ELSEVIER

www.elsevier.nl/locate/farmac

Il Farmaco 55 (2000) 484–491

IL FARMACO

Building a pharmacophore model for a novel class of antitubercular compounds☆

F. Manetti [a], F. Corelli [a,*1], M. Biava [b], R. Fioravanti [b], G.C. Porretta [b], M. Botta [a,*2]

[a] *Dipartimento Farmaco Chimico Tecnologico, Università degli Studi di Siena, Via Aldo Moro snc, 53100 Siena, Italy*
[b] *Dipartimento di Studi di Chimica e Tecnologia delle Sostanze Biologicamente Attive, Università degli Studi di Roma 'La Sapienza', Piazzale Aldo Moro 5, 00185 Rome, Italy*

Received 8 March 2000; accepted 12 June 2000

Abstract

Starting from a set of 32 antitubercular compounds, for the first time a three-dimensional pharmacophore model has been derived through a computational approach based on CATALYST software. The model proved to be able to identify compounds belonging to classes of molecules already reported as antitubercular agents. © 2000 Elsevier Science S.A. All rights reserved.

Keywords: Pharmacophore model; Antitubercular agents; Database search

1. Introduction

Tuberculosis is an important human pathology that has recently re-emerged despite the existence of an effective chemotherapy. In addition, AIDS co-infection coupled with some societal features (homelessness, poor nutrition, and so on) are changing the tubercular organism into a more formidable killer. As a consequence, the World Health Organization has reported that more people died of tuberculosis in 1995 than in any other year in history [1].

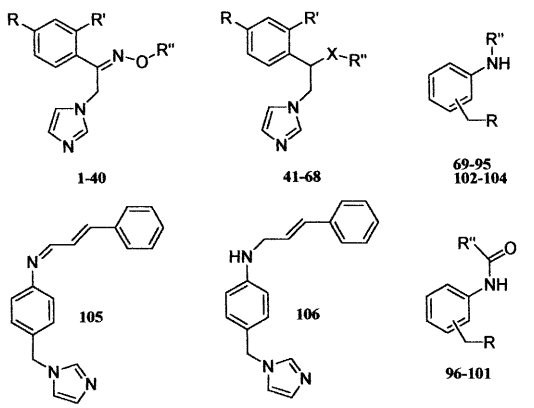

Chart 1. General structures of the antitubercular compounds **1**–**106** used in this study.

Although many structurally diverse compounds showing antitubercular activity are known and an extensive knowledge of the detailed mechanisms of action of current antimycobacterial drugs, such as isoniazid, ethionamide, ethambutol, rifampin, is now available [2], no pharmacophore model has yet been suggested for antitubercular compounds. During the last few years we have been synthesizing a number of imidazole derivatives **1**–**106** [3–8], which have shown interesting activity against *Mycobacterium tuberculosis 103471* and *Mycobacterium avium 103317* (Chart 1, Table 1). On the basis of these findings, we decided to develop a pharmacophore model, which could account for the antimycobacterial activity of this novel class of antitubercular agents, on the assumption that all the structurally related compounds **1**–**106** may act through the same biological mechanism. Accordingly, we have performed a two step computational study in order to (i) generate a three-dimensional pharmacophore model for antitubercular compounds; and (ii) use this model to perform a search on chemical databases with the aim of identifying interesting structural templates. Thus, leading by synthetic optimization, to compounds that possibly could act against mycobacteria [9].

☆ Part of this work was presented as an oral communication at the Italian–Hungarian–Polish Joint Meeting on Medicinal Chemistry, Giardini Naxos-Taormina, Italy, September 28–October 1, 1999.

[1] *Corresponding author.
[2] *Corresponding author.

PII: S0014-827X(00)00072-0

2. Results and discussion

The pharmacophore generation, which represents the first step of our computational work, uses the program CATALYST [14] which provides two different methods. The first method is usually applied to develop three-dimensional pharmacophore models, starting from a collection of molecules possessing a range of diversity both in structure and activity with the last spanning at least five orders of magnitude. The second method, called HipHop or common feature hypothesis generation [15], is able to generate pharmacophore models only by identification of the common chemical features shared by the molecules and their relative alignment to the common feature set, without considering biological data.

The structural analogy between the studied compounds, combined with their small activity range of only two orders of magnitude, led to the conclusion that the common feature hypothesis generation was the most suitable method to be applied to this class of compounds. In fact, although the biological activity of these compounds represent a good result from the medicinal chemistry point of view, the activity range seems to be too narrow to generate a statistically significant quantitative model able to correlate the structural features of these compounds with their biological data. For this reason, the first pharmacophore generation method was considered inappropriate.

