



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Leveraging Structural Information for the Discovery of New Drugs - Computational Methods

T. B. Nguyen, S. E. Wong, F. C. Lightstone

March 29, 2011

Methods in Molecular Biology: Structure Based Drug Discovery

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

Leveraging Structural Information for the Discovery of New Drugs – Computational Methods

Toan B. Nguyen, Sergio E. Wong and Felice C. Lightstone

Physical and Life Sciences Directorate

Lawrence Livermore National Laboratory

7000 East Avenue

Livermore, CA 94550

Abstract

Escalating problems with drug resistance continue to compromise the effectiveness of commercial antibiotics, necessitating the search for novel classes of antimicrobial agents. To circumvent problems with resistance, a multi-target single-pharmacophore approach has been employed to discover inhibitors that maintain a balanced activity against multiple target enzymes. In this chapter we examine the application of computational techniques, in particular, structure-based drug design approaches, to design new dual-targeting antibacterial agents against bacterial topoisomerases.

Keywords: Structure-based drug design, Structure-activity relationships, multi-target single pharmacophore, virtual screening, docking, DNA gyrase.

1. Introduction

Almost seventy years ago, penicillin was used to save the life of an infected patient suffering from *streptococcal* sepsis (1). Since then, many new classes of antibiotics have been discovered and developed. However, bacteria, the champions of evolution, have adapted and developed resistance against frontline antibiotics, such as vancomycin, methicillin, fluoroquinolones and macrolides (2-7). A growing clinical concern is that bacteria are becoming increasingly multidrug-resistant where even new antibiotics, such as linezolid (8) and daptomycin (9), are encountering significant resistance. Consequently, the unmet medical need caused by prevailing bacterial drug resistance has renewed interest in the discovery and development of new classes of antibiotics (10) with novel mechanisms of action (11-13). Pharmaceutical researchers are pursuing a multi-target single-pharmacophore approach (14) to fight antibiotic resistance. This new strategy focuses on the design of inhibitors that maintain well-balanced activity against multiple target enzymes from multiple pathways, like bacterial topoisomerase IV and DNA gyrase, dramatically reducing the probability of resistance incidence (15).

The utilization of microbial genomics (16) for target identification, advances in x-ray crystallography for 3D target-ligand interactions, and exploitation of computational approaches for designing and developing (17, 18) novel small molecules should accelerate the discovery of new antibacterial agents with the desired broad spectrum efficacy and the low potential for resistance development.

This chapter discusses the application of computational structure-based drug design methods in the discovery of novel small molecule inhibitors against well-validated targets involved in bacterial DNA replication (19, 20) and cell division; the type IIA topoisomerases, gyrase B

(GyrB) and topoisomerase IV (ParE). The focus will be on specific ligand-protein interactions observed in publically available crystallographic structures of novobiocin and/or adenylyl-imidodiphosphate (ADPNP) bound to the *E. coli* GyrB and ParE subunits as the initial guide for inhibitor discovery and optimization.

2. The design of antibacterial agents that inhibit the function of type IIA topoisomerases - GyrB and/or ParE

2.1 Background

There are many classes of antibiotics that target different aspects of bacterial function. Two important classes of antibiotics, including the coumarins (*e.g.* novobiocin) and fluoroquinolones (*e.g.* ciprofloxacin), target proteins involved in DNA replication, specifically the type IIA family of topoisomerases, DNA gyrase and topoisomerase IV (19, 20). These enzymes are structurally homologous proteins and exist as heterotetramers *in vivo*, comprising 2 GyrA and 2 GyrB subunits for DNA gyrase and 2 ParC and 2 ParE subunits for topoisomerase IV. Type II topoisomerases possess both DNA cleavage and ligation functionalities, and an ATP-dependent clamp that work in a concerted fashion to resolve DNA catenates and supercoils during replication. The GyrA and ParC subunits catalyze DNA cleavage and re-ligation *via* an intermediate where the 5'-ends of the DNA chain are covalently bound to conserved tyrosine residues. The GyrB and ParE subunits are responsible for passage of a separate DNA strand through the cut DNA strands in an ATP-dependent manner. At the cleavage site, DNA becomes single stranded due to a four base-pair staggered break within the binding cavity, allowing the fluoroquinolone inhibitors to bind to the DNA-topoisomerase complex. The drug-DNA-enzyme

ternary complex locks the tetrameric complex in a catalytically non-competent state (21), arresting DNA replication. The accumulation of single- and double-stranded breaks irreversibly damage the bacterial chromosome (22, 23).

Although dual-targeting in principle, fluoroquinolones appear to kill *via* single targets in different bacterial classes. In *E. coli*, a gram-negative bacteria, fluoroquinolones kill mainly *via* inhibition of DNA gyrase, while in gram-positive organisms like *S. aureus*, fluoroquinolones appear to operate primarily *via* inhibition Topoisomerase IV (24). Coumarin antibiotics, on the other hand, which target the ATP-binding domains of the GyrB and ParE subunits, are poor dual targeting agents that operate mainly *via* the inhibition of GyrB. However, the high degree of similarity at the sequence and structural level between GyrB and ParE suggests that the design of potent dual targeting agents against the GyrB and ParE domains should be possible (25, 26).

2.2 Structural features of the type IIA topoisomerases

A number of X-ray structures (27) of the ATPase domain of DNA-gyrase complexed with phosphoaminophosphonic acid-adenylate ester (ADPNP) and novobiocin are available (Figure 1). The detailed analysis of these structures demonstrates that the ligands bind similarly to a key aspartic acid side-chain and a conserved water molecule in the binding cavity. In both cases, each ligand donates a hydrogen bond to the aspartate side chain and accepts a hydrogen bond from the conserved water molecule. Figure 2 illustrates the hydrogen-bond network observed for the adenine of the ADPNP and the carbamate moiety of novobiocin. Novobiocin is an ATP-competitive inhibitor that overlaps partially with the adenine moiety of ATP (27). This overlapping competitive binding area is the focus of novel inhibitor design.

[Insert Figures 1 & 2 here]

Recently, Bellon *et. al* from Vertex (27) reported the structure of the ATP-binding domain of *E. coli* ParE complexed with adenylyl-imidodiphosphate (ADPNP) at 2.0 Å resolution as well as a ParE novobiocin complex at 2.1 Å resolution (Figure 3). A comparison of the GyrB and ParE structures demonstrate strikingly similarity within the ATP-binding sites. Superposition of the *E. coli* ParE and GyrB structures shows the amino acids in the active site overlap closely. The ParE and the corresponding *E. coli* GyrB residues (in parenthesis) are as follows when ADPNP is bound (Figure 4): Y5 (Y5), E38 (E42), N42 (N46), E46 (E50), D69 (D73), M74 (I78), K99 (K103), Y105 (Y109), T163 (T165), Q332 (Q335), and K334 (K337). The authors also point out that the specific functional roles of GyrB vs. ParE residues in the ATP active site are transferable. Similarly, the key residues of both proteins that interact with novobiocin (Figure 5) are as follows: E46 (E50), D69 (D73), R72 (R76), M74 (I78), D77 (D81), I90 (I94), R132 (R136), and T163 (T165). Moreover, Figures 4 and 5 illustrate two important features: **1**) the hydrogen-bond interactions with D1069 observed for the adenine scaffold of the ADPNP also engage the carbamate moiety of novobiocin, and **2**) the orientation of these ligands with respect to the binding site environment – the polar phosphate group from the tail end of the ADPNP hydrogen bonds with K1099 & K1334, while the carbonyl-ester of the middle ring and the hydroxyl of the tail end of novobiocin hydrogen bonds with R1132 and D1077, respectively.

[Insert Figure 3 here]

To maintain dual targeting activity, the authors report that the key structural difference between GyrB and ParE is a single amino acid change from Ile78 to Met74, respectively. Charifson *et al* also report (28) that the binding site of ParE is narrower than that of GyrB because the switch from Ile78 in GyrB to Met74 in ParE, thus dictating the co-planarity requirement of the ligands to bind in the cavity. Moreover, our in-house crystallographic structures data [unpublished

results] for both targets complexed with our ligands suggest that the co-planarity requirement of the ligands is not solely dictated by Met74 of ParE or Ile78 of GyrB but also by the key aspartate side chain and conserved crystallographic water molecule in the binding site. Further analysis of the complexes shows the distance from the C β of the Met or the Ile to the centroid of our ligand is 3.489Å for ParE and 3.523Å for GyrB. The resultant 0.034-Å shorter distance in the ParE binding cavity seems small but is consistent with the selectivity profile of ParE with respect to the ligand's chemotype. This agrees well with Charifson's observation. However, when the ligands are removed, the calculated volumes of the whole binding site cavities for ParE and GyrB structures are 79.8 Å³ and 74.4 Å³, respectively. The overall volume of the ligand-binding cavity of ParE appears to be larger than that of GyrB by about 5.4 Å³.

[Insert Figures 4 & 5 here]

2.3 Identification of the “hot spots” in the GyrB binding cavity

Once a binding cavity is characterized, interaction “hot spots” within the binding cavity need to be identified. These “hot spots” are specific locations within the binding cavity that would potentially be exploited to interact with the inhibitor pharmacophore. Thus, the “hot spots” can be used for fragment-based screening and lead identification. One method to identify these “hot spots” is to use a variation of virtual screening (see Section 2.4). Schechner *et al* (29) screened the binding site of GyrB with small fragments. Combinations of these fragments are then used as the basis for the design of novel ligands. They used the Multiple Copy Simultaneous Search (MCSS) technique developed by Karplus *et. al* (30), where all chemical fragment conformations are generated and randomly distributed in the binding cavity, removing fragments that occupy the same positions. The fragment library contained twenty-three functional groups, representing

various chemical characteristics, such as charged, polar, hydrophobic, aromatic and aliphatic groups. Multiple structural states of the 24 kDa subdomain of GyrB identified crystallographically (to encompass active-site flexibility) were used to collect unbiased results. Overall, the protein conformations did not affect the functional group binding. However, binding energies and specific orientations of small functional groups were affected. The tightest binding pharmacophore among the groups tested was a phenol, forming a hydrogen bond with the key Asp73 in the binding cavity. Other functional groups were identified to target the phosphate binding site. A deep hydrophobic binding pocket near the ATP binding site was also identified and suggested a path for optimization of the inhibitor scaffolds identified in the pharmacophore search.

2.4 Computationally aided structure-based drug design

Structure-based drug design (SBDD) takes advantage of structural information about a protein target to identify and optimize lead compounds. The protein target structure is often derived from X-ray diffraction data, but may also be an NMR-based structure or based on a homology model (31-33). SBDD still relies on all the available information about the target and its biological role. Experimental data help to guide the computational chemist as to the most relevant calculations to drive optimization. This section aims to broadly describe the role of computational approaches in SBDD for lead identification and optimization.

At the lead compound identification stage, virtual screening searches large chemical databases to eliminate non-binders. When large compound libraries are screened, i.e., in the range of 100K or larger, the speed of the calculations is paramount, and this consideration leads to using relatively simple and fast docking algorithms. Each compound is assigned a score that is used to produce a

sorted list with the most promising predicted compounds at the top. Success at this stage is not focused on predicting accurate binding affinities, but rather on enrichment of strongly binding compounds ('hits'). Docking programs are sometimes evaluated for robustness by seeding libraries with known binding compounds to see if they appear at the top of the ranked list. Zhou *et. al* (34) used enrichment to compare the performance of Glide (35), DOCK (36), and GOLD (37).

After initial virtual screening of a large database, top-scored compounds are re-ordered using more accurate scoring methods. Re-scoring aims to correct some of the deficiencies in docking scores ignored in the initial stage. Effects to consider for re-scoring may include solvation (32) and accounting for flexibility of the protein target upon ligand binding. Two common re-scoring methods are Molecular Mechanics Generalized Born Surface Area (MM-GBSA) (38) and Linear Interaction Energy (LIE) (39). For some well studied cases, such as kinases (33) and DHFR (40), the protein has larger motions that respond to the ligand binding, making flexible conformations of the protein necessary for more accurate binding predictions.

