Using SAR and QSAR analysis to model the activity and structure of the quinolone-DNA complex

Belsis Llorente, Fabrice Leclerc, Robert Cedergren

Pergamon

Bioorganic & Medicinal Chemistry, Vol. 4, No. 1, pp. 61–71, 1996

0968-0896/96 $15.00+0.00

0968-0896(95)00172

Using SAR and QSAR Analysis to Model the Activity and Structure of the Quinolone–DNA Complex

Belsis Llorente,† Fabrice Leclerc and Robert Cedergren*
Département de Biochimie, Université de Montréal, Montréal, PQ H3C 3J7, Canada

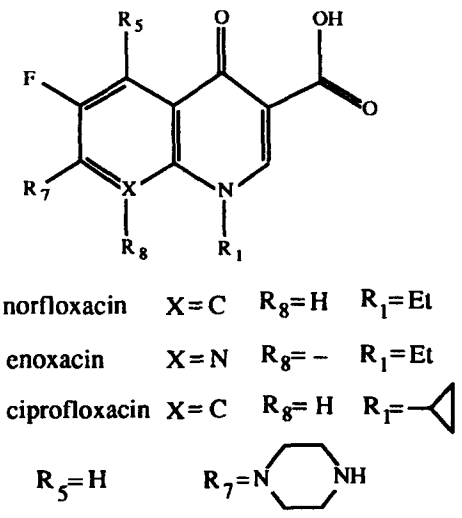
Abstract—A set of 78 quinolone derivatives were used in a structure–activity study to identify structural features correlating with antibacterial activity. Distinct combinations of functional properties were identified for Gram-negative and Gram-positive bacteria. 3-D Quantitative structure–activity relationship (QSAR) studies identified specific hydrophobic, topologic and electronic properties of the molecules for both in vitro and in vivo activities. From these results, a three-dimensional model of a DNA–quinolone complex was built using molecular modeling techniques. It was based on the intercalation of quinolone into the double helix of DNA. We conclude that the intercalation model is consistent with most available data on the structure of the quinolone complex. This predicted structure is stabilized by the binding of magnesium ion with the sp^2 oxygens present in quinolone, a phosphate and a purine base of the DNA. Substituents R_1 and R_7 are predicted to make hydrophobic interactions in the major and minor groove of DNA, respectively. R_7 Could also form hydrogen bonds with amino groups of guanines and the aspartic acid residue at position 87 in DNA gyrase subunit A.

Introduction

The quinolones shown in Figure 1 are broad spectrum antibiotics which act by inhibiting bacterial deoxyribonucleic acid (DNA) gyrase required for initiation and propagation of DNA synthesis. Although DNA gyrase is composed of an A and B subunit, the inhibitory effect of quinolones is mediated via the catalytic A subunit.[1] In spite of the large number of quinolone analogues which have been synthesized and tested reliable structural information on the requirements for drug binding, including the identity of the binding site itself, is sparse. Indeed, it is not known with certainty whether DNA or the protein is the primary target, although both factors are necessary for inhibition.

Currently three structural models have been put forward to account for the action of quinolones. These models differ widely in the manner in which the components of the complex are juxtaposed and which functional groups are implicated, although they all require a direct interaction of the drug with DNA. Differences in the models arise also based on the use of a single- or a double-stranded DNA target. The Shen[2,3] model involves hydrogen-bond type interactions between single-stranded DNA and the quinolone as well as a stacked dimerization of the drug. The more recent model of Palumbo suggests that the drug is intercalated between nucleotides in a region of single-stranded DNA; stabilization of this interaction is assured by a Mg^{2+} bridge between the quinolone and a phosphate of the DNA backbone.[4] This latter point may be essential since quinolones have a low affinity for DNA under many conditions. Finally, the group of Hurley[5] proposes an attractive and well documented double-stranded DNA intercalation model whose stabilization is assisted by a Mg^{2+} bridge as in the Palumbo model. This proposal is strengthened by physical studies and suggests that the role of DNA gyrase is to partially unwind the DNA helix thereby permitting access by the drug.

Generally, successful modeling is rooted in the ability to identify and thereafter incorporate or predict the role of specific functional groups in the formation of the complex between drug and target. In the case of quinolones, attempts at discerning these key structural



1 norfloxacin X=C R_8=H R_1=Et

2 enoxacin X=N R_8=– R_1=Et

3 ciprofloxacin X=C R_8=H R_1=

R_5=H R_7=

Figure 1. Examples of typical quinolones.

†On leave from the Center of Pharmaceutical Chemistry, Apartado postal 16042, Havana, Cuba.

features have been complicated by the proliferation of analogues and their various efficiencies against different bacteria. However, data from a variety of compounds demonstrating a common activity have often been the basis of analysis by structure–activity relationship (SAR) calculations. Structural correlation of gyrase inhibition by quinolones have been previously attempted, and QSAR studies have shown that hydrophobic, electronic, and steric parameters of quinolones play equally important roles in their biological activity.[6–9] Nevertheless, correlation of specific groups with activity are often complicated by the multidimensional nature of the data; that is, conformational preferences of the analogues, factors concerned with in vivo bioavailability due to membrane effects, biostability, and the innate ability to partake in the DNA–gyrase complex.

Due to the economic potential of quinolone-based therapies as well as fundamental issues pertaining to DNA interactions with small ligands, we initiated an extensive SAR and QSAR study of the quinolone analogues in order to more firmly establish the structure and properties of key functional groups. These data together with energy-based molecular modeling studies have allowed us to evaluate different aspects of the drug–DNA complex and to propose a detailed three-dimensional model of the interaction. This model is based on a primary target of double-stranded DNA and is supported by interactions which involve the key functional groups pointed to by the SAR and QSAR studies.

Results

Structure–activity relationships

Activity classes were defined for active Gram-negative inhibitors (G^-), for active Gram-positive inhibitors (G^+) and for compounds with poor inhibition of both bacterial groups (P), see Experimental. Classification of the molecules in the training set of 67 compounds (Table 1) required the definition of several tasks generated by varying the parameters. Generally, the suggested tolerance values for the parameters in the APEX program were used; these provide that centers with properties within 20% of each other are considered equal. However, the strong dependence of classification on the tolerance of the *distance between centers* parameter prompted us to experiment with this value (Dr Eric Vorpagel, private communication). Different absolute tolerance values from 0.4 to 1.6 Å were tested for equivalence in distance. The value of 1.2 Å obtained the least number of incorrectly classified molecules in the training set and was therefore used throughout.

