Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 76 (1) 85–99 (2011)
JSCS–4102
UDC 547.575'461.5:66.097.8+
519.711+54–124:57–188
Original scientific paper

MTD–CoMSIA modelling of HMG-CoA reductase inhibitors

DANIEL M. DUDA-SEIMAN[1], SPERANȚA AVRAM[2], SILVIA MANCAȘ[1], VALENTIN CAREJA[3], CORINA DUDA-SEIMAN[4]*, MIHAI V. PUTZ[4] and DAN CIUBOTARIU[5]

[1]Department of Medical Ambulatory, Medical Emergencies, University of Medicine and Pharmacy "Victor Babes", Timisoara, [2]Department of Physiology and Biophysics, University of Bucharest, Faculty of Biology, Bucharest, [3]Institute of Chemistry "Coriolan Drăgulescu" of the Romanian Academy, Timisoara, [4]Department of Chemistry, University of West Timisoara, Faculty of Chemistry-Biology-Geography, Timişoara and [5]Faculty of Pharmacy, University of Medicine and Pharmacy "Victor Babeş", Timisoara, Timisoara, Romania

(Received 1 June, revised 23 August 2010)

Abstract: The 3D quantitative structure–activity relationship for a series of hydroxymethylglutaryl-CoA (HMG-CoA) reductase inhibitors based on the pyrrolylethyl-tetrahydropyranone scaffold was examined using the Minimal Topological Difference (MTD) method and comparative molecular similarity index analysis (CoMSIA). The studied compounds were of the tetrahydro-4-hydroxy--6-[2-(1*H*-pyrrol-1-yl)ethyl]-2*H*-pyran-2-one type. In clinical practice, HMG--CoA reductase inhibitors are usually referred to by the generic name statins. The analysis performed using the MTD method showed that voluminous substituents produce a significant biological activity (R^2_{CV} = 0.677 > 0.5; *SEECV* = = 0.319), while the CoMSIA method added useful information regarding the influence of the steric, electrostatic, hydrophobic, hydrogen bond donor, and acceptor properties on biological activity (R^2_{CV} = 0.60; r^2 = 0.98).

Keywords: statins; molecular modelling; correlations; 3D-quantitative structure–biological activity.

INTRODUCTION

Among cholesterol-lowering drugs, statins (hydroxymethylglutaryl-CoA reductase inhibitors) manage effectively artherogenic dyslipidemia,[1] with immediate and long-term consequences on cardiovascular morbidity and mortality.[2]

Statins inhibit competitively the enzyme hydroxymethylglutaryl-CoA reductase (HMG-CoA-reductase)[3] and they play a fundamental role in the modern concept of artherothrombotic cardiovascular disease prevention.[4] Statins have realized a significant reduction of global cardiovascular risk, as well as of cardio-

* Corresponding author. E-mail: cori_mam@yahoo.com
doi: 10.2298/JSC100601019D

vascular mortality.[2] The absolute vascular benefit of statins is obtained when the level of low-density-lipoproteins (LDL – "bad cholesterol") is reduced below 70 mg dL^{-1}.[5,6] Thus, in the ARBITER trial, one-year aggressive statin therapy led to a 48.5 % reduction in LDL-cholesterol to 76 mg dL^{-1}, being associated with a significant regression of the carotid intima-medial thickness (p = 0.03).[5,7] Achieving the above mentioned LDL goals with aggressive statin therapies, HDL-cholesterol levels improved significantly and coronary arteriosclerosis may regress.[8]

The HMG-CoA-reductase inhibition has a double action mechanism: decreasing hepatic cholesterol synthesis and increasing LDL catabolism.[9]

Besides correcting the lipid fractions, statins show other beneficial effects on the arterial wall: they modulate the endothelial function and they have anti-inflammatory and anti-proliferative effects, thereby stabilizing the atheroma plaque and preventing thrombogenesis.[10] Due to these now well-established effects of statins, it has been suggested that they could be a choice of treatment in a large variety of diseases, including dementia and autoimmune disorders.[11]

Hydroxymethylglutaryl-CoA reductase inhibitors were well established as the cornerstone of pharmacoprevention of atherosclerotic and atherothrombotic arterial disease by means of primary and secondary cardiovascular prevention trials.[12]

Molecular modelling requires not only good algorithms but also a close integration with experiment and a better understanding of the underlying physical and biological principals involved.[13] Thus, molecular modelling studies would allow a more profound documentation of ligand (L)–receptor (R) interactions and a better specification of optimal structural requirements. The present study investigates the biologic activity of HMG-CoA reductase under the interaction with a representative series of statins by employing the 3D-minimal topological difference (3D-QSAR MTD)[14] and comparative molecular similarity indices analysis (CoMSIA) methods.[15]

Background methods

The MTD method. The MTD method aims to provide a description of the molecular stereochemistry (of both receptor and effectors, *i.e.*, drug molecules) by indicating the presence or absence of atoms from the considered M_i molecule, i = 1,...,N, in the vertices of the hypermolecule, H (which describes the receptor).[14,16,17] The hypermolecule is obtained by an approximately atom-by-atom superposition of all M_i molecules, neglecting the hydrogen atoms. The hypermolecule H can be considered as a topological network in which the vertices correspond to atoms, while the edges are derived from the corresponding chemical bonds. If molecule i occupies vertex j, this may be accounted for by x_{ij} = 1,

while otherwise $x_{ij} = 0$. Then, the minimal steric difference, MTD_i, of molecule i with respect to the receptor is calculated according to the formula:[14]

$$MTD_i = s + \sum_j \varepsilon_j x_{ij} \quad (1)$$

with $\varepsilon_j = -1$, 0 or +1 for vertices attributed to the receptor cavity (beneficial), to the exterior (irrelevant) and to the receptor walls (detrimental), respectively; s is the total number of cavity vertices. Thus, the degree of steric misfit for the molecule M_i with respect to the receptor, *i.e.*, MTD_i, is defined as the sum of the number of unoccupied cavity vertices of H and the number of occupied wall vertices of H.