2.4.1 Structure-based virtual screening

Virtual screening is a process where chemical structures are evaluated *in silico* against the target protein structure *via* docking and/or a 3D pharmacophore model. This approach enables the identification of the biologically relevant molecules more efficiently when the virtual library is large (>5 million) than a random experimental high throughput screening. Figure 6 illustrates a typical workflow for a virtual screen with a protein receptor as a target, from which a library of compounds is screened for a subset of compounds for further analysis. This section surveys a limited number of docking programs for virtual screening and their performance for studying

topoisomerase inhibitors against GyrB and/or ParE with emphasis on the identification of novel leads and improvement of these leads. The related *in silico* techniques such as target, ligand preparation, scoring functions, and molecular dynamics are herein broadly discussed.

[Insert Figure 6 here]

2.4.1.1 Target preparation

Several factors should be considered when preparing protein targets for molecular docking. Frequently, the structures used for the protein targets are derived from x-ray crystallography experiments. Within these structures, small molecules, such as water molecules, ions, co-factors, and crystallization agents can appear in the binding site. In other cases, regions of the protein may be unresolved or disordered. Also, because of the experimental limitations, hydrogen atoms are not usually resolved, meaning that protonation states of titratable groups must be assigned. Each of these issues needs to be addressed in the context of the relevant biology and the desired goals for the molecular docking. For example, conserved water molecules may play a key role in mediating contacts between a ligand and its protein receptor (41, 42). Certainly, this is the case for novobiocin (27). Thus, including key crystallographic water molecules can improve docking results (43, 44). However, the water positions may be different for each putative ligand, and in some cases displacing key waters with a ligand (45) may be more favorable. Other molecules to consider as necessary in the binding site include co-factors, metals and ligands of metals so that a more biologically relevant protein environment is represented. Most importantly, each of the residues in the binding site need to be completely represented. If portions of the binding site are missing in the original crystal structure, then homology modeling or loop prediction methods need to be used to complete the binding site structure.

2.4.1.2 Fundamentals of molecular docking

Proteins are well known to be flexible and conformationally dynamic. Typically, in molecular docking the protein targets are held fixed during virtual screening to minimize the computational cost. The key routine for docking programs is to **1)** generate poses for the ligand within the binding site and **2)** score each ligand pose as the predicted binding mode for that ligand.

In addition to the internal conformational space, the ligand rotational and transitional degrees of freedom are explored to generate a set of poses. The number of low energy conformers increases exponentially with the number of rotatable bonds, so efficiency in the search algorithm is imperative. To this end, several algorithms have been developed with modified protocols including: incremental construction of the ligand [DOCK (36), FlexX (46)], genetic algorithms that optimize the conformation and orientation of the ligand simultaneously in the binding pocket [AutoDock (47, 48), Fitted (49)], Monte Carlo based algorithms [MCDOCK (50), Ligandfit (51), QXP (52)], pre-generation of low energy ligand conformers before introducing them into the binding pocket [FLOG (53), FRED (54)], high temperature molecular dynamics simulations to generate random ligand conformations that are rigidly translated into the binding site [Cdocker (51), ICM (55)], matching ligand functional properties with hot spots in the binding site [Libdock (51)], calculated distance geometry for intraligand and ligand-sphere interactions [DockIt (56)], or exhaustive search algorithms [Glide (35)]. Many codes combine a number of the algorithms to meet the user's needs [Dock Vision (57, 58), MolDock (59), SurFlex (60), MoeDock (61)]. The major approaches are mentioned with common docking methods as examples only. Neither the details of available methods nor the completeness of the list of docking codes are the intention for this survey. For a more complete discussion on pose generation, the reader is referred to the review by Brooijmans and Kuntz (62).

The next step is to assign a score for each pose to: **1)** select the best pose for each compound and **2)** rank ligands based on their scores. Key challenges for the scoring functions are to account for the conformational entropy penalty of the ligand, the ligand desolvation, and the intramolecular interaction upon binding. While in principle we formally understand the theoretical basis to calculate the free energy of binding (63), in practice calculating free energy remains a research problem that requires massive (hundreds of cpu hours) computer power per ligand (see for example the work by Wang *et. al* (64) and Mobley *et. al* (65)). Due to the large number of ligands that need scores in a typical virtual screening campaign, simplified scoring functions assign pose scores. The scoring functions are a compromise between accuracy and computational efficiency.

Three broad categories of scoring functions exist: physics (force-field) based, knowledge-based and empirical scoring functions. Force-field based scoring functions base their scores on a function that explicitly lists physical interactions such as electrostatics and van der Waals. For example, the scoring function in AutoDock (47) originally took its parameters from the Amber force field (66).

Another approach is to consider a database of ligand-protein complexes and calculate the likelihood of forming specific contacts between ligands and proteins. A score based on this ‘knowledge’ is assigned by summing the likelihood of each contact a ligand pose has with a protein. This type of scoring is called knowledge-based scoring (67, 68).

In the case of an empirical scoring function, the strategy is to fit the scoring function to reproduce binding data from a set of training data. The development of this type of function often starts with either a force-field or knowledge-based scoring function. Each contribution is

weighed to reproduce binding affinity data (69, 70). For example, the later versions of Autodock (3.0 and above) introduced a scoring function that weighed the Amber force field parameters to better reproduce binding data.

Strategies exist to make the molecular docking results more reliable or at least represent experimental data better. Tirado-Rives *et al* (71) have recently introduced a conformer-focusing contribution to the uncertainty of predicting affinities. This contribution was used in most available docking programs such as Discovery Studio (51), MOE (61), and Glide (35) to account for the penalty of conformational change from the unbound state to the bound state of the ligand.

A number of studies (72-79) have evaluated several docking programs and scoring functions for different major targets such as kinases, metalloenzymes, serine proteases, nuclear receptors and DNA gyrase. Warren *et al* (80) from GSK reported that these programs performed well for some targets but poorly for others, in terms of the correct identification of actives or enrichment. The report showed that the best-to-least performers were FlexX, GOLD, LigandFit, QXP, Dockit, FRED, DOCK, Glide and MoeDock for GyrB. Schulz-Gasch *et al* (81) at Hoffman LaRoche reported their findings using FlexX, Glide and FRED towards GyrB, such that FlexX and Glide predicted the inhibitors well, while FRED's performance was the least among three programs in virtual screening. Generally, the docking programs – Glide, FlexX, DOCK and GOLD - were the most popular for using in lead identification by pharmaceutical companies, and their performance is not consistent across target classes. For example, the top performing programs for kinases were Glide, DOCK, ICM, LigandFit, FlexX, GOLD, respectively, and for metalloenzymes the order was FlexX, LigandFit, Glide, FRED and ICM. Additionally, there are other programs developed by pharmaceutical teams with comparable performance.

Other validation studies suggest that combining several scores into a consensus score may be better than any individual scoring function alone (82-86). Another popular strategy to make the scoring better is to re-examine the top scoring compounds. Because of the known short-comings of the rapid high-throughput virtual screening scoring functions, one strategy is to use more complex, accurate methods to re-score these compounds. Additionally scoring functions can be improved by including solvation effects *via* an implicit solvent treatment such as Poisson Boltzmann, or its approximation: generalized Born. Scoring *via* a molecular mechanics Poisson Boltzmann (MMPBSA) or MMGBSA can help improve binding affinity predictions (38, 87-90). The variation in performance of docking programs and scoring functions suggests that molecular docking is an active area of research with room for improvements (91).

2.4.1.3 Flexible receptor docking

To quickly screen large virtual libraries effectively, the protein target is typically fixed. However, cases exist where the protein needs to respond to the binding of substrates to achieve efficient binding. In the case of kinases (92, 93), DHFR (94, 95), aldose reductase (96, 97), and members of the enolase superfamily (98), the protein backbone moves substantially to adapt to small molecule binding. The conformational changes may introduce new van der Waals, electrostatic and hydrogen bond contacts or clear space for the ligand to fully enter the binding site. Addressing these changes can considerably affect docking pose generation and scoring. By simultaneously sampling the degrees of freedom in both the protein and the ligand, large changes in conformation can be explored and modeled more accurately. Certainly, a better representation of the conformational space should result in a better prediction, but it has an increased computational cost. The total conformational space available for exploration increases exponentially with each independent degree of freedom, regardless of whether it belongs to the

protein or ligand. If no experimental data exists for a ligand-protein complex structure where the protein has conformationally responded to ligand binding, then this flexible protein docking method becomes more difficult to justify.

Some conformational changes can be relatively small, where only a few side-chains rearrange to accommodate ligand binding. While small, these changes may have a large impact on binding. Consider, for example, orienting polar groups to optimize a hydrogen bond network. If only a few residues move, sampling their degrees of freedom along with the ligand conformational space is the most practical approach. Some programs, such as AutoDock (47) and Glide (35) include options to specify flexible residues in their induced fit calculations. AutoDock performs the search of all degrees of freedom in one run while Glide undergoes a cycle of optimization of docking the ligand and optimizing flexible side-chains (99).

Another approach is to run molecular dynamics (MD) simulations on the free protein target to explore the conformational space available to the protein (relaxed complex method). The key advantage is that the receptor degrees of freedom are calculated once instead of for each ligand individually. While relatively new, some work has shown this approach is viable. Wong *et al* (100) performed MD simulations of the receptor and docked into trajectory snapshots. They found improved rankings of known ligands. Combining molecular dynamics and docking has also been proposed by other authors (101). Note, however, that large conformational changes, particularly those involving an entropic barrier, will be difficult to sample using molecular dynamics. Therefore, this approach is really best suited for relatively small changes at the side-chain level.

Larger conformational changes, whether local (loop rearrangements) or global (domain hinge motions), are more challenging to capture computationally. However, some groups have shown progress in this area. Wong and Jacobson (102) were able to predict loop rearrangements involved in ligand binding of up to 15 residues in length. They showed that the enrichment of known ligands improved when a database was docked against the predicted structures. In special cases, experimental evidence greatly simplifies the problem. For example, if a domain-domain rearrangement involves the two domains acting as rigid bodies around a small, flexible hinge region, then one can include only the relevant degrees of freedom in the docking calculation (103). Reducing the effective number of degrees of freedom is a viable approach, if one can identify them *a priori*.

Post-processing of molecular dynamics or more advanced sampling techniques can guide the selection of key receptor degrees of freedom. For example, normal mode or principal component analysis of a molecular dynamics simulation is one approach to identify large-scale concerted motions (104). The result of this type of analysis is a list of concerted motions that individually account for a fraction of the motions in the molecular dynamics simulation. The top 2 or 3 motions (normal modes or principal components) likely include the majority of variation in the conformational space.

An alternative to predicting large induced fit changes is to use a holo crystal structure where the changes are evident. While smaller changes may remain, even in domain-domain hinge regions, holo structures will likely capture the large conformational change while small perturbations may be adjusted at the side-chain level as discussed in this section above.

2.4.2 Identification of new lead compounds by virtual screening

The aim of virtual screening is to identify actives and inactive compounds. In lead identification accurate classification of hits as active compounds needs to be confirmed experimentally. Several groups have used virtual screening to find new lead compounds that bind to GyrB. Boehm *et. al* at Roche (26) used the binding interactions of the novobiocin-GyrB complex structure, two key hydrogen bond interactions of the ligand with an aspartate side chain and with a conserved crystallographic water molecule, as the required binding motif (Figure 7) to virtually screen a commercially available database and an in-house collection. They used LUDI and CATALYST (51) software to identify 3000 compounds, from which 150 hits were confirmed and clustered into 14 classes. Seven classes (Figure 8) were validated as novel GyrB inhibitors: phenols, 2-aminotriazines, 4-amino-pyrimidines, 2-amino-pyrimidines, pyrrolopyrimidines, indazoles and 2-hydroxymethyl-indoles. The molecular weights of these “lead-like” hits were lower than that of novobiocin, and their activities were within 2-3 orders of magnitude higher than that of this antibiotic compound. Of these lead-like hits, the Indazole A complexed with the 24kDa N-terminal fragment of GyrB structure (Figure 9) was used to explore the potential van der Waals interactions of the ligand with the lipophilic surface of Ile78, Pro79 and Ile94. The resulting activity of compound B indicated 10 times more potency than novobiocin, and the binding hypothesis was confirmed by the x-ray structure, where the indazole moiety hydrogen bonded with the postulated Asp73 and the conserved water molecule. The mercaptobenzoic acid side chain pi-stacked with the Glu50-Arg76 salt bridge and made a hydrogen bonding interaction to Arg136 and the benzyloxy side chain interacted with the lipophilic area of Ile78-Pro79-Ile94.