The centers and their properties listed in Table 2 allowed the identification of 10, 25, and 14 biophores for the G^-, G^+, and P activity classes, respectively. Inversely, when these biophores are used for classification of the training set, 90% of the G^- class are recovered, whereas 91% and 88% of the compounds in the G^+ and P classes are found. In order to identify common and distinguishing features in the G^+ and G^- classes, the frequency of appearance was determined for each of the 15 centers in the identified biophores (Fig. 2). Comparison of the frequencies in the G^- class with those of the G^+ illustrates a great similarity with the exception of the frequencies of centers 12, 13, and 15 which deal with the R_7 substituent. On the other hand, the presence of the carbonyl and carboxyl oxygens, the nitrogens at C_1 and R_7 positions (centers 1 and 7) and the two 6-member condensed ring system are strong predictors of good activity against both types of bacteria.

In the light of the differences observed between G^- and G^+ classes, it seemed prudent to ascertain the consistency of the division between Gram-negative and Gram-positive bacteria with regard to the activity of quinolones. The data were thus separated into 10 activity classes; that is, an active and an inactive class for each of the five bacteria. From the biophores identified for the five active classes, the frequency of all centers in the biophores of each class was calculated as in the case of the experiments with the G^- and G^+ classes. The graphic representation of these data is shown in Figure 3. Although differences in biophores among each of the bacterial types are evident, it was very encouraging to observe very similar profiles of the frequency values for the three Gram-negative and two Gram-positive bacteria. *Pseudomonas aeruginosa* displays the largest variance among the Gram-negative bacteria (Fig. 3a), a likely consequence of its more distant phylogenetic relationship to the two other members of this class. The absence of center 15, the substituent R_7, in biophores generated in Gram-negative bacteria (Figs 2 and 3a) is observed in both qualitative analysis.

Biophores derived from the training set were then used to classify the test set (Table 1). To this end, the test set was scanned for the presence of biophores from each activity class that permit the spatial superposition of at least 60% of the non-hydrogen atoms of the compounds in that class. When biophores from more than one class were present in the test compound, the class of the most probable biophore was assigned (Table 3). As shown, all compounds in the test set are correctly classified into the G^-, G^+, and P classes with the exception of p3 whose placement is complicated by the fact that the probability values associated with biophores contained in both G^- and G^+ classes do not differ significantly.

The good classification of compounds into G^- and G^+ classes suggests that some structural features were identified as important and specific to each activity class. Among these, centers 12, 13, and 15 are the most distinctive for classes G^- and G^+. Since these centers are contained in the R_7 substituent, requirements for this part of the molecule must be different for

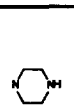
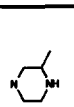
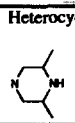
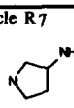
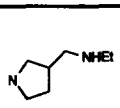
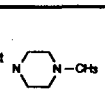
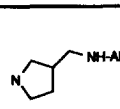
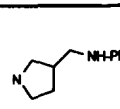
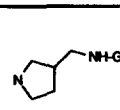
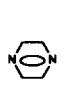
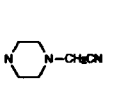
Gram-negative and Gram-positive bacteria. The biophores specific to G^- class generally contain the piperazine ring (center 7, 12, and 13) and thus represent a structural pattern strongly correlating with good activity. On the other hand, pyrrolidine derivatives are more important in the G^+ class.

3-D-QSAR analysis

3-D-QSAR Models were created for the minimum inhibitory concentration (MIC) assay for all five bacteria and the in vitro DNA cleavage activity of purified DNA gyrase from *Escherichia coli*. Biophores that were found in at least 85% of the compounds in the training set were used to obtain mathematical expressions which predict the biological activity of compounds in both the training and test sets. The variables, their coefficients and statistical significance tests of the mathematical model which best represented the data for each bacterium are shown in Table 4. In the case of the Gram-negative bacteria biophore 2 (Figs 4a, 4b, and 4c) give the most satisfactory prediction of activity, whereas, biophore 395 and 173 predicted best the activity of the quinolones in Gram-positive and in in vitro tests (Figs 4d, 4e, and 4f). The confidence intervals for the regression coefficients of all variables were higher than 95%, guaranteeing that the equations produced a good statistic fit of the data. The probability of a chance correlation for all models approached zero. The variable themselves have the value of one or zero depending on the presence or absence, respectively, of H-acceptor ability, hydrophobicity, and ring structure of centers. In spite of the fact that the correlation coefficients in the six QSAR models vary from 0.76 to 0.86, the biological activity of quinolones are quite well explained by these equations when the structural heterogeneity of the analogue collection and the complexity of biological phenomena is taken into account.

Comparison of the coefficients in Table 4 brings to light several properties of the quinolone inhibition. First, comparisons within the Gram-negative and Gram-positive classes show no major differences. However, coefficients determined from the in vitro assay using the *E. coli* enzyme, especially with regard to the requirement for hydrophobicity (the X_2 Coefficient), differ significantly from those determined

Table 1. Chemical structures of the 78-member compound set

Quinolone nucleus R_5,R_8,R_1	Heterocycle R7											
											
H,H,Et	a1	a2	a3	a4	a5	q4					q7	
H,F,Et	b1	b2	b3	b4	b5	q23						
H,Cl,Et	c1	c2	c3	c4	c5							
H,H,c-C_3H_5	f1	f2	f3	f4	f5							
H,F,c,-C_3H_5	g1	g2	g3	g4	g5		22d	22h				
H,Cl,c-C_3H_5	h1	h2	h3	h4	h5							
H,H,$F_2C_6H_4$		d2										
H,Oflox[a]				e4	e5							
NH_2,H,Et	i1				i5							
NH_2,F,Et	j1	j2	j3	j4	j5							
NH_2,Cl,Et	k1		k3		k5							
NH_2,H,c-C_3H_5	n1	n2		n4	n5							
NH_2,F,c-C_3H_5	o1	o2	o3	o4	o5					24f		
NH_2,Cl,c-C_3H_5	p1	p2	p3	p4	p5							
NH_2,H,$F_2C_6H_4$		l2										
NH_2,Oflox				m4	m5							
OH,F,Et	q1	q2	q3	q4	q5							
OH,H,c-C_3H_5	r1			r4								
OH,F,c-C_3H_5	s1	s2		s4								
H,-,Et[b]	q10											
H,-,c-C_3H_5										21i		

Circled compounds composed the test set; all others were in the training set.
[a]Substituents R_1 and R_8 are linked like in ofloxacin compound.[16]
[b]The substituent at position 8 is a nitrogen rather than a carbon.

