the estrogen receptor as a target, we also investigate if the additional solutions obtained in the flexible site are likely to be of interest or merely noise.

## RESULTS AND DISCUSSION

**A. Carboxypeptidase A as a Validation Target.** A large amount of prior research has focused on identification of flexible proteins for the development and validation of flexible docking algorithms. We analyzed the ten protein families investigated during the development of the flexible docking algorithm FlexE<sup>24</sup> to ascertain if they would be suitable candidates for the validation of Reflex. We found carboxypeptidase A to be an excellent target for a validation study of Reflex as it shows significant ( $>2$  Å all atom RMSD) movement of the side chain(s) of at least one active site residue, without significant concomitant movement of the protein backbone. The carboxypeptidases are significant as they are involved in the heat shock response associated with infection in warm-blooded animals and hydrolysis of  $\beta$ -lactam antibiotics such as penicillin. The study of these proteins is therefore important in developing an understanding of how to combat virulence.<sup>25</sup>

Sixteen carboxypeptidase A X-ray crystal structures were investigated: 5CPA,<sup>26</sup> 1ARL,<sup>27</sup> 1YME,<sup>27</sup> 3CPA,<sup>28</sup> 4CPA,<sup>29</sup> 6CPA,<sup>30</sup> 7CPA,<sup>31</sup> 8CPA,<sup>31</sup> 1BAV\_A,<sup>32</sup> 1BAV\_B,<sup>32</sup> 1BAV\_C,<sup>32</sup> 1BAV\_D,<sup>32</sup> 1CBX,<sup>33</sup> 1CPS,<sup>34</sup> 2CTB,<sup>35</sup> and 2CTC.<sup>35</sup> (1BAV is a tetrameric protein structure; for the purpose of this investigation each monomer of 1BAV is classed as a distinct structure (A-D)). The relative side-chain orientations of the twenty-nine amino acids highlighted as flexible by the initial research,<sup>24</sup> His 69, Arg 71, Arg 127, Arg 145, Ser 162 – Thr 164, Ser 194 – Ser 199, Leu 203, Tyr 206, Tyr 208, Ile 243, Ile 247 – Ile 255, Thr 268, Glu 270, and Thr 274, were examined. Most of these residues display only a small change in position across the crystal structures and were therefore not included in this investigation.

One residue however, Tyr 248, appears to adopt two distinct conformations dependent upon the presence of a ligand in the carboxypeptidase A active site (Figure 1). When a ligand is bound in the active site (e.g., 1BAV\_C) the Tyr 248 side chain is oriented inward toward the cavity with the



**Figure 1.** Comparison between ligand free (1YME, cyan) and ligand bound (1BAV\_C, magenta) carboxypeptidase A, complexed with BIP, illustrating the different rotameric positions for Tyr 248.

phenolic oxygen atom pointing at the ligand. In the ligand-free protein (e.g., 1YME) the aromatic ring and phenolic oxygen of the Tyr 248 side chain point away from the active site toward solvent. Such a distinct change in side-chain conformation is amenable to being modeled in Reflex. Also, there is little backbone movement upon the change in side-chain orientation for Tyr 248. Hence, carboxypeptidase A could be used as a validation target for Reflex, focusing upon the movement of the Tyr 248 residue. The aim of this relatively simple validation is to assess whether or not Reflex is capable of modeling a single side-chain movement successfully for this target, and hence whether it will be of use in further studies which have more complex systems with a greater number of flexible residues.

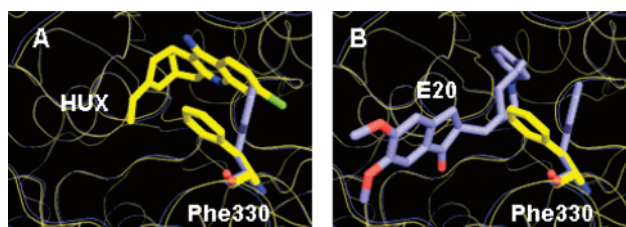
**Benchmarking of ScreenScore.** If it is assumed that the conformation of Tyr 248 in carboxypeptidase A changes upon binding of a ligand because there is an energetic reward for doing so, it is important to illustrate that the scoring function in Reflex is capable of modeling this. Hence, an experiment to compare ScreenScore values returned for the highest resolution ligand-bound structure (1BAV\_C) and the highest resolution ligand free structure (1YME) containing a transposed ligand in the active site was carried out. It was expected that the holo Tyr 248 ‘in’ conformation (1BAV\_C) would achieve a better (more negative) ScreenScore than the ligand transposed into the apo Tyr 248 ‘out’ conformation (1YME). The ScreenScore value of the ligand-bound crystal structure (1BAV\_C) was found to be  $-15.97$  units. The 2-benzyl-1-iodopropanoic acid (BIP) ligand present in 1BAV\_C was then transposed directly into the 1YME active site prior to a score being calculated. Having ensured that this transposition did not result in any ligand-protein clash, the ScreenScore for 1YME with BIP transposed into the active site was obtained, returning a value of  $-6.97$  units. As the ScreenScore returned for the native ligand bound 1BAV\_C structure is approximately double that of the ligand transposed 1YME structure we can have confidence that Reflex recognizes that the true ligand-bound protein conformation for carboxypeptidase A is favored over a spurious protein conformation complexed with the same ligand. This reassures us that at the simplest level ScreenScore is able to attribute higher scores to crystallographically observed protein conformations over alternatives.

**Generating the Ligand-Bound Protein Conformation from a Ligand Free Conformation.** Manual rotation of the Tyr 248 side chain in the apo (1YME) crystal structure, to yield a conformation as near as possible to the Tyr 248

conformation in the holo (1BAV\_C) crystal structure, gave an all atom RMSD for the side chain of 1.09 Å. Numerous attempts to automatically reproduce an approximation of the ligand-bound protein conformation (1BAV\_C) starting from apo form via the iterative recalculation and modification of  $\chi$  angles for the Tyr 248 side chain proved unsuccessful. It was postulated that this was due to a steric clash between Tyr 248 in the new 'in' position and the side chains of neighboring Ile 247 and Gln 249 residues. By allowing side-chain flexibility for Ile 247 and Gln 249 in addition to Tyr 248 during the experiment, Reflex was able to produce a good approximation of the true ligand-bound protein conformation. An all atom RMSD for the side chains of Ile 247, Tyr 248, and Glu 249 between the ligand-bound crystal structure (1BAV\_C) and the Reflex generated model was 1.12 Å. These results give confidence that Reflex is able to accurately model the most likely conformation of a ligand-bound protein for carboxypeptidase A when seeded with a different protein conformation.

