0026-895X/88/060852-11$02.00/0

MOLECULAR PHARMACOLOGY, 34:852-862

An Artificial Intelligence Approach to the Study of the Structural Moieties Relevant to Drug-Receptor Interactions in Aldose Reductase Inhibitors

GILLES KLOPMAN and ERDEM BUYUKBINGOL
Case Western Reserve University, Department of Chemistry, Cleveland, Ohio 44106
Received November 6, 1987; Accepted August 5, 1988.

SUMMARY

The computer-automated structure evaluation program has been used to study 482 compounds relevant to the inhibition of the aldose reductase enzyme. Major activating/inactivating fragments were generated *automatically*. The significance of these molecular descriptors with respect to the activity of the compounds is discussed.

The aldose reductase enzyme, AR, involved in the sorbitol pathway (Fig. 1), which is an important mechanism in the regulation of mammalian glucose metabolism, has been found to play a physiologically significant role in the initiation of diabetic complications (1, 2). Therefore, over the past few years, considerable efforts have been made by several research groups to design inhibitors of this enzyme. The enzyme is primarily found in the corneal epithelium, retina, optic nerve, placenta, brain, kidney, muscle, and sperm (3–5).

Using NADPH as a cofactor, the AR enzyme catalyzes the transformation of glucose into sorbitol, which, in diabetes, is believed to accumulate in certain tissues such as nerve, kidney, pancreas, retina, and lens. Increased concentration of sorbitol can cause damage to these tissues, leading to diabetic complications such as microangiopathy, nephopathy, neuropathy, retinopathy, cataractogenesis, and corneal epitheliopathy (1, 3). Tissue destruction and cell damage are also believed to result from a change in the osmotic process within the cell (2, 6). Under normal glycemic conditions, AR has very little affinity for glucose (7). Therefore, glucose is primarily converted by hexokinase (one of the enzymes of the sorbitol pathway) into substrates for the glycolytic pathway whereas only small amounts are actually converted to sorbitol (4, 7). This is not so under diabetic conditions, in which the latter process becomes predominant. Although the exact mechanism of the polyol pathway is still largely unknown, it is known that some AR enzyme inhibitors are apparently able to reduce the flux through the sorbitol pathway under pathophysiological conditions in animals as well as in humans (3, 8, 9).

Since the early discoveries of inhibitors, such as aliphatic fatty acids (10) and tetramethylene glutaric acid (11), many natural as well as synthetic compounds have been studied, under both experimental and clinical conditions, for the treatment of the aforementioned complications. One class of inhibitory compounds found to be effective is the flavonoids, which are derived from many natural sources and exhibit a broad range of bioactivity (5). Numerous analogs have been prepared in the hope of improving their pharmacological profile with respect to the inhibition of the AR enzyme (8, 12, 13). However, few of the flavonoids have sufficient activity to warrant further investigation. Their low water solubility as well as their inability to penetrate biological membranes such as the blood-retina barrier are additional deterrents to further evaluation of these compounds.

Numerous other molecules have been screened for potential activity as AR inhibitors. The most potent compounds were found to be mostly synthetic and of great structural diversity (3, 8, 14–35). Among these, AY-20,263 was found to be particularly effective under *in vivo* conditions; however, here again the low solubility prevented a more general use (1). The solubility problem was somewhat overcome by modifying the structure to give AY-22,284 (alrestatine) (14). Although solubility increased for this structure, oral administration was shown to result in an overall lower therapeutic efficacy. Another entry based upon previous results was tolrestat (15). This was followed by the development of several novel compounds including ICI-128,436, ONO-2235 (16, 17), and sorbinil (18); the latter was found to have substantial beneficial effects in the treatment of diabetic symptoms. In fact, many of the biochemical perturbations associated with diabetic complications, including changes in *myo*-inositol, glutathione, ATP, and Na-K-ATPase were somewhat reduced by the administration of sorbinil (36).

Our purpose in undertaking this study was to find whether some general guidelines in the design of AR inhibitors could be

Support by the Office of Naval Research through its Selected Research Opportunities Program (N00014-84-K-0090) and the Environmental Protection Agency is highly appreciated.

ABBREVIATIONS: AR, aldose reductase; CASE, computer-automated structure evaluation; QSAR, quantitative structure-activity relationship.

obtained from a structure-activity study based on the diverse molecules known to have some inhibitory activity (Fig. 2). This was made possible by the availability of the CASE program (37), a computer program capable of automatically selecting the structural features required for optimal biological activity in a diverse database. This report discusses our efforts to determine the bioactive feature of the inhibitory compounds that may affect recognition and binding to the AR enzyme.

Methods

The method followed in the present study has been described in earlier publications (37–45). The general algorithm is as follows. 1) Construct a data base including about an equal number of active and inactive inhibitors of the AR enzyme. 2) Scale the activity values suitable to the CASE format. 3) Tabulate and analyze the fragments that are generated from the structure of the molecules of the data base. 4) Eliminate irrelevant data and determine causalities. 5) Select statistically significant fragments and derive a QSAR.

The CASE methodology involves the "fragmentation" of the chemical structure of the molecules of the data base into units of 3 to 10 heavy atoms, together with their associated hydrogens. Statistical analysis, based on a binomial distribution, results in a set of fragments believed to be related to activity. Biophores are derived from biologically active compounds whereas biophobes are derived from inactive ones (38, 39). After the data base has been analyzed, new molecular structures can be submitted to the program and evaluated for potential activity. Overall, the program acts as a 'learning machine' (40), and the data base can be updated when more data become available, thus providing a better base for prediction. The CASE logic has been successfully applied to a number of biological activities and in each case, molecular features relevant to the observed activity were obtained (37–45).

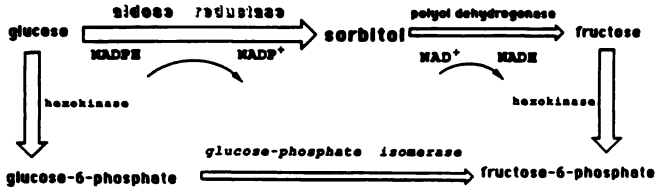


Fig. 1. Sorbitol pathway.

Results

In order to identify molecular features responsible for AR inhibitory activity, we have compiled a data base consisting of 482 compounds for which relevant data exists. Forty-eight compounds were randomly excluded from the data base and kept for later submission to the program as a test set. The remaining 434 compounds were coded for computer input and used as the training data base. All of the compounds included in the training and test sets, are listed in Table 1. The test compounds are marked with an asterisk. The code required to enter the molecules into the CASE program is the Klopman line notation code, which is similar to Wiswesser's notation but generally simpler to use and more appropriate for computer entry (39). The criterion used to judge the inhibitory activity was the IC_{50} value, which is the concentration of drug required for 50% inhibition of the rat lens AR enzyme activity.

Compounds were classified as inactive (−), marginally active (+), active (++), very active (+++), and extremely active (++++). The boundary between the inactive and active classes was chosen on the basis of the natural clustering of the experimental IC_{50} values. For the classification of data in terms of biological activity, compounds with IC_{50} values less than 6.0×10^{-6} M were listed as active; compounds with values between 6.0 and 10^{-6} M and 20.0×10^{-6} M were classified as marginally active, and those with values greater than 20.0×10^{-6} M were

