Journal of Computer-Aided Molecular Design, **15:** 859–872, 2001.
KLUWER/ESCOM

Neuronal nicotinic receptor agonists: a multi-approach development of the pharmacophore

O. Nicolotti[a], M. Pellegrini-Calace[a], A. Carrieri[a], C. Altomare[a], N.B. Centeno[b], F. Sanz[b] & A. Carotti[a,*]

[a]*Dipartimento Farmacochimico, Università di Bari, Via Orabona, 4, 70125 Bari, Italy;* [b]*Research Group on Biomedical Informatics, Institut Municipal d'Investigació Mèdica, Universitat Pompeu Fabra, Carrer Dr. Aiguader 80, E-08003 Barcelona, Spain*

Received 4 April 2001; accepted 10 August 2001

Key words: Catalyst, CoMFA, DISCO, nicotinic agonists, pharmacophore model, QXP

Summary

Based on the results obtained with different automated computational approaches as applied to the study of eleven high-affinity agonists of the neuronal nicotine acetylcholine receptor (nAChR), belonging to different chemical classes, new relevant features were detected which complement the existing pharmacophores. Convergent results from DISCO (Distance Comparison), QXP (Quick Explore), Catalyst/HipHop, and MIPSIM (Molecular Interaction Potential Similarity) allowed us to identify and locate, in a well defined spatial arrangement, three geometrically independent key structural features: (i) a positively charged nitrogen atom for ionic or hydrogen bond interactions, (ii) a lone pair of the pyridine nitrogen or a specific lone pair of a carbonyl oxygen, as a hydrogen bond acceptor, and (iii) a centre of a hydrophobic area generally occupied by aliphatic cycles. The pharmacophore presented herein, along with predictive 2D and 3D QSAR models recently developed in our group, could represent valuable computational tools for the design of new nAChR agonists having therapeutical potential.

Introduction

Nicotine acetylcholine receptors (nAChRs) are members of the ligand-gated ion channel superfamily (LGIC) that play a key role in the signal transmission between cells at the nerve/muscle synapses [1]. They are largely distributed both in the peripheral and central nervous systems [2]. There is increasing evidence from animal and human studies which demonstrates the involvement of nAChRs in high brain functions [3] and in severe neurodegenerative disorders. Indeed, a deficit of nAChRs is considered an important hallmark of Alzheimer's disease (AD), and nicotine itself has been proven beneficial in the symptomatic treatment of AD, showing neuroprotective effects in several *in vivo* models [4]. Preclinical research points to the therapeutic potential of some nAChR agonists also for the symptomatic treatment of Parkinson's disease (PD) [5]. Implications of nAChRs in a number of other relevant physiological and pathological processes, like appetite, schizophrenia, epilepsy, depression and analgesia, have also been suggested [6, 7].

nAChRs are composed of five transmembrane protein subunits, which assemble to form the ion channel [8]. Each subunit consists of four transmembrane α-helical domains (M_1-M_4). The inside wall of the channel is made up of five α-helices coming from the M_2 domains of each five subunits. Various subunits (α, β, γ, δ and ε) have been identified in different nAChRs subtypes [4]. In the mammalian skeletal muscle the pentameric structure is constituted by two α-subunits and one each of β, δ, and γ (ε). In the central nervous system (CNS) of rodents two main classes [2] of nAChRs have been identified by autoradiographic analysis and regional binding studies, namely recep-

*To whom correspondence should be addressed. E-mail: carotti@farmchim.uniba.it

tors with high affinity for acetylcholine (ACh), nicotine and cytisine (agonist sites), and receptors with high affinity for the antagonist α-bungarotoxin (Bgt) but low affinity for ACh and nicotine (α-Bgt sites). The first class of nAChRs (agonist sites), predominant in the CNS, shows mainly $\alpha_4\beta_2$ subunits, whereas a homooligomeric $5\alpha_7$ combination characterises the α-Bgt sites. Other neuronal nAChRs, classified as no $\alpha_4\beta_2$ and no α_7, have been also identified.

Many agonist sites are localised at the presynaptic level and are thought to modulate the release of neurotransmitters, such as ACh, GABA, glutamic acid and dopamine, involved in learning and memory processes. α-Bgt sites are involved in cognition and neuroprotection, although their CNS distribution is still unclear.

