Functionality Maps of the ATP Binding Site of DNA Gyrase B: Generation of a Consensus Model of Ligand Binding

Martina Schechner,[†] Finton Sirockin,[†,§] Roland H. Stote,*[,‡] and Annick P. Dejaegere*[,†]

Département de Biologie et Génomique Structurales, IGBMC UMR 7104, Ecole Supérieure de Biotechnologie de Strasbourg, Boulevard S. Brant BP 10413, F-67412 Illkirch Cedex, France, and Laboratoire de Chimie Biophysique, ISIS UMR 7006, Université Louis Pasteur, 8, Allée Gaspard Monge BP 70028, F-67083 Strasbourg Cedex, France

Received December 1, 2003

The Multiple Copy Simultaneous Search method (MCSS) was used to construct consensus functionality maps for functional group binding in the ATP binding site of DNA gyrase B. To account for the conformational flexibility of the protein active site, which involves small side chain fluctuations as well as large-scale loop motions, the calculations were done for three different conformations of the 24 kDa subdomain of DNA gyrase B. A postprocessing procedure that employs a continuum dielectric model to include solvent effects was used to calculate the binding free energy for every functional group. These results were ranked according to their affinity for DNA gyrase B and clustered using a new procedure based on van der Waals contacts that is better adapted for cases where multiple conformations are being considered. A total of 23 different functional groups were tested. The results gave consensus maps that indicate those functional group binding sites that are insensitive to the specific protein conformation. The maps also demonstrate that functional groups other than those found in the known ligands may bind competitively in the binding sites of known ligands. This suggests numerous scaffolds that can be used in the development of new ligands for the ATP and coumarinic binding sites in DNA gyrase B. Finally, the calculations show the existence of alternative binding sites near the known binding sites that could be targeted in the rational design for new inhibitors.

I. Introduction

DNA gyrase is a bacterial type II topoisomerase specific to prokaryotes such as *Escherichia coli, Staphylococcus aureus, Bacillus subtilis, Borrelia burgdorferi,* and *Pseudomonas putida*[1] that catalyzes the negative supercoiling of double-stranded circular DNA.[2] It belongs to a family of ATPases known as GHKL;[3] within this family of proteins are several important pharmaceutical targets for anti-infectious and cancer treatments. While similar to eukaryotic topoisomerases II, DNA gyrase shows enough differences so that certain antibiotics act specifically against it.[4] As a consequence, DNA gyrase is an important target for the development of new antibacterial therapeutics.[5]

Coumarins are natural compounds that inhibit the ATPase activity of DNA gyrase B; novobiocin and chlorobiocin are the most prominent members of this class of compounds (cf. Figure 1). Their K_i values range from 10^{-7} to 10^{-9} M, making these compounds potent inhibitors. However, they have not been very successful as drugs because of several problems. First, they have low solubility in water. Second, because of very poor permeability, they do not act very well against Gram-negative bacteria, and finally, they are toxic in eukaryotes.[6] In addition, coumarin-sensitive organisms develop resistance because of their ability to transfer the carbamoyl group of the novobiose sugar moiety from position 3 to position 2.[7] Because of the emerging resistance to current antibiotics, as well as to their many side effects, there is great interest in designing new ligands for DNA gyrase that function as antibacterial compounds.[6]

A number of experimental crystal structures of the ATPase domain of DNA gyrase are known and therefore can be used in a rational approach for the development of new ligands. Structures of the 43 kDa domain of DNA gyrase B in complex with the ATP analogue 5′-adenylyl-β-γ-imidodiphosphate (ADPNP)[10,11] as well as the structure of the 24 kDa subunit in complex with different inhibitors have been determined.[12,13] A comparison of the different crystal structures shows that the 24 kDa subunit (referred to in this work as p24) contains two loops that are implicated in the ATPase activity. In Figure 2a, it can be seen that the substrate and the inhibitor partially overlap each other and that the structures of the two complexes are very similar with two exceptions. First, residues 83−86, which make up one of the active site loops, were unresolved in the structure of the chlorobiocin complex, and second, residues 99−119, which make up the second loop, are in a different conformational state relative to their conformation in the 43 kDa domain/ADPNP complex. In the structure with ADPNP, this loop is folded over the active site, while in the chlorobiocin structure, this loop is in an extended conformation.

Essential for structure-based drug design is an understanding of the physicochemical basis of molecular recognition including the effects of large-scale conformational changes on ligand binding. Through the

* To whom correspondence should be addressed. For R.H.S.: phone, +33 3-90-245-128; fax, +33 3-90-245-126; e-mail, rstote@isis.u-strasbg.fr. For A.P.D.: phone, +33 3-90-244-721; fax, +33 3-90-244-718; e-mail, annick@esbs.u-strasbg.fr.

[†] Ecole Supérieure de Biotechnologie de Strasbourg.

[‡] Université Louis Pasteur. UMR 7006 du CNRS.

[§] Current address: Novartis Pharm AG, Postfach CH-4002, Basel, Switzerland.

10.1021/jm0311184 CCC: $27.50 © 2004 American Chemical Society
Published on Web 07/29/2004

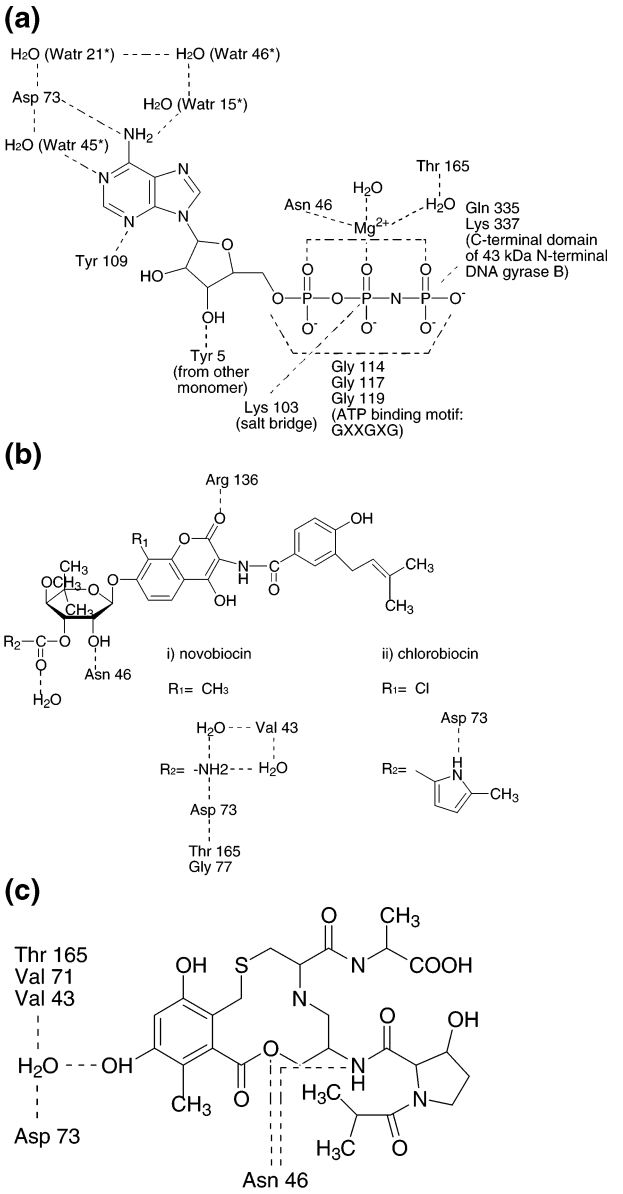

Figure 1. Interactions of the DNA gyrase B ligands. Dashed lines represent hydrogen bonds except when indicated differently: (a) ATP; (b) coumarins; (i) novobiocin; (ii) chlorobiocin. The pyrrole moiety of chlorobiocin makes additional hydrophobic contacts with Val43, Ala47, Val71, Asp73, Ile78, Val120, Thr165, and Val167[13]). (c) Cyclothialidine GR122222X.[7] Water molecules indicated with an asterisk (∗) are numbered according to ref 31.

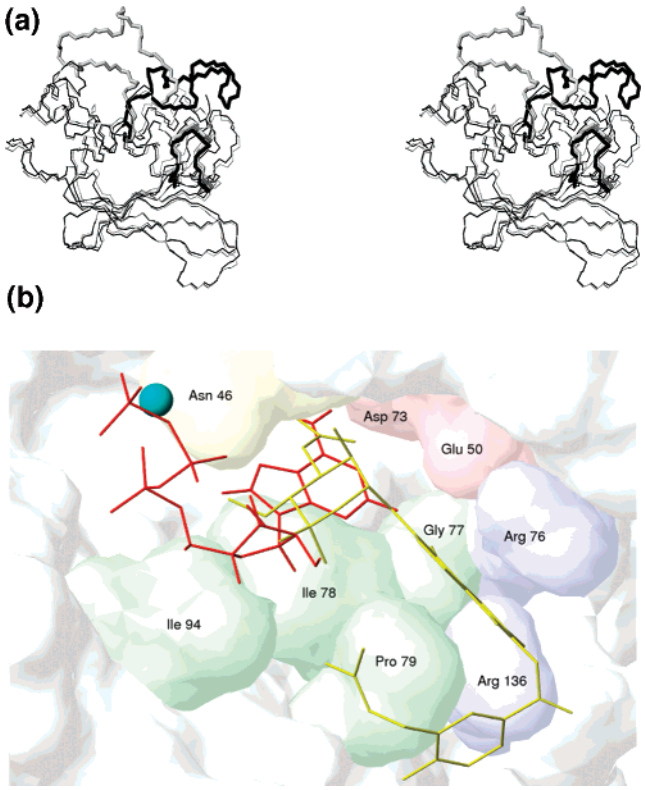

Figure 2. (a) Stereo representation of the 24 kDa subdomain of DNA gyrase B. The "closed" crystal structure is shown in black, and the "open" crystal structure from the complex with chlorobiocin is shown in gray. The first and second loops are represented as thick lines. In the open conformation, the active site is accessible to ligand binding. Superposition was done on all atoms except for loop residues 80−90 and 99−120. The ligands ADPNP and chlorobiocin have been omitted for clarity. (b) Solvent accessible surface of the active site of the ATP/coumarin binding site in the 24 kDa subdomain of DNA gyrase B (cf. part a). This orientation is used in all subsequent figures. Magnesium is shown as a cyan sphere, ATP is in red, and chlorobiocin is in yellow. Important active site residues of DNA gyrase B are labeled by their residue name and sequence number.

widespread use of in silico methods for lead identification and optimization, factors that were routinely neglected in early studies (such as solvation and protein flexibility) are now being recognized as important, and efforts are being made to incorporate them in computational studies.[14−16]

In this paper, we study the binding of small ligands to the *E. coli* DNA gyrase (EC 5.99.1.3) and explicitly consider protein flexibility as well as solvent effects in the docking procedure. The results obtained give a consensus pharmacophore mapping for the gyrase ATP binding site that can be used for lead identification or optimization of existing compounds. We use the Multiple Copy Simultaneous Search (MCSS) method to construct a consensus map for functional group binding on the ATP binding site of DNA gyrase. Several refinements were incorporated into the standard approach. First, a more precise determination of the protonation state of DNA gyrase under physiological conditions was made using a continuum dielectric model and finite difference techniques prior to the MCSS calculations. Second, to account for the multiple conformations of the active site loops, the calculations were repeated for three different conformations instead of for a single (experimental) structure. Multiple structures can be obtained from different crystal structures taken from the Protein Data Bank (http://www.pdb.org)[17] or from molecular dynamics simulations. This removes artifacts that may result from the use of a single conformation in the docking procedure. In the present case, the differences in conformation lie mainly in the large-scale differences in the loop conformations around the active site. And third, a postprocessing procedure using continuum electrostatics to account for solvent effects was used to calculate the free energy of binding of every functional group to its minimum energy position as determined by the MCSS calculations. This provides a better approximation to the binding free energy of the functional groups. This model was recently tested in a side-by-side comparison to results from NMR

experiments on FKBP2.[18] The results demonstrate that the use of this solvation model removes many false minima and generates results in excellent agreement with experiment.

Using these refinements, we were able to produce consensus binding site maps of the ATP binding site that identified functional group binding sites that are less sensitive to both small and large-scale conformational changes. This is an important consideration in the design of new ligands. The results also suggest numerous small-molecule compounds that can be used as scaffolds in the design of new ligands. These compounds were shown to have good binding affinities in and around the ATP binding site of DNA gyrase B. Finally, from the consensus functionality maps, a new hydrophobic binding site that is relatively independent of the active site conformation was located adjacent to the binding sites of the known ligands. These calculations provide detailed information on the binding of pharmacologically relevant ligands to the ATP binding site. These results can be used in the development of new binding agents for DNA gyrase B.

II. Methods

The MCSS method allows one to get exhaustive information of possible binding sites and orientations for small chemical or functional groups in a known protein structure.[19] The functional groups used are chemical fragments of larger organic molecules with different chemical character, such as charged, polar, aromatic, hydrophobic, and aliphatic groups. The result of such an analysis is a map of the active site that indicates regions of favorable interactions for the different functional groups, which in turn could suggest possible directions for lead compound screening or optimization. The MCSS method has been described in detail in earlier publications.[19,20] In the present manuscript, we discuss those aspects that are particular to the calculations done here on the DNA gyrase.