**de novo Ligand Design to Carboxypeptidase A.** SkelGen was used to design structures to the active site of the 1BAV\_C (holo) crystal structure, incorporating the flexibility of Ile 247, Tyr 248, and Gln 249 via the Reflex algorithm. 100 structures were requested, and 94 unique structures were returned, with the remaining six structures being duplicates. These structures were similar to known carboxypeptidase A inhibitors in terms of molecular weight and, in some cases, chemical functionality and had ScreenScore values ranging between -20 and -10 units. As a general strategy, a filter is applied to SkelGen output to ensure the structures returned have ScreenScore values comparable to a known inhibitor in the active site, either extracted from cocrystal data or from docking runs. In this case the values compare favorably to that obtained by BIP in the holo crystal structure (-15.97). To assess the impact of flexibility on compound design to this site, each of the 94 structures generated in the partially flexible protein was rescored in both of the static crystal structures (1BAV\_C and 1YME), using -10 units (the poorer ScreenScore value) as a threshold for success. The rescoring process permitted torsional flexibility of compounds as well as movement of up to 2 Å from the Reflex generated position, thereby allowing each compound to optimize interactions with the protein. It was postulated that if a compound generated in a flexible active site could not achieve the threshold value of -10 units when rescored in both of the static structures (1BAV\_C and 1YME), the solution was produced as a consequence of incorporating flexibility in the active site and is unlikely to have been produced if the de novo ligand design experiment had utilized only the two static structures. 26 of the 94 structures rescored in this manner failed to achieve the threshold ScreenScore value of -10 units in both 1BAV\_C and 1YME despite having achieved this in the flexible site. We believe this illustrates the extent to which the exploration of chemical space can be enhanced by considering inherent protein flexibility.

**B. Acetylcholinesterase.** Previous studies have shown that Reflex is a valuable tool in accurate docking of known inhibitors to the acetylcholinesterase active site,<sup>21</sup> an important therapeutic target due to its involvement in diseases such as myasthenia gravis and Alzheimer's disease.<sup>36</sup> The aim of the study described below is to explore the influence of the



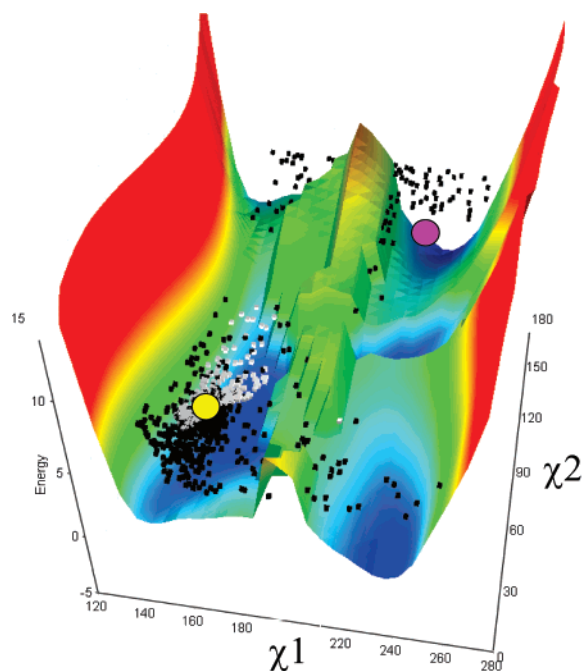
**Figure 2.** Illustration of the flexible nature of Phe 330 in acetylcholinesterase. A) 2-(-)-huprine X (HUX), the complexed ligand in 1E66 (yellow) could not bind in the same orientation in 1EVE (blue) due to a steric clash with the Phe 330 side chain. B) Aricept, the ligand of 1EVE (blue), could not bind in the same orientation to 1E66 (yellow) because of a steric clash with Phe 330.

choice of starting conformation on the accessible protein conformations.

The PDB was mined for acetylcholinesterase structures, and an analysis of side-chain flexibility undertaken. At the time of writing, 30 acetylcholinesterase structures were publicly available for *Torpedo californica*, which has a 55% sequence similarity with the human homologue.<sup>37</sup> After excluding those complexes containing covalently bound inhibitors or mutated forms of the receptor and structures with a resolution of greater than 2.5 Å, 10 structures remained. Each structure was aligned to 1EA5,<sup>38</sup> the apo form of the protein with the highest resolution (1.8 Å). A movement in the backbone atoms at Gly 117 compared to 1EA5 for Gly 117 of 1GPN<sup>40</sup> was observed. Since Reflex is unable to model such a movement, 1GPN was removed from the study. The remaining structures showed good backbone alignment with 1EA5, the RMSDs ranging from 0.17 Å to 0.4 Å. The residue with the largest range of movement in the active site was found to be Phe 330, as exemplified by comparing the crystal structures 1EVE<sup>36</sup> and 1E66<sup>40</sup> (Figure 2). This comparison shows that the ligand Aricept (Figure 2B) could not be bound in the same orientation in the 1E66 crystal structure, as in its native protein 1EVE. Similarly, 2-(-)-huprine X (HUX), the ligand bound in 1E66 (Figure 2A), could not be transposed to 1EVE without incurring severe steric clashes with Phe 330. Movement of the Phe 330 side chain is independent of backbone movement, making acetylcholinesterase an ideal system to investigate with Reflex.

**de novo Design to Acetyl Cholinesterase.** Structure generation experiments were undertaken using two different starting conformations for the Phe 330 side chain. The first rotamer starting conformation was identical to that of Phe 330 in the 1E66 crystal structure (rot1), and the second corresponded to the orientation of the Phe 330 side chain in the 1EVE crystal structure (rot2). 600 structures were requested in both experiments, with the proviso that structures had to achieve a threshold ScreenScore of -45 or better, a value comparable to the scores of HUX and Aricept in their native crystal structures. The 1E66 (rot1) experiment generated 465 structures with a mean ScreenScore value of  $-51.9 \pm 6.0$ . The 1EVE (rot2) experiment generated 529 structures with a similar mean ScreenScore of  $-52.2 \pm 5.7$ . SkelGen failed to generate structures with penalties below the threshold set in 135 attempts in case of the rot1 experiment and in 71 attempts in case of the rot2 experiment. In these failed attempts, high penalties may have arisen





**Figure 3.** Energy landscape for side-chain rotation of Phe 330 in acetylcholinesterase. The energy is expressed as the difference in ScreenScore energy between the reference X-ray crystal structure, 1E66, and the new protein conformation. The conformations of Phe 330 generated from the 1E66 X-ray crystal structure (gray rot1) and the Phe 330 conformations generated from structures from the 1EVE X-ray crystal structure (black rot2) are projected onto the  $x$ - $y$  plane. The reference conformations for Phe 330 in the 1EVE (magenta circle) and 1E66 (yellow circle) crystal structures are shown.

because of intramolecular clashes in the ligand or protein, ligand–receptor clashes, failure to reach the ScreenScore threshold, or any combination of these factors.

