

## 3D QSAR study of hypolipidemic asarones by comparative molecular surface analysis

Tomasz Magdziarz,<sup>a</sup> Bożena Łozowicka,<sup>b</sup> Rafał Gieleciak,<sup>a</sup> Andrzej Bąk,<sup>a</sup>  
Jarosław Polański<sup>a,\*</sup> and Zdzisław Chilmonczyk<sup>b,c</sup>

<sup>a</sup>Department of Organic Chemistry, Institute of Chemistry, University of Silesia, PL-40-006 Katowice, Poland

<sup>b</sup>Department of Organic Chemistry, Institute of Chemistry, University of Białystok, J. Piłsudskiego 11/4, PL-15-443 Białystok, Poland

<sup>c</sup>National Institute of Public Health, Chełmska 30/34, 00-725 Warszawa, Poland

Received 3 August 2005; revised 30 September 2005; accepted 6 October 2005

Available online 3 November 2005

**Abstract**—Three-dimensional quantitative structure–activity relationship (3D QSAR) modeled for  $\alpha$ -asarone derivatives using the comparative molecular surface analysis (CoMSA) allowed us to reveal a correlation between the activity of these compounds and the electrostatic potential at the molecular surface. The grid formalism (s-CoMSA) allowed us to indicate a pharmacophore that is of key importance for compound activity. The CoMSA formalism coupled with the iterative variable elimination method gives a highly predictive model.

© 2005 Elsevier Ltd. All rights reserved.

### 1. Introduction

Atherosclerosis and subsequent cardiovascular diseases still remain one of the major death risks in the industrialized and developing countries. The initiation of atherosclerosis is most likely caused by inflammatory responses, hyperlipidemia, and blood clotting factors. Many factors were found to be associated with the development of atherosclerosis and cardiovascular diseases.<sup>1–5</sup> The link between elevated cholesterol and coronary heart diseases (CHD) has been clearly established, and clinical trials have found that a 1% reduction in serum total cholesterol reduced CHD risk by 2%.<sup>6</sup>

Basically, cells (except for hepatic and ileum cells) do not synthesize cholesterol de novo but obtain it from blood and the cholesterol that accumulates in atherosclerotic lesions originates, primarily, in plasma lipoproteins.<sup>7</sup> The lowering of abnormally elevated levels of atherogenic lipoproteins (chylomicrons, very low density lipoproteins—VLDL-C, and low density lipoproteins—LDL-C) is now accepted as the first line of approach to the treatment of hyperlipidemic patients.<sup>8</sup>

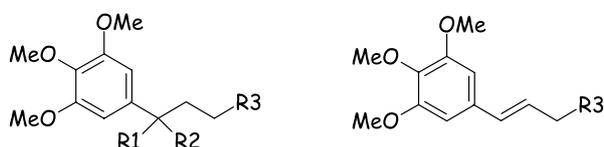
Increasing attention is also being focused on other lipoprotein fractions, such as high-density lipoprotein cholesterol (HDL-C) and triglycerides (TG), as additional potential targets of therapy. Elevated serum TG combined with low HDL-C, a condition often associated with smaller, dense LDL particles, is frequently referred to as atherogenic dyslipidemia or the ‘lipid triad.’<sup>9</sup>

At present, the therapy of lipid metabolism disorders is based on the use of drugs having hypolipidemic activity like fibric acid derivatives, anion exchange resins which sequester the bile acids, the inhibitors of 3-hydroxymethyl-coenzyme A (3-HMG Co-A) reductase, an enzyme involved in de novo sterol synthesis, probucol, lifestrol, and many others.<sup>10,11</sup>  $\alpha$ -Asarone (**1**) is an active component of *Acorus calamus* Linn, *Acorus gramineus* Soland, and *Guaeteria gaumeri*.<sup>12</sup>  $\alpha$ -Asarone and its analogues are known to be endowed with hypolipidemic activity<sup>13,14</sup> and have been the subject of many pharmacological,<sup>15–17</sup> toxicological,<sup>18,19</sup> synthetic,<sup>20</sup> and QSAR studies.<sup>21,22</sup>

Different  $\alpha$ -asarone structural features were found to influence the hypolipidemic activity. Chamorro et al. examined analogues possessing a dimethoxylated unconjugated propenyl side chain (mice) and Cruz et al.<sup>23</sup> analogues with saturated side chain (rats). Analysis of a length of a hydrophobic chain has shown that the most active analogues had the shortest side chain.

**Keywords:**  $\alpha$ -Asarone analogues; Hypolipidemic activity; QSAR study.

\* Corresponding author. Tel.: +48 32 3591197; fax: +48 32 2599978; e-mail addresses: [bozena@uwb.edu.pl](mailto:bozena@uwb.edu.pl); [polanski@us.edu.pl](mailto:polanski@us.edu.pl); [chilmon@il.waw.pl](mailto:chilmon@il.waw.pl)



$R_1, R_2=O$ ;  $R_1=H, R_2=OH$ ;  $R_3=amine$

Figure 1. Molecular formulae of compounds 7–13.

During our work on  $\alpha$ -asarone analogues, we synthesized several new compounds and examined their hypolipidemic activity.<sup>21,24–26</sup> We also examined a relationship between the hypolipidemic activity and molecular features. It appeared that the asarone activity could be described, at least in part, using the pseudo- or mini-receptor models.<sup>21,27</sup> However, we could not have indicated any clear molecular rule controlling the activity of these compounds.<sup>24–26</sup>

Some data concerning the compounds 1–40 were preliminarily reported<sup>27</sup> however, no experimental details are available for compounds 7–13. Thus, in the present publication we describe the synthesis and biological evaluation of compounds 7–13 (Fig. 1, Scheme 1).

We concentrate on the understanding of the structural basis of the compounds' pharmacological activity. Since our previous models did not completely explain the structure–activity relationship, we investigated 3D QSAR by comparative molecular surface analysis, a novel method described in our previous publications.<sup>28–36</sup> We have found that this method indicates a clear molecular basis for the activity.

## 2. Methods

### 2.1. Data sets for the analysis

The chemical structures of the  $\alpha$ -asarone derivatives are shown in Table 1. Hypolipidemic activity reported was measured (see Section 2.2, Table 1). As the activity data in all our QSAR calculation, we used the atherogenic index  $I_{TG/HDL}$ . It is calculated by the ratio of the triglyceride concentration (TG) [mmol/L] to the HDL

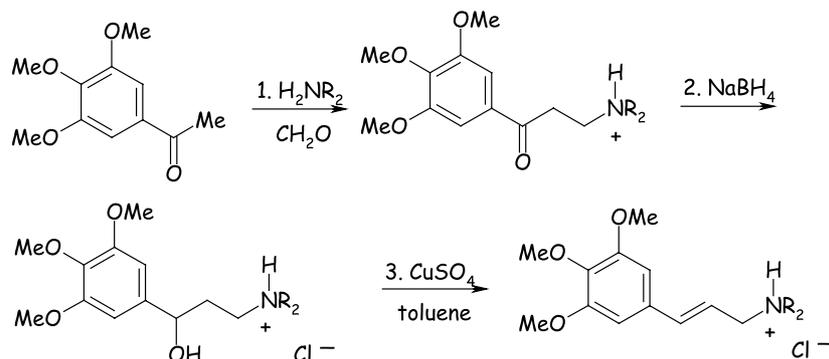
cholesterol concentration (HDL-C) [mmol/L]. The atherogenic index for the analyzed series ranges from 0.10 (high activity) to 3.28 (low activity) (Table 2).

