ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, Sept. 1993, p. 1807–1815
0066-4804/93/091807-09$02.00/0

Vol. 37, No. 9

Anti-*Mycobacterium avium* Activity of Quinolones: Structure-Activity Relationship Studies

GILLES KLOPMAN,[1]* SHAOMENG WANG,[1]† MICHAEL R. JACOBS,[2] AND JERROLD J. ELLNER[2]

Chemistry Department[1] and School of Medicine,[2] Case Western Reserve University, Cleveland, Ohio 44106

Received 5 April 1993/Returned for modification 13 May 1993/Accepted 7 June 1993

The structures and inhibitory activities of 88 quinolones, previously studied as potential in vitro inhibitors of 14 selected strains of *Mycobacterium avium* complex, were examined in an effort to identify a quinolone with optimal activity towards all strains. A MULTICASE structure-activity relationship analysis of the inhibitory activities of these 88 quinolones against 14 strains of *M. avium* was performed and led to the identification of a number of structural constraints required to overcome the resistance of most of the strains. Our data suggested that the increased resistance of the strains was probably not due to a specific resistance mechanism but rather due to gradual limitation of the constraints imposed on the structure of the quinolones. This increasing structural selectivity could be produced either at the level of cell membrane penetration or at the level of interaction with the DNA gyrase receptor site. On the basis of these findings, a number of new quinolones holding the promise of superior activity are currently being evaluated in vitro and in vivo to determine the clinical relevance of our observations.

In the study reported in the accompanying paper, 88 structurally different quinolones were evaluated for their potential in vitro activities against 14 selected strains of *Mycobacterium avium-M. intracellulare* complex (9). These strains reacted differently to the quinolones; some were inhibited by the majority of the drugs, while others were substantially resistant to them. Our analysis permitted us to classify the strains as to their general susceptibilities to inhibition by quinolones. It was found that a continuum existed between the least and most active drugs. In this paper, we report the results of a structure-activity relationship (SAR) study of these results with the MULTIple Computer Automated Structure Evaluation (MULTICASE) program (5) to gain an understanding of the structural requirements needed to be fulfilled in order to design either a drug that would be active against even the most resistant strains or a combination of drugs that would achieve the same results.

The MULTICASE methodology. The MULTICASE method is a hierarchical computer automated structure evaluation program (5). Although it shares many features with its predecessor CASE (Computer Automated Structure Evaluation program) (4), it differs from CASE in a great many ways. Both of these methods automatically identify molecular substructures that have a high probability of being relevant to or responsible for observed biological activity. Both methods require a learning data base which contains the structures and activities of a set of molecules to initiate the analysis.

In both the CASE and MULTICASE approaches, the structures of each of the compounds of the data base are fragmented into all possible linear substructures, which can be as simple as two heavy atoms or as complex as needed. All the fragments generated from all the active and inactive compounds are then submitted to a discriminant analysis in order to assess the significance of each of the fragments and to identify the most relevant "biological functionalities." In both approaches, some properties, such as the partition coefficient between *n*-octanol and water (log P) (8), aqueous solubility, quantum mechanics parameters, and graph indices, are computed for each molecule (5–8).

Activating features (biophores). In the CASE approach, the significant substructures, the log P, the quantum mechanics parameters, and graph indices are used as potential parameters for a regression analysis to generate a linear quantitative SAR (QSAR). Subsequently, the QSAR equation can be used to quantitatively determine the relevance of the parameters to the observed biological activity. It can also be used to predict the activity of new, untested compounds. In the MULTICASE approach, the algorithm performs the analysis in a hierarchical way rather than in a one-step regression analysis. In the first part, the MULTICASE program identifies only true biophores, i.e., those fragments found to have an unquestionable relation to activity. This is done by first selecting the substructure that has the highest probability of being responsible for activity, as judged by the binomial probability that its observed distribution among active and inactive molecules is not due to chance. Those molecules containing this substructure are then eliminated from the data set, and the remaining compounds are submitted to a new analysis. This procedure is then repeated until either (i) the entire set is eliminated (i.e., enough structural features have been found to account for the activity of the entire data set) or (ii) all statistically relevant substructures have been identified and the remaining data cannot be explained by statistically significant descriptors.

The molecules are thus separated into subclasses based on the presence of each of the biophores. For each subclass, a new analysis is performed to produce the modulators capable of modifying the activity of each of the biophores. The system is thus hierarchical in that modulators are important only in the context of molecules containing the primary biophore. These modulators may offset the activity of the

* Corresponding author.

† Present address: Laboratory of Medicinal Chemistry, National Cancer Institute, Bethesda, MD 20892.