Inspection of the Reflex-generated protein conformers from the rot1 experiment showed that there was a limited range of movement for the Phe 330 side chain. A plot of the distribution of  $\chi_1$  and  $\chi_2$  dihedral angles for Phe 330 in the rot1 experiment shows only one outlier (Figure 3);  $\chi_1$  and  $\chi_2$  are within a range of 10 degrees from the starting conformation, which is the global minimum. A similar plot for the distribution of  $\chi_1$  and  $\chi_2$  dihedral angles for Phe 330 in the rot2 experiment shows a far greater range of  $\chi_1$  and  $\chi_2$  dihedral angles, including many in the rot1 minimum, indicating that the algorithm tends to produce the global minimum energy conformation regardless of starting conformation.

This behavior can be explained by considering how the algorithm functions. When generating a structure, the SkelGen algorithm attempts to optimize a 19-parameter objective function, which includes terms describing steric clash and hydrogen-bond interactions. During de novo structure generation, SkelGen has many options available to modify the structure that can potentially impact on the objective function, e.g., the addition of a new fragment, deletion of an existing fragment, and roto-translational movements of fragments. Such changes are only accepted if they decrease the objective function or fall within the Metropolis condition. A structure is output if the final penalty associated with optimization of the objective function is below a predetermined acceptance threshold value. In Reflex mode the objective function is modified to include a parameter representing the combined

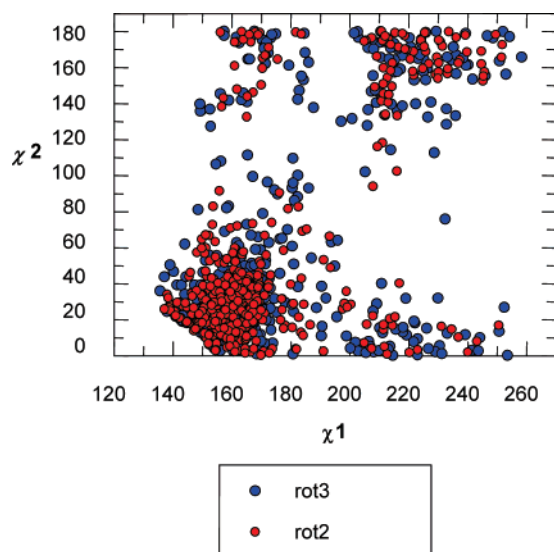
conformational and hydration energy of the flexible residue side chains, calculated with reference to the starting protein conformation (usually the crystal structure). As before, if the change in side-chain orientation results in decreasing the overall penalty associated with the objective function, or falls within the criteria described by the Metropolis condition, the move is accepted. For Phe330 in acetylcholinesterase the hydration energy term is large, resulting in high-energy barriers in the energy landscape. This was not observed for carboxypeptidase A, where the energy landscape between the two conformations of Tyr 248 is flatter.

In the rot1 data set the starting conformation of Phe 330 is at a local minimum on the energy surface. Displacement from this position incurs a large increase in the side-chain conformational + hydration energy term of the objective function, which results in a final value of the objective function which is above the acceptance threshold, so no structure is output. As a consequence, Phe 330 becomes trapped in an energy minimum. In the rot2 experiment the starting conformation of Phe 330 is not in a local minimum, and hence displacement of Phe 330 does not increase the objective function by as much, which makes acceptance of the conformational change to the protein more likely. Figure 3 shows the energy surface describing movement of the Phe 330 residue in terms of its dihedral angles. An energy minimum is found between  $\chi_1 = 120$ – $160^\circ$  and  $\chi_2 = 60$ – $150^\circ$ . This illustrates that selection of an initial protein conformation is an important factor in generation of diverse protein conformations.

The degree of sampling can be controlled by the adjustment of a parameter that sets an energy threshold above the reference protein crystal structure. Increasing this threshold allows side-chain conformational changes to be more easily accepted, and hence chemical structures are generated in a wider range of protein conformations. In the experiment described above this energy window was set to zero, and because rot1 was a low-energy conformation it had very little room to maneuver. By increasing the energy threshold for the 1E66 structure to that of 1EVE, Reflex is able to produce a very similar distribution of active site conformational space from the 1E66 Phe 330 starting geometry as was generated in the rot2 experiment. We have labeled this experiment the rot3 data set. Figure 4 shows the similar distribution of the  $\chi$  angles for both the rot2 and rot3 data sets. This demonstrates that increasing the energy threshold overcomes the problem of the initial protein conformation being in a local or global minimum.

All of the structures generated from the rot1 and rot2 experiments were rescored in the static 1EVE and 1E66 structures, as described for carboxypeptidase A. The number of structures unique to Reflex was defined as those structures that do not achieve the ScreenScore threshold value in either of the static crystal structures (1E66 or 1EVE). This corresponds to 8.5% (85) of the total structures generated from both runs, 7.5% (35) from the rot1 experiment and 9.5% (50) from the rot2 experiment, thus confirming that runs starting from the different static structures yield compounds that are only obtained when Reflex is used.