Structures. The present MCSS calculations were done using three different structures of the p24 subunit. This permits one to calculate a consensus functionality map that accounts for conformational changes in the active site due to protein flexibility. In all calculations, only residues 12−220, of the p24 subunit (ATP binding site) were used. The first 11 residues were not included in these calculations because this segment extends over to the other monomer during dimerization overlapping the active site, but since dimerization is inhibited by most known inhibitors, these inhibitors do not interact with this segment. The structure of the p24 subunit was extracted from the 2.0 Å resolution structure of the 43 kDa domain of gyrase B complexed to ADPNP complex determined by Wigley and co-workers.[10] In this structure, the positions of both flexible loops were determined; however, the position of the Mg^{2+} ion, known to be important for binding, was absent. Coordinates for the Mg^{2+} ion were obtained from the 2.3 Å resolution structure of the 43 kDa gyrase/ADPNP determined by Brino et al.[11] The Mg^{2+} ion was inserted into the structure of Wigley and co-workers. This structure was energy-minimized and used in a molecular dynamics simulation as described below (Molecular Dynamics Simulations). The initial and final structures from the simulation were subsequently used in the MCSS study and are referred to as XATP and MDATP, respectively. The third structure used in the MCSS study was the 1.9 Å resolution structure of the p24 subunit complexed to chlorobiocin[13] and is referred to as XCLO. This structure was missing coordinates for one of the flexible loops (residues 83−86). This loop was modeled into the structure by superposing the 43 kDa domain/ADPNP structure and the p24/chlorobiocin structures. The missing coordinates for residues Pro83, Glu84, Glu85, and Gly86 were then inserted into the chlorobiocin structure. Comparing the structure of the modeled loop to its structure in the recently determined structure of the 43 kDa subunit complexed with novobiocin,[21] in which the conformation of this loop conformation has been resolved, shows that the modeled conformation is consistent with the experimental structure. There are two mutations in the sequence of the p24/chlorobiocin structure with respect to the 43 kDa domain/ADPNP structure (Val111Met and Asp198Asn), but these do not affect the calculations presented here.

These structures were chosen for their different active site conformations. In the structure referred to as the XCLO structure, the second active site loop has an open and extended conformation, while the XATP and MDATP structures are both derived from the structure of DNA gyrase B p43 dimer complexed to ADPNP,[10] where the active site loops are folded over the ATP binding site. The crystallographic structures of the p24 domain complexed with inhibitors such as novobiocin (PDB code: 1aj6.pdb[34]) were not used in the present calculations since neither of the two active site loops were structurally resolved and the structure of the complex with cyclothialidine[12] is not available in the Protein Data Bank.

Continuum Dielectric Calculations. The p24 subdomain contains a large number of charged amino acids and 11 His residues. Since the experimental structures were determined at near-neutral pH, it is likely that several His residues may be protonated. A total of five histidines are in or close to the active site and can therefore considerably influence the results of the MCSS calculation if they are not protonated in a physically realistic manner. To better model the physiological state of the protein, macroscopic continuum dielectric models, in which the protein is represented as a low dielectric in an aqueous high dielectric, have been employed to calculate the pK_a values of titrating groups as described in Schaefer et al.[22] By use of these methods, excellent results have been obtained in the prediction of pK_a shifts induced by charge perturbations and in the prediction of absolute pK_a values in small proteins.[23] In this work, a continuum dielectric model was used to calculate the total charge of the protein/ATP/Mg^{2+} complex at pH 7 and the probable protonation states for all the His residues, which can have single or double protonation involving either the Nδ or the Nϵ tautomer. Except for the N- and C- terminal ends, which were included in the pK_a calculation, all other titratable residues were assigned their standard protonation states at pH 7. This calculation employed the UHBD program[24] and the scripts developed by M. Schaefer (personal communication). The calculations used an internal dielectric constant of 4 and an external dielectric constant of 80. The

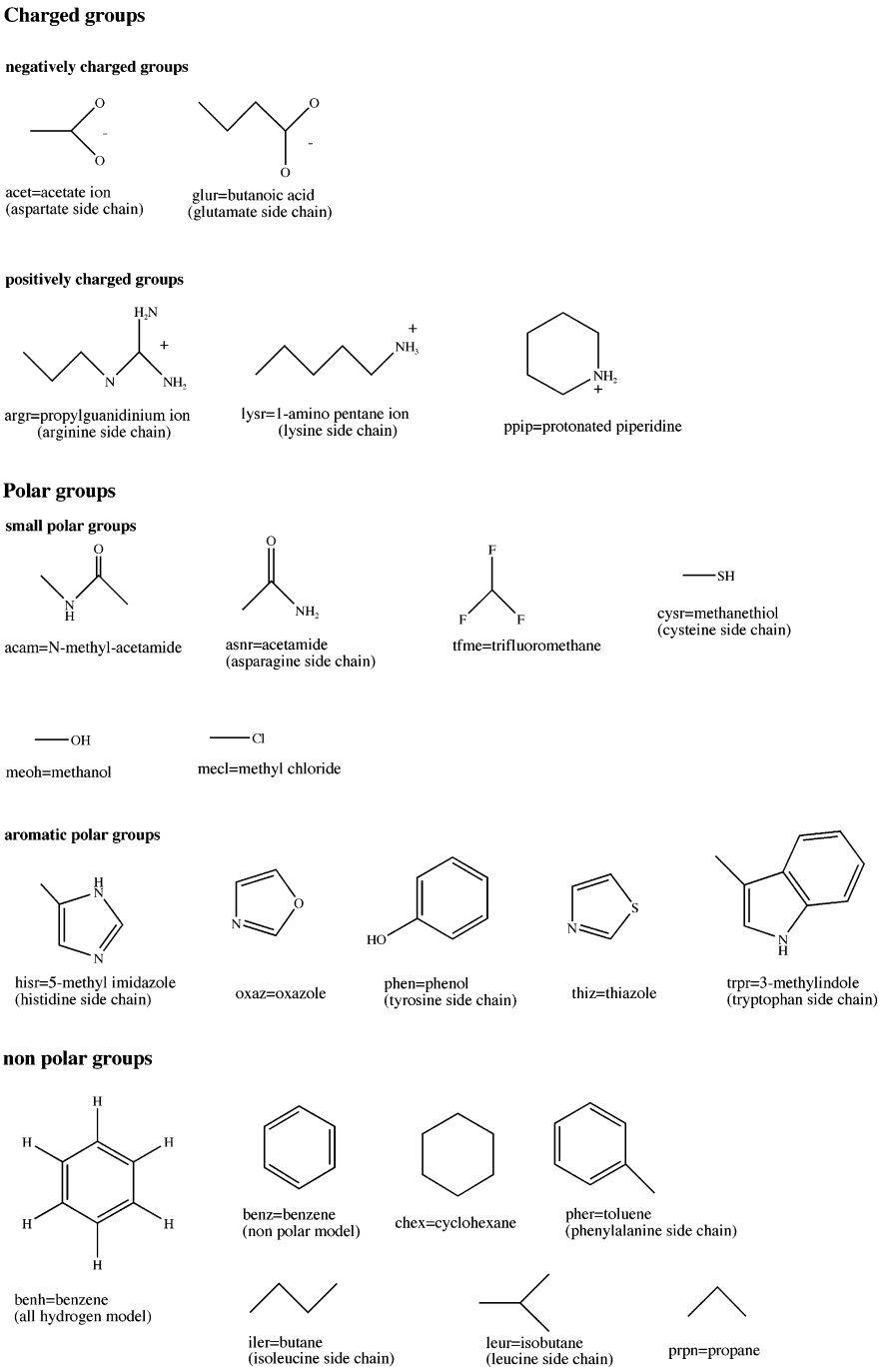

Figure 3. Functional groups used for the MCSS calculations. The abbreviations of the different groups are the same as used in the text.

calculation resulted in the following protonation states of the histidines: histidine 37 was singly protonated at its Nδ1 atom, histidines 38, 99, 116, 147, and 217 were doubly protonated histidines, and histidines 55, 64, 83, 141, and 215 were singly protonated at their Nϵ2 atoms. This protonation state was used in both the molecular dynamics simulations and the MCSS calculations.

Molecular Dynamics Simulations. Molecular dynamics simulations of the p24 subunit complexed to ATP and a Mg^{2+} ion and in the protonation state determined by the continuum dielectric calculations were done using the CHARMM program (version 27[25]) and the all-atom protein−nucleic acid force field.[26] The calculation was limited to the p24 subunit in order to reduce the computational requirements by minimizing the number of atoms in the simulation. The simulations were done primarily to generate an alternative conformation for the MCSS calculations. Hydrogen atoms were added using the HBUILD facility[27] in the CHARMM program. Harmonic constraints were placed on the protein heavy atoms, and the entire system was energy-minimized with 1000 steps of steepest descent minimization. A distance-dependent dielectric constant of $4r$ (r = interatomic distance) and a nonbond cutoff of 13.0 Å were used for the minimization. For the molecular dynamics simulation, the protein was solvated in a 56 Å square box of TIP3P water, periodic boundary conditions were imposed, and a dielectric constant of 1 was used in the treatment of electrostatic interactions. The SHAKE algorithm was used to constrain hydrogen

Table 1. Summary of Results for the Functional Groups Used for the MCSS Calculations[a]

group[b]	no. of minima[c]	$E^{solv\ elec}$ [d] (kcal/mol)	CHARMM energy[e] (kcal/mol)		$\Delta G^{binding}$ [f] (kcal/mol)				minima with favorable $\Delta G^{binding}$ (<0)[f]
			lowest	highest	lowest	second	third	highest	
(1) XATP Structure									
nonpolar groups									
benh	150	−2.0	−22.1	0.0	−11.5	−11.0	−9.6	34.4	120
benz	179	0.0	−12.3	−2.6	−12.1	−11.3	−10.6	31.3	141
chex	472	0.0	−11.7	3.0	−12.1	−10.6	−9.9	42.9	223
iler	720	0.0	−10.4	−1.7	−11.0	−10.9	−10.9	44.5	478
leur	422	0.0	−10.1	−2.0	−11.3	−10.2	−10.0	36.5	309
pher	647	0.0	−32.9	−3.2	−14.1	−13.7	−13.7	41.9	483
prpn	471	0.0	−8.8	1.2	−10.3	−10.1	−9.7	35.3	346
polar groups									
acam	478	−9.1	−28.2	3.0	−12.1	−11.7	−11.6	23.6	416
asnr	498	−11.5	−24.8	2.8	−12.3	−11.9	−11.6	22.7	423
cysr	145	−1.1	−9.8	−0.1	−6.5	−6.4	−6.4	22.7	92
mecl	145	−2.2	−15.4	1.5	−7.8	−7.7	−7.6	17.1	112
meoh	299	−7.5	−26.8	1.7	−14.5	−12.3	−11.5	19.5	239
tfme	403	−2.2	−20.7	−0.1	−13.9	−13.8	−13.7	19.2	365
aromatic polar groups									
hisr	498	−5.5	−46.9	1.9	−13.5	−12.0	−11.9	25.9	417
oxaz	874	−3.2	−31.7	2.9	−11.2	−11.0	−10.8	27.2	690
phen	581	−7.4	−39.5	2.2	−17.6	−16.7	−13.4	38.9	446
thiz	660	−2.9	−32.9	2.3	−11.1	−10.9	−10.7	23.1	507
trpr	1031	−3.2	−24.9	−0.1	−16.1	−15.8	−15.3	44.8	755
positively charged groups									
argr	1250	−95.9	−68.8	2.7	−13.0	−13.0	−12.3	105.7	383
lysr	1008	−100.2	−57.9	1.6	−6.7	−5.6	−5.1	104.3	67
ppip	533	−73.6	−50.6	2.7	−8.7	−8.4	−8.4	92.8	139
negatively charged groups									
acet	150	−71.7	−58.4	1.2	−14.1	−12.7	−12.5	74.6	72
glur	429	−71.4	−62.0	0.6	−9.3	−9.3	−8.8	83.4	211
(2) MDATP Structure									
nonpolar groups									
benh	65	−2.0	−13.0	3.0	−11.7	−11.6	−10.8	32.5	46
benz	85	0.0	−12.8	−3.2	−11.5	−11.5	−11.3	5.8	64
prpn	248	0.0	−8.5	2.6	−10.1	−9.9	−9.6	3.6	173
polar groups									
acam	310	−9.1	−35.5	2.7	−14.4	−12.6	−12.5	5.7	223
asnr	311	−11.5	−29.3	2.9	−12.4	−11.1	−9.9	27.2	241
meoh	220	−7.5	−25.6	2.7	−10.7	−8.9	−8.7	10.9	167
tfme	240	−2.2	−19.6	2.1	−16.5	−14.9	−13.3	18.7	214
aromatic polar groups									
oxaz	350	−3.2	−21.3	2.7	−11.9	−11.0	−11.0	22.2	294
phen	362	−7.4	−27.5	0.7	−12.6	−11.1	−11.1	34.7	254
thiz	340	−2.9	−20.6	2.4	−11.7	−11.0	−10.4	22.4	265
positively charged groups									
lysr	995	−100.2	−62.2	2.9	−7.3	−7.0	−5.7	111.4	26
negatively charged groups									
acet	120	−71.7	−54.6	2.7	−9.3	−9.1	−8.0	67.2	49
(3) XCLO Structure									
nonpolar groups									
benz	36	0.0	−11.5	1.4	−10.8	−10.4	−10.0	11.6	23
prpn	79	0.0	−8.3	−2.5	−9.4	−9.1	−8.7	12.8	54
polar groups									
acam	114	−9.0	−49.8	−0.7	−15.2	−14.4	−12.5	14.9	93
aromatic polar groups									
oxaz	115	−3.1	−19.2	1.2	−12.7	−12.2	−11.5	12.5	88
phen	86	−7.1	−37.8	−0.5	−18.3	−12.0	−11.4	8.1	61
thiz	105	−2.9	−18.8	−2.8	−10.2	−9.9	−9.5	7.7	78
trpr	101	−3.1	−22.9	−5.3	−13.3	−12.5	−12.4	5.4	81
positively charged groups									
lysr	424	−100.2	5.4	−2.2	−2.4	−2.3	−1.2	66.7	7
negatively charged groups									
acet	58	−71.7	−48.4	1.0	−9.6	−8.7	−8.3	60.2	15
glur	74	−71.4	−46.6	−0.5	−9.8	−9.7	−9.6	68.3	20

[a] Results are given for the XATP (1), MDATP (2), and XCLO (3) structures. [b] See Figure 3 for the functional group definition and structure. [c] Number of minima that have an energy lower than 3 kcal/mol when the energy is evaluated using the CHARMM energy function (see text). [d] Solvation free energy of the ligand calculated by numerical solution of the linear Poisson−Boltzmann equation. [e] The range of energies within the collection of minima calculated using the CHARMM energy function. [f] Binding free energy calculated using eq 3 (see text). The range of values is indicated as the lowest three free energies and the highest binding free energy within the collection of minima.

heavy atom bond distances. The water was first equilibrated for 20 ps at 300 K with the protein−ligand complex fixed, the constraints were removed, and the entire system was heated from 150 to 300 K for 10 ps and again equilibrated for 20 ps. The simulation continued for another 500 ps at 300 K. The simulations

contain 17 178 atoms, including 3255 atoms for the protein/ligand/ions and 13 923 atoms for the water contained in the 56 Å box. Two configurations were extracted from the molecular dynamics simulation for the MCSS calculations. The first one was the energy-minimized crystal structure used for the initial conformation for the molecular dynamics simulations and having the protonation state determined by the continuum dielectric calculations described above. The second structure was taken from the end of the 500 ps production run of the molecular dynamics simulation. The rmsd of all backbone atoms between these two structures was 1.37 and 2.06 Å for the side chain atoms. These two structures, as well as the structure derived from the crystal structure of the DNA gyrase chlorobiocin structure, were prepared for the MCSS calculations as described in the next section.