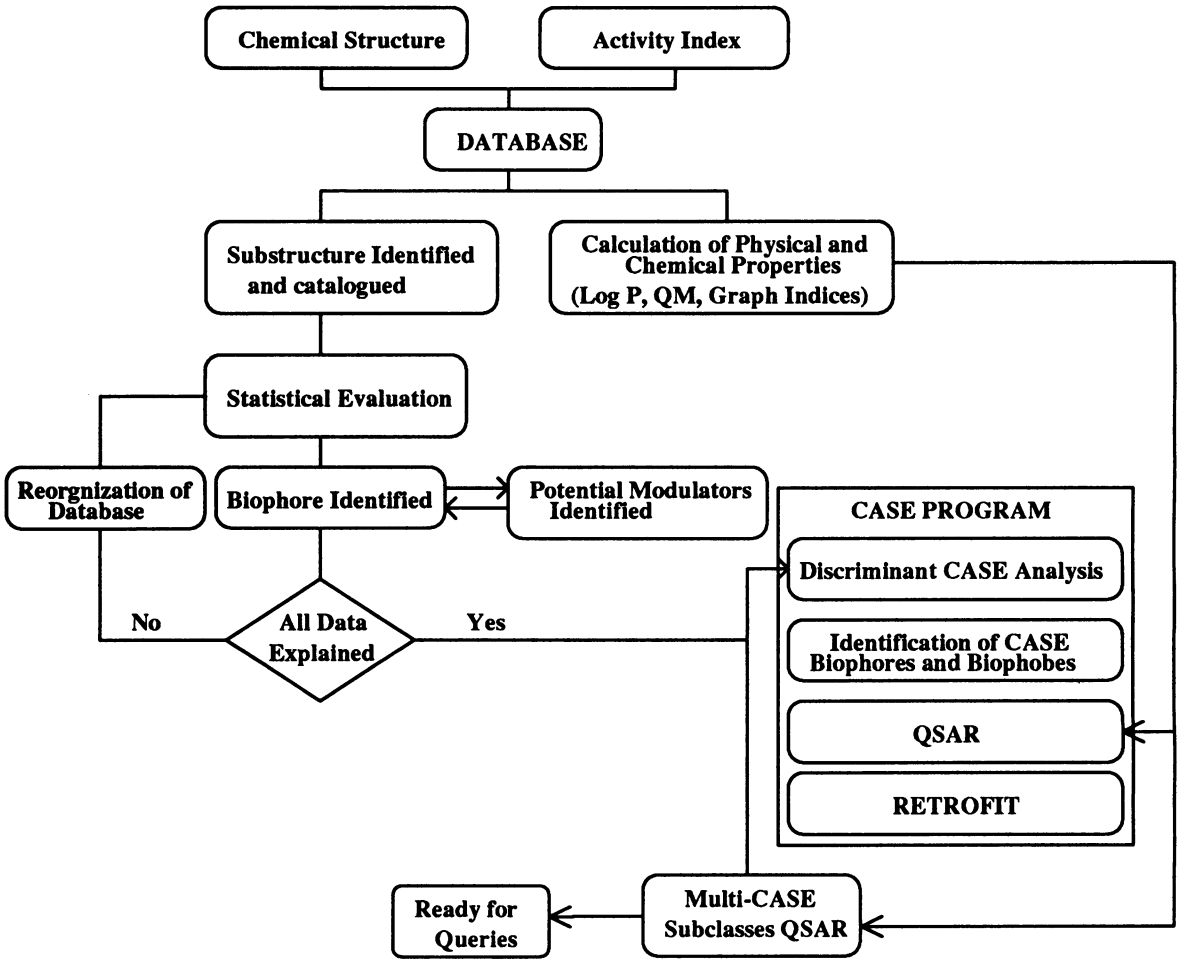


FIG. 1. The MULTICASE program. QM, quantum mechanics.

biophore by decreasing the activity, in which case they will be called biophobes, or by increasing it, in which case they will be called synergistic. In the latter case, it is possible that the modulator itself is an essential part of the biophore. This is particularly true if most of the compounds containing the biophore also contain the modulator. It should be noted that unless the biophore is embedded in the modulator, the relative positions of the biophore and its modulator are not recognized by the program.

Deactivating features (biophobes). We often see in our studies some compounds that contain a biophore yet show little or no activity. The MULTICASE algorithm rationalizes this as follows: (i) an active compound has to have a biophore embedded in its structure and (ii) a compound that does not show any activity either lacks any activating feature (absence of a biophore) or has one or more biophores but also contains a strongly deactivating feature. These deactivating features are called biophobes.

When a new molecule is submitted for testing, the MULTICASE program will search its structure for the existence of any known biophores. If it does not find one, the molecule will be called inactive by default. However, if it does find one, it will then search for the presence of modulators to arrive at a projected value for the biophore's potency.

Overall, it can be said that MULTICASE deals with several sets of congeneric systems. The main difference between these and conventional congeneric data bases is that the commonality among the molecules is based on a rational evaluation of their structures rather than on an arbitrary choice of common structural features. Figure 1 illustrates the operation of the MULTICASE program.

MATERIALS AND METHODS

The data bases were generated from MICs of 88 agents tested against 14 strains with various resistances (9). The conventional approach to construct a data base for SAR or QSAR study is to take the minimum concentration of each compound required to inhibit the growth of 50% of strains as the activity input (6). However, we felt that we would gain more understanding of the biological events if we were to treat each strain as a separate data base and use actual MICs.

In testing an agent multiple times, the reproducibility of the observed MICs can vary up to fourfold. Conventional QSAR approaches would not derive stable models if such widely varying activity values from a single strain were to be used as input. However, both CASE and MULTICASE have the ability to accommodate some inaccuracies in the experimental data since they use a statistical discriminant analysis to distinguish between the substructures that are relevant and those that are irrelevant to the observed activ-

TABLE 1. Distributions of active and inactive agents in the 14 data bases and importance of top biophores in the various data bases

Strain	No. (%) of agents active at 16 μg/ml	% of active agents containing the top biophore
TMC 1403	56 (63.6)	95
PI 2/8	56 (63.6)	95
1695757	48 (54.5)	69
1779564	46 (52.2)	70
PI 44/4	41 (46.6)	45
1958339	38 (43.2)	53
PI 2/6	29 (32.9)	45
TMC 1461	29 (32.9)	37
34540W	27 (30.7)	48
PI 12/39	22 (25.0)	41
1760694	22 (25.0)	53
1988557	21 (23.9)	45
1915112	19 (21.6)	53
1772733	19 (21.6)	53

ity. Therefore, they can accommodate these variations and provide useful information about the key structural features which determine the activity of the quinolones against strains with different susceptibilities.