The attributions of $\varepsilon_j = -1$, 0 or +1 to the vertices j are performed according to an optimization procedure (see below). One starts from an initial attribution of j vertices, ε_j^0, and changes these ε_j settings towards an increase in the correlation coefficient, r^2, by measuring the degree of residues minimization for a predicted equation of the type:

$$A_i = \alpha - \beta MTD_i \quad (2)$$

The resulting picture of H (meaning cavity, wall and irrelevant zone) obtained by the optimization procedure constitutes a hypothetical predicted steric receptor map.

The CoMSIA method. The CoMSIA method was developed to improve the limitations of the steric and electrostatic fields in comparative molecular field analysis (CoMFA).[15] Being an extension of the CoMFA approach, the CoMSIA method incorporates five different property fields: steric, electrostatic, hydrophobic, and hydrogen bond donor and acceptor properties. In CoMSIA, the screening interactions between the inhibitors and a probe atom having a radius of 1 Å, charge of +1 and hydrophobicity of +1 are considered. Using the CoMSIA method, similarly to the CoMFA method, the aim is to divide the contribution of the steric, electrostatic, hydrophobic, and hydrogen bond donor and acceptor properties and to evaluate these contributions with respect to the biological activity. This indicates which factors have the largest influence on the receptor binding affinities.[15,18]

The aim of the CoMSIA method employed here is to use predictive 3D QSAR models in order to understand the enzyme inhibitor interaction at the active site of HMG-CoA reductase and, in particular, to pinpoint which structural features are responsible for its selectivity. Future work based on the models developed here will therefore allow the design of selective inhibitors of HMG-CoA reductase with superior clinical profiles for the treatment of cardiovascular diseases.

The CoMSIA method was employed to evaluate various descriptors in the context of inhibition of HMG-CoA reductase, such as the electrostatic, steric,

hydrophobic and donor effects. As opposed to MTD (and other QSAR methods), the CoMSIA method enables the determination of the individual contribution of each descriptor to the biological activity of HMG-CoA reductase inhibitors.

COMPUTATIONAL DETAILS

The general structure of the studied series of hydroxymethylglutaryl-CoA reductase inhibitors is shown in Fig. 1.

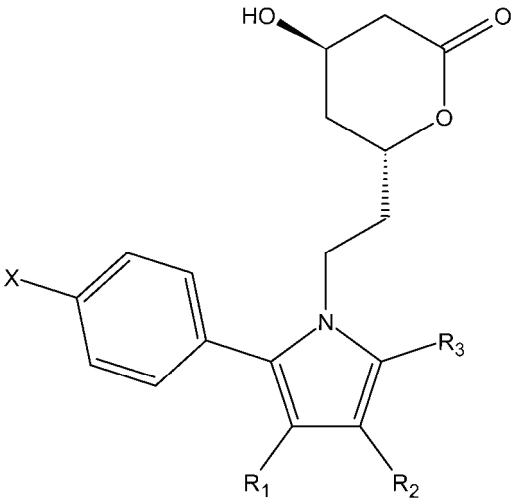

Fig 1. General structure of the statins.

The statins inhibitory activity data were selected considering the criteria of low values for the biological activity, IC_{50}, and through the existence of a high number of different substituent groups (*e.g.*, phenyl, CO_2CH_3, 2-pyridyl, 3-pyridyl, methyl, *etc*.).[19]

HMG-CoA reductase was expressed using the common logarithm, log ($1/IC_{50}$), for the inverse of experimental concentration of a compound leading to 50 % inhibition (IC_{50}), see Table I.

Note that the absolute *R*/*S* configurations for the compounds listed in Table I were not considered for the MTD calculations.

The local minimum potential energy was determined within Polack-Ribiere conjugate--gradient method with a convergence of the total root-mean-square (*RMS*) gradient as 0.01 kcal $mol^{-1} \cdot Å^{-1}$. The method allows geometry optimization (energy minimization) through finding the Cartesian coordinates $\mathbf{r}_i$ of a molecular structure that represent the potential energy minimum, $\partial V / \partial \mathbf{r}_i = 0$, *i.e.*, when the interatomic forces are minimum. This is searched by cycles with default conjugate gradient directions having the *RMS* of the current configuration of atoms that is close to zero on local minimums (see above); the geometry optimization results in a new structure at a minimum for which the atomic coordinates and energy can be examined. After energy minimization, the Gasteiger–Marsili partial charges[15,20,21] were used for the CoMSIA methodology.