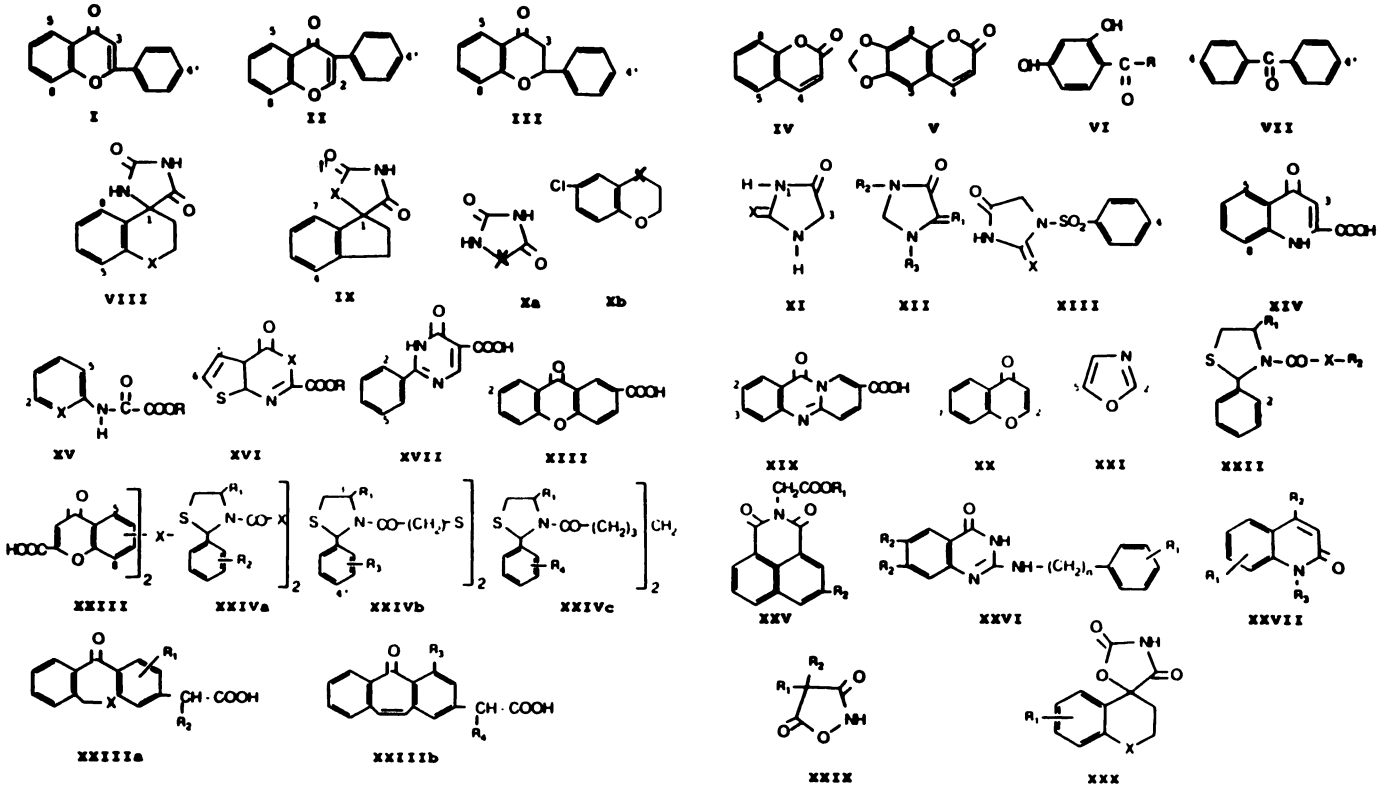


Fig. 2. Chemical structures used for the aldose reductase data base.

TABLE 1

Compounds considered for the CASE program evaluation of AR inhibition

*, Test compounds. OR, *O*-rhamnose; OG, *O*-glucose; OGa, *O*-galactose; OGR, *O*-glucose-rhamnose; OGA, *O*-glucose-apiose; OGaR, *O*-galactose-rhamnose; ORu, *O*-rutinose; OG2M, *O*-glucose-2-malonate; OR2Ac, *O*-rhamnose-2-acetate; ORG, *O*-rhamnose-glucose; A, OCH_2—CH(OH)—CH_2—O—; B, O—CH_2—CH(OH)—CH_2—O—; C, O—CH_2—CH(OH)—CH_2—O—; Me, methyl; Et, ethyl; Ph, phenyl. −, inactive, $IC_{50} > 20.0 \times 10^{-6}$; +, marginally active, $IC_{50} > 6.0 \times 10^{-6}$; ++, active, $IC_{50} > 2.0 \times 10^{-6}$; +++, very active, $IC_{50} > 0.6 \times 10^{-6}$; ++++, extremely active, $IC_{50} < 0.6 \times 10^{-6}$.

Structure I	Experimental	Calculated	Ref.	Structure I	Experimental	Calculated	Ref.
1. 5,7,4′,5′—OH,3,6—OMe	++++	++++	21	42. 3,5,7,3′,4′—OH	+++	++	12
2. 5,4′—OH,6,7—OMe	+++	+++	21	43. 3,5,7,2′,4′—OH	+	+	12
3. 5—3′,4′—OH,6,7—OMe	++++	++++	21	44. 5,7,3′,4′—OH,3—ORu	++	++	12
4. 5,3′,4′—OH,3,6,7—OMe	++++	++++	21	45. 5,7,3′,4′,OH,3—OG	+++	++++	12
5. 5,4′—OH,6,7,3′—OMe	+	++	21	46. 5,7,3′,4′—OH,3—OGa	+++	++++	12
6. 5—OH,6,7—OMe,4′—OG	−	−	21	47. 5,7,3′,4′—OH,3—OG2M	+++	++++	12
7. 5,3′—OH,6,7—OMe,4′—OG	+	+++	21	48. 5,7,3′,4′—OH,3—OR	++++	++++	12
8. 5—OH,6,7,3′—OMe,4′OG	++	−	21	49. 5,7,3′,4′—OH,3—OR2Ac	++++	++++	12
9. 3′,4′—OH,5,6,7—OMe,3—COMe	+	+	21	*50. 3,5,7,2′,3′,4′—OH	++	−	12
*10. 5,6,7—OH,8—OMe	+	−	21	51. 5,7,2′,3′,4′—OH,3—OR	++++	++++	12
11. 5,6,7,4′—OH,8—OMe	+++	+++	21	52. 3,5,7,3′,4′—OH,8—OG	+	+	12
12. 5,6,7,4′—OH,8,3′—OMe	+++	++	21	53. 3,5,3′,4′—OH,7—OEtOH	++	+	12
13. 8—OMe,5,6,7,3′,4′—OCOMe	++++	−	21	54. 3—ORG,5,3′,4′—OH,7—OEtOH	−	+	12
14. 5,6—OH,7,8—OMe	+	−	21	55. 3—ORG,5—OH,7,3′,4′—OEtOH	−	−	12
15. 5,6,4′—OH,7,8—OMe	+++	+++	21	56. 3—ORG,5,7,3′,4′—OEtOH	−	−	12
16. 5,6,4′—OH,7,8,3′—OMe	−	++	21	57. 5,7,4′—OH,3—ME	−	+	27
17. 5,7,4′—OH,6,8—OMe	++++	+++	21	58. 3,5,7—OH	−	−	27
18. 5,7—OH,6,8,4′—OMe	+	−	21	59. 3—OCOMe,7,3′,4′,—OSO_3H	++++	+++	8
19. 5,7,4′—OH,6,8,3′—OMe	++	++	21	*60. 6,7,8,3′,4′—OH	++++	++++	25
*20. 5,7—OH,6,8,3′,4′—OMe	+	−	21	61. 7,8,3′,4′—OH,6—OMe	++++	++++	25
21. 5,7—OH,6,8,3′—OMe,4′—OG	−	−	21	62. 7,3′,4′—OH,6,8—OMe	++++	++++	25
22. 5,4′—OH,6,8,3′—OMe,7—OG	+	++	21	63. 3′,4′—OH,6,7,8—OMe	++++	++++	25
23. 6—OH,5,7,8—OMe	+	−	21	64. 6,7,8,3′,4′—OH,3—OMe	++++	++++	25
24. 6,4′—OH,5,7,8—OMe	++++	++++	21	65. 7,8,3′,4′—OH,3,6—OMe	+++	++++	25
25. 6,3′,4′—OH,5,7,8—OMe	++++	++++	21	66. 7,3′,4′—OH,3,6,8—OMe	++++	++++	25
26. 6,4′—OH,5,7,8,3′—OMe	++	+++	21	67. 3′,4′—OH,3,6,7,8—OMe	++++	++++	25
27. 5,4′—OH,6,7,8—OMe	++++	+++	21	68. 6,8,3′,4′—OH,5—OMe	++++	++++	25
28. 5,4′—OH,6,7,8,3′—OMe	−	++	21	69. 5,8,3′,4′—OH,6—OMe	++++	++++	25
29. 4′—OH,5,6,7,8,3′—OMe	++++	++++	21	*70. 8,3′,4′—OH,5,6—OMe	++++	++++	25
*30. 4′—OH,5,6,7,8,3′—OMe	+	+++	21	71. 5,3′,4′—OH,6,8—OMe	++++	++++	25
31. 5,7,3′,4′—OH,6—OMe,3—OR	++++	++++	21	72. 6,3′,4′—OH,5,8—OMe	++++	++++	25
32. 5,7—OH	+	−	12	73. 3′,4′—OH,5,6,8—OMe	++++	++++	25
33. 5,7,4′—OH	++++	++++	12	74. 6,8,3′,4′—OH,3,5—OMe	++++	++++	25
34. 5,7—OH,4′—OMe	++	+	12	75. 6,3′,4′—OH,3,5,8—OMe	++++	++++	25
35. 5,4′—OH,7—OGA	+	+++	12	76. 3′,4′—OH,3,5,6,8—OMe	++++	++++	25
36. 3,5,7,4′—OH	+	+	12	77. 6,7,8,3′,4′—OH,5—OMe	++++	++++	25
37. 3,5,7—OH,4′—OMe	+	+	12	78. 7,8,3′,4′—OH,5,6—OMe	++++	++++	25
38. 5,4′—OH,3—OGaR,7—OR	+	−	12	79. 6,8,3′,4′—OH,5—CH_2Ph,7—OMe	++++	++++	25
39. 3,7,3′,4′—OH	+++	+++	12	*80. 8,3′,4′—OH,5,6,7—OMe	++++	++++	25
*40. 5,7,3′,4′—OH,8—OG	+++	++++	12	81. 3′,4′—OH,5,6,7,8—OMe	++++	++++	25
41. 5,7,3′,4′—OH	+++	++++	12	82. 6—COOH	++++	++++	13

Structure II	Experimental	Calculated	Ref.	Structure III	Experimental	Calculated	Ref.
83. 7—OH,6—OMe	−	−	21	88. 5—OH,7—ORG	−	−	12
84. 5,4′—OH,7,2′,5′—OMe	−	−	21	89. 5,4′—OH,7—ORG	+	−	12
85. 5,5′—OH,7,2′,4′—OMe	−	−	21	*90. 5,7,3′,4′—OH	++	−	12
86. 5,7,4′—OH	−	−	27	91. 3,5,7,3′,4′—OH	+	−	12
87. 5,7—OH,4′—OMe,2—COOEt	++	−	12	92. 5,7,3′—OH,4′—OMe	−	−	12
				93. 5,3′—OH,4′—OMe,7—ORu	−	−	12
				94. 5,7,4′—OH	−	−	27
				95. 3—OMe	+	−	27