MCSS Calculations. Preparation of the Protein Structures Used for MCSS Calculations. The structures used in the MCSS calculations are described above. The PARM19 parameters and topologies for the polar hydrogen parameter set[28] were used for the MCSS calculations, and the protonation states for the histidine residues were those that were determined for the p24/ADPNP structure (see Continuum Dielectric Calculations). The structures were prepared in the following way. All water molecules, the magnesium ion, the ATP or the coumarin inhibitor, and the hydrogen atoms were removed. The polar hydrogen atoms were added using the HBUILD facility[27] in the CHARMM program.[25] The three structures (XATP, MDATP, and XCLO) were each energy-minimized using the steepest descent algorithm (700 steps) followed by the conjugate gradient algorithm (1000 steps) and used in the MCSS calculations.

The functional groups are initially placed randomly over the active site within a 20 Å sphere centered on Gly 102. This sphere included all residues within the p24 subdomain implicated in ADPNP/ATP, novobiocin, cyclothialidine, and chlorobiocin binding and catalysis. To better accommodate the extended second loop in the XCLO structure, the functional groups were placed randomly in a box with dimensions 17.4 Å × 19.6 Å × 16.5 Å. For each calculation, 10 000 replicas (or copies) of a particular functional group were randomly distributed over the DNA gyrase B subunit using a minimal protein−ligand distance of 1.2 Å in order to avoid overlap of the functional group with the protein. After removal of overlapping functional groups, an rmsd cutoff of 0.2 Å for minima at the same positions was applied and an initial energy cutoff of 500 kcal/mol was used. At every step of the energy minimization, the rmsd criterion for overlapping groups was used and the cutoff energy was reduced gradually to a final value of 3 kcal/mol in the last step.

The positions of the functional groups were simultaneously energy-minimized with 500 steps using the steepest descent algorithm, followed by a second cycle of 300 steps of steepest descent and finishing with 20 cycles of up to 500 steps using the conjugate gradient algorithm. This yields a maximum of 10 800 steps of minimization. For each cycle, the minimization stops when the energy gradient is less than 0.0001 kcal mol^{-1} $Å^{-1}$. This calculation was repeated for each of the functional groups. The protein was fixed for all MCSS calculations. The calculations were done on Pentium based computers running the LINUX operating system, as well as on an SGI ORIGIN2000.

Functional Groups. The functional groups used in the present calculations are shown in Figure 3. The nonpolar groups used are cyclohexane (chex), propane (prpn), and butane (iler) as a model for the isoleucine side chain and isobutane (leur) as a model for the leucine side chain and two models for benzene (benh, benz). With the exception of benzene (benh), none of these groups bear partial charges in the force field, and therefore, they do not interact electrostatically with the protein. However, during postprocessing, the protein desolvation can contribute an electrostatic term to the binding free energy (cf. eq 3). An all-atom model for benzene was added to the standard set of polar hydrogen functional groups in the MCSS library in order to determine whether a 12-site model (benh), where the carbon and hydrogen atoms bear small partial charges (C atom, +0.115 e^-; H atom, −0.115 e^-), would give different results than the reduced-atom 6-site model (benz) that has no partial charges. For the nonpolar group results, we only discuss the benzene groups (6- and 12-site models) in detail, since the functional group maps of the other nonpolar groups superpose with the benzene groups. Because there was a strong similarity among the results for the different functional groups within a particular class of compounds, not all functional groups were tested for all structures (see Table 1 for a listing of groups that were tested for the different protein conformations).

The polar groups encompass all the functional groups that have partial charges and therefore can have electrostatic interactions with the protein while being globally neutral. They cover a large range of compounds including small polar groups such as acetamide (model for the Asn side chain), methanol, methane thiol (Cys side chain model), methyl chloride, and trifluoromethane. The *N*-methylacetamide (acam) was tested for all three structures XATP, MDATP, and XCLO. The other small polar groups, such as acetamide (asnr), trifluoromethane (tfme), and methanol (meoh), were only tested for the XATP and MDATP structures.

Aromatic compounds used in the MCSS calculations include oxazole, thiazole, 5-methylimidazole (which is a model for the histidine side chain), phenol (tyrosine side chain), 3-methylindole (tryptophan side chain), and benzene (which has been discussed in the nonpolar section; see above). With the exception of benzene, these aromatic molecules can form hydrogen bonds.

The negatively charged groups used were acetate ion (acet) and butanoic acid (glur), which mimic the side chains of aspartic and glutamic acid, respectively. The positively charged groups tested were the 1-aminopentane ion (lysr) and the propylguanidinium ion (argr), which mimic lysine and arginine side chains, and a protonated piperidine (ppip) (cf. Figure 3). In the MDATP and XCLO structures only the lysr functional group was examined.

Postprocessing of the MCSS Minima. MCSS calculates a binding energy to rank each functional group minimum. This energy is the sum of an internal group energy term and an interaction energy between

the functional group and the protein, eq 1:[18,19]

$$\Delta E_{\mathrm{MCSS}}^{\mathrm{binding}} = \Delta E^{\mathrm{functional\ group}} + \Delta E_{\mathrm{vdW}}^{\mathrm{intermolecular}} + \Delta E_{\mathrm{electrostatic}}^{\mathrm{intermolecular}} \quad (1)$$

where

$$\Delta E^{\mathrm{functional\ group}} = \Delta E_{\mathrm{bonding}}^{\mathrm{functional\ group}} + \Delta E_{\mathrm{vdW}}^{\mathrm{functional\ group}} + \Delta E_{\mathrm{electrostatic}}^{\mathrm{functional\ group}} \quad (2)$$

The $\Delta E^{\mathrm{functional\ group}}$ term is the change in the internal energy of the fragment upon binding relative to its isolated equilibrium conformation; in the MCSS minimization, the fragments are flexible. $\Delta E_{\mathrm{bonding}}^{\mathrm{functional\ group}}$ is the change in the internal energy (bonds, angles, and torsions). The internal energy terms are included in order to avoid unrealistic group conformations.[20] The $\Delta E_{\mathrm{vdW}}^{\mathrm{functionalgroup}}$ term is the change in intrafragment van der Waals energy, and $\Delta E_{\mathrm{electrostatic}}^{\mathrm{functionalgroup}}$ is the change in intrafragment vacuum Coulombic energy. $\Delta E_{\mathrm{vdW}}^{\mathrm{functionalgroup}}$ is calculated with the standard PARAM19 parameters (cutoff of 8.0 Å and the VSHIFT truncation function); $\Delta E_{\mathrm{electrostatic}}^{\mathrm{functionalgroup}}$ is evaluated using an infinite cutoff and a distance-dependent dielectric of $1r$ (r = interatomic distance) for all groups. The $\Delta E_{\mathrm{MCSS}}^{\mathrm{binding}}$ term includes, in addition to the $\Delta E^{\mathrm{functional\ group}}$ term, the nonbonded interaction terms (van der Waals energy, $\Delta E_{\mathrm{vdW}}^{\mathrm{intermolecular}}$, and electrostatic energy, $\Delta E_{\mathrm{electrostatic}}^{\mathrm{intermolecular}}$) for protein−group interactions.

Following the MCSS calculations using the above energy function, a postprocessing procedure was done where the binding free energies of each functional group minimum was estimated by the following equation:[18,29,30]

$$\Delta G^{\mathrm{binding}} = \Delta E^{\mathrm{functional\ group}} + \Delta E_{\mathrm{vdW}}^{\mathrm{intermolecular}} + \Delta G_{\mathrm{elect}}^{\mathrm{intermolecular}} + \Delta G_{\mathrm{elect.desolvation}}^{\mathrm{protein}} + \Delta G_{\mathrm{elect.desolvation}}^{\mathrm{functional\ group}} + \Delta G_{\mathrm{np}}^{\mathrm{complex}} \quad (3)$$

The first two terms, $\Delta E^{\mathrm{functional\ group}}$ and $\Delta E_{\mathrm{vdW}}^{\mathrm{intermolecular}}$, are calculated as described above (see eqs 1 and 2). The next three terms ($\Delta G_{\mathrm{elect}}^{\mathrm{intermolecular}} + \Delta G_{\mathrm{elect.desolvation}}^{\mathrm{protein}} + \Delta G_{\mathrm{elect.desolvation}}^{\mathrm{functional\ group}}$) take into account the electrostatic free energy of solvation, and the last term, ($\Delta G_{\mathrm{np}}^{\mathrm{complex}}$), accounts for the nonpolar contribution to the free energy of binding.

The electrostatic terms arise from the following facts. The charges on the solute polarize the solvent, thus influencing the electrostatic energy of a molecular assembly in two ways: the interactions between solute partial charges are screened and the solvent reaction field interacts directly with each solute charge (self-energy). The continuum electrostatic free energy of solvation is the sum of the screening effect and the direct interaction of each solute charge with the solvent.[30]

The shielded intermolecular interaction, $\Delta G_{\mathrm{elect}}^{\mathrm{intermolecular}}$, the protein desolvation free energy, $\Delta G_{\mathrm{elect.desolvation}}^{\mathrm{protein}}$, and functional group desolvation free energy, $\Delta G_{\mathrm{elect.desolvation}}^{\mathrm{functional\ group}}$, all contribute to the electrostatic energy term. $\Delta G_{\mathrm{elect.desolvation}}^{\mathrm{protein}}$ and $\Delta G_{\mathrm{elect.desolvation}}^{\mathrm{functional\ group}}$ contain the self (or Born) free energy term and the intramolecular shielding. For completely apolar groups without any partial charges in the PARAM19 force field, $\Delta G_{\mathrm{elect}}^{\mathrm{intermolecular}}$ and $\Delta G_{\mathrm{elect.desolvation}}^{\mathrm{functional\ group}}$ are zero.

The term for the nonpolar contribution, $\Delta G_{\mathrm{np}}^{\mathrm{complex}}$, is the sum of the cavitation energy and solute−solvent van der Waals energy. It is estimated to be proportional to the loss in solvent accessible surface (SAS)[30] in the following way:

$$\Delta G_{\mathrm{np}}^{\mathrm{complex}} = \gamma(\mathrm{SAS}^{\mathrm{complex}} - (\mathrm{SAS}_{\mathrm{isolated}}^{\mathrm{protein}} + \mathrm{SAS}_{\mathrm{isolated}}^{\mathrm{functional\ group}})) \quad (4)$$

The proportional factor γ can be interpreted as the vacuum−water microscopic surface tension. Its assigned value is 0.025 kcal mol^{-1} Å^{-2}.[30] The solvent accessible surfaces of the complex and of the isolated protein and fragment are calculated using the algorithm of Lee and Richards implemented in the CHARMM program. A probe sphere radius of 1.4 Å is used.[30] The electrostatic free energies of solvation were calculated using the UHBD program, which solves the linearized Poisson−Boltzmann (LPB) equation using a finite-difference procedure.[30] The LPB equation gives good estimates of the electrostatic free energies of solvation in macromolecules.[30] Scripts developed by M. Schaefer (personal communication) were used in these calculations.

It should be noted that eq 1 is designed to be valid for the relative ranking of the binding energies of the MCSS minima for a given functional group. It neglects terms, such as the configurational, rotational, and translational entropy loss on binding, that contribute to the relative binding free energies of different ligands and to the absolute binding energy but that are expected to be very similar for a given ligand in different positions.

Clustering Algorithm. After the MCSS calculation and the postprocessing procedure described above, there still remains a large number of functional group minima with favorable free energies of binding. To systematically analyze the large amount of data generated by the calculation, we developed a new clustering algorithm.[18] Rather than using the usual position criterion (rmsd criterion) for clustering the minima, the new algorithm uses an energy criterion where the clustering is done on the basis of the van der Waals interaction energy between the functional group and the protein residues. The van der Waals interaction is short-ranged, so only residues close to a minimum make a significant contribution. This allows the clustering of functional group minima by the protein amino acids with which they interact. Amino acids that make a van der Waals contribution to the interaction energy above a certain user-specified cutoff are included in the so-called "cluster signature" (see Table 2 for examples). This approach is better adapted when considering multiple conformations where the rmsd is less revealing than the vdW contacts between the protein and the ligand. The algorithm will be described in more detail elsewhere (Sirockin et al.; manuscript in preparation).