[Insert Figures 7, 8 & 9 here]

Oblak *et al* (105) also used LUDI to identify low-molecular weight fragments of 2-aminobenzimidazole and indolin-2-one as potential GyrB inhibitors from the MDL-databases

(Figure 10). Because DNA gyrase is a member of the GHKL ATPase superfamily (106), it is structurally related to the larger family of kinases, where the indolinone hit is known as a potent inhibitor of many kinases (92) much like tyrosine, JNK, casein, etc.

Focusing on the previously unexplored structural pocket at the dimer interface of subunit A and a small area of the ATP binding pocket on subunit B targeted by coumarin and cyclothialidine, Ostrov *et. al* (107) used DOCK (v5.1.0) to screen 140,000 small molecules and identified the following chemotypes (Figure 11), such as methylxanthines, imidazole carboxamides, pyrimidine nitrous amides, benzothiazolinones, aminonaphthoquinones, mandelic acids, aminobenzoacetic acids, and imidocarbonodithioic diamides as the potential scaffolds for novel GyrB antibacterial leads targeting fluoroquinolone-resistant strains.

[Insert Figures 10 & 11 here]

2.5 Structure-based lead optimization

The main goal of lead optimization is to start with initial leads, selected among the computational and experimental screen hits, and optimize them to yield strongly binding, highly specific and effective compounds. Once a lead is identified, more accurate computational methods can be employed to optimize the ligand. A structure-based approach focuses on improving binding affinity by suggesting perturbations to the lead compound. One approach is to use some of the re-scoring tools such as MMGBSA or LIE methods. Another approach is to use free energy methods that are more general and potentially more accurate, but substantially more computationally expensive.

MMGBSA and LIE are both post processing methods that can be useful in lead optimization. Early in its development MMGBSA involved producing a thermodynamic ensemble of the

protein-ligand complex and averaging the interaction energy of all the structures. An alternate approach is to calculate the binding energy for a single protein-ligand structure using the MM-GBSA energy model. Typically, the protein-ligand input to the single-point MMGBSA rescoring is the product of a docking calculation. The distinction between the two approaches is important because producing a thermodynamic ensemble of the protein-ligand complex is computationally intensive. There are two types of studies common in the literature: **1)** those that focus on reproducing published data and **2)** those that use these methods to discover new inhibitors. While these two methods are well known in academia, their use in pharmaceutical development is relatively new. Publications using MMGBSA or LIE for drug discovery have appeared in the literature only in the last 5-10 years.

De Amorim *et al* (108) present an in depth discussion of the methods and applications of the linear interaction energy method (LIE). A retrospective study, using data for inhibitors of CDK2, Lck and p38, was used to fit parameters for a LIE model with continuum electrostatics. Using this method, Kolb *et al* (109) found novel, low micromolar inhibitors of EphB4 and CDK2.

Rastelli *et al* (110) performed a retrospective study using known DHFR inhibitors to test the usefulness of MMGBSA and MMPBSA in a virtual screening tool. They found that the methods were able to discriminate between binders and decoys in this diverse set. They found similar results when they performed ensemble averaging over snapshots of an MD trajectory versus single point calculations on a single protein-ligand complex structure. Bag *et al* (111) used MMGBSA estimates of binding free energy to discover novel inhibitors of DHFR. The correlation between the IC₅₀s of the inhibitors and the MMGBSA predictions was ~ 0.79. Hendricksen *et al* (112) identified novel bacterial histidine biosynthesis inhibitors using a

combined MMGBSA/MMPBSA approach. Most importantly, the authors considered side-chain flexibility by running an MD simulation and averaging MMGBSA interaction energies from trajectory snapshots. They tested the compounds on a whole-cell assay directly without any initial enzymatic assay because of their confidence in the improved predictive power of these approaches. A combination of docking and MMPBSA ranking yielded new inhibitors of dengue virus methyltransferase (113).

Structure based drug design is particularly well suited for optimizing binding affinity, which frequently translates to higher potency. However, the effectiveness of a drug not only depends on potency but on other properties and biological activities, including but not limited to solubility, octanol/water partition coefficient, affinity for molecular pumps, ability to penetrate relevant tissues, lifetime in the blood stream, toxicology and pharmacokinetic properties. Thus, many have coupled a number of methods together to maximize the design effectiveness. To this end, the understanding of the structure-activity relationships (SAR) is very important when applying structure-based drug design methods for drug discovery. The rational design of experiments to generate biological data is a crucial step, providing possible comparisons of predictions and experimental results. The quantitative structure-activity relationships approach (QSAR) is a valuable drug design tool that can offer solutions based on the statistical analysis of relationships between chemical structure and biological activity in a quantitative and mechanism-based manner. Herein, the single-targeting GyrB approach is reported for the novobiocin and cyclothialidine lead compound generation and optimization by using structure-based and SAR-based techniques. Subsequently, the dual-targeting GyrB/ParE approach aimed at minimizing target-based resistance development is exemplified with the work from other groups that identified newer chemotypes beyond coumarins (*e.g.* novobiocin).

Musicki *et al* (114) utilized the x-ray structures of a 24-kDa N-terminal fragment of GyrB to identify a larger hydrophobic region consisting of Val94 and Phe95, where the methyl groups of novobiocin were located. These two amino acids are conserved across twelve Gram-positive strains and in *E. coli*. Using the combined information, they modified the methyl groups of novobiocin to spirocyclopentyl moiety derivatives with lower molecular weights and improved physiochemical properties and pharmacokinetic profiles. The spirocyclopentylnoviose (RU79115) is shown in Figure 12 with a MIC₅₀ of 0.08 µg/mL, while the MIC₅₀ of novobiocin is in the range of 0.3-40 µg/mL.

[Insert Figure 12 here]

Angehrn *et al* (115) from Hoffman-La Roche used structure-activity relationships derived from various substitution patterns and x-ray structural analyses to identify the simple hydroxylated benzyl sulfide moiety as the active structural feature in cyclothialidine, a potent gyrase inhibitor. Inhibitors with this moiety demonstrated *in vitro* activity against Gram-positive bacteria. The best activities were shown by the 14-membered lactone C, which incorporated the benzyl sulfide as well as other features that improved pharmacokinetic properties and lipophilicity (Figure 13). Lactone C exhibits *in vivo* efficacy with an ED₅₀ of 25 mg/kg against a *S. aureus* septicemia model and overcomes resistance against other marketed antibiotics drugs.

[Insert Figure 13 here]

Scientists at Vertex used the information from the structures of ParE complexed with ADPNP and novobiocin (described in section 2.2) to design a series of aminobenzimidazoles (28) such as VRT-125853 and 752586 that demonstrate dual inhibition of GyrB and ParE (Figure 14). The K_i (µM) values of VRT-125853 against *E. coli* were 0.015 and 0.68 for GyrB and ParE,

respectively. Similarly, the K_i (μM) values for VRT-752586 were < 0.004 (GyrB) and 0.023 (ParE). Resistance incidence studies showed very low rates of resistance, consistent with the dual targeting profile of the inhibitors. As with novobiocin and ATP, the aminobenzimidazole core exploits the anchoring H-bonding interactions with the conserved Asp residue (D73) and structural water molecule that is coordinated by D73 and T165.

[Insert Figure 14 here]

In a similar vein, Pfizer (116) and Evotec (117) teams reported a series of imidiazolo- and triazolo-pyridines as dual inhibitors of bacterial GyrB and ParE (Figures 15, 16). These synthetically challenging pyridines exhibit the same shape and pharmacophore as the aminobenzimidazoles developed by the Vertex group. The two groups reasoned that their 5,6-fused heterocyclic scaffolds could provide different SAR and optimization opportunities with alternative ADME, pharmacology, toxicology and efficacy profiles. The authors pointed out that achievement of dual-targeting across the series was challenging and only a few isolated molecules were observed with balanced, dual enzyme activity. The IC_{50} (μM) values of imidazolo-pyridines against *Spn* reported by the Pfizer team for compound D were 1.91 (GyrB) and 1.38 (ParE), and those of compound G were 0.117 (GyrB) and 0.147 (ParE). The IC_{50} (μM) values of triazolo-pyridines against *E.coli* reported by the Evotec team were 9.64 (GyrB, no activity measured for ParE) for compound M and 0.042 (GyrB) and 11 (ParE) for compound Z.

[Insert Figure 15 & 16 here]

In recent years there has been a paradigm shift in antibiotic discovery from the development of single-target single-pharmacophore agents to multi-target single-pharmacophore agents. This trend is exemplified by the efforts directed towards the discovery of novel agents that inhibit the

function of bacterial topoisomerases as described in this survey. Undoubtedly, structural biology and computational techniques have helped to attain the promising results.

3 Concluding Remarks

The availability of high resolution crystal structures of GyrB and ParE, and the opportunity for dual target inhibition to minimize the potential resistance development have catalyzed numerous efforts to discover novel antibacterial agents against these targets. In conjunction with crystal structures, docking-based virtual screening applications have demonstrated effectiveness as a quick and inexpensive alternative to high throughput screening. Of the available methods, some have demonstrated reasonable accuracy and continue to be improved. At the time this chapter is being written there are more than 30-million unique compounds available in commercial libraries in which to search for leads. Invariably, *in silico* screening will identify numerous potential inhibitor scaffolds beyond the molecules surveyed here. The combination of computational structure-based drug design methods and multi-objective ADMET structure-activity relationships are valuable and effective drug design tools that identify and optimize lead compounds for tight binding, high specificity and effectiveness. The chemical space diversity of dual-target topoisomerase inhibitors continues to expand beyond the coumarin and quinolone classes by utilizing these drug development tools.

In summary, the drug discovery process has steadily become more information driven. Thus, using the knowledge of drug resistance gained from the structure-function-activity relationships combined with continuing improvements in computational structure-based methods should help feed drug discovery pipelines with novel inhibitors with the desired properties for development into the next generation of antibiotics.

4 Acknowledgements

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Release number LLNL-JRNL-476952.

5 References

1. Lax E (2004) *The Mold On Dr. Florey's Coat: The Story of The Penicillin Miracle* (Henry Holt & Co., New York) p 307 pp.
2. Talbot GH, Bradley J, Edwards JJE, Bilbert D, Scheld M, & Barlett JG (2006) Bad Bugs Need Drugs: An Update on the Development Pipeline from the Antimicrobial Availability Task Force of the Infectious Diseases Society of America. *Clin. Infect. Dis.* 42:657-668.
3. Aspa J, Rajas O, Rodriguez de Castro F, Blanquer J, Zalacain R, Fenoll A, de Celes R, Vargas A, Salvanes FR, Espana PP, Rello J, & Torres A (2004) Drug-resistant pneumococcal pneumonia: clinical relevance and related factors. *Clin. Infect. Dis.* 38:787-798.
4. Barrett CT & Barrett JF (2003) Antibacterials: are the new entries enough to deal with the emerging resistance problems? *Curr. Opin. Biotech.* 14:621-626.
5. Jacobs MR (2003) Worldwide trends in antimicrobial resistance among common respiratory tract pathogens. *Pediatr. Infect. Dis. J.* 22:S109-S119.
6. Livermore DM (2004) The need for new antibiotics. *Clin. Microbiol. Infect.* 10:(Suppl. 4) 1-9.
7. Pottumarthy S, Fritsche TR, & Jones RN (2005) Comparative activity of oral and parenteral cephalosporins tested against multidrug-resistant *Streptococcus pneumoniae*: report from the SENTRY Antimicrobial Surveillance Program (1997–2003). *Diagn. Microbiol. Infect. Dis.* 51:147-150.
8. Mutnick AH, Enne V, & Jones RN (2003) Linezolid resistance since 2001: SENTRY Antimicrobial Surveillance Program. *Ann. Pharmacother.* 37:769-774.