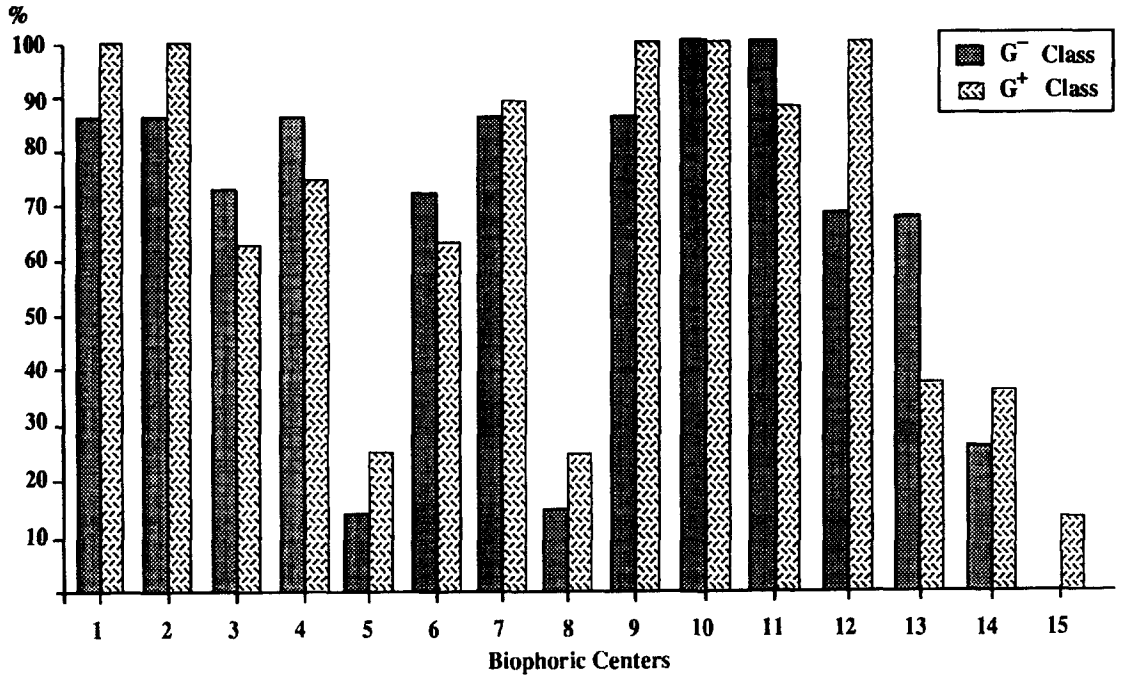
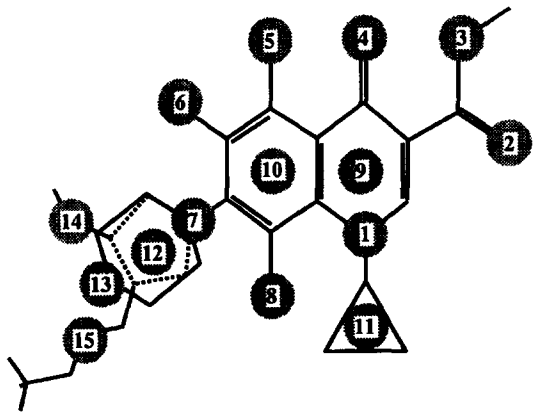


Figure 2. Frequency of appearance of the 15 centers in biophores identified for classes G^+ and G^-.

from *E. coli* in vivo. A major difference between MIC and gyrase cleavage activity is also observed by the presence of a hydrophobic region (eq 6) in substituent R_7. The different requirements of R_7 for activity in G^- and G^+ classes, as well as the in vitro assay suggest strongly that this substituent plays an important role in bioavailability, an aspect of which would be cell penetration. This notion is reinforced by the fact that the major differences in coefficients are between classes rather than within them and that these two classes of bacteria differ by their membrane composition. On the other hand, differences in biostability or metabolism of the quinolones in different bacteria can not be ruled out at the present. Furthermore, the common hydrophobic pathway used by Gram-negative and Gram-positive bacteria for lipid bilayer penetration is likely the basis for the presence of Total_Hydrophobicity and Quadr_Hydrophobicity in eqs 1–5. The importance of hydrophobicity for cell penetration is also underlined by the difference between the correlation coefficients for these variables in eqs 6 and 1–5 (see Table 4) and agrees with previous results obtained by Bazile et al.[10] However the conclusion of these authors concerning the influence of molecular size for *Staphylococcus aureus* during cell penetration does not agree with our results. The relationship between Total_Refractivity and MIC in eqs 1–5 shows that the steric bulk effect is important for both kind of bacteria.

Table 2. The 15 atomic centers and their properties used for biophores



Biophoric center		Charge	H-acceptor	Properties H-donor	Ring-size	π-electrons
1	N	x	-	-	-	-
2, 4	O	x	x	-	-	-
3	O	x	x	x	-	-
5	N, O	x	-	x	-	-
6	F	x	x	-	-	-
7	N	x	-	-	-	-
8	F, Cl	-	x	-	-	-
9, 10, 11, 12[a]	Rings	-	-	-	x	x
13, 14, 15	N	x	x	x	-	-

[a]Piperazine and pyrrolidine rings.

The quantitative prediction capacity of these 3-D-QSAR models was evaluated by calculating the biological activity of each compound in the test set using the six equations. In the test set the correlation coefficients between experimental and predicted values are 0.87, 0.86, 0.82, 0.77, 0.89, and 0.60 for eqs 1–6, respectively, with probability greater that 99.99% that they are not zero. Table 5 exhibits the prediction in the test set and the standard deviation values (SD) for each QSAR model. The predicting ability of these equations can be judged by the low standard deviation

values which are comparable to the RMSP values from the cross validation test in Table 4.

3-D Model building

The convincing evidence presented by the Hurley group[5] on the implication of double-stranded DNA prompted us to evaluate different possible interactions using this form of DNA as the target molecule. We have thus constructed a model of quinolone interaction in which the quinolone is intercalated between the base pairs using the crystal structure[11] of the complex between DNA and doxorubicin (an anthracycline analogue) as a template in order to insert the drug into a preformed intercalation binding site (Fig. 5). Two configurations were tried: the two aromatic rings of quinolone were superimposed on either the a/b or the b/c rings of doxorubicin while maintaining the maximum overlap of the aromatic rings. The two stacking patterns were discriminated by steric constraints imposed by the two hydrophobic R_1 and R_7 substituents. Thus, in one orientation the R_7 substituent is in the major groove, and in the other it is in the minor groove (vice versa for the R_1 substituent).