**C. The Estrogen Receptor (ER).** SkelGen has previously been used to design compounds with affinity for ER,<sup>41</sup> an important target in combating ailments such as breast cancer, osteoporosis, and cardiovascular disease.<sup>42,43</sup> During the



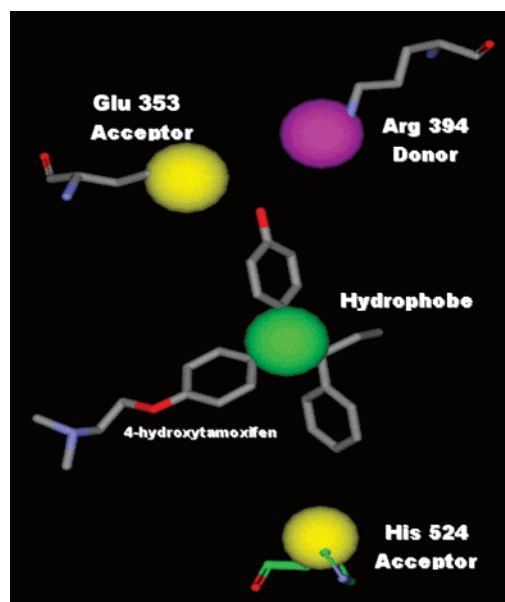
**Figure 4.** Side-chain distribution for structures generated with 1e66 starting conformation with a reference energy window of 7.1 screenscore units above the crystal structure reference energy (138.0), labeled as rot3. Rot2 side-chain distributions are shown for comparison.

initial validation studies of SkelGen seven different ER crystal structures were used to generate a variety of potential ER ligands.<sup>41</sup> It became apparent that potential ligands produced by one crystal structure would not have been produced if an alternative crystal structure had been used, and it was hypothesized that this was due to differences in conformation of side chains between structures.

The five active ER compounds from the original SkelGen investigation (**1–5** Table 1) were used as a small data set to investigate this hypothesis, using the highest resolution ER crystal structure available (PDB accession code 3ERT)<sup>44</sup> as the starting point.

Two questions needed to be addressed: 1) which of the active compounds (**1–5**) could not fit into the 3ERT crystal structure and 2) which residues in the ER active site have flexible side chains.

**Active Compounds not Accessible from the 3ERT Crystal Structure.** Two experiments were carried out to determine if the ER compounds **1–5** could achieve a nominal ScreenScore value of  $-25.8$  units (equal to the score of the cocrystallised ligand in 3ERT) in the static active-site of



**Figure 5.** The pharmacophore used to guide ligand generation in the initial ER study. Donor atoms in the ligand must satisfy (directed) acceptor points on Glu-353 and His-524 (yellow), while an acceptor atom is required to interact with the (directed) donor of Arg-394 (magenta). A hydrophobic atom must also be present at the ligand T-junction (green).

3ERT, irrespective of the identity of the crystal structure in which they were originally generated. Compound **1** was initially generated in 3ERT, **2** and **4** were derived from the 1QKT<sup>45</sup> crystal structure, compound **3** was produced using 1ERR,<sup>46</sup> and compound **5** resulted from experiments with 1GWQ.<sup>47</sup>

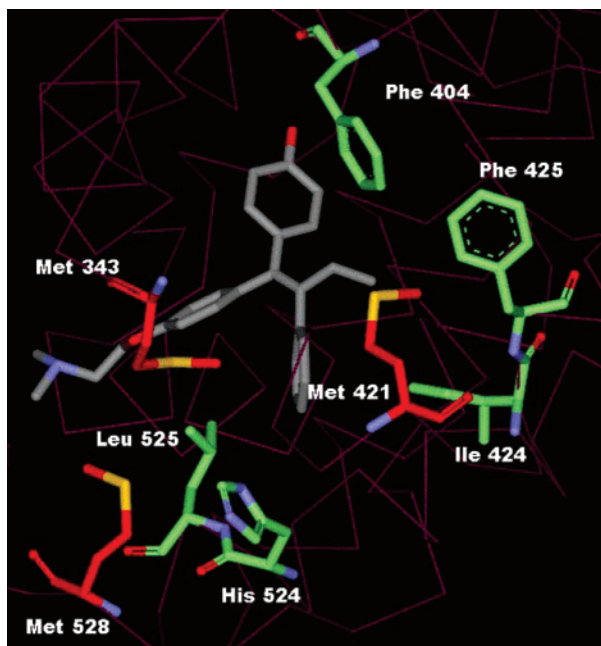
In the first experiment, using the site-point file (Figure 5) as a guide, the ER compounds were rescored as rigid bodies in the exact orientations generated by the parent crystal structure but transposed into the 3ERT active site. Compounds **3**, **4**, and **5** failed to achieve the threshold ScreenScore value of  $-25.8$  units when rescored using these criteria.

In the second experiment, the ER compounds were allowed roto-translational and conformational movement within the active site, in an effort to help optimize protein–ligand interactions and achieve the threshold ScreenScore value. Despite this, structures **3** and **4** both still failed to achieve the threshold ScreenScore value of  $-25.8$  units in the static

**Table 1.** Five Active ER Compounds Produced in the Initial SkelGen Validation Exercise<sup>a</sup>

<b>1 (3ERT)</b>	<b>2 (1QKT)</b>	<b>3 (1ERR)</b>
<b>4 (1QKT)</b>	<b>5 (1GWQ)</b>	

<sup>a</sup> The ER crystal structure that each compound was generated in is given in parentheses.

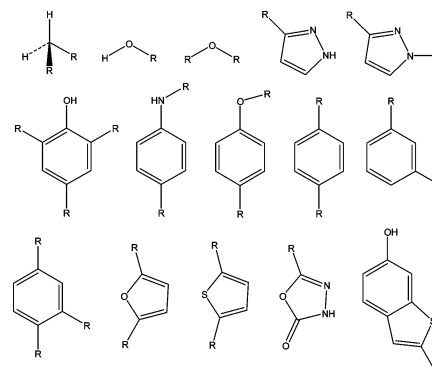


**Figure 6.** Residues with side chains considered flexible in the Reflex studies of the ER structure 3ERT. Ideal residues are highlighted with green carbon atoms (Phe 404, Ile 424, Phe 425, His 524, and Leu 525). Nonideal residues, included because of the impact they potentially have on free movement of ideal residues, have red carbon atoms (Met 343, Met 421, and Met 528). The position of 4-hydroxytamoxifen (gray), the ligand complexed with ER in the 3ERT crystal structure, is shown in relation to these residues.

3ERT active site. Hence, had the static 3ERT crystal structure been the only structure used in the initial design experiment for ER compounds, it is unlikely that compounds **3** and **4** would have been produced.