Structure IV	Experimental	Calculated	Ref.	Structure IV	Experimental	Calculated	Ref.
96. 3—OH	−	−	21	102. 6,7—OH	+++	−	12
97. 3—OH,6—OMe	−	−	21	103. 6—OG,7—OH	−	−	12
98. 4—OH,3,7—OMe	−	−	21	104. 4—OH,7—OEtOPh	+++	+++	22
99. 4—Me,7—OMe	−	−	21	105. 4—OH,6—OPrOPh	++++	+++	22
*100. 3—Ph,4—OH,7—OMe	−	−	21	106. 4—OH,7—A	++	+++	22
101. 3,8—COOH,5—OMe	−	−	21	107. 3—NO_2,4—OH,7—OPrOPh	++	+	22

TABLE 1 Cont'd

Structure IV	Experimental	Calculated	Ref.	Structure IV	Experimental	Calculated	Ref.
108. 3—NO_2,4—OH—7—B	++++	++++	22	121. 4—CH_2COOH,7—OH	++++	+++	29
109. 3—NO_2,4—OH,7—C	++++	++++	22	122. 4—CH_2COOMe,7—OH	−	+	29
*110. 3—NO_2—4—OH,7—A	+++	++++	22	123. 4—CH_2CONH_2,7—OH	+	+	29
111. 6—OH,7—OMe	−	−	31	124. 4—CH_2COOH,7—OAc	−	+++	29
112. 4—Me	−	−	29	125. 4—CH_2COOH,7—OMe	++++	+++	29
113. 4—Me,6—OH	−	−	29	126. 4—COOH,7—OMe	+	−	29
114. 4—Me,7—OH	−	−	29	127. 4—CH=CHCOOH,7—OMe	++	−	29
115. 4—Me,7—OAc	−	−	29	128. 4—CH_2CH_2COOH,7—OMe	+++	++	29
116. 4—Me,7—Et	−	−	29	129. 4—CH_2COOH,7Cl	++++	+++	29
117. 4—Me,7—Cl	−	−	29	*130. 4—CH_2COOH,7—Et	++++	+++	29
118. 4—Me,5,7—OH	+	−	29	131. 4—CH_2COOH,5,7—OH	++++	+++	29
119. 4—Me,7,8—OH	+	−	29	132. 4—CH_2COOH,5,6— —$(CH_2)_4$—	++++	+++	29
*120. 4—CH_2COOH	++++	+++	29	133. 4—CH_2COOH,7,8— —$(CH_2)_4$—	++++	+++	29

Structure V	Experimental	Calculated	Ref.	Structure VI	Experimental	Calculated	Ref.
134. 4—OH	+	+++	21	*140. 3—Prydinyl	+	−	27
135. 4—Me	−	−	21	141. CH_2—CO—Ph	++	−	27
136. 3—CN	−	−	21	142. Cyclohexanyl	++	−	27
137. 3—COOH	−	−	21	143. $C(CH_3)_3$	+	−	27
138. 3—Ph,4—OH	−	−	21	144. CH_2—Cl	−	−	27
139. 3—Me,4—OH	−	−	21	145. $(CH_2)_6$—CH_3	−	−	27
				146. CH_3	−	−	27

Structure VII	Experimental	Calculated	Ref.	Structure VII	Experimental	Calculated	Ref.
147. 2,4—OH,4′—Me	++++	+++	27	153. 2,2′—OH,4,4′—OMe	−	−	27
148. 2,4,2′,4′—OH	++++	++++	27	154. 2—OH	+	−	27
149. 2,4,4′—OH	+++	++++	27	155. 3,5—Cl,4—OH	−	−	27
*150. 2,4,2′—OH,4′—OMe	+++	++	27	156. 4—OH,2′—COOH	−	−	27
151. 2,4,2′,3′,4′—OH	++	+++	27	157. 2′,6—OH	−	−	27
152. 2,4—OH	−	−	27	158. 3,4,6,3′,4′,5′—OH	++	−	28

Structure VIII	Experimental	Calculated	Ref.	Structure VIII	Experimental	Calculated	Ref.
159. X=CH_2	−	−	30	173. 7—Br X=O	++++	+++	30
*160. 5—OMe X=CH_2	+	−	30	174. 5—Cl X=O	+++	++	30
161. 6—OMe X=CH_2	−	−	30	175. 5,7—Cl X=O	++++	++++	30
162. 7—OMe X=CH_2	−	−	30	176. 6,7—Cl X=O	++++	+++	30
163. 6,7—OMe X=CH_2	−	−	30	177. 7—OMe X=O	++	−	30
164. 3—Me,7—F X=O	++++	+++	27	178. 7—Me X=O	++	−	30
165. 3—Me,5—NO_2,7—F X=O	++++	++++	27	179. 7—tBu X=O	−	−	30
166. 3—Me,5—NH_2,7—F X=O	++++	++++	27	*180. 7—SO_2NMe_2	−	−	30
167. 3—Me,7—F,3′—$(CH_2)_2$—$N(Me)_2$ X=O	−	−	27				
168. 3—Me,7—Cl X=O	++++	++++	27	181. X=NH	−	−	30
169. 3—Me,5—NH_2,7—Cl X=O	++++	++++	27	182. 6—F X=NH	−	−	30
*170. X=O	+	−	30	183. 6—OMe X=NH	+	−	30
171. 7—F X=O	++++	−	30	184. 5,6—OMe X=NH	+++	−	30
172. 7—Cl X=O	++++	+++	30	185. X=O	+	−	24

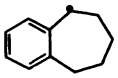
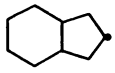
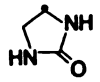
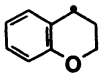
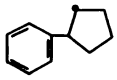
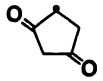
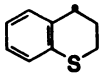
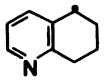
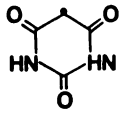
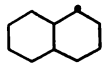
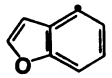
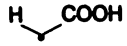
	Structure Xa	Experimental	Calculated	Ref.		Structure Xa	Experimental	Calculated	Ref.		Structure Xb	Experimental	Calculated	Ref.
186.		−	−	30	*190.		−	−	30	194.		+	+++	30
187.		+	−	30	191.		−	−	30	195.		−	−	30
188.		+	+	30	192.		+	−	30	196.		−	−	30
189.		−	−	30	193.		−	−	30	197.		++	+++	30

TABLE 1 Cont'd

Structure XI	Experimental	Calculated	Ref.
198. X═O	–	–	20
199. 3,3—Ph X═O	–	–	20
*200. 3═CBrPh X═O	–	–	20
201. 3═CClPh X═O	–	–	20
202. 1—Et,3—Ph X═O	–	–	20
203. 1—Ph,3—$(CH_2)_3$Me X═O	–	–	20
204. 1—COPh,3═CHPh X═O	–	–	20
205. 1—SO_2Ph,3═CHPh X═O	–	–	20
206. 1,4—COPh,3═CHPh X═O	–	–	20
207. 1,4—Ph,3═$CH(CH_2)_2$Me X═O	–	–	20
208. 1,4—Ph,3—$(CH_2)_3$Me X═O	–	–	20
209. 1—Ph,3—$(CH_2)_3$Me,4—CH_2Ph X═O	–	–	20
*210. 3,4—Ph X═O	–	–	20
211. 1—Ph,4—CH_2Ph X═O	–	–	20
212. 1—CH_2Ph,4—N═CHPh X═O	–	–	20
213. 1,4—COPh X═O	–	–	20
214. 4—COPh X═O	–	–	20
215. 4—SO_2Ph X═O	+++	+++	20
216. X═S	–	–	20
217. 3═CHPh—p—Cl X═S	–	–	20
218. 3═CHPh—p—OMe X═S	–	–	20
219. 1—Ph,3═CHPh X═S	–	–	20
*220. 1,4—Ph,3═CHPh X═S	–	–	20
221. 1,4—COPh,3═CHPh X═S	–	–	20
222. 1—Ph,3—$(CH_2)_3$Me X═S	–	–	20
223. 1—Ph X═S	–	–	20
224. 1,4—Ph X═S	–	–	20
225. 1,4—COPh X═S	–	–	20
226. 4—COPh X═S	–	–	20
227. 4—SO_2Ph X═S	+	+	20
228. 4—COMe X═S	–	–	20

Structure XII	Experimental	Calculated	Ref.
229. 3═CHpH—p—Cl,4—SMe	–	–	20
*230. 1—Me,3═CHPh—p—OMe,4—SMe	–	–	20
231. 1—SO_2Ph,3═CHPh,4—SMe	–	–	20
232. 1—COPh,3═CHPh,4—SEt	–	–	20
233. 1—SO_2Ph,3═CHPh,4—SEt	–	–	20
234. 1—Ph,3═CHPh,4—SCH_2COEt	–	–	20
235. 3═CHPh,4—SCH(Me)Et	–	–	20
236. 4—Me X═S	++	+	20
237. 4—Br X═S	++	+	20
238. 4—Cl X═S	++	+	20
239. 4—OMe X═S	+	+	20
*240. 2—NO_2 X═S	+++	++	20
241. 3—NO_2 X═S	++	++	20
242. 4—NO_2 X═S	+	++	20
243. 4—AcNH X═S	++	+	20
244. 4—Me X═O	+++	+++	20
245. 4—Br X═O	+++	+++	20
246. 4—Cl X═O	+++	+++	20
247. 4—OMe X═O	+++	+++	20
248. 2—NO_2 X═O	+++	++++	20
249. 3—NO_2 X═O	+++	++++	20
*250. 4—NO_2 X═O	+++	++++	20