Since each computational run is limited to no more than 32 compounds, a selection within the original set of molecules was performed. In order to enhance the probability of finding a significant model, it is important to maximize the information content of the newly generated set. For this purpose, the following two rules were applied: (i) inclusion of the most active compounds in the set, as CATALYST pays particular attention to these molecules when it generates the chemical feature space (see Section 4); (ii) maximization of the kinds and relative positions of the chemical features shared by the molecules because the program recognizes the molecules as collections of chemical features, not as assemblies of atoms or bonds. Accordingly, compounds **1**, **8**, **10**, **15**, **16**, **19**, **23**, **28**, **33**, **36**, **38**, **40**, **44**, **46**, **49**, **56**, **58**, **63**, **67**, **68**, **70**, **74**, **76**, **78**, **81**, **86**, **88**, **95**, **97**, **103**, **105**, and **106** were selected as 'the training set'.

Experimental data on the biologically relevant conformations of the 32 selected compounds (for example, atomic coordinates derived from X-ray crystallographic studies of their complexes with the putative receptor) was not available, so a molecular mechanics approach was applied to build the conformational models to be used for pharmacophore generation. The 'best conformer' analysis implemented in CATALYST has been applied to collect a representative set of conformers chosen within a range of energetically reasonable conformations for each compound. In particular, all conformers within a range of 20 kcal mol^{-1} with respect to the global minimum, have been employed to build a set of pharmacophore hypotheses. On the basis of the summary of the computational run (Table 2) provided by the program, a selection was made among the ten best pharmacophores found in order to obtain a new and more complete pharmacophore hypothesis. Thus, two hypotheses (namely the first and the third, showing different sets of functions with two features in common) were found for which the same conformer of one of the most active compounds **63** maps as the best fit. The superimposition of the two hypotheses led to the identification of three pairs of chemical features located at the same spatial positions. In particular, both the hypotheses show a hydrophobe accommodating the *para* methyl group on the phenyl moiety of compound **63** and an aromatic ring feature able to fit one portion of the naphthyl ring of the same ligand. The third pair of features located almost exactly at the same coordinates is an aromatic ring from hypothesis 1 and a hydrophobe from hypothesis 3. We decided to choose the aromatic ring feature and discard the hydrophobic one, because many compounds of the original set show, in this position, an aromatic group that could be a key element for antimycobacterial activity.

Finally, the two complementary four-feature hypotheses were merged into a new single five-feature hypothesis (the final pharmacophore model), shown in Fig. 1 with compound **63**, taken as a representative example, superimposed on it. The calculated energy of the conformer shown is 5.33 kcal mol^{-1} above the lowest energy conformation. The chemical functionalities of the hypothesis are all matched by the chemical groups of the molecule: the imidazole ring fits the region of the hydrogen bond acceptor feature (HBA); the phenyl portion of the molecule and the methyl substituent of the carbon atom at the *para* position are located inside an aromatic ring (RA1) and a hydrophobic feature (HY1), respectively; the naphthyl moiety maps to both the second hydrophobic (HY2) and the second aromatic ring (RA2) features of the hypothesis.

Inspection of isomeric pairs of compounds indicated that there were two possible modes in which the isomers could be aligned to the hypothesis. Fig. 2 shows the mapping of two isomers (**40** and **36**, *Z*- and *E*-isomer, respectively, taken as representative examples of the whole set of geometric isomers) on the five-feature hypothesis. Both isomers have the same alignment of their phenyl portions on the hypothesis. In fact, the benzyl group bound to the oxygen atom of the oxime moiety (right portion of Fig. 2) fulfills the RA2 aromatic ring feature and the second phenyl group of the molecules (on the left in Fig. 2) is mapped to the RA1 aromatic ring feature of the hypothesis. An additional

