# A Genetic-Function-Approximation-Based QSAR Model for the Affinity of Arylpiperazines toward $\alpha_1$ Adrenoceptors

Laura Maccari,[†,‡] Matteo Magnani,[†] Giovannella Strappaghetti,[§] Federico Corelli,[†] Maurizio Botta,*[,†] and Fabrizio Manetti[†]

Dipartimento Farmaco Chimico Tecnologico, Università degli Studi di Siena, Via Alcide de Gasperi, 2, I-53100 Siena, Italy, and Dipartimento di Chimica e Tecnologia del Farmaco, Università di Perugia, Via del Liceo 1, I-06123 Perugia, Italy

Received January 27, 2006

The genetic function approximation (GFA) algorithm has been used to derive a three-term QSAR equation able to correlate the structural properties of arylpiperazine derivatives with their affinity toward the $\alpha_1$ adrenoceptor ($\alpha_1$-AR). The number of rotatable bonds, the hydrogen-bond properties, and a variable belonging to a topological family of descriptors ($\chi$) showed significant roles in the binding process toward $\alpha_1$-AR. The new model was also compared to a previous pharmacophore for $\alpha_1$-AR antagonists and a QSAR model for $\alpha_2$-AR antagonists with the aim of finding common or different key determinants influencing both affinity and selectivity toward $\alpha_1$- and $\alpha_2$-AR.

## INTRODUCTION

α adrenoceptors (ARs) belong to the family of G-protein coupled receptors (GPCRs) that constitute an important class of integral cell-membrane proteins, characterized by seven transmembrane domains. Because of their fundamental role in cellular signaling, α-ARs are key targets for the pharmaceutical development of drugs for several diseases.

Originally, the prefixes $\alpha_1$ and $\alpha_2$ were proposed to indicate classical postsynaptic α-ARs on vascular smooth muscle and presynaptic neuronal α-ARs on noradrenergic nerve terminals, respectively.[1] However, the identification of $\alpha_2$-ARs either outside noradrenergic terminals or at postsynaptic sites made it necessary to apply the prefixes $\alpha_1$ and $\alpha_2$ on the basis of specificity toward drugs, such as radioligands, that have been used for their characterization.[2,3]

As demonstrated by DNA cloning and sequencing techniques, GPCRs are a receptor family built on a common structural scheme. As a consequence, the search of selective ligands seems to be a very difficult task. On demonstration of that, there are many compounds able to interact with both $\alpha_1$- and $\alpha_2$-ARs,[4] because such receptors share many common features.[5,6] Interestingly, such similarities imply the existence of only a few ligand features responsible for discriminating affinity toward different α-ARs and make the discovery of a selective drug a remarkable challenge for medicinal chemistry. In general, previously, SAR studies highlighted the important role played by a protonated nitrogen allowing the ligand to interact with an aspartic residue of the binding site.[7d,7f,8] In addition to this essential feature, among the remaining molecular portions, ligands show several common structural moieties able to modulate affinity. This experimental evidence suggests that few molecular variables are able to account for a higher affinity toward one α-AR type compared to that toward the other.

Recently, $\alpha_1$-AR antagonists have attracted much attention for the treatment of benign prostatic hyperplasia-related lower urinary tract symptoms,[9] sexual dysfunctions,[10] hypertension, and cardiovascular diseases,[11] while $\alpha_2$-AR antagonists have been studied for their activity in a variety of diseases such as pain,[12] depression,[13] anxiety,[14] and obesity.[15]

The discovery of a lead candidate or a hit molecule worthy of further development includes characterization of both the affinity/activity and selectivity profiles toward such targets. As a consequence, the identification of features responsible for enhancing binding to the receptor of interest and for avoiding undesired side effects coming from interactions with different GPCRs is a critical issue. In this context, the identification of a molecule able to selectively interact with only one of the α adrenergic terminals is of great interest to the pharmaceutical industry.

In principle, with the aim of finding good ligands for a macromolecular target, when experimental protein structures are available (i.e., from X-ray crystallography or NMR studies), they are commonly used to investigate details of the receptor−ligand interaction mode and ligand affinity data and, subsequently, to rationally design novel chemical structures with improved properties as potential lead compounds (the so-called structure-based drug design). Unfortunately, no atomic-level structure exists for any of the α-ARs. Nevertheless, sequence analysis has been done and homology models have been obtained in the past on the basis of bacteriorhodopsin or rhodopsin structures.[7] Whereas homology modeling results afforded stimulating speculations about ligand binding, a direct relationship between structural properties and the biological outcome of a set of compounds may be also investigated through the ligand-based drug design approach. This modeling technique usually applies regression analyses to identify pharmacophoric elements or

* Corresponding author phone: ++39 (0) 577-234306; fax: ++39 (0) 577-234333; e-mail: botta@unisi.it.
† Università degli Studi di Siena.
‡ Present address: Drug Design and Information Technologies, Sienabiotech S.p.A., Via Fiorentina, 1, I-53100 Siena, Italy.
§ Università di Perugia.

10.1021/ci060031z CCC: $33.50 
Published on Web 04/18/2006

descriptor variables upon which the biological activity of a set of compounds depends. The resulting models are, in many cases, able to explain the variance of biological data, in both a qualitative and quantitative sense.

During the course of our research, because of the lack of experimental structural data on α-ARs, we resorted to a ligand-based ligand design approach, building a three-dimensional pharmacophoric model (by means of Catalyst software)[16] able to rationalize the relationships between structural features and $\alpha_1$-AR affinities for a set of arylpiperazine derivatives.[17] The model has been also successfully applied in database searching and the rational design of new molecules. As a result, many new compounds have been synthesized and tested following suggestions coming from the model, leading to the identification of molecules with improved affinity toward $\alpha_1$-AR.[18]

Moreover, we have recently developed a robust QSAR model for the whole set of our own arylpiperazine derivatives synthesized up to that moment, which describes the characteristics required for high $\alpha_2$-AR affinity.[19]

Both models described above were intended to be a part of a more comprehensive study aimed at gaining further knowledge on the interaction pathway between α-ARs and their ligands and increasing the predictive power on affinity and selectivity toward both $\alpha_1$- and $\alpha_2$-ARs. In this context, an additional QSAR model has been built, able to explore the quantitative relationships between structures and $\alpha_1$-AR affinities, thus leading to a more comprehensive view on compounds belonging to this structural class. Moreover, this study was also planned to sort out the requirements (from the physicochemical point of view) of arylpiperazinyl derivatives for finely tuning the affinity toward $\alpha_1$-AR and, together with the above-cited QSAR model for $\alpha_2$-AR, to predict the affinity profile toward either $\alpha_1$- or $\alpha_2$-ARs.

**Pharmacophoric Models Versus QSAR Models. General Considerations.** The pharmacophoric model previously generated for $\alpha_1$-AR antagonists, constituted by a positive ionizable feature representing the basic nitrogen atom, three hydrophobic groups, and a hydrogen-bond acceptor moiety, is consistent with the more general model accepted for $\alpha_1$-AR antagonists, which expects the central positive charge on the ligand structure and a polar moiety and an aromatic feature on the two opposite edges.[8c] It is also able to calculate activity data on the basis of how well the molecules match the pharmacophoric features. However, such a pharmacophoric model suffers from several problems.[17] In fact, pharmacophoric elements represented by size-, shape-, and chemical-function-based features do not consider more specific aspects of molecules, such as the electronic properties (i.e., dipole moment, polarizabilities, etc.) or the fine 2D molecular topology (i.e., structural connectivity, branching, etc.). Moreover, such a model does not deal with excluded volume problems.[20]

On the basis of such considerations, the identification of a pharmacophoric model may be intended to drive the earliest understanding of the ligand binding to the macromolecular target, but it may leave unexplored key elements involved in the binding events. Consequently, different approaches or techniques (such as typical QSAR methods) are usually required to better investigate steric and electronic influences exerted by ligands and affecting their binding properties.

For this purpose, a classical QSAR approach was applied, which started from the calculation of an extremely varied range of variables and then identified a limited number of them, able to describe the biological outcome of interest. In further detail, to explore all of the different properties that an arylpiperazinyl derivative should have to bind $\alpha_1$-AR and to fit the severe ligand−protein interaction requirements, different families of descriptors were calculated, each of them being able to depict a certain aspect of compounds (i.e., structural, electronic, quantum mechanical, spatial, topological, and thermodynamic features). A typical statistical pretreatment for all of the data generated was performed, and the genetic function approximation (GFA) algorithm was applied to calculate QSAR equations. Moreover, to allow for a better interpretation of such models, the partial least-squares (PLS) analysis was employed, leading to the understanding of how each molecular descriptor is able to influence the quantitative structure−affinity relationships detected by the model.

The final QSAR model was compared to the Catalyst pharmacophoric model, to define possible complementarities and divergences.

Finally, information encoded by both the QSAR and the pharmacophoric model for $\alpha_1$-AR blockers, integrated with findings derived from a QSAR equation calculated from a similar set of $\alpha_2$-AR inhibitors, allowed the identification of molecular determinants that could be modified to obtain high affinity and selectivity toward $\alpha_1$-AR.

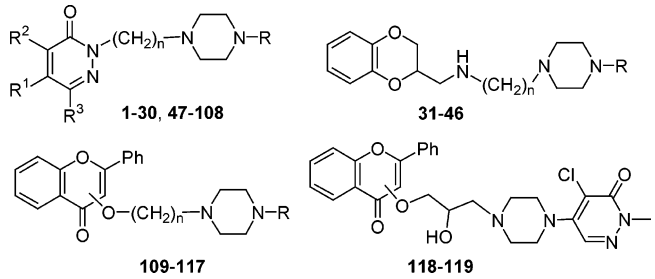
**Computational Details for the Development of the QSAR Model for $\alpha_1$-AR Antagonists.** A QSAR analysis has been performed on the basis of the following steps: (i) data set collection, (ii) the entry and alignment of molecules, (iii) the computation of descriptors, (iv) training set selection, (v) variable selection and QSAR model generation, (vi) QSAR model validation, and (vii) PLS analysis of the final model.

The building and alignment of molecular structures was performed by means of Catalyst software, while descriptor calculation and QSAR model generation were performed using the software Cerius2.[21] Finally, PLS analysis was performed by means of Minitab 13.[22]

*Data Set Collection.* Structure and affinity data of 119 arylpiperazines collected from the literature are reported in Table 1. Such compounds share a common structural pattern, being characterized by an arylpiperazinyl moiety linked through an alkyl spacer of variable length (from two to eight methylene units) to variously functionalized hydrophobic groups. Affinity values toward $\alpha_1$-AR (expressed as inhibition constants $K_i$, see the Supporting Information), ranging from 0.052 to 986 nM, were transformed into $-\log K_i$ values for subsequent calculations.