The huge structural and functional complexity of nAChRs has stimulated a vast research aimed at finding out their 3D arrangement and therefore to enhance the understanding of the structure-function relationships. Unfortunately, 3D information on nAChRs [9] is still defective to be used in the structure-based design of new ligands. Actually, in recent years, some 3D data on the overall topologies of nAChR have been reported for nAChRs of the skeletal muscle at the neuromuscular junction ($2\alpha_1\beta_1\delta\gamma/\varepsilon$), but not for neuronal nAChRs, which have been instead widely investigated by molecular modelling approaches, voltage-clamp, photoaffinity labelling, site-directed mutagenesis and other molecular biology techniques [10].

During the last forty years, several investigations have been performed to define nicotinic pharmacophores. Among the most important ones, Beers and Reich [11] found that nicotinic agents (Figure 1, antagonists **1–3** and agonists **7–8**) are characterised by an onium moiety and a hydrogen bond (HB) acceptor atom, the distance between them being 4.8 Å. Sixteen years later, Sheridan and co-workers examined four nicotinic agonists (Figure 1, cmps **6–9**) and developed a pharmacophore by using the Distance Geometry approach [12]. They found three salient features, i.e., a basic nitrogen atom (corresponding to the pyrrolidine nitrogen in nicotine), a HB acceptor atom (e.g., the pyridine nitrogen of nicotine or the carbonyl oxygen of cytisine), and a third point representing the centroid of the pyridine ring of nicotine or the carbonyl carbon atom of cytisine or of other agonists. Based on X-ray crystallographic data of two nicotinic ligands (**7** and **8**), in 1989 Barlow and Johnson [13] suggested that the agonist activity requires a charged nitrogen (i.e., an onium site) and a planar area on the receptor able to recognise an aromatic ring or a double bond through π-π stacking or hydrophobic interactions. A further relevant contribution in this field has been given by Livingstone et al. in 1996 [14], who analysed seven semirigid agonists (Figure 1, cmps **4–5** and **7–11**) through a comparative analysis of the gnomonic projections of the surface molecular properties and developed a pharmacophore showing the following key features: (i) a cationic head, (ii) a centroid of a pyridine ring or the carbon atom of a carbonyl group, and (iii) a dummy atom indicating the location on the receptor of a functional group likely to make a HB with a pyridyl nitrogen or a carbonyl oxygen. A lipophilic region, close to the positions 3′ and 4′ of the pyrrolidine ring of nicotine, was considered as an important modulator of the agonistic activity, but had not been taken into account in the derivation of the pharmacophore model. Just one year ago, Barreiro et al. [15] published an interaction model, based on fundamental cation–π and hydrogen bond interactions, that is obtained by using a simplified minireceptor modelling technique. A quantitative model of the interaction energies, including the desolvation energy of the ligands, has been proposed for all ligand–minireceptor complexes examined. More recently, Tonder et al. [16] have reported a pharmacophore for agonists at the $\alpha_4\beta_2$ nAChRs, that can be reasonably considered as a *trait d'union* among all those previously published. Three pharmacophoric features have been determined, i.e., (i) a site point (**a**), corresponding to a protonated nitrogen atom, (ii) a site point (**b**) corresponding to an electronegative atom capable of forming a HB, and (iii) the centre of a heteroaromatic ring or of a C=O bond (**c**). These features have to be within narrow ranges of interatomic distances (7.3–8.0 Å and 6.5–7.4 Å for **a–b** and **a–c** distances, respectively) and angles (the angle measured between the interatomic distance vectors **a–b** and **a–c** should range from 30.4 to 35.8°). As a reliability test of the proposed pharmacophore model (i.e., the self-consistency of ligand alignment), the same authors have carried out a Comparative Molecular Field Analysis (CoMFA) on the whole set of 19 agonists but poor PLS models, both in terms of fitting and predictive capacity have been obtained. Better results came only upon the application of the region focusing and the dropping of a certain number of ligands.

Despite the efforts to derive consistent nAChR pharmacophores, it was our opinion that the models proposed so far suffer from some limitations, such as, for example: (i) the mixing of agonists and antagonists

in the sets of Beers and Reich [11] and Tonder [16]; (ii) the relatively low number of ligands analysed in most of the previous studies; (iii) related geometric features taken as independent pharmacophoric points in the set of Sheridan [12] and, though to a lesser extent, in the Livingstone's [14] and Tonder's [16] sets; (iv) no definition of the hydrogen bond directionality, especially for carbonyl-containing ligands.