### 2.2. Hypolipidemic activity

Wistar rats weighing 300–320 g male were bought from Nofer Institute of Occupational Medicine, Łódź, Poland. The animals were equally divided into groups of five animals each and maintained at temperature of  $22 \pm 2$  °C, 45–80% relative humidity, and every 12 h periods changing of light and darkness. All rats were fed a high cholesterol diet (Murigran enriched with cholesterol 1%, sodium cholate 0.2%, and olive oil 5%) for 7 days (Murigran-Łomna near Warsaw). Rats fed with laboratory chow for the same duration as above were used as noncholesterol control group. Compounds, diluted in oil, were administered through gastric intubation at 80 mg/kg once a day for the duration of the experiment. Group receiving clofibrate (150 mg/kg) served as positive control. Compounds, diluted in oil, were adjusted so that the rats were administered a volume of 5 mL/kg of body weight. Rats in the control group received a similar volume of vehicle. At the end of seven-day period, each animal was fasted for 16 h and anesthetized. Blood samples were collected through ocular puncture and centrifuged at 3000 rpm. Total cholesterol (TC), HDL-cholesterol (HDL-C), and triglycerides (TG) were determined using Alpha Diagnostics kits on Clinic System—700 Beckman. All the data were statistically analyzed by Student's *t* test. All the results were calculated against noncholesterol and cholesterol control groups.

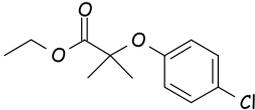
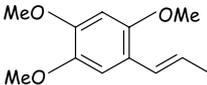
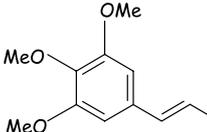
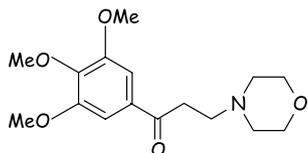
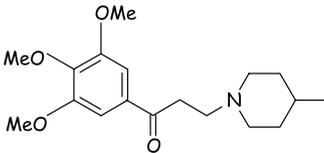
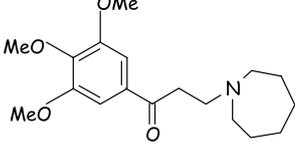
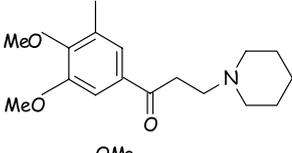
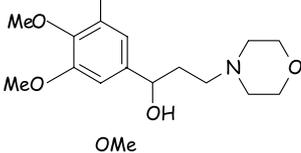
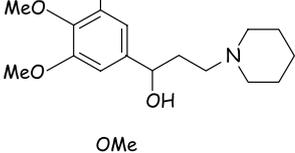
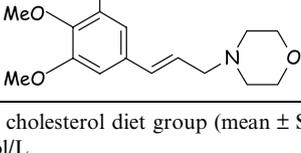
### 2.3. CoMFA analysis and molecular alignment

All modeling work was performed using the Sybyl 6.2 software package run on Silicon Graphics O2 workstation. The initial geometry was optimized using the standard Tripos force field (POWELL method) with 0.005 kcal/mol energy gradient convergence criterion and a distant-dependent dielectric constant. Charges were calculated using the Gasteiger–Marsilli method implemented in Sybyl. We used the FIT option of the Sybyl to align the compounds analyzed. Parent  $\alpha$ -asarone, that is, a common fragment for all molecules, was chosen as a template. Alignment was carried out



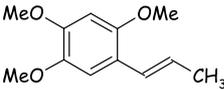
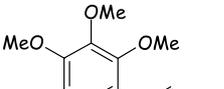
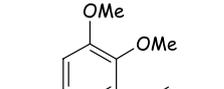
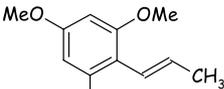
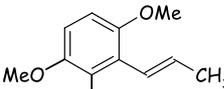
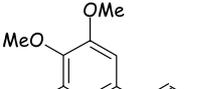
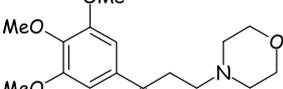
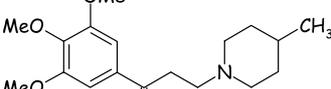
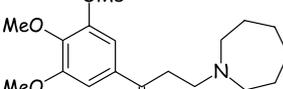
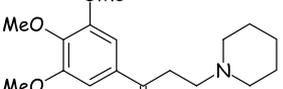
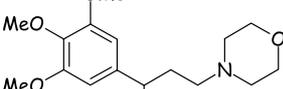
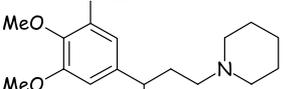
Scheme 1. General procedure synthesis of compounds 7–13.

**Table 1.** Structures of  $\alpha$ -asarone analogues and hypolipidemic activity<sup>a</sup>

| Entry                      | Structure   | TC                  | HDL-C                | LDL-C               | TG                  |
|----------------------------|---|---------------------|----------------------|---------------------|---------------------|
| Noncholesterol diet 7–13   | ...   | $-61 \pm 0.16^{**}$ | $+152 \pm 0.05^{**}$ | $-91 \pm 0.07^*$    | $-39 \pm 0.02^{**}$ |
| Cholesterol diet (CD) 7–13 | ...   | $100 \pm 0.06^b$    | $100 \pm 0.06^c$     | $100 \pm 0.33^d$    | $100 \pm 0.29^e$    |
| CLO + CD                   |    | $-23 \pm 0.14^{**}$ | $+24 \pm 0.06^{**}$  | $+7 \pm 0.14$       | $-27 \pm 0.12^{**}$ |
| 1 + CD                     |    | $+2.3 \pm 0.24$     | $+57 \pm 0.06^{**}$  | $-43 \pm 0.17^{**}$ | $+75 \pm 0.30^{**}$ |
| 6 + CD                     |    | $-16 \pm 0.28$      | $+56 \pm 0.06^{**}$  | $-54 \pm 0.14^{**}$ | $+8 \pm 0.16$       |
| 7 + CD                     |    | $+7 \pm 1.16$       | $+71.0 \pm 0.16^*$   | $-7.4 \pm 1.13$     | $+54 \pm 0.31$      |
| 8 + CD                     |   | $-10 \pm 0.84^{**}$ | $+35 \pm 0.07^{**}$  | $-15.6 \pm 0.88$    | $-16 \pm 0.19^{**}$ |
| 9 + CD                     |  | $+18 \pm 0.52^{**}$ | $+68 \pm 0.09^*$     | $+12 \pm 0.64$      | $+12 \pm 0.09^{**}$ |
| 10 + CD                    |  | $+6 \pm 0.58^{**}$  | $+65 \pm 0.11^*$     | $-5 \pm 0.50^{**}$  | $+26 \pm 0.18$      |
| 11 + CD                    |  | $0 \pm 1.59$        | $+36 \pm 0.09$       | $-2 \pm 1.62^*$     | $-23 \pm 0.18$      |
| 12                         |  | $+13 \pm 0.25^*$    | $+10 \pm 0.06^{**}$  | $+20 \pm 0.28^*$    | $-15 \pm 0.10$      |
| 13                         |  | $+13 \pm 0.38$      | $+26 \pm 0.09$       | $+12 \pm 0.54$      | $+16 \pm 0.33$      |

<sup>a</sup> Expressed as a percentage of the cholesterol diet group (mean  $\pm$  SD),  $n = 6$ .<sup>b</sup> Cholesterol diet group 3.07 mmol/L.<sup>c</sup> Cholesterol diet group 0.31 mmol/L.<sup>d</sup> Cholesterol diet group 2.44 mmol/L.<sup>e</sup> Cholesterol diet group 0.69 mmol/L; for compounds 7–13.\* Significantly different from the result for the cholesterol diet group control at  $p < 0.05$ .\*\* Significantly different from the result for the cholesterol diet group control at  $p < 0.01$ .