*Entry and Alignment of Molecules.* All of the compounds of the data set were built using the two- and three-dimensional sketcher of the software Catalyst. A representative family of conformations was generated for each molecule by applying the same settings used for generating the original pharmacophore model. Thus, the CHARMM force field implemented in Catalyst,[23] together with the Poling algorithm, and the best quality conformational analysis were applied for the conformational search.[24] The Poling function prevented conformations from being too close together and allowed for the exploration of conformational space of

**Table 1.** Compounds **1**−**119** Used to Generate and Validate the QSAR Model for $\alpha_1$-AR Antagonists




| compd | ref, name in ref[a] | p$K_i$ [b] | p$K_{i(\alpha1)}$/p$K_{i(\alpha2)}$ [c] | R[d] | R$^{1}$[d] | R$^{2}$[d] | R$^{3}$ | *n* | sub.[e] | RB[f] | HBA[g] | CHI[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 18b, 5a | 6.33 (6.94) | 0.64 | j | H | H | a | 2 | | 5 | 5 | 1.06 |
| **2**[i] | 18b, 5b | 7.32 (7.06) | 0.96 | j | H | H | a | 3 | | 6 | 5 | 1.00 |
| **3** | 18b, 5c | 7.61 (7.22) | 0.42 | j | H | H | a | 4 | | 7 | 5 | 0.98 |
| **4** | 18b, 5d | 7.63 (7.40) | 0.76 | j | H | H | a | 5 | | 8 | 5 | 0.98 |
| **5**[i] | 18b, 5e | 7.53 (7.57) | 0.67 | j | H | H | a | 6 | | 9 | 5 | 0.97 |
| **6** | 18b, 5f | 8.34 (7.76) | 0.73 | j | H | H | a | 7 | | 10 | 5 | 0.97 |
| **7** | 18b, 2a | 7.48 (7.54) | 1.03 | i | H | H | b | 2 | | 6 | 5 | 1.42 |
| **8** | 18b, 2b | 8.19 (7.67) | 1.39 | i | H | H | b | 3 | | 7 | 5 | 1.37 |
| **9** | 18b, 2c | 8.96 (7.82) | 1.16 | i | H | H | b | 4 | | 8 | 5 | 1.35 |
| **10** | 18b, 4a | 6.63 (7.05) | 0.80 | i | H | H | a | 2 | | 6 | 5 | 0.99 |
| **11** | 18b, 2d | 8.51 (7.99) | 0.99 | i | H | H | b | 5 | | 9 | 5 | 1.33 |
| **12** | 18b, 2e | 8.19 (8.17) | 1.25 | i | H | H | b | 6 | | 10 | 5 | 1.33 |
| **13**[i] | 18b, 2f | 8.77 (8.35) | 1.72 | i | H | H | b | 7 | | 11 | 5 | 1.33 |
| **14**[i] | 18b, 3a | 7.15 (7.44) | 0.58 | j | H | H | b | 2 | | 5 | 5 | 1.49 |
| **15** | 18b, 3b | 7.82 (7.56) | 0.47 | j | H | H | b | 3 | | 6 | 5 | 1.43 |
| **16** | 18b, 3c | 8.89 (7.72) | 0.65 | j | H | H | b | 4 | | 7 | 5 | 1.41 |
| **17** | 18b, 3d | 8.35 (7.89) | 1.45 | j | H | H | b | 5 | | 8 | 5 | 1.41 |
| **18**[i] | 18b, 3e | 7.82 (8.07) | 0.32 | j | H | H | b | 6 | | 9 | 5 | 1.40 |
| **19** | 18b, 3f | 9.00 (8.25) | 1.01 | j | H | H | b | 7 | | 10 | 5 | 1.40 |
| **20**[i] | 18b, 4b | 6.94 (7.18) | 0.77 | i | H | H | a | 3 | | 7 | 5 | 0.94 |
| **21** | 18b, 6a | 7.64 (8.21) | 1.68 | i | H | H | c | 2 | | 6 | 4 | 1.51 |
| **22**[i] | 18b, 6b | 8.49 (8.33) | 1.73 | i | H | H | c | 3 | | 7 | 4 | 1.45 |
| **23** | 18b, 6c | 9.05 (8.48) | 1.35 | i | H | H | c | 4 | | 8 | 4 | 1.43 |
| **24** | 18b, 6d | 8.17 (8.65) | 1.15 | i | H | H | c | 5 | | 9 | 4 | 1.42 |
| **25** | 18b, 6e | 8.55 (8.83) | 1.57 | i | H | H | c | 6 | | 10 | 4 | 1.42 |
| **26**[i] | 18b, 6f | 8.96 (9.01) | 2.44 | i | H | H | c | 7 | | 11 | 4 | 1.42 |
| **27** | 18b, 4c | 7.32 (7.33) | 0.92 | i | H | H | a | 4 | | 8 | 5 | 0.92 |
| **28** | 18b, 4d | 7.43 (7.49) | 1.06 | i | H | H | a | 5 | | 9 | 5 | 0.90 |
| **29** | 18b, 4e | 7.28 (7.67) | 0.40 | i | H | H | a | 6 | | 10 | 5 | 0.90 |
| **30**[i] | 18b, 4f | 8.22 (7.85) | 1.11 | i | H | H | a | 7 | | 11 | 5 | 0.90 |
| **31**[i] | 18c, 10 | 6.01 (6.87) | −0.78 | m | | | | 2 | | 7 | 5 | 0.68 |
| **32** | 18c, 3 | 7.78 (7.16) | 0.91 | i | | | | 3 | | 8 | 5 | 0.77 |
| **33**[i] | 18c, 4 | 6.94 (6.84) | 0.29 | j | | | | 2 | | 6 | 5 | 0.81 |
| **34** | 18c, 5 | 6.90 (7.05) | −0.08 | j | | | | 3 | | 7 | 5 | 0.83 |
| **35** | 18c, 6 | 6.70 (7.05) | −0.03 | l | | | | 2 | | 6 | 4 | 0.50 |
| **36** | 18c, 7 | 6.83 (7.23) | 0.41 | l | | | | 3 | | 7 | 4 | 0.50 |
| **37** | 18c, 8 | 6.38 (6.43) | −0.24 | n | | | | 2 | | 6 | 5 | 0.45 |
| **38** | 18c, 9 | 6.46 (6.61) | −1.13 | n | | | | 3 | | 7 | 5 | 0.45 |
| **39** | 18c, 11 | 6.14 (7.05) | −0.87 | m | | | | 3 | | 8 | 5 | 0.68 |
| **40** | 18c, 12 | 6.44 (6.58) | −1.08 | k | | | | 2 | | 6 | 5 | 0.59 |
| **41**[i] | 18c, 13 | 6.69 (6.77) | 0.62 | k | | | | 3 | | 7 | 5 | 0.59 |
| **42** | 18c, 14 | 6.57 (6.10) | 0.04 | h | | | | 2 | | 6 | 7 | 1.14 |
| **43**[i] | 18c, 15 | 6.53 (6.31) | −0.32 | h | | | | 3 | | 7 | 7 | 1.16 |
| **44**[i] | 18c, 16 | 6.75 (6.07) | −0.05 | o | | | | 2 | | 7 | 6 | 0.47 |
| **45**[i] | 18c, 17 | 6.48 (6.25) | −0.73 | o | | | | 3 | | 8 | 6 | 0.47 |
| **46** | 18c, 2 | 7.43 (6.97) | 0.26 | i | | | | 2 | | 7 | 5 | 0.76 |
| **47** | 17, 2a | 7.80 (7.99) | 1.41 | i | d | Cl | H | 2 | | 11 | 8 | 2.47 |
| **48**[i] | 17, 2j | 8.00 (8.06) | 0.59 | j | d | Cl | H | 4 | | 12 | 8 | 2.38 |
| **49** | 17, 2k | 8.35 (8.22) | 0.81 | j | d | Cl | H | 5 | | 13 | 8 | 2.36 |
| **50**[i] | 17, 2l | 8.38 (8.40) | 0.78 | j | d | Cl | H | 6 | | 14 | 8 | 2.35 |
| **51** | 17, 2m | 8.57 (8.58) | 0.44 | j | d | Cl | H | 7 | | 15 | 8 | 2.35 |
| **52** | 17, 2n | 8.25 (8.76) | 0.61 | j | d | Cl | H | 8 | | 16 | 8 | 2.35 |
| **53**[i] | 17, 1a | 7.42 (7.73) | 1.52 | i | H | H | d | 2 | | 11 | 7 | 1.76 |
| **54** | 17, 1b | 8.14 (7.84) | 1.54 | i | H | H | d | 3 | | 12 | 7 | 1.70 |
| **55** | 17, 1c | 9.22 (7.99) | 2.01 | i | H | H | d | 4 | | 13 | 7 | 1.67 |
| **56** | 17, 1d | 7.89 (8.15) | 0.73 | i | H | H | d | 5 | | 14 | 7 | 1.65 |
| **57**[i] | 17, 1e | 8.15 (8.33) | 1.01 | i | H | H | d | 6 | | 15 | 7 | 1.65 |
| **58** | 17, 2b | 7.84 (8.05) | 1.23 | i | d | Cl | H | 3 | | 12 | 8 | 2.37 |
| **59** | 17, 1f | 8.64 (8.51) | 1.30 | i | H | H | d | 7 | | 16 | 7 | 1.65 |
| **60** | 17, 1g | 7.74 (7.62) | 1.31 | j | H | H | d | 2 | | 10 | 7 | 1.83 |
| **61**[i] | 17, 1h | 8.17 (7.73) | 1.67 | j | H | H | d | 3 | | 11 | 7 | 1.76 |