III. Results and Discussion

Throughout this section, we make frequent reference to the known binding sites identified from the crystal

Table 2. Adenine Binding Site[a]

cluster	ΔG^{bind} (kcal/mol)	cluster signature
		(i) Close to Asp 73 Aromatic Groups
oxazole		
XATP_oxaz 29	−9.2	Asn46 Ala47 Glu50 Asp73 Gly75 Arg76 Gly77 Ile78 Tyr109 Thr165
MDATP_oxaz 30	−8.8	Asn46 Ala47 Glu50 Asp73 Arg76 Gly77 Ile78 Tyr109 Thr165
XCLO_oxaz 5	−11.3	Glu50 Asp73 Gly75 Arg76 Gly77 Ile78 Thr165
phenol		
XATP_phen 2	−16.7	Asn46 Ala47 Glu50 Asp73 Gly75 Gly77 Ile78 Pro79 Lys103 Tyr109 Gly164 Thr165
MDATP_phen 169	−4.1	Asn46 Ala47 Glu50 Asp73 Arg76 Gly77 Ile78 Ile78 Tyr109 Thr165
XCLO_phen 1	−18.3	Asn46 Ala47 Glu50 Asp73 Gly75 Arg76 Gly77 Ile78 Pro79 Gly164 Thr165
thiazole		
XATP_thiz 54	−7.2	Asn46 Glu50 Asp73 Gly75 Arg76 Gly77 Ile78 Tyr109 Thr165
MDATP_thiz 67	−6.8	Asn46 Glu50 Gly77 Ile78 Lys103 Tyr1099
MDATP_thiz 99	−6.0	Val43 Asn46 Ala47 Glu50 Asp73 Gly77 Ile78 Tyr109 Thr165
XCLO_thiz 1	−10.2	Asn46 Glu50 Asp73 Gly75 Arg76 Gly77 Ile78 Thr165
3-methylindole		
XCLO_trpr 13	−10.4	Asn46 Ala47Asp49 Glu5 Asp73 Gly75 Arg76 Gly77 Ile78 Pro79 Ile94 Arg136 Thr165
XCLO_trpr 14	−10.2	Asn46 Ala47 Glu50 Asp73 Gly75 Arg76 Gly77 Ile78 Pro79 Arg136 Thr165
		Small Polar Groups
NMA		
XCLO_acam 1	−15.2	Asn46 Ala47 Asp49 Glu50 Asp73 Gly75 Arg76 Ile78 Pro79
acetamide		
XATP_asnr 9	−9.1	Asn46 Glu50 Asp73 Gly75 Arg76 Ile78 Thr165
methanol		
XATP_meoh 2	−12.3	Gly75 Arg76 Ile78 Thr165
MDATP_meoh 27	−5.4	Asn46 Glu50 Thr165
		(ii) Around C2−H2 of the Adenine Moiety Small Polar Groups
NMA		
XATP_acam 9	−10.4	Asn46 Glu50 Arg76 Gly77 Ile78 Pro79 Lys103 Tyr109 Thr165
MDATP_acam 7	−10.8	Asn46 Glu50 Arg76 Gly77 Ile78 Pro79 Tyr109 Arg136
XCLO_acam 31	−6.2	Glu50 Arg76 Gly77 Ile78 Pro79Tyr109 Arg136
acetamide		
XATP_asnr 17	−8.7	Glu50 Arg76 Gly77 Ile78 Tyr109
methanol		
MDATP_meoh 48	−4.4	Glu50 Arg76 Ile77 Tyr109
		(iii) Around Adenine's NH_2 Group Small Polar Groups
NMA		
XATP_acam 177	−6.5	Asn46 Ala47 Glu50 Asp73 Gly77 Ile78 Tyr109 Thr 165
MDATP_acam 13	−10.2	Val43 Asn46 Ala47 Glu50 Val71 Asp73 Gly77 Ile78 Val120 Thr165 Val167
XCLO_acam 25	−6.5	Val43 Asn46 Ala47 Asp49 Glu50 Asp73 Ile78 Val120 Thr165 Val167
acetamide		
XATP_asnr 312	−3.6	Val43 Asn46 Ala47 Val71 Asp73 Ile78 Thr165 Val167
MDATP_asnr 80	−5.7	Val43 Asn46 Ala47 Glu50 Asp73 Gly77 Ile78 Tyr109 Thr165
methanol		
MDATP_meoh 36	−5.1	Asn46 Ala47 Glu50 Asp73 Gly77 Ile78 Thr165
trifluoromethane		
XATP_tfme 18	−10.9	Asn46 Ala47 Glu50 Asp73 Gly75 Arg76 Gly77 Ile78 Pro79 Tyr109 Thr165
MDATP_tfme 28	−10.9	Asn46 Glu50 Arg76 Gly77 Ile78 Pro79 Tyr109 Arg136
		(iv) On Top of Adenine Ring Small Polar Groups
acetamide		
MDATP_asnr 18	−7.8	Asn46 Ala47 Glu50 Gly77 Ile78 Asp84 Tyr109

[a] Important vdW contacts made by adenine in the crystal structure (XATP): Asn46 Ala47 Glu50 Asp73 Arg76 Gly77 Ile78 Pro79 Ile94 Lys103 Tyr109 Val120 Thr165.

structures as well as to specific residues in the binding site. For that reason, we present here a brief overview of the important interactions made between DNA gyrase with ADPNP (ATP) and the inhibitors for which structures of the complex have been determined. Interactions of the different inhibitors with the protein residues are shown in Figure 1. Important residues that interact with the ATP include Asp73 and Tyr109 both interacting with the adenine ring, Lys103 and glycines 114, 117, and 119 stabilizing the phosphate moiety, and residues Asn46 and Thr165. Asn46 is one of the key residues in the binding of inhibitors, forming hydrogen bonds to the novobiose sugar and the substituted cyclothialidinic resorcinol ring, respectively. In the complex with ADPNP/ATP, Asn46 stabilizes the phosphate moiety via a magnesium atom (cf. Figure 1a).

Residues that comprise the second loop make important contributions to ATP binding by stabilizing the phosphate moiety. In the crystal structure of the 43 kDa complex with the ATP analogue, this loop is folded over the active site, while in the crystal structures with the inhibitor, this loop is in an open, extended conformation and not implicated in chlorobiocin binding. Interactions related to coumarin binding involve residues Val43, Gly77, and Arg136, which mediate an important hydrogen bond to the coumarin ring. The pyrrole ring of the chlorobiocin is completely buried in a hydrophobic pocket formed by Val43, Val71, Val120, Val167, and

Ile78.[13] The isopentenyl moiety "wraps around" Pro79 and away from the solvent.[13]

For the ensemble of structures used in this study, Table 1 summarizes the energetic results for functional groups tested before and after the postprocessing procedure. The results show that in all cases, postprocessing both reduces the number of minima that are to be included in the subsequent clustering and modifies their energetic ranking. All minima that had free energies greater than zero were eliminated and not included in the subsequent clustering. These results show that the postprocessing procedure can be used, in some cases, to vastly reduce the number of cluster minima found and to remove a large number of "false positive" results for the binding site.

After the postprocessing of the MCSS minima, the clustering procedure was done. In the following sections, we present an analysis of these results for the various binding pockets that can be utilized for ligand design or optimization. The results also demonstrate how these pockets are affected by the dynamic nature of the binding site. We will present data that indicate that different compounds can be found to bind in existing binding pockets and that new pockets can be exploited for inhibitor binding. The detailed results are presented in Tables 2−6, which summarize the different clusters found for the different functional groups. The notation used for a particular cluster denotes the particular protein structure (XATP, MDATP, or XCLO), the functional group (see Figure 3), and the ranking of the best minimum in that cluster. Tables 2−6 also give the best binding free energy as well as the cluster signature, or list of significant van der Waals contacts that the cluster members make. Each member of the cluster makes contact with all but at most two residues in the signature list.

Adenine Binding Pocket. In the crystal structure of DNA gyrase with ADPNP, the adenine group makes several important hydrogen bonds to Asp 73, crystal waters, and Tyr 109.[11,13,31] Crystal structures of the chlorobiocin and novobiocin complexes show that they interfere with ATP binding by partially overlapping this site (cf. Figure 1b). In particular, the novobiose sugar overlaps a site that prevents binding of the adenine ring of ADPNP. The novobiose sugar is, by itself, inactive in antibacterial and antigyrase activity, and it is probably the coumarin ring that directs the sugar moiety to its position in gyrase.[32] In the crystal structure of the p24 subdomain complex with the cyclic peptide cyclothialidine from *Streptomyces*, the resorcinol ring takes the same position as the adenine ring in the complex with ATP. In Table 2, the amino acids from the p24 domain that make van der Waals contact with the adenine ring in the crystal structure are listed; these include Asn46, Ala47, Glu50, Asp73, Arg76, and Gly77, among others. Several of these residues are part of the flexible loops, for example, Lys 103 and Tyr 109, and are not near the adenine binding pocket in the open-loop conformation.

Analysis of the adenine binding site after the MCSS calculations indicates that a variety of aromatic functional groups have well-ranked minima in all three structures. Representative positions for aromatic groups in the adenine binding pocket are shown in Figures 4, 5, and 6a,b. Despite the fact that second-loop residues

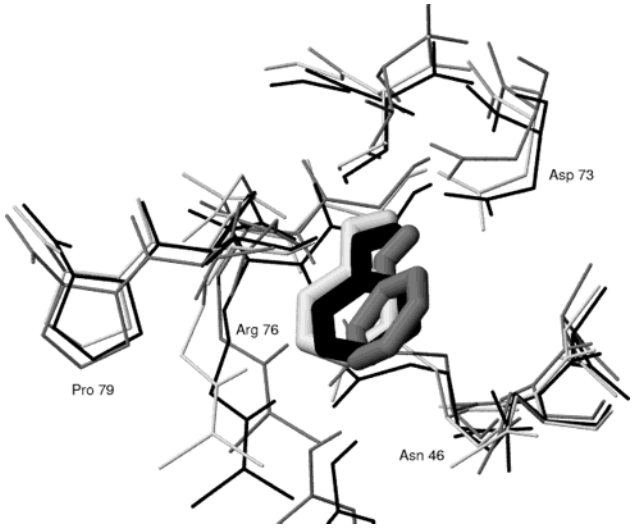

Figure 4. Consensus binding site for the phenol group in the three DNA gyrase structures. Important residues of the three different DNA gyrase B conformations are shown as lines (black (XATP), dark-gray (MDATP), light-gray (XCLO)). $\Delta G^{binding}$ in kcal/mol is shown in parentheses: XCLO_phen 1 (−18.3), MDATP_phen 169 (−4.1), XATP_phen 2 (−16.7). In all three structures, the phenol makes hydrogen bonds with the gyrase residue Asp73.

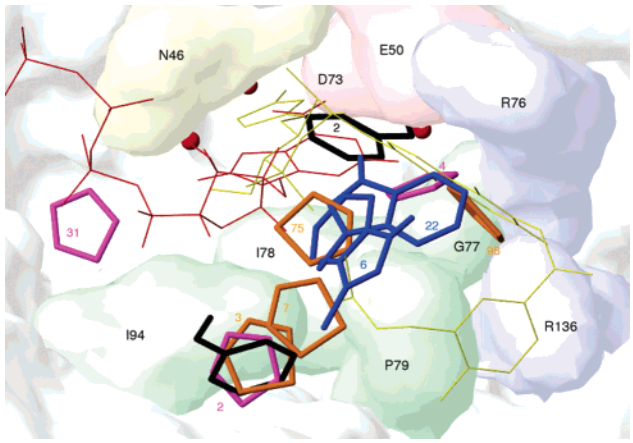

Figure 5. Representative aromatic clusters for the groups in the active site of the XATP structure are shown here. Oxazole (oxaz) clusters are shown in magenta, thiazole (thiz) clusters in orange, phenol (phen) clusters in black, and methylindole (trpr) clusters in blue. Clusters are labeled according to their postprocessing rank. Their $\Delta G^{binding}$ in kcal/mol are shown in parentheses: phen 2 (−16.7), phen 4 (−13.2), thiz 3 (−10.7), thiz 7 (−9.8 kcal/mol), thiz 75 (−6.5), thiz 98 (−5.9), oxaz 2 (−11.0), oxaz 4 (−10.6), oxaz 31 (−9.2), trpr 6 (−15.0), and trpr 22 (−13.1). ATP is shown in red, clorobiocin is shown in yellow, and the three crystal waters are shown as red spheres. Second-loop residues are left out for a better view of the active site.

make contact with the adenine ring in the crystal structure, several residues make significant contacts to aromatic functional groups irrespective of the loop orientation; several of these functional groups have well-ranked minima in all three structures. A striking example is phenol, which is found to hydrogen-bond to Asp73 in all three structures (see Figure 4). As indicated in Table 2, functional group clusters with favorable binding energies are found in all three structures. The cluster signatures for these three all include residues Asn46, Ala47, Glu50, Asp73, Ile78, and Thr165,