9. Schwartz BS, Ngo PD, & Guglielmo BJ (2008) Daptomycin treatment failure for vancomycin-resistant *Enterococcus faecium* infective endocarditis: impact of protein binding? *Ann. Pharmacother.* 42:289-290.
10. Cassell GH & Mekalanos J (2001) Development of antimicrobial agents in the era of new and reemerging infectious diseases and increasing antibiotic resistance. *J. Am. Med. Assoc.* 285:601-605.
11. Spratt BG (1994) Resistance to antibiotics mediated by target alterations. *Science* 264:388-393.
12. Ng EY, Trucksis M, & Hooper DC (1996) Quinolone resistance mutations in topoisomerase IV: relationship to the flqA locus and genetic evidence that topoisomerase IV is the primary target and DNA gyrase is the secondary target of fluoroquinolones in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 40:1881-1888.
13. Chopra I (1998) Protein synthesis as a target for antibacterial drugs: current status and future opportunities. *Expert Opin. Investig. Drugs* 7:1237-1244.
14. Silver L (2007) Multi-targeting by monotherapeutic antibacterials. *Nat. Rev. Drug Dis.* 6:41-55 and references cited therein.
15. Strahilevitz J & Hooper DC (2005) Dual targeting of topoisomerase IV and gyrase to reduce mutant selection: direct testing of the paradigm by using WCK-1734, a new fluoroquinolone, and ciprofloxacin. *Antimicrob. Agents Chemother.* 49:1949-1956.
16. Mills SD (2003) The role of genomics in antimicrobial discovery. *J. Antimicrob. Chemother.* 51:749-752.
17. Schneider G & Bohm HJ (2002) Virtual screening and fast automated docking methods. *Drug Discov. Today* 7:64-70.

18. Bailey D & Brown D (2001) High-throughput chemistry and structure-based design: survival of the smartest. *Drug Discov. Today* 6:57-59.
19. Maxwell A (1997) DNA gyrase as a drug target. *Trends Microbiol.* 5:102-109.
20. Drlica K & Zhao XL (1997) DNA gyrase, topoisomerase IV, and the 4-quinolones. *Microbiol. Mol. Biol. Rev.* 61:377-392.
21. Wang JC (1996) DNA topoisomerases. *Annu. Rev. Biochem.* 65:635-692.
22. Hiasa H, Yousef DO, & Marians KJ (1996) DNA strand cleavage is required for replication fork arrest by a frozen topoisomerase-quinolone-DNA ternary complex. *J. Biol. Chem.* 271:26424-26429.
23. Brino L, Urzhumtsev A, Mousli M, Bronner C, Mitschler A, Oudet P, & Moras D (2000) Dimerization of Escherichia coli DNA-gyrase B provides a structural mechanism for activating the ATPase catalytic center. *J. Biol. Chem.* 275:9468-9475.
24. Hooper DC (1999) Mechanisms of quinolone resistance. *Drug Resistance Updates* 2:38-55.
25. Bradbury BJ & Pucci MJ (2008) Recent advances in bacterial topoisomerase inhibitors. *Current Opinion in Pharmacology* 8:574-581.
26. Boehm H-J, Boehringer M, Bur D, Gmuender H, Huber W, Klaus W, Kostrewa D, Kuehne H, Luebbers T, Meunier-Keller N, & Mueller F (2000) Novel inhibitors of DNA Gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization. A promising alternative to random screening. *J. Med. Chem.* 43:2664-2674.
27. Bellon S, D. PJ, Wei Y, Hayakawa K, Swenson LL, Charifson PS, Lippke JA, Aldape R, & Gross CH (2004) Crystal Structures of Escherichia coli Topoisomerase IV ParE Subunit (24 and 43 Kilodaltons): a Single Residue Dictates Differences in Novobiocin Potency

against Topoisomerase IV DNA Gyrase. *Antimicrobial Agents and Chemotherapy* 48:1856-1864 and the references cited therein.

28. Charifson PS, Grillot A, Grossman TH, Parsons JD, Badia M, Bellon S, Deininger DD, Drumm JE, Gross CH, Letiran A, Liao Y, Mani N, Nicolau DP, Perola E, Ronkin S, Shannon D, Swenson LL, Tang Q, Tessier PR, Tian S, Trudeau M, Wang T, Wei Y, Zhang H, & Stamos D (2008) Novel Dual-targeting Benzimidazole Urea Inhibitors of DNA Gyrase and Topoisomerase IV Possessing Potent Antibacterial Activity: Intelligent Design and Evolution through the Judicious use of Structure-guided Design and Structure-activity Relationships. *J. Med. Chem.* 51:5243-5263.
29. Schechner M, Sirockin F, Stote RH, & Dejaegere AP (2004) Functionality Maps of the ATP Binding Site of DNA Gyrase B: Generation of a Consensus Model of Ligand Binding. *J. Med. Chem.* 47:4373-4390.
30. Miranker A & Karplus M (1991) Functionality maps of binding sites: a multiple copy simultaneous search method. *Prot.: Struct. Funct. and Genet.* 11:29-34.
31. McRobb FM, Capuano B, Crosby IT, Chalmers DK, & Yuriev E (2010) Homology Modeling and Docking Evaluation of Aminergic G Protein-Coupled Receptors. *Journal of Chemical Information and Modeling* 50(4):626-637.
32. Oshiro C, Bradley EK, Eksterowicz J, Evensen E, Lamb ML, Lanctot JK, Putta S, Stanton R, & Grootenhuis PDJ (2004) Performance of 3D-Database Molecular Docking Studies into Homology Models. *Journal of Medicinal Chemistry* 47(3):764-767.
33. Ferrara P & Jacoby E (2007) Evaluation of the utility of homology models in high throughput docking. *Journal of Molecular Modeling* 13(8):897-905.

34. Zhou Z, Felts A, Friesner R, & Levy R (2007) Comparative Performance of Several Flexible Docking Programs and Scoring Functions: Enrichment Studies for a Diverse Set of Pharmaceutically Relevant Targets. *Journal of Chemical Information and Modeling* 47(4):1599-1608.
35. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, & Shenkin PS (2004) Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry* 47(7):1739-1749.
36. Moustakas D, Lang P, Pegg S, Pettersen E, Kuntz I, Brooijmans N, & Rizzo R (2006) Development and validation of a modular, extensible docking program: DOCK 5. *Journal of Computer-Aided Molecular Design* 20(10):601-619.
37. Jones G, Willett P, Glen RC, Leach AR, & Taylor R (1997) Development and Validation of a Generic Algorithm for Flexible Docking. *J. Mol. Biol.* 267:727-748.
38. Guimarães CRW & Cardozo M (2008) MM-GB/SA Rescoring of Docking Poses in Structure-Based Lead Optimization. *Journal of Chemical Information and Modeling* 48(5):958-970.
39. Hansson T, Marelius J, & Åqvist J (1998) Ligand binding affinity prediction by linear interaction energy methods. *Journal of Computer-Aided Molecular Design* 12(1):27-35.
40. Osborne MJ, Schnell J, Benkovic SJ, Dyson HJ, & Wright PE (2001) Backbone Dynamics in Dihydrofolate Reductase Complexes: Role of Loop Flexibility in the Catalytic Mechanism†. *Biochemistry* 40(33):9846-9859.
41. Mancera R (2007) Molecular modeling of hydration in drug design. *Current Opinion in Drug Discovery and Development* 10:275-280.

42. Wong SE & Lightstone FC (2011) Accounting for water molecules in drug design. *Expert Opinion on Drug Discovery* 6(1):65-74.
43. Huang N & Shoichet BK (2008) Exploiting Ordered Waters in Molecular Docking. *Journal of Medicinal Chemistry* 51(16):4862-4865.
44. Minke WE, Diller DJ, Hol WGJ, & Verlinde CLMJ (1999) The Role of Waters in Docking Strategies with Incremental Flexibility for Carbohydrate Derivatives: Heat-Labile Enterotoxin, a Multivalent Test Case. *Journal of Medicinal Chemistry* 42(10):1778-1788.
45. Michel J, Tirado-Rives J, & Jorgensen WL (2009) Energetics of Displacing Water Molecules from Protein Binding Sites: Consequences for Ligand Optimization. *Journal of the American Chemical Society* 131(42):15403-15411.
46. Rarey M, Kramer B, Lengauer T, & Klebe G (1996) A Fast Flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* 261:470-489.
47. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, & Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry* 19(14):1639-1662.
48. Jones G, Willett P, & Glen RC (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *Journal of Molecular Biology* 245(1):43-53.
49. Corbeil CR & Moitessier N (2009) Docking Ligands into Flexible and Solvated Macromolecules. 3. Impact of Input Ligand Conformation, Protein Flexibility, and Water molecules on the Accuracy of Docking Programs. *J. Chem. Inf. Model.* 49:997-1009.
50. Liu M & Wang S (1999) MCDOCK: A Monte Carlo simulation approach to the molecular docking problem. *Journal of Computer-Aided Molecular Design* 13(5):435-451.

51. Accelrys Discovery Studio Suites, Accelrys, Inc, 10188 Telesis Court, Suite 100 San Diego, CA 92121, USA.
52. McMartin C & Bohacek RS (1997) QXP, Powerful, rapid computer algorithms for structure-based drug design. *J. Comput. Aided. Mol. Design* 11:333-344.
53. Miller MD, Kearsley SK, Underwood DJ, & Sheridan RP (1994) FLOG: A system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure. *Journal of Computer-Aided Molecular Design* 8(2):153-174.
54. FRED: Fast Rigid Exhaustive Docking, Version 2.2.5, OpenEye Scientific Software, Inc., Santa Fe, NM, USA.
55. Abagyan RT, & Kuznetsov, M. R., (1994) A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* 15:488-506.
56. Blaney JM & Dixon JS, DockIt, Metaphorics, LLC, Mission Viejo, CA.
57. Hart TN & Read RJ (1992) A multiple-start Monte Carlo docking method. *Proteins* 13:206-222.
58. Hart TN, Ness, R. S., & Read, R. J., (1997) Critical evaluation of the research docking program for the CASP2 challenge. *Proteins*, Suppl 1:205-209.
59. Thomsen R & Christensen MH (2006) MolDock: A New Technique for High-Accuracy Molecular Docking. *J. Med. Chem.* 49(11):3315-3321.
60. Jain AN (2003) Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* 46:499-511.
61. MOE: Molecular Operating Environment, Chemical Computing Group Inc., 1010 Sherbrooke, Street West, Suite 910, Montreal, Quebec, Canada.

62. Brooijmans N & Kuntz ID (2003) Molecular Recognition and Docking Algorithms. *Annual Review of Biophysics and Biomolecular Structure* 32(1):335-373.
63. Gilson M, Given J, Bush B, & McCammon JA (1997) The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophysical Journal* 72(3):1047-1069.
64. Wang J, Deng Y, & Roux B (2006) Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials. 91(8):2798-2814.
65. Mobley DL, Graves AP, Chodera JD, McReynolds AC, Shoichet BK, & Dill KA (2007) Predicting Absolute Ligand Binding Free Energies to a Simple Model Site. *Journal of Molecular Biology* 371(4):1118-1134.
66. Morris GM, Goodsell DS, Huey R, & Olson AJ (1996) Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4. *Journal of Computer-Aided Molecular Design* 10(4):293-304.
67. Muegge I (2000) A knowledge-based scoring function for protein-ligand interactions: Probing the reference state. *Perspectives in Drug Discovery and Design*, 20:99-114.
68. Gohlke H, Hendlich M, & Klebe G (2000) Knowledge-based Scoring Function to Predict Protein-Ligand Interactions. *Journal of Molecular Biology* 295:337-356.
69. Eldridge M, Murray C, Auton T, Paolini G, & Mee R (1997) Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of Computer-Aided Molecular Design* 11(5):425-445.

70. Korb O, Stutzle T, & Exner TE (2009) Empirical Scoring Functions for Advanced Protein–Ligand Docking with PLANTS. *Journal of Chemical Information and Modeling* 49(1):84-96.
71. Tirado-Rives J & Jorgensen WL (2006) Contribution of conformer focusing to the uncertainty in predicting free energies for protein-ligand binding. *J. Med. Chem.* 49:580-5884.
72. Bursulaya BD, Totrov M, Abagyan R, & Brooks III CL (2003) Comparative study of several algorithms for flexible ligand docking. *J. Comput. Aided. Mol. Design* 17:755-763.
73. Kellenberger E, Rodrigo J, Muller P, & Rognan D (2004) Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* 57:225-242.
74. Klon AE, Glick M, Thoma M, Acklin P, & Davies JW (2004) Finding more needles in the haystack: a simple and efficient method for improving high-throughput docking results. *J. Med. Chem.* 47:2743-2749.
75. Perola E, Walters WP, & Charifson PS (2004) A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* 56:235-249.
76. Muege I & Enyedy I (2004) Virtual screening for kinase targets. *J. Curr. Med. Chem.* 11:693-707.
77. Cummings MD, DesJarlais RL, Gibbs AC, Mohan V, & Jaeger EP (2005) Comparison of automated docking programs as virtual screening tools. *J. Med. Chem.* 48:962-976.
78. Kontoyianni M, Sokol GS, & McClellan LM (2005) Evaluation of library ranking efficacy in virtual screening. *J. Comput. Chem.* 26:11-22.