**Table 2.**  $\alpha$ -Asarone analogues and atherogenic index— $I_{TG/HDL}$  data

| Compound | Molecular formulae  | $I_{TG/HDL}$ | TG (mmol/L) | HDL-C (mmol/L) |
|----------|---|--------------|-------------|----------------|
| 1        |    | 2.10         | 2.48        | 1.18           |
| 2        |    | 1.27         | 1.41        | 1.11           |
| 3        |    | 1.56         | 1.58        | 1.01           |
| 4        |    | 1.34         | 1.31        | 0.98           |
| 5        |    | 2.35         | 1.83        | 0.78           |
| 6        |   | 1.31         | 1.53        | 1.17           |
| 7        |  | 2.00         | 1.06        | 0.53           |
| 8        |  | 1.38         | 0.58        | 0.42           |
| 9        |  | 1.48         | 0.77        | 0.52           |
| 10       |  | 1.71         | 0.87        | 0.51           |
| 11       |  | 1.26         | 0.53        | 0.42           |
| 12       |  | 1.74         | 0.59        | 0.34           |

(continued on next page)

Table 2 (continued)

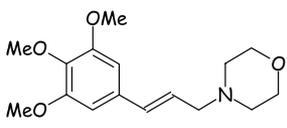
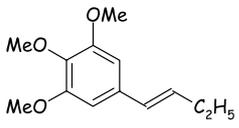
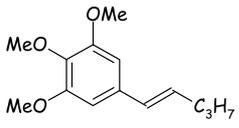
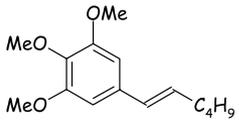
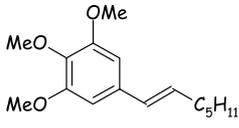
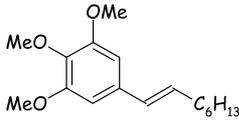
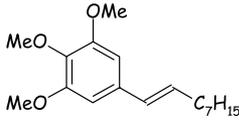
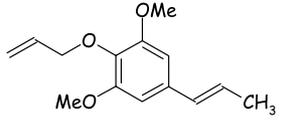
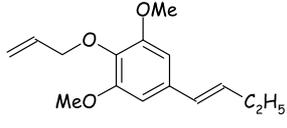
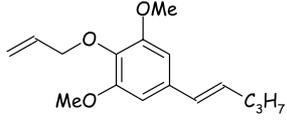
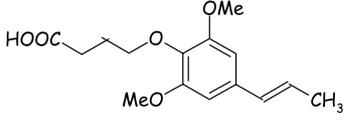
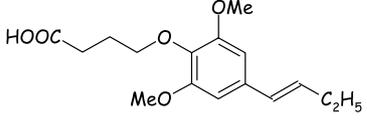
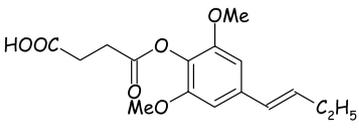
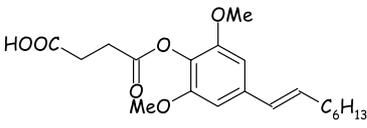
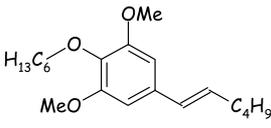
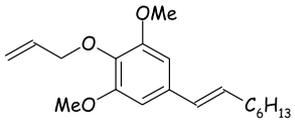
| Compound | Molecular formulae  | $I_{TG/HDL}$ | TG (mmol/L) | HDL-C (mmol/L) |
|----------|---|--------------|-------------|----------------|
| 13       |    | 2.05         | 0.8         | 0.39           |
| 14       |    | 0.36         | 0.21        | 0.59           |
| 15       |    | 0.45         | 0.24        | 0.53           |
| 16       |    | 0.16         | 0.09        | 0.57           |
| 17       |    | 0.27         | 0.19        | 0.71           |
| 18       |   | 0.13         | 0.07        | 0.56           |
| 19       |  | 0.14         | 0.08        | 0.57           |
| 20       |  | 0.47         | 0.2         | 0.43           |
| 21       |  | 0.15         | 0.1         | 0.66           |
| 22       |  | 0.10         | 0.07        | 0.73           |
| 23       |  | 0.45         | 0.29        | 0.64           |
| 24       |  | 0.22         | 0.14        | 0.64           |

Table 2 (continued)

| Compound | Molecular formulae | $I_{TG/HDL}$ | TG (mmol/L) | HDL-C (mmol/L) |
|----------|--------------------|--------------|-------------|----------------|
| 25       |                    | 0.13         | 0.09        | 0.69           |
| 26       |                    | 0.85         | 0.4         | 0.47           |
| 27       |                    | 0.77         | 0.33        | 0.43           |
| 28       |                    | 1.90         | 1.71        | 0.90           |
| 29       |                    | 3.15         | 1.04        | 0.33           |
| 30       |                    | 1.52         | 0.67        | 0.44           |
| 31       |                    | 1.79         | 0.61        | 0.34           |
| 32       |                    | 3.28         | 0.95        | 0.29           |
| 33       |                    | 0.45         | 0.21        | 0.47           |
| 34       |                    | 0.56         | 0.18        | 0.32           |
| 35       |                    | 1.98         | 0.99        | 0.5            |
| 36       |                    | 2.58         | 1.01        | 0.4            |

(continued on next page)

Table 2 (continued)

| Compound | Molecular formulae  | $I_{TG/HDL}$ | TG (mmol/L) | HDL-C (mmol/L) |
|----------|---|--------------|-------------|----------------|
| 37       |  | 1.92         | 0.96        | 0.5            |
| 38       |  | 1.96         | 1.06        | 0.54           |
| 39       |  | 1.28         | 0.77        | 0.6            |
| 40       |  | 0.80         | 0.43        | 0.54           |

by superimposing the atoms of benzene ring. We used a single conformation for each molecule, which provided an alignment illustrated in [Supplementary Materials Figure 1](#). The steric (Lennard-Jones) and electrostatic fields around the set of compounds were sampled with the probe atoms:  $sp^3$  carbon (charge +1 and 0) and hydrogen (charge +1), on the rectangular grid that encompasses all aligned molecules (with margin of 3.0–4.0 Å). For each molecule the energies of 1769 grid points were calculated with 2 Å spacing in a lattice of  $16 \times 10 \times 12$ . We kept a convention to truncate the steric and electrostatic values at the level of 30.0 kcal/mol.

#### 2.4. CoMSA analyses

For the calculation of shape descriptors we applied both grid (s-CoMSA) and neural formalisms described in our previous publications.<sup>28–36</sup> Thus, each 3D molecular representation is placed in its own virtual cubic grid and molecular surface is calculated, respectively. The electrostatic potential is calculated for the points randomly sampled on the molecular surface and a mean value of the electrostatic potential corresponding to the respective points found in each grid cell is used to describe this cell. Grid cells are unfolded into vectors and vectors describing all molecules of the series are aligned into a matrix. Grid cells that are empty for all molecules in the series analyzed are eliminated and the resulted matrix was used for further calculations using the PLS method. Alternatively, a CoMSA version with Kohonen self-organizing neural network (SOM-CoMSA) was used for comparison.

#### 2.5. PLS analysis

The obtained vectors were processed by the PLS analysis with a leave-one-out cross-validation procedure. The

PLS procedures were programmed within the MATLAB environment (MATLAB).<sup>37</sup>

A PLS model was constructed for the centered data and its complexity was estimated on the basis of the leave-one-out cross-validation procedure (CV). In the leave-one-out CV, one repeats the calibration  $m$  times, each time treating the  $i$ th left-out object as the prediction object. The dependent variable for each left-out object is calculated on the basis of the model with one, two, three, etc., factors.

The root mean square error of CV for the model with  $j$  factors is defined as

$$RMSECV = \sqrt{\frac{\sum(\text{obsd} - \text{pred})^2}{m}}$$

where obsd denotes the assayed value; pred is the predicted value of dependent variable, which ranges from 1 to  $m$ . Model with  $k$  factors, for which RMSECV reaches a minimum, is considered as an optimal one.

We used the performance metrics that are accepted and widely used in CoMFA analyses, that is, cross-validated  $q_{cv}^2$ ,  $s$ , RMS, and SDEP.

#### 2.6. Iterative variable elimination

In our previous publications, we have shown that uninformative variable elimination (UVE)<sup>38</sup> as well as its modifications, that is, modified UVE (m-UVE) and iterative variable elimination (IVE), can be used in 3D and 4D QSAR schemes.<sup>30,36,39</sup> This enables the identifications of the molecular areas important for the interactions with biological receptors or enzymes. In the current calculation, we used iterative variable elimination (IVE-PLS)<sup>30</sup> that is a modification of the UVE

algorithm based on the analysis of the regression coefficients calculated by the PLS method. PLS allows presenting the relation between the  $Y$  answer and  $X$  predictors in a form of

$$Y = Xb + e,$$

where  $b$  is a vector of the regression coefficients and  $e$  the vector of the errors.