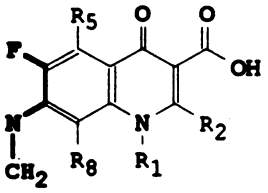
RESULTS AND DISCUSSION

Distribution of active and inactive agents. We evaluated 14 data bases, one for each of the strains tested, against 88 quinolones. It is assumed that a quinolone with a MIC of ≥32.0 μg/ml is unlikely to achieve good activity in macrophages or in vivo. Hence, all quinolones with MICs of ≥32.0 μg/ml were regarded as inactive in our analysis, while those with MICs of ≤16 μg/ml are regarded as active. Table 1 shows the distribution of active and inactive agents for the 14 data bases, using the endpoints defined above.

As can be seen, strains TMC 1403 and PI 2/8 are the two most susceptible strains; more than 63% of the 88 quinolones tested are active against these two strains. Strains 1695757 and 1779564 are very susceptible strains; more than 50% of the 88 compounds are active against these two strains. Strains 1958339 and PI 44/4 have medium susceptibility; about 43% of the 88 compounds are found to be active against these two strains. Strains PI 2/6, TMC 1461, and 34540W are resistant strains; only about 30% of the 88 compounds are active against these strains. The five remaining strains, 1988557, PI 12/39, 1760694, 1772733, and 1915112, are very resistant; only one-fifth to one-fourth of the 88 compounds are active against these strains.

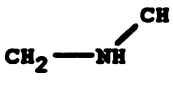
MULTICASE analysis. Biophores were identified by MULTICASE analyses for each of the 14 data bases derived from the 14 strains. The biophores identified as relevant to one of the most susceptible strains, TMC 1403, and those relevant to one of the most resistant strains, 1760694, are shown in Fig. 2 and 3, respectively.

Comparisons between the top biophores. The most significant activating fragments from all 14 data bases identified by the MULTICASE program, referred to as the top biophores, are shown in Fig. 4. As shown in Table 1, they account for the activities of 37 to 95% (average, 57%) of the active compounds in the data bases. An interesting observation is that for the two data bases corresponding to the most susceptible strains, TMC 1403 and PI 2/8, the top biophore (Fig. 4, structure I; referred to hereafter as biophore 4-I) is found in 95% of the active compounds in the data base. This indicates that wide structural variations are tolerated and that the drug receptor is not very selective in these strains. The top biophores for the data bases corresponding to the two very susceptible strains, 1695757 and 1779564, still account for more than two-thirds of the active compounds in the data bases. However, for the next 10 less susceptible strains, the top biophores identified can only explain about half of the active compounds. This probably indicates that the in vitro activity of the quinolones against the most susceptible strains is controlled by a single factor. It is not clear whether this factor is the penetration of the quinolones through the cell wall or the binding affinity to DNA gyrase (1–3, 10–12), but the top biophore (4-I) identified for the two

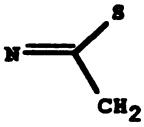


Biophore # 1 (A:I=53:11, prob=100%)

Biophore # 2 (A:I=10:0, prob=99%)

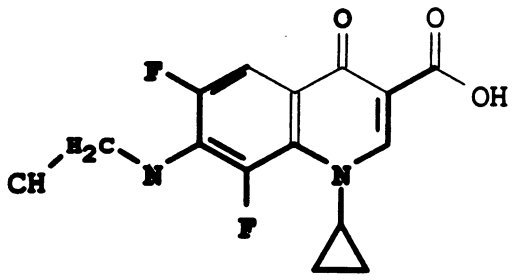


Biophore # 3 (A:I=9:0, prob=98%)

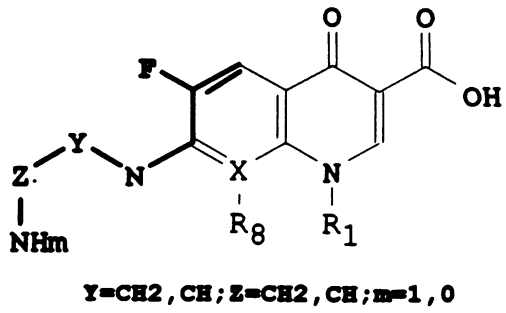


Biophore # 4 (A:I=2:0, prob=60%)

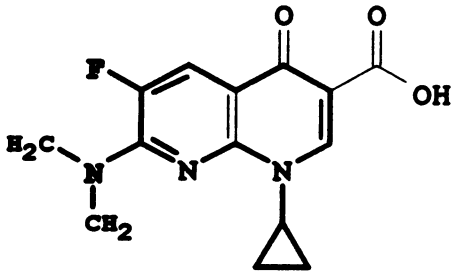
FIG. 2. Biophores identified by the MULTICASE program for the data base corresponding to strains TMC 1403 and PI 2/8. Biophores are shown by boldface lines and boldface characters in the molecules. The other parts of the molecules, unless explicitly shown, are unspecified. A:I, ratio of active to inactive compounds; prob, probability that the fragment is indeed responsible for the observed activity.