The steric field, the electrostatic field energies, as well as the hydrophobic and donor characters of each inhibitor were calculated at the intersection points of a regularly spaced grid (2Å) in a grid box surrounding the molecules. Sybyl/CoMSIA software was used to

calculate the electrostatic and steric potential functions within the Tripos force field,[15] while using an sp^3 carbon atom with a +1 charge as the probe atom.

TABLE I. The series of tetrahydro-4-hydroxy-6-[2-(1*H*-pyrrol-1-yl)ethyl]-2*H*-pyran-2-one molecules with the experimental 50 % inhibition concentrations, IC_{50}, on the HMG-CoA reductase together with the employed $A_{observed} = -\log (IC_{50} / \mu mol\ dm^{-3}) = 6 - \log (IC_{50} / mol\ dm^{-3})$ activity in the actual study[19]

Molecular labels[a]	Substituents				IC_{50} μmol dm^{-3}	A_{obs}
	X	R_1	R_2	R_3		
1	F	H	H	*i*-Pr	0.230	6.638
3a	H	Ph	CO_2Et	CH_3	4.000	5.398
3b	H	Ph	CO_2Et	Et	0.890	6.051
3c	H	Ph	CO_2Et	*i*-Pr	0.170	6.77
3d	F	CO_2CH_3	CO_2CH_3	*i*-Pr	0.180	6.745
3e	F	CO_2Et	CO_2Et	*i*-Pr	0.350	6.456
3f	F	CO_2Et	Ph	*i*-Pr	0.050	7.301
3g	F	Ph	CO_2Et	*i*-Pr	0.200	6.699
3h	F	Ph	CO_2CH_2Ph	*i*-Pr	0.040	7.398
3i	F	Ph	CONHPh	*i*-Pr	0.025	7.602
3j	F	4-CN-Ph	CO_2Et	*i*-Pr	0.280	6.553
3k	F	CH_3	CH_3	*i*-Pr	0.140	6.854
3l	F	Ph	H	*i*-Pr	0.347	6.46
3m	F	2-pyridyl	H	*i*-Pr	0.046	7.337
3n	F	3-pyridyl	H	*i*-Pr	0.071	7.149
3o	F	4-pyridyl	H	*i*-Pr	0.310	6.509
3p	F	H	Ph	*i*-Pr	0.120	6.921
30a	F	Cl	Cl	*i*-Pr	0.028	7.553
30b	F	Br	Br	*i*-Pr	0.028	7.553
30c	F	$COCF_3$	H	*i*-Pr	0.800	6.097
33	F	Ph	CONHPh	*i*-Pr	0.007	8.155

[a]nomenclature from ref. 19

The geometry of the statin molecules under study was optimized using molecular mechanics and quantum chemistry methods. The structures obtained by minimizing the internal energy with the aid of the MM+ force field were used as input data for the program AM1 (Austin Model 1).[13] An energy cut-off of 30 kcal mol^{-1} was used for both electrostatic and steric contributions. The relatively high rigidity of the statins structure justified the molecular construction in bi-dimensional space (2D) using the classic formulas, where the optimized structures **1** to **33** were overlain with the HyperChem programme package, using molecular mechanics techniques of the MTD algorithm in the superposition process for designing the hypermolecule H.

The hydrogen atoms were not considered and common vertices of the 3D molecular graphs in Fig. 2b were specified in the parent structure of the hypermolecule (numbered vertices).

In the superposition process, those conformations from the minimal step interval +1 kcal assuring a maximal coverage, *i.e.*, a minimal number of considered (numbered) vertices, were considered.

Regression analysis was performed by the Partial Least Squares (PLS) algorithm within Sybyl 7.2. The cross-validation procedure was also used in order to assess the reliability of the

obtained results. This allows for the separation of steric factors (measured by MTD_i values) from the others, such as hydrophobic, electrical, *etc*., which possibly control the biological response. The leave-one-out cross-validation method using the SAMPLS (a program created by Brucle Bush at DuPont which vastly accelerates cross-validation calculations for PLS analyses involving COMFA, COMSIA fields) was used.[22] The optimal number of components for the final 3D-QSAR equation was chosen based on the highest q_2 (cross-validated r^2) value.[16] Field points with a standard deviation below 2.0 were dropped from the PLS calculation. The statistical results q_2 (cross-validated r^2),[23,24] fitted correlation coefficient r^2, Fisher test (F), SEE (standard error of estimate)[23,25] and the electrostatic, steric, hydrophobic and donor contributions[15,25] were calculated as implemented in Sybyl 7.2.