Table 1
Chemical and biological properties of compounds **1–106**

Compd.	X	Isomer	R	R′	R″	MIC (μmol ml^{-1}) [a]
1 [b]		*Z*	H	H	H	7.95×10^{-2}
2		*Z*	Cl	H	H	3.39×10^{-2}
3		*Z*	F	H	H	1.46×10^{-1}
4		*Z*	Cl	Cl	H	1.18×10^{-1}
5		*Z*	H	H	*E*-cinnamyl	2.52×10^{-2}
6		*Z*	Cl	H	*E*-cinnamyl	2.27×10^{-2}
7		*Z*	F	H	*E*-cinnamyl	2.38×10^{-2}
8 [b]		*Z*	Cl	Cl	*E*-cinnamyl	2.07×10^{-2}
9		*Z*	H	H	*p*-biphenylmethyl	2.17×10^{-2}
10 [b]		*Z*	Cl	H	*p*-biphenylmethyl	1.99×10^{-2}
11		*Z*	F	H	*p*-biphenylmethyl	2.07×10^{-2}
12		*Z*	Cl	Cl	*p*-biphenylmethyl	1.83×10^{-2}
13		*Z*	H	H	1-naphthylmethyl	2.34×10^{-2}
14		*Z*	Cl	H	1-naphthylmethyl	2.13×10^{-2}
15 [b]		*Z*	F	H	1-naphthylmethyl	2.22×10^{-2}
16 [b]		*Z*	Cl	Cl	1-naphthylmethyl	1.95×10^{-2}
17		*E*	H	H	H	3.97×10^{-2}
18		*E*	Cl	H	H	3.39×10^{-2}
19 [b]		*E*	F	H	H	7.29×10^{-2}
20		*E*	Cl	Cl	H	1.18×10^{-1}
21		*E*	H	H	*E*-cinnamyl	2.52×10^{-2}
22		*E*	Cl	H	*E*-cinnamyl	2.27×10^{-2}
23 [b]		*E*	F	H	*E*-cinnamyl	2.38×10^{-2}
24		*E*	Cl	Cl	*E*-cinnamyl	2.07×10^{-2}
25		*E*	H	H	*p*-biphenylmethyl	2.17×10^{-2}
26		*E*	Cl	H	*p*-biphenylmethyl	1.99×10^{-2}
27		*E*	F	H	*p*-biphenylmethyl	2.07×10^{-2}
28 [b]		*E*	Cl	Cl	*p*-biphenylmethyl	1.83×10^{-2}
29		*E*	H	H	1-naphthylmethyl	2.34×10^{-2}
30		*E*	Cl	H	1-naphthylmethyl	2.13×10^{-2}
31		*E*	F	H	1-naphthylmethyl	2.22×10^{-2}
32		*E*	Cl	Cl	1-naphthylmethyl	1.95×10^{-2}
33 [b]		*E*	H	H	benzyl	1.72×10^{-3}
34		*E*	Cl	H	benzyl	1.53×10^{-3}
35		*E*	F	H	benzyl	1.29×10^{-2}
36 [b]		*E*	Cl	Cl	benzyl	2.22×10^{-2}
37		*Z*	H	H	benzyl	3.43×10^{-3}
38 [b]		*Z*	Cl	H	benzyl	1.53×10^{-3}
39		*Z*	F	H	benzyl	2.59×10^{-2}
40 [b]		*Z*	Cl	Cl	benzyl	1.39×10^{-3}
41	NH		H	H	*E*-cinnamyl	5.27×10^{-2}
42	NH		Cl	H	*E*-cinnamyl	9.47×10^{-2}
43	NH		F	H	*E*-cinnamyl	2.49×10^{-2}
44 [b]	NH		Cl	Cl	*E*-cinnamyl	2.15×10^{-2}
45	NH		H	H	*p*-biphenylmethyl	1.13×10^{-2}
46 [b]	NH		Cl	H	*p*-biphenylmethyl	5.16×10^{-3}
47	NH		F	H	*p*-biphenylmethyl	1.07×10^{-2}
48	NH		Cl	Cl	*p*-biphenylmethyl	1.89×10^{-2}
49 [b]	NH		H	H	1-naphthylmethyl	2.44×10^{-2}
50	NH		Cl	H	1-naphthylmethyl	2.21×10^{-2}
51	NH		F	H	1-naphthylmethyl	9.26×10^{-2}
52	NH		Cl	Cl	1-naphthylmethyl	8.08×10^{-2}
53	O		H	H	*E*-cinnamyl	1.32×10^{-2}
54	O		Cl	H	*E*-cinnamyl	2.37×10^{-2}
55	O		CH_3	H	*E*-cinnamyl	1.26×10^{-2}
56 [b]	O		Cl	Cl	*E*-cinnamyl	2.15×10^{-2}
57	O		H	H	*p*-biphenylmethyl	2.26×10^{-2}
58 [b]	O		Cl	H	*p*-biphenylmethyl	5.15×10^{-3}
59	O		CH_3	H	*p*-biphenylmethyl	5.43×10^{-3}
60	O		Cl	Cl	*p*-biphenylmethyl	9.47×10^{-3}
61	O		H	H	1-naphthylmethyl	3.05×10^{-3}
62	O		Cl	H	1-naphthylmethyl	1.10×10^{-2}

(*Continued on next page*)

Table 1 (*Continued*)