**Table 1** (Continued)

| compd | ref, name in ref[a] | p$K_i$ [b] | p$K_{i(\alpha1)}$/p$K_{i(\alpha2)}$ [c] | R[d] | R1[d] | R2[d] | R3 | *n* | sub.[e] | RB[f] | HBA[g] | CHI[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **62** | 17, 1i | 9.10 (7.88) | 1.94 | j | H | H | d | 4 | | 12 | 7 | 1.74 |
| **63** | 17, 1j | 8.15 (8.05) | 1.29 | j | H | H | d | 5 | | 13 | 7 | 1.73 |
| **64** | 17, 1k | 8.08 (8.23) | 1.22 | j | H | H | d | 6 | | 14 | 7 | 1.72 |
| **65**[i] | 17, 1l | 7.82 (8.42) | 0.97 | j | H | H | d | 7 | | 15 | 7 | 1.73 |
| **66** | 17, 3c | 7.48 (8.02) | 1.12 | i | e | Cl | H | 2 | | 8 | 7 | 2.49 |
| **67**[i] | 17, 3a | 7.69 (8.08) | 1.52 | i | e | Cl | H | 3 | | 9 | 7 | 2.39 |
| **68**[i] | 17, 3e | 7.47 (8.20) | 1.41 | i | e | Cl | H | 4 | | 10 | 7 | 2.33 |
| **69** | 17, 2c | 8.37 (8.17) | 1.73 | i | d | Cl | H | 4 | | 13 | 8 | 2.31 |
| **70** | 17, 3g | 8.72 (8.35) | 1.09 | i | e | Cl | H | 5 | | 11 | 7 | 2.30 |
| **71** | 17, 3i | 8.39 (8.53) | 0.78 | i | e | Cl | H | 6 | | 12 | 7 | 2.30 |
| **72** | 17, 3k | 8.72 (8.71) | 2.44 | i | e | Cl | H | 7 | | 13 | 7 | 2.29 |
| **73**[i] | 17, 3d | 8.41 (7.92) | 1.59 | j | e | Cl | H | 2 | | 7 | 7 | 2.56 |
| **74** | 17, 3b | 8.46 (7.97) | 1.39 | j | e | Cl | H | 3 | | 8 | 7 | 2.45 |
| **75** | 17, 3f | 7.72 (8.09) | 1.17 | j | e | Cl | H | 4 | | 9 | 7 | 2.39 |
| **76** | 17, 3h | 8.44 (8.25) | 0.70 | j | e | Cl | H | 5 | | 10 | 7 | 2.37 |
| **77**[i] | 17, 3j | 8.24 (8.43) | 1.06 | j | e | Cl | H | 6 | | 11 | 7 | 2.37 |
| **78** | 17, 3l | 7.94 (8.61) | 0.50 | j | e | Cl | H | 7 | | 12 | 7 | 2.37 |
| **79** | 17, 2d | 8.41 (8.32) | 0.59 | i | d | Cl | H | 5 | | 14 | 8 | 2.29 |
| **80** | 17, 4e | 6.74 (7.96) | 0.64 | i | H | H | e | 4 | | 10 | 6 | 1.64 |
| **81** | 17, 4g | 7.36 (8.13) | 0.77 | i | H | H | e | 5 | | 11 | 6 | 1.62 |
| **82**[i] | 17, 4i | 7.75 (8.30) | 0.75 | i | H | H | e | 6 | | 12 | 6 | 1.62 |
| **83** | 17, 4k | 7.26 (8.49) | 0.24 | i | H | H | e | 7 | | 13 | 6 | 1.62 |
| **84**[i] | 17, 4d | 7.89 (7.70) | 1.31 | j | H | H | e | 3 | | 8 | 6 | 1.73 |
| **85** | 17, 4f | 8.10 (7.86) | 0.48 | j | H | H | e | 4 | | 9 | 6 | 1.70 |
| **86** | 17, 4h | 7.88 (8.03) | 0.90 | j | H | H | e | 5 | | 10 | 6 | 1.69 |
| **87**[i] | 17, 2e | 8.82 (8.49) | 0.37 | i | d | Cl | e | 6 | | 15 | 8 | 2.28 |
| **88** | 17, 4j | 8.23 (8.21) | 0.60 | j | H | H | e | 6 | | 11 | 6 | 1.69 |
| **89**[i] | 17, 4l | 8.25 (8.39) | 0.62 | j | H | H | e | 7 | | 12 | 6 | 1.69 |
| **90** | 17, 2f | 8.85 (8.68) | 0.52 | i | d | Cl | H | 7 | | 16 | 8 | 2.28 |
| **91**[i] | 17, 2g | 8.46 (8.86) | 0.81 | i | d | Cl | H | 8 | | 17 | 8 | 2.28 |
| **92** | 17, 2h | 7.23 (7.88) | 0.70 | j | d | Cl | H | 2 | | 10 | 8 | 2.54 |
| **93** | 17, 2i | 7.56 (7.94) | 0.90 | j | d | Cl | H | 3 | | 11 | 8 | 2.43 |
| **94** | 18a, 10 | 9.27 (9.04) | 0.43 | i | Cl | G | H | 7 | | 12 | 8 | 3.23 |
| **95**[i] | 18a, 11 | 9.37 (9.34) | 0.75 | p | Cl | G | H | 7 | | 13 | 8 | 3.32 |
| **96** | 18a, 12 | 9.49 (9.59) | 1.62 | q | Cl | G | H | 7 | | 13 | 8 | 3.54 |
| **97** | 18a, 13 | 9.26 (8.88) | 0.46 | i | f | Cl | H | 7 | | 17 | 8 | 2.30 |
| **98** | 18a, 14 | 9.37 (9.17) | 0.68 | p | f | Cl | H | 7 | | 18 | 8 | 2.39 |
| **99**[i] | 18a, 15 | 9.59 (9.43) | 1.09 | q | f | Cl | H | 7 | | 18 | 8 | 2.61 |
| **100**[i] | 18a, 16 | 9.24 (8.97) | 1.15 | p | d | Cl | H | 7 | | 17 | 8 | 2.38 |
| **101**[i] | 18a, 2 | 9.30 (9.00) | 0.90 | p | e | Cl | H | 7 | | 14 | 7 | 2.39 |
| **102** | 18a, 3 | 10.28 (9.26) | 1.03 | q | e | Cl | H | 7 | | 14 | 7 | 2.61 |
| **103** | 18a, 4 | 9.43 (8.90) | 0.52 | i | Cl | e | H | 7 | | 13 | 7 | 2.46 |
| **104** | 18a, 5 | 9.64 (9.20) | 0.54 | p | Cl | e | H | 7 | | 14 | 7 | 2.56 |
| **105**[i] | 18a, 6 | 10.10 (9.45) | 0.92 | q | Cl | e | H | 7 | | 14 | 7 | 2.78 |
| **106**[i] | 18a, 7 | 8.77 (8.81) | −0.30 | i | g | Cl | H | 7 | | 12 | 8 | 3.03 |
| **107** | 18a, 8 | 9.38 (9.11) | 0.71 | p | g | Cl | H | 7 | | 13 | 8 | 3.13 |
| **108** | 18a, 9 | 9.51 (9.36) | 1.03 | q | g | Cl | H | 7 | | 13 | 8 | 3.35 |
| **109** | 18d, 1 | 8.23 (8.00) | 2.25 | i | | | | 2 | 3 | 8 | 5 | 1.62 |
| **110**[i] | 18d, 2 | 8.11 (8.14) | 1.83 | i | | | | 3 | 3 | 8 | 8 | 1.79 |
| **111** | 18d, 3 | 8.24 (7.49) | 2.15 | i | | | | 2 | 7 | 7 | 5 | 1.22 |
| **112**[i] | 18d, 4 | 8.11 (7.70) | 2.14 | i | | | | 3 | 7 | 6 | 7 | 1.59 |
| **113** | 18d, 5 | 8.14 (7.50) | 1.89 | i | | | | 2 | 6 | 8 | 8 | 1.79 |
| **114**[i] | 18d, 7 | 7.26 (7.13) | 2.26 | h | | | | 2 | 3 | 7 | 5 | 1.66 |
| **115** | 18d, 8 | 7.12 (7.29) | 2.12 | h | | | | 3 | 3 | 7 | 7 | 2.01 |
| **116** | 18d, 10 | 6.69 (6.84) | 1.69 | h | | | | 3 | 7 | 6 | 7 | 2.03 |
| **117**[i] | 18d, 11 | 6.24 (6.62) | 1.05 | h | | | | 2 | 6 | 8 | 5 | 1.24 |
| **118** | 18d, 12 | 7.00 (6.66) | 0.94 | | | | | | 7 | 7 | 5 | 1.22 |
| **119** | 18d, 13 | 6.36 (6.66j) | 0.87 | | | | | | 6 | 7 | 7 | 1.62 |

[a] Reference reporting that compound and the compound name in that reference. [b] In parentheses, estimated/predicted p$K_i$ values. [c] Ratio between p$K_i$ affinity values toward $\alpha_1$- and $\alpha_2$-AR. [d] a: 1-imidazolyl. b: 1-benzimidazolyl. c: 1-indolyl. d: 4-[2-(2-methoxyphenoxy)ethyl]piperazin-1-yl. e: 4-(2-furoyl)piperazin-1-yl. f: 4-[2-(2-ethoxyphenoxy)ethyl]piperazin-1-yl. g: 4-[2-(1,4-benzodioxan)]piperazin-1-yl. h: 2-methyl-4-chloropyridazin-3(2*H*)-one-5-yl. i: *o*-methoxyphenyl. j: *o*-chlorophenyl. k: *o*-fluorophenyl. l: phenyl. m: *p*-methoxyphenyl. n: 2-pyridinyl. o: 2-furoyl. [e] Substitution pattern on the flavone nucleus. [f] Values assumed by the Rotlbonds (RB) descriptor. [g] Values assumed by the Hbond acceptor (HBA) descriptor. [h] Values assumed by the CHI-V-12_PC (CHI) descriptor. [i] Compounds belonging to the training set.

molecules within a user-defined energy threshold. The best quality analysis was considered as the method of choice for systems with flexible rings. Conformational diversity was emphasized by selection of the conformers that fell within a 20 kcal/mol range above the lowest-energy conformation. For each compound, the number of conformers generated by the algorithm was found to be <1000. The pharmacophoric model for $\alpha_1$-AR antagonists was then used to define the alignment criterion. In particular, the best-fitting conformer of each molecule was selected and, in turn, used to calculate different families of molecular descriptors.

**Table 2.** Statistical Parameters Relative to the First Eight Principal Components (PC), Resulting from PCA Performed on the Whole Set of 119 Arylpiperazine Compounds and 311 Descriptors

| | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 |
|---|---|---|---|---|---|---|---|---|
| eigenvalue | 108.71 | 32.58 | 18.64 | 17.48 | 8.11 | 7.37 | 6.48 | 4.88 |
| proportion | 0.463 | 0.139 | 0.079 | 0.074 | 0.035 | 0.031 | 0.028 | 0.021 |
| cumulative | 0.463 | 0.601 | 0.681 | 0.755 | 0.789 | 0.821 | 0.848 | 0.869 |

*Computation of Descriptors.* The superposition assembly was imported into Cerius2. Molecular descriptors belonging to different categories (conformational, electrotopological, electronic, information-content, quantum mechanical, spatial, structural, thermodynamic, and topological) were calculated for all 119 molecules. The resulting descriptors (470) were reduced to 311 by the rejection of variables with null variance.

*Training Set Selection.* The selection of training set compounds was performed taking into account both the structural features and activity profile of the whole data set.[25] In detail, to maximize the structural diversity and to maintain the original distribution profile of affinity data, the training set was selected on the basis of the simultaneous optimization of multiple properties of a subset of compounds, which is a common concept in the field of library design.[25] Selection of the training set was carried out by means of the C2.LibProfile and C2.Diversity modules, implemented in Cerius2. This selection method is based on the specification of an objective function and its following optimization. In our case, the objective function consisted of two terms, the first one taking into account the structural diversity of selected compounds and the second one accounting for their activity profile. Structural diversity was evaluated by the MaxMin diversity function that drives the compound selection step in such a way as to maximize the minimum squared distance of each point (compound) with respect to all other points in the selected subset of molecules. For this purpose, the principal component analysis (PCA) technique transformed the original set of 311 descriptors into a reduced number (8) of orthogonal variables (principal components), able to explain 87% of the variance of the original variables. Table 2 reports the statistical parameters of the eight principal components.

On the other hand, the second term of the objective function was defined on the basis of the affinity profile histogram relative to the whole set of compounds. It consists of a penalty function which drives the selection of a subset of compounds characterized by an activity distribution similar to that shown by the reference histogram (reporting the whole data set).

A genetic algorithm was then applied to reach the best value of the objective function (optimization step), leading to the simultaneous optimization of both diversity and penalty functions.

As a result, 40 compounds (a third of the whole data set) were selected as the training set, while the remaining 79 compounds constituted the test set. Both the training and test sets uniformly spanned the descriptor space, as demonstrated by Figure 1, showing compounds of the training and test sets in the space of the first three principal components (explaining approximately the 70% variance). Moreover, as expected, the activity distribution profile of both the training and test sets (Figure 2B and Figure 2C, respectively) is very similar to that of the whole data set (Figure 2A).