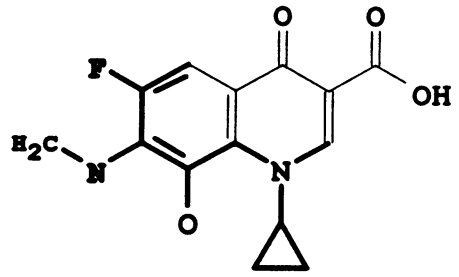
Biophore # 1 (A:I=9:4, prob.=100%)



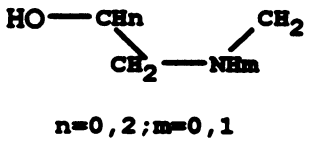
Biophore # 2 and its expanded biophores (A:I=8:0, prob.=100%)



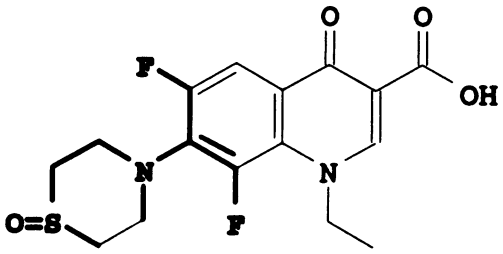
Biophore # 3 (A:I=2:0, prob.=94%)



Biophore # 4 (A:I=2:0, prob.=94%)



Biophore # 5 (A:I=2:0, prob.=94%)



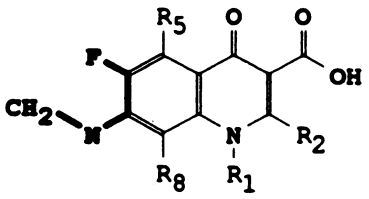
Biophore # 6 (A:I=1:0, prob=75%)

FIG. 3. Biophores identified by the MULTICASE program for the data base corresponding to strain 1760694. Biophores are shown by boldface lines and boldface characters in the molecules. The other parts of the molecules, unless explicitly shown, are unspecified. A:I, ratio of active to inactive compounds; prob, probability that the fragment is indeed responsible for the observed activity.

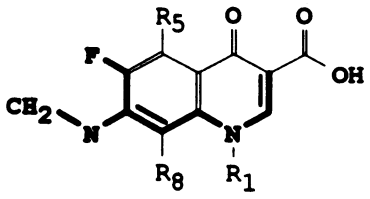
most susceptible strains should be highly correlated with this determining factor.

Interestingly, when we compared the top biophores for all of the strains, from susceptible to resistant, we found that as the susceptibility of the strains decreases, the structure of the top biophore becomes more constrained (Fig. 4). This is seen by the fact that as the strains become resistant, the top biophores become more complex but retain all the structural features needed to achieve activity in the more susceptible strains. Indeed, biophore 4-I is embedded in biophores 4-II to 4-VI. Biophore 4-II is embedded in 4-III to 4-VI, biophore III is embedded in 4-IV to 4-VI, etc. At no time do we observe a sudden change in the nature of the biophore that could be associated with the initiation of a new mechanism triggered by the increased resistance.

The top biophore (4-I) for the most susceptible strains, TMC 1403 and PI 2/8, shows that the most important feature determining the activity of a quinolone against *M. avium-M. intracellulare* complex is the presence of a fluorine at position C-6 and the existence of a CH_2 group attached to a tertiary nitrogen at position C-7. If a quinolone bears these structural features, it will have an 83% chance of being active. The other substituents of the quinolone backbone (R1, R2, R5, and R8) are not as important. However, they may still play some role in increasing or decreasing the activity of a compound. This is in line with previous studies which have shown that the fluoro group at position C-6 seems to improve both the DNA gyrase complex binding (2- to 17-fold) and cell penetration (1- to 70-fold) (1). In all 64 compounds containing the top biophore (4-I), the tertiary nitrogen atom at position 7 appears 63 times in a ring system with more than four members. PD 136576 (no. 14 in Fig. 2 of reference 11), the only compound with a linear C-7 substituent, is inactive. It was observed in previous studies that quinolones with small or linear C-7 substituents generally possess only moderate to weak biological activity (1). This seems to be confirmed by our results. Although it is not clear to us what role the tertiary nitrogen in a ring system at the C-7 position plays, it seems that this is a necessary requirement for a quinolone to achieve some activity against *M.*

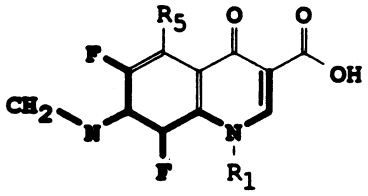


Top Biophore
for strains TMC 1403 and PI 2/8



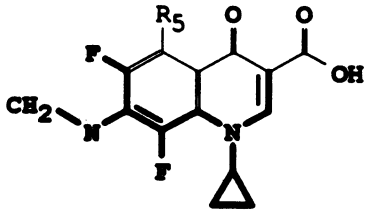
Top Biophore
for strains 1695757 and 1779564

(I)



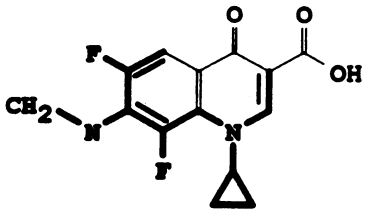
Top Biophore
for strain 1958339

(II)



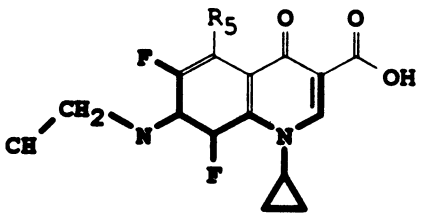
Top Biophore
for strains PI 2/6 and TMC 1461

(III)



Top Biophore
for strain 34540W

(IV)



Top Biophore
for strains 198855, PI 12/39,
1915112,1772733,1760694

(V) (VI)

FIG. 4. Top biophores identified by MULTICASE for the data bases.

avium strains. The top biophore (4-II) for strains 1695757 and 1779564, as expected, imposes the same structural requirements as biophore 4-I. In addition, biophore 4-II specifies that position 8 should bear a substituent. This biophore also suggests that N-1 and C-2 should not be fused into a ring.