79. Chen H, Lyne PD, Giordanetto F, Lovell T, & Li J (2006) On evaluating molecular docking methods for pose prediction and enrichment factors. *J. Chem. Inf. Model.* 46:401-415.
80. Warren GL, Andrews CW, Capelli A-M, Clarke B, LaLonde J, Lambert MH, Lindvall M, Nevins N, Semus SF, Senger S, Tedesco G, Wall I, D., Woolven JM, Peishoff CE, & Head MS (2006) A critical assessment of docking programs and scoring functions. *J. Med. Chem.* 49:5912-5931.
81. Schulz-Gasch T & Stahl M (2003) Binding site characteristics in structure-based virtual screening: Evaluation of current docking tools. *J. Mol. Model* 9:47-57.
82. Wang R, Lu Y, Fang X, & Wang S (2004) An Extensive Test of 14 Scoring Functions Using the PDBbind Refined Set of 800 Protein–Ligand Complexes. *Journal of Chemical Information and Computer Sciences* 44(6):2114-2125.
83. Wang R, Lu Y, & Wang S (2003) Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *Journal of Medicinal Chemistry* 46(12):2287-2303.
84. Clark RD, Strizhev A, Leonard JM, Blake JF, & Matthew JB (2002) Consensus scoring for ligand/protein interactions. *Journal of Molecular Graphics and Modelling* 20(4):281-295.
85. Gohlke H & Klebe G (2001) Statistical potentials and scoring functions applied to protein-ligand binding. *Current Opinion in Structural Biology* 11(2):231-235.
86. Charifson PS, Corkery JJ, Murcko MA, & Walters WP (1999) Consensus Scoring: A Method for Obtaining Improved Hit Rates from Docking Databases of Three-Dimensional Structures into Proteins. *Journal of Medicinal Chemistry* 42(25):5100-5109.

87. Graves AP, Shivakumar DM, Boyce SE, Jacobson MP, Case DA, & Shoichet BK (2008) Rescoring Docking Hit Lists for Model Cavity Sites: Predictions and Experimental Testing. *Journal of Molecular Biology* 377(3):914-934.
88. Zhong S, Zhang Y, & Xiu Z (2010) Rescoring ligand docking poses. *Current Opinion in Drug Discovery and Development* 13:326-334.
89. Lyne PD, Lamb ML, & Saeh JC (2006) Accurate Prediction of the Relative Potencies of Members of a Series of Kinase Inhibitors Using Molecular Docking and MM-GBSA Scoring. *Journal of Medicinal Chemistry* 49(16):4805-4808.
90. Srivastava M, Singh H, & Naik PK (2010) Molecular Modeling Evaluation of the Antimalarial Activity of Artemisinin Analogues: Molecular Docking and Rescoring using Prime/MM-GBSA Approach. *Current Research Journal of Biological Sciences* 2(2):83-102.
91. Cole JC, Murray CW, Nissink JW, Taylor RD, & Taylor R (2005) Comparing protein-ligand docking programs is difficult. *Proteins* 60:325-332.
92. Noble M, Barrett P, Endicott J, Johnson L, McDonnell J, Robertson G, & Zawaira A (2005) Exploiting structural principles to design cyclin-dependent kinase inhibitors. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* 1754(1-2):58-64.
93. Cavasotto CN & Abagyan RA (2004) Protein Flexibility in Ligand Docking and Virtual Screening to Protein Kinases. *Journal of Molecular Biology* 337(1):209-225.
94. Bowman AL, Lerner MG, & Carlson HA (2007) Protein Flexibility and Species Specificity in Structure-Based Drug Discovery: Dihydrofolate Reductase as a Test System. *Journal of the American Chemical Society* 129(12):3634-3640.

95. Then RL (2004) Antimicrobial Dihydrofolate Reductase Inhibitors - Achievements and Future Options: Review. *Journal of Chemotherapy* 16:3-12.
96. Klebe G, Krämer O, & Sotriffer C (2004) Strategies for the design of inhibitors of aldose reductase, an enzyme showing pronounced induced-fit adaptations. *Cellular and Molecular Life Sciences* 61(7):783-793.
97. Sotriffer CA, Krämer O, & Klebe G (2004) Probing flexibility and “induced-fit” phenomena in aldose reductase by comparative crystal structure analysis and molecular dynamics simulations. *Proteins: Structure, Function, and Bioinformatics* 56(1):52-66.
98. Babbitt PC, Hasson MS, Wedekind JE, Palmer DRJ, Barrett WC, Reed GH, Rayment I, Ringe D, Kenyon GL, & Gerlt JA (1996) The Enolase Superfamily: A General Strategy for Enzyme-Catalyzed Abstraction of the α -Protons of Carboxylic Acids, *Biochemistry* 35(51):16489-16501.
99. Fernández-Recio J, Totrov M, & Abagyan R (2002) Soft protein–protein docking in internal coordinates. *Protein Science* 11(2):280-291.
100. Wong CF, Kua J, Zhang Y, Straatsma TP, & McCammon JA (2005) Molecular docking of balanol to dynamics snapshots of protein kinase A. *Proteins: Structure, Function, and Bioinformatics* 61(4):850-858.
101. Alonso H, Bliznyuk AA, & Gready JE (2006) Combining Docking and Molecular Dynamics Simulations in Drug Design. *Medicinal Research Reviews* 26(5):531-568.
102. Wong S & Jacobson MP (2008) Conformational selection in silico: Loop latching motions and ligand binding in enzymes. *Proteins: Structure, Function, and Bioinformatics* 71(1):153-164.

103. Sandak B, Wolfson HJ, & Nussinov R (1998) Flexible Docking Allowing Induced Fit in Proteins: Insights From an Open to Closed Conformational Isomers. *Proteins: Structure, Function, and Bioinformatics* 32:159-174.
104. Cavasotto CN, Kovacs JA, & Abagyan RA (2005) Representing Receptor Flexibility in Ligand Docking through Relevant Normal Modes. *Journal of the American Chemical Society* 127(26):9632-9640.
105. Oblak M, Gdadolnik SG, Kotnik M, Jerala R, Filipic M, & Solmajer T (2005) In silico fragment-based discovery of indolin-2-one analogues as potent DNA gyrase inhibitors. *Bioorganic & Medicinal Chemistry Letters* 15:5207-5210.
106. Dutta R & Inouye M (2000) GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem. Sci.* 25:24-28.
107. Ostrov DA, Prada JAH, Corsino PE, Finton KA, Le N, & Rowe TC (2007) Discovery of novel DNA Gyrase inhibitors by high-throughput virtual screening. *Antimicrobial Agents and Chemotherapy* 51:3688-3698.
108. De Amorim HLN, Caceres RA, & Netz PA (2008) Linear Interaction Energy (LIE) Method in Lead Discovery and Optimization. *Current Drug Targets* 9:1100-1105.
109. Kolb P, Huang D, Dey F, & Caflisch A (2008) Discovery of Kinase Inhibitors by High-Throughput Docking and Scoring Based on a Transferable Linear Interaction Energy Model. *Journal of Medicinal Chemistry* 51(5):1179-1188.
110. Rastelli G, Rio AD, Degliesposti G, & Sgobba M (2010) Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *Journal of Computational Chemistry* 31(4):797-810.

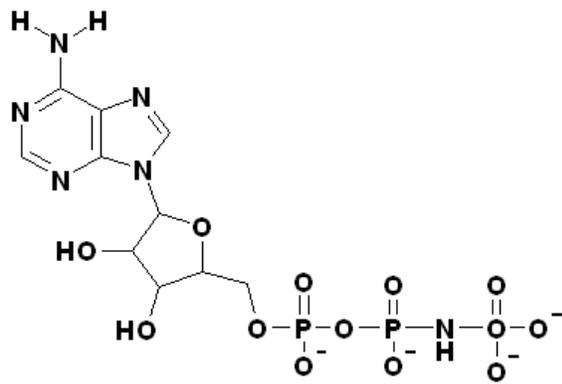
111. Bag S, Tawari NR, Degani MS, & Queener SF (2010) Design, synthesis, biological evaluation and computational investigation of novel inhibitors of dihydrofolate reductase of opportunistic pathogens. *Bioorganic & Medicinal Chemistry* 18(9):3187-3197.
112. Henriksen ST, Liu J, Estiu G, Oltvai ZN, & Wiest O (2010) Identification of novel bacterial histidine biosynthesis inhibitors using docking, ensemble rescoring, and whole-cell assays. *Bioorganic & Medicinal Chemistry* 18(14):5148-5156.
113. Podvinec M, Lim SP, Schmidt T, Scarsi M, Wen D, Sonntag L-S, Sanschagrín P, Shenkin PS, & Schwede T (2010) Novel Inhibitors of Dengue Virus Methyltransferase: Discovery by in Vitro-Driven Virtual Screening on a Desktop Computer Grid. *Journal of Medicinal Chemistry* 53(4):1483-1495.
114. Musicki B, Periers A, Laurin P, Ferroud D, Benedetti Y, Lachaud S, Chatreaux F, Haesslein J, Iltis A, Pierre C, Khider J, Tessot N, Airault M, Demassej J, Dupuis-Hamelin C, Lassaing P, Bonnefoy A, Vicat P, & Klich M (2000) Improved antibacterial activities of coumarin antibiotics bearing 5', 5'-dialkylnoviose: Biological activity of RU79115. *Biorg. Med. Chem. Lett.* 10:1695-1699.
115. Angehrn P, Buchmann S, Funk C, Goetschi E, Gmuender H, Hebeisen P, Kostrewa D, Link H, Luebbers T, Masciadri R, Nielsen J, Reindl P, Ricklin F, Schmitt-Hoffmann A, & Theil FP (2004) New antibacterial agents derived from the DNA gyrase inhibitor cyclothialidine. *J. Med. Chem.* 47:1487-1513.
116. Starr JT, Sciotti RJ, Hanna DL, Huband M, D., Mullins LM, Cai H, Gage JW, Lockard M, Rauckhorst MR, Owen RM, Lall MS, Tomilo M, Chen H, McCurdy SP, & Barbachyn MR (2009) 5-(2-Pyrimidinyl)-imidazol[1,2-a]pyridines are antibacterial agents targeting the

ATPase domains of DNA gyrase and topoisomerase IV. *Biorg. Med. Chem. Lett.* 19:5302-5306.

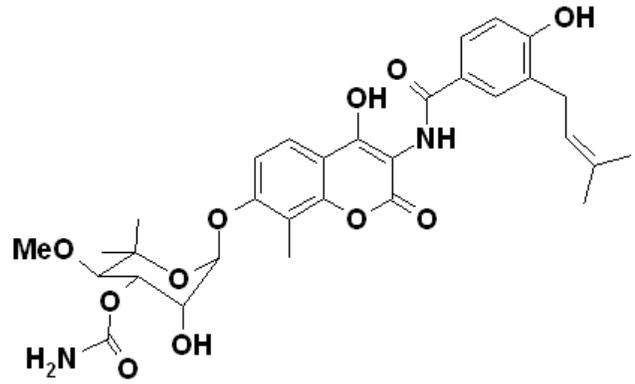
117. East SP, White CB, Barker O, Barker S, Bennett J, Brown D, Boyd EA, Brennan C, Chowdhury C, Collins I, Convers-Reignier E, Dymock BW, Fletcher R, Haydon DJ, Gardiner M, Hatcher S, Ingram P, Lancett P, Mortenson P, Papadopoulos K, Smee C, Thomaides-Brears HB, Tye H, Workman, J., & Czaplewski LG (2009) DNA gyrase (GyrB/topoisomerase IV (pare) inhibitors: Synthesis and antibacterial activity. *Biorg. Med. Chem. Lett.* 19:894-899.