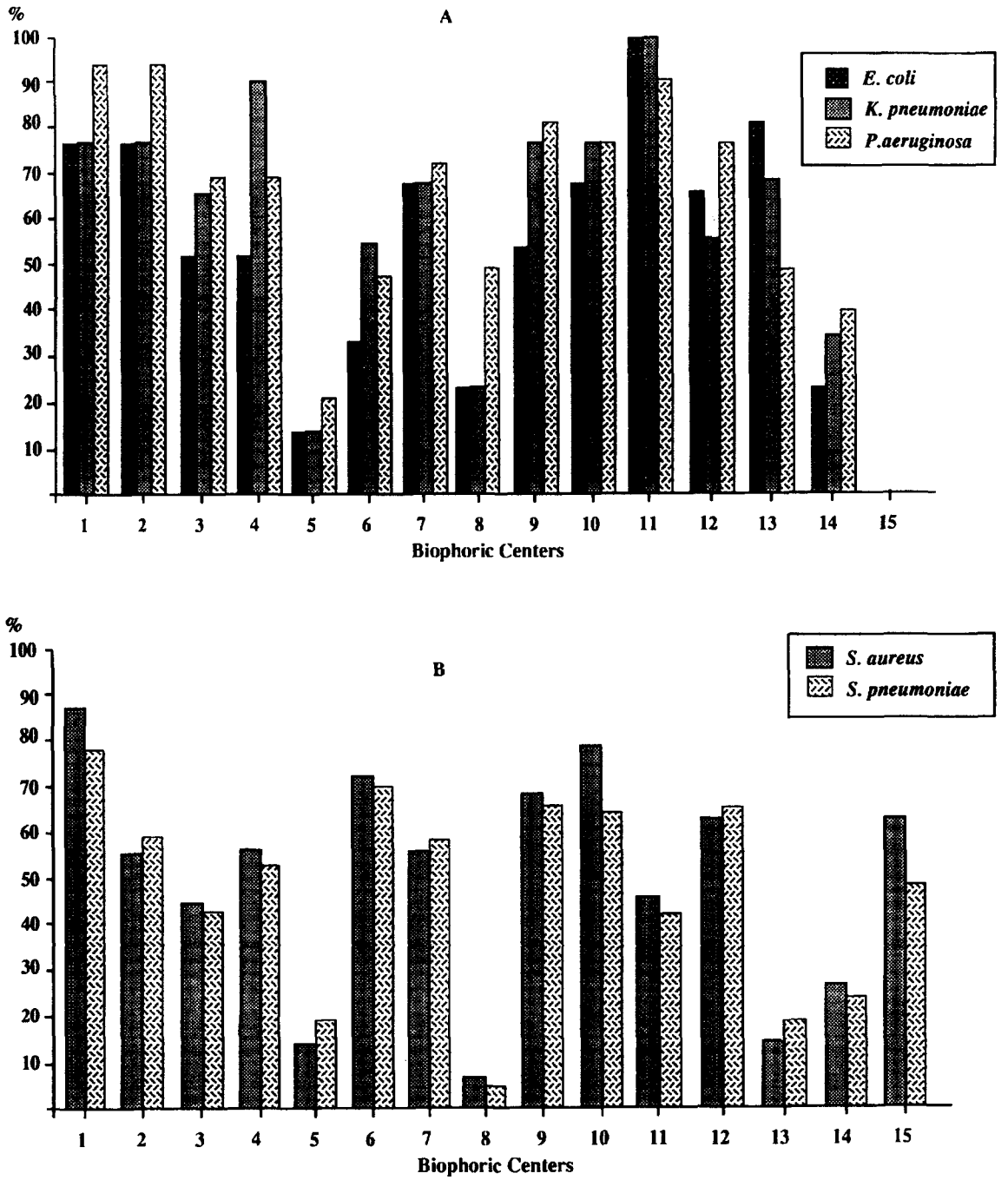


Figure 3. Frequency of appearance of the 15 centers in biophores identified for (A) Gram-negative and (B) Gram-positive bacteria.

Table 3. Qualitative classification of the test set

Compounds Test set	Observed class	Predicted class
b3	P	P
c5	G^+	G^+
i5	G^+	G^+
p3	G^+	G, G^{+a}
p4	G^-, G^+	G^-, G^+
q2	P	P
q3	P	P
q5	G^+	G^+
q10	P	P
q23	P	P
q30	P	P

[a]A false positive classification error in class G^-.

In this way the R_7 and R_1 substituents determine the nature of the hydrophobic interactions involved in the stabilization of the complex. The two orientations were then evaluated in relation to their ability to bind Mg^{2+}, a key interaction which was demonstrated experimentally to play an important role in the formation of the complex. Only the orientation of the drug that involves hydrophobic interactions of R_7 in the minor groove allows the formation of a favorable magnesium ion binding site composed of the two sp^2 oxygens (the ketone and carboxylic group functions of quinolone), two oxygens of a phosphate group, a nitrogen N^7 and oxygen O^6 of a neighboring purine nucleotide, and its corresponding phophate oxygens. The presence of two hydrophobic groups, R_1 and R_7, in the two grooves of the DNA helix permits the firm anchoring of the intercalated drug, which is stabilized additionally by stacking interactions.

The sequence of the gyrase

In order to bring to this discussion possible roles of the A subunit of the gyrase protein, we have compared and aligned the sequence of these subunits from 4 Gram-negative and 5 Gram-positive bacteria. The quinolone resistance determining region (QRDR) is composed of amino acids 67 to 106.[12,13] Position 83, a serine in the *E. coli* gyrase A, has been shown particularly critical for the acquisition of quinolone resistance. A *gyrA* mutation at this position (Ser to Leu) is sufficient to confer quinolone and fluoroquinolone resistance in both the Gram-positive *S. aureus, Staphylococcus epidermis*, and Gram-negative bacteria *C. jejuni,* and *P. aeruginosa*,[12] whereas, two concomitant mutations at 83 and 87 (an aspartate to glutamate) are necessary for *E. coli* resistance to fluoroquinolones.[12] The presence of a hydrophilic serine residue in position 84 in Gram-positive bacteria *(S. aureus, S. epidermis, Streptomyces coelicolor)* instead of a hydrophobic residue (alanine) as in Gram-negative bacteria constitutes a major difference in this region between the two groups. The increased hydrophobicity of this region in Gram-negative bacteria could be related to their requirement of higher hydrophobicity as determined by the QSAR studies.