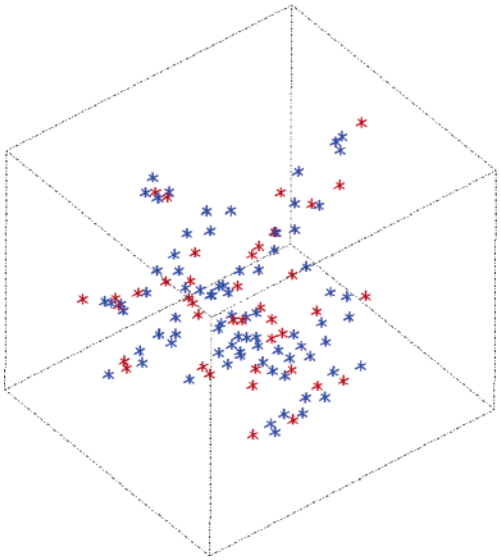


**Figure 1.** Training set (red crosses) and test set (blue crosses) in the space of the first three principal components, accounting for about 70% of variance of the whole data set. Both training and test sets are uniformly distributed in the descriptor space.

*Selection of Variables and QSAR Model Generation.* The descriptor set was further reduced by removing those descriptors found to be not significantly correlated to $-\log K_i$ at the 95% confidence level ($p > 0.05$). After these screening procedures, a total of 238 descriptors (about 50% of the original set) remained available for model generation.

Because of the large number of descriptors, the GFA algorithm of the genetic algorithm module of Cerius2 was used to efficiently search the wide solution space and to identify the most suitable subsets of descriptors in order to build robust QSAR models. The GFA method[26] uses a genetic algorithm to search the space of possible QSAR models, referring to the lack-of-fit (LOF) score (see the Supporting Information) to estimate the fitness of each model. Therefore, the process tries to evolve a population of equations that best fit data of the training set.

Moreover, a distinctive feature of GFA is that it produces a population of models, instead of a single model, thus providing useful additional information. In fact, predictions could be improved by averaging the results of multiple models, and the frequency of use of descriptors in the population of equations could be useful to reveal the descriptors relevant to activity prediction. In addition, GFA is a good technique to find subsets of variables best correlating with activity. These subsets could not be discovered by techniques which select variables incrementally, thus evaluating only their individual predictivity.

Given the large number of descriptors (238), the size of the initial population of equations was raised from 100 (the default) to 300, to efficiently search the equation space. All of the initial equations consisted of four terms, in addition to a constant. The equation length was not fixed, allowing it

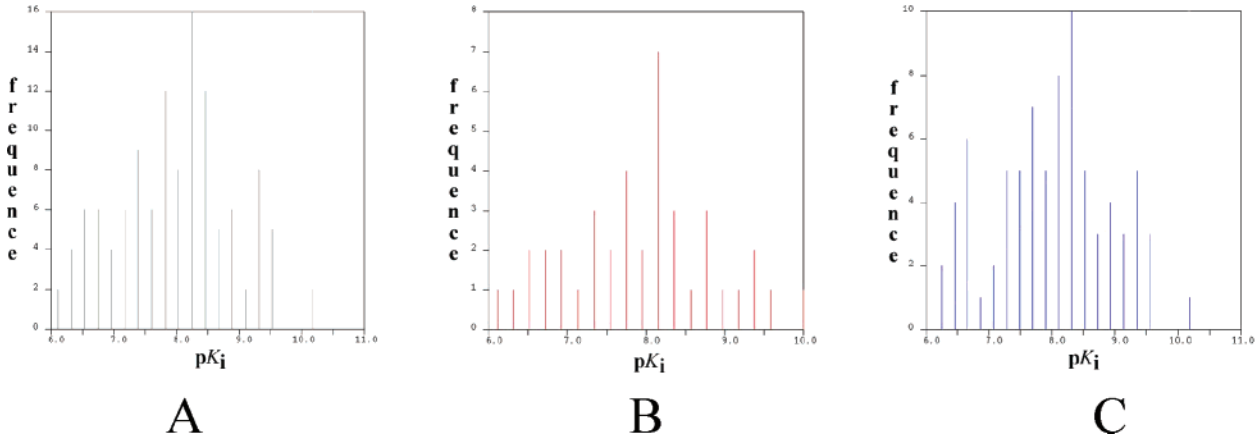



**Figure 2.** Activity profile histograms for the whole data set (A), training set (B), and test set (C).

**Table 3.** Overview of the Ten Best QSAR Models Obtained by Means of the GFA Algorithm

| | | statistical parameters | | | |
|---|---|---|---|---|---|
| no. | QSAR equation[a] | LOF | $R^2$ | $F$ | $Q^2$ |
| 1 | aff. = 7.61 − 0.56 "Hbond acceptor" + 0.18 "Rotlbonds" + 1.15 "CHI-V-12_PC" | 0.21 | 0.86 | 70.85 | 0.82 |
| 2 | aff. = 7.78 + 1.33 "CHI-V-13_PC" − 0.57 "Hbond acceptor" + 0.19 "Rotlbonds" | 0.21 | 0.85 | 70.74 | 0.82 |
| 3 | aff. = 9.25 + 0.16 "Rotlbonds" − 0.38 "Shadow-nu" − 0.62 "Hbond acceptor" + 1.47 "CHI-V-13_PC" | 0.22 | 0.88 | 64.52 | 0.84 |
| 4 | aff. = 9.73 − 0.59 "Hbond acceptor" + 0.21 "Rotlbonds" + 1.78 "CHI-V-14_PC" − 0.09 "Shadow-Xlength" | 0.22 | 0.88 | 64.37 | 0.84 |
| 5 | aff. = 7.38 + 0.18 "Rotlbonds" − 0.54 "Hbond acceptor" + 1.01 "CHI-V-11_PC" | 0.22 | 0.85 | 69.55 | 0.82 |
| 6 | aff. = 9.46 − 0.63 "Hbond acceptor" + 1.71 "CHI-V-14_PC" + 0.17 "Rotlbonds" − 0.41 "Shadow-nu" | 0.22 | 0.88 | 63.92 | 0.84 |
| 7 | aff. = 9.42 − 0.08 "Shadow-Xlength" − 0.58 "Hbond acceptor" + 1.52 "CHI-V-13_PC" + 0.19 "Rotlbonds" | 0.22 | 0.88 | 63.82 | 0.84 |
| 8 | aff. = 7.85 − 0.57 "Hbond acceptor" + 1.52 "CHI-V-14_PC" + 0.20 "Rotlbonds" | 0.22 | 0.85 | 68.10 | 0.82 |
| 9 | aff. = 7.17 − 0.54 "Hbond acceptor" + 0.18 "Rotlbonds" + 0.91 "CHI-V-10_PC" | 0.22 | 0.85 | 67.58 | 0.81 |
| 10 | aff. = 8.90 + 1.26 "CHI-V-12_PC" + 0.16 "Rotlbonds" − 0.34 "Shadow-nu" − 0.60 "Hbond acceptor" | 0.23 | 0.88 | 61.85 | 0.83 |

[a] The term aff. represents $-\log K_i$ where $K_i$ is the affinity of compounds toward $\alpha_1$-AR, expressed in M concentrations. $R^2$, LOF, $F$, and $Q^2$ are the coefficient of determination, the lack-of-fit value, the value from F test, and the cross-validated coefficient of determination, respectively.

to vary during the evolution. Only linear terms were used to construct equations, thus avoiding the insertion of spline and quadratic terms. The mutation probability, consisting of the addition of a new term, was set to 0.5, while the smoothing parameter ($d$) of the LOF expression was set to 2 (1 is the default), to bias toward smaller models. The least-squares regression method was applied, while evolution was allowed to proceed for 50 000 crossovers.

The output of GFA calculations consisted of 300 equations ranked according to their fitness (LOF) value. The first 10 best-scored equations reported in Table 3 shared very similar information content in terms of statistical parameters and types of descriptors. In particular, they could be classified in three-term and four-term equations and contain common descriptors, such as Rotlbonds (RB), Hbond acceptor (HBA), and a variable belonging to the CHI family. The additional descriptor of the four-term equations belongs to the Shadow family.[27] Although four-term models (eqs 3, 4, 6, 7, and 10) showed $R^2$ and $Q^2$ values (0.88 and 0.83−0.84, respectively) that were slightly better with respect to the remaining models (with $R^2$ and $Q^2$ values of 0.85−0.86 and 0.81−0.82, respectively), the last were characterized by improved predictive power toward test set affinity data. On the basis of these findings, three-term models were chosen, and among eqs 1 and 2 (both showing a LOF value of 0.21 and a very similar composition), the first one was definitely selected as the final model, also on the basis of its better $R^2$ with respect to eq 2. The model showed either a good fit to the experimental data or a good predictive power, being characterized by the following statistical parameters: $R^2 = 0.86$, LOF = 0.21, $F = 70.85$, $Q^2 = 0.82$, and $s = 0.38$. Estimated versus observed $-\log K_i$ values, as well as the residual plot for training set compounds, are shown in Figure 3 (A and B, respectively).

*QSAR Model Validation.* The predictive power of the QSAR model was also assessed by both internal and external validation tests. Internal validation comprised both a cross-validation and a randomization test. The $Q^2$ value (0.82) resulting from a leave-one-out (LOO) procedure suggested a good predictive power. To complement the study based on LOO, leave-several-out cross-validation runs were carried out for $n = 2$, 8, 14, 20, 26, 32, and 38, $n$ being the number of groups which were left out during the cross-validation procedure. The results of 10 cross-validations for each value of $n$ are shown in Figure 4. With the exception of $Q^2 = 0.72$ found by using only two groups, all of the remaining cross-validations produced mean values of $Q^2 > 0.80$.

The randomization test was performed by repeatedly scrambling the affinity data, which were in turn used to generate QSAR models, whose scores were compared with the score of the original QSAR equation generated with nonrandomized data. In detail, a total of 99 models were

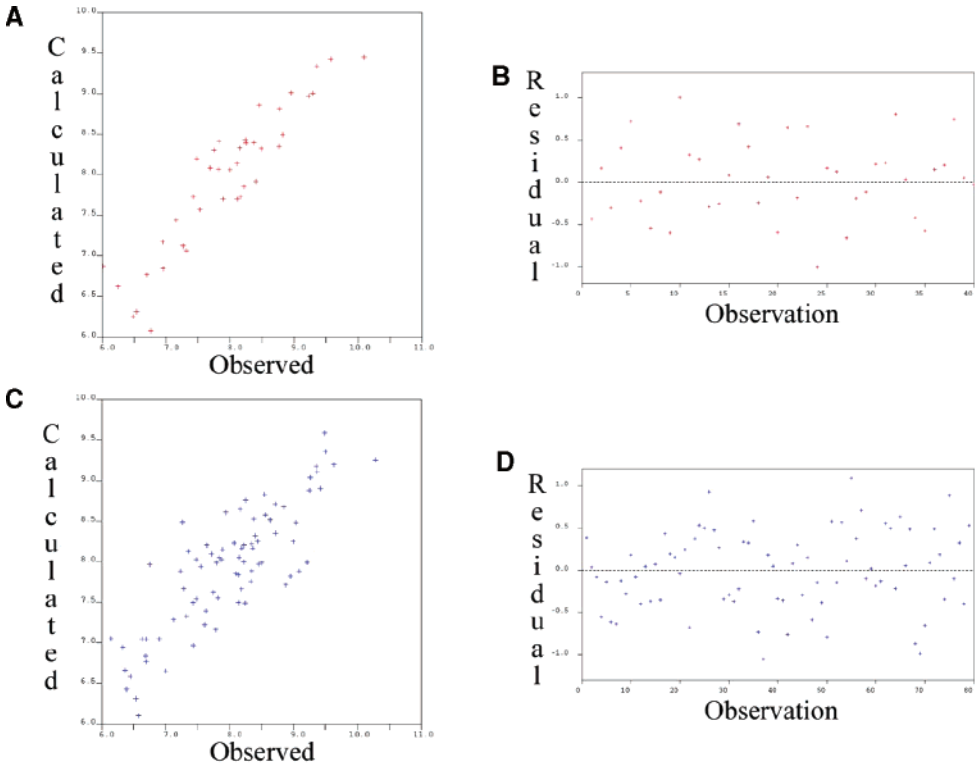



**Figure 3.** Calculated (estimated and predicted) versus observed p$K_i$ activity values for training- (A; $R^2 = 0.86$) and test-set (C; $R^2 = 0.70$) compounds, respectively. Residual plots: residual versus indices of training-set (B) and test-set (D) observations.