**Flexible Residues in the ER Active Site.** It was obvious from superposition of the ER crystal structures used in the original study (3ERT, 1L2I, 1QKT, 1GWR, 1GWQ, 1ERR, and 1PCG)<sup>44–49</sup> that several of the active site residues have flexible side chains. Both backbone and side-chain displacements for the active-site residues were assessed. Eight residues (Thr 347, Asp 351, Leu 354, Phe 404, Ile 424, Phe 425, His 524, Leu 525) were found to show a consistently higher displacement of their side chains compared to their backbones and are potentially interesting for use in a Reflex experiment. These eight residues cluster in three areas of the ER active site. Phe 404, Ile 424, and Phe 425 occupy an area opposite the antagonist channel in a ‘buried’ region of the active site. Examination of the crystal structures reveals that these residues are able to move and open up a hydrophobic pocket in the active site that some ligands are able to exploit. As such, they are ideal for inclusion in a Reflex run. However, Met 421, which displays a significant degree of backbone displacement in addition to side-chain flexibility, impinges upon these residues. Hence, Met 421 is also included as a flexible residue.

Two more residues with large side-chain displacements, His 524 and Leu 525, are found at the ‘bottom’ of the active site, with His 524 often interacting with ligands. Two methionine residues (Met 343 and Met 528) initially deemed unsuitable for inclusion in the experiment due to large movements in the position of their backbone atoms appear to impact upon the local environment of these residues and



**Figure 7.** The 15 fragments used in generation of ER compounds (including compounds **1–5**) in the near-completion design runs to the static and flexible active site of the 3ERT crystal structure. R refers to the points at which connection to a fragment is allowed during the SkelGen run.

could potentially restrict movement of the His 524 and Leu 525 side chains. Met 343 and Met 528 were therefore included as flexible residues in the Reflex runs.

Finally, Thr 347, Asp 351, and Leu 354 line the top of the ‘antagonist channel’. According to the initial de novo design experiment, none of the novel ER compounds (**1–5**) are believed to bind in this region.<sup>41</sup> For this reason, they were not included in this investigation.

Therefore, a total of eight residues were allowed to flex for the purpose of this Reflex investigation: Phe 404, Ile 424, Phe 425, His 524, and Leu 525 and three methionine residues, Met 343, Met 421, and Met 528 (Figure 6).

**Structure Regeneration in a Single, Flexible ER Crystal Structure.** The aim of this experiment was to show that it was possible to generate all of the novel ligands (**1–5**) that were originally produced using a number of selected ER crystal structures using only the highest resolution ER crystal structure (3ERT) by incorporating flexibility in the side chains of eight active site residues, Phe 404, Ile 424, Phe 425, His 524, Leu 525, Met 421, Met 343, and Met 528, as an example of the power of Reflex. In actuality it would be unrealistic to use a single ER crystal structure to attempt to generate compounds that have been derived from the distinct agonist and antagonist structures, given that there is substantial active site rearrangement upon binding of the different ligand types and high quality structural data available for both conformations. As such this is merely a validation and an exercise in how well Reflex can cope with such a challenging situation. A pharmacophore restraint was used to guide the design (Figure 5). This required any ligand produced to form hydrogen bonds with Glu 353, Arg 394, and His 524 and have a hydrophobic atom occupying the position between the ‘main’ binding site and the ‘antagonist channel’ leading to the I2 helix in the buried region of the protein. To increase the probability of regenerating compounds **1–5**, a streamlined fragment set was used. This fragment library contained the 15 fragments that SkelGen had used to generate compounds **1–5** (Figure 7). Constraints on the minimum and maximum numbers of fragments used to generate the structures were implemented. Restricting the input in this manner allows the strategy to be run to near completion. Why is the strategy run to ‘near completion’? SkelGen uses a stochastic process to generate structures, thereby generating a large number of unique compounds at

**Table 2.** Number of Copies of Each of the Novel ER Compounds (1–5) Generated in the Reflex Runs Incorporating Flexible Residues in 3ERT and a Restricted Fragment Library Input of 15 Fragments<sup>a</sup>

molecule	no. of times generated
<b>1</b>	185
<b>2</b>	1503
<b>3</b>	61
<b>4</b>	2
<b>5</b>	2261

<sup>a</sup> The experiment generated 25 000 structures. Italicized compounds are those that did not achieve the threshold ScreenScore value of  $-25.8$  units in the static 3ERT protein.


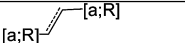

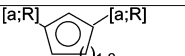
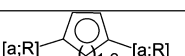
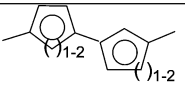
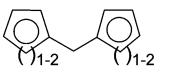
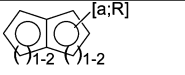
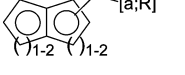
the start of a run. But as time continues, duplicates of structures which have already been generated begin to appear, while the number of “new structures” generated is reduced. After infinite time, SkelGen will have generated all possible structures that can be generated for a particular strategy. Running SkelGen to completion is therefore not realistically possible, and a definition of “near completion” has to be considered. For this work we have defined “near completion” as the point at which three or less novel structures are being produced per 1000. Running the strategy to near-completion gives a true reflection of the differences between the structures available to the static and flexible sites, with the caveat that we assume all extra data in the Reflex run is generated exclusively under Reflex conditions.

Near-completion for the static site was achieved after 5000 structures had been generated, with 81 unique structures produced in total. After generating 25 000 structures under Reflex conditions 475 unique structures had been generated, but near-completion had still not been achieved at this stage. This is six times the number of structures produced compared to results from the static site. It could be argued that this increase is due to the Reflex experiment being allowed to run for a longer time than the static experiment. However, if only the first 5000 structures generated in the Reflex run are considered, 295 unique structures are still observed—3.5 times more structures than the equivalent run in the static site. It is clear that far more structures are generated under Reflex conditions.

**Compounds not Accessible to the Static ER (3ERT) Active Site.** Previously we have shown that two of the SkelGen generated ER compounds, **3** and **4**, are not able to achieve a nominal ScreenScore threshold value of  $-25.8$  units in the static 3ERT active site. Attempts to regenerate ER ligands **1–5** using a streamlined fragment set, and running to near-completion, with a static 3ERT active site failed to produce compounds **3** and **4**, although compounds **1**, **2**, and **5** were produced.