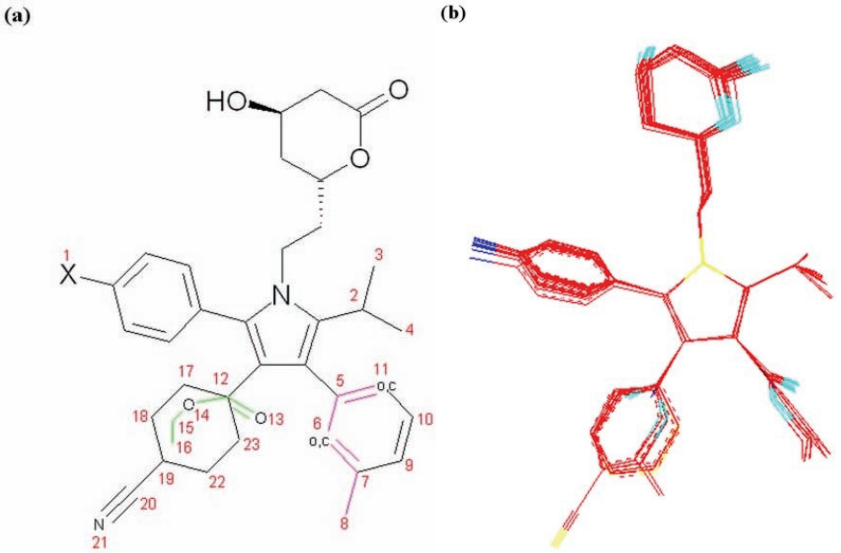

Fig. 2. a) 2D-Hypermolecule for the superimposed statins of Table I, excepting **3h**, (±)-**3i**, **30c**, (+)-**33** and (–)-**33** molecules, with the marked vertices. The green lines indicate a carboxyl or ester group which in space is located in a perpendicular plane to the benzene hexagon plane, while with pink lines, the same group type partially superposed (on vertices 5, 6, 7, 11) upon the benzene hexagon plane is indicated. The molecules with higher flexibility (**3a**, **3b**, **3d** and **3e** of Table I) had a position outside the benzene plane and could introduce supplementary vertices. However, they were partly "laid down" upon benzene and after the rotations, they become the conformers **3a70**, **3b70**, **3d70** and **3e70** (see Table II); consequently, the number of vertices decreased and the hypermolecule was realized. A similar procedure could not be applied to the molecules in the green zone because of steric misfit and of increased rotation impediments. b) The same hypermolecule as in a) but as a 3D-representation.

RESULTS AND DISCUSSION

The starting map S^0 for the MTD optimization procedure,[17] which was obtained by an inspection of vertices found preferentially in molecules with high and respectively low activity for all the compounds given in Table I, looks like:

$$S^0 = \begin{cases} -1: 1,2,3,4,5,12 \\ 0: 9,10,13,14,15,16,20,21 \\ +1: 6,7,8,11,17,18,19,22,23 \end{cases} \quad (3)$$

It is based on the 2D-representation of the standard benchmark molecule **3a** in Table I.

In MTD, the starting map S^0 is automatically chosen using the interquartile interval. The structures were sorted by their inhibitory potency in parallel with descending activity, given in Table I: occupied vertices in all (or in almost all) the most active molecules (usually the first quarter) are considered in the cavity (–1), and those in the last quarter (molecules with the weakest activity) are considered in the wall (+1). The others, in the middle of the series, are irrelevant (0). Note that the S^0 coefficients were set at the beginning according to the recorded activity (the most/least influential), while the S^* coefficients paralleling the cross-validation process are automatically delivered.

The studied series of statins shows an increased conformational degree of liberty. Thus, initially, the obtained hypermolecule contained the superposition of all 21 molecules in the series; some of them introduced in a disadvantageous way many more supplementary vertices. These facts imposed the removal of four molecules from the studied series, namely **3h**, **3i**, **30c** and **33**.

In this way, the obtained results are based on a statins series containing the remaining 17 compounds; from them four did not superpose well, namely the molecules **3a**, **3b**, **3d** and **3e** in Table I, see Fig. 2a.

However, on closer inspection, it was observed that some bonds in the mentioned molecules, due to their high conformational flexibility, could undergo some rotations. Hence they may be forced to better superimpose on others when appropriate torsions are applied; actually, referring to just one valence angle, which was changed by about 21–34° in each molecule the energy required in the torsion process was computed, see Table II.

TABLE II. The four molecules of Table I for which smooth low energy torsion were performed in order to overlap with the rest of the minimal energy conformations of molecules in the hypermolecule of Fig. 2

Molecule	Angle, °	Torsion, °	Heat of formation, kcal/mol	Energy difference, kcal/mol
3a	36.18	33.82	–6674.52	0.96
3a70	70	–	–6673.56	–
3b	35.79	34.21	–6955.2	0.86
3b70	70	–	–6954.34	–
3d	42.2	27.8	–6403.09	0.73
3d70	70	–	–6402.36	–
3e	48.71	21.29	–6965.08	0.77
3e70	70	–	–6964.31	–

If the resulting torsion energy were appreciable, say over 1.0 kcal mol^{-1}, the respective molecule must be excluded from the computation series, since it would represent a departed conformation from the congeneric series in focus. Accordingly, the results in Table II clearly indicate that the above-mentioned molecules with high conformational flexibility may be superposed upon the others when slightly rotated. In this way, the 3D-hypermolecule of Fig. 2b was built.

The optimized 3D structures were used for overlapping. From this point of view, the methodology of the MTD method may be considered as 2.5D. Such a superposition procedure is the basis for the description of molecular stereochemistry in the MTD method. Equivalent atoms with inter-distances less than 0.5 Å are considered to occupy the same vertex; the inter-distance 0.5 Å limit was justified by an analysis of the dependence of the Van der Waals interaction energies on the interatomic distances.[17]

Vertices occupancy x_{ij} (the junctions) of each molecule in the hypermolecule H of Fig. 2a, together with resulting molecular MTD parameters, are provided in Table III.