Structure XIV	Experimental	Calculated	Ref.
251.	+	–	22
252. 6—COMe	+	–	22
253. 6—SPh	++	++	22
254. 6—SO_2Ph	–	–	22
255. 7—COMe	+	++	22
256. 7—SPh	+++	+++	22
257. 7—SO_2Ph	++	++	22
258. 8—COMe	+	–	22
259. 8—SPh	++	++	22
*260. 5—NHCOCOOH,8—Me	+	–	22
261. 5—CN,6—NHCOCOOH	+++	–	22
262. 7—NHCOCOOH,8—Me	++++	++	22

Structure XV	Experimental	Calculated	Ref.
263. 2—$COCON(Me)_2$ X═CH R═H	–	–	22
264. 2—NHCOCOOH X═N R═N	–	–	22
265. 2—NO_2,4—COOH,5—OH X═CH R═Et	++++	+	22
266. X═NH R═Et	–	–	22
267. 6—nC_6H_{13} X═NH R═H	+	–	22
268. 6—Et X═NH R═H	–	–	22
269. 5—Me,6—nC_6H_{13} X═NH R═H	++	–	22
*270. 5,6— —(CH_2)— X═NH R═H	+	–	22
271. 5—Me,6—Et X═NH R═H	+	–	22
272. 5—Me,6—COMe X═NH R═H	–	–	22
273. 6—nC_6H_{13} X═O R═Et	+	–	22

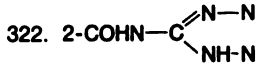
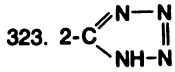
Structure XVII	Experimental	Calculated	Ref.
274.	+	–	22
275. 2—F	–	–	22
276. 2—OH	++	–	22
277. 2—OMe	–	–	22
278. 2—SMe	+	+	22
279. 2—OEt,5—SO_2NH_2	–	–	22
*280.	–	+	19
281. 2—SO_3H	+	–	19
282. 2—SO_2NH_2	–	–	19
283. 2—SO_2NHMe	+	–	19

Structure XVIII	Experimental	Calculated	Ref.
284. 2—$SO_2N(Me)_2$	++++	+++	19
285. 2—SO_2N(Me)Et	++	+++	19
286. 2—SO_2NH—i—Pr	–	–	19
287. 2—SO_2N(Me)—i—Pr	++	+++	19
288. 2—SO_2N(Me)—i—Pr	++	+++	19
289. 2—$SO_2NC_4H_8$	+	–	19
*290. 2—$SO_2NC_4H_8O$	–	–	19
291. 2—$SO_2NHCH_2CH_2OH$	+	–	19
292. 2—$SO_2N(Me)CH_2CH_2OH$	++++	+++	19
293. 2—$SO_2N(CH_2CH_2OH)_2$	+	+++	19
294. 2—SCH_2CH_2OH	+	–	19
295. 2—$SOCH_2CH_2OH$	+	–	19

Structure XVIII	Experimental	Calculated	Ref.
296. 2—$SO_2CH_2CH_2OH$	+++	–	19
297. 2—$SOCH_2CH_2OMe$	–	–	19
298. 2—SOMe	–	–	19
299. 2—CH(OH)Me	–	–	19
*300. 2—CH(OMe)Me	–	–	19
301. 2—$OCH_2CH(OH)CH_2OH$	+	–	22
302. 2—$OCH_2CH(OH)CH_2SMe$	–	–	22
303. 2—$OCH_2CH(OH)CH_2SOMe$	–	–	22
304. 2—$OCH_2CH(OH)CH_2SOPh$	–	–	22

Structure XIX	Experimental	Calculated	Ref.
305.	–	–	22
306. 2—OH	–	–	22
307. 2—COOH	+	–	22
308. 2—OMe	–	–	22
309. 2—OCH_2CH_2OH	–	–	22
*310. 2,3—OH	+	–	22
311. 3—OH	+++	++	22
312. 3—COOH	–	–	22
313. 3—OMe	–	–	22
314. 3—Me	–	–	22

TABLE 1 Cont'd

Structure XX	Experimental	Calculated	Ref.
315. 2—COOH	–	–	13
316. 2—COOH,6—Me	–	–	13
317. 2—COOH,6—Cl	–	–	13
318. 2—COOH,6—OPh	–	–	13
319. 2—COOH,7—Me	–	–	13
*320. 2—COOH,7—OH	++	++	13
321. 2—COOEt,7—OH	++	++	13
322. 	+	–	13
323. 	++	–	13
324. 2—COO—i—Pr,7—Oh	++	++	3
325. 2—COO—Bzl,7—OH	++	++	3

Structure XXI	Experimental	Calculated	Ref.
326. 2—COOEt,5—Me	–	–	26
327. 2—COOEt,5—Ph	–	–	26
328. 2—COOEt,5—pF—Ph	–	–	26
329. 2—COOEt,5—pCl—Ph	–	–	26
*330. 2—COOEt,5—pNO_2—Ph	–	–	26
331. 2—$CONHNH_2$,5—Me	–	–	26
332. 2—$CONHNH_2$,5—Ph	–	–	26
333. 2—NHCOOMe,5—Ph	–	–	26
334. 2—NHCOOEt,5—Me	–	–	26
335. 2—NHCOOEt,5—Ph	–	–	26
336. 2—$NHCOOCH_2$Ph,5—Me	–	–	26
337. 2—$NHCOOCH_2$Ph,5—Ph	+	–	26
338. 2—$NHCOOCH_2$Ph,5—pF—Ph	+	–	26
339. 2—$NHCOOCH_2$Ph, 5—pCl—Ph	–	–	26
*340. 2—$NHCOOCH_2$—pCl—Ph,5—pCL—Ph	–	–	26
341. 2—$NHCOOCH_2$—pCl—Ph,5—Ph	–	–	26
342. 2—$NHCOOCH_2$Ph,5—pNO_2—Ph	–	–	26
343. 2—NH_2,5—Ph	–	–	26
344. 2—NHCONHMe,5—Ph	–	–	26
345. 2—NHCONHPh,5—Ph	+	–	26
346. 2—$NHCOOCH_2$Ph,5—PhPh	–	–	26

Structure XXII		Experimental	Calculated	Ref.
R	X			
347. R_1,R_2—COOH 2—OH	$(CH_2)_4$	+	++++	28
348. R_1,R_2—COOH 2—OH	$(CH_2)_5$	++++	++++	28
349. R_1,R_2—COOH 2—OH	$(CH_2)_6$	++++	++++	28
*350. R_1,R_2—COOH 2—OH	$(CH_2)_7$	++++	++++	28
351. R_1,R_2—COOH 2—OH	$(CH_2)_8$	++++	++++	28
352. R_1,R_2—COOH 2—OH	$(CH_2)_2S(CH_2)_2$	++++	++++	28
353. R_1,R_2—COOH 2—OH	$(CH_2)_2O(CH_2)_2$	+++	++++	28
354. R_1,R_2—COOH 3—NO_2	$(CH_2)_7$	++++	++++	28
355. R_1,R_2—COOH 3—F	$(CH_2)_7$	+++	++++	28
356. R_1,R_2—COOH 2—Cl,5—NO_2	$(CH_2)_7$	+++	++++	28
357. R_1,R_2—COOH 2—OH	$(CH_2)_3$	–	++++	28
358. R_1,R_2—COOH 3—NO_2	$(CH_2)_7$	++++	++++	28
359. R_1—COOH 3—NO_2	$(CH_2)_7$	++++	++++	28
*360. R_1—COOH,R_2—CONHMe,3—NO_2	$(CH_2)_7$	++++	++++	28
361. R_1—COOH,, 2—OH $(CH_2)_2$ R_2—$S_2(CH_2)_2$—COOH		++++	++++	28
362. R_2—COOH, 3—NO_2	$(CH_2)_7$	++	+++	28

Structure XXIII,		Experimental	Calculated	Ref.
X				
363. 5—O—$CH_2CH(OH)CH_2$—O—		–	–	22
364. 6—O—$CH_2CH(OH)CH_2$—O—		+	–	22
365. 7—O—$CH_2CH(OH)CH_2$—O—		+	–	22

Structure XXIVa			Experimental	Calculated	Ref.
R_1	R_2	X			
366. COOH	2—OH	$(CH_2)_4$	+++	+++	28
367. COOH	2—OH	$(CH_2)_5$	++++	+++	28
368. COOH	2—OH	$(CH_2)_6$	++++	+++	28

Structure XXIVa			Experimental	Calculated	Ref.
R_1	R_2	X			
369. COOH	2—OH	$(CH_2)_7$	++++	+++	28
*370. COOH	2—OH	$(CH_2)_8$	++++	+++	28
371. COOH	2—OH	$(CH_2)_2S(CH_2)_2$	+++	+++	28

Structure XXIVa			Experimental	Calculated	Ref.
R_1	R_2	X			
372. COOH	2—OH	$(CH_2)_2O(CH_2)_2$	+++	+++	28
373. COOH	3—NO_2	$(CH_2)_6$	++++	++++	28
374. COOH	3—NO_2	$(CH_2)_7$	++++	++++	28
375. COOH	3—NO_2	$(CH_2)_8$	++++	++++	28