The top biophore (4-III) for strain 1958339 clearly shows that a fluorine at position 8 is normally better than any other known substituent. All the compounds bearing this biophore satisfy all the requirements imposed by the top biophores 4-I and 4-II for the four most susceptible strains. The top biophore (4-IV) for the next two less susceptible strains, PI 2/6 and TMC 1461, indicates that a cyclopropyl group is the preferred substituent at the N-1 position. All the compounds containing the top biophore (4-V) for the next less susceptible strain (34540W) satisfy all the requirements imposed by the previous biophores for the more susceptible strains. In addition, biophore 4-V indicates that position 5 should not bear a substituent.

The same fragment was identified as the top biophore (4-VI) for the five most resistant strains (1988557, PI 12/39, 1915112, 1772733, and 1760694). This biophore emphasizes the importance and the nature of the group attached to the tertiary nitrogen at position 7. When compared with the top biophore (4-V) for strain 34540W, it seems that the top biophore (4-VI) for these five most resistant strains does not create any requirement about a substituent at a C-5. However, examination of the R5 group shows that 9 of the 13 compounds containing the biophore have a hydrogen at C-5. The other types of groups attached to position 5 are $N(CH_3)_2$ (in one inactive compound), CH_3 (in two active compounds), and NH_2 (in one very active compound). This indicates that NH_2 and CH_3 groups are acceptable for C-5 substitution. Overall, we find that molecules containing the biophore 4-VI should be effective against all the strains of *M. avium-M. intracellulare* complex, including the most resistant strains.

It is clear from the above observations that as the strains become more resistant, the biophore required for activity

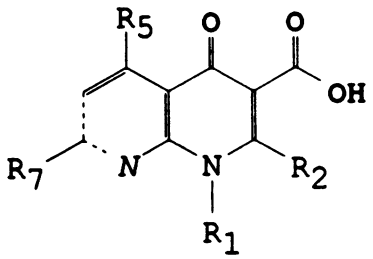


FIG. 5. One deactivating feature found in all 14 data bases. The biophobe is shown by italics and dotted lines. Other parts of the molecule, unless explicitly specified, can be any group. The ratio of active to inactive compounds was 0:7.

becomes more restrictive. This may indicate that the drug receptor site is becoming more selective as the strains become more resistant. However, whether this selectivity is produced by the penetration requirement or by the interaction with the DNA gyrase (1–3, 10–12) is still an open question. Nevertheless, it appears that the increased resistance is not due to the appearance of a new phenomenon but rather to a gradual limitation of the constraints imposed on the structure.

Biophobes. A biophobe (deactivating fragment) was found in all 14 data bases, as shown in Fig. 5. This inactivating feature was found in seven compounds, all of which are inactive against all the strains. Another deactivating feature was found in the data bases of resistant strains, such as strains PI 2/6, TMC 1403, 34540W, 1988557, PI 12/39, 1760694, 1915112, and 1772733, as shown in Fig. 6. This biophobe appears in 16 compounds, all of which are inactive against resistant strains. Figure 7 shows that a quinolone with an sp^2 nitrogen at position 8 and an ethyl group attached to N-1 will be unlikely to show any activity against the less susceptible strains. We also found that one fragment, shown in Fig. 7, gradually becomes significantly deactivating as the susceptibility of the strains decreases. This fragment shows that the nature of the R1 group becomes increasingly important in preventing the activity of a quinolone against more resistant strains. The presence of this fragment suggests that a quinolone with an ethyl group at position 1 would be unlikely to achieve substantial activity against resistant strains. PD 115311 (no. 5 in Fig. 2 of reference 9) is the only compound in the data base which has an ethyl as its R1 group but still shows some activity against the resistant strains.

QSARs. MULTICASE allows QSARs as well as qualitative SARs to be found. The QSARs resulting from this

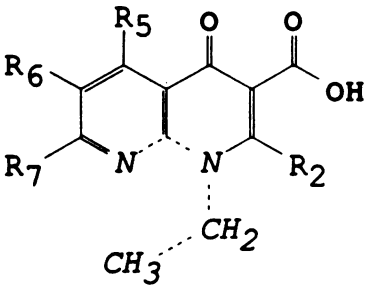


FIG. 6. A deactivating feature found in eight data bases. The biophobe is shown by italics and dotted lines. Other parts of the molecule, unless explicitly specified, can be any group. The ratio of active to inactive compounds was 0:16.

analysis can then be used to provide quantitative activity predictions for untested compounds. However, before any predictions can be made, the validity of the derived models must be established in order to build confidence that the models have predictive value (5). Indeed, evaluation of the prediction potential of a SAR or a QSAR model is one of the most important aspects of SAR studies. The quality of a QSAR model cannot be established by the fact that the data base can be satisfactorily retrofitted.

One way to evaluate the prediction potential of a QSAR model is to randomly separate the data into two sets, one learning set and one test set. Normally the learning set contains many more molecules than the test set. The learning set then is used to establish a QSAR model, and the activities of the molecules of the test set are calculated. The predictive value of the model can then be judged by evaluating how well the model predicted the activities of the molecules of the test set.

In this study, we proceeded to evaluate the predictive power of the QSAR models as we were testing the agents. We originally selected one-third of the available compounds (the first 33 compounds of Fig. 2 of reference 9). When submitted to MULTICASE analysis, 14 QSAR models corresponding to the 14 strains were obtained (model I).

The results of comparisons between the recalculated and the observed MICs for the 33 compounds in the learning data bases are shown in Table 2. The results in Table 2 essentially indicate how well the data had been retrofitted by the QSAR equations. As can be seen, between 85 and 100% of the compounds, the average being 97% for all the strains, have been correctly recalculated. Notice that the concordances for three strains, 1779564, 1958339, and PI 12/39, are significantly lower than those for other strains.