TABLE III. Vertices occupancy x_{ij} (junctions), together with the MTD parameters computed with Eq. (1), MTD-predicted activities, MTD residual values, with respect to those observed from Table I, for all 17 molecules of the hypermolecule H of Fig. 2

Molecule	Vertex																							*MTD*	A_{MTD}	A_{obs}–A_{MTD}
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23			
1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	6.648	–0.0102
3a	0	1	0	0	1	1	1	1	0	0	1	1	0	0	0	0	1	1	1	0	0	1	1	11	5.691	–0.2942
3b	0	1	0	1	1	1	1	1	0	0	1	1	0	0	0	0	1	1	1	0	0	1	1	10	6.01	0.0398
3c	0	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	1	1	0	0	1	1	9	6.329	0.4398
3d	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	8	6.648	0.0958
3e	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	9	6.329	0.1258
3f	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	6	7.286	0.0148
3g	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	1	1	0	0	1	0	8	6.648	0.04982
3j	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	1	1	1	1	1	1	8	6.648	–0.0962
3k	1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	6	7.286	–0.4332
3l	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	0	0	1	1	7	6.967	–0.5082
3m	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	0	0	1	1	7	6.967	0.3698
3n	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	0	0	1	1	7	6.967	0.1808
3o	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	0	0	1	1	7	6.967	–0.4592
3p	1	1	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	7	6.967	–0.0472
30a	1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	6	7.286	0.2658
30b	1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	6	7.286	0.2658

As such, the MTD method reveals the structural requirements necessary to have a significant biological activity[26,27] as the hypothetically hypermolecule H of Fig. 2 is predicted. Through conformational analysis calculations, the MTD

method prescribes that the optimized map at the receptor level should maximally fit with the statins ligand map:

$$S^* = \begin{cases} -1: 3,4,5,9,10,12,17,18,20 \\ 0: 1,6,13,14,15,16,19,21 \\ +1: 2,7,8,11,22,23 \end{cases} \quad (4)$$

whereas the predicted activity correlation equation unfolds as:

$$A_i = 9.2 - 0.319 MTD_i,\ i = 1,...,N = 17 \quad (5)$$

with the statistical factors: $r = 0.853$, $r^2 = 0.728$, $EV = r_{adj}^2 = 0.70984$, $PRESS = 1.288$, $SEE = 0.293$ and $F = 40.142$, as the correlation, squared correlation, explained variance (or adjusted correlation), prediction sum of squares, standard error of estimate and the Fisher test, respectively.

The values given in Table III qualify the actual MTD approach among acceptable 3D-QSAR models since both the observed and predicted values provide the same mean value of 6.761, while the mean of the residual activities goes to zero. Moreover, aiming for the internal validation procedure, the leave-half--out procedure (*i.e.*, splitting the original set of data into even and odd sets and considering, alternatively, one as the observed data and computing the remaining data leading to the final cross-validation scores) was undertaken, leading to the results:

$$PRESS_{CV} = \sum_{i=1}^{N} \left(A_i^{obs} - A_i^{MTD\text{-}CV} \right)^2 = 1.529$$

$$R_{CV}^2 = 1 - \frac{PRESS_{CV}}{\sum_{i=1}^{N} \left(A_i^{obs} - \overline{A^{obs}} \right)^2} = 0.677 > 0.5 \quad (6)$$

$$SEE_{CV} = \sqrt{\frac{PRESS_{CV}}{N-2}} = 0.319 \cong \frac{1}{9}\left(A_{obs}^{max} - A_{obs}^{min} \right)$$

suggesting the present MTD analysis is of significant value since the R_{CV}^2 value lies above the threshold of 0.5 for the statistical limit of cross-validation, while the cross-validation standard error of estimate behaves like the 9th part of all the observed activity range of the HMG-CoA reductase inhibitors in Table I.

However, it may be argued that the MTD regression seems to be over-determined: while each vertex in H represents a degree of freedom, there are 23 vertices considered against only 17 correlated molecules. In this regards, it is worth noting that since the MTD method is based on the non-null occupancies of the vertices of the hypermolecule, the maximum cardinal of occupancy among the series does not exceed the number of molecules in the series, see Table I and Eq. (4), from which it appears that the non-zero attributions (ε_j) do not exceed 15

out of 23 possibilities, revealing in fact only 15 degrees of freedom for 17 molecules, thereby confirming that the present approach is well-defined.

For completion of the 3D-activity analysis, the inhibitors listed in Table I were further used to design a CoMSIA model, the results of which are presented in Table IV. This time, the statistical significance of the model was assessed throughout by the leave-one-out cross-validated PLS analysis running with four principal components, which lead to a R^2_{CV} cross-validated correlation coefficient as high as 0.60, of comparable order with that obtained in a previous MTD analysis.