Structure XXIVa			Experimental	Calculated	Ref.
R_1	R_2	X			
376. COOH	3—F	$(CH_2)_7$	++++	++++	28
377. COOH	3—CN	$(CH_2)_7$	++++	++++	28
378. COOH	2—Cl,5—NO_2	$(CH_2)_7$	+++	++++	28

Structure XXIVb		Experimental	Calculated	Ref.
R_3				
379. COOH	2—OH	++++	+++	28
*380. COOH	2—NO_2	++++	++++	28

Structure XXIVc,		Experimental	Calculated	Ref.
R_4				
381. H	3—NO_2	–	+	28
382. COOMe	3—NO_2	++	++	28

Structure XXV		Experimental	Calculated	Ref.
R_1	R_2			
383. H	H	+++	++	8
384. H	N=C=S	++++	++	31
385. H	N_3	++	++	31
386. Me	N=C=S	+	–	31
387. Me	N_3	–	–	31

Structure XXVI			Experimental	Calculated	Ref.
R_1	R_2	n			
388. H	H	1	–	–	32
389. H	H	0	–	–	32
*390. H	4—OH	0	–	–	32
391. H	4—COOH	0	–	–	32
392. H	4—CO_3—H	0	–	–	32
393. H	3—OH, 4—COOH	0	–	–	32
394. OMe	H	1	–	–	32
395. OMe	H	0	–	–	32
396. OMe	4—COOH	0	–	–	32

TABLE 1 Cont'd

Structure XXVII						Structure XXVII					
R_1	R_2	R_3	Experimental	Calculated	Ref.	R_1	R_2	R_3	Experimental	Calculated	Ref.
397. H	Me	H	–	–	33	404. 8—OH	H	CH_2COOH	++	++	33
398. 6—OMe	H	H	–	–	33	405. 6—OH	H	CH_2COOH	++	++	33
399. 8—OCOMe	H	H	–	–	33	406. 6—NO_2	Me	CH_2COOH	++++	+++	33
*400. H	Me	Me	–	–	33	407. H	Me	CH(Me)COOH	–	–	33
401. H	Me	CH_2COOEt	–	–	33	408. 6—OMe	H	CH(Me)COOH	–	–	33
402. H	Me	CH_2COOH	++	++	33	409. 8—OH	H	CH(Me)COOH	–	–	33
403. 6—OMe	H	CH_2COOH	+++	++	33	*410. H	Me	CH_2CH_2COOH	–	++	33

Structure XXVIIIA			Experimental	Calculated	Ref.	Structure XXVIIIb		Experimental	Calculated	Ref.
R_1	R_2	X				R_3	R_4			
411. H	H	CH_2	+++	+++	34	*420. H	H	+++	+++	34
412. 1—OH	H	CH_2	++++	++++	34	421. H	Me	++	+++	34
413. 4—OH	H	CH_2	+++	++	34	422. OH	H	+++	++	34
414. 6—OH	H	CH_2	++	++	34	423. OH	Me	++	+	34
415. 7—OH	H	CH_2	+++	++	34					
416. 8—OH	H	CH_2	+++	++++	34					
417. H	H	O	++++	+++	34					
418. H	H	S	+++	+++	34					
419. 4—OH	Me	CH_2	++++	+	34					

Structure XXIX		Experimental	Calculated	Ref.	Structure XXIX		Experimental	Calculated	Ref.
R_1	R_2				R_1	R_2			
424. Me	4—AcNH—Ph	–	–	23	429. H	3—(CF_3)—Ph	–	–	23
425. H	2—Pyridyl	–	–	23	*430. H	CH_2CH_2—SO—Ph	–	–	23
426. H	3—Pyridyl	–	–	23	431. H	CH_2CH_2—SO_2—Ph	–	–	23
427. H	2—(OMe)—Ph	–	–	23	432. H	CH_2CH_2—O—Ph	–	–	23
428. H	3—(OMe)—Ph	–	–	23					

Structure XXX		Experimental	Calculated	Ref.	Alkaloids	Experimental	Calculated	Ref.
R_1	X							
433. H	O	+	–	24	443. Nandazurine	–	–	27
434. 6—Cl	O	++++	+++	24	444. Domesticine	–	–	27
435. 6—F	O	++++	+++	24	445. Nantenine	–	–	27
436. 6—Br	O	++++	++++	24	446. Berberine	–	–	27
437. 6,8—Cl_2	O	++++	++++	24	447. Palmatine	–	–	27
438. 6—Cl, 8—Me	O	++++	++++	24	448. Canadine	–	–	27
439. H	S	+++	+	24	449. Tetrahydropalmatine	–	–	27
*440. 6—Cl	S	++++	++++	24	*450. Papaverine	–	–	27
441. 6—F	S	+++	++++	24	451. Demethylpapaverine	–	–	27
442. H	CH_2	+	–	24				

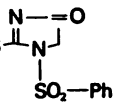
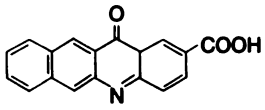
Miscellaneous	Experimental	Calculated	Ref.
452. Furo[3′,2′:6,7]isoflavone	–	–	21
453. 8,9—Dihydroxy—3—metoxy coumestan	–	–	21
454. Veneen	–	–	12
455. Hesperidine chalcone	–	–	12
456. Chlorogenic acid	++++	++++	12
457. D—Cathechin (2*R*:3*S*)—5,7,3′,4′—tetrahydroxy flavan—3—ol	–	–	12
458. *p*—OH—Ph—SO_2—Ph—p—OH	+	–	27
459. *p*—OH—Ph—S—Ph—p—OH	–	–	27
*460. 	–	–	20
461. 4—Oxo—4H—[1]—benzothiero[2,3*b*]pyran	++	+	13
462. 4—Oxo—4H—[1]—benzofuro[3,2*b*]pyran	–	–	13
463. 	–	–	22
464. 11—Oxo—11*H*—pyrido[2,1*b*]quinazoline—2—tetrazyl	++	–	22
465. Salicylic acid	–	–	34
466. Acetylsalicylic acid	–	–	8
467. Indomethacine	++	+++	34
468. Flufenamic acid	–	–	34
469. Ketoprofen	–	–	34

TABLE 1 Cont'd

Miscellaneous	Experimental	Calculated	Ref.
*470. Sulindac sulphide	+++	+++	8
471. Sulindac	+++	+++	8
472. Tetramethyleneglutaric acid	++++	++++	30
473. Tolrestat (AY – 27,773)	++++	++	15
474. ONO 2235	++++	++	17
475. ICI 128,436	+++	+++	16
476. WF 3681	+	+	35
477. 1,4,6,9—tetrahydro—4,6—dioxo—pyrido[3,2*g*]quinoline	+	–	3
478. 5—(3—Etoxy,4—butoxyphenyl)thiazolidine—2,4—dione	+	–	3
479. 1—Hydroxyh—10,11—dihydro—5—oxo—5H—dibenzo[*a*,*d*]cyclohepten—2—yl acetic acid lac-ton	+++	+++	34
*480. 2—Methyl—3—(4—carboxyphenyl)—4(3*H*)quinalinone	–	–	32
481. 2—Oxo—pyridine—1—acetic acid	–	++	33
482. 1—Naphtylacetic acid	–	+++	29