Compd.	X	Isomer	R	R′	R″	MIC (μmol ml^{-1}) [a]
63 [b]	O		CH_3	H	1-naphthylmethyl	1.46×10^{-3}
64	O		Cl	Cl	1-naphthylmethyl	2.52×10^{-3}
65	O		H	H	*p-tert*-buthylbenzyl	2.99×10^{-3}
66	O		Cl	H	*p-tert*-buthylbenzyl	2.71×10^{-3}
67 [b]	O		CH_3	H	*p-tert*-buthylbenzyl	2.87×10^{-3}
68 [b]	O		Cl	Cl	*p-tert*-buthylbenzyl	2.48×10^{-3}
69		*ortho*	1-imidazolyl		*p*-Cl-phenylmethyl	2.69×10^{-2}
70 [b]		*meta*	1-imidazolyl		*p*-Cl-phenylmethyl	1.34×10^{-2}
71		*para*	1-imidazolyl		*p*-Cl-phenylmethyl	1.34×10^{-2}
72		*ortho*	1-imidazolyl		*p*-biphenylmethyl	2.36×10^{-2}
73		*meta*	1-imidazolyl		*p*-biphenylmethyl	1.18×10^{-2}
74 [b]		*para*	1-imidazolyl		*p*-biphenylmethyl	2.36×10^{-2}
75		*ortho*	1-*N*-methylpiperazinyl		*p*-Cl-phenylmethyl	9.78×10^{-2}
76 [b]		*meta*	1-*N*-methylpiperazinyl		*p*-Cl-phenylmethyl	9.78×10^{-2}
77		*para*	1-*N*-methylpiperazinyl		*p*-Cl-phenylmethyl	9.78×10^{-2}
78 [b]		*ortho*	1-*N*-methylpiperazinyl		*p*-biphenylmethyl	1.08×10^{-2}
79		*meta*	1-*N*-methylpiperazinyl		*p*-biphenylmethyl	2.16×10^{-2}
80		*para*	1-*N*-methylpiperazinyl		*p*-biphenylmethyl	1.08×10^{-2}
81 [b]		*ortho*	1-morpholinyl		*p*-Cl-phenylmethyl	2.50×10^{-2}
82		*meta*	1-morpholinyl		*p*-Cl-phenylmethyl	5.01×10^{-2}
83		*para*	1-morpholinyl		*p*-Cl-phenylmethyl	1.01×10^{-1}
84		*ortho*	1-morpholinyl		*p*-biphenylmethyl	8.94×10^{-2}
85		*meta*	1-morpholinyl		*p*-biphenylmethyl	2.23×10^{-2}
86 [b]		*para*	1-morpholinyl		*p*-biphenylmethyl	1.12×10^{-2}
87		*ortho*	1-thiomorpholinyl		*p*-Cl-phenylmethyl	1.03×10^{-1}
88 [b]		*meta*	1-thiomorpholinyl		*p*-Cl-phenylmethyl	2.57×10^{-2}
89		*para*	1-thiomorpholinyl		*p*-Cl-phenylmethyl	1.03×10^{-1}
90		*ortho*	1-thiomorpholinyl		*p*-biphenylmethyl	8.55×10^{-2}
91		*meta*	1-thiomorpholinyl		*p*-biphenylmethyl	8.55×10^{-2}
92		*para*	1-thiomorpholinyl		*p*-biphenylmethyl	8.55×10^{-2}
93		*ortho*	1-thiomorpholinyl		*E*-cinnamyl	9.88×10^{-2}
94		*meta*	1-thiomorpholinyl		*E*-cinnamyl	9.88×10^{-2}
95 [b]		*para*	1-thiomorpholinyl		*E*-cinnamyl	2.47×10^{-2}
96		*ortho*	1-thiomorpholinyl		*p*-Cl-phenyl	9.23×10^{-2}
97 [b]		*meta*	1-thiomorpholinyl		*p*-Cl-phenyl	2.31×10^{-2}
98		*para*	1-thiomorpholinyl		*p*-Cl-phenyl	9.23×10^{-2}
99		*ortho*	1-thiomorpholinyl		*p*-biphenyl	8.25×10^{-2}
100		*meta*	1-thiomorpholinyl		*p*-biphenyl	8.25×10^{-2}
101		*para*	1-thiomorpholinyl		*p*-biphenyl	8.25×10^{-2}
102		*ortho*	1-thiomorpholinyl		*E*-cinnamoyl	9.47×10^{-2}
103 [b]		*meta*	1-thiomorpholinyl		*E*-cinnamoyl	2.37×10^{-2}
104		*para*	1-thiomorpholinyl		*E*-cinnamoyl	9.47×10^{-2}
105 [b]						5.57×10^{-2}
106 [b]						3.46×10^{-2}

[a] Biological tests have been performed according to the methods described in references [10–13].
[b] Compounds used to generate the pharmacophore model.