Discussion

In the SAR data, the very high probability of biophores encompassing two oxygen atoms (centers 2 and 4, Fig. 1) shows that they serve an essential role and are therefore likely to be involved in interactions with the receptor. The importance of two condensed 6-member rings (center 9 and 10) which were also identified as essential for the activity may be related to their high π-electron density, a key characteristic of molecules involved in stacking interactions with DNA. The nitrogen atom (center 7) in R_7 is also an important center for activity (see Figs 2 and 3), since its π-population and HOMO properties are present in eqs 1–6. This result suggests that an atom with high electron density directly attached to the aromatic ring at position C_7 is desirable for high antibacterial activity. The high electron density of this atom could increase the stacking interactions with DNA bases by increasing the π-electron density of the aromatic ring system or implicate the R_7 substituent in an ionic interaction with the gyrase protein. This last suggestion is also supported by mutants of *gyrA* that decrease the quinolone binding affinity.[14,15] The presence of a hydrophobic or a ring moiety at the N^1 position in eqs 1–6 gives additional evidence of the importance of hydrophobic forces for drug–receptors interactions.

The role of the above factors in drug–receptor interactions is well explained by the DNA–quinolone model of Figure 5, since the main structural features reside in the anchoring of the drug by hydrophobic interactions of the R_1 substituent in the major groove and the R_7 substituent in the minor groove. The importance of a hydrophobic region in R_7 for gyrase cleavage (see eq 6, Table 4) also suggests the existence of hydrophobic interactions of this substituent within the minor groove. The tertiary amino group of the piperazine ring in R_7 can be stabilized in a sequence-specific manner by hydrogen-bonding of guanine amino groups distributed in the minor groove on both sides of the base stacking plane.

Stabilization of the complex by the binding of the magnesium ion is a particularly strong prediction of this model, since the magnesium binding pocket would be composed of sp^2 oxygens from the carboxyl and carbonyl groups of the quinolone and the nitrogen N^7 of a purine, the oxygen O^6 of guanine and the neighboring phosphate oxygens. In order to establish a more accurate magnesium binding model, a short molecular dynamics simulation (21 ps) was performed on a DNA–quinolone complex surrounded by explicit solvent and counterions (sodium ions). The magnesium ion was preferentially coordinated by water molecules which in turn were firmly associated with these atoms which constitute the binding pocket. The average distance between the magnesium ion and the

coordination sites is 3.7 Å so in order for the ion to remain hexa-hydrated, four water molecules of the hydration shell would have to associate with the four electronegatively charged atoms previously mentioned.

This prediction is in agreement with the binding mode found in one of the known magnesium ion binding sites in phenylalanine tRNA determined by X-ray crystallography. In this site, the magnesium ion binds in the major groove of two successive Watson–Crick paired guanines (anticodon stem) and in the proximity of two phosphate groups on the opposite strand of the RNA helix. The arrangement of negatively charged atoms around magnesium is such that it is at the center of mass of the 4 strong coordination sites: the O^6 and N^7 of guanine, and the two sp^2 oxygens of quinolone. The fact that only one of the two oxygens of the carboxyl group is at the coordination site is consistent with the QSAR result obtained for the inhibition of *E. coli* DNA gyrase (Biophore 173, Fig. 4). With only the sp^2 oxygen of the carboxyl group, this biophore contrasts with the biophores determined using in vivo activities for which there are two oxygen centers. This analysis also supports the idea of a different role for the carboxyl group in DNA binding of quinolone and bioavailability.

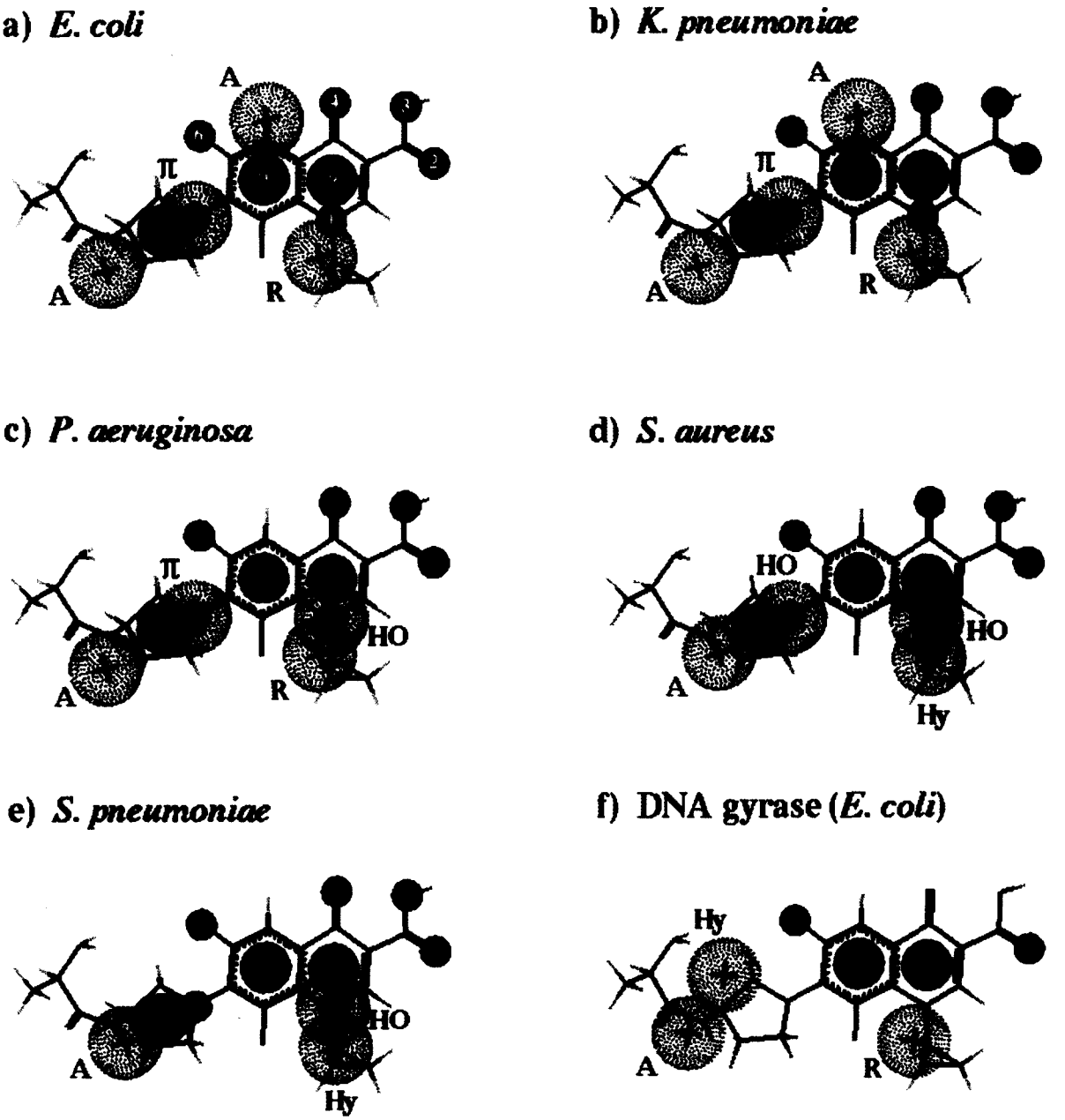


Figure 4. 3-D-QSAR models for the calculation of activities in different bacteria and in vitro. The center present in biophores are represented by blue spheres. The reference numbers of these centers are shown in (a). The properties associated with each center are listed in Table 2. The properties associated with secondary sites, represented by red dot spheres, consist of π-electron density (π) hydrogen-bond acceptor sites (A), hydrophobic center (Hy), frontier orbital indices: HOMO (HO), and ring centroids (R).