We then used these relationships (model I) to predict the activities of the remaining untested 55 compounds. A comparison between the predicted and the observed MICs for the 55 compounds is shown in Table 3.

As can be seen, the concordance between the predicted and the experimental results for the new compounds is lower than that observed (Table 2) for the compounds in the learning data base. This is to be expected since the compounds in the learning data base had been used by the program to establish the QSAR parameters. Of the 55 compounds of the test set, the activities of 58 to 87% of the agents (average, 74%) were correctly predicted by the MULTICASE program. Considering the small size of the learning data base (only 33 compounds), the results are considered satisfactory and show the learning ability of the MULTICASE program. Examination of Table 2 also shows that the concordance for active compounds decreases as the susceptibilities of the strains decrease. In contrast, the concordance for the inactive compounds increases as the susceptibilities of the strains decrease. This is to be expected since the number of active compounds in the learning data base becomes fewer as the susceptibilities of the strains decrease. With only very few active compounds in the learning data bases for strains 1915112 and 1772733, it is impossible for the program to gain enough knowledge to make good predictions for active compounds. The same argument can be applied to the most susceptible strains, TMC 1403 and PI 2/8 S4, for which there were insufficient inactive compounds in the learning data bases. It seems that 20 active and 20 inactive compounds are needed in a learning data base for the program to be able to give reliable predictions for both active and inactive molecules. The reason for the poor predictions for strains 1779564 and 1958339 is not clear.

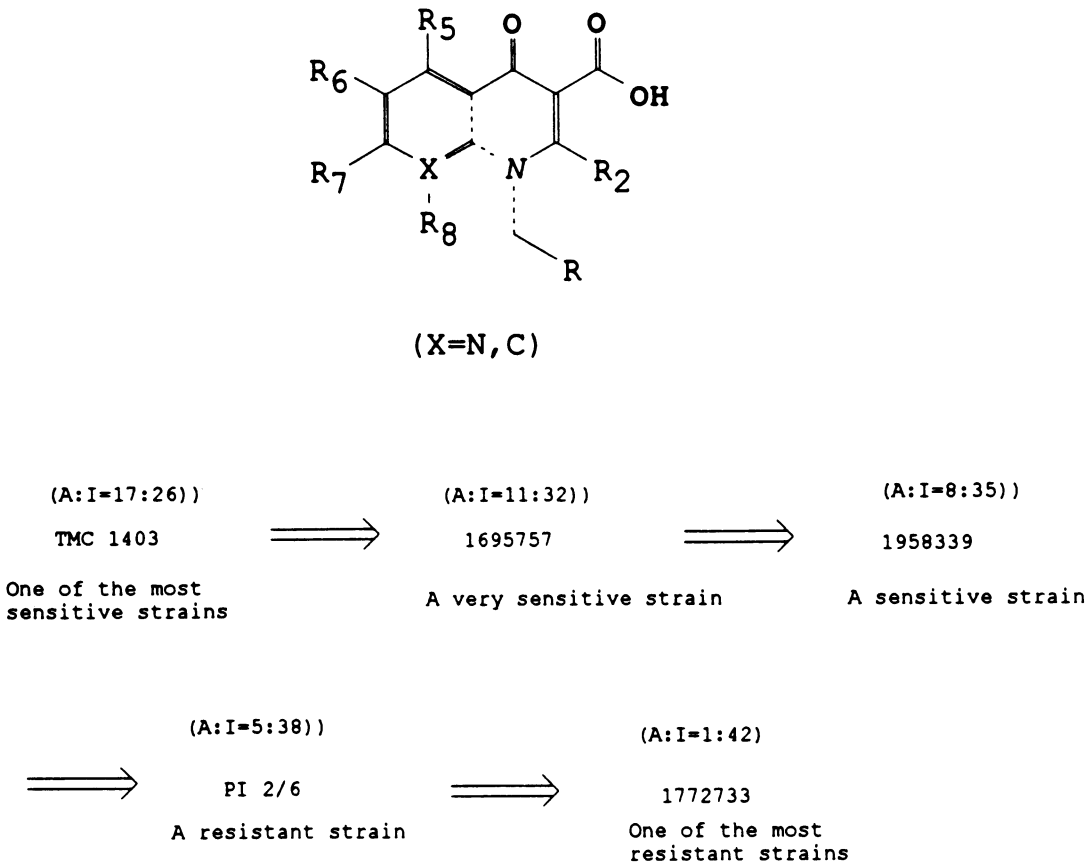


FIG. 7. A fragment that becomes a very significant deactivating feature as the susceptibility of the strain decreases. The fragment is shown by italics and dotted lines. Other parts of the molecule, unless explicitly specified, can be any group. A:I, ratio of active to inactive compounds.

Another one-third of the 88 compounds were then selected (a total of 29 compounds, from PD 107522 to PD 143289 of Fig. 2 of reference 9) for analysis (model II). These compounds were selected on the basis that MULTICASE issued a warning signal for them, meaning that they contain substructures that did not exist in the original learning set. A new MULTICASE analysis that now included the 62 compounds (the original 33 plus the new 29) in the learning data bases was then performed. Table 4 shows the summary of comparisons between the calculated and observed MICs for these 62 compounds. As can be seen, 99% concordance was found for the retrofit. Model II was then used to predict the activities of the remaining untested 26 molecules. Table 5 shows the results of a comparison between the predicted and subsequently measured MICs for these 26 compounds.