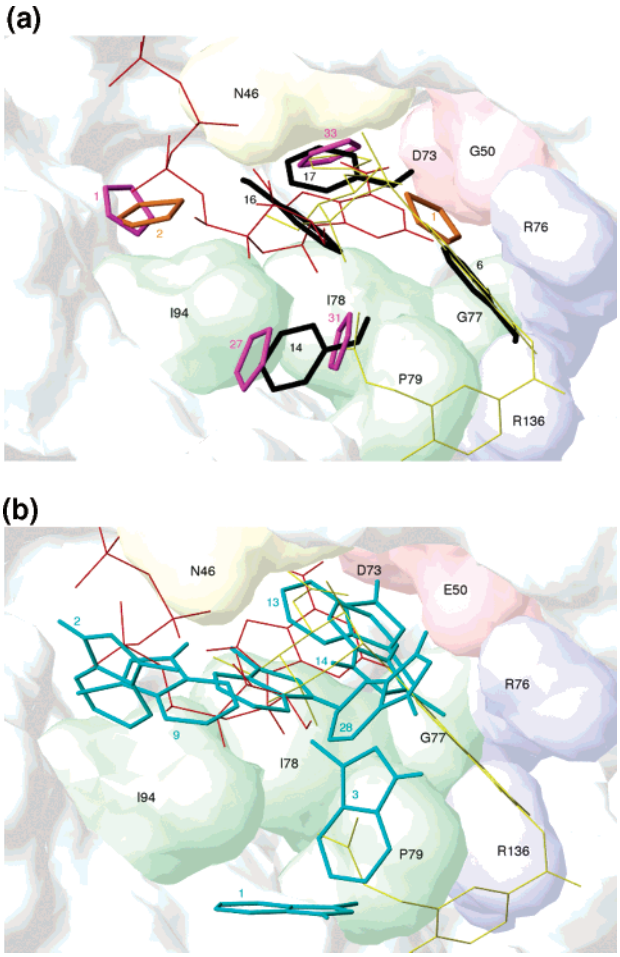

Figure 6. (a) Representative aromatic clusters found in the active site of the XCLO structure. Oxazole clusters are drawn in magenta, thiazole clusters in orange, and phenol clusters in black. The clusters are labeled according to their postprocessing rank. The clusters are ($\Delta G^{binding}$ in kcal/mol in parentheses) phen 6 (−10.9), phen 14 (−8.7), phen 16 (−8.3), phen 17 (−7.7), oxaz 1 (−12.7), oxaz 27 (−8.0), oxaz 31 (−7.9), oxaz 33 (−7.9), thiz 1 (−10.2), thiz 2 (−9.9). ATP is shown in red, and clorobiocin is shown in yellow. The best phen cluster (phen 1) is shown in Figure 4. Second-loop backbone atoms are left out for better visualization. (b) Representative trpr clusters for the XCLO structure. The methylindole (trpr) clusters are shown in cyan. The clusters are labeled according to their postprocessing rank. The $\Delta G^{binding}$ (kcal/mol) are shown in parentheses: trpr 1 (−13.3), trpr 2 (−12.5), trpr 3 (−12.4), trpr 6 (−11.7), trpr 9 (−10.8), trpr 13 (−10.4), trpr 14 (−10.2), trpr 28 (−7.4). ATP is shown in red, and clorobiocin is shown in yellow. Second-loop backbone atoms are left out for better visualization.

meaning that the binding site for this functional group is not strongly affected by the change in loop conformation. Moreover, in all three structures, the hydroxyl oxygen of the phenol group overlaps the oxygen atom from crystallographic water oxygen (see Figure 5); this water mediates an important hydrogen bond to Asp73[13,31] in both the ATP and novobiocin bound structures. This suggests that certain functional groups could replace the crystallographic water.

The binding affinity calculated for phenol by eq 3 is among the most favorable for all groups tested, which makes phenol interesting both as a scaffold and as a side chain in inhibitor design. Interestingly, one molecule in a series of inhibitors developed by Boehm and co-workers[31] used phenol as a scaffold. The inhibitor

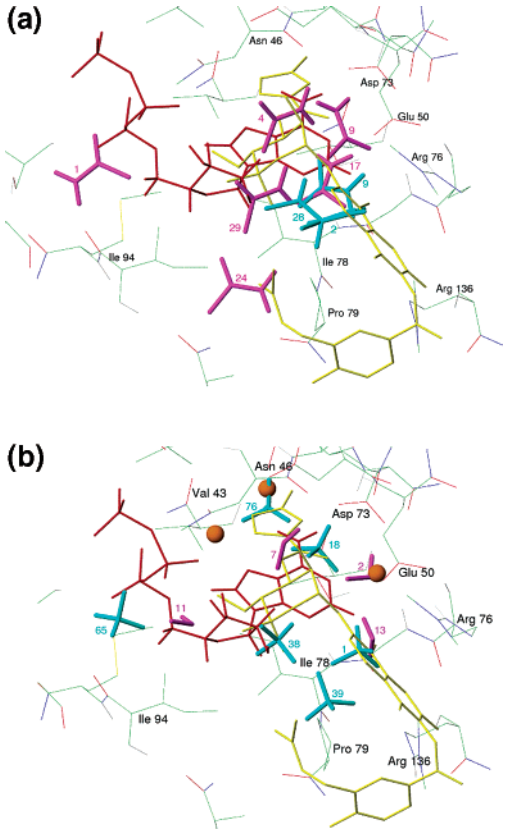

Figure 7. (a) Representative clusters for the NMA (acam, in cyan) and acetamide (asnr, in magenta) groups in the active site of the XATP structure are shown here. Clusters are labeled according to their postprocessing rank. Their $\Delta G^{binding}$ in kcal/mol are the following: acam 2 (−11.7), acam 9 (−10.2), acam 28 (−9.3), asnr 1 (−12.3), asnr 4 (−11.5), asnr 9 (−9.8), asnr 17 (−8.7), asnr 24 (−8.1), asnr 29 (−7.9). ATP is shown in red, and chlorobiocin is shown in yellow. Second-loop residues are left out for a better view of the active site. (b) Representative clusters for the methanol (meoh, in magenta) and trifluoromethane (tfme, in cyan) groups in the active site of the XATP structure are shown here.The clusters are labeled according to their postprocessing rank. Their $\Delta G^{binding}$ in kcal/mol are the following: meoh 2 (−12.3), meoh 7 (−9.3), meoh 11 (−7.6), meoh 13 (−7.5), tfme 1 (−13.9), tfme 18 (−10.9), tfme 38 (−10.0), tfme 39 (−10.0), tfme 65 (−9.3), and tfme 76 (−8.9). ATP is shown in red, and clorobiocin is shown in yellow. The three crystal waters around the adenine ring are represented as red spheres.

cyclothialidine also places a similar aromatic ring close to Asp73 (cf. Figure 1c).

Other small aromatic functional groups such as oxazole and thiazole are also found to bind in the adenine pocket, making contacts similar to those made by phenol (see Figure 6a for an example with thiazole and Table 2 for a detailed listing of the contacts made by oxazole and thiazole in the adenine binding pocket). The binding of the smaller aromatic groups (as evaluated by eq 3) is generally less favorable than it is for phenol; i.e., the best $\Delta G^{binding}$ values for these groups are about −10 vs −18 kcal/mol for phenol. This difference is mainly due to less favorable van der Waals and electrostatic protein−ligand interactions for oxazole and thiazole. This suggests that H bond donor groups are especially favorable around Asp73.

For the larger aromatic group 3-methylindole, the analogue of tryptophan (trpr), favorable positions are found for the XCLO structure only, and it is found to

Table 3. Phosphate Binding Site[a]

cluster	ΔG^{bind} (kcal/mol)	cluster signature
		Aromatic Groups
oxazole		
XATP_oxaz 31	−9.2	Asn46 Ile94 Ala100 Gly101 Gly117 Val118 Gly119 Val120 Ser 121
MDATP_oxaz 167	−5.5	Asn46 Ile94 Ala100 Gly101 Val118 Val120
MDATP_oxaz 272	−1.3	Val 43 Asn46 Ala47 Asp73 Ile78 Thr165 Val167
XCLO_oxaz 1	−12.7	Asn46 Ile94 Val97 Leu98 His99 Ala100 Val118 Gly119 Val120 Ser121
XCLO_oxaz 80	−2.1	Asn46 Ile94 His99 Gly116 Val118 Gly119 Val120
phenol		
XCLO_phen 3	−11.4	Asn 46 Ile94 Val97 His99 Ala100 Val118 Gly119 Val120 Ser121
thiazole		
XATP_thiz 208	−4.3	Asn46 Ile94 Ala100 Gly117 Val118 Gly119 Val120 Ser121
MDATP_thiz 159	−4.7	Asn46 Ile94 Ala100 Gly101 Val118 Val120
XCLO_thiz 2	−9.9	Ile94 Val97 Leu98 His99 Ala100 Gly119 Val120 Ser121
3-methylindole		
XCLO_trpr 2	−12.5	Asn46 Ile94 Val97 Leu98 His99 Ala100 Phe104 Gly117 Val118 Gly119 Val120 Ser121
		Nonpolar Groups
benzene		
XCLO_benz 7	−8.9	Ile94 Val97 Leu98 His99 Ala100 Val118 Gly119 VAl120 Ser121
XCLO_benz 15	−5.1	Asn46 His99 Phe104 Gly117 Val118 Gly119
		Small Polar Groups
NMA		
XCLO_acam 15	−7.2	Asn46 Ile94 Val97 Leu98 His99 Ala100 Phe104 Gly117 Val118 Gly119 Val120 Ser121
XCLO_acam 26	−6.4	Asn46 Ile94 His99 Ala100 Gly117 Val118 Gly119 Val120 Ser121
XCLO_acam 38	−5.5	Glu42 Asp45 Asn46 Asp49 His99 Gly117 Val118 Gly119
XCLO_acam 44	−4.8	Ile94 Val97 His99 Ala100 Gly118 Val120 Ser121
acetamide		
XATP_asnr 1	−12.3	Ile94 Ala100 Gly101 Gly117 Val118 Gly119 Val120 Ser121
MDATP_asnr 10	−8.5	Ile94 Ala100 Gly101 Gly117 Val118 Val119 Val120 Ser121
MDATP_asnr 39	−6.8	Gly101 Gly102 Lys103 Phe104 His116 Gly117
methanol		
XATP_meoh 23	−5.8	Ile94 Ala100 Val118 Gly119 Val120 Ser121
XATP_meoh 24	−5.7	Ile94 Gly117 Gly119 Val120LE94 GLY117 GLY119 VAL120
MDATP_meoh 14	−6.0	Ile94 Val118 Val120 Ser121
trifluoromethane		
XATP_tfme 65	−9.3	Asn46 Ile94 Val97 Ala100 Gly101 Gly117 Val118 Gly119 Val120 Ser121
MDATP_tfme 51	−10.0	Ile94 Ala100 Gly101 Gly117 Val118 Gly119 Val120 Ser121
MDATP_tfme 116	−7.6	Glu42 Asp45 Asn46 Asp49 Lys103 Gly113 Gly114
		Negatively Charged Clusters
acetate ion		
XATP_acet 1	−14.1	Glu42 Asn46 Gly114 Leu115 His116 Gly117 Val118 Gly119 Val120
XATP_acet 19	−5.1	Asn46 Ala100 Gly101 Gly102 Lys103 Phe104 Gly114 Leu115 His116 Gly117 Val118
XATP_acet 25	−4.0	Asn46 Ala100 Gly101 Gly102 Gly114 Leu115 Gly117
MDATP_acet 3	−8.0	Glu42 Asn46 Ile94 Ala100 Gly101 Gly117 Val118 Gly119 Val120 Ser121
MDATP_acet 9	−5.8	Glu42 Asn46 Gly101 Lys103 Gly114 Leu115 His116 Gly117 Val118 Val120
butanoic acid		
XATP_glur 1	−9.3	Glu42 Asn46 Lys103 Gly114 Leu115 His116 Gly117 Val118 Gly119 Val120
XATP_glur 2	−9.3	Met25 Tyr26 His99 His 116 Val118
XATP_glur 13	−7.7	Glu42 Gly113 GLy114 Leu115 His116 GLy117 Val118 Gly119
XATP_glur 36	−5.5	Asn46 Lys103 Gly113 Leu115 His116 Gly117 Val118

[a] Important vdW contacts made by phosphate in the crystal structure (XATP): Glu42 Asn46 Ile94 Ala100 Gly101 Gly102 Gly113 Gly114 Leu115 His116 Gly117 Val118 Gly119 Val120 Ser121.

bind deep in the ATP binding pocket (Figure 6b). Postprocessing eliminates all 3-methylindole clusters found deep inside the ATP binding pocket in the ATP bound structure (XATP); this group was not studied for the MDATP structure. In the XCLO structure, two binding modes for 3-methylindole are found partially overlapping the adenine ring, in positions very similar to those found for the phenol clusters (XCLO_trpr 13 and 14, Figure 6b). Their binding, however, is not as favorable as that of phenol in the same position because of a less favorable electrostatic interaction with the protein.

Besides aromatic groups, small polar groups also bind favorably in the adenine binding pocket. However, they do not occupy the entire adenine binding pocket because of their smaller size. Clusters of polar groups are found in several positions around the adenine binding pocket, making hydrogen bonds to polar residues lining the pocket.

There are several polar and charged amino acids that make favorable interactions with the polar groups. As can be seen from Table 2 and Figure 7a,b, *N*-methylacetamide, acetamide, and methanol all bind favorably in the vicinity of Asp73, an important amino acid in substrate and inhibitor binding. It can be seen from Figure 7b that the hydroxyl of a methanol superposes exactly onto the oxygen of a crystallographic water molecule, as was observed for the phenol hydroxyl. This suggests again that this water molecule can be displaced in designed inhibitors. The $\Delta G^{binding}$ values computed for this binding site are very favorable (−15 to −10 kcal/mol), considering the small size of the ligands, which is due to favorable electrostatic interactions.