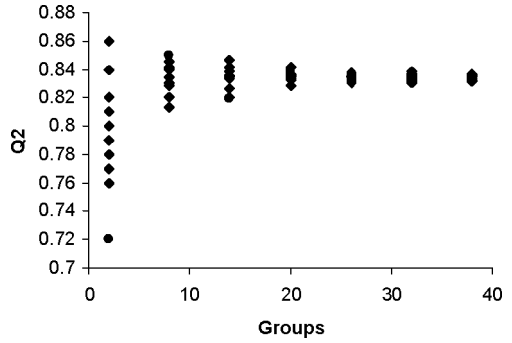



**Figure 4.** *n*-Fold cross-validation for the three-variable eq 1.

generated using permuted data, corresponding to a 99% confidence level. None of the equations resulting from randomized data showed an *r* value higher than 0.92, found for the nonrandom model (Figure 5). Moreover, 0.92 was significantly higher than the mean value of *r* values found for models obtained from random trials (mean $r = 0.44$, SD $= 0.16$).

An external validation of the model was also performed through the prediction of affinity data of 79 molecules belonging to the test set. A good correlation ($R^2 = 0.70$) was found between predicted and experimental $-\log K_i$ values (Figure 3C). Remarkably, 55 compounds (nearly 70% of test set molecules) were predicted within 0.50 log units (Figure 3D and Table 1).

*Model Interpretation by PLS Analysis.* Considering that the PLS analysis of a previously calculated QSAR model has been proven to be of a great help in its understanding,[28] we have performed such an analysis on compounds belonging to our original training set, using only descriptors found as independent variables (or predictors) in the original QSAR model (see the Supporting Information).

## RESULTS AND DISCUSSION

Among all of the descriptors calculated for the set of arylpiperazine derivatives, the QSAR analysis found that affinity toward $\alpha_1$-AR depends on the number of hydrogen-bond acceptor (HBA) features, the number of rotatable bonds (RB), and the topological variable CHI-V-12_PC (CHI), a Kier−Hall valence-modified 12th-order path-cluster molecular connectivity index.[29] In particular, ligand affinity is predicted by the model to be enhanced by an increase of RB and CHI values,[30] while an inverse relationship exists with the number of HBA groups, as shown in the following equation representing the final QSAR model:

$$-\log K_i = 7.61 - 0.56\ (\pm 0.08)\ \text{“Hbond acceptor”} + 0.18\ (\pm 0.02)\ \text{“Rotlbonds”} + 1.15\ (\pm 0.13)\ \text{“CHI-V-12\_PC”}$$

Both RB and CHI relate to the size of molecules. However, they account for two different aspects of molecules under study. In fact, whereas RB depends on the flexibility of molecules, CHI accounts for the connectivity of ligand structures. Moreover, their 59% correlation (Table 4) suggests that an increase in the first descriptor is not necessarily linked to an increase in the second descriptor.

On the basis of the QSAR equation, molecules with a lower number of HBAs are estimated/predicted to have a higher affinity. As an example, compounds **23**, **84**, and **43** show an increasing number of HBAs and a corresponding decrease in affinity toward $\alpha_1$-AR.

It is important to point out that, despite the usual efforts made to understand how biological data depend on the single descriptor of a QSAR model, each statistical model must be considered as a whole, and the concomitant action of all three descriptors, together with their coefficient signs, have to be

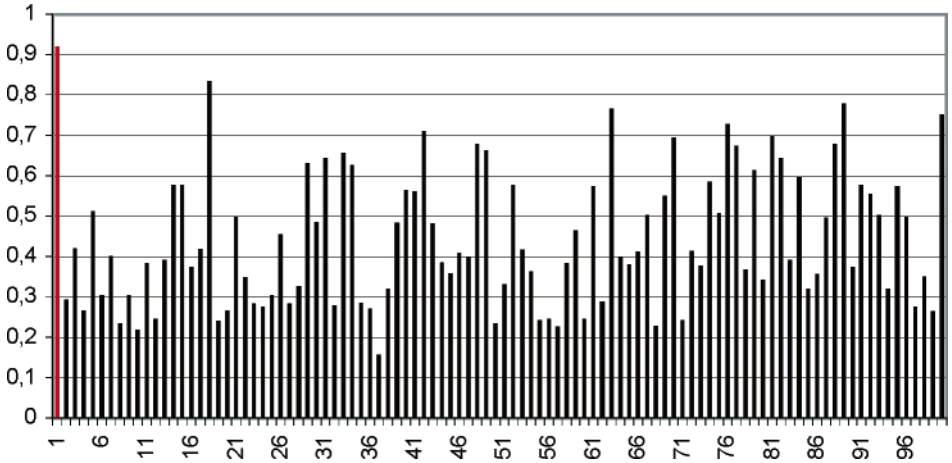


**Figure 5.** GFA randomization test. The red bar (left) shows the *r* value for the nonrandom model, while the remaining 99 black bars show *r* values for models derived from permuted data.

**Table 4.** Correlation Matrix Showing the Pearson Product Moment Correlation Coefficient between Each Pair of Variables[a]

| | RB | HBA |
|---|---|---|
| HBA | 0.629 (0.396) | 1.000 (1.000) |
| CHI | 0.594 (0.353) | 0.758 (0.574) |

[a] Pearson coefficients measure the degree of linear relationship between two variables. The coefficient is negative (maximum allowed value of −1) if one variable tends to increase as the other decreases. Conversely, if the two variables tend to increase together, the correlation coefficient is positive (maximum allowed value of +1). The corresponding $R^2$ value is reported within parentheses.

**Table 5.** Summary of the PLS Analysis for the QSAR Model

| components | $R^2$ | $Q^2$ [a] | PRESS | X variance | error SS |
|---|---|---|---|---|---|
| 1 | 0.55 | 0.49 | 18.06 | 0.77 | 16.13 |
| 2 | 0.84 | 0.80 | 7.12 | 0.88 | 5.82 |
| 3 | 0.85 | 0.82 | 6.32 | 1.00 | 5.14 |

[a] Leave-one-out cross-validation.

**Table 6.** X Weights for the PLS Analysis for the QSAR Model

| | component 1 | component 2 | component 3 |
|---|---|---|---|
| RotlBonds | 0.60 | 0.47 | −0.68 |
| HBA | 0.59 | −0.92 | −0.07 |
| CHI-V-12_PC | 0.61 | −0.02 | 0.72 |

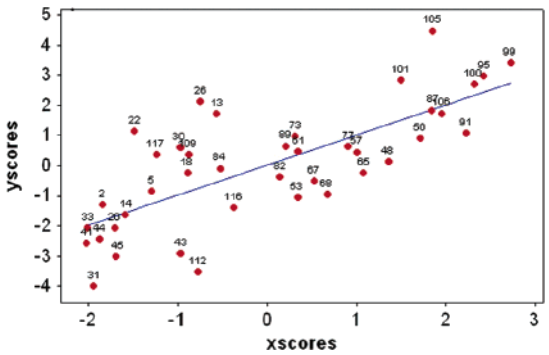



**Figure 6.** Score plot for the first component of the PLS model.

kept in mind. This consideration implies that the estimated/predicted outcome of the regression equation is given by a weighted effect exerted at the same time by HBA, RB, and CHI values. In other words, coefficients of the regression equation are not always directly interpretable, because they are embodied into a composite picture of three determinants (the descriptors) acting on the model.

However, to better understand the influence of each single determinant on the QSAR captured by the model, a PLS analysis was performed (see the Supporting Information), indicating that the optimal number of components was three, thus confirming that the model was not overfitted. Moreover, PLS analysis also showed that the activity of a compound badly estimated by a given component could be corrected by a different component of the model, which is represented by a different combination of the descriptors in terms of coefficients and signs. The results are summarized in Tables 5 and 6, which provide statistical parameters and X weights, respectively, for the three valid components. Each of them was a linear combination of the three descriptors reported in the QSAR equation. Thus, the X weights represented the contribution of each descriptor to a given component.

In further detail, component 1, explaining 55% of the variance of biological data (out of a total of 85% explained by the first three components, Table 5), was able to capture most of the information regarding compounds of the training set. In fact, both the upper left and lower right regions of the score plot (Figure 6) were not significantly populated, most of the compounds being located close to the diagonal axis.

As shown in Table 6, the three descriptors were equally weighted in component 1, and their coefficients had a positive sign, indicating that an increase in their values led to higher activity. Accordingly, **100** and **87**, characterized by a flexible structure (high number of RBs) and also showing high values for HBA and CHI variables, were estimated to be highly active compounds, in agreement with their experimental data. On the other hand, **33**, **14**, and **41**, whose affinities were well estimated by the model, were much less active and showed values for all three descriptors significantly lower than those found for **100** and **87**. The same trend was followed by **61**, **82**, **77**, having medium values for RB, HBA, and CHI descriptors, corresponding to medium affinity, correctly estimated by the model.

However, component 1 poorly estimated compounds whose p$K_i$ values were inversely proportional to the values of HBA, RB, and CHI. As an example, **105**, a very active compound with values for the three descriptors (especially

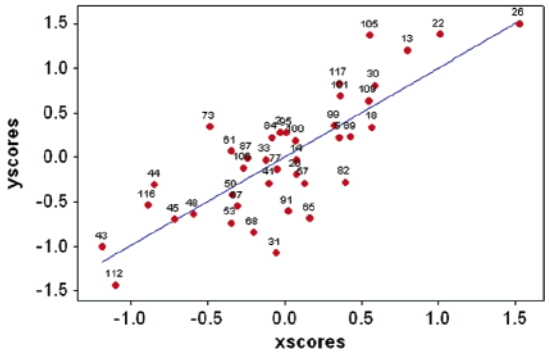



**Figure 7.** Score plot for the second component of the PLS model.

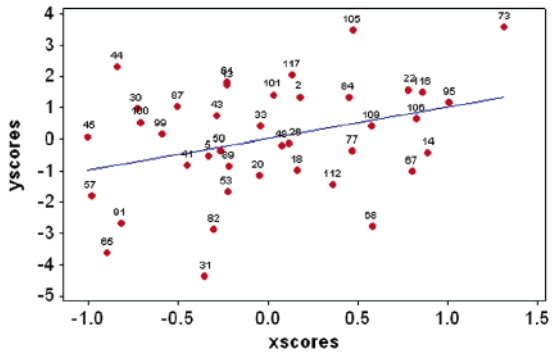



**Figure 8.** Score plot for the third component of the PLS model.

for HBA) lower with respect to the other highly active arylpiperazines, was underestimated. Similarly, **13**, **22**, and **26**, showing relatively high p$K_i$ values, were underestimated on the basis of their low number of HBA features. On the contrary, **112**, **31**, **45**, and **43** were overestimated.