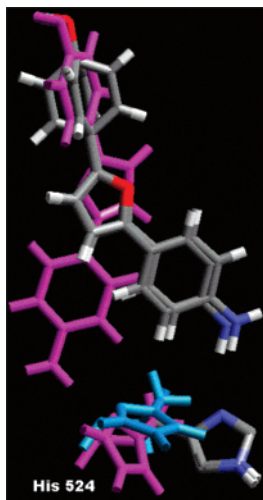
Under Reflex conditions both compounds **3** and **4** (along with the three compounds accessible to the static site: **1**, **2**, and **5**) were generated (Table 2). Compound **3** is generated 61 times, with ScreenScore values around  $-33$  units,

**Table 3.** SMARTS<sup>50</sup> Clustering for the Static and Reflex Runs<sup>a</sup>

Cluster	SMARTS	Depiction	Static Run	Reflex Run	% Static Run	% Reflex Run
2	<chem>[!R][!R]a1aaa(aa1)a2a[a;R1][a;R1]aa2</chem>		0	17	0	6.2
3	<chem>[a;R][!R]~[!R][a;R]</chem>		2	12	2.5	4.4
4	<chem>a1aaa(aa1)a2aa3aaaaa3a2</chem>		13	32	16	11.7
5	<chem>[a;R]-[a;R][a;R][a;R]-[a;R]</chem>		20	84	24.7	30.7
7	<chem>[a;R]-[a;R][a;R][a;R][a;R]-[a;R]</chem>		26	61	32.1	22.3
11	<chem>[!R]-a[a;R1][a;R1]a-[a;R][a;R][a;R]-[!R]</chem>		5	34	6.2	12.4
13	<chem>[a;R1][a;R1][a;R1][a;R1][a;R1]-[!R]-[a;R1][a;R1][a;R1][a;R1][a;R1]</chem>		11	16	13.6	5.8
16	<chem>[a;R1][a;R2]([a;R2])[a;R1].[a;R]-[a;R]</chem>		2	10	2.5	3.6
17	<chem>[R1][a;R2]([a;R2])[R1].[a;R]-[!R]-[a;R]</chem>		2	8	2.5	2.9

<sup>a</sup> 63% of the static generated structures fall in the active classes (4, 5, 7) and 65% of the Reflex generated structures are represented in the active SMART classes. Due to the fact that a limited fragment set was employed for these experiments, only 9 of the original 22 clusters were occupied (including the clusters 4, 5, and 7 that contained the active compounds 1–5 and cluster 2 that was only accessible in the Reflex experiment).





**Figure 8.** Binding modes obtained for compound **4** in the Reflex experiment (gray) compared to the binding mode obtained in the static 1QKT crystal structure (magenta). His 524 conformations in 1QKT (magenta) and 3ERT (cyan) crystal structures are also shown for comparison purposes.

comparable to the best ScreenScore achieved when compound **3** was generated in the 1ERR crystal structure (−36.7 units).

Only two instances of molecule **4** were produced in 25 000 Reflex runs, both in a very different binding mode to that observed for the structure produced in the static 1QKT structure (Figure 8). The different binding orientation is not surprising given the large differences in active site architecture, especially protein backbone movement, between the antagonist structure 3ERT used in the reflex experiments and the agonist 1QKT structure used to generate molecule **4** in the original set of experiments. The different binding modes are thought to arise primarily due to a clash between compound **4** and the position of the side chain of Leu 525. In the 3ERT crystal structure, the side chain of Leu 525 impinges upon the active site to a greater extent than in the 1QKT crystal structure due to a difference in the position of the protein backbone. Reflex is unable to model such a change. This would prevent the ligand orientation observed for the static run being duplicated in the Reflex run, as it would invoke a severe steric clash. It is also possible that the value of the side-chain energy window parameter prevented a 1QKT-like conformation from being accepted by the objective function. Once again, however, the ScreenScore values obtained for the Reflex experiment (ca. −33 units) compare favorably to the values obtained in the static run in the 1QKT structure (−35 units) despite the marked difference in binding orientation.

Thus, it is possible to produce all of the known ER compounds (**1–5**) using a single crystal-structure, if flexibility of the side chains of relevant residues is accounted for and a constrained fragment set is employed. While the experiment was considered successful, using standard SkelGen filtering criteria, which only examines the top ten percent of structures based upon Complementarity Screen Score, would have resulted in three potent compounds being discarded. This is potentially offset by there being additional high ranking active structures in the Reflex generated output, although this can only be verified through synthesis and testing. Earlier work on ER showed that the SkelGen

generated structures could be grouped in 22 classes based on similarity between chemical substructures.<sup>41</sup> Nine of the classes were accessible to the streamlined fragment set, of which three contained compounds that have ER affinity (compounds **1–5**). A comparison between the relative proportion of structures occupying ‘active’ and ‘nonactive’ clusters for static and flexible structure generation experiments gives an indication as to whether or not the additional structures generated by the Reflex algorithm are potentially useful or whether they are just noise. We find that Reflex and static-site experiments yield a similar proportion of structures that occupy the ‘active’ and ‘nonactive’ clusters (63% occupying the ‘active’ clusters for the static investigation and 65% occupying the ‘active’ clusters in the flexible experiment) when using a limited fragment set. This suggests that Reflex generated output is similar to the SkelGen output and that the additional structures generated by Reflex are not merely noise. Therefore Reflex can be a valuable tool for de novo drug design, particularly if only a small number of conformers are available for the target protein.

## CONCLUSIONS

We have demonstrated that, for carboxypeptidase A, Reflex is able to distinguish between the observed ligand-bound protein conformation and an alternative and is able to score the empirical ligand-bound conformation more highly than the alternative. In addition it is also possible to reproduce an accurate representation of the true ligand-bound conformation of carboxypeptidase A using Reflex, when starting from a different conformation of the protein. Also, we have gained some insight into the importance of incorporating inherent protein flexibility in de novo ligand design by illustrating that 28% of the structures generated in a relatively small Reflex run would not have been generated with identical parametrization using only the static crystal structures for this protein family.

When flexible de novo ligand design experiments are carried out in acetylcholinesterase it has been shown that 8.5% of the structures produced are uniquely generated because of the inclusion of a flexibility term for the Phe 330 side chain. This work has also highlighted the dependence of Reflex output on the initial protein conformation and stressed the importance of correctly parametrizing experiments to prevent side-chain conformations becoming trapped in local energy minima.

Extending previous work on the ER protein, we have shown that all five of the previously described active ER compounds **1–5** could be generated from a single, high-resolution crystal structure using Reflex. Our studies have highlighted eight residues in ER that could impact significantly on ligand binding and illustrated that by incorporating side-chain flexibility for these residues we are able to regenerate all of the ER ligands using only one crystal structure as input. By restricting the number of input fragments in the fragment library to fifteen we were able to run the static site and Reflex experiments extensively. This revealed that Reflex runs can generate six times more unique solutions than runs focusing on static active sites, thus allowing a more comprehensive sampling of chemical space. Also, by clustering the Reflex output with regards to chemical substructure using predetermined criteria, we have



demonstrated that there is no significant increase in terms of percentage occupancy in those clusters that are known to contain inactive compounds, when compared with the results from the static experiment.