All the above prompted us to undertake a new investigation of the pharmacophore for nAChR agonists, starting from the examination of a larger set of ligands and taking into account features physicochemically well defined and truly unrelated from a geometrical point of view. For this purpose an exhaustive perusal of literature data was carried out and a set of agonists, selected according to the principles of the experimental design [17], was analysed by different computational approaches.

Methods

Binding data and selection of nAChR ligands

Only binding affinity data (espressed as K_i) determined on rat brain membranes were considered. K_i values coming from different sources were normalised with respect to the K_i value of 1.41 nM reported by Damaj et al. [18] for (*S*)-nicotine. The selection of suitable ligands was done according to the following criteria: (i) high binding affinity ($pK_i \geq 8.15$), (ii) regular distribution of data over the range of affinity explored, (iii) presence of the same molecules already investigated in previous pharmacophore modelling (**7**, **8** and **10**) or of slightly modified molecules (**13** and **19**), and (iv) structural diversity and limited molecular flexibility (**12, 14, 15, 16, 17** and **18**). According to these criteria, eleven nicotinic agonists (**7, 8, 10**, and **12–19** in Figure 1) were chosen among 90 high activity candidates ($pK_i \geq 8.15$) present in a large data set recently analysed by us trough 2D and 3D QSAR studies [19]. Such a selection was made looking at the diversity of the 90 ligands in a property hyperspace defined by the first three principal components obtained from the Principal Component Analysis (PCA) of a matrix of 235 molecular descriptors, calculated with the program Dragon 1.0 [20], which account for constitutional, geometrical, topological, electronic, aromaticity, and other empirical properties. Actually, Dragon calculates 853 descriptors, and then automatically reduces the descriptor matrix (235 columns in our case), by eliminating constant variables and highly interrelated variables, the latter ones detected by a multivariate correlation analysis. In our study, PCA of the Dragon-computed descriptors was performed to be guaranteed of adequately sampling the whole space of physicochemical properties of the most active agonists.

Computational tools

Molecular models of the protonated ligands **12–19** were constructed with standard bond distances and angles using the SYBYL software package (Ver. 6.6; Tripos Assoc. St. Louis, MO, USA). For ligands **7, 8**, and **10** crystallographic data were retrieved from the Cambridge Structural Database [21] (ref. code NICOTI, FITPIH and DIZPUX, respectively). Full geometry optimization was performed for all the selected ligands with the PM3 Hamiltonian using the parameter set included in the MOPAC (ver. 6.0) suite of programs [22].

The software DISCO and Catalyst/HipHop, used for pharmacophore development, are distributed by Tripos Assoc. St. Louis, MO-USA and Molecular Simulations Inc., San Diego, CA-USA, respectively. QXP, the other program utilized for pharmacophore development, is part of the Flo96 software package available upon request from Colin Mc Martin (cmma@ix.net.com) 302 suite 1B, Lindsley Drive Morris Township, NJ 07690, USA. The progam MIPSIM Ver 1.1 utilized for the comparison of Molecular Interaction Potentials, has been developed by F. Sanz and co-workers at the University of Pompeu Fabra, Barcelona (Spain) and can be downloaded by non profit institutions from the following URL http://www1.imim.es/modeling/mipsim .

All the calculations were performed on a Silicon Graphics O2 R10000 workstation.

Results and discussion

Initialising nAChR pharmacophore development

In a search for the key molecular determinants of high affinity, we first recovered the cationic centre, present in all the previous models, and then, as the second essential feature we looked for both a more precise definition and a better spatial location of a HB acceptor site. As a third relevant feature, geometrically unrelated with the previous ones, we searched for a common hydrophobic region. Indeed, similarly placed