chlorine atom at the *para* position of the latter phenyl group is able to locate itself within the HY1 hydrophobic feature very close to RA1. The key difference between the *Z* and *E* isomers is shown in the mapping and involves the imidazolylmethyl portion of the compounds. The *E*-isomer **36** places the imidazole moiety very close to the HBA feature of the hypothesis in such a way so as the N3 nitrogen atom of the imidazole group is able to map the HBA feature. Conversely, the imidazole ring of the *Z*-isomer **40** reaches the HY2 hydrophobic feature of the hypothesis. This result suggests that the HY2 feature fulfills a region of the space that is very important for the biological activity (**40** is 16 times more active than **36**). In addition, although the chemical functionalities are different for each generated

hypothesis, an HY or RA feature is present in the region occupied by HY2. Therefore, it is likely that a hydrophobic feature in this position may be a representative portion of the common pharmacophore for this class of compounds. More quantitative considerations are not possible because of the very similar values of antimycobacterial activity of these molecules. Moreover, HipHop method does not take into account relative biological data of compounds and thus is not intended for predicting activities but only to highlight the common chemical features shared by the molecules.

The five feature pharmacophore hypothesis was then used as a three-dimensional query to perform a database search in order to find other structural motifs that fulfill the functional and spatial constraints of the model. This method constitutes a powerful way for quickly finding new potential lead compounds in a medicinal chemistry project. In particular, NCI (National Cancer Institute), Maybridge, MiniBioByte and Sample databases (provided by MSI along with CATALYST 4.0), containing a total number of approximately 178 600 compounds, were searched using the pharmacophore model as a query. As a result of this search, 2249 compounds (about 1.25% of total compounds) were retrieved and then pruned according to a molecular weight cutoff set to 600. This filter was applied to avoid excluded volume problems. In fact, compounds that possess the appropriate chemical functions for a good fit onto the pharmacophore hypothesis may also

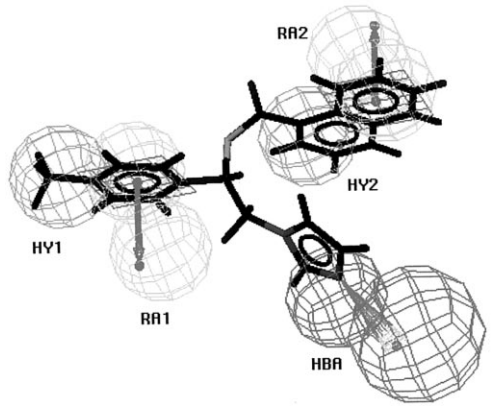

Fig. 1. Compound **63** mapped to the five-feature hyothesis (the final pharmacophore model) obtained by superimposition and merging of hypothesis 1 and hypothesis 3. RA, atomatic ring feature; HY, hydrophobic feature; HBA, hydrogen bond acceptor feature.

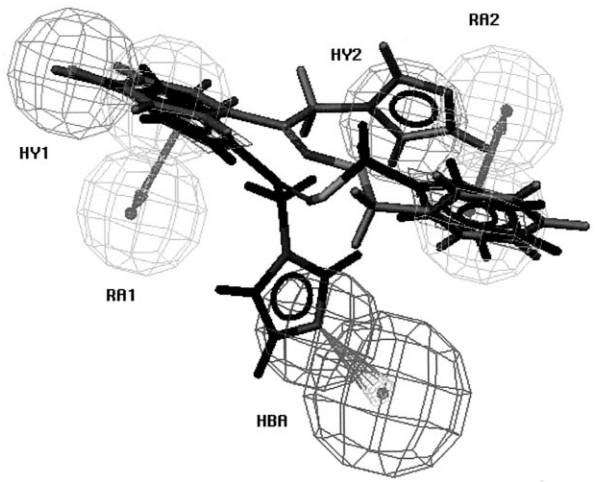

Fig. 2. Compound **40** (*Z*-isomer, grey) and compound **36** (*E*-isomer, black) mapped on the five-feature hypothesis. RA, atomatic ring feature; HY, hydrophobic feature; HBA, hydrogen bond acceptor feature.

Table 2
Summary of properties associated to the best three pharmacophore hypotheses found by a common feature hypothesis run

Number[a]	Composition[b]	Ranking score[c]	Direct (DH) and partial hits (PH)[d]
1	RRZZ	251.988	DH: 1111111111111111 101110011111111 PH: 0000000000000000 0000000100000000
2	RRZZ	245.949	DH: 1111111111111111 101110011111111 PH: 0000000000000000 0000000100000000
3	RZZH	243.065	DH: 1111111111111111 101110101111111 PH: 0000000000000000 0100010100000000

[a] Progressive numbers of the pharmacophore hypotheses.

[b] R: aromatic ring feature; Z: hydrophobic feature; H: hydrogen bond acceptor feature.

[c] The higher the ranking score, the less likely it is that the molecules in the set fit the hypothesis by a chance correlation. The best hypothesis shows the highest value.