Component 2, accounting for 29% of the total variance of the dependent variable (Table 5), was expected to have a strong influence in estimating affinity, similarly to component 1. Moreover, the corresponding score plot reported in Figure 7, showing many data points close to the diagonal axis, demonstrated the good estimating properties of such a component. In detail, HBA is the descriptor with the highest weight and a negative sign (Table 6), suggesting that higher p$K_i$ values should be associated with lower numbers of HBA features. On this basis, component 2 contributed to an adjustment of affinity values overestimated (or underestimated) by component 1 for compounds showing many (or few) hydrogen-bonding capabilities and low activity (or high activity). In fact, compounds **43**, **45**, and **112**, with high HBA values, significantly overestimated by component 1, were estimated well by component 2. In a similar way, **13**, **22**, and **26**, characterized by a low HBA descriptor value and underestimated by component 1, were located in a position close to the diagonal axis of the graph for component 2, thus compensating for their underestimated affinity by the first component.

In summary, component 2 tended to roughly estimate medium-active compounds, such as **65** (overestimated) or **73** (underestimated), whereas these compounds were well estimated by the first component. Compounds **31** and **105** turned out to be poorly predicted by the second component also.

Component 3 generally showed poor estimating properties toward training set compounds. The only exception was represented by **95**, which was estimated slightly better with respect to the first two components (Figure 8).

Although component 3 showed poor estimating properties and was not expected to add much more information to the SAR already extracted by the first two components (it accounted only for 1% of the total variance encoded in the independent variable), it is important to note that it was able to capture 12% of the X variance, similarly to component 2 (Table 5).

Component 3 regulated the activity of all of the molecules only on the basis of their flexibility and molecular connectivity. Weights and signs of coefficients for the third component (Table 6) suggested that compounds with a higher value of the connectivity index and a lower number of rotatable bonds should have a higher activity, while HBA showed a nonsignificant contribution in defining affinity values.

In summary, PLS analysis allowed for a better interpretation of the overall QSAR model, in terms of both the sign and coefficient of each variable. Although it makes interpretation more difficult, such an analysis provided information that could not be evident from just reading the QSAR equation and allowed for a clearer understanding of the role played by each variable in describing the affinity of each compound.

Compounds **105** and **31**, poorly estimated by all three components, are worthy of further consideration. In fact, they are characterized by descriptor values very similar to those of **101** and **41**, respectively. However, there are differences in affinity, **101** being 0.80 log units less-active than **105**, while **41** shows an affinity 0.68 log units better than that of **31**. On this basis and considering that affinity values for **101** and **41** are well-estimated by PLS components, compounds **105** and **31** were not expected to be well-predicted by the model. Therefore, the relevant difference between their predicted and actual affinity values (0.65 and 0.86 log units, respectively) led us to consider them as outliers. As a consequence, descriptors different from those present in the model could be necessary to explain the properties buried into them. Otherwise, it may also be supposed that an error affecting their experimentally determined biological affinity occurred.

Interestingly, the QSAR model was not based on properties that may be influenced by 3D arrangement of the arylpiperazine derivatives. In fact, the variables appearing in the model are two-dimensional descriptors, independent of the conformations used during calculations. In agreement with such a result, an attempt to generate a molecular-field-based model for the same set of compounds was also performed, without any satisfactory result. In fact, several PLS models were generated by means of the software Golpe,[31] starting from different combinations of Grid[32] molecular interaction fields calculated for C3 (a probe simulating a methyl group), N1 (an amide NH probe), and an O probe (corresponding to a carbonyl oxygen atom). Such models, while characterized by high $R^2$ and $Q^2$ values for the training set compounds, showed a very poor predictive power on test set compounds (SDEP values between 0.70 and 0.80), thus suggesting overfitting. Furthermore, their PLS pseudo-coefficient plots did not allow for finding any correlation between molecular steric/electrostatic fields and affinity toward $\alpha_1$-AR, being contour maps characterized by small and scattered polyhedra,

unsuited for identifying molecular regions critical for activity. This finding could further suggest that conformation-related properties were not involved in the steps leading this class of compounds to bind the $\alpha_1$-AR. Moreover, it represented an indirect validation of the previously developed pharmacophoric model, in the sense that size-, shape-, and chemical-function-based pharmacophoric features seemed to be sufficient to account and rationalize affinity values of arylpiperazines toward $\alpha_1$-AR. In other words, no additional property of the molecules under study (such as properties encoded by electronic, thermodynamic, and quantum mechanical descriptors) seems to be required to develop a structure−activity relationship for this class of compounds.

To support this hypothesis, it is also important to point out that a QSAR analysis performed on the same set of compounds bearing the most basic nitrogen atom of the piperazine ring in the protonated form led to models containing the same descriptors reported in the current QSAR equation and characterized by very similar statistical parameters. This finding was in agreement with results coming from both the QSAR analysis on the unprotonated compounds and that developed for $\alpha_2$-AR antagonists, both of them taking into account any descriptor related to the charged nitrogen, which is known to strongly interact with an Asp residue of helix 3, through a charge-reinforced hydrogen-bond interaction.

Such additional calculations suggested that the role played by the protonatable nitrogen was not influential in modulating affinity values toward $\alpha_1$-AR. In fact, as already hypothesized,[33] the protonated nitrogen of a putative ligand and the Asp residue of the protein were responsible for a long-range electrostatic interaction occurring during an early step of the recognition process. Such an interaction was a constant contribution to the ligand binding, because of the basic nitrogen that was a common feature of all of the compounds under study. As a consequence, this contribution was not expected to be accounted for by a QSAR model. On the contrary, the model identified molecular features discriminating among the ligands that were supposed to influence the short-range interactions following the first recognition step. These short-range acting forces were in turn related to the modulation of both affinity and selectivity among $\alpha_1$- and $\alpha_2$-ARs and, even more finely, among their subtypes.

The current QSAR model for $\alpha_1$-AR antagonists led to suggestions that are also in good agreement with those pointed out using a previous model derived from a smaller set of arylpiperazines (14 compounds), reporting $\alpha_1$-AR affinity "as dependent on the length of the alkyl spacer and on the presence of hydrogen bond acceptors".[34]

To summarize, both the Catalyst and GFA models underlined the significant role of a few chemical functionalities in the binding process toward $\alpha_1$-AR. (i) The alkyl chain acting as a spacer between the arylpiperazine and the terminal heterocyclic moieties plays a fundamental role in defining affinity. In fact, while the pharmacophore for $\alpha_1$-AR antagonists suggested four to seven methylene groups as the optimal length to perfectly fit all the pharmacophoric features, the QSAR model presented in this paper demonstrated the importance of the flexibility of molecules, mainly exerted by the spacer. (ii) The hydrogen-bonding properties of molecules also turned out to be important for affinity. The QSAR model emphasized that hydrogen-bonding features were involved in defining $\alpha_1$-AR's affinity, but the negative sign associated with this descriptor in the QSAR equation defined a limit to the number of HBAs that could be present in a molecule to have high affinity (the pharmacophoric model included one hydrogen-bond acceptor group among the five pharmacophoric features). Moreover, the negative sign of the HBA descriptor acted as a balancing factor for the positive signs of both RB and CHI determinants.

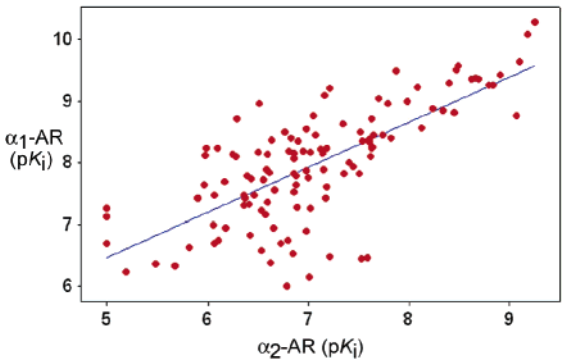



**Figure 9.** Relationship between p$K_i$ values for $\alpha_1$- and $\alpha_2$-ARs for 119 arylpiperazine compounds considered in this study.

**Comparison between QSAR Models for $\alpha_1$- and $\alpha_2$-AR Antagonists.** The aim of this work was also to identify and analyze the (few) discriminating functionalities which determine selectivity within the family of α-ARs. For this purpose, the following simple linear regression between the binding affinities toward $\alpha_1$- and $\alpha_2$-AR was calculated for the 119 arylpiperazine derivatives under study (Figure 9):

$$pK_i(\alpha_1) = 2.80 + 0.73\ (\pm 0.06) \times pK_i(\alpha_2)$$

with $R^2 = 0.51$, $s = 0.65$, and $F = 125.49$, where $R^2$ is the coefficient of determination, $s$ represents the standard error, and $F$ is the value of the F test.

Such an equation indicates that, within the whole series of compounds, vectors describing binding affinity toward $\alpha_1$-AR were able to explain 51% of the total variation in affinity toward $\alpha_2$-AR, thus leaving the remaining 49% unexplained. In other words, $\alpha_1$-AR affinity shows a poor correlation with the binding affinity for $\alpha_2$-AR. To support this hypothesis, it is important to note that a SAR analysis of the compounds in Table 1 does not allow for finding particular trends in the selectivity between $\alpha_1$- and $\alpha_2$-AR. In fact, none of the structural features (such as substitution pattern and alkyl chain length) seem to be related to the $p(K_{i(\alpha1)}/K_{i(\alpha2)})$ value. These findings allow for investigating the motif responsible for differences between the two affinity profiles.

In this context, a QSAR model previously built to account for $\alpha_2$-AR affinity[19] highlighted the importance of MW, AC2, AC9, AC27, SssCH2, and SC3C descriptors.[35] In particular, the activity increases as MW, AC2, and AC9 increase and decreases as the other descriptors increase. As a consequence, such a model estimates a raise in p$K_i$ for $\alpha_2$-AR as the molecular weight or alkyl spacer length increases or when a benzodioxan moiety is included in the molecule. The effect of the first two variables is just the same as that calculated by the QSAR model for $\alpha_1$ affinity presented in this paper (RB and CHI descriptors), while the benzodioxane moiety

is generally related to lower affinity toward $\alpha_1$-ARs, as can be noted through a simple SAR analysis of the data set under study (Table 1). Conversely, affinity toward $\alpha_2$-AR decreases whenever the SC3C value increases. The effect of this descriptor on the p$K_i$ for $\alpha_2$-AR can be compared with the relationship found between CHI and $\alpha_1$-AR. Both SC3C and CHI descriptors are representative of molecular connectivity, but they exert an opposite effect on affinity values. On the other hand, SssCH2 affects the affinity in a way such that the higher the number of methylene units (also including those not belonging to the alkyl spacer between the arylpiperazine and the terminal heterocyclic moiety), the lower the p$K_i$ for $\alpha_2$-AR. This descriptor seems to be in close relationship with AC2, both of them being related to the length of the alkyl spacer. Notably, they inversely affect p$K_i$ values. It may be concluded that, if the lengthening of the alkyl spacer, as accounted by AC2, leads to an increase of affinity values for $\alpha_2$-AR, the concomitant but opposite effect exerted by SssCH2 introduces a correcting term so that the further presence of alkyl substituents on the arylpiperazinyl or in the terminal heterocyclic moiety causes a decrease of p$K_i$. Such an effect is not accounted for by the current QSAR model for $\alpha_1$-AR antagonists, simply showing that an increase in the number of rotatable bonds (thus a lengthening of any alkyl chain) leads to greater values for $\alpha_1$-AR p$K_i$ values. Finally, regarding the hydrogen-bonding properties of the molecules under study, the presence of imidazolylpyridazinone and piridinyl-piperazine derivatives, which show hydrogen-bonding acceptor features, seems to decrease $\alpha_2$-AR affinity values. In contrast, the current QSAR model for $\alpha_1$-AR antagonists, together with the previous pharmacophoric model, showed that a limited number of hydrogen-bond acceptors is related to high $\alpha_1$-AR affinity.