Figures

Figure 1



Adenylyl-imidodiphosphate (ADPNP)



Novobiocin

Figure 2

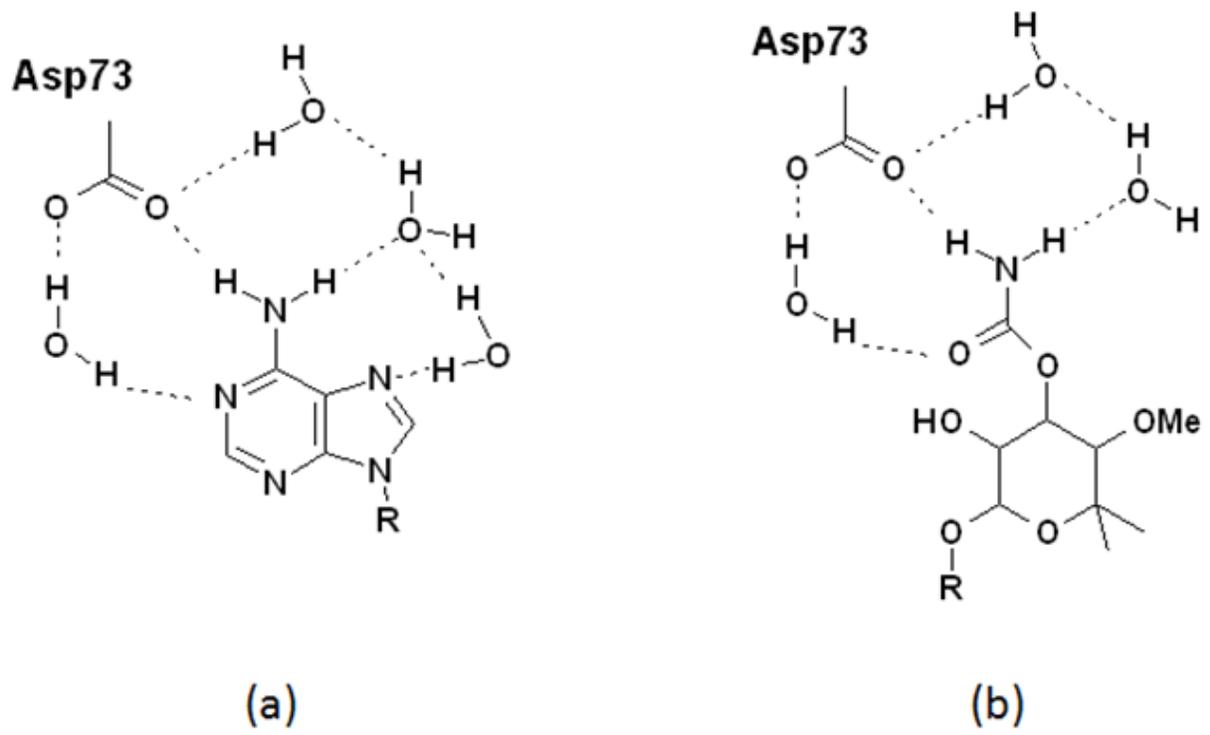


Figure 3

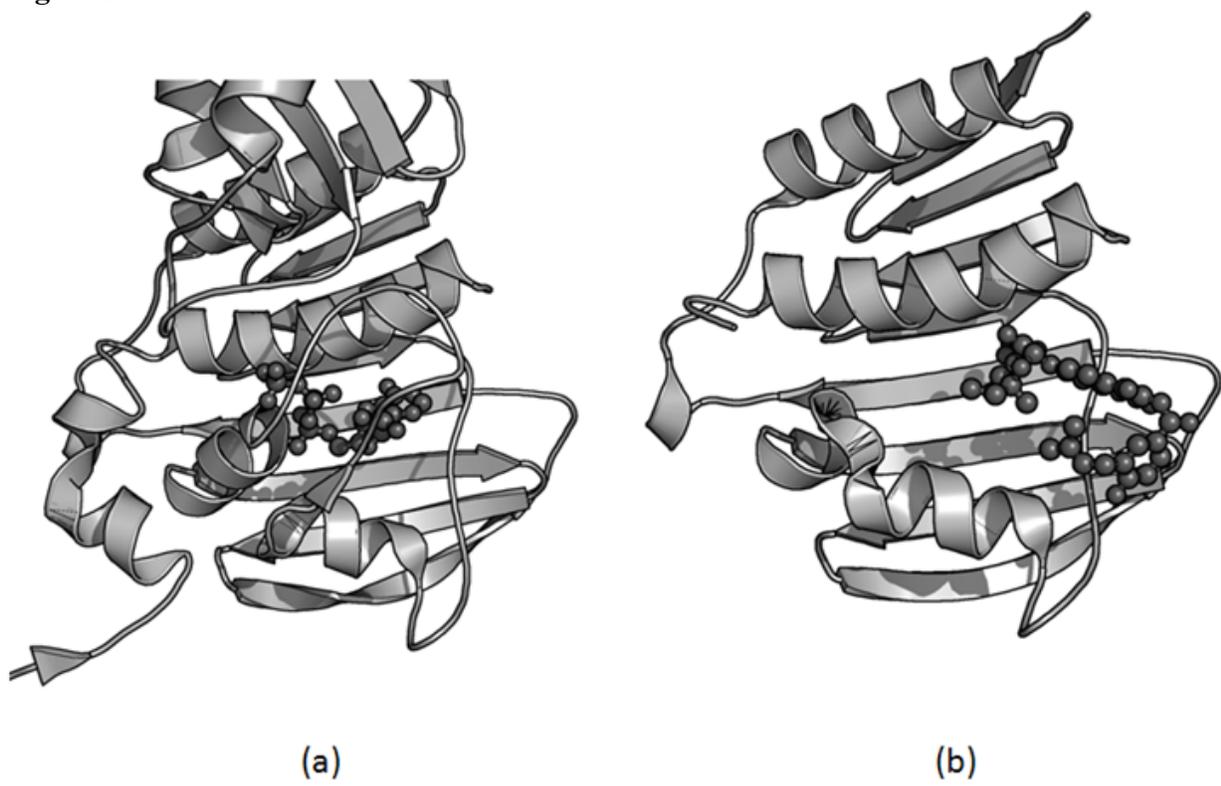


Figure 4

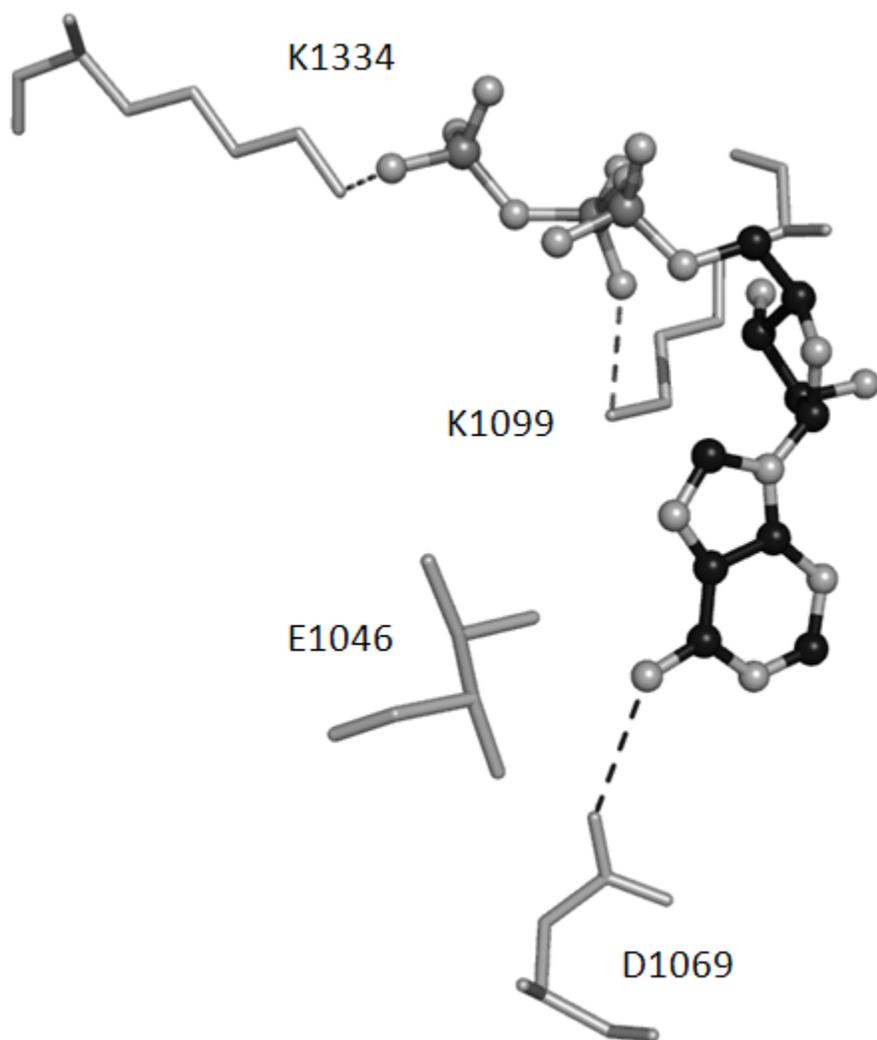


Figure 5

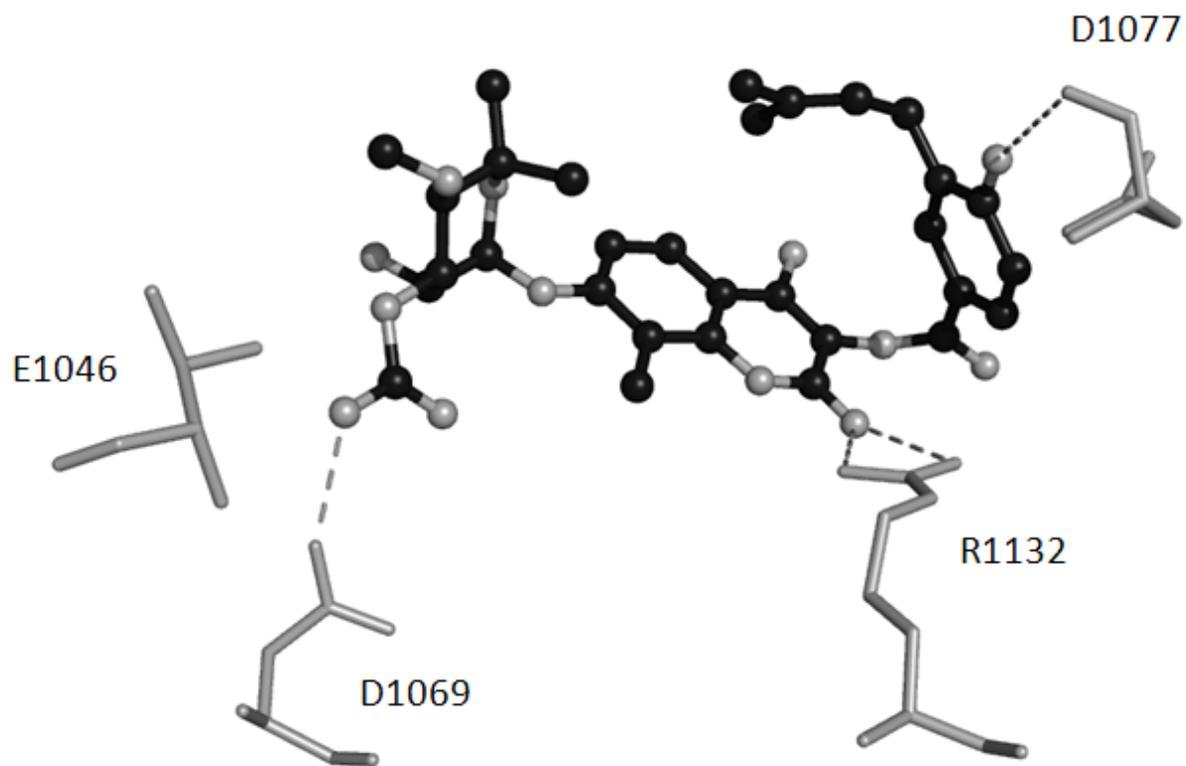


Figure 6

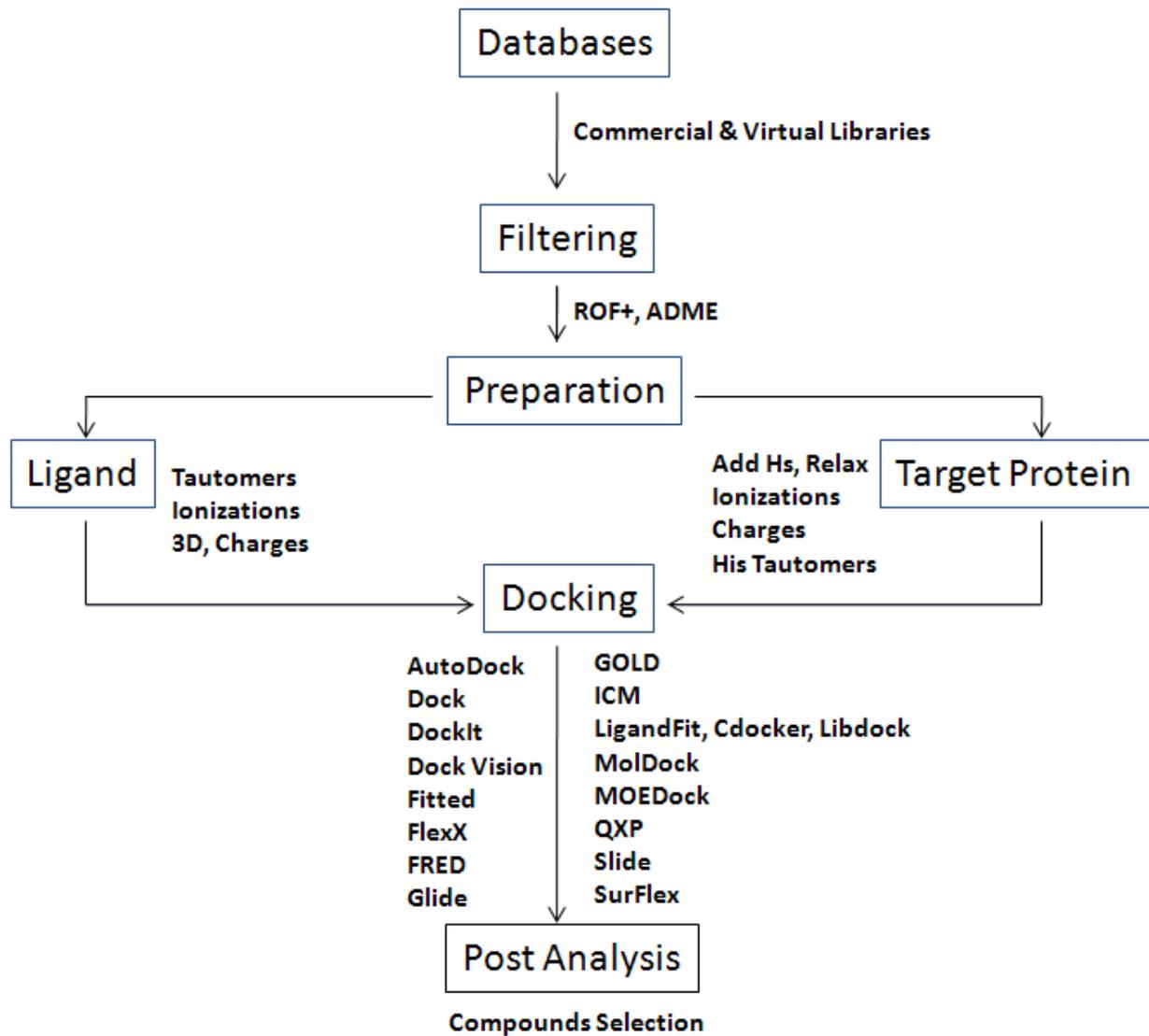


Figure 7

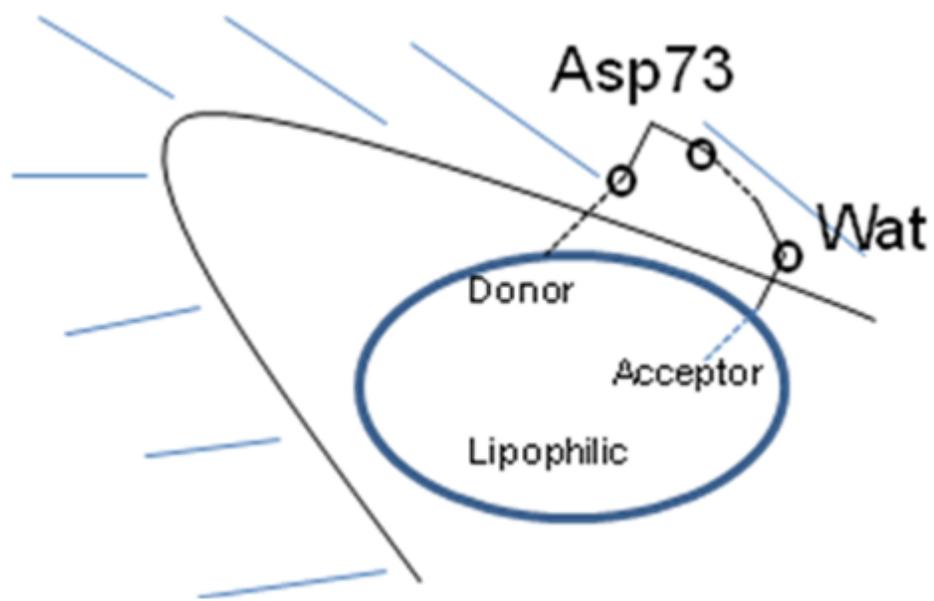
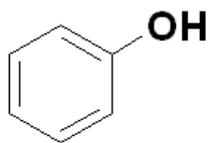
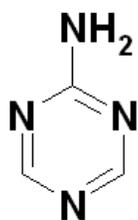