# ACKNOWLEDGMENT

The authors thank Ian Alberts, Nik Todorov, Bill Harris, Louise Birch, and Philip Dean for useful discussions and advice in both the work leading up to and in the preparation of this document.

# REFERENCES AND NOTES

- (1) Kuntz, I. D. Structure-based Strategies for Drug Design and Discovery. *Science* **1992**, 257, 1078–1082.
- (2) Günther, J.; Bergner, A.; Hendlich, M.; Klebe, G. Utilising Structural Knowledge in Drug-Design Strategies: Applications Using Relibase. *J. Mol. Biol.* **2003**, 326, 621–636.
- (3) Blundell, T. L.; Jhoti, H.; Abell, C. High Throughput Crystallography for Lead Discovery in Drug Design. *Nat. Rev. Drug Discovery* **2002**, 1, 45–54.
- (4) Klebe, G. Recent Developments in Structure-based Drug-design. *J. Mol. Med.* **2000**, 78, 269–281.
- (5) Rhodes, G. *Crystallography Made Clear*, 2nd ed.; Academic Press: San Diego, CA, 2000.
- (6) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
- (7) Stowell, M. H. B.; Miyazawa, A.; Unwin, N. Macromolecular Structure Determination by Electron Microscopy: New Advances and Recent Results. *Curr. Opin. Struct. Biol.* **1998**, 8, 595–600.
- (8) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, 28, 235–242.
- (9) Davis, A. M.; Teague, S. J.; Kleywegt, G. J. Applications and Limitations of X-ray Crystallographic Data in Structure-based Ligand and Drug Design. *Angew. Chem., Int. Ed.* **2003**, 42, 2718–2736.
- (10) Zhu, J.; Fan, H.; Lui, H.; Shi, Y. Structure-based Ligand Design for Flexible Proteins: Application of new F-DycoBlock. *J. Comput.-Aided Mol. Des.* **2001**, 15, 979–996.
- (11) Carlson, H. A.; McCammon, J. A. Accommodating Protein Flexibility in Computational Drug Design. *Mol. Pharmacol.* **2000**, 57, 213–218.
- (12) Anderson, A. C.; O'Neil, R. H.; Surti, T. S.; Stroud, R. M. Approaches to Solving the Rigid Receptor Problem by Identifying a Minimal Set of Flexible Residues During Ligand Docking. *Chem. Biol.* **2001**, 8, 445–457.
- (13) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, 261, 470–489.
- (14) Lorber, D. M.; Schoichet, B. K. Flexible Ligand Docking Using Conformational Ensembles. *Protein. Sci.* **1998**, 7, 938–950.
- (15) Totrov, M.; Abagyan, R. Flexible Protein-ligand Docking by Global Energy Optimization in Internal Coordinates. *Proteins: Struct., Funct., Genet.* **1997**, 29, 215–220.
- (16) Fradera, X.; Knegtel, R. M. A.; Mestres, J. Similarity-driven Flexible Ligand Docking. *Proteins: Struct., Funct., Genet.* **2000**, 40, 623–636.
- (17) Makino, S.; Ewing, T. J. A.; Kuntz, I. D. DREAM++: Flexible Docking Program for Virtual Combinatorial Libraries. *J. Comput.-Aided Mol. Des.* **1999**, 13, 513–532.
- (18) Budin, N.; Majeux, N.; Caffisch, A. Fragment-Based Flexible Ligand Docking by Evolutionary Optimisation. *Biol. Chem.* **2001**, 382, 1365–1372.
- (19) Stahl, M.; Todorov, N. P.; James, T.; Mauser, H.; Böhm, H.-J.; Dean, P. M. A Validation Study on the Practical Use of Automated de novo Design. *J. Comput.-Aided Mol. Des.* **2002**, 16, 459–478.
- (20) Källblad, P.; Todorov, N. P.; Willems, H. M. G.; Alberts, I. L. Receptor Flexibility in the in silico Screening of Reagents in the S1' Pocket of Human Collagenase. *J. Med. Chem.* **2004**, 47, 2761–2767.
- (21) Alberts, I. L.; Todorov, N. P.; Källblad, P.; Dean, P. M. Ligand Docking and Design in a Flexible Receptor Site. *QSAR Comb. Sci.* **2005**, 24, 503–507.
- (22) Dunbrack, R. L., Jr.; Karplus, M. Backbone-dependent Rotamer Library for Proteins. Application to Side-chain Prediction. *J. Mol. Biol.* **1993**, 230, 543–574.
- (23) Xiang, Z.; Honig, B. Extending the Accuracy Limits of Prediction for Side-chain Conformations. *J. Mol. Biol.* **2001**, 311, 421–430.
- (24) Claussen, H.; Buning, C.; Rarey, M.; Lengauer, T. FlexE: Efficient Molecular Docking Considering Protein Structure Variations. *J. Mol. Biol.* **2001**, 308, 377–395.
- (25) Christianson, D. W.; Lipscomb, W. N. Carboxypeptidase A. *Acc. Chem. Res.* **1989**, 22, 62–69.
- (26) Rees, D. C.; Lewis, M.; Lipscomb, W. N. Refined Crystal Structure of Carboxypeptidase A at 1.54 Å Resolution. *J. Mol. Biol.* **1983**, 168, 367–387.
- (27) Greenblatt, H. M.; Feinberg, H.; Tucker, P. A.; Shoham, G. Carboxypeptidase A: Native, Zinc-Removed and Mercury-replaced Forms. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, 54, 289–305.
- (28) Christianson, D. W.; Lipscomb, W. N. X-ray Crystallographic Investigation of Substrate Binding to Carboxypeptidase A at Subzero Temperature. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, 83, 7568–7572.
- (29) Rees, D. C.; Lipscomb, W. N. Refined Crystal structure of the Potato Inhibitor Complex of Carboxypeptidase A at 2.5 Å Resolution. *J. Mol. Biol.* **1982**, 160, 475–498.
- (30) Kim, H.; Lipscomb, W. N. Crystal Structure of the Complex of Carboxypeptidase A with a Strongly Bound Phosphonate in a New Crystalline Form: Comparison with Structures of Other Complexes. *Biochemistry* **1990**, 29, 5546–5555.
- (31) Kim, H.; Lipscomb, W. N. Comparison of the Structures of Three Carboxypeptidase A-Phosphonate Complexes Determined by X-Ray Crystallography. *Biochemistry* **1991**, 30, 8171–8180.
- (32) Massova, I.; Martin, P.; deMel, S.; Tanaka, Y.; Edwards, B.; Mobashery, S. Crystallographic and Computational Insight on the Mechanism of Zinc-ion-dependent Inactivation of Carboxypeptidase A by 2-benzyl-3-iodopropanonate. *J. Am. Chem. Soc.* **1996**, 118, 12479–12480.
- (33) Mangani, S.; Carloni, P.; Orioli, P. Crystal Structure of the Complex Between Carboxypeptidase A and the Biproduct Analog Inhibitor L-benzylsuccinate at 2.0 Å Resolution. *J. Mol. Biol.* **1992**, 223, 573–578.
- (34) Cappalonga, A. M.; Alexander, R. S.; Christianson, D. W. Structural Comparison of Sulfodiimine and Sulfonamide Inhibitors in Their Complexes with Zinc Enzymes. *J. Biol. Chem.* **1992**, 267, 19192–19197.
- (35) Teplyakov, A.; Wilson, K. S.; Orioli, P.; Mangani, S. High-Resolution Structure of the Complex Between Carboxypeptidase A and L-Phenyl Lactate. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1993**, 49, 534–540.
- (36) Kryger, G.; Silman, I.; Sussman, J. L. Structure of Acetylcholinesterase Complexed with E2020 (Aricept): Implications for the Design of New anti-Alzheimer Drugs. *Structure* **1999**, 7, 297–307.
- (37) Bartolucci, C.; Perola, E.; Pilger, C.; Fels, G.; Lamba, D. Three-dimensional Structure of a Complex of Galanthamine (Nivalin) with Acetylcholinesterase from Torpedo Californica: Implications for the Design of New anti-Alzheimer Drugs. *Proteins: Struct., Funct., Genet.* **2001**, 42, 182–191.
- (38) Harel, M.; Weik, M.; Silman, I.; Sussman, J. L. Native Acetylcholinesterase (E. C. 3.1.1.7) From Torpedo Californica at 1.8 Å Resolution. To be published.
- (39) Dvir, H.; Jiang, H. L.; Wong, D. M.; Harel, M.; Chetrit, M.; He, X. C.; Jin, G. Y.; Yu, G. L.; Tang, X. C.; Silman, I.; Bai, D. L.; Sussman, J. L. X-ray Structures of Torpedo Californica Acetylcholinesterase Complexed with (+)-Huperzine A and (–)-Huperzine B: Structural Evidence for an Active Site Rearrangement. *Biochemistry* **2002**, 41, 10810–10818.
- (40) Dvir, H.; Wong, D. M.; Harel, M.; Barril, X.; Orozco, M.; Luque, F. J.; Muñoz-Torrero, D.; Camps, P.; Rosenberry, T. L.; Silman, I.; Sussman, J. L. 3D Structure of Torpedo Californica Acetylcholinesterase Complexed with Huprine X at 2.1 Å Resolution: Kinetic and Molecular Dynamic Correlates. *Biochemistry* **2002**, 41, 2970–2981.
- (41) Firth-Clark, S.; Willems, H. M. G.; Williams, A.; Harris, W. Generation and Selection of Novel Estrogen Receptor Ligands. *J. Chem. Inf. Model.* **2006**, 46, 642–647.
- (42) Meegan, M. J.; Lloyd, D. G. Advances in the Science of Estrogen Receptor Modulation. *Curr. Med. Chem.* **2003**, 10, 181–210.
- (43) Sato, M.; Grese, T. A.; Dodge, J. A.; Bryant, H. U.; Turner, C. H. Emerging Therapies for the Prevention or Treatment of Postmenopausal Osteoporosis. *J. Med. Chem.* **1999**, 42, 1–24.
- (44) Shiau, A. K.; Barstad, D.; Loria, P. M.; Cheng, L.; Kushner, P. J.; Agard, D. A.; Greene, G. L. The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of this Interaction by Tamoxifen. *Cell* **1998**, 95, 927–937.
- (45) Gangloff, M.; Ruff, M.; Eiler, S.; Duclaud, S.; Wurtz, J. M.; Moras, D. Crystal Structure of a Mutant hERα Ligand-Binding Domain Reveals Key Structural Features for the Mechanism of Partial Agonism. *J. Biol. Chem.* **2001**, 276, 15059–15065.
- (46) Brzozowski, A. M.; Pike, A. C. W.; Dauter, Z.; Hubbard, R. E.; Bonn, T.; Engström, O.; Ohman, L.; Greene, G. L.; Gustafsson, J.-Å.; Carlquist, M. Molecular Basis of Agonism and Antagonism in the Oestrogen Receptor. *Nature* **1997**, 389, 753–758.