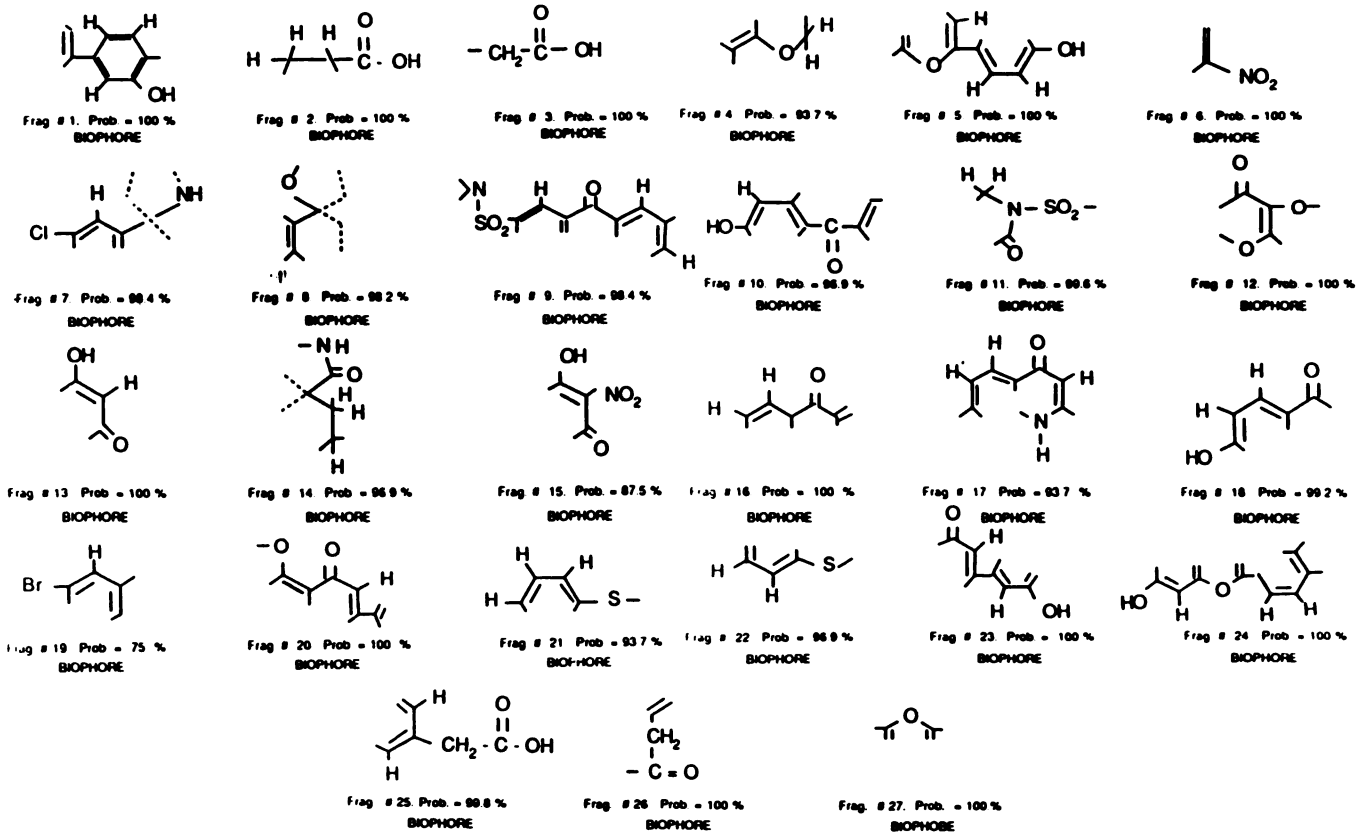


Fig. 3. Descriptors of aldose reductase inhibition generated by the CASE program. *Dotted lines* are added as needed to indicate content in which the fragments exist in the relevant molecules. *Dangling lines* indicate non-H substituents or group.

considered as inactive. The inactive cluster contained 184 compounds, the marginal cluster involved 65 compounds and the active one consisted of 185 compounds. Submission of the training data base to the CASE program resulted in the *automatic* generation of a total of 120 uncorrelated fragments. Of these 120 relevant fragments, 103 were found to be activating and 17 inactivating. The 120 fragments, together with the theoretically calculated log P values (45, 46) were then used as potential descriptors in a multivariate stepwise regression analysis. The procedure resulted in a regression equation incorporating the 26 biophores and 1 biophobe found to be the most relevant to the quantitative evaluation of aldose reductase inhibitory activity:

$$-\log_{10}(IC_{50}) = 4.57 + 0.23\,F_1 + 1.25\,F_2 + 1.08\,F_3 + 0.78\,F_4 + 0.33\,F_5 + 0.54\,F_6 + 1.43\,F_7 + 1.29\,F_8 + 1.53\,F_9 + 0.31\,F_{10} + 0.79\,F_{11} + 1.55\,F_{12} + 0.61\,F_{13} + 0.71\,F_{14} + 1.43\,F_{15} + 0.37\,F_{16} + 0.84\,F_{17} + 0.81\,F_{18} + 1.39\,F_{19} + 0.49\,F_{20} + 0.33\,F_{21} + 0.60\,F_{22} + 0.33\,F_{23} + 0.73\,F_{24} + 0.52\,F_{25} + 0.43\,F_{26} - 0.33\,F_{27}$$

Where $N = 434$, $R = 0.850$, $F_{(27,406)} = 42.07$, and $s = 0.59$.

Seventy-two percent of the variance within the data base was accounted for by the regression equation and the standard deviation of residuals was found to be 0.59. The activating fragments (biophores), nos. 7, 9, 12, 15, and 19, are found to be the most potent descriptors, as seen from the fact that their regression coefficients are the highest, i.e., 1.43, 1.53, 1.55, 1.43, and 1.39, respectively. The molecular fragments used by the QSAR equation are shown in Fig. 3, together with the probability that their presence is indeed relevant to AR inhibitory potency. The correlation is significant, at the 95% confidence

level, inasmuch as the F test is considerably better than needed by our criteria to eliminate chance correlations (47).

In order to evaluate the quality of the observed relationship, we submitted for analysis the test data set, consisting of the 48 compounds that were previously excluded from the training data base. The results of the predictions can be seen in Table 1. These compounds are preceded by an asterisk. Of the 21 active compounds, 18 were predicted correctly to be active whereas only 3 were misclassified as inactives. One of the marginal compounds was predicted to be active whereas the other 7 were classified as inactive. Of the 19 inactive compounds, only 1 was predicted to be active and 1 to be marginally active. The remaining 17 were found to be true inactives. These results seem to indicate that the predictive capability of the CASE program for identifying AR inhibitors is good. This is quite remarkable considering the diverse nature of the data base.

Discussion

Previous studies of the mechanism of AR inhibition have shown that certain electronic and steric requirements are necessary for biological activity. Most of these were based on computer modeling and molecular orbital calculations of a limited number of compounds (1, 48). It has been proposed that the relevant structural features of the AR enzyme inhibitors consist of two planar hydrophobic regions with hydrogen bonding substituents and an area containing a group such as a carbonyl or a thiocarbonyl, capable of undergoing reversible nucleophilic attack (3). It was further suggested that the amino acid residue of the AR enzyme active site responsible for the formation of a reversible tetrahedral intermediate is a tyrosine residue (3).

The proposal leading to tetrahedral intermediate formation is generally accepted, in spite of the fact that several alternative roles for the carbonyl group can be postulated. For example, it could be involved in a charge transfer bridge between an acidic group, such as the imine hydrogen of arginine or histidine, and a basic group, such as the hydroxyl of tyrosine at the enzyme inhibitory site (49). Alternatively, a transition state mechanism including the carbonyl group and an active site of the enzyme could be formed if the CONH— group is positioned appropriately (50). Whatever its function, however, the presence of a carbonyl group is one of the essential functional requirements for this type of drug-enzyme interaction (48). This view is supported in our study by the fact that a carbonyl functionality is found in many of the activating fragments identified by the CASE program, i.e., in 13 out of the 26 activating fragments. A more detailed analysis seems to show that the carbonyl group should preferentially be part of a conjugated ring system. Indeed, only one fragment (no. 10) includes a noncyclic carbonyl group, which is located between two aromatic rings. Considerable environmental variation exists around the carbonyl group in the other 12 fragments. For example, fragments 13 and 15 are similar. They are mostly found in 4-hydroxycoumarines. However, the regression coefficient corresponding to fragment 15, with a NO_2 substituent, is considerably larger than that of 13. This seems to confirm the hypothesis that electron-withdrawing groups α to the carbonyl carbon greatly enhance the activity. However, molecules containing fragment 12, consisting of an alkoxy group, are also active, nearly as active as the corresponding compounds containing fragment 15. It is therefore questionable whether the increase is related to an increased charge density on the carbonyl carbon, because the OR_1 group is usually electron donating. It does not seem to be linked to the bulkiness of the aliphatic group either, because the effect is observed whether R_1 is a methyl, methoxy, rhamnose, glucose, or a galactose group.

Fragments 11 and 14 both have a nitrogen atom adjacent to the carbonyl carbon and so have some degree of similarity. Their influence on the reactivity of the carbonyl group is probably identical and it is thus not surprising to find that their regression coefficient is roughly the same. When one evaluates the compounds that contain the —CONR— group, one can find that a decrease in activity is associated with N-alkylation. In contrast, electron-withdrawing groups attached to the nitrogen atom, as in fragment 11, increase the activity. This is consistent with the fact that substitution α of the carbonyl group by electron-withdrawing substituents may lead to optimal interaction for the formation of a tetrahedral intermediate at the inhibition site. The affinity of the inhibitor compound on the AR enzyme may be affected by changes in the aromatic substitution pattern. We find that to be active, molecules can possess halogen (nos. 7 and 19), hydroxyl (nos. 1, 5, 10, 18, 23, and 24), and nitro (no. 6) substituents on their aromatic rings. The presence of Cl and Br substituents generally results in improved inhibitory activity. This may be due to their ability to act as hydrogen-bonding acceptor and, when added at the proper location, may lead to more selective drugs; the preferred substitution location being in a meta position with respect to the attachment to the rest of the ring structure. The introduction of a nitro group into the aromatic ring system markedly increases the AR inhibitory activity. Fragment 6 is found in 42 compounds, of which 41 are active. In all compounds containing this fragment, the nitro group is again located at the *meta*-position. Somewhat similar evidence is observed with fragment 9. Here again, we find that the electron-withdrawing SO_2 group at the *meta*-position contributes to an enhanced activity, leading us to conclude that, in general, electron-withdrawing substituents at this particular position contribute to the inhibition of activity.