Table 4. Regression coefficients for variables in 3-D-QSAR models obtained for all bacteria

Equation	Biophore (*n*)	Bacteria	X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	R	F	RMSA	RMSP
1	2 (66)	*E coli* MIC	8.58	1.22 (R_7)	1.9 5	−0.36	−0.07	−0.5 (R_7) −0.45 (R_5)	0.87 (R_1)			0.82	16.50	0.39	0.44
2	2 (66)	*K. pneu* MIC	8.62	1.57 (R_7)	1.9 5	−0.36	0.08	−0.59 (R_7) −0.42 (R_5)	0.96 (R_1)			0.85	22.59	0.36	0.40
3	2 (66)	*P. aure.* MIC	8.37	1.36 (R_7)	1.8 3	−0.36	−0.08	−0.79 (R_7)	0.89 (R_1)	−3.05 (N_1)		0.86	23.49	0.35	0.37
4	395 (60)	*S. aure.* MIC	7.05		0.85		−0.04	−0.62 (R_7)		−5.49 (N_1) −2.26 (R_{78})	−0.70 (R_1)	0.76	12.25	0.47	0.52
5	395 (60)	*S. pneu.* MIC	7.87		0.85		−0.05	−0.62 (R_7)		−6.43 (N_1)	−1.1 (R_7)	0.78	16.45	0.59	0.69
6	173 (66)	*E. coli* GC	4.93		0.25		−0.03	−0.21 (R_7)	0.69 (R_1)		−0.56 (R_7)	0.79	19.57	0.28	0.29

n is the number of compounds in the equation; R is the correlation coefficient; F is a significance test; RMSA is the standard error of the estimate; and RMSP is the standard error of activity prediction calculated using cross validation. Variables included in the equations are X_0, Independent term of the equation; X_1, π-Population; X_2, Total_hydrophobicity; X_3, Quadratic_hydrophobicity; X_4, Total_refractivity; X_5, H-Acceptor presence; X_6, Ring presence; X_7, HOMO; and X_8, Hydrophobic presence.

The fluorine and the secondary amino group of the piperazine ring remain accessible in this model for interactions possibly with the QRDR region of the gyrase. The conserved serine 83 in the DNA gyrase subunit A could be a target for electrostatic interactions with the fluorine atom. Similarly, the conserved acid (aspartic or glutamic) residue in position 87 is a possible candidate for interactions involving the side chain with one of the two amino groups of the piperazine ring. Models can also be obtained with quinolones where the piperazine ring has been replaced by the pyrrolidine ring.

Substituents at positions C^5 and C^8 do not seem to participate directly in drug–receptor interactions, although they effect the planar structure of quinolone and therefore the disposition of substituents R_1 and R_7. We selected some compounds with modifications in R_5 and R_8 to study their influence on the planar structure of the two condensed rings. The dihedral angle of the R_5 substituent in compounds 1f–4, 1h–1, 1n–2, 1o–1, 1r–4, and 1s–4 was calculated. The values of the angles were 0.1°, −2.7°, −8.4°, −13.3°, 0.1°, and −0.1°, respectively, demonstrating that the addition of a 5-amino group and an fluorine or chlorine atom at position 8 affected the planarity of the benzyl and pyridine rings respectively due to steric constraints. This modification in the geometry of the aromatic rings could play a role in the adjustment of the drug to its receptor or to proteins during cell penetration. Figure 6 summarizes possible interactions involved in the DNA–gyrase–quinolone complex derived from the quinolone intercalation model.

Conclusions

Due to the strong experimental evidence suggesting a role for magnesium ions in the antibacterial activity of these drugs and for intercalation of quinolones into

Table 5. Prediction of activity for compounds in the test set

Compounds Test Set	Eq 1		Eq 2		Eq 3		Eq 4		Eq 5		Eq 6	
	E. coli		*K. pneumoniae*		*P. aeruginosa*		*S. aureus*		*S. pneumoniae*		*E. coli*	
	Obsd	Calcd	Obsd	Calcd	Obsd	Calcd	Obsd	Calcd	Obsd	Calcd	Obsd	Calcd
b3	3.26	3.44	2.96	3.09	1.47	2.12	2.36	2.48	1.76	2.16	1.98	1.90
c5	3.30	2.94	2.99	2.59	1.50	2.02	3.60	3.39	3.90	3.52	2.20	2.13
i5	2.67	2.95	2.08	2.55	1.18	1.67	2.67	3.16	2.97	2.89	1.70	1.95
p3	3.91	3.77	2.38	2.83	1.48	2.05	3.61	3.75	3.61	3.83	1.31	2.44
p4	4.18	4.17	2.69	2.52	1.50	1.89	3.58	3.50	3.58	3.64	1.28	2.34
q2	2.96	3.20	3.31	3.38	2.41	2.38	1.77	2.39	1.47	2.01	1.87	2.02
q3	2.68	3.19	3.88	3.84	2.68	2.93	1.78	2.59	0.88	2.19	1.88	1.99
q5	2.99	2.87	2.66	2.85	1.47	2.13	2.99	2.84	3.29	2.78	2.20	2.07
q10	3.51	3.64	3.51	3.33	2.60	2.63	2.01	2.45	2.01	2.26	1.81	2.32
q23	3.24	3.41	2.94	3.07	2.64	2.18	2.94	2.18	2.05	1.88	2.15	1.89
q30	2.95	3.00	2.65	2.64	1.75	1.90	2.35	2.18	1.75	1.92	1.55	2.00
Standard Deviation	0.24		0.24		0.44		0.45		0.50		0.53	