In summary, the evaluation and comparison of the QSAR models generated for our set of arylpiperazine derivatives provide additional understanding for $\alpha_1$- and $\alpha_2$-AR selectivity, leading to a few significant conclusions: (i) The length of the alkyl spacer gives a strong signal in correlating with $\alpha_1$- and $\alpha_2$-AR affinity. In fact, longer chains lead to higher p$K_i$ values toward both ARs. (ii) A further introduction of alkyl substituents onto the arylpiperazinyl scaffold has a detrimental effect on $\alpha_2$-AR affinity, while having no consequences on $\alpha_1$-AR affinity. (iii) An increase of the molecular connectivity index values is predicted to increase the affinity toward $\alpha_1$-AR, but not toward $\alpha_2$-AR. (iv) A low number of hydrogen-bond acceptor features seems to be related to high $\alpha_1$-AR affinity but low $\alpha_2$-AR affinity.

The proposed considerations may be useful to address the design of novel adrenoceptor antagonists aimed at obtaining a precise biological outcome upon receptor binding, thus avoiding undesired side effects.

## CONCLUSIONS

A great deal of work is generally done to define ligand key features responsible for high affinity toward a certain target. For this purpose, a classical QSAR approach was applied with the aim of defining chemical and structural properties that characterize $\alpha_1$-AR blockers. Both internal and external validation tests demonstrated the good predictive power of the current QSAR model, which also appeared to be unaffected by overfitting and chance correlations.

The output of this study was also compared and integrated with a previous pharmacophoric model developed for a smaller set of $\alpha_1$-AR antagonists, allowing for better defining details of the SAR for a huge set of arylpiperazine derivatives. In particular, 2D descriptors seemed to be correlated with the affinity profile of this class of compounds, the flexibility, the molecular connectivity, and the hydrogen-bond acceptor features being the key elements significantly defining their binding properties.

Finally, a comparison was done between the current QSAR model for $\alpha_1$-AR antagonists and a QSAR equation, which modeled the $\alpha_2$-AR inhibitory activity for a similar set of compounds. As a result, determinants able to distinguish between $\alpha$-ARs were identified and discussed.

In summary, results from the analysis of a pharmacophoric model and QSAR equations were combined to rationalize the $\alpha_1$- and $\alpha_2$-AR binding profiles of arylpiperazine compounds, leading to the identification of structural and physicochemical requirements for enhancing affinity toward a specific $\alpha$-adrenergic receptor type in a qualitative and quantitative way.

Moreover, the awareness and understanding of the descriptors involved in both the affinity and selectivity properties of our compounds could provide a great opportunity to know the features, where on the ligand structures they occur, and how they affect the biological data upon binding to $\alpha_1$- and $\alpha_2$-ARs. It is clear that SAR information extracted from this kind of analysis could give suggestions for the drug design process. As a consequence, the outcome of this study could be kept as a guide for the further development of arylpiperazinyl compounds with high affinity and selectivity properties toward the target of interest.

## ACKNOWLEDGMENT

We are indebted to Molecular Discovery for the GRID code. We thank Prof. Gabriele Cruciani (University of Perugia) for the use of the program GOLPE in his lab. Prof. Martin G. Ford (University of Porthsmouth) is also acknowledged for allowing us to use the program Minitab.

**Supporting Information Available:** Additional information on biological data and computational details for the development of the QSAR model. This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES AND NOTES

(1) Langer, S. Z. Presynaptic Regulation of Catecholamine Release. *Biochem. Pharmacol.* **1974**, *23*, 1793−1800.

(2) Wood, C. L.; Arnett, C. D.; Clarke, W. R.; Tsai, B. S.; Lefkowitz, R. J. Subclassification of α-Adrenergic Receptors by Direct Binding Studies. *Biochem. Pharmacol.* **1979**, *28*, 1277−1282.

(3) Hoffman, B. B.; Lefkowitz, R. J. α-Adrenergic Receptor Subtypes. *N. Engl. J. Med.* **1980**, *302*, 1390−1396.

(4) (a) Dong, J.; Mrabet, O.; Moze, E.; Li, K.; Neveu, P. J. Lateralization and Catecholaminergic Neuroimmunomodulation: Prazosin, an $\alpha_1$/$\alpha_2$-Adrenergic Receptor Antagonist, Suppresses Interleukin-1 and Increases Interleukin-10 Production Induced by Lipopolysaccharides. *NeuroImmunoModulation* **2002**−**2003**, *10*, 163−168. (b) Artigues-Varin, C.; Richard, V.; Varin, R.; Mulder, P.; Thuillez, C. $\alpha_2$-Adrenoceptor Ligands Inhibit $\alpha_1$-Adrenoceptor-Mediated Contraction of Isolated Rat Arteries. *Fundam. Clin. Pharmacol.* **2002**, *16*, 281−287. (c) Krege, S.; Goepel, M.; Sperling, H.; Michel, M. C. Affinity of Trazodone for Human Penile $\alpha_1$- and $\alpha_2$-Adrenoceptors. *BJU Int.* **2000**, *85*, 959−961. (d) Mayer, P.; Brunel, P.; Chaplain, C.; Piedecoq, C.; Calmel, F.; Schambel, P.; Chopin, P.; Wurch, T.; Pauwels, P. J.; Marien, M.; Vidaluc, J. L.; Imbert, T. New Substituted 1-(2,3-Dihydrobenzo[1,4]dioxin-2-ylmethyl)piperidin-4-yl Derivatives with

$\alpha_2$-Adrenoceptor Antagonist Activity. *J. Med. Chem.* **2000**, *43*, 3653−3664. (e) Quaglia, W.; Pigini, M.; Tayebati, S. K.; Piergentili, A.; Giannella, M.; Marucci, G.; Melchiorre, C. Structure−Activity Relationships in 1,4-Benzodioxan-Related Compounds. 4. Effect of Aryl and Alkyl Substituents at Position 3 on α-Adrenoreceptor Blocking Activity. *J. Med. Chem.* **1993**, *36*, 1520−1528.

(5) Strader, C. D.; Sigal, I. S.; Dixon, R. A. Structural Basis of *β*-Adrenergic Receptor Function. *FASEB J.* **1989**, *3*, 1825−1832.

(6) Strader, C. D.; Sigal, I. S.; Dixon, R. A. Genetic Approaches to the Determination of Structure−Function Relationships of G Protein-Coupled Receptors. *Trends Pharmacol. Sci.* **1989**, Dec. Suppl., 26−30.

(7) (a) Evers, A.; Klabunde, T. Structure-Based Drug Discovery Using GPCR Homology Modeling: Successful Virtual Screening for Antagonists of the $\alpha_{1A}$ Adrenergic Receptor. *J. Med. Chem.* **2005**, *48*, 1088−1097. (b) Leonardi, A.; Barlocco, D.; Montesano, F.; Cignarella, G.; Motta, G.; Testa, R.; Poggesi, E.; Seeber, M.; De Benedetti, P. G.; Fanelli, F. Synthesis, Screening, and Molecular Modeling of New Potent and Selective Antagonists at the $\alpha_{1d}$ Adrenergic Receptor. *J. Med. Chem.* **2004**, *47*, 1900−1918. (c) Kinsella, G. K.; Rozas, I.; Watson, G. W. Computational Development of an $\alpha_{1A}$-Adrenoceptor Model in a Membrane Mimic. *Biochem. Biophys. Res. Commun.* **2004**, *324*, 916−921. (d) Carrieri, A.; Centeno, N. B.; Rodrigo, J.; Sanz, F.; Carotti, A. Theoretical Evidence of a Salt Bridge Disruption as the Initiating Process for the $\alpha_{1d}$-Adrenergic Receptor Activation: A Molecular Dynamics and Docking Study. *Proteins* **2001**, *43*, 382−394. (e) Greasley, P. J.; Fanelli, F.; Scheer, A.; Abuin, L.; Nennniger-Tosato, M.; De Benedetti, P. G.; Cotecchia, S. Mutational and Computational Analysis of the $\alpha_{1b}$-Adrenergic Receptor. Involvement of Basic and Hydrophobic Residues in Receptor Activation and G Protein Coupling. *J. Biol. Chem.* **2001**, *276*, 46485−46494. (f) Montesano, F.; Barlocco, D.; Dal Piaz, V.; Leonardi, A.; Poggesi, E.; Fanelli, F.; De Benedetti, P. G. Isoxazolo-[3,4-*d*]-pyridazin-7-(6*H*)-ones and Their Corresponding 4,5-Disubstituted-3-(2*H*)-pyridazinone Analogues as New Substrates for $\alpha_1$-Adrenoceptor Selective Antagonists: Synthesis, Modeling, and Binding Studies. *Bioorg. Med. Chem.* **1998**, *6*, 925−935. (g) Wetzel, J. M.; Salon, J. A.; Tamm, J. A.; Forray, C.; Craig, D.; Nakanishi, H.; Cui, W.; Vaysse, P. J.; Chiu, G.; Weinshank, R. L.; Hartig, P. R.; Branchek, T. A.; Gluchowski, C. Modeling and Mutagenesis of the Human $\alpha_{1a}$-Adrenoceptor: Orientation and Function of Transmembrane Helix V Sidechains. *Recept. Channels* **1996**, *4*, 165−177. (h) Fanelli, F.; De Benedetti, P. G. Computational Modeling Approaches to Structure−Function Analysis of G Protein-Coupled Receptors. *Chem. Rev.* **2005**, *105*, 3297−3351.

(8) (a) Cavalli, A.; Fanelli, F.; Taddei, C.; De Benedetti, P. G.; Cotecchia, S. Amino Acids of the $\alpha_{1B}$-Adrenergic Receptor Involved in Agonist Binding: Differences in Docking Catecholamines to Receptor Subtypes. *FEBS Lett.* **1996**, *399*, 9−13. (b) Porter, J. E.; Hwa, J.; Perez, D. M. Activation of the $\alpha_{1b}$-Adrenergic Receptor Is Initiated by Disruption of an Interhelical Salt Bridge Constraint. *J. Biol. Chem.* **1996**, *271*, 28318−28323. (c) De Marinis, R. M.; Wise, M.; Hieble, J. P.; Ruffolo, R. R., Jr. In *The $\alpha_1$ Adrenergic Receptor*; Ruffolo, R. R., Jr., Ed.; Humana Press: Clifton, NJ, 1987; pp 211−265.