A hydroxyl group attached to an aromatic ring appears in six out of eight fragments. The ubiquitous occurrence of hydroxyl groups on the aromatic ring system of active molecules has previously been linked to its ability to hydrogen bond to the AR inhibitory site. The resulting conformational stabilization has been proposed as an explanation for the observed activity enhancement (3). As can be seen from the hydroxy-bearing fragments 10 and 18, when the carbonyl group is attached to the aromatic ring, the *para*-hydroxyl group acts to increase the negative charge on the carbonyl carbon atom through electron delocalization. This would probably decrease the binding activity of the carbonyl function and, hence, such fragments were not expected to be found in the activating fragment pool. However, *p*-hydroxy-bearing fragments do emerge as activating in the QSAR equation. Furthermore, the electron-delocalization effect cannot occur when a hydroxyl group is *meta* on the phenyl ring, as for example in fragment 23 (Fig. 4). Thus, one may conclude that the favorable effect of a hydroxyl group is probably not due to electron delocalization. On the other hand, if the role of the hydroxyl group is to hydrogen bond to the enzyme, then this observation would make sense. Indeed, the conjugation with the carbonyl group will enhance the lability of the hydroxyl proton and thus contribute to its ability to enter into such a hydrogen-bonding complex. Further support for this hypothesis comes from the observation that, when the hydroxyl groups are replaced by methoxy groups, substantial

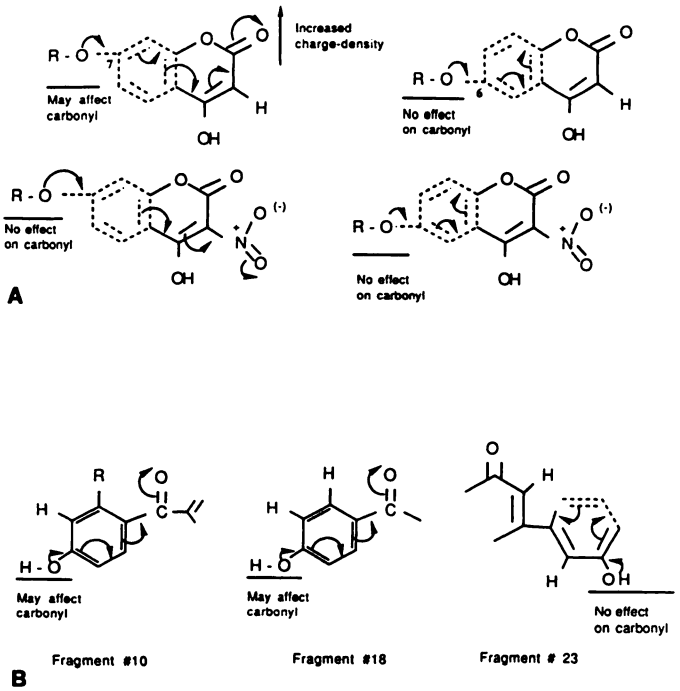


Fig. 4. A, Delocalization effect of the substituents on the carboxyl charge density. B, Delocalization effect of phenol group on the carbonyl function.

activity decreases are observed (21, 27, 48). However, the replacement of the hydroxyl group by a hydrogen does not necessarily decrease the activity, e.g., morphine versus codeine and 5-hydroxytryptemine versus 5-MeOT.

In previous studies of the AR inhibitor site, the prototype of inhibitory compounds was seen to consist of two parallel aromatic rings that can influence the binding to the AR enzyme (3, 48, 49). The results of the present study appear to be consistent with this hypothesis. In fragments 7, 9, 16, 17, and 18, the phenyl portion of the molecule is in a rigid structure whereas in fragments 1, 5, 10, 19, 21, 22, and 23 the aromatic ring is not constrained. In the case of fragment 24, one aromatic ring is constrained and the other is not. It is thus possible, but not proven, that the conformational flexibility of the compounds with respect to the spatial arrangement of the aromatic rings might lead to increased binding affinity.

One of the results obtained from the current program suggests that the presence of a carboxyl function is also important for the AR inhibitory activity. The carboxyl moiety, which is ionized at physiological pH, is preferably located in a side chain, as in fragments 2, 3, and 25. Previously it was proposed that the inhibitory potency of compounds containing acidic groups is not decreased with a decrease in charge-transfer potency (48). This can be explained by the fact that the esterified form of the carboxyl group behaves as a carbonyl functionality in the formation of the tetrahedral intermediate. To the best of our knowledge, there is not much information available at the present time concerning the manner of interaction of the carboxyl group on the enzyme, except that the acidity of the inhibitor drug may be involved in an interaction with arginine, lysine, and/or histidine residues at the cationic site of the AR enzyme (7). On the other hand, it is reported that the carboxyl group may be involved in an interaction with a tyrosine moiety at the substrate site (52). Although our results suggest that acidity is not the only factor involved, we believe that a polar group within the side chain can contribute to the activity. In this case, it is thought that the appropriately placed carboxylic group may be responsible for binding at the cationic arginine, lysine, or possibly histidine residues located near the binding cleft of the AR enzyme.

It has been shown before that certain antiallergy compounds, which have AR inhibitory activity, bear structural and electronic similarities to the AR inhibitory compounds (22). Many of the potent antiallergy agents are dicarboxylic acids (53). This may indicate that more than one of the amino acid residues may contribute to the binding of carboxylic acid moieties. Nevertheless, the existence of acidic fragments indicates that the incorporation of suitable substituents without masking of carboxyl groups may be the crucial step in the explanation of the interaction. From the relevant fragments, we find that chain lengths of one or two methylene units between the aromatic moiety and the carboxyl group are optimum for interaction. A more detailed knowledge of the effect of the spatial orientation of the carboxyl group on the AR enzyme may be of fundamental importance to the understanding of the reaction mechanism between inhibitors and the AR enzyme.

Another interesting factor is that a second carbonyl function is not essential for AR inhibitory activity although it appears in many of the inhibitory compounds. It may be that the second carbonyl group promotes receptor affinity by orientation of binding.

Only one inactivating biophobe, fragment 27, was identified by the QSAR equation. This fragment consists of an oxygen atom in a ring in which both α positions are substituted. This structure is to be contrasted to biophores 5 and 24, which contain the same structural moiety, except that the heterocyclic ring is part of a fused ring system. Finally, it is pertinent to note that the partition coefficient was not selected as a suitable descriptor for enzyme inhibition, in spite of the fact that it was available for selection during the stepwise selection of relevant descriptors.

Conclusion

A large number of potential inhibitors of AR enzyme were submitted to the CASE program in an effort to find structural features relevant to activity. Twenty-six biophores and one biophobe were found to be responsible for the overall inhibition of AR activity. In addition to determining the contribution of each individual descriptor to the AR inhibitory activity, we discuss possible modes of action on the basis of the descriptors generated by the CASE program.

The critical relationship between enzyme inhibitory activity and the structure of the inhibitors is believed to depend on the relative position of complementary regions within the inhibitor molecule. Although the CASE program does not give much information about potential restrictions on the spatial arrangement of the fragments, a critical examination of the context in which the fragments exist in the molecules of the data base can help evaluate the appropriateness of mechanistic hypothesis involving such spatial or stereochemical constrains.

The present study addressed the question of whether or not it is possible to organize and understand the pharmacophore requirements of AR enzyme inhibitors. It is apparent from this comparative evaluation that the desired AR inhibitory activity profile can be reliably predicted for a wide variety of chemical structures on the basis of a small number of activating and inactivating descriptors.