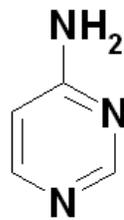
Figure 8



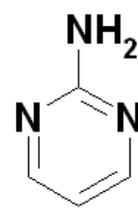
Phenol



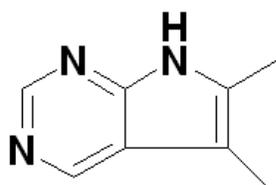
2-NH₂-Triazines



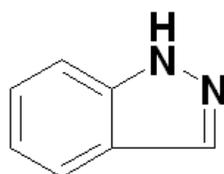
4-NH₂-Pyrimidines



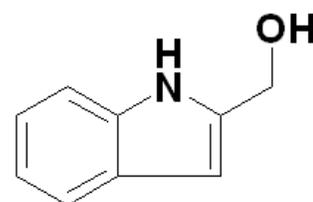
2-NH₂-Pyrimidines



Pyrroloyrimidines



Indazoles



2-HOCH₂-Indoles

Figure 9

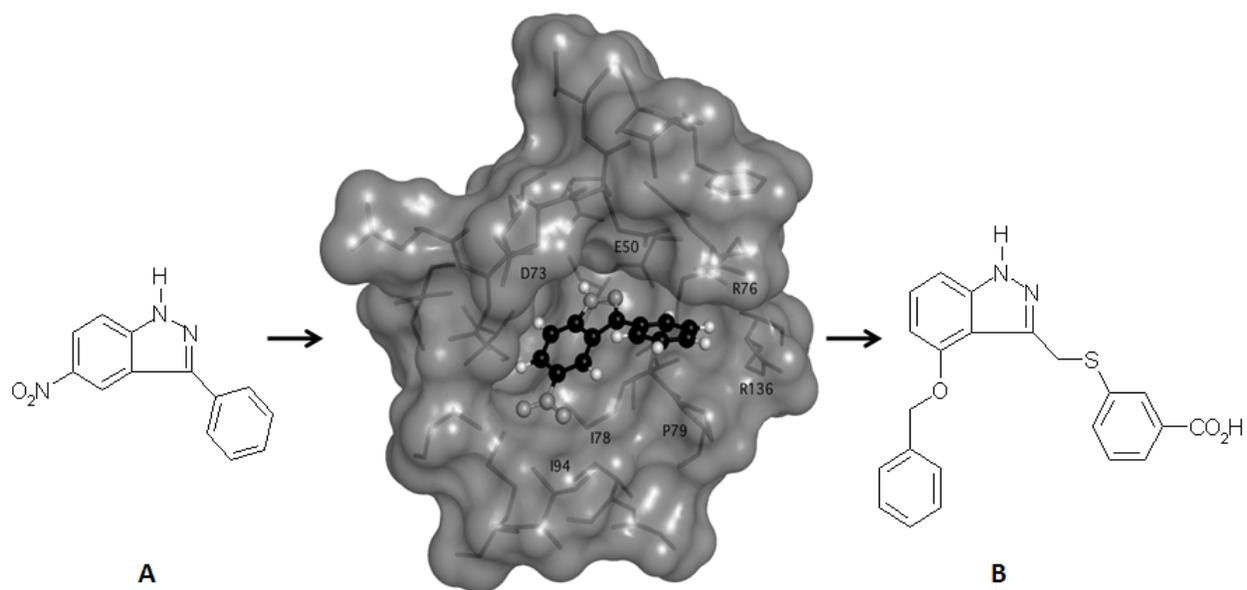


Figure 10

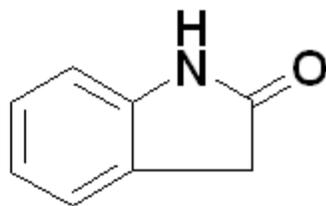
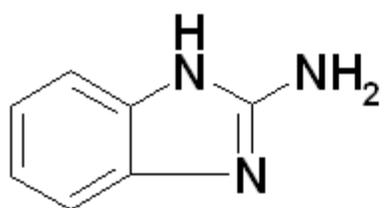