Thus, the UVE algorithm analyzes the reliability of the mean( $b$ )/ $s(b)$  ratio (where  $s(b)$  means standard deviation of  $b$ ). Then, only the variables of the 'relative' high mean( $b$ )/ $s(b)$  ratio are included into the final PLS model. Instead of a single-step UVE procedure we used here an iterative algorithm based on the abs(mean( $b$ )/ $s(b)$ ) criterion to find the variables to be eliminated. This procedure includes:

1. Standard PLS analysis applied to analyze the matrices yielded from the s-CoMSA procedure with the leave-one-out cross-validation to estimate the performance of the PLS model ( $q_{cv}^2$ ),
2. The elimination of the matrix column of the lowest abs(mean( $b$ )/ $s(b)$ ) value,
3. Standard PLS analysis of the new matrix without the column eliminated in step 2,
4. Iterative repetition of steps 1–3 to maximize the LOO  $q_{cv}^2$  parameter.

All procedures were programmed within the MATLAB environment (MATLAB).

### 3. Chemistry

The compounds employed in this study **7–13** (Fig. 1) were synthesized according to Scheme 1. The compounds were prepared by a modified method of Blike and Burckhalter.<sup>40</sup> Thus, an appropriately substituted amine was reacted with three equivalents of paraformaldehyde in absolute ethanol with the addition of equivalent of concentrated HCl to form an intermediate iminium salt, which was then reacted with the substituted trimethoxyphenone to afford the desired product (Scheme 1). This reaction works very well for small amines. Although the compounds were obtained in good yields, the amount of product sharply decreases as the size of the amine becomes bigger. Usually, the reaction time ranges from 2 to 4 h; however; it needs to be increased with the increasing amine bulkiness. The best results were obtained using ethanol/HCl solution and the products formed can be efficiently purified by recrystallization from ethanol/acetone (9:1).

### 4. Results and discussion

The hypolipidemic activity of compounds **7–13** was tested on male rats fed with cholesterol diet against clofibrate as a reference drug. Compounds **7–13** exhibited diversified hypolipidemic activity in rats. The most active compound **8** elevated the HDL-C level by 35% and diminished TC, LDL-C, and TG levels by 10%,

15.6%, and 16%, respectively (as compared to clofibrate with +24%, –23%, +7%, and –27% respective levels). Moreover, compounds **11** and **12** elevated HDL-C while diminishing TG levels.

The description of the molecular shape by spatial sectors was originally proposed by Purcel and Testa, and further improved by Motoc.<sup>41</sup> In this method, a molecule is separated into partitions of the spatial regions either filled or unfilled by atoms or groups of atoms of certain volumes. Using a similar idea but with improved formalism, we have developed the CoMSA method and proved that it can be a powerful tool for 3D QSAR modeling. Hasegawa reported some further CoMSA modifications.<sup>42–44</sup>

The statistical results of several CoMSA analyses are summarized in Table 3. For the comparison, we performed also standard CoMFA calculations. Cross-validated  $q_{cv}^2$  values obtained using these methods range from 0.58 (CoMFA) to  $q_{cv}^2 = 0.60$  (SOM-CoMSA). This proves a correlation between descriptors and the hypolipidemic activity of the asarone series.

CoMFA calculations provide the best results for the H<sup>+</sup> atomic probe, which points that electrostatic interactions limit compound activity (Table 3, model 1). All steric fields provided unproductive models (the best model can be obtained for CH<sub>3</sub>(0) probe;  $q_{cv}^2 = 0.35$ ).

All further results are reported for the s-CoMSA method. Standard method provides a relatively low  $q_{cv}^2$  value of 0.53 (Table 3 model 2a). However, the exclusion of three compounds (**13**, **29**, and **32**) increases model quality to  $q_{cv}^2 = 0.69$  (s-CoMSA model—Table 3 model 2b). As the experience in 3D QSAR modeling indicates that  $q_{cv}^2$  alone cannot be a sufficient indicator of the model quality, we performed further model validation. The asarone series are split into two subseries. The compounds were sampled into these subseries by performing the Kennard–Stone calculations.<sup>45</sup> Then a respective model calculated for first subseries was used for the activity predictions in the second, test group and the SDEP error is calculated. The results and samplings are reported in Table 3. This indicates reasonable predictivity. Figure 2a–d illustrates the relationships between the cross-validated (predicted) and observed values of atherogenic index for the various models reported in Table 3.

Figure 3 (see also Supplementary Materials Figure 2) illustrates the surface areas of the key importance for the compound activity as indicated by s-CoMSA model 2a. We used that model because it includes all compounds. The respective color-coding allows us to identify the influence of the respective points sampled on the molecular surface by the combination of the electrostatic potential value and a value of the  $b$  weight in the PLS model. The variables contributing to the activity on a level close to 0 (near 90% of the points sampled) are omitted. Such an illustration suggests the key pharmacophore for the asarones investigated. Thus, an area near the central aromatic ring substituted with an alkoxy

**Table 3.** The results of CoMFA and CoMSA 3D QSAR modeling

| Compound  | Model                          | $q_{cv}^2$ <sup>a</sup> | $r^{2j}$ | $s$  | RMS            | SDEP              |
|-----------|--------------------------------|-------------------------|----------|------|----------------|-------------------|
| <b>1</b>  | CoMFA <sup>b</sup>             | 0.58                    | 0.71     | 0.57 | — <sup>f</sup> | — <sup>f</sup>    |
| <b>2a</b> | s-CoMSA                        | 0.53                    | 0.76     | 0.62 | 0.42           | — <sup>f</sup>    |
| <b>2b</b> | s-CoMSA <sup>c</sup>           | 0.69                    | 0.90     | 0.45 | 0.24           | — <sup>f</sup>    |
| <b>3a</b> | s-CoMSA-IVE-PLS <sup>g</sup>   | 0.92                    | 0.98     | 0.28 | 0.13           | — <sup>f</sup>    |
| <b>3b</b> | s-CoMSA-IVE-PLS <sup>c,h</sup> | 0.94                    | 0.97     | 0.20 | 0.13           | — <sup>f</sup>    |
| <b>4</b>  | s-CoMSA                        | 0.55                    | 0.92     | 0.54 | 0.19           | 0.80 <sup>d</sup> |
| <b>5</b>  | SOM-CoMSA                      | 0.60                    | 0.90     | 0.51 | 0.21           | 0.64 <sup>e</sup> |
| <b>6a</b> | s-CoMSA <sup>i</sup>           | 0.45                    | 0.85     | 0.71 | 0.34           | 0.49              |
| <b>6b</b> | s-CoMSA-IVE-PLS <sup>i</sup>   | 0.75                    | 0.86     | 0.48 | 0.33           | 0.30              |

<sup>a</sup> LOO cross-validated values; all compounds are included in the model.

<sup>b</sup> Calculation was performed with H(+1) as a probe atom.

<sup>c</sup> After the exclusion of three compounds: **13**, **29**, and **32**.

<sup>d</sup> Test set (K–S subset selection): **1:5**, **7**, **11**, **15**, **23**, **28**, **29**, **32**, **36**.

<sup>e</sup> Test set (K–S subset selection): **14:17**, **19**, **22**, **24**, **25**, **29**, **32**, **35**, **37**, **38**.