TABLE IV. The predicted and residual biological CoMSIA activities of the statin derivatives of Table I, grouped in training and test sets

Compound index	A_{CoMSIA} (predicted)	A_{obs}–A_{CoMSIA} (residual)
Training set		
1	6.57	0.06
3a	5.46	–0.07
3b	5.96	0.09
3d	6.77	–0.03
3f	7.33	–0.03
3g	6.71	–0.02
3h	7.33	0.06
3i	7.52	0.08
3j	6.58	–0.03
3k	6.84	0.01
3l	6.49	–0.04
3o	6.53	–0.03
3p	7.01	–0.09
30a	7.36	0.19
30b	7.63	–0.08
30c	6.08	0.01
33	8.23	–0.08
Testing set		
3c	6.03	0.73
3e	6.94	–0.49
3m	6.28	1.05
3n	6.34	0.8

Taking into account that the HMG-CoA reductase inhibitors display large molecular diversity through the many types of chemical substituents, the leave-one-out cross-validation method is here found to be the appropriate statistical technique. In addition, for a non-cross-validated PLS analysis, a suitable fitted correlation coefficient (r^2) of 0.98 was obtained, as revealed in Table V.

However, by comparing the statistical parameters of the CoMSIA and MTD methods, *i.e.*, $R^2_{CV\text{-}MTD}$ = 0.677106 with $R^2_{CV\text{-}CoMSIA}$ = 0.6, although the first was based on the leave-half-out, whereas the second on the leave-one-out algo-

rithm, permits the conclusion that the MTD method gives a similar robust model as the CoMSIA method. Such a comparison is justified by the fact that both models were computed on the same training set, *i.e.*, with 17 molecules, albeit slightly different, because the CoMSIA method did not discriminate the molecules **3h**, **3i**, **30c** and **33**, leading to their exclusion from training set, while the MTD analysis did it due to the revealed geometrical reasons. In this way, the MTD analysis was somewhat superior from the computational point of view; this superiority was also emphasized by the considerable lower standard error of estimate for CoMSIA with respect to the MTD analysis, see Table V and the statistical factor for Eq. (5), respectively. Nevertheless, the MTD method gains because of the starting configuration of the intuitive 3D hypermolecule, producing in the end similar statistical results. Yet, for CoMSIA analysis, the value of R^2_{CV} =0.60 is by over 0.3 units different to the determination coefficient of the model r^2 = 0.98 , compared to the difference of 0.05 between the same quantities in the case of the MTD model. Such a behaviour indicates that the present CoMSIA description could suffer either from an over-fitted circumstance by the presence of irrelevant independent variables and/or by the presence of outliers; at the same time, the MTD algorithm, although employing a 2.5D analysis, provides better self-consistent statistical factors, this being an argument for further MTD development in molecular design.

TABLE V. Synopsis of the statistical CoMSIA-PLS analysis used to evaluate the predictive quality for the biological activity of the statins derivatives of Table III used as HMG-CoA reductase inhibitors

Statistical quantity	Value
Number of molecules in the training set	17
Cross-validated R^2_{CV}	0.6
r^2	0.98
SEE (standard error of estimate)	0.08
Fisher test	249.58

Furthermore, the predictive power of CoMSIA was tested on the inhibitors included in the molecular test set given in Table IV.

From residuals given in Table IV, it is clear that the predictive power of CoMSIA was quite poor outside the training range. These results limit the usefulness of the CoMSIA analysis to the sample molecules under consideration. However, useful information from the training set is revealed, indicating the development of new inhibitors starting from **3k** and **30c** as templates with considerable low residuals with respect to those predicted by MTD analysis, see Tables IV and III, respectively.

Moreover, because the CoMSIA method predicts biological activity using electrostatic, steric, hydrophobic and donor descriptors, it is important to con-

sider their contributions to the inhibitory activities of the statins. The contribution of steric, electrostatic and donor descriptors to the inhibitory activities of the statins were just 0.154, 0.153 and 0.287 respectively, while hydrophobic contribution was found significantly higher, 0.407.

Finally, in order to properly compare the biostatistics of the MTD and CoMSIA methods, the recent Steiger test is considered here.[28] It is based on checking whether the null hypothesis according to which the dispersions of the (molecular) populations investigated by MTD and CoMSIA are equal:

$$\sigma_{\mathrm{MTD}} = \sigma_{\mathrm{CoMSIA}} \tag{7a}$$

(against the alternative $\sigma_{\mathrm{MTD}} \neq \sigma_{\mathrm{CoMSIA}}$) parallels inclusion in the confidence interval of the root-mean-square standardized effect (*RMSSE*),

$$\Psi = \sqrt{\frac{\lambda}{(p-1)n}} \tag{7b}$$

of the non-central Fisher-related statistics:

$$f_{\mathrm{CoMSIA}}^{\mathrm{MTD}} = \frac{SEE_{\mathrm{MTD}}}{SEE_{\mathrm{CoMSIA}}} \tag{7c}$$

in terms of the non-centrality parameter λ, when n observations per group (p) are statistically performed. In the present case, p = 17 molecules (as the whole group) and n = 1 (as the individual members studied for their bioactivity); while noting that for the Fisher test:

$$F_{\mathrm{CoMSIA}}^{\mathrm{MTD}} = \left(f_{\mathrm{CoMSIA}}^{\mathrm{MTD}} \right)^2 = 13.414$$

with the degrees of freedom $v_1 = v_2 = 16$, the 95 % interval for λ ranges from 70.617 to 385.227. Hence, the application of the confidence interval transformation in accordance with the inversion confidence interval principles,[28] provides the validation result:

$$f_{\mathrm{CoMSIA}}^{\mathrm{MTD}} \in \left(\Psi_{\mathrm{lower}}, \Psi_{\mathrm{upper}} \right), \tag{8a}$$

with:

$$f_{\mathrm{CoMSIA}}^{\mathrm{MTD}} = 3.66; \ \Psi_{\mathrm{lower}} = 2.1; \ \Psi_{\mathrm{upper}} = 4.91. \tag{8b}$$

In this way, the hypothesis (7a) is confirmed and therefore the bioequivalence in assessing the activity modelling by both the MTD and CoMSIA approaches on the series of molecules given in Table I.