References

1. Kador, F. P., W. G. Robinson, Jr., and J. H. Kinoshita. The pharmacology of aldose reductase inhibitors. *Annu. Rev. Pharmacol. Toxicol.* **25**:691–714 (1985).
2. Kador, P. F., Y. Akagi, and J. H. Kinoshita. The effect of aldose reductase and its inhibition on sugar cataract formation. *Metabolism* **35**(Supp. 1):15–19 (1986).
3. Kador, F. P., J. H. Kinoshita, and N. E. Sharpless. Aldose reductase inhibitors: a potential new class of agents for the pharmacological control of certain diabetic complications. *J. Med. Chem.* **28**:841–849 (1985).
4. Lipinski, C. A., and N. J. Hutson. Aldose reductase inhibitors as a new approach to the treatment of diabetic complications. *Annu. Rep. Med. Chem.* **19**:169–177 (1984).
5. Varma, S. D. Inhibition of aldose reductase by flavonoids: possible attenuation of diabetic complications, in *Plant Flavonoids in Biology and Medicine: Biochemical, Pharmacological, and Structure-Activity Relationships*, Alan R. Liss, New York, 343–358 (1986).
6. Kinoshita, J. H. Mechanism initiating cataract formation. Proctor lecture. *Invest. Ophthalmol.* **13**:713–724 (1974).
7. Kinoshita, J. H., S. Fukushi, F. P. Kador, L. O. Menola. Aldose reductase in diabetic complications of the eye. *Metabolism*, **28**(Suppl. 1):462–469 (1979).
8. Chadury, P. S., J. Cabrera, H. R. Juliani, and S. D. Varma. Inhibition of human lens aldose reductase by flavonoids, sulindac and indomethacine. *Biochem. Pharmacol.* **32**:1995–1998 (1983).
9. O'Brien, M. M., P. J. Schofield, and M. R. Edwards. Inhibition of human brain aldose reductase and hexonate dehydrogenase by alrestatin and sorbinil. *J. Neurochem.* **39**:810–814 (1982).
10. Hayman, S., and J. H. Kinoshita. Isolation and properties of lens aldose reductase. *J. Biol. Chem.* **240**:872–882 (1965).
11. Jedziniak, J. A., and J. H. Kinoshita. Activators and inhibitors of aldose reductase. *Invest. Ophthalmol.* **10**:357–366 (1971).
12. Varma, S. D., and J. H. Kinoshita. Inhibition of lens aldose reductase by flavonoids: their possible role in the prevention of diabetic cataracts. *Biochem. Pharmacol.* **25**: 2505–2513 (1976).
13. Kador, P. F., and N. E. Sharpless. Structure-activity studies of aldose reductase inhibitors containing the 4-oxo-4*H*-chromen ring system. *Biophys. Chem.* **8**:81–85 (1978).
14. Gabbay, K. H., N. Spack, S. Lod, H. J. Hirsch, and A. Ackil. Aldose reductase inhibition: studies with alrestatin., *Metabolism.* **28**(Suppl. 1):471–476 (1979).
15. Sestanj, K., F. Bellini, S. Fung, N. Abraham, A. T. L. Humber, N. Simard-Duquesne, and D. Dvornik. N-[[5-trifluoromethyl)-6-methoxy-1-naphthanenyl]thiomethoxyl]-N-methylglycine (tolrestat), a potent, orally active aldose reductase inhibitor. *J. Med. Chem.* **27**:256–261 (1984).
16. Stribling, D., D. J. Mirrlees, H. E. Harrison, and D. C. N. Earl. Properties of ICI-128,436, a novel aldose reductase inhibitor, and its effects on diabetic complications in the rat. *Metabolism* **34**:336–344 (1985).
17. Terashima, H., K. Hama, R. Yamamoto, M. Tsuboshima, R. Kikkawa, I. Hatanaka, and Y. Shigeta. Effects of a new aldose reductase inhibitor on various tissues in vitro. *J. Pharmacol. Exp. Ther.* **229**:226–230 (1984).
18. Peterson, M. J., R. Sarges, C. E. Aldinger, and D. P. MacDonald. CP-45,634: a novel aldose reductase inhibitor that inhibits polyol pathway activity in diabetic and galactosemic rats. *Metabolism.* **28**(Suppl. 1):456–461 (1979).
19. Pfister, J. R., W. E. Wymann, J. M. Mahoney, and L. D. Waterbuty. Synthesis and aldose reductase inhibitory activity of 7-sulfamoylxanthone-2-carboxylic acids. *J. Med. Chem.* **23**:1264–1267 (1980).
20. Inagaki, K., I. Miwa, T. Yashiro, and J. Okuda. Inhibition of aldose reductases from rat and bovine lenses by hydantoin derivatives. *Chem. Pharm. Bull. (Tokyo)* **30**:3244–3254 (1982).
21. Okuda, J., I. Miwa, K. Inagaki, T. Horie, and M. Nakayama. Inhibition of aldose reductases from rat and bovine lenses by flavonoids. *Biochem. Pharmacol.* **31**:3807–3822 (1982).
22. Kador, P. F., N. E. Sharpless, and J. D. Goosey. Aldose reductase inhibition by anti-allergic drugs. *Enzymology of Carbonyl Metabolism: Aldehyde Dehydrogenase and Aldo/Keto Reductase.* Alan R. Liss, New York, 243–259 (1982).
23. Richon, A. B., M. E. Maragoudakis, and J. S. Wasvary. Isoxazolidine-3,5-diones as lens aldose reductase inhibitors. *J. Med. Chem.* **25**:745–747 (1982).
24. Schnur, R. C., R. Sarges, and M. J. Peterson. Spirooxazolidinedione aldose reductase inhibitors. *J. Med. Chem.* **25**:1451–1454 (1982).
25. Okuda, J., I. Miwa, K. Inagaki, T. Horie, and M. Nakayama. Inhibition of aldose reductase by 3′,4′-dihydroxyflavones. *Chem. Pharm. Bull.* **32**:767–772 (1984).
26. Tanimoto, T., H. Fukuda, J. Kawamura, M. Nakao, U. Shimada, A. Yamada, and C. Tanaka. Inhibition of aldose reductases from rabbit lens by oxazole derivatives. *Chem. Pharm. Bull. (Tokyo)* **32**:1032–1039 (1984).
27. Nakai, N., Y. Fujii, K. Kobashi, and K. Nomura. Aldose reductase inhibitors: flavonoids, alkaloids, acetophenones, benzophenones and spirohydantoins of chroman. *Arch. Biochem. Biophys.* **239**:491–496 (1985).
28. Kato, E., K. Yamamoto, T. Baba, T. Watanabe, Y. Kawahima, M. Horiuchi, M. Oya, T. Iso, and J. Iwao. Synthesis and aldose reductase inhibitory activity of *N*-acylthiazolidinecarboxylic acids. *Chem. Pharm. Bull. (Tokyo)* **33**:5341–5350 (1985).
29. Brubaker, A. N., J. Deruiter, and W. L. Whitmer. Synthesis and rat lens aldose reductase inhibitory activity of some benzypyran-2-ones. *J. Med. Chem.* **29**:1094–1099 (1986).
30. Sarges, R., and M. Peterson. Sorbinil: a member of the novel class of spirohydantoin aldose reductase inhibitors. *Metabolism* **35**(Suppl. 1):101–104 (1986).
31. Ares, J. J., P. F. Kador, and D. D. Miller. Synthesis and biological evaluation of irreversible inhibitors of aldose reductase. *J. Med. Chem.* **29**:2384–2389 (1986).
32. Deruiter, J., A. N. Brubaker, J. Millen, and T. N. Riley. Design and synthesis of 2-(arylamino)-4-(3H)-quinazolinones as novel inhibitors of rat lens aldose reductase. *J. Med. Chem.* **29**:627–629 (1986).
33. Deruiter, J., A. N. Brubaker, W. L. Whitmer, and J. J. Stein, Jr. Synthesis and aldose reductase inhibitory activity of substituted 2-oxoquinoline-1-acetic acid derivatives. *J. Med. Chem.* **89**:2024–2028 (1986).
34. Cerelli, M. J., D. L. Curtis, J. D. Dunn, P. H. Nelson, T. M. Peak, and L. D. Waterbury. Antiinflammatory and aldose reductase inhibitory activity of some trycyclic arylacetic acids. *J. Med. Chem.* **29**:2347–2351 (1986).
35. Uchida, I., Y. Itoh, T. Namiki, M. Nishikawa, and M. Hashimoto. Structure and synthesis of WF-3681, a novel aldose reductase inhibitor *Tetrahedron Lett.* **27**:2015–2018 (1986).
36. Yeh, L., C. E. Rafford, T. A. Beyer, and N. J. Hutson. Effects of the aldose reductase inhibitor sorbinil on the isolated cultured rat lens. *Metabolism* **35**(Suppl. 1):4–9 (1986).
37. Klopman, G. Artificial intelligence approach to structure-activity studies: computer automated structure evaluation of biological activity of organic molecules. *J. Am. Chem. Soc.* **106**:7315–7321 (1984).
38. Klopman, G., and R. Contreras. Use of artificial intelligence in structure-activity correlations of anticonvulsant drugs. *Mol. Pharmacol.* **27**:86–93 (1985).
39. Klopman, G., and M. MacGonigal. Computer simulation of physical-chemical properties of organic molecules. 1. Molecular system identification. *J. Chem. Inf. Comput. Sci.* **21**:48–52 (1981).
40. Klopman, G., M. R. Frierson, and H. S. Rosenkranz. Computer analysis of toxicological data bases: mutagenicity of aromatic amines in *Salmonella* tester strains. *Environ. Mutagen.* **7**:625–644 (1985).
41. Klopman, G., and A. N. Kalos. Quantitative structure-activity relationships of *beta*-adrenergic agents: applications of the computer automated structure evaluation (CASE) technique of molecular fragment recognition. *J. Theor. Biol.* **118**:199–214 (1986).
42. Klopman, G., and O. T. Macina. Computer automated structure evaluation of antileukemic 9-anilinoacridines. *Mol. Pharmacol.* **31**:457–476 (1987).
43. Klopman, G., and O. T. Macina. Use of the computer automated structure evaluation program in determining quantitative structure-activity relationships within hallucinogenic phenylalkylamines. *J. Theor. Biol.* **113**:637–648 (1985).
44. Klopman, G., O. T. Macina, E. J. Simon, and J. M. Hiller. Computer automated structure evaluation of opiate alkaloids. *J. Mol. Struct.* **134**:299–308 (1986).
45. Klopman, G., K. Namboodiri, and M. Schochet. Simple method of computing the partition coefficient. *J. Comput. Chem.* **6**:28–38 (1985).
46. Klopman, G., and R. E. Venegas. Case study of *in vivo* inhibition of sparteine monooxygenase. *Acta Pharm. Jugosl.* **36**:189–209 (1986).
47. Klopman, G., and N. A. Kalos. Causality in structure-activity studies. *J. Comput. Chem.* **6**:492–506 (1985).
48. Kador, P. F., and N. E. Sharpless. Pharmacophor requirements of the aldose reductase inhibitor site. *Mol. Pharmacol.* **24**:521–531 (1983).
49. Kador, P. F., J. H. Kinoshita, and N. E. Sharpless. The aldose reductase inhibitor site. *Metabolism* **35**(Suppl. 1):109–113 (1986).
50. Lienhard, G. E. Enzymatic catalysis and transition-state theory. *Science (Wash. D. C.)* **180**:149–154 (1973).
51. Doughty, C. C., S. M. Lee, S. Conrad, and S. Shade. Kinetic mechanism and structural properties of lens aldose reductase. *Prog. Clin. Biol. Res.* **114**:223–242 (1982).
52. Cairns, H., D. Cox, K. J. Gould, A. H. Ingall, and J. L. Suschitzky. New antiallergic prano[3,2-*g*]quinoline-2,8-dicarboxylic acids with potential for the topical treatment of asthma. *J. Med. Chem.* **28**:1832–1842 (1985).
53. Andrews, P. R., D. L. Craik, and J. L. Martin. Functional group contributions to drug-receptor interactions. *J. Med. Chem.* **27**:1648–1657 (1984).

Send reprint requests to: Gilles Klopman, Case Western University, Chemistry Department, Cleveland, OH 44106.