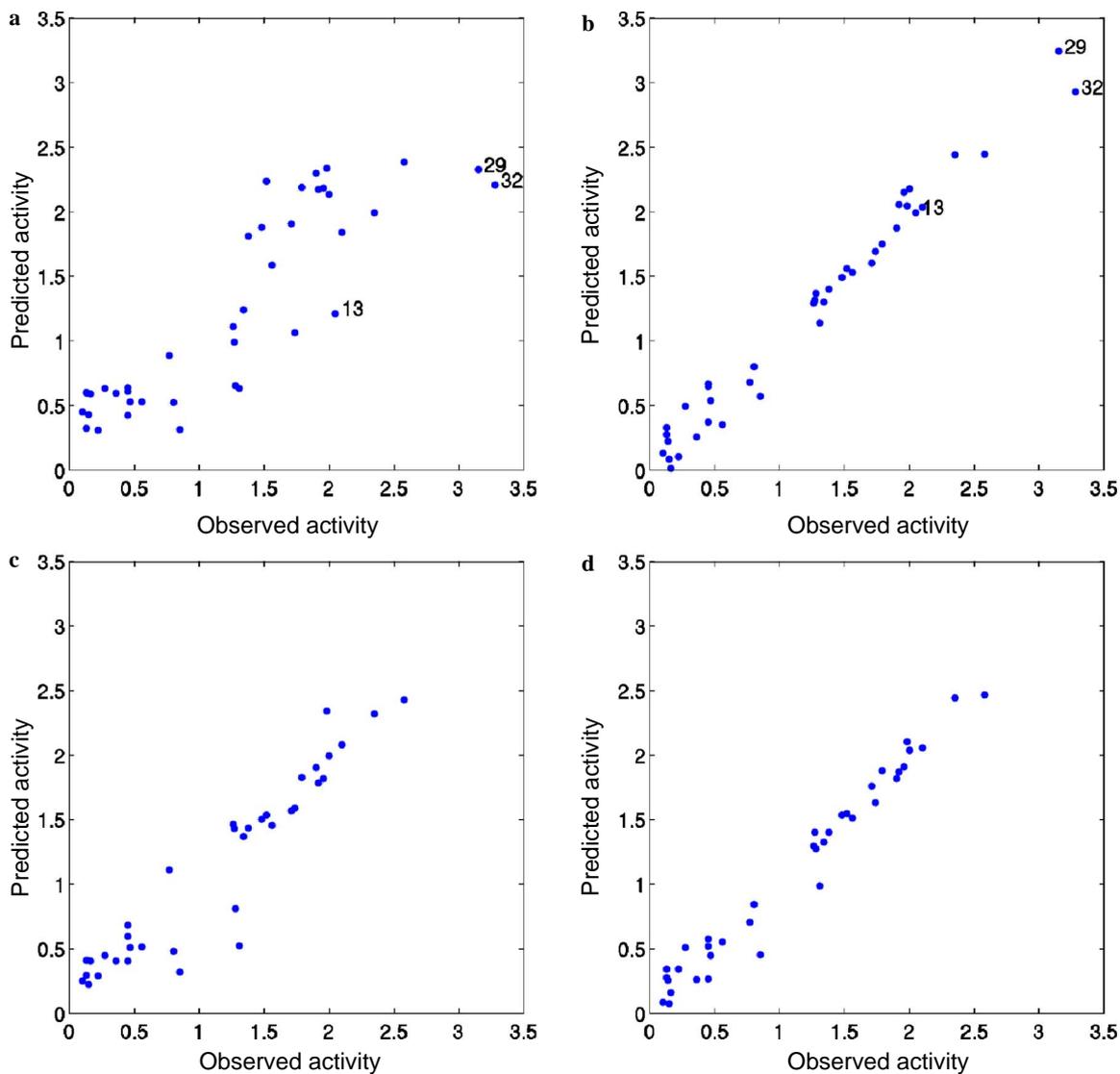
<sup>f</sup> Not tested.

<sup>g</sup> After IVE-PLS starting from model 2a.

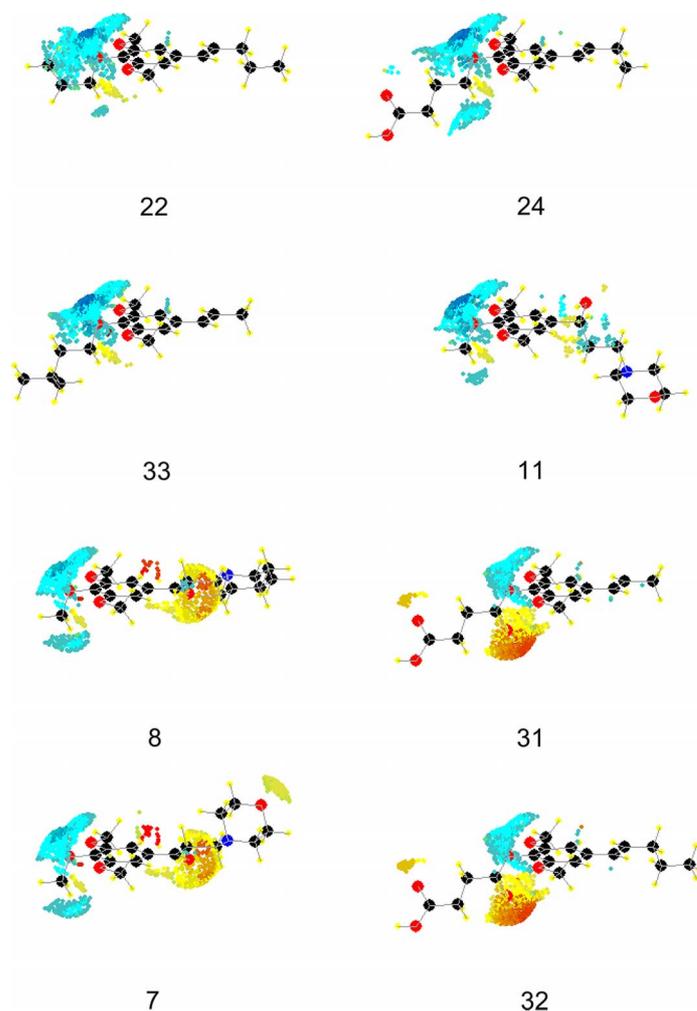
<sup>h</sup> After IVE-PLS starting from model 2b.

<sup>i</sup> Test set **33:40**.

<sup>j</sup> Fitted statistics.



**Figure 2.** The relationship between the cross-validated (predicted) and observed values of atherogenic index for the various models reported in Table 3; (a) model 2a, (b) model 3a, (c) model 2b and (d) model 3b.

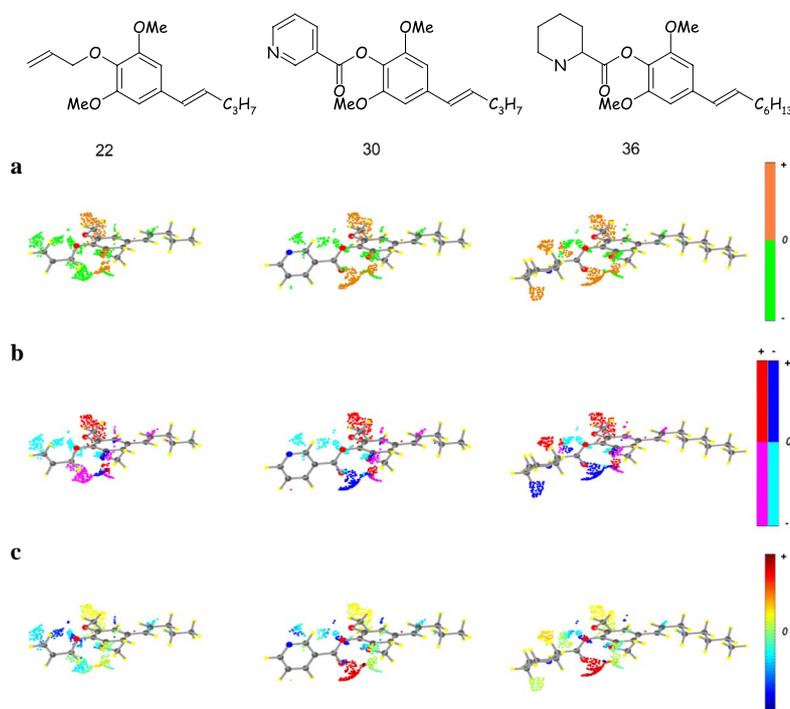


**Figure 3.** Molecular surface areas of the highest electrostatic contribution to the compound activity according to model 2b, sorted by the decreasing activity of some illustrative molecules (see also Supplementary Materials Figure 2). Blue indicates positive, and yellow and red negative influence on the activity. Details in text.