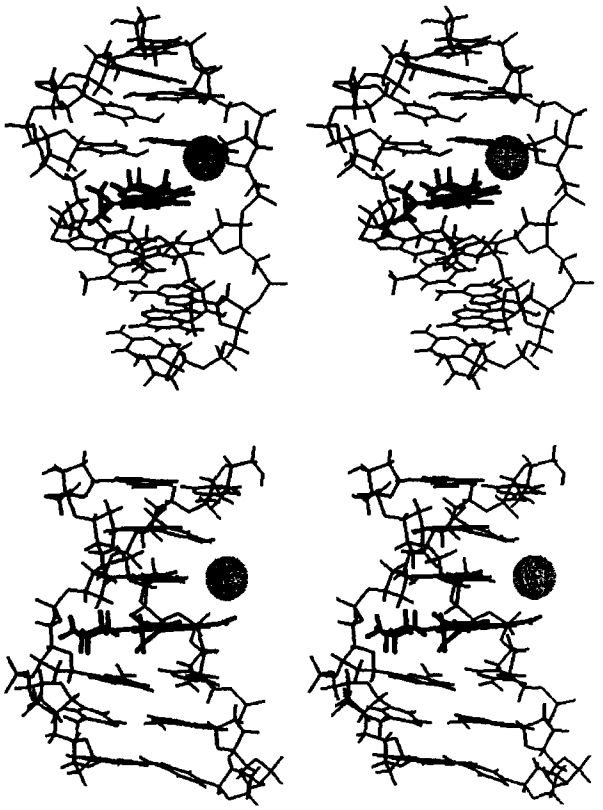


Figure 5. Molecular model for DNA–quinolone complex based on intercalation.

DNA helix, our model was based heavily on these aspects (Fig. 5). This model details stabilization by the binding of magnesium ion with the sp^2 oxygens of quinolone, a phosphate, and purine base in the DNA. Substituents R_1 and R_7 are involved in hydrophobic interactions in the major and minor groove of DNA, respectively. The gyrase A mutations suggest that amino acids in positions 83 and 87 contribute to the binding specificity by forming direct interactions with some regions of quinolones (R_6 and R_7), or by changing the interface gyrase A/gyrase B. R_7 Could also form hydrogen bonds with a phosphate group in the DNA and the aspartic amino acid at position 87 in DNA gyrase subunit A.

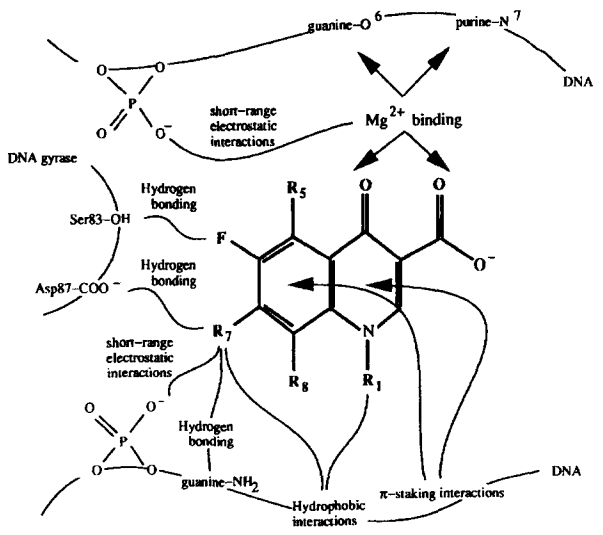


Figure 6. Summary of interactions inferred from modeling and structure–activity analyses.

Our model does not consider the possibility that two quinolone molecules could be involved in two different aspects of complex formation as previously suggested.[5] Although experimental evidence does weigh in favor of this possibility, neither SAR or QSAR techniques are useful to sort out two different functions in the grouping of compounds. Even though this is clearly a shortcoming of this type of analysis, the ability of identifying key functional groups in a collection of molecules with the same activity can not be underestimated in its application to drug design and model building. The correlation between qualitative, quantitative, and modeling data obtained here suggests that the knowledge gained by this study may be applied to the design and evaluation of novel antimicrobial agents. Clearly this applies to the R_7 substitutent which we have identified as being important not only for the activity, but also for distinguishing between bacteria.

Experimental

Biological activity and compounds

A 78-member compound set composed of various substitutions at positions N^1, C^5, C^7, and C^8 of the basic quinolone nucleus was assembled.[16,17,18] Of these, 67 compounds were included in the training and 11 in the test set. Training set selection was performed to assure representability of all functional groups in the total set. The minimum inhibitory concentration in μg mL^{-1} (MIC) against three Gram-negative, *E. coli, Klebsiella pneumoniae* and *P. aeruginosa* and two Gram-positive, *S. aureus* and *S. pneumoniae* bacteria was the measure of biological activity. Since these values deal with whole cells, they are referred to in vivo studies. The gyrase cleavage μg mL^{-1} was employed to study the potency of the compounds against bacterial DNA gyrase obtained from *E. coli.* The threshold values used to define active and inactive compounds varied for each bacteria in order to correct for the difference in overall drug sensitivity of the individual bacteria. Thus, compounds that showed no activity at a concentration of less than 0.1 μg mL^{-1} for *E. coli* and *K. pneumonia* were considered inactive, whereas a threshold of 0.8 μg mL^{-1} was used for more recalcitant *Pseudomonas* and Gram-positive bacteria. For the purpose of QSAR analysis, the reported activity was divided by the molecular weight of the compound and considered as log(1/molar MIC) or log(1/molar).

Logical-combinatorial approach

Analysis was performed by the APEX-3-D module of the Biosym InsightII system (Biosym Technologies, San

Diego, CA, U.S.A.) which is based on a combinatorial-logical analysis for the classification and prediction of biological activity proposed by Golender et al.[19] In this methodology, structural or electronic centers in the molecule are user-defined atoms, pseudoatoms such as ring centers, hydrophobic regions or hydrogen bonding sites. Based on the quinolone structure, 15 such centers were used in the present study, and the attributes of these centers which were found significant by the APEX program are shown in Table 2. Biophores are local arrays of such centers which are common to a class of molecules having a similar biological activity; biophores are automatically identified by APEX.[20] 3-D-QSAR Models are created using the three-dimensional biophoric patterns which are present in subsets of the molecules under study.

Input to APEX consists of chemical structures, a task definition file, and the biological activity information. The task file defines chemical or physical properties of centers. The chemical structures were built using the InsightII Builder module. The energy of each compound was minimized using the molecular mechanics program Discovery. The coordinates of these structures were used for further energy optimization with the quantum mechanical method AM1 incorporated in the molecular orbital package MOPAC/AMPAC.[21] Atoms were assumed to exist in the singlet state and as a neutral species.

Confidence levels of biophore prediction is based on $P(_k/f_r)$, where P is the Bayesian probability that a compound possessing the biophore f_r belongs to the activity class A_k; and the reliability P_{rk} which is the probability of a non-chalance occurrence of the biophore. The posit of a biophore is based on a threshold value assigned for each of these statistics. The quality of the model is evaluated by the probability of the placement of a compound in the appropriate activity group. In addition, an efficiency coefficient can be calculated which gives a relationship between the number of compounds that are active and those that are predicted active by the activity prediction system.