[d] Each number in the DH and PH rows corresponds to a particular compound used to generate the pharmacophore hypotheses. DH and PH indicate whether (1) or not (0) a molecule mapped every features or all but one features in the hypothesis, respectively, giving an idea of how well each compound is able to fit to a particular hypothesis. For example, the alignment of the ninth, tenth and fourteenth compounds (numeration is from right to left in both the DH and PH rows) to the first hypothesis is not the best possible (DH values equal to zero mean that these compounds are unable to map all the features of this hypothesis). On the other hand, if we consider the third hypothesis, an optimal alignment for the ninth compound (DH = 1) and a good alignment for the tenth and fourteenth compounds (PH = 1, in both cases; these compounds map all but one hypothesis features) can be highlighted. These considerations provide some help to choose the hypotheses to be compared, superimposed and merged in order to obtain more complex and complete pharmacophore hypotheses.

have some functionalities that bump into a portion of the receptor preventing close contacts of the binding functions. On the basis of this consideration, we chose 600 as an optimal cutoff to reduce the occurrence of this problem. The remaining 1837 compounds (about 1.03% of the total) were submitted to a geometric fit calculation and ranked by descending fit numbers.

As a result, in addition to mapping some compounds structurally similar to the training set (see Chart 2), the pharmacophore model has also retrieved compounds belonging to different structural classes of well known antitubercular agents, assuming that they bind to the receptor in the same way as the molecules of the training set. In particular, miconazole and (*E*)-oxiconazole, identified by the database search, showed an antitubercular activity of 1.03×10^{-2} μg ml^{-1} and 9.32×10^{-3} μg ml^{-1}, respectively (determined accordingly to the methods described in references [11–13]). Moreover, a variety of diarylthioureas, thiosemicarbazones, and isoniazid derivatives were also stored (Chart 3). There is some experimental evidence showing that the above cited classes of molecules act as antitubercular compounds on the basis of the same mechanism. In particular, a number of diarylthioureas have shown

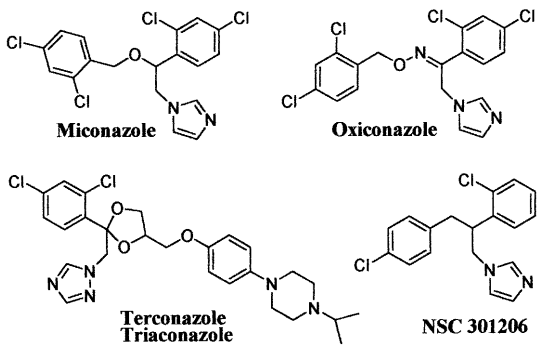

Chart 2. Representative examples of compounds found in the database search and structurally related to the molecules used to generate the pharmacophore hypothesis.

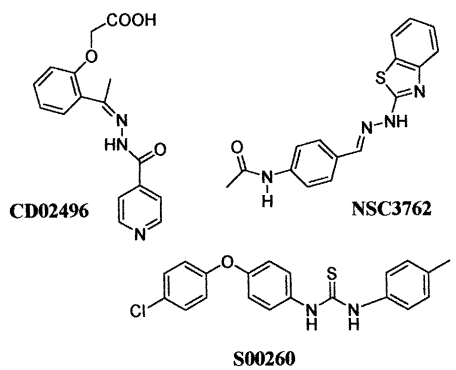

Chart 3. Representative examples of compounds found in the database search and belonging to classes of molecules known as mycolic acid biosynthesis inhibitors such as isoniazid, thiosemicarbazone, and diarylthiourea derivatives.

considerable activity against *M. tuberculosis* and one of these (4-4′-diisoamyloxydiphenylthiourea) has proved to strongly inhibit mycolic acid synthesis [16,17], as isoniazid and ethionamide do [18]. Moreover, though the actual mode of action of thiacetazone (a member of the thiosemicarbazones) is unknown, it has been suggested that also thiacetazone, like ethionamide, might inhibit mycolic acid biosynthesis [19] since some thiacetazone-resistant strains of *M. tuberculosis* exhibit cross-resistance to ethionamide [20].

A subsequent investigation aimed at checking the presence of additional thioureas, thiosemicarbazones, and isoniazid derivatives within the searched databases, highlighted some interesting compounds such as thiacetazone ({*N*-[4-[[(aminothiomethyl)hydrazono]-methylene]phenyl]-}acetamide), thiambutosine ({*N*-(4-butoxyphenyl)-*N*′-[4-(dimethylamino)phenyl]-}thiourea), and isoniazid, not identified during the database search. On the other hand, the pharmacophore model retrieved many compounds structurally related to those reported in Chart 3, all characterized by a higher number of features with respect to that of thiacetazone, thiambutosine, and isoniazid. In the light of these last findings, the database search missed the identification of the most simple compounds (thiacetazone, thiambutosine, and isoniazid) because the pharmacophore model described here and used for the search corresponds to an overall explanation of the common chemical features present in the most active compounds reported in this paper. Thus, compounds unable to fit one or more features of the model have been rejected by the search.