functionality provides a negative contribution, i.e., increases the activity (blue-colored sections), while a negatively charged carbonyl oxygen in the side chain generally decides a lower activity, clearly decreasing the activity as illustrated by the yellow and red molecular surface areas. This rule can be proved by the examination of the structures given in Table 2. It is worth noticing that if a hydroxyl function replaces a carbonyl oxygen, for example, compound **11**, we do not observe any activity decrease similar to that effected by a carbonyl-group, for example, compound **10**. Thus, our analysis reveals the molecular basis for the hypolipidemic activity of asarones. In Figure 4, we show the results obtained after application of additional filters. This allowed us to obtain highly predictive 3D QSARs (model 3a  $q_{cv}^2 = 0.92$ ,  $s = 0.28$ , RMS = 0.13 and 3b  $q_{cv}^2 = 0.94$ ,  $s = 0.20$ , RMS = 0.13, respectively); however it does not give so clear molecular illustration. Thus, in Figure 4 we give examples of such plots for compound, of high, medium, and low activity. In Figure 4a, the sectors contributing to the activity are divided into two subsets. Green sector increases the activity; orange—decreases. The plots are now more specific for individual compounds, for example, the contribution of the alliloxyl side

chain of compound **22** is pronounced more clearly than the areas closer to the aromatic unit. However, the key pharmacophore near the central aromatic unit and the disadvantageous influence of carbonyl oxygen are preserved. Moreover, in Figure 4b we coded by colors the different influences of the electro-negative and -positive surface sectors. This reveals the disadvantageous contribution of the electronegative carbonyl oxygen (dark blue) and electropositive methoxyl (red). The plots shown in Figure 4c allow for the estimation of the relative contribution of the sectors indicated. It is worth noticing that the contour interaction plots displayed in Figures 3 and 4 are completely different from that which resulted from CoMFA modeling. Unlike in CoMFA, each plot is characteristic for individual compound. This enables a clear differentiation of the low and high activity attributes. In particular, this provides also a novel insight into the molecular basis for the asarone activity.

Since overfitting is an important problem in QSAR modeling, we further validated the quality of the CoM-SA models by the so-called model randomization. We investigated this by generating 100 random permutations of the activity column. We used model **6a** as an

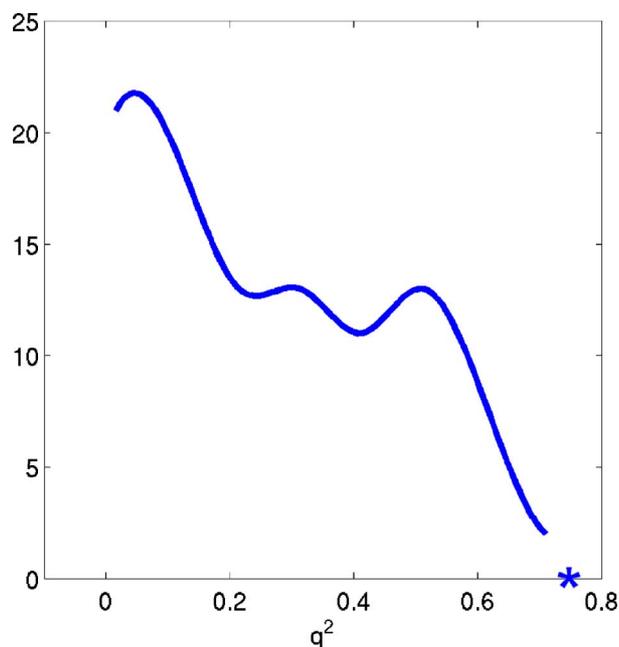


**Figure 4.** Molecular surface sectors indicated by s-CoMSA-IVE-PLS (model 3b from Table 2) for compounds of the highest, medium, and lowest activity; (a) color codes indicate a sign of activity) change: orange—decreases the activity, green—increases the activity, (b) combination of the electrostatic potential sign and the sign of the *b* weight in the model: +/+ (red, decreases the activity), +/- (cyan—increases the activity), +/- (magenta, increases the activity), -/+ (dark blue, decreases the activity), and (c) color codes indicate a relative value of activity change: warm color (red)—decreases the activity, cool colors (blue)—increases the activity, respectively.

initial form to perform IVE in such random models. Thus, each model with random activity column was processed by the IVE procedure and  $q_{cv}^2$  was measured. Figure 5 illustrates the histogram of calculated  $q_{cv}^2$ . Red asterisk points to the  $q_{cv}^2$  value of the unmodified model, that is, model 6b. Clearly, the majority of models have a relatively low  $q_{cv}^2$ , significantly lower than the value calculated for a model with actual compound activity. This validates the statistical significance of the model 6b where the activity is properly arranged according to the actual activity reported in Table 2.

## 5. Conclusions

Compounds 7–13 exhibited diversified hypolipidemic activity in rats. The most active compound 8 elevated HDL-C level by 35% and diminished TC, LDL-C, and TG levels by 10%, 15.6%, and 16%, respectively (as compared to clofibrate with +24%, -23%, +7%, and -27% respective levels). Although some pseudo- or mini-receptor models were reported for asarones previously, this did not reveal any clear molecular basis for controlling the activity of these compounds. The present analysis employing the CoMSA allowed us to reveal a correlation between the activity of these compounds and the electrostatic potential at the molecular surface. The grid formalism (s-CoMSA) allowed us also to indicate a pharmacophore that is of key importance for the compounds' activity. The CoMSA formalism coupled with the IVE (CoMSA-IVE) allowed us also to obtain highly predictive models.



**Figure 5.** The histogram of  $q_{cv}^2$  computed for 100 randomized CoMSA models (random activity arrangement). An asterisk indicates a  $q_{cv}^2$  value for the proper activity arrangement. Details in text.

## 6. Experimental

### 6.1. Chemical methods and reagents

Melting points were determined with a Köffler apparatus of the Bötius type and are uncorrected. <sup>1</sup>H and

$^{13}\text{C}$  NMR spectra were recorded on a Bruker AC 200F spectrometer using  $\text{CDCl}_3$  solution with TMS as internal standard (chemical shifts in  $\delta$  ppm). IR spectra were recorded on a Nicolet Magna 550 FTIR Spectrometer in chloroform solutions. The UV spectra were collected on a Hewlett-Packard UV-vis Diode Array Spectrophotometer 8452A. Mass spectra were obtained at 70 eV with an AMD-604 spectrometer. The reaction products were isolated by column chromatography or flash chromatography performed on a silica gel 70–230 mesh ASTM (Merck, Darmstadt, Germany). Thin-layer chromatograms were developed on aluminum TLC sheets precoated with silica gel F<sub>254</sub> (Merck, Darmstadt, Germany). The spots were visualized with 50% sulfuric acid after heating. All the solvents were dried and freshly distilled prior to use. Starting materials and reagents were purchased from Aldrich Chemical Co. (Steinheim, Germany).

## 6.2. General procedure for 3-amin-4-yl-1-(3,4,5-trimethoxyphenyl)propan-1-one (7–10)

The mixture of 0.046 mol of hydrochloride cycloamine, 2.0 g (0.066 mol) of paraformaldehyde, 2 g 3,4,5-trimethoxyacetophenone (0.040 mol), 2.4 mL concd HCl, and 20 mL absolute ethanol in a 100 mL three-necked flask was refluxed for 2 h. Then 1.2 g (0.040 mol) paraformaldehyde was added. The reaction mixture was refluxed for 5 h. To a hot mixture 50 ml acetone was added and refluxed for 15 min. After cooling down to room temperature, the reaction mixture was evaporated under reduced pressure to dryness. The product was crystallized from ethanol/acetone 9:1 (v/v).