From Table 2, polar groups are also stabilized by H bonds to Glu50 (for example, *N*-methylacetamide no. 9 in Figure 7a), to the main chain of Gly77, and to Asn46 (for example, methanol no. 7 in Figure 7b). Note that the oxygen of methanol no. 7 in Figure 7b (using the closed-loops structures) is placed close to the hydroxyl group of the novobiose sugar in the novobiocine inhibitor, making similar interactions to Asn46. These interactions are thus conserved in different structures. In the closed-loop conformation, polar groups are also stabilized in the adenine binding pocket by interactions with Lys103 and Tyr109. However, these stabilizing interactions are not present in the open-loop conformation.

These results suggest that the adenine binding site provides a good environment for aromatic group binding making significant interactions with residues Asn46, Gly75, Asp76, Gly77, Ile78, and Thr165. This suggests that aromatic groups other than those found in the known inhibitors could serve as possible scaffolds. Small polar functional groups bind preferably at the edges of the adenine ring binding pocket, making them interesting groups that can be used to link the adenine binding pocket to other pockets in the active site. These results also suggest that inhibitors that displace the active site bridging water molecules can be developed. Slight changes in the active site can lead to slight differences in the detailed binding or large changes in binding free energies; however, the use of several structures allows for the identification of consensus binding sites that are found in the three protein conformations. The large variety of small polar groups suggests that their binding is less specific, thus opening up larger possibilities for linker groups.

Ribose and Phosphate. Besides the adenine ring, the binding pocket of the natural ligand must accommodate the ribose and phosphate of ATP. Several residues from the mobile second loop have an important role in binding these functional groups;[10] thus, these binding interactions are likely to be affected by the conformational changes occurring in the protein. In the natural ligand ATP, the binding of the triphosphate tail has an important functional role. The crystal structures show that the phosphate moiety of the ATP is stabilized by interactions with the GXXGXG motif of the second loop, which undergoes significant conformational change between the ATP bound structure and the inhibitor bound structure. Even though the large difference in conformation of the second loop affects the phosphate binding site, and therefore the functional group mapping at this site, our results show that consensus sites can be found.

Consensus binding sites for several polar aromatic groups were found in the phosphate binding site. The results for these groups are presented in Table 3 and Figures 5 and 6a,b. Oxazole has favorable clusters in all three structures situated around the α-phosphate moiety of ATP (Figure 5 XATP structure; Figure 6a, XCLO structure; data not shown for structure MDATP); contact residues include Asn46, Ile94, Ala100, Val118, Gly119, and Val120. Favorable minima for thiazole are also found in the XCLO (Figure 6a) and XATP structure (data not shown). Some of these contacts are similar to those made by the adenine group. It can be seen from

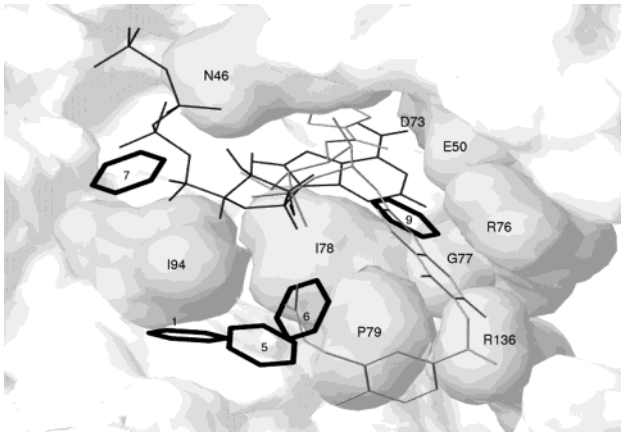

Figure 8. Most favorable benzene (in black) clusters in the active site of the XCLO structure. The clusters are labeled according to their postprocessing rank. $\Delta G^{binding}$ in kcal/mol in parentheses are the following: benz 1 (−10.8), benz 5 (−8.9), benz 6 (−8.9), benz 7 (−8.9), benz 9 (−8.1). ATP is shown in black, and clorobiocin is shown in gray. Second-loop residues are left out for better visualization.

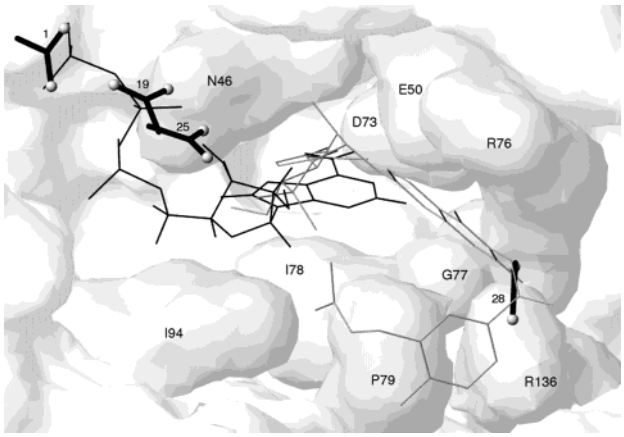

Figure 9. Representative clusters for the negatively charged acetate ion (in black, with oxygen atoms as gray spheres) in the XATP active site are shown here labeled by their postprocessing rank. Their $\Delta G^{binding}$ in kcal/mol are the following: acet 1 (−14.1), acet 19 (−5.9), acet 25 (−4.0), acet 28 (−3.4). ATP is shown in black, and clorobiocin is shown in gray.

Table 3 that these small polar groups are in contact with residues situated at both edges of the second loop, such as, for example, Ala100, Val118, Gly119, and Val120. The fact that in all three structures this binding site can be occupied by small polar aromatic groups suggests that the phosphate binding site could be a target for ligands based on aromatic compounds.

Larger aromatic groups such as benzene (Figure 8) and tryptophan (Figure 6b) are also found around the α-phosphate binding site but only in the XCLO structure, suggesting that these functional group sites would be useful if one is targeting the open structure in the rational design of new ligands.

In the closed-loop conformation, a binding pocket favorable for negatively charged ligands is formed, particularly around the oxygens of the γ-phosphate. Negatively charged groups such as the acetate ion (acet) and butanoic acid (glur) have some of the best binding free energies at the phosphate binding site in the MDATP and XATP structures. Although not a consensus binding site because no favorable cluster minima were found for the negatively charged groups in the

Table 4. Pyrrole Binding Site[a]

cluster	ΔG^{bind} (kcal/mol)	cluster signature
	Aromatic Groups	
oxazole		
XCLO_oxaz 33	−7.9	Val43 Asn46 Ala47 Val71 Gln72 Asp73 Thr165 Met166 Val167
phenol		
XCLO_phen 17	−7.7	Val43 Asn46 Ala47 Glu50 Val71 Asp73 Ile78 Val120 Thr165 Val167
3-methylindole		
XCLO_trpr 30	−7.3	Val42 Asn46 Ala47 Glu50 Val71 Asp73 Ile78 Val120 Thr165 Val167
	Small Polar Groups	
NMA		
XATP_acam 277	−4.7	Val43 Asn46 Ala47 Glu50 Asp73 Ile78 Thr165
MDATP_acam 13	−10.2	Val43 Asn46 Ala47 Glu50 Val71 Asp73 Gly77 Ile78 Val120 Thr165 Val167
XCLO_acam 2	−14.4	Val43 Asn46 Ala47 Glu50 Val71 Gln72 Asp73 Ile78 Val120 Thr165 Met166 Val167
acetamide		
XATP_asnr 312	−3.6	Val43 Asn46 Ala47 Val71 Asp73 Ile78 Thr165 Val167
MDATP_asnr 30	−7.2	Val43 Asn46 Ile78 Val120 Val165
trifluoromethane		
XATP_tfme 76	−8.9	Val43 Asn46 Ala47 Val71 Gln72 Asp73 Ile78 Thr165 Met166 Val167
MDATP_tfme 49	−10.1	Val43 Ala47 Val71 Gln72 Asp73 Thr165 Met166 Val167
	Nonpolar Groups	
benzene		
XATP_benz 161	6.8	Val 43 Asn 46 Ala 47 Asp73 Val120 Thr165 Val 167
MDATP_benz 64	−1.7	Val43 Ala47 Asp73 Val120 Thr165 Val167
MDATP_benz 72	2.7	Val43 Asn46 Ala47 Asp73 Gly77 Ile78 Val120 Thr165 Val167
XCLO_benz 20	−2.6	Val43 Asn46 Ala47 Val71 Asp73 Ile78 Val120 Thr165 Val167

[a] Important vdW contacts made by pyrrole in the XCLO crystal structure: Val43 Ala47 Val71 Asp73 Ile78 Val120 Thr165 Val167.

XCLO structure, we present the results for the negatively charged groups given the importance of this binding site.

The results for binding of negatively charged groups in the phosphate binding site are presented in Table 3 and Figure 9. As expected, negatively charged functional groups are stabilized in the phosphate binding sites in the closed-loop conformation but not in the open-loop conformation. It can be seen from Figure 9 that a very favorable position for the acetate group is found superposed onto the γ-phosphate binding site in the XATP structure (acetate ion cluster no. 1). This cluster has a very favorable binding free energy of −14.1 kcal/mol. Three butanoic acid clusters with favorable binding are also found there (see Table 3). There is thus a very favorable electrostatic environment at this position that compensates for the large desolvation energy of the charged group. Negatively charged groups are also found close to the binding site of the α- and β-phosphates (cf. Figure 9), with favorable binding free energies but not as favorable as for the γ-phosphate binding site (around −5 kcal/mol; cf. Table 3). In the ATP bound structure, the β-phosphate is coordinated by an Mg^{2+} ion, which is not included in the present calculations.

Many clusters for acetate and butanoic acid form hydrogen bonds to residues in the second loop of the XATP and MDATP structures that stabilize the charges. In particular, hydrogen bonds are made to Leu115, His116, Gly117, Val118, and Gly119. These residues are part of the ATP binding motif GXXGXG in loop 99−119 interacting with the phosphate moiety of ATP.[3] The strongest electrostatic interaction for the cluster superposed onto the γ-phosphate position (XATP_acet 1; cf. Figure 9) is found with the charged His116; other acetate ion clusters also interact electrostatically with this residue. In addition to highlighting an important interaction, these results also underline the importance of preliminary pK_a calculations before MCSS runs. These calculations show that in this particular conformation, the His is protonated when complexed with ATP. In recent calculations, we showed that in the absence of ATP in the binding site, His116 would be unprotonated in the closed conformation.[33] Clusters XATP_acet 19 and 25 form hydrogen bonds with Lys 103, which is involved in a salt bridge with the β-phosphate of ADPNP in the crystal structure (cf. Figure 2a). Thus, the MCSS calculations show that small negatively charged ligands can mimic the interactions of the triphosphate tail.

The data for the acetate group further demonstrate the importance of using multiple structures. Two clusters, MDATP_acet 1 and XATP_acet 19, were indeed found in equivalent positions, namely, lying on the ATP β-phosphate binding site, yet their binding free energies differ by more than 4 kcal/mol (−9.3 kcal/mol (MDATP) vs −5.1 kcal/mol (XATP)). It is known that small fluctuations in the protein structure can change significantly the electrostatic binding, and using several structures helps identify relevant interactions that can appear less favorable if only one structure is used.

In contrast to the two closed-loop conformations, none of the negatively charged minima found in the XCLO structure make contact with residues 113−119, which mediate important stabilizing interactions for the negatively charged groups in the ATP bound structure (cf. Figure 3a). The electrostatic interactions make a significant contribution to the free energy of binding of the negatively charged minima in the XATP/MDATP structures. However, without the stabilizing interactions of the second loop, the electrostatic interactions of the negatively charged functional groups are not sufficiently high to overcome the high desolvation energies. These results further emphasize that in considering ligand optimization, desolvation of the functional group must be considered in addition to functional group−protein interactions.

Aside from the ATP binding pocket discussed above, gyrase binds functional groups that partially overlap ATP and make use of different binding pockets in the protein to achieve efficient interactions. In what follows,

Table 5. Coumarin Binding Site[a]

cluster	ΔG^{bind} (kcal/mol)	cluster signature
	Aromatic Groups	
phenol		
XATP_phen 30	−9.3	Glu50 Arg76 Gly77 Ile78 Pro79 Tyr109 Arg136
MDATP_phen 228	−1.8	Glu 50 Arg76 Ile78 Pro79 Ser108 Tyr 109 Arg136
XCLO_phen 6	−10.9	Asn46 Asp49 Glu50 Arg76 GLy77 Ile78 Pro79 Arg136
thiazole		
XATP_thiz 15	−9.3	Glu50 Arg76 Gly77 Pro79 Ser108 Tyr109
XATP_thiz 98	−5.9	Glu50 Arg76 Gly77 Pro79 Ser108 Arg136
XCLO_thiz 1	−10.2	Asn46 Glu50 Asp73 Gly75 Arg76 Gly77 Ile78 Thr165
oxazole		
XATP_oxaz 4	−10.6	Glu50 Arg76 Gly77 Ile78 Pro79 Gly102 Ser108 Tyr109 Arg136
MDATP_oxaz 181	−5.3	Glu50 Arg76 Gly77 Ile78 Pro79 Ser108 Tyr109 Arg136
XCLO_oxaz 30	−8.0	Glu50 Arg76 Gly77 Ile78 Pro79
3-methylindole		
XATP_trpr 22	−13.1	Glu50 Arg76 Pro79 Ile94 Gly101 Gly102 Ser108 Tyr109
XCLO_trpr 28	−7.4	Glu50 Arg76 Gly77 Ile78 Pro79 Ile94 Arg136
	Nonpolar Groups	
benzene		
XCLO_benz 9	−8.1	Glu50 Arg76 Gly77 Ile78 Pro79 Arg136
XATP_benz 51	−5.9	Glu50 Arg76 Gly77 Ile78 Tyr109 Thr165
MDATP_benz 49	−3.8	Pro79 Gly102 Ser108 Tyr109
	Small Polar Groups	
trifluoromethane		
XATP_tfme 1	−13.9	Glu50 Arg76 Gly77 Ile78 Pro79 Ser 108 Tyr109 Arg136
MDATP_tfme 28	−10.9	Asn46 Glu50 Arg76 Gly77 Ile78 Pro79 Tyr109 Arg136
NMA		
XATP_acam 58	−8.4	Glu50 Arg76 Ile78 Pro79 Tyr109
MDATP_acam 217	−0.5	Glu50 Arg76 Gly77 Pro79 Tyr 109 Arg136
XCLO_acam 3	−12.5	Glu50 Asp73 Arg76 Gly77 Ile78 Pro79 Arg136 Thr165
XCLO_acam 21	−6.7	Arg76 Pro79 Arg136 Glu137
	Negatively Charged Groups	
acetate ion		
XATP_acet 28	−3.4	Arg76 Pro79 Arg136 Glu137 Thr163
MDATP_acet 14	−4.4	Arg76 Pro79 Arg136
XCLO_acet 1	−8.9	Ala53 His55 Arg76 Pro79 Arg136
butanoic acid		
XATP_glur 41	−5.1	Arg76 Pro79 Ser108 Arg136
XCLO_glur 4	−6.9	Pro79 Thr90 Gly81 Arg136 Glu137

[a] Important vdW contacts made by the coumarin ring in the crystal structure (XCLO): Glu50 Arg76 Gly77 Ile78 Pro79.

we describe these pockets and the small ligands that are found to bind favorably to them.