- (47) Wärnmark, A.; Treuter, E.; Gustafsson, J.-Å.; Hubbard, R. E.; Brzozowski, A. M.; Pike, A. C. W. Interaction of Transcriptional Intermediary Factor 2 Nuclear Receptor Box Peptides with the Coactivator Binding site of Estrogen Receptor  $\alpha$ . *J. Biol. Chem.* **2002**, *277*, 21862–21868.
- (48) Shiau, A. K.; Barstad, D.; Radek, J. T.; Meyers, M. J.; Nettles, K. W.; Katzenellenbogen, B. S.; Katzenellenbogen, J. A.; Agard, D. A.; Greene, G. L. Structural Characterization of a Subtype-Selective Ligand Reveals a Novel Mode of Estrogen Receptor Antagonism. *Nat. Struct. Biol.* **2002**, *9*, 359–364.
- (49) Leduc, A.-M.; Trent, J. O.; Wittliff, J. L.; Bramlett, K. S.; Briggs, S. L.; Chirgadze, N. Y.; Wang, Y.; Burris, T. P.; Spatola, A. F. Helix-stabilised Cyclic Peptides as Selective Inhibitors of Steroid Receptor-Coactivator Interactions. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11273–11278.
- (50) James, C. A.; Weininger, D.; Delany J. *Daylight Theory Manual*; Daylight Chemical Information Systems Inc. <http://www.daylight.com/dayhtml/doc/theory/theory.toc.html> (accessed April 2005).

CI700282U