On average, a 76% concordance for the 14 strains was found (range, 69 to 85%). It seems that the prediction results of model II are only slightly better than those of model I (73%). However, if one compares the prediction results of models II and I for the last 26 compounds, one finds that model I can, on average, correctly predict the activities of only 69% of the 26 compounds. In contrast, model II can correctly predict the activities of 76% of the 26 compounds. As can be seen, some improvement of the predictive power was achieved by expanding the learning data bases, which indicates that the program has learned to some extent from the additional 29 compounds with new substructures existing in them. After the in vitro MICs of the last 26 compounds

TABLE 2. Comparisons between recalculated and observed MICs of the 33 agents of the learning data bases (model I)

Strain	No. of agents				Overall CP[e] (%)
	Obs(+)/ calc(+)[a]	Obs(+)/ calc(−)[b]	Obs(−)/ calc(+)[c]	Obs(−)/ calc(−)[d]	
TMC 1403	22/22	22/0	11/1	11/10	97
PI 2/8	22/22	22/0	11/1	11/10	97
1695757	17/17	17/0	16/0	16/16	100
1779564	17/14	17/3	16/2	16/14	85
PI 44/4	8/8	8/0	25/0	25/25	100
1958339	12/12	12/12	21/4	21/17	88
PI 2/6	12/11	12/0	21/0	21/21	97
TMC 1461	11/11	11/0	22/0	22/22	100
34540W	10/10	10/0	23/0	23/23	100
PI 12/39	10/7	10/3	23/1	23/22	88
1760694	8/8	8/0	25/0	25/25	100
1988557	9/9	9/0	24/0	24/24	100
1915112	7/7	7/0	26/0	26/26	100
1772733	6/6	6/0	27/0	27/27	100

[a] The number of agents which were both observed and predicted to be active.

[b] The number of agents which were observed to be active but were predicted to be inactive.

[c] The number of agents which were observed to be inactive but were predicted to be active.

[d] The number of agents which were both observed and predicted to be inactive.

[e] The overall percentage correctly predicted for each strain.

TABLE 3. Comparisons between the predicted and observed MICs of the 55 agents not used for Table 2

Strain	No. of agents				Overall CP[e] (%)
	Obs(+)/ pred(+)[a]	Obs(+)/ pred(−)[b]	Obs(−)/ pred(+)[c]	Obs(−)/ pred(−)[d]	
TMC 1403	34/30	34/4	21/12	21/9	71
PI 2/8	34/30	34/4	21/14	21/7	67
1695757	31/26	31/5	24/12	24/12	69
1779564	28/23	28/5	27/16	27/11	62
PI 44/4	29/15	29/14	26/6	26/24	71
1958339	22/15	22/7	33/16	33/17	58
PI 2/6	16/13	16/3	39/12	39/27	73
TMC 1461	18/16	18/2	37/12	37/25	75
34540W	17/14	17/3	38/9	38/29	78
PI 12/39	12/8	12/4	43/3	43/40	87
1760694	15/12	15/3	40/9	40/31	78
1988557	13/8	13/5	42/4	42/38	84
1915112	12/8	12/4	43/7	43/36	80
1772733	13/8	13/5	42/7	42/35	78

[a] See footnote *a* of Table 2.
[b] See footnote *b* of Table 2.
[c] See footnote *c* of Table 2.
[d] See footnote *d* of Table 2.
[e] See footnote *e* of Table 2.

TABLE 5. Comparisons between predicted and observed MICs of the 26 agents not used for Table 4 (model II)

Strain	No. of agents				Overall CP[e] (%)
	Obs(+)/ pred(+)[a]	Obs(+)/ pred(−)[b]	Obs(−)/ pred(+)[c]	Obs(−)/ pred(−)[d]	
TMC 1403	19/18	19/1	7/4	7/3	81
PI 2/8	19/18	19/1	7/4	7/3	81
1695757	17/16	17/1	9/5	9/4	77
1779564	15/14	15/1	11/6	11/5	73
PI 44/4	16/12	16/4	10/3	10/7	73
1958339	14/10	14/4	12/3	12/9	73
PI 2/6	10/7	10/3	16/5	16/11	69
TMC 1461	11/9	11/2	15/3	15/12	81
34540W	9/7	9/2	17/6	17/11	69
PI 12/39	7/6	7/1	19/5	19/14	77
1760694	9/9	9/0	17/6	17/11	77
1988557	8/6	8/2	18/5	18/13	73
1915112	7/5	7/2	19/4	19/15	77
1772733	8/6	8/2	18/2	18/16	85

[a] See footnote *a* of Table 2.
[b] See footnote *b* of Table 2.
[c] See footnote *c* of Table 2.
[d] See footnote *d* of Table 2.
[e] See footnote *e* of Table 2.

(from PD 111752 to PD 144881 of Fig. 2 of reference 9) were added, the final model (model III) was established by including all 88 compounds in the learning data bases. Although we do not have any additional compounds available to evaluate the predictive power of the final model, we expect that the average concordance between the predicted and observed MICs for new compounds should be greater than the 76% average agreement found by model II.

These results show that the MULTICASE program not only was able to explain the data in the learning set but also has the potential for predicting the activities of new, untested compounds. This gives us confidence that the results of our analysis provide a solid base to evaluate new structures and help select promising compounds for synthesis.