Pyrrole Binding Pocket. The pyrrole ring in chlorobiocin contributes to the higher affinity of this inhibitor compared to novobiocin.[13,34] The pyrrole ring makes contact with residues that are involved in the binding of the adenine ring, but because of its smaller size, it is buried deeper in the protein and occupies a cavity lined by residues Val43, Ala47, Val71, Asp73, Ile78, Val120, and Thr165. The MCSS calculations locate several consensus clusters in the vicinity of the pyrrole moiety. The groups that bind favorably in the pyrrole pocket are mainly small polar groups such as NMA and trifluoromethane (Figure 7b). Their cluster signatures include many of the residues that are in the signature of the pyrrole ring, and their binding energies range from −14.4 to −3.6 kcal/mol (see Table 4). Small polar aromatics such as oxazole (Figure 6a) were found to bind only in the XCLO structure. Larger aromatics or purely hydrophobic groups are not found to bind favorably in the pyrrole pocket because of a large desolvation penalty of the protein. It must be noted that for small buried cavities inside proteins, the results of the continuum electrostatic calculations depend critically on the probe radius used to define the solvent accessible surface of the protein. With a small probe radius, a buried cavity will be considered accessible and filled with a high dielectric in the apo protein, while with a large probe radius, the cavity will be considered as a low dielectric region in the apo protein. This in turn affects the desolvation penalty of the protein upon binding of the ligand. For sake of consistency, we used the same probe radius (1.4 Å) for all the calculations presented in this paper. With this probe radius, the pyrrole pocket is treated as solvent accessible in the apo protein. Even though this pocket most likely contains water in the unbound protein, it is obviously an approximation inherent in the continuum dielectric method to treat it as bulk water.[15] The results for the buried pyrrole pocket should thus be treated with caution.

Coumarin Binding Pocket. A principal interaction of the coumarinic ring observed in the crystal structure is due to its stacking on the Arg76 side chain. Aromatic groups are found close to the position of the coumarin rings in two orientations: either parallel or perpendicular to the coumarin cycle. For example, benzene (Figure 8), thiazole (for XATP_thiz 98, see Figure 5; for XCLO-_thiz 1, see Figure 6a), oxazole (XCLO_oxaz 30 and MDATP_oxaz 181 (not shown)), phenol (XCLO_phen 6 in Figure 6a), and 3-methylindole (XCLO_trpr 28 in Figure 6b) all are found having the same orientation and signature (see Table 5) as the coumarinic rings, while aromatic ligands oriented perpendicular to the coumarin ring are shown in Figure 5 (for example, XATP_oxaz 4 and XATP_trpr 22 or XATP_thiz 15 (not shown)). Most of these clusters are stabilized by hydro-

gen bonds to the side chain of Arg136. These hydrogen bonds are some of the most important in coumarin binding to DNA gyrase.[13]

Good energy minima are also found for some small polar goups at several positions in the coumarin binding site. A consensus site for the *N*-methylacetamide minima was around the coumarinic ring in all three structures, and the details are given in Table 5. A very favorable position for *N*-methylacetamide binding perpendicular to the coumarin ring is shown in Figure 7a.

Several negatively charged consensus clusters are found around the coumarin ring binding pocket as well. All these clusters make hydrogen bonds to Arg136 and are found around the coumarinic C═O bond, which also accepts a hydrogen bond from Arg136. The groups include the acetate ion (aspartate side chain; cf. Figure 9) and butanoic acid (glutamate side chain) (see Table 5). Interestingly, trifluoromethane is found superposed onto the chlorine atom of chlorobiocin in two clusters with good binding free energy (see Figure 7b).

Identification of a New Hydrophobic Pocket. The previous results show that the principal interactions in and around the ATP binding site in DNA gyrase B can be determined by the MCSS method when employing the approach presented in the Methods. By use of this approach, the present calculations have identified a hydrophobic binding pocket close to the binding sites of ATP and the coumarinic inhibitors (cf. Figure 10). This binding pocket is composed of residues Pro79, His83, Ala90, Val93, Ile94, Ala100, Gly101, and Gly102 (cf Figure 10). Although some of these residues are implicated in stabilizing the isopentenyl group of the coumarin inhibitors, the MCSS calculations show that there is a deeper binding pocket formed by the residues mentioned above that may, if exploited in a rational design, open the way for more specific, better binding inhibitors. In all three gyrase conformations (XATP, MDATP, and XCLO), aromatic cluster minima with very good binding free energies are found in this pocket. Binding free energies for benzene clusters are in the −12 to −10 kcal/mol range (see minima XCLO_benz 1 in Figure 8), and all the clusters have similar signatures (see Table 6). Minima for aromatic groups such as thiazole, oxazole, phenol, and 3-methylindole are found with binding free energies between −13 and −8 kcal/mol. Some of these clusters are shown in Figure 5 (XATP_thiz 3, XATP_ozaz 2, and XATP_phen 4), Figure 6a (XCLO_phen 14, XCLO_oxaz 27, XCLO_trpr 1), Figure 6b (XCLO_trpr 1), and Figure 10 (XATP_benz 1, XATP_benz 2). Especially favorable is a cluster of 3-methylindole (trpr) in the XCLO structure that contains the best binding minimum with $\Delta G^{\text{binding}} = -13.3$ kcal/mol (Figure 7b, XCLO_trpr 1) and the phenol cluster in the XATP structure that contains a minimum with $\Delta G^{\text{binding}} = -13.2$ kcal/mol (Figure 5, XATP_phen 4).

The results from the calculations suggest that a ligand in or around the ATP binding site that inserts a strong binding functional group into this site could be an effective inhibitor. For example, the high-affinity inhibitors for gyrase B described by Boehm et al[31] make polar contacts with Asp73 and Arg136; these residues contact the coumarinic moiety in the coumarin inhibitor/protein crystal structures. The inhibitors described by

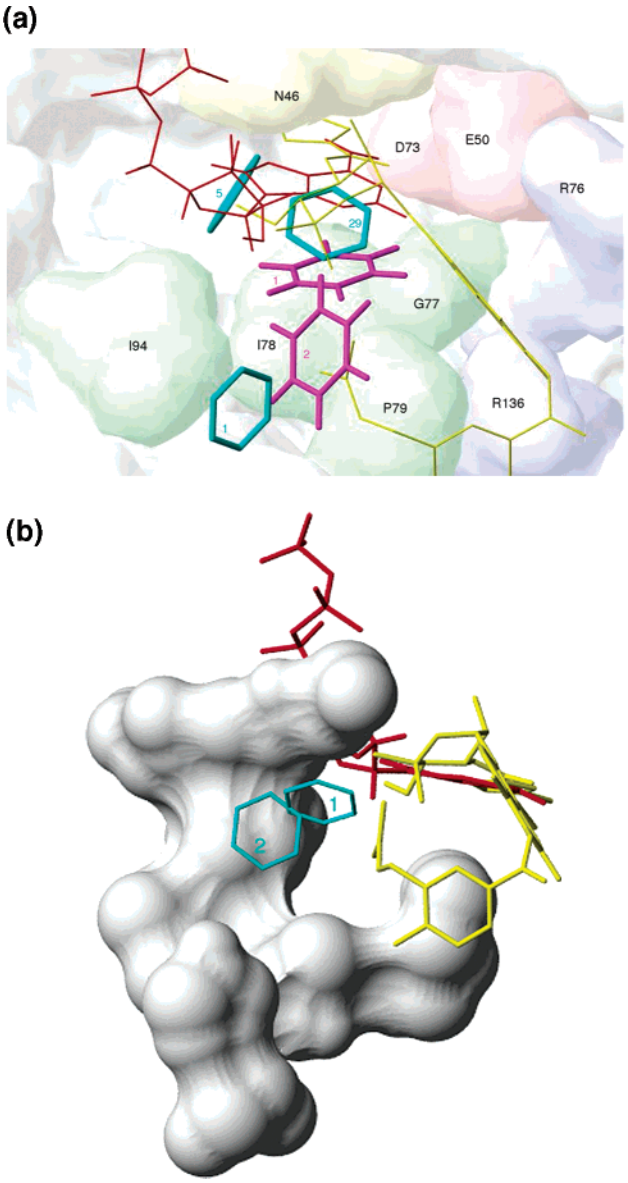

Figure 10. (a) Best benz clusters in the active site of the MDATP structure. Two very favorable benh (magenta) clusters at positions where no favorable benz (cyan) clusters could be found are also shown. The clusters shown are ($\Delta G^{\text{binding}}$ in kcal/mol in parentheses) benz 1 (−11.5), benz 5 (−10.1), benz 29 (−6.1), benh 1 (−11.7), and benh 2 (−11.6). ATP is shown in red, and chlorobiocin is shown in yellow. The second loop is not shown for better view of the active site. (b) Zoom on the hydrophobic pocket identified by the MCSS procedure with the two best benz minima in the XATP structure (with ATP in red and chlorobiocin in yellow) ($\Delta G^{\text{binding}}$ in kcal/mol in parentheses): benz 1 (−12.1), benz 2 (−11.3).

Boehm et al. contain a phenyl group that extends over Ile78; however, the MCSS calculations suggest that *if* this group could be inserted deeper into this hydrophobic pocket, the binding could be further enhanced and a binding mode significantly different from those of existing coumarin inhibitors could be developed. These results also demonstrate that MCSS calculations can identify potential binding sites outside the known binding site defined by experimental structures.

Linking the Hydrophobic Pocket to the Interior of the ATP Binding Site. Several possibilities for linking this hydrophobic pocket to the ATP/coumarin binding pocket of DNA gyrase are suggested by the

Table 6. Clusters in the Hydrophobic Binding Pocket

cluster	ΔG^{bind} (kcal/mol)	cluster signature
	Aromatic Groups	
oxazole		
XATP_oxaz 2	−11.0	Pro79 His83 Ala90 Val93 Ile94 Ala100 Gly101 Gly102
MDATP_oxaz 8	−9.6	Ala90 Ile94 Gly101 Gly102 Phe104
MDATP_oxaz 15	−9.4	Gly77 Ile78 Pro79 Ala90 Ile94 Gly101 Gly102 Tyr109
XCLO_oxaz 27	−8.0	His83 Ile94 Ala100 Gly101
phenol		
XATP_phen 4	−13.2	Pro79 His83 Ala90 Val93 Ile94 Ala100 Gly101 Gly102
MDATP_phen 3	−11.1	Ile78 Pro79 Ala90 Ala91 Ile94 Gly101 Gly102 Phe104
XCLO_phen 14	−8.3	Ile78 Pro79 His83 Ala90 Ile4 Ala100
thiazole		
XATP_thiz 3	−10.7	Pro79 His83 Ala90 Val93 Ile94 Ala100 Gly101 Gly102
MDATP_thiz 8	−9.6	Ile78 Pro79 Ala90 Ile94 Gly101 Gly102
XCLO_thiz 8	−8.2	Ile78 Pro79 His83 Ala90 Ile94
3-methylindole		
XATP_trpr 1	−16.1	Pro79 His 83 Ala90 Val93 Ile94 Ala100 Gly101 Gly102
XCLO_trpr 1	−13.3	Pro79 His93 Ala90 Val93 Ile94 Ala100 Gly101
	Nonpolar Groups	
benzene		
XATP_benz 1	−12.1	His83 Ala90 Val93 Ile94 Ala100 Gly101 Gly102
XATP_benz 2	−11.3	His83 Ala90 Val93 Ile 94 Ala100 Gly101 Gly102
XCLO_benz 1	−10.8	His83 Ala90 Val93 Ile94 Ala100 Gly101
MDATP_benz 1	−11.5	Pro79 Ala90 Ile94 Gly101 Gly102 Phe104
	Linking the Hydrophobic Pocket to the ATP Binding Site	
thiazole		
XATP_thiz 75	−6.5	Pro79 Ile94 Gly101 Gly102 Lys103 Tyr109
MDATP_thiz 17	−8.9	Asn46 Gly77 Ile78 Pro79 Ile94 Tyr109

MCSS calculations. Consensus binding sites for aromatic molecules positioned midway between the ATP/coumarin binding sites and the hydrophobic side pocket interacting with nonpolar residues such as Ile78 were found. Several examples are shown in Figure 5 (XATP_thiz 7 and 75) and in Figure 6b (XCLO_trpr 3). Instead of wrapping around Pro79 as it may occur in the coumarin inhibitors, these clusters suggest that a more direct linking of the hydrophobic binding pocket to the coumarinic/ATP binding pocket might be possible.