Figure 11

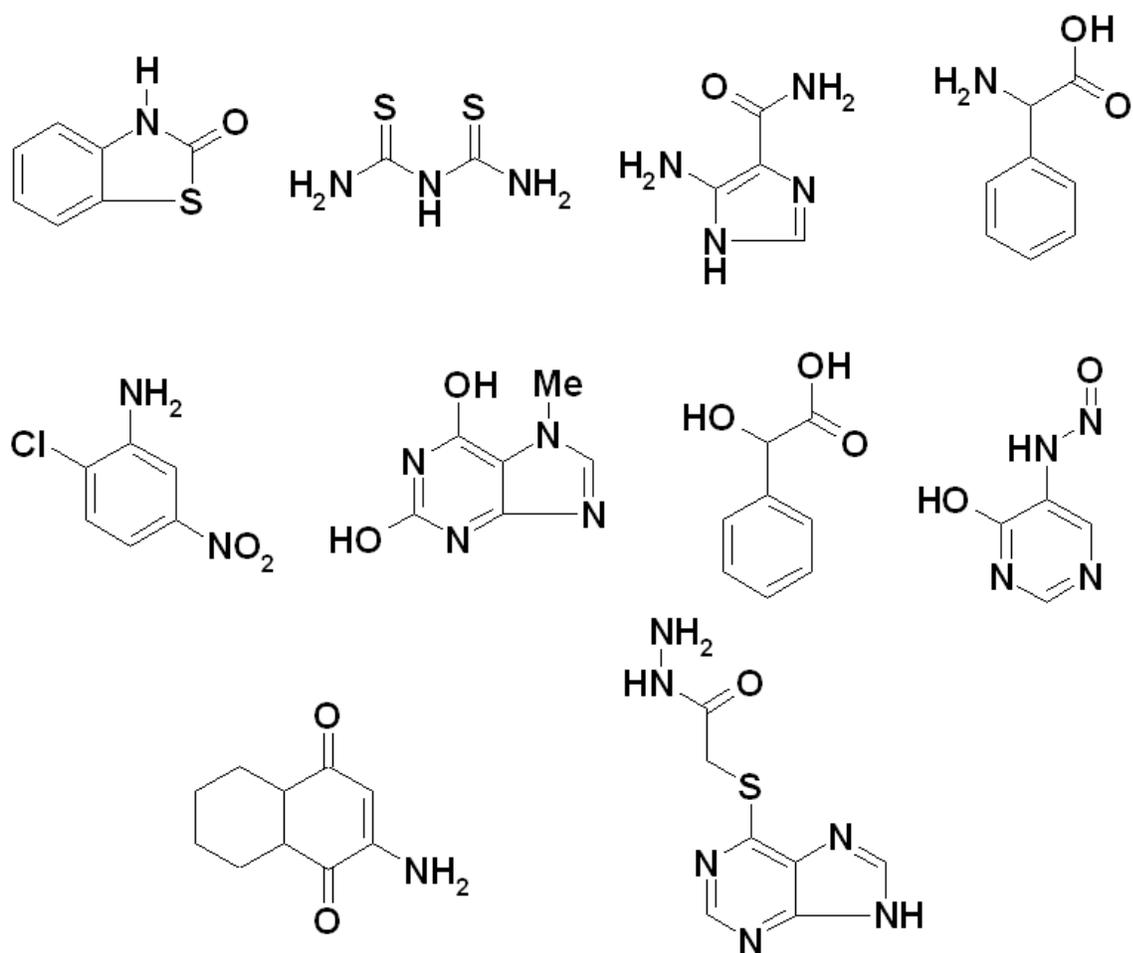


Figure 12

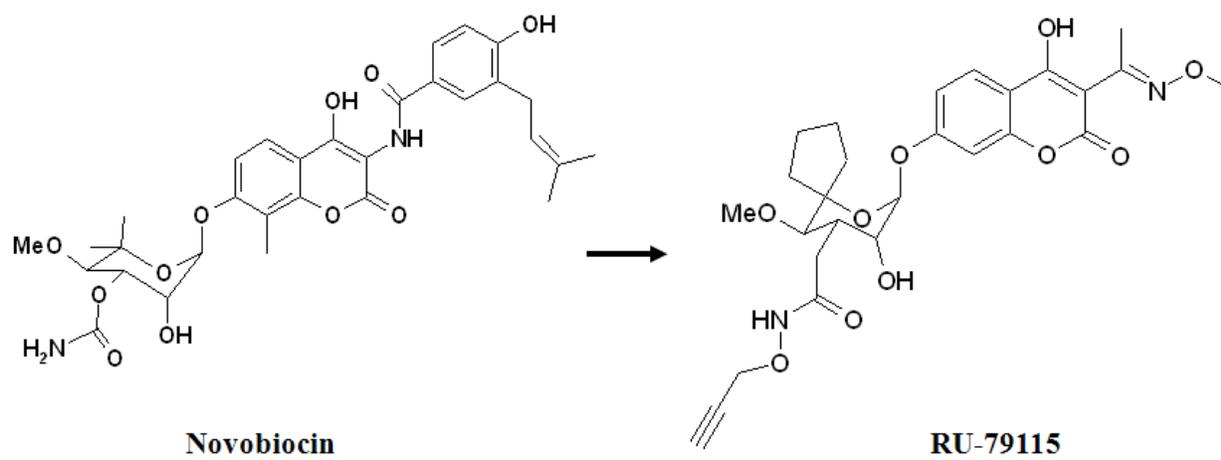


Figure 13

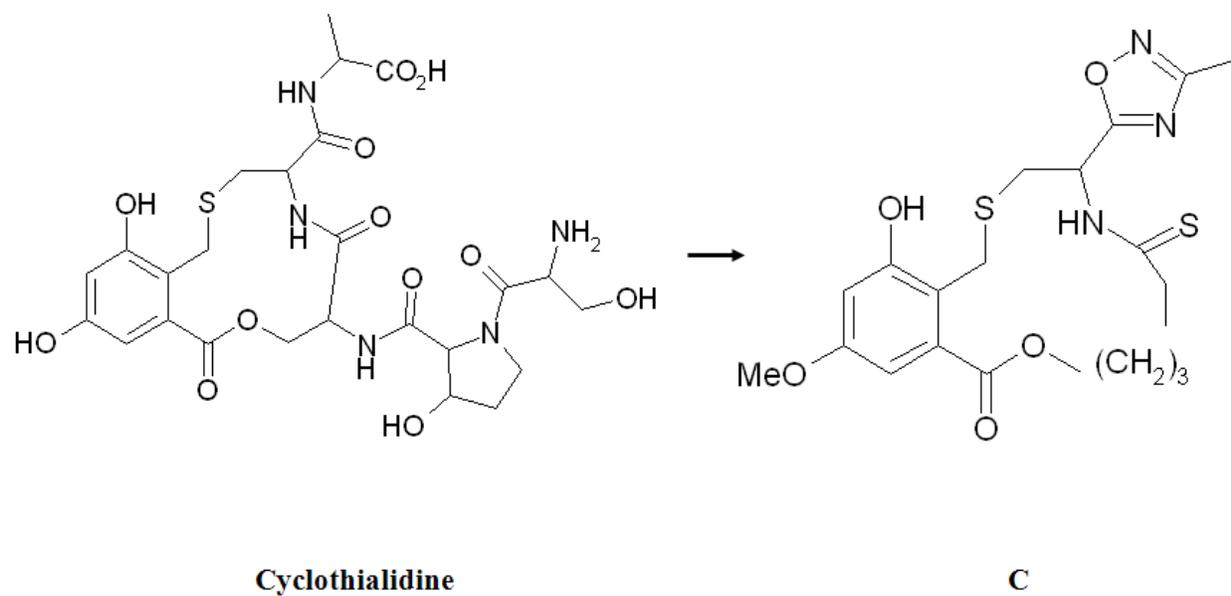
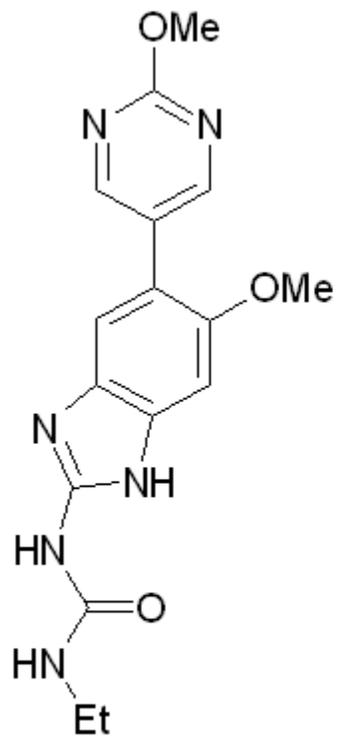
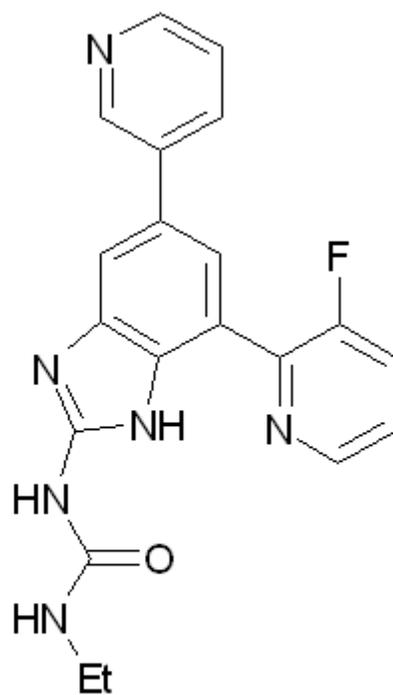


Figure 14



VRT-125853



VRT-752586

Figure 15

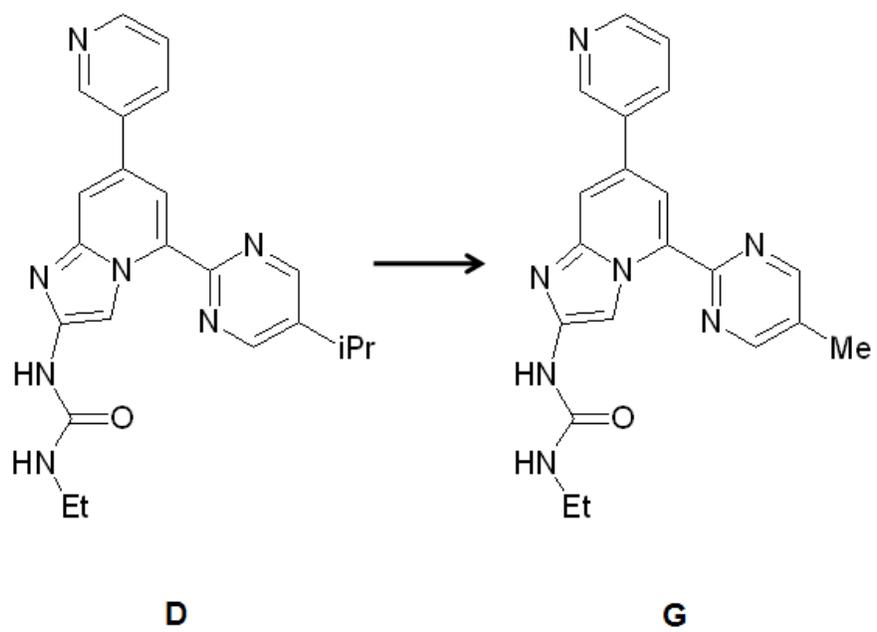


Figure 16

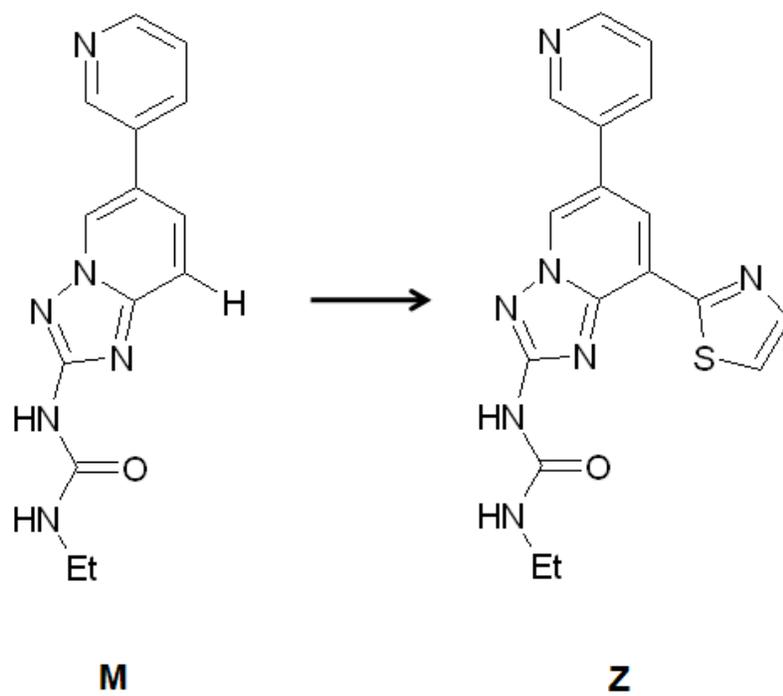


Figure Legends

Figure 1: Chemical structures of the substrate and a natural product inhibitor of type IIA topoisomerases; adenylyl-imidodiphosphate (ADPNP) and novobiocin.

Figure 2: The “anchoring” hydrogen-bond network used in GyrB and ParE to bind: **(a)** the adenine moiety of ADPNP and **(b)** the carbamate moiety of novobiocin.

Figure 3: Crystal structures of *E. coli* ParE. **(a)** A 43-kDa fragment (PDB code 1S16) complexed with ADPNP at 2.0 Å resolution. **(b)** A 24-kDa fragment (PDB code 1S14) complexed with novobiocin at 2.1 Å resolution (27).

Figure 4: The ATP binding site of *E. coli* ParE with ADPNP drawn in balls and sticks, and the protein shown in cylinders with key amino acids side chains labeled. The illustration shows the hydrogen-bond interactions between the adenine of ADPNP with the key amino acid D1069 and its orientation with respect to the binding environment - the polar phosphate group from the tail end of the ADPNP hydrogen bonds with K1099 & K1334.

Figure 5: The novobiocin binding mode from the crystal structure of *E. coli* ParE with novobiocin shown in balls and sticks and key amino acid side chains labeled. The illustration shows the hydrogen-bond interactions of the carbamate moiety of novobiocin with the key amino acid D1069 and its orientation with respect to the binding environment - the carbonyl-ester of the middle ring and the hydroxyl of the tail end of novobiocin hydrogen-bonds with R1132 and D1077, respectively.

Figure 6: An illustration of a typical workflow used for *in silico* screening.

Figure 7: An illustration of the binding hypothesis used for *in silico* screening.

Figure 8: Seven classes of lead compounds as novel inhibitors of DNA GyrB (26).

Figure 9: Structure-guided design of indazole B for maximal inhibitor-receptor interactions (26).

Figure 10: Chemical structures of DNA GyrB lead compounds; aminobenzimidazole and indolin-2-one, respectively (105).

Figure 11: Chemical structures identified by virtual screening using DOCK (107).

Figure 12: Potent inhibitor (RU-79115) optimized from the natural product novobiocin (114).

Figure 13: A design of compound C from structural guidance and SAR (115).

Figure 14: Chemical structures of aminobenzimidazoles with dual *E. coli* GyrB/ParE activity.

Figure 15: Chemical structures of imidazolo-pyridines reported by the Pfizer team.

Figure 16: Chemical structures of triazolo-pyridines reported by the Evotec team.