## 6.3. 3-Morpholin-4-yl-1-(3,4,5-trimethoxyphenyl)propan-1-one (7)

Mp 199–201 °C; IR ( $\text{CHCl}_3$ ,  $\text{cm}^{-1}$ )  $\nu$ : 2700–2400; 1697; 1200; 1118; 1131;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 200 MHz)  $\delta$ : 11.3 (s, 1H); 7.2 (s, 2H); 4.2 (m, 2H); 4.1 (m, 2H); 3.9 (s, 9H); 3.7 (m, 2H); 3.4 (m, 4H); 2.9 (m, 2H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 50 MHz)  $\delta$ : 194 (C); 153 (C); 143 (C); 130 (C); 105 (CH); 63 ( $\text{CH}_2$ ); 60 ( $\text{CH}_3$ ); 56 ( $\text{CH}_3$ ); 52 ( $\text{CH}_2$ ); 32 ( $\text{CH}_2$ ); MS  $m/z$  100 ( $\text{M}^+$ , 100).

## 6.4. 3-(4-Methylpiperidin-1-yl)-1-(3,4,5-trimethoxyphenyl)propan-1-one (8)

Mp 153–155 °C; IR ( $\text{CHCl}_3$ ,  $\text{cm}^{-1}$ )  $\nu$ : 2700–2400; 1684, 1200, 1133;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 200 MHz)  $\delta$ : 12.2 (s, 1H); 7.3 (s, 2H); 3.9 (s, 6H); 3.8 (s, 3H); 3.7 (t, 2H); 3.5 (m, 4H); 2.7 (m, 2H); 1.8 (m, 5H); 1.1 (d, 2H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 50 MHz)  $\delta$ : 194 (C); 152 (C); 142 (C); 130 (C); 102 (CH); 66 ( $\text{CH}_2$ ); 60 ( $\text{CH}_3$ ); 56 ( $\text{CH}_3$ ); 52 ( $\text{CH}_2$ ); 32 ( $\text{CH}_2$ ); MS  $m/z$  112 ( $\text{M}^+$ , 100).

## 6.5. 3-Azepan-1-yl-1-(3,4,5-trimethoxyphenyl)propan-1-one (9)

Mp 171–173 °C; IR ( $\text{CHCl}_3$ ,  $\text{cm}^{-1}$ )  $\nu$ : 2700–2400; 1680, 1202, 1131, 1131;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 200 MHz)  $\delta$ : 12.2 (s, 1H); 7.2 (s, 2H); 3.9 (s, 6H); 3.8 (s, 3H); 3.7 (m, 2H); 3.6 (m, 4H); 2.9 (m, 2H); 1.8 (m, 4H); 1.6 (m,

2H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 50 MHz)  $\delta$ : 194 (C); 152 (C); 142 (C); 130 (C); 105 (CH); 52 ( $\text{CH}_2$ ); 60 ( $\text{CH}_3$ ); 56 ( $\text{CH}_3$ ); 45 ( $\text{CH}_2$ ); 33 ( $\text{CH}_2$ ); 26 ( $\text{CH}_2$ ); 24 ( $\text{CH}_2$ ); 22 ( $\text{CH}_2$ ); MS  $m/z$  112 ( $\text{M}^+$ , 100).

## 6.6. 3-Piperidin-1-yl-1-(3,4,5-trimethoxyphenyl)propan-1-one (10)

Mp 203–204 °C; IR ( $\text{CHCl}_3$ )  $\nu$  2700–2400; 1697, 1200, 1131  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 200 MHz)  $\delta$ : 11.4 (s, 1H); 7.2 (s, 2H); 3.9 (s, 6H); 3.8 (s, 3H); 3.7 (m, 2H) 3.6 (m, 4H); 2.6 (m, 2H); 2.1 (m, 2H); 2.1 (m, 2H); 1.7 (t, 2H); 1.5 (m, 2H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 50 MHz)  $\delta$ : 194 (C); 152 (C); 142 (C); 130 (C); 105 (CH); 53 ( $\text{CH}_2$ ); 60 ( $\text{CH}_3$ ); 56 ( $\text{CH}_3$ ); 32 ( $\text{CH}_2$ ); 22 ( $\text{CH}_2$ ); 21 ( $\text{CH}_2$ ); MS  $m/z$  112 ( $\text{M}^+$ , 100).

## 6.7. Synthesis of 3-morpholin-4-yl-1-(3,4,5-trimethoxyphenyl)propan-1-ol (11) 3-piperidin-1-yl-1-(3,4,5-trimethoxyphenyl)propan-1-ol (12)

3.5 g (0.010 mol) of hydrochloride 3-morpholin-4-yl-1-(3,4,5-trimethoxyphenyl)propan-1-one or 3-piperidin-1-yl-1-(3,4,5-trimethoxyphenyl)propan-1-one was dissolved in 2 ml water. After adding (20 mL) 6 M NaOH, the mixture was extracted three times with ether (3 × 30 mL). The ether extract was dried over  $\text{MgSO}_4$  and evaporated in vacuo. The crude product was dissolved in 30 mL methanol. The reaction mixture was dropped to a cold mixture of 440 mg (0.010 mol)  $\text{NaBH}_4$  dissolved in 20 mL methanol and 20 mL water. The reaction mixture was stirred for 2 h at room temperature and 15 min at 45 °C. The organic solvent was removed and 15 ml of 6 N NaOH was added. The aqueous solution was extracted with chloroform (3 × 30 mL). The organic extract were combined and washed with water and dried over  $\text{MgSO}_4$ . The mixture was filtered off and concentrated under a reduced pressure. After evaporation of solvent, the residue was purified by silica gel column chromatography (elution with hexane/ether; 1:1; v/v) and crystallized from an ethanol/acetone 9:1 (v/v). W = 63.5%.

## 6.8. 3-Piperidin-1-yl-1-(3,4,5-trimethoxyphenyl)propan-1-ol

Mp 65–68 °C; IR ( $\text{CHCl}_3$ ,  $\text{cm}^{-1}$ )  $\nu$ : 3181, 1227, 1132, 1112;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 200 Hz)  $\delta$ : 6.6 (s, 2H); 4.9 (t, 1H); 3.9 (s, 6H); 3.8 (s, 6H); 2.7 (m, 4H); 2.4 (m, 2H); 1.9 (m, 2H); 1.6 (m, 4H); 1.5 (m, 2H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 50 Hz)  $\delta$ : 153 (C); 139 (C); 136 (C); 106 (CH); 80 (CH); 54 ( $\text{CH}_2$ ); 52 ( $\text{CH}_2$ ); 33 ( $\text{CH}_2$ ); 22 ( $\text{CH}_2$ ); 60 ( $\text{CH}_3$ ); 57 ( $\text{CH}_3$ ); MS  $m/z$  98 ( $\text{M}^+$ , 100).

## 6.9. 3-Morpholin-4-yl-1-(3,4,5-trimethoxyphenyl)propan-1-ol

Mp 86–88 °C; IR ( $\text{CHCl}_3$ )  $\nu$ : 3190, 1227, 1132, 1119, 1127  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 200 MHz)  $\delta$ : 6.6 (s, 2H); 4.9 (t, 1H); 3.9 (s, 9H); 3.7 (t, 4H); 2.7 (m, 4H); 2.5 (m, 4H); 1.9 (m, 2H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 50 MHz)  $\delta$ : 154 (C); 140 (C); 138 (C); 102 (CH); 75 (CH); 68 ( $\text{CH}_2$ ); 61 ( $\text{CH}_3$ ); 58 ( $\text{CH}_2$ ); 57 ( $\text{CH}_3$ ); 53 ( $\text{CH}_2$ ); 33( $\text{CH}_2$ ); MS  $m/z$  100 ( $\text{M}^+$ , 100).

### 6.10. Synthesis of 4-[(2E)-3-(3,4,5-trimethoxyphenyl)prop-2-enyl]morpholine (13)

A solution of crude alcohol, 2.7 g (0.009 mol) 3-morpholin-4-yl-1-(3,4,5-trimethoxyphenyl)propan-1-ol in 50 ml of dry toluene was treated with 3 g 0.02 mol of anhyd CuSO<sub>4</sub>. The reaction mixture was heated for 2 h at reflux. After the evaporation of the solvent, the residue was crystallized from ethanol/acetone 9:1 (v/v).