In conclusion, we suggest that the proposed pharmacophore hypothesis represents a model able to accommodate a variety of derivatives corresponding to the precursors of well known antitubercular drugs such as thiacetazone, thiambutosine, and isoniazid.

With these considerations, our pharmacophore model, which was able to identify thiourea, thiosemicarbazone and isoniazid derivatives in the database search, seems to be a promising model to describe the structural properties required for antitubercular activityor, at least, inhibition of mycolic acid biosynthesis in mycobacteria.

3. Conclusions

By the application of a two step computational protocol, for the first time we have proposed a general pharmacophore model for antitubercular compounds that might interfere with the biosynthesis of mycolic acid. This model has now been used to retrieve from databases of known molecules some structures that all contain the chemical features shared by the original set

of compounds. Some of these stored compounds belong to classes of molecules already reported as antitubercular agents (diarylthioureas [16,17], thiosemicarbazones [19], and isoniazid [18] derivatives) acting by the same mechanism of action based on the inhibition of the mycolic acid synthesis. On the basis of these results, the pharmacophore will hopefully lead to the identification of new antitubercular compounds by synthetic modification and optimization of the templates found in the database search.

4. Computational methods

A set of 32 derivatives with biological data ranging from 1.39×10^{-3} μmol ml^{-1} (compound **40**) to 1.46×10^{-1} μmol ml^{-1} (compound **3**), was used in this study. Due to the fact that the biological evaluation of all the chiral compounds was carried out using racemic mixtures, while the biological activity of chiral compounds is usually due to one of the enantiomers, it was arbitrarily decided to model the chiral compounds **41**–**68** with undefined chirality. Thus, allowing the pharmacophore model generation procedure to choose which configuration of the asymmetric carbon atom common to compounds **41**–**68** was most appropriate. This decision can be justified on the basis that a number of chiral compounds (namely **44**, **46**, **49**, **56**, **58**, **63**, **67**, and **68**) were included in the training set and compound **63** was used as one of the five 'reference compounds' (see below). *S* enantiomer of the eight compounds has been chosen by the program to fit the proposed pharmacophore model. To the best of our knowledge, there are no experimental data supporting the hypothesis that *S* enantiomers are more active than *R* enantiomers. In fact, although some antifungal imidazole derivatives have been recently tested as enantiopure compounds [21], none of them have been evaluated for antitubercular activity.

The compounds were built using CATALYST 2D–3D sketcher and a representative family of conformations were generated for each molecule using the Poling algorithm [22] and the 'best conformational analysis' method. This approach has been selected because during the analysis of the conformation, it takes into account the spatial arrangement of chemical features instead of the arrangement of atoms. In addition, this also provides the best possible conformational coverage within CATALYST. This approach represents the method of choice when the conformations are to be used for hypothesis generation, as in this case. Conformations were selected that fell within 20 kcal mol^{-1} range above the lowest energy conformation found.

The training set of molecules, with their associated conformational models, were submitted to common feature hypothesis generation producing potential pharmacophore models by generating alignments of common chemical features. To perform the alignment of the molecules, instead of using only the minimum energy conformers found, all the conformers generated for each compound have been included in the calculations. The chemical functions (features) used in this generation step included hydrogen bond donor (HBD), hydrogen bond acceptor lipid (HBA), hydrophobic (HY), positive ionizable (PI), and aromatic ring (RA) [23]. Hydrogen bond acceptor lipid includes basic nitrogen which is not considered in the hydrogen bond acceptor feature. An aromatic ring function was used in addition to the generic hydrophobic function in order to emphasize that an aromatic interaction is likely to be an important feature because all the compounds show one or more aromatic rings within their structure. Moreover, the aromatic ring function allows consideration of directionality of the interactions with the receptor. Due to the size of the molecules, the minimum permitted interfeature spacing was reduced to 1 Å. In fact, a default value of about 3 Å works well for most medium to large molecules, but is not good for small molecules that do not have many features or for hypotheses where features are close together.

The five most active compounds (**40**, **63**, **38**, **34**, and **33** ranked by decreasing activity) were used to derive common-feature based alignments and considered as 'reference compounds' (except for this classification, biological data in the analysis were not used) specifying a 'Principal' value of 2 and a 'MaxOmitFeat' value of 0. HipHop uses these values to determine which molecule should be considered to build hypothesis space and how many features in the final hypotheses must map to the chemical features in each compound, respectively. If Principal is set to 2, the chemical feature space of the conformers of such a compound is used to define the initial set of potential hypotheses. Principal equal to 1 means that this compound must map onto the pharmacophore hypotheses previously defined. A MaxOmitFeat value of 0 associated with the five reference compounds forces them to map all the features of each pharmacophore hypothesis. Principal and MaxOmitFeat values for the remaining 27 compounds of the training set were set to 1 and 2, respectively.