Besides the aromatic clusters, small polar groups were also found at positions between the ATP/coumarin and hydrophobic binding pocket (data not shown; see Table 6). The results suggest that diverse functional groups could be used to link the hydrophobic pocket directly to the novobiose and/or coumarin site.

V. Conclusion

The emergence of bacterial strains resistant to available antibiotics is a major public health concern, and there is a growing need to develop new antibacterial compounds. DNA gyrase, a procaryote specific type II topoisomerase, has attracted considerable attention from the pharmaceutical industry, which has focused primarily on developing analogues of coumarin inhibitors, a class of high-affinity natural inhibitors of gyrase B. A large number of modified coumarins have been synthesized and tested;[35−37] however, these drugs have proven to be toxic.

This study provides a comprehensive functional group mapping of the ATP binding site of DNA gyrase that provides a significant amount of data that can be exploited in lead identification and optimization, which in turn can spur the development of new antibiotic compounds. The results from the calculations identified binding sites of small chemical fragments that are predicted to bind with high affinity irrespective of the detailed conformations of the flexible active site loop. For example, the calculations suggest the possibility of using aromatic ligands to target the phosphate binding site (Figures 5, 6, and 8). The functional groups used in the calculations can serve as scaffolds for new compounds. The calculations also showed the existence of a deep hydrophobic binding pocket near the ATP binding site (Figure 10). Such sites are of interest especially in the optimization of existing compounds that are known to bind in the vicinity of the pocket.

In a recent study, Boehm et al. showed that screening for small ligands (so-called needle screening) can identify low molecular weight compounds (MW < 300) that bind to the gyrase B ATP binding site and that these small ligands can then be optimized to high-affinity compounds using structural information available on the binding site. The method developed in this paper is particularly useful in this kind of approach, where it can help find the "needle" compounds and subsequently refine them to higher affinity.

Additionally, the work incorporated new methodological developments that are of general interest to those involved in computational drug design. For example, the protonation states of the active site residues were determined prior to docking using continuum electrostatics.[23,38] The protonation states of titrating groups in proteins are important for their stability and function, but in drug design studies, residues such as histidines are most often taken as neutral by default. Our data show that some histidines of the gyrase are charged, and this in turn influences protein−ligand interactions. The introduction of a postprocessing step that incorporates a solvation correction to the computed binding energy represents another important modification to the standard protocol. Although the importance of solvation on binding free energies is now well recognized[22] and models that incorporate solvent corrections in binding studies have been proposed and used by other authors,[29,30] this approach is not used routinely in com-

putational drug design. Finally, protein flexibility was explicitly taken into account by using several protein structures. Conformations generated from molecular dynamics simulations were used in conjunction with the multiple experimental crystal structures in calculating the functional maps. The results show that the use of alternative conformations can result in significant changes in binding energies, which in turn can provide a set of novel positions and orientations for small functional groups. A new clustering algorithm using an energy criterion rather than the more usual rms position criterion was used for this project.[18] The protocol was tested and shown to be both reliable and efficient. It is especially useful when comparing results from multiple protein conformations.

The results of the present work show that it is computationally feasible to obtain a detailed pharmocophore mapping of a binding site using a realistic description of intermolecular interactions incorporating pK_a effects, ligand and protein desolvation, and protein flexibility. The results presented here can be used for pharmacophore-based screening of databases,[39] de novo design, and lead optimization.

Acknowledgment. The authors thank Prof. D. Wigley and Dr. D. Moras for providing crystal structures prior to publication and Dr. V. Lamour for helpful discussions. This work was supported by the Centre National de la Recherche Scientifique (CNRS), the Institut National de la Santé et de la Recherche Médicale (INSERM), and the Université Louis Pasteur de Strasbourg. We thank the Region Alsace for support of this work. Computer time was supplied by the Centre Informatique National l'Enseignement Supérieur (CINES), the Institut du Développement et des Ressources en Informatique Scientifique (IDRIS), and the Centre d'Études du Calcul Parallèle et de la Visualisation (CECPV).

Appendix

Abbreviations. ADPNP, 5′-adenylyl-β-γ-imidodiphosphate; CHARMM, Chemistry at Harvard Macromolecular Mechanics; MCSS, multiple copy simultaneous search; p24, 24 kDa subdomain of DNA gyrase B; p43, 43 kDa subdomain of DNA gyrase B; UHBD, University of Houston Brownian Dynamics Program; XATP, energy-minimized conformation of the 24 kDa subdomain of DNA gyrase used for the MCSS calculations; MDATP, conformation of the 24 kDa subdomain of DNA gyrase after the molecular dynamics simulation used for the MCSS calculations; XCLO, conformations of the 24 kDa subdomain of DNA gyrase B from the complex with chlorobiocin; rmsd, root mean square difference.

References

(1) Caron, P. R.; Wang, J. C. Alignment of primary sequences of DNA topoisomerases. *Adv. Pharmacol. (San Diego)* **1994**, *B29*, 271−297.
(2) Wang, J. C. Moving one DNA double helix through another by a type II DNA topoisomerase: the story of a simple molecular machine. *Q. Rev. Biophys.* **1998**, *31*, 107−144.
(3) Dutta, R.; Inouye, M. GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem. Sci.* **2000**. *25*, 24−28.
(4) Huang, W. M. Type II DNA Topoisomerase genes. *Adv. Pharmacol.* **1994**, *A29*, 201−225.
(5) Maxwell, A.; Lawson, D. The ATP-binding site of type II topoisomerases as a target for antibacterial drugs. *Curr. Top. Med. Chem.* **2003**, *3*, 283−303.
(6) Maxwell, A. DNA gyrase as a drug target. *Trends Microbiol.* **1997**, *5*, 102−109.
(7) Lewis, R. J.; Tsai, F. T. F.; Wigley, D. B. Molecular mechanisms of drug inhibition of DNA gyrase. *BioEssays* **1996**, *18*, 661−671.
(8) Morais-Cabral, J. H.; Jackson, A. P.; Smith, C. V.; Shikotra, N.; Maxwell, A.; Liddington, R. C. Crystal structure of the breakage−reunion domain of DNA gyrase. *Nature* **1997**, *388*, 903−906.
(9) Kampranis, S. C.; Gormley, N. A.; Tranter, R.; Orphanides, G.; Maxwell, A. Probing the binding of coumarins and cyclothialidines to DNA gyrase. *Biochemistry* **1999**, *38*, 1967−1976.
(10) Wigley, D. B.; Davies, G. J.; Dodson, E. J.; Maxwell, A.; Dodson, G. Crystal structure of an N-terminal fragment of the DNA gyrase B protein. *Nature* **1991**, *351*, 624−629.
(11) Brino, L.; Urzhumtsev, A.; Mousli, M.; Bronner, C.; Mitschler, A.; Oudet, P.; Moras, D. Dimerization of *Escherichia coli* DNA-gyrase B provides a structural mechanism for activating the ATPase catalytic center. *J. Biol. Chem.* **2000**, *275*, 9468−9475.
(12) Lewis, R. J.; Singh, O. M.; Smith, C. V.; Skarzynski, T.; Maxwell, A.; Wonacott, A. J.; Wigley, D. B. The nature of inhibition of DNA gyrase by the coumarins and the cyclothialidines revealed by X-ray crystallography. *EMBO J.* **1996**, *15*, 1412−1420.
(13) Tsai, F. T. F.; Singh, O. M.; Skarzynski, T.; Wonacott, A. J.; Weston, S.; Tucker, A.; Pauptit, R. A.; Breeze, A. L.; Poyser, J. P.; O'Brien, R.; Ladbury, J. E.; Wigley, D. B. The high-resolution crystal structure of a 24-kDa gyrase B fragment from *E. coli* complexed with one of the most potent coumarin inhibitors, clorobiocin. *Proteins* **1997**, *28*, 41−52.
(14) Carlson, H. A. Protein flexibility and drug design: how to hit a moving target. *Curr. Opin. Chem. Biol.* **2002**, *6*, 447−452.
(15) Mattos, C. Protein−water interactions in a dynamic world. *Trends Biochem. Sci.* **2002**, *27*, 203−208.
(16) Carlson, H.; McCammon, J. Accommodating protein flexibility in computational drug design. *Mol. Pharmacol.* **2000**, *57*, 213−218.
(17) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindylaov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(18) Sirockin, F.; Sich, C.; Improta, S.; Schaefer, M.; Saudek, V.; Froloff, N.; Karplus, M.; Dejaegere, A. Structure activity relationship by NMR and by computer: A comparative study. *J. Am. Chem. Soc.* **2002**, *124*, 11073−11084.
(19) Miranker, A.; Karplus, M. Functionality maps of binding sites: a multiple copy simultaneous search method. *Proteins* **1991**, *23*, 472−490.
(20) Joseph-McCarthy, D.; Hogle, J. M.; Karplus, M. Use of the Multiple Copy Simultaneous Search (MCSS) method to design a new class of picornavirus capsid binding drugs. *Proteins* **1997**, *29*, 32−58.
(21) Lamour, V.; Hoermann, L.; Jeltsch, J. M.; Oudet, P.; Moras, D. An open conformation of the thermus thermophilus gyrase B ATP-binding domain. *J. Biol. Chem.* **2002**, *41*, 7217−7223.
(22) Schaefer, M.; van Vlijmen, H. W. T.; Karplus, M. Electrostatic contributions to molecular free energies in solution. *Adv. Protein Chem.* **1998**, *51*, 1−57.
(23) van Vlijmen, H. W. T.; Schaefer, M.; Karplus, M. Improving the accuracy of protein pK_a calculations: conformational averaging versus the average. *Proteins* **1998**, *33*, 145−158.
(24) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. Electrostatics and diffusion of molecules in solution: Simulations with the University of Houston Brownian Dynamics Program. *Comput. Phys. Commun.* **1991**, *62*, 187−197.
(25) Brooks, B. R.; Bruccoleri, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: a program for macromolecular energy minimization and dynamics calculations. *J. Comput. Chem.* **1982**, *4*, 187−217.
(26) MacKerell, A. D. J.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, S. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential form molecular modeling and dynamics studies of proteins. *J. Phys. Chem.* **1998**, *102*, 3586−3616.
(27) Brunger, A. T.; Karplus, M. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. *Proteins* **1988**, *4*, 148−156.
(28) Neria, E.; Fischer, S.; Karplus, M. Simulation of activation free energies in molecular systems. *J. Chem. Phys.* **1996**, *105*, 1902−1921.
(29) Caflisch, A. Computational combinatorial ligand design: Application to human alpha thrombin. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 372−396.
(30) Caflisch, A.; Schramm, H. J.; Karplus, M. Design of dimerization inhibitors of HIV-1 aspartic proteinase: A computer-based combinatorial approach. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 161−179.

(31) Boehm, H.; Boehringer, M.; Bur, D.; Gmuender, H.; Huber, W.; Klaus, W.; Kostrewa, D.; Kuehne, H.; Luebbers, T.; Meunier-Keller, N.; Mueller, F. 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*. **2000**, *43*, 2664−2674.

(32) Althaus, I. W.; Dolak, L.; Reusser, F. Coumarins as inhibitors of bacterial DNA gyrase. *J. Antibiot.* **1988**, *41*, 373−376.

(33) Schechner, M.; Dejaegere, A. P.; Stote, R. H. Effects of loop conformation on pK_a and ligand binding in DNA gyrase B. *Int. J. Quantum Chem*. **2004**, *98*, 378−387.

(34) Holdgate, G. A.; Tunnicliffe, A.; Ward, W. H.; Weston, S. A.; Rosenbrock, G.; Barth, P. T.; Taylor, I. W.; Paputit, R. A.; Timms, D. The entropic penalty of ordered water accounts for weaker binding of the antibiotic novobiocin to a resistant mutant of DNA gyrase: A thermodynamic and crystallographic study. *Biochemistry* **1997**, *36*, 9663−9673.

(35) Ferroud, D.; Collard, J.; Klich, M.; Dupuis-Hamelin, C.; Mauvais, P.; Lassaigne, P.; Bonnefoy, A.; Musicki, B. Synthesis and biological evaluation of coumarincarboxylic acids as inhibitors of gyrase B. L-Rhamnose as an effective substitute for L-noviose. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2881−2886.

(36) Laurin, P.; Ferroud, D.; Klich, M.; Dupuis-Hamelin, C.; Mauvais, P.; Lassaigne, P.; Bonnefoy, A.; Musicki, B. Synthesis and in vitro evaluation of novel highly potent coumarin inhibitors of gyrase B. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2079−2084.

(37) Laurin, P.; Ferroud, D.; Schio, L.; Klich, M.; Dupuis-Hamlin, C.; Mauvais, P.; Lassaigne, P.; Bonnefoy, A.; Mucischi, B. Structure−activity relationship in two series of aminoalkyl substitued coumarin inhibitors of gyrase B. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2875−2880.

(38) Schaefer, M.; Sommer, M.; Karplus, M. pH-dependence of protein stability: Absolute electrostatic free energy differences between conformations. *J. Phys. Chem. B* **1997**, *101*, 1663−1683.

(39) Joseph-McCarthy, D.; Alvarez, J. C. Pharmacophore-based molecular docking to account for ligand flexibility. *Proteins* **2003**, *51*, 172−188.

JM0311184