Mp 85–86 °C; IR (CHCl<sub>3</sub>, cm<sup>-1</sup>) ν: 3181, 1227, 1132, 1112; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 200 Hz) δ: 6.86 (s, 2H); 6.1 (d, 1H, *J* = 15.8 Hz); 5.9 (dt, 1H, *J* = 6.6 Hz, *J* = 15.75 Hz); 3.85 (s, 6H); 3.8 (s, 6H); 3.38 (m, 4H); 2.2 (dt, 2H, *J* = 6.65 Hz, *J* = 7.04 Hz); 2.1 (m, 2H); <sup>13</sup>C NMR (CDCl<sub>3</sub>, 50 MHz) δ: 154 (C); 138 (C); 133 (CH); 128 (CH); 120 (C), 104 (CH); 69 (CH<sub>2</sub>); 66 (CH<sub>2</sub>); 60 (CH<sub>3</sub>); 56 (CH<sub>3</sub>); 53 (CH<sub>2</sub>).

### Acknowledgments

The authors thank Professor Johann Gasteiger of the University of Erlangen-Nürnberg, Germany, both for facilitating access to the CORINA, PETRA, SURFACE, and KMAP programs. The financial support of the KBN Warsaw under Grants No. KBN P05F 01617, 4P05F019 19, KBN 4T09A 088 25, and PBZ 040 P04/08 is gratefully acknowledged. R.G. thanks the Foundation for Polish Science for an individual grant.

### Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.bmc.2005.10.014](https://doi.org/10.1016/j.bmc.2005.10.014).

### References and notes

- Anderson, K. M.; Wilson, P. W. F.; Odell, P. M.; Kannel, W. B. *Circulation* **1991**, *83*, 356.
- Van Gaal, L. F.; Hang, A.; Steijaert, M.; De Leeuw, I. *Int. J. Obes.* **1995**, *19*, S21.
- McCully, K. S. *Nat. Med.* **1996**, *2*, 386.
- Brown, B.; Zhao, X.; Sacco, D.; Albers, J. J. *Circulation* **1993**, *87*, 1781.
- Badimon, J. J.; Fuster, V.; Badimon, L. *Circulation* **1992**, *86*, II-86–III-94.
- Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) *JAMA* **2001**, *285*, 2486.
- Chilmonczyk, Z.; Siluk, D.; Kaliszan, R. *Exp. Opin. Ther. Pat.* **2001**, *11*, 1301.
- Levine, G.; Keaney, J.; Vita, J. N. *Engl. J. Med.* **1995**, *332*, 512.
- Szapary, P. O.; Rader, D. J. *Am. Heart. J.* **2004**, *148*, 211.
- Eghdamian, B.; Ghose, K. *Drugs Today* **1988**, *35*, 79.
- Farmer, J. A.; Gotto, A., Jr. *Adv. Pharmacol.* **1966**, *35*, 79.
- Enriquez, R.; Chávez, M.; Jáuregui, F. *Phytochemistry* **1980**, *19*, 2024.
- Chamorro, G.; Garduno, L.; Sanchez, A.; Labarrios, F.; Salazar, M.; Martinez, E.; Diaz, F.; Tamariz, J. *Drug Dev. Res.* **1998**, *43*, 105.
- Labarrios, F.; Garduno, L.; Vidal, M.; Garcia, R.; Salazar, M.; Martinez, E.; Diaz, F.; Chamorro, G.; Tamariz, J. *J. Pharm. Pharmacol.* **1999**, *51*, 1.
- Dandiya, P. C.; Sharma, J. D. *Ind. J. Med. Res.* **1962**, *50*, 46.
- Dandiya, P. C.; Menon, M. K. *Br. J. Pharmacol.* **1963**, *20*, 436.
- Belova, L.; Alibekov, S.; Baginskaya, A.; Sokolov, S.; Pokrovskaya, G.; Stikhin, V.; Trumpe, T.; Gorodnyuk, T. *Farmak. Toksikol.* **1985**, *48*, 17.
- Morales, R.; Madrigal, B.; Mercader, M.; Cassani, M.; Gonzalez, G.; Chamorro, G.; Salazar, M. *Mutat. Res.* **1992**, *279*, 269.
- Salazar, M.; Salazar, S.; Ulloa, V.; Mendoza, T.; Pages, N.; Chamorro, G. *J. Toxicol. Clin. Exp.* **1992**, *12*, 149.
- Diaz, F.; Contreras, L.; Flores, R.; Tamariz, J.; Labarrios, F.; Chamorro, G.; Munoz, H. *Org. Prep. Proc. Int.* **1992**, *24*, 78.
- Filipek, S.; Łozowicka, B. *Acta Pol. Pharm.* **2000**, *57*, 106.
- Cruz, M.; Salazar, M.; Garciafigueroa, Y.; Hernandez, D.; Diaz, F.; Chamorro, G.; Tamariz, J. *Drug Dev. Res.* **2003**, *60*, 186.
- Cruz, A.; Garduno, L.; Salazar, M.; Martinez, E.; Jimenez-Vazquez, H.; Diaz, F.; Chamorro, G.; Tamariz, J. *Arzneim.-Forsch./Drug Res.* **2001**, *51*, 535.
- Poplawski, J.; Łozowicka, B.; Dubis, A.; Lachowska, B.; Witkowski, S.; Siluk, D.; Petruszewicz, J.; Kaliszan, R.; Cybulski, J.; Chilmonczyk, Z.; Strzalkowska, M. *J. Med. Chem.* **2000**, *43*, 3671.
- Łozowicka, B.; Filipek, S.; Dubis, A. *Eur. J. Med. Chem.* (submitted).
- Łozowicka, B.; Poplawski, J.; Dubis, A.; Chilmonczyk, Z.; Cybulski, J.; Kita, K.; Kobes, S.; Filipek, S. *Drug Dev. Res.* (submitted).
- Chilmonczyk, Z.; Siluk, D.; Kaliszan, R.; Łozowicka, B.; Poplawski, J.; Filipek, S. *Pure Appl. Chem.* **2001**, *73*, 1445.
- Polanski, J.; Walczak, B. *Comp. Chem.* **2000**, *24*, 615.
- Polanski, J.; Gieleciak, R.; Bak, A. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 184.
- Polanski, J.; Gieleciak, R. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 656.
- Polanski, J.; Gieleciak, R.; Wyszomirski, M. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1754.
- Polanski, J.; Gieleciak, R.; Wyszomirski, M. *Dyes Pigments* **2004**, *62*, 63.
- Polanski, J.; Gieleciak, R. *Mol. Diversity* **2003**, *7*, 45.
- Polanski, J.; Gieleciak, R.; Bak, A. *Comb. Chem. High Throughput Screen.* **2004**, *7*, 793.
- Polanski, J.; Bak, A.; Gieleciak, R.; Magdziarz, T. *Molecules* **2004**, *9*, 1148.
- Polanski, J.; Gieleciak, R.; Magdziarz, T.; Bak, A. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1423.
- MATLAB 5.0 program. Available from: The Mathworks Inc., Natick, MA, USA, <http://www.mathworks.com>.
- Centner, V.; Massart, D. L.; de Noord, O. E.; de Jong, S.; Vandeginste, B. M. V.; Sterna, C. *Anal. Chim. Acta* **1996**, *330*, 1.
- Polanski, J.; Bak, A. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 208.
- Blicke, F.; Bruckhalter, J. *J. Am. Chem. Soc.* **1942**, *64*, 451.
- Testa, B.; Purcell, W. P. *Eur. J. Med. Chem.* **1978**, *13*, 509; Motoc, I. Molecular Shape Descriptors. In *Steric Effects in Drug Design*; Charton, M., Motoc, I., Eds.; Akademie: Berlin, 1983; pp 93–105.
- Hasegawa, K.; Matsuoka, S.; Arakawa, M.; Funatsu, K. *Comput. Chem.* **2002**, *26*, 583.

43. Hasegawa, K.; Morikami, K.; Shiratori, Y.; Ohtsuka, T.; Aoki, Y.; Shimma, N. *Chemometr. Intell. Lab. Syst.* **2003**, *69*, 51.
44. Hasegawa, K.; Matsuoka, S.; Arakawa, M.; Funatsu, K. *Comput. Biol. Chem.* **2003**, *27*, 381.
45. Kennard, R. W.; Stone, L. A. *Technometrics* **1969**, *11*, 137.