3-D-QSAR models were achieved by multivariate linear regression analyses of biophoric patterns. To select only the models having good predictive properties, the estimation of the probability of chance correlation[22] and calculation of the predicted sum of squared residuals were used. The cross-validation method was used to analyze the 3-D-QSAR models prediction capacity.

Molecular modeling techniques

3-D Models of the drug–DNA complex were constructed by molecular building techniques in the Biopolymer module of InsightII. The molecular structures were optimized by molecular mechanics computations performed on a Indigo2 R4400 workstation using the molecular mechanics package Discover (version 94.0) interfaced to the cff93 force field. The screening effect of solvent was simulated using a distance dependent dielectric constant $\varepsilon(r)=r$. The structures were refined using both the steepest descent energy minimizer (SD) and the conjugate gradients algorithm (CG). The SD algorithm was used until the maximum derivative was less than 10 kcal mol^{-1} $Å^{-1}$ then the CG algorithm was applied until the maximum derivative was less than 0.5 kcal mol^{-1} $Å^{-1}$. No non-bonded cutoff was used. During minimization, distance restraints that applied a quadratic function with a selected force constant of 20.0 kcal $Å^{-1}$ and a defined distance value of 2.9 Å between respective heavy atoms involved in hydrogen bonding were used to preserve the base pairing between the nucleotides of the DNA duplex.

The initial system (for the model in Figure 5) was the hexamer d(GpCpCpGpTpG) with 8 sodium counterions. Counterions were initially placed at 5 Å from the phosphorous atom along the O—P—O bisector. Water molecules were used in order to have 12 Å of water around the solute in all directions. The inclusion of counterions was done in order to maintain a neutral charge for the whole system including the magnesium ion. Water molecules were submitted to a rough optimization until the maximum derivative was less than 50 kcal mol^{-1} $Å^{-1}$ followed by 1 ps of molecular dynamics (MD) simulation at 300 K and constant volume. The counterions, DNA and quinolone were held fixed to their initial position. A MD simulation of 3 ps was performed on the solvent with the DNA–quinolone complex fixed at 10 K and constant volume. The solvent and counterions (sodium and magnesium ions) were heated by a gradual increase of the temperature from 10 to 300 K in steps of 50 K (3 ps at each temperature). Finally, the constraints imposed on the DNA–quinolone complex were removed and the complete system was optimized by energy-minimization until the maximum derivative was less than 1.0 kcal mol^{-1} $Å^{-1}$. The distance constraints described previously to maintain the base-pairings were imposed during the optimization.

Acknowledgments

We would like to thank Dr Eric Vorpagel for helpful insights. Robert Cedergren is a Fellow of the Evolutionary Biology Program of the Canadian Institute for Advanced Research. This work was supported by MRC of Canada.

References

1. Reece, R. J.; Maxwell, A. *Crit. Rev. Biochem. Mol. Biol.* **1991**, *26*, 335.

2. Shen, L. L.; Baranowski, J.; Pernet, A. *Biochemistry* **1989**, *28*, 3879.

3. Shen, L. L.; Mitscher, L.; Sharma, P.; O'Donnell, T.; Chu, D.; Cooper, C.; Rosen, T.; Pernet, A. *Biochemistry* **1989**, *28*, 3886.

4. Palumbo, M.; Gatto, B.; Zagotto, G.; Pald, G. *Trends Microbiology* **1993**, *1*, 232.

5. Fan, J.; Sun, D.; Yu, H.; Kerwin, S. M.; Hurley, L. H. *J. Med. Chem.* **1995**, *38*, 408.

6. Koga, H. In *Structure–Activity-Relationships-Quantitative Approaches*; Fujita, T., Ed.; Nankodo: Tokyo, 1982; pp 177–202.

7. Domagala, J. M.; Heifetz, C. L.; Hutt, M. P.; Mich, T. F.; Nichols, J. B.; Solomon, M.; Worth, D. F. *J. Med. Chem.* **1988**, *31*, 991.

8. Ohta, M.; Koga, H. *J. Med. Chem.* **1991**, *34*, 131.

9. Llorente, B.; Rivero, N.; Carrasco, R.; Martinez, R.; Estrada, E. *Quant. Act. Struct. Relat.* **1994**, *13*, 419.

10. Bazile, S.; Moreau, N.; Bouzard, D.; Essiz, M. *Antim. Agents Chemotherapy* **1992**, *36*, 2622.

11. Cirilli, M.; Bachechi, F.; Ughetto, G.; Colonna, F. P.; Capobianco, M. L. *J. Mol. Biol.* **1993**, *230*, 878.

12. Cambau, E.; Gutman, L. *Drugs* **1993**, *45*, 15.

13. Yoshida, H.; Bogaki, M.; Nakamura, M.; Nakamura, S. *Antim. Agents Chemotherapy* **1990**, *34*, 1271.

14. Willmott, C.; Maxwell, A. *Antim. Agents Chemotherapy* **1993**, *37*, 126.

15. Heisig, P.; Shedletzky, H.; Falkenstein-Paul, H. *Antim. Agents Chemotherapy* **1993**, *37*, 696.

16. Domalaga, J.; Badges, A. J.; Culbertson, T. P.; Gambino, L.; Hagen, S. E.; Karrick, G.; Porter, K.; Sanchez, J. P.; Sesnie, J. A.; Spense, F. G.; Szotek, D.; Wemple, J. *J. Med. Chem.* **1991**, *34*, 1142.

17. Sanchez, J. P.; Domalaga, J.; Heifetz, C.; Priebe, S.; Sesnie, J. A.; Trehan, A. *J. Med. Chem.* **1992**, *35*, 1764.

18. Domalaga, J.; Hanna, L.; Heifetz, C.; Hutt, M.; Mich, T.; Sanchez, J. P.; Solomon, M. *J. Med. Chem.* **1986**, *29*, 394.

19. Golender, V. E.; Rozenblit, A. B. In *Logical and Combinatorial Algorithms for Drug Design*; Research Studies: U.K., 1983.

20. APEX-3-D, User Guide, Version 1.4, Biosym Technologies, 1993.

21. Stewart, J. J. P.; MOPAC: A General Molecular Orbital Package, Version 6.0, QCPE #455.

22. Topliss, J. G.; Edwards, R. P. *J. Med. Chem.* **1979**, *22*, 1238.

(*Received in the U.S.A.* 1 *August* 1995; *accepted* 16 *October* 1995)