(9) (a) Lowe, F. C. Role of the Newer Alpha$_1$-Adrenergic-Receptor Antagonists in the Treatment of Benign Prostatic Hyperplasia-Related Lower Urinary Tract Symptoms. *Clin. Ther.* **2004**, *26*, 1701−1713. (b) Milani, S.; Djavan, B. Lower Urinary Tract Symptoms Suggestive of Benign Prostatic Hyperplasia: Latest Update on α-Adrenoceptor Antagonists. *BJU Int.* **2005**, *95*, 29−36. (c) Roehrborn, C. G.; Schwinn, D. A. $\alpha_1$-Adrenergic Receptors and Their Inhibitors in Lower Urinary Tract Symptoms and Benign Prostatic Hyperplasia. *J. Urol. (Hagerstown, MD, U. S.)* **2004**, *171*, 1029−1035.

(10) Lowe, F. C. Treatment of Lower Urinary Tract Symptoms Suggestive of Benign Prostatic Hyperplasia: Sexual Function. *BJU Int.* **2005**, *95*, 12−18.

(11) (a) Steers, W. D.; Kirby, R. S. Clinical Ease of Using Doxazosin in BPH Patients with and without Hypertension. *Prostate Cancer Prostatic Dis.* **2005**, *8*, 152−157. (b) Tanoue, A.; Koshimizu, T. A.; Shibata, K.; Nasa, Y.; Takeo, S.; Tsujimoto, G. Insights into $\alpha_1$ Adrenoceptor Function in Health and Disease from Transgenic Animal Studies. *Trends Endocrinol. Metab.* **2003**, *14*, 107−113. (c) Chalothorn, D.; McCune, D. F.; Edelmann, S. E.; Tobita, K.; Keller, B. B.; Lasley, R. D.; Perez, D. M.; Tanoue, A.; Tsujimoto, G.; Post, G. R.; Piascik, M. T. Differential Cardiovascular Regulatory Activities of the $\alpha_{1B}$- and $\alpha_{1D}$-Adrenoceptor Subtypes. *J. Pharmacol. Exp. Ther.* **2003**, *305*, 1045−1053.

(12) Leiphart, J. W.; Dills, C. V.; Levy, R. M. $\alpha_2$-Adrenergic Receptor Subtype Specificity of Intrathecally Administered Tizanidine Used for Analgesia for Neuropathic Pain. *J. Neurosurg.* **2004**, *101*, 641−647.

(13) Sanacora, G.; Berman, R. M.; Cappiello, A.; Oren, D. A.; Kugaya, A.; Liu, N.; Gueorguieva, R.; Fasula, D.; Charney, D. S. Addition of the $\alpha_2$-Antagonist Yohimbine to Fluoxetine: Effects on Rate of Antidepressant Response. *Neuropsychopharmacol.* **2004**, *29*, 1166−1171.

(14) (a) Cahir, M.; King, D. J. Antipsychotics Lack $\alpha_{1A/B}$ Adrenoceptor Subtype Selectivity in the Rat. *Eur. Neuropsychopharmacol.* **2005**, *15*, 231−234. (b) Grinshpoon, A.; Valevski, A.; Moskowitz, M.; Weizman, A. Beneficial Effect of the Addition of the 5-HT$_{2A/2C}$ and $\alpha_2$ Antagonist Mianserin to Ongoing Haloperidol Treatment in Drug-Resistant Chronically Hospitalized Schizophrenic Patients. *Eur. Psychiatry* **2000**, *15*, 388−390.

(15) Velliquette, R. A.; Ernsberger, P. The Role of $I_1$-Imidazoline and $\alpha_2$-Adrenergic Receptors in the Modulation of Glucose Metabolism in the Spontaneously Hypertensive Obese Rat Model of Metabolic Syndrome X. *J. Pharmacol. Exp. Ther.* **2003**, *306*, 646−657.

(16) *Catalyst*, version 4.8; Accelrys, Inc.: San Diego, CA.

(17) Barbaro, R.; Betti, L.; Botta, M.; Corelli, F.; Giannaccini, G.; Maccari, L.; Manetti, F.; Strappaghetti, G.; Corsano, S. Synthesis, Biological Evaluation, and Pharmacophore Generation of New Pyridazinone Derivatives with Affinity toward $\alpha_1$- and $\alpha_2$-Adrenoceptors. *J. Med. Chem.* **2001**, *44*, 2118−2132.

(18) (a) Betti, L.; Corelli, F.; Floridi, M.; Giannaccini, G.; Maccari, L.; Manetti, F.; Strappaghetti, G.; Botta, M. $\alpha_1$-Adrenoceptor Antagonists. 6. Structural Optimization of Pyridazinone-Arylpiperazines. Study of the Influence on Affinity and Selectivity of Cyclic Substituents at the Pyridazinone Ring and Alkoxy Groups at the Arylpiperazine Moiety. *J. Med. Chem.* **2003**, *46*, 3555−3558. (b) Betti, L.; Botta, M.; Corelli, F.; Floridi, M.; Giannaccini, G.; Maccari, L.; Manetti, F.; Strappaghetti, G.; Tafi, A.; Corsano, S. $\alpha_1$-Adrenoceptor Antagonists. 4. Pharmacophore-Based Design, Synthesis, and Biological Evaluation of New Imidazo-, Benzimidazo-, and Indoloarylpiperazine Derivatives. *J. Med. Chem.* **2002**, *45*, 3603−3611. (c) Barbaro, R.; Betti, L.; Botta, M.; Corelli, F.; Giannaccini, G.; Maccari, L.; Manetti, F.; Strappaghetti, G.; Corsano, S. Synthesis and Biological Activity of New 1,4-Benzodioxan-arylpiperazine Derivatives. Further Validation of a Pharmacophore Model for $\alpha_1$-Adrenoceptor Antagonists. *Bioorg. Med. Chem.* **2002**, *10*, 361−369. (d) Betti, L.; Floridi, M.; Giannaccini, G.; Manetti, F.; Paparelli, C.; Strappaghetti, G.; Botta, M. Design, Synthesis, and $\alpha_1$-Adrenoceptor Binding Properties of New Arylpiperazine Derivatives Bearing a Flavone Nucleus as the Terminal Heterocyclic Molecular Portion. *Bioorg. Med. Chem.* **2004**, *12*, 1527−1535. (e) Betti, L.; Zanelli, M.; Giannaccini, G.; Manetti, F.; Schenone, S.; Strappaghetti, G. Synthesis of New Piperazine-pyridazinone Derivatives and Their Binding Affinity toward $\alpha_1$-, $\alpha_2$-Adrenergic and 5-HT$_{1A}$ Serotoninergic Receptors. *Bioorg. Med. Chem.* **2006**, *14*, 2828−2836.

(19) Salt, D. W.; Maccari, L.; Botta, M.; Ford, M. G. Variable Selection and Specification of Robust QSAR Models from Multicollinear Data: Arylpiperazinyl Derivatives with Affinity and Selectivity for $\alpha_2$-Adrenoceptors. *J. Comput.-Aided Mol. Des.* **2004**, *18*, 495−509.

(20) As an example, some of the arylpiperazines under study are characterized by molecular portions exceeding the pharmacophoric model. They occupy regions of space where any feature of the model lies, and thus, they are not accounted by the model itself. However, such portions could be involved in steric clashes with the receptor counterpart. Such a limitation of the pharmacophoric modeling approach has been, in part, overcome with the introduction of the use of excluded volumes in Catalyst that supposes the (partial) knowledge of the ligand binding site to be applied.

(21) *Cerius2*, version 4.8; Accelrys Inc.: San Diego, CA.

(22) *Minitab*; Minitab Inc.: State College, PA, 2003.

(23) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMm: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.

(24) (a) Smellie, A.; Teig, S. L.; Towbin, P. Poling: Promoting Conformational Coverage. *J. Comput. Chem.* **1995**, *16*, 171−187. (b) Smellie, A.; Kahn, S. D.; Teig, S. L. An Analysis of Conformational Coverage. 1. Validation and Estimation of Coverage. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 285−294. (c) Smellie, A.; Kahn, S. D.; Teig, S. L. An Analysis of Conformational Coverage. 2. Applications of Conformational Models. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 295−304.

(25) Gillet, V. J.; Khatib, W.; Willet, P.; Fleming, P. J.; Green, D. V. S. Combinatorial Library Design Using a Multiobjective Genetic Algorithm. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 375−385.

(26) Rogers, D.; Hopfinger, A. J. Application of Genetic Function Approximation to Quantitative Structure−Activity Relationships and Quantitative Structure−Property Relationships. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 854−866.

(27) Shadow descriptors are calculated by projecting the molecular surface on three mutually perpendicular planes, *XY*, *YZ*, and *XZ*, as described in Rohrbaugh, R. H.; Jurs, P. C. Descriptions of Molecular Shape Applied in Studies of Structure/Activity and Structure/Property Relationships. *Anal. Chim. Acta* **1987**, *199*, 99−109.

(28) Stanton, D. T. On the Physical Interpretation of QSAR Models. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1423−1433.
(29) (a) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis*; John Wiley & Sons: New York, 1986. (b) Kier, L. B.; Hall, L. H. Molecular Connectivity. VII. Specific Treatment of Heteroatoms. *J. Pharm. Sci.* **1976**, *65*, 1806−1809.
(30) Higher CHI values are associated with an increased molecular complexity in terms of both alkyl chain length and substitution pattern. On this basis, such a variable (with a positive sign in the QSAR equation) is able to account for the general trend found for affinity values that are very high for alkyl spacers shorter than an ethyl moiety and decrease for polymethylene chains up to eight carbon atoms.
(31) *GOLPE 4.5.12*; Multivariate Infometric Analyses: Perugia, Italy, 1999.
(32) *GRID 21*; Molecular Discovery Ltd: Pinner, Middlesex, U. K.
(33) Menziani, M. C.; Montorsi, M.; De Benedetti, P. G.; Karelson, M. Relevance of Theoretical Molecular Descriptors in Quantitative Structure−Activity Relationship Analysis of $\alpha_1$-Adrenergic Receptor Antagonists. *Bioorg. Med. Chem.* **1999**, *7*, 2437−2451.
(34) Patané, E.; Pittalà, V.; Guerrera, F.; Salerno, L.; Romeo, G.; Siracusa, M. A.; Russo, F.; Manetti, F.; Botta, M.; Mereghetti, I.; Cagnotto, A.; Mennini, T. Synthesis of 3-Arylpiperazinylalkylpyrrolo[3,2-*d*]-pyrimidine-2,4-dione Derivatives as Novel, Potent, and Selective $\alpha_1$-Adrenoceptor Ligands. *J. Med. Chem.* **2005**, *48*, 2420−2431.
(35) AC descriptors simply report the number of that atom type in a molecule. In particular, AC2 is referred to the number of CH2R2 carbon atom types (where R indicates any group linked through carbon), which is in turn related to the length of the alkyl spacer. AC9, counting the number of carbon atoms belonging to the CHRX2 type (where X represents any heteroatom), is always associated with the presence of the benzodioxane moiety in our set of molecules, whereas AC27 (carbon atom belonging to aromatic heterocycle) is present only in imidazolyl-pyridazinone and pyridinyl-piperazine compounds. SssCH2 is an electrotopological descriptor related to the carbon atom belonging to a methylene unit involved into two single bonds. It is mainly correlated to the alkyl chain acting as a spacer but also to eventual alkyl chains, such as substituents of the arylpiperazinyl moiety. Finally, SC3C (Subgraph Count Index) encodes for a particular cluster of three heavy atoms, thus related to molecule connectivity.

CI060031Z