Acknowledgements

This work has been partially supported by contribution of the 'Istituto Pasteur-Fondazione Cenci Bolognetti' Università di Roma 'La Sapienza' and the National Tuberculosis Projects 'ISS-Ministero della Sanità — Grant No. 96/D/T48'. The Italian CNR,

Progetto Strategico 'Modellistica Molecolare di Sistemi Molecolari Complessi' is gratefully acknowledged.

References

[1] TB: Groups at risk, WHO report on the tuberculosis epidemic. World Health Organization, Geneve, Switzerland, 1996.
[2] I. Chopra, P. Brennan, Tuberc. Lung Dis. 2 (1998) 89–98.
[3] R. Fioravanti, M. Biava, G.C. Porretta, G. Lampis, C. Maullu, R. Pompei, Med. Chem. Res. 9 (1999) 249–266.
[4] R. Fioravanti, M. Biava, G.C. Porretta, M. Artico, G. Lampis, D. Deidda, R. Pompei, Med. Chem. Res. 7 (1997) 87–97.
[5] R. Fioravanti, M. Biava, G.C. Porretta, G. Lampis, D. Deidda, R. Pompei, Med. Chem. Res. 9 (1999) 162–175.
[6] M. Biava, R. Fioravanti, G.C. Porretta, G. Sleiter, A. Ettorre, D. Deidda, G. Lampis, R. Pompei, Med. Chem. Res. 7 (1997) 228–250.
[7] M. Biava, R. Fioravanti, G.C. Porretta, G. Sleiter, A. Ettorre, D. Deidda, G. Lampis, R. Pompei, Med. Chem. Res. 8 (1998) 523–541.
[8] M. Biava, R. Fioravanti, G.C. Porretta, D. Deidda, G. Lampis, R. Pompei, Farmaco 54 (1999) 721–727.
[9] New horizons in the treatment of tuberculosis are described in detail in: C.E. Barry, Biochem. Pharmacol. 54 (1997) 1165–1172, and in reference 13.
[10] J.E. Hawkin, R.J. Wallace, Jr., B.A. Brown, Antibacterial susceptibility test: mycobacteria, in: A. Balows, W.J. Hausler, Jr., K.L. Hermann, H.D. Isenberg, H.J. Shadomy (Eds.), Manual of Clinical Microbiology, ASM, Washington, DC, 1991, pp. 1138–1152.
[11] H. David, Bacteriology of the mycobacterioses. DHEW Publication No. (CDC) 768316, CDC, Mycobacteriology Branch, Atlanta, GA, 1976.
[12] National Commitee for Clinical Laboratory Standards. Antimycobacterial susceptibility testing, Proposed Standard M24, Villanova, PA, 1990.
[13] B.A. Brown, J.M. Swenson, R.J. Wallace, Jr., Broth microdilution test for rapidly growing mycobacteria, in: Clinical Microbiology Procedures Handbook, vol. I, ASM, Washington, DC, 1992.
[14] CATALYST 4.0 is a MSI product, San Diego, CA, USA.
[15] D. Barnum, J. Greene, A. Smellie, P Sprague, J. Chem. Inf. Comp. Sci. 36 (1996) 563–571.
[16] F.G. Winder, P.B. Collins, D. Whelan, J. Gen. Microbiol. 66 (1971) 379–380.
[17] B. Phetsuksiri, A.R. Baulard, A.M. Cooper, D.E. Minnikin, J.D. Douglas, G.S. Besra, P.J. Brennan, Antimicrob. Agents Chemother. 43 (1999) 1042–1051.
[18] D.A. Rozwarski, G.A. Grant, D.H.R. Barton, V.R. Jacobs, J.C. Sacchettini, Science 279 (1998) 98–102.
[19] F.G. Winder, The biology of the mycobacteria, in: C. Ratledge, J. Stanford (Eds.), Physiology, Identification and Classification, Academic Press, London, 1982, pp. 353–438 vol. 1.
[20] X.-S. Pan, J. Ambler, S. Mehtar, M. Fischer, Antimicrob. Agents Chemother. 40 (1996) 2321–2326.
[21] S.Y. Ablordeppey, P. Fan, J.H. Ablordeppey, L. Mardenborough, Curr. Med. Chem. 6 (1999) 1151–1195.
[22] A. Smellie, S. Teig, P. Towbin, J. Comp. Chem. 16 (1995) 171–187.
[23] J. Greene, S. Kahn, H. Savoy, S. Teig, J. Chem. Inf. Comp. Sci. 34 (1994) 1297–1308.