CONCLUSIONS

QSAR studies enable the identification and removal of molecular structures with no therapeutic potential. The provided data allow structural features es-

sential for an increased biological activity to be identified, suggesting that there are certain structural requirements for a statin to have an increased biological potential. The present 3D-QSAR study modelled a series of 21 molecules of the tetrahydro-4-hydroxy-6-[2-(1*H*-pyrrol-1-yl)ethyl]-2*H*-pyran-2-one type against the HMG CoA-reductase enzyme by MTD and CoMSIA algorithms.

The 3D-QSAR MTD-CoMSIA study was performed in order to obtain a more detailed insight into the structure–activity relationships of HMG-CoA reductase inhibitors, beyond the usual features apparent from pharmacophore models. The observed biological activities of HMG CoA-reductase inhibitors were selected from a published report. Molecular modelling of the inhibitors was performed in Sybyl 7.2. The minimum potential energy for statins derivatives was calculated by the conjugate–gradient method, convergence of 0.01, Tripos force field. Finally, Gasteiger–Marsili partial charges were loaded. The partial least squares (PLS) algorithm within Sybyl 7.2 was used to evaluate the statistic parameters R^2_{CV}, *PRESS*, *SEE*, r^2, Fisher test (*F*). Statistically significant models were derived within both the MTD and CoMSIA models: while MTD allows a better geometrical description and control of the optimum hypermolecular 3D structure, CoMSIA provides better residuals for the identified template molecules. Moreover, it was demonstrated by the recently established Steiger test that both provide compatible biostatistics.

However, it is noteworthy that the present study enriches in some respect a previous report on the 3D-QSAR-CoMSIA approach for studying the inhibitory activity of statin moieties,[29] through performing molecular alignment by systematic MTD analysis, while imposing more restrictive convergence criteria on the gradient fields and yielding smoother better MTD-statistics.

While, generally, it is difficult to propose a clear conclusion regarding the steric configuration because of the diversity of the studied compounds, it appears that a better activity is predicted for substitutes that are more voluminous. The study showed that the relative activity in the studied series is given not only by the presence of the substituents on the nitrogen atom of the main ring, but also by the presence of 4-fluorophenyl and isopropyl substituents in positions 2 and 5 of the same ring, see Fig. 1. In addition, the activity is strongly influenced by the presence and the type of the substituents in positions 3 and 4 of the main ring: electrophilic atoms or groups of atoms determine a variation of the activity depending on the electrophilic potential and on the geometry (steric conformation) of the substitute.

It is believed that these results could be very useful for the development of new HMG-CoA reductase inhibitors belonging to statins classes, if good superposition (employing the MTD method) and minimization (employing the CoMSIA method) of the molecules belonging to the statistic model are achieved.

Acknowledgements. The authors are truly indebt to the anonymous referees for their constructive criticisms that considerably refined the initial arguments in the manuscript.

И З В О Д

MTD–CoMSIA МОДЕЛОВАЊЕ ИНХИБИТОРА HMG-CoA РЕДУКТАЗЕ

DANIEL M. DUDA-SEIMAN[1], SPERANŢA AVRAM[2], SILVIA MANCAŞ[1], VALENTIN CAREJA[3], CORINA DUDA-SEIMAN[4], MIHAI V. PUTZ[4] и DAN CIUBOTARIU[5]

[1]*Department of Medical Ambulatory, Medical Emergencies, University of Medicine and Pharmacy "Victor Babes", Timisoara,* [2]*Department of Physiology & Biophysics, University of Bucharest, Faculty of Biology, Bucharest,* [3]*Institute of Chemistry "Coriolan Drăgulescu" of the Romanian Academy, Timisoara,* [4]*Department of Chemistry, University of West Timisoara, Faculty of Chemistry-Biology-Geography, Timişoara и* [5]*Faculty of Pharmacy, University of Medicine and Pharmacy "Victor Babeş", Timisoara, Timisoara, Romania*

Испитиване су 3Д квантитативне релације између структуре и активности за серију инхибитора хидроксиметилглутарил-CoA (HMG-CoA)-редуктазе, базираној на основу пиролилетил-тетрахидропиранона. Радило се применом методе минималне тополошке разлике (*Minimal Topological Difference*, MTД) и анализом индекса компаративне молекулске сличности (CoMSIA). Проучавана су једињења тетрахидро-4-хидрокси-6-[2-(1*H*-пирол-1-ил)етил]--2*H*-пиран-2-она. У клиничкој пракси, инхибитори HMG-CoA-редуктазе се обично називају статинима. Анализа MTD методом показала је да волуминозни супституенти имају знатну биолошку активност (R^2_{CV} = 0,677 > 0,5; *SEECV* = 0,319), док метода CoMSIA даје корисне додатне информације о утицају на биолошку активност стерних, електростатичких и хидрофобних особина, као и донорских и акцепторских утицаја водоничне везе (R^2_{CV} = 0,60; r^2 = = 0.98).