TABLE 4. Comparisons between observed and calculated MICs of 62 agents (model II)

Strain	No. of agents				Overall CP[e] (%)
	Obs(+)/ pred(+)[a]	Obs(+)/ pred(−)[b]	Obs(−)/ pred(+)[c]	Obs(−)/ pred(−)[d]	
TMC 1403	37/37	37/0	25/2	25/23	97
PI 2/8	37/36	37/1	25/0	25/25	98
1695757	31/31	31/0	31/1	31/1	98
1779564	30/30	30/0	32/0	32/32	100
PI 44/4	21/21	21/0	41/0	41/41	100
1958339	24/24	24/0	38/1	38/37	98
PI 2/6	19/19	19/0	43/0	43/43	100
TMC 1461	18/18	18/0	44/0	44/44	100
34540W	18/18	18/0	44/1	44/43	98
PI 12/39	14/14	14/0	48/0	48/48	100
1760694	13/13	13/0	49/0	49/49	100
1988557	14/14	14/0	48/0	48/48	100
1915112	12/12	12/0	50/0	50/50	100
1772733	11/11	11/0	51/0	51/51	100

[a] See footnote *a* of Table 2.
[b] See footnote *b* of Table 2.
[c] See footnote *c* of Table 2.
[d] See footnote *d* of Table 2.
[e] See footnote *e* of Table 2.

Indeed, if a newly designed compound is predicted to have good MICs against most of the strains, it could be selected for synthesis and experimental evaluations.

Conclusions. Our in vitro anti-*M. avium* activity study of 88 quinolones has enabled us to identify a number of quinolones with in vitro activities better than or comparable to that of ciprofloxacin, the most effective known agent against *M. avium-M. intracellulare* complex. At this stage it is not clear whether these structures can achieve high activities in macrophages and in vivo. A SAR study using the MULTICASE program allowed us to identify a number of key structural features relevant to the activity of quinolones against *M. avium-M. intracellulare* complex. Moreover, we found that as the susceptibility of these strains decreases, the key structural features determining the activity become more restrictive. QSAR studies showed that the derived QSAR models can correctly explain 97% of the data of the learning data bases. Incremental validation studies indicate that the resulting QSAR models, on average, may correctly predict the activity of up to 80% of new, untested compounds.

On the basis of these results, the structures of other potent quinolones can be proposed. We are in the process of evaluating the in vitro and in vivo activities of a number of such agents to determine the relevance of our observations to clinical application.

ACKNOWLEDGMENTS

This work was supported in part by the National Institute of Allergy and Infectious Diseases as a National Cooperative Drug Discovery Group for Opportunistic Infections grant (AI 30189 SRC).

We thank the Parke-Davis Company for providing samples of the quinolones used in this study.

REFERENCES

1. **Domagala, J. M., L. D. Hanna, C. L. Heifetz, M. P. Hutt, T. F. Mich, J. P. Sanchez, and M. Solomon.** 1986. New structure-activity relationships of the quinolone antibacterials using target enzyme. The development and application of a DNA gyrase assay. J. Med. Chem. **29**:394–404.
2. **Gellert, M., K. Mizuuchi, M. H. O'Dea, T. Itoh, and J. Tomi-**

zawa. 1977. Nalidixic acid resistance: a second genetic character involved in DNA gyrase activity. Proc. Natl. Acad. Sci. USA **74:**4772–4776.

3. **Hogberg, T., I. Khanna, S. D. Drake, L. A. Mitscher, and L. L. Shen.** 1984. Structure-activity relationships among DNA gyrase inhibitors. Synthesis and biological evaluation of 1,2-dihydro-4,4-dimethyl-1-oxo-2-naphthalenecarboxylic acid as 1-carbabioisoesters of oxolonic acid. J. Med. Chem. **27:**306–310.
4. **Klopman, G.** 1984. Artificial intelligence approach to structure-activity studies. Computer automated structure evaluation of biological activity of organic molecules. J. Am. Chem. Soc. **106:**7315–7320.
5. **Klopman, G.** 1992. MULTICASE: a hierarchical computer automated structure evaluation program. QSAR **11:**176–184.
6. **Klopman, G., O. T. Macina, M. E. Levinson, and H. S. Rosenkranz.** 1987. Computer automated structure evaluation of quinolone antibacterial agents. Antimicrob. Agents Chemother. **31:**1831–1840.
7. **Klopman, G., C. Raychaudhury, and R. V. Henderson.** 1988. A new approach to structure-activity using distance information content of graph vertices. A study with phenylalkylamines. Math. Comput. Modelling **11:**635–640.
8. **Klopman, G., and S. Wang.** 1991. A computer automated structure evaluation (CASE) approach to the calculation of partition coefficients. J. Comput. Chem. **12:**1025–1032.
9. **Klopman, G., S. Wang, M. R. Jacobs, S. Bajaksouzian, K. Edmonds, and J. J. Ellner.** 1993. Anti-*Mycobacterium avium* activity of quinolones: in vitro activities. Antimicrob. Agents Chemother. **37:**1799–1806.
10. **Koga, H., A. Itoh, S. Murayama, S. Suzue, and T. Irikura.** 1980. Structure activity relationships of antibacterial 6,7- and 7,8-disubstituted 1-alkyl-1,4-dihydro-4-oxoquinoline-3-carboxylic acids. J. Med. Chem. **23:**1358–1363.
11. **Morita, J., K. Watabe, and T. Komano.** 1984. Mechanism of action of new synthetic nalidixic acid-related antibiotics: inhibition of DNA gyrase supercoiling catalyzed by DNA gyrase. Agric. Biol. Chem. **38:**663–668.
12. **Sugino, A., C. L. Peebles, K. N. Kreuzer, and N. R. Cozzarelli.** 1977. Mechanism of action of nalidixic acid: purification of *Escherichia coli* nalA gene product and its relationship to DNA gyrase and a novel nicking-closing enzyme. Proc. Natl. Acad. Sci. USA **22:**1238–1244.