(Примљено 1. јуна, ревидирано 23. августа 2010)

REFERENCES

1. R. Y. A. Mukhtar, J. Reid, J. P. D. Reckless, *Int. J. Clin. Practice* **59** (2005) 239
2. B. Genser, W. März, *Clin. Res. Cardiol*. **95** (2006) 393
3. D. J. Maron, S. Fazio, M. F. Linton, *Circulation* **101** (2000) 207
4. C. J. Vaughan, A. M. Gotto, *Circulation* **110** (2004) 886
5. C. L. Grines, *J. Interven. Cardiol*. **19** (2006) 3
6. C. I. Ardern, P. T. Katzmarzyk, I. Janssen, T. S. Church, S. N. Blair, *Circulation* **112** (2005) 1478
7. A. J. Taylor, S. M. Kent, P. J. Flaherty, C. L. Coyle, T. T. Markwood, M. N. Vernalis, *Circulation* **106** (2002) 2055
8. S. E. Nissen, S. J. Nicholls, I. Sipahi, P. Libby, J. S. Raichlen, C. M. Ballantyne, J. Davignon, R. Erbel, J. C. Fruchart, J. C. Tardif, P. Schoenhagen, T. Crowe, V. Cain, K. Wolski, M. Goormastic, E. M. Tuzcu, *J. Am. Med. Assoc*. **295** (2006) E1
9. J. A. Tobert, *Circulation* **76** (1987) 534
10. J. G. Robinson, B. Smith, N. Maheshwari, H. Schrott, *J. Am. Coll. Cardiol*. **46** (2005) 1855
11. J. C. LaRosa, *J. Am. Coll. Cardiol*. **46** (2005) 1863
12. A. J. Taylor, S. M. Kent, P. J. Flaherty, L. C. Coyle, T. T. Markwood, M. N. Vernalis, *Circulation* **106** (2002) 2055
13. A. R. Leach, *Molecular Modelling Principles and Applications*, 2nd ed., Pearson Prentice Hall, Harlow, 2001

14. Z. Simon, A. Chiriac, S. Holban, D. Ciubotariu, I. Mihalas, *Minimum Steric Difference. The MTD Method for QSAR Studies*, Wiley Res. Stud. Press, Letchworth, NY, 1984
15. G. Klebe, U. Abraham, T. Mietzner, *J. Med. Chem*. **34** (1994) 4130
16. I. Niculescu-Duvăz, D. Ciubotariu, Z. Simon, N. Voiculetz, in *Modeling of Cancer Genesis and Prevention*, N. Voiculetz, A. T. Balaban, I. Niculescu-Duvăz, Z. Simon, Eds., CRC Press, Boca Raton, FL, 1991, p. 157
17. D. Ciubotariu, V. Gogonea, M. Medeleanu, in *QSAPR/QSAR Studies by Molecular Descriptors*, M. V. Diudea, Ed., NOVA Science, Huntington, NY, 2001, p. 281
18. E. R. Malinowski, D. G. Howery, *Factor Analysis in Chemistry*, Wiley, New York, 1980
19. B. D. Roth, C. J. Blankley, A. W. Chucholowski, E. Ferguson, M. L. Hoefle, D. F. Orwine, R. S. Newton, C. S. Sekerke, D. R. Sliskovic, C. D. Stratton, M. W. Wilson, *J. Med. Chem*. **34** (1991) 357
20. M. E. Suh, S.-Y. Park, H.-J. Lee, *Bull. Korean Chem. Soc*. **23** (2002) 417
21. J. Boström, M. Böhm, K. Gundertofte, G. Klebe, *J. Chem. Inf. Comput. Sci*. **43** (2003) 1020
22. T. I. Oprea, in *Quantitative chemical structure–biological activity relationships. The MTD method*, A. Chiriac, D. Ciubotariu, Z. Simon, Eds., Mirton Publishing House, Timisoara, 1996, p. 221 (in Romanian)
23. M. Akamatsu, *Curr. Top. Med. Chem*. **2** (2002) 1381
24. S. J. Cho, A. Tropsha, *J. Med. Chem*. **38** (1995) 1060
25. *SYBYL Theory Manual*, version 6.4, Tripos Associates Inc., 1699 South Hanley Road, St. Louis, MO, 1988
26. C. Duda-Seiman, D. Duda-Seiman, D. Dragoş, M. Medeleanu, V. Careja, M. V. Putz, A. M. Lacrămă, A. Chiriac, R. Nuţiu, D. Ciubotariu, *Int. J. Mol. Sci*. **7** (2006) 537
27. C. Duda-Seiman, D. Duda-Seiman, M. V. Putz, D. Ciubotariu, *Digest. J. Nanomat. Biostruct*. **2** (2007) 207
28. J. H. Steiger, *Psychological Methods* **9** (2004) 164
29. R. Thilagavathi, R. Kumar, V. Aparna, M. E. Sobhia, B. Gopalakrishnan, A. K. Chakraborti, *Bioorg. Med. Chem. Lett.* **15** (2005) 1027.