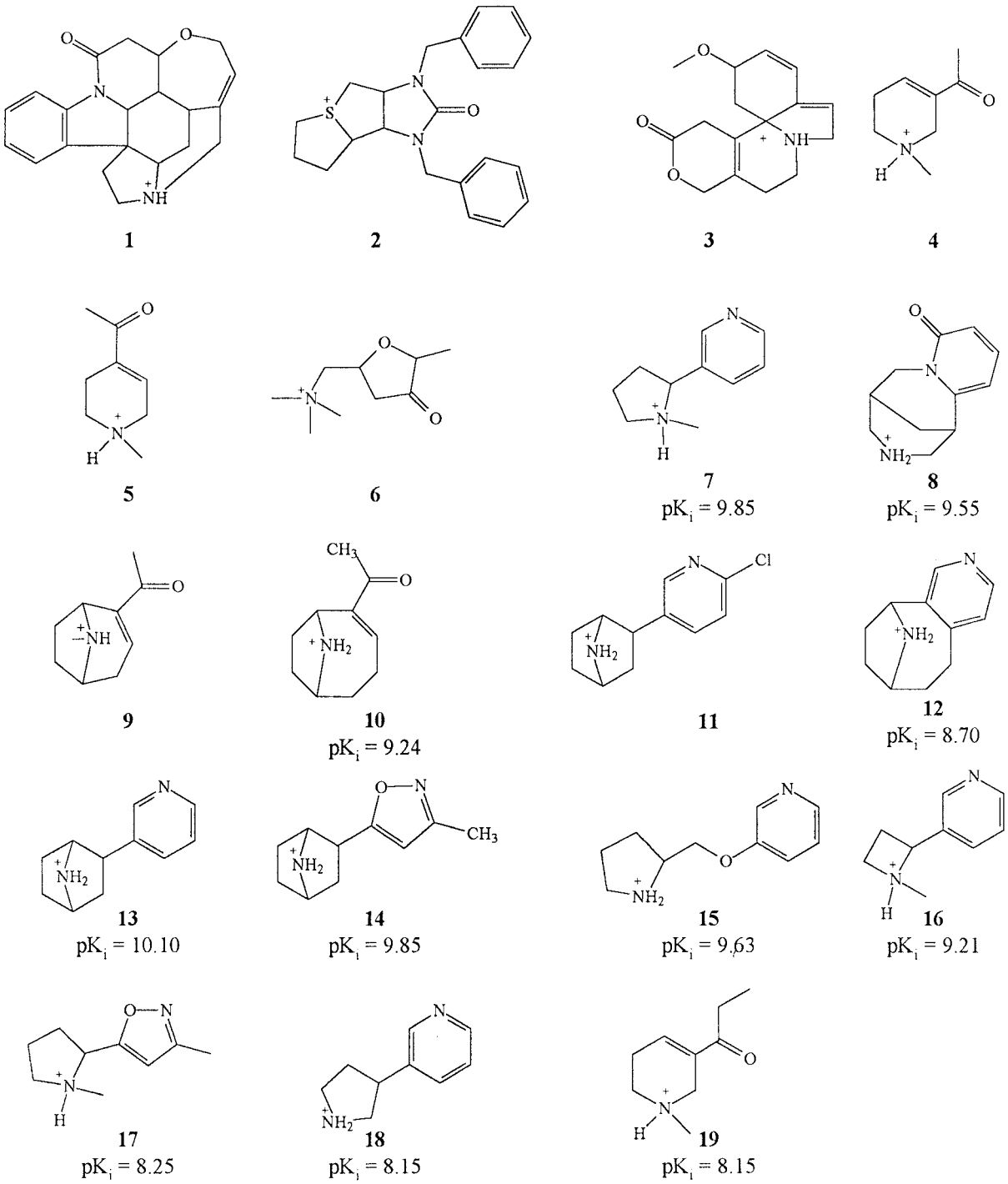

Figure 1. nAChR ligands: strychnine (**1**), trimethapan (**2**), β-erythroidine (**3**), arecolone (**4**), isoarecolone (**5**), muscarone (**6**), (*S*)-nicotine (**7**), (-)-cytisine (**8**), ferruginine (**9**), (+)-anatoxin a (**10**), epibatidine (**11**), pyridohomotropane (PHT) (**12**), (1*R*, 2*R*, 4*S*)-deschloroepibatidine (**13**), (1*R*, 2*S*, 4*S*)-epiboxidine (**14**), (*S*)-nor-A-84543 (**15**), (*S*)-2-(3-pyridyl)azetidine (**16**), (*S*)-ABT-418 (**17**), (*S*)-isonornicotine (**18**), homoarecolone (**19**); chiral configurations and pK_i values are reported only for ligands used for the development of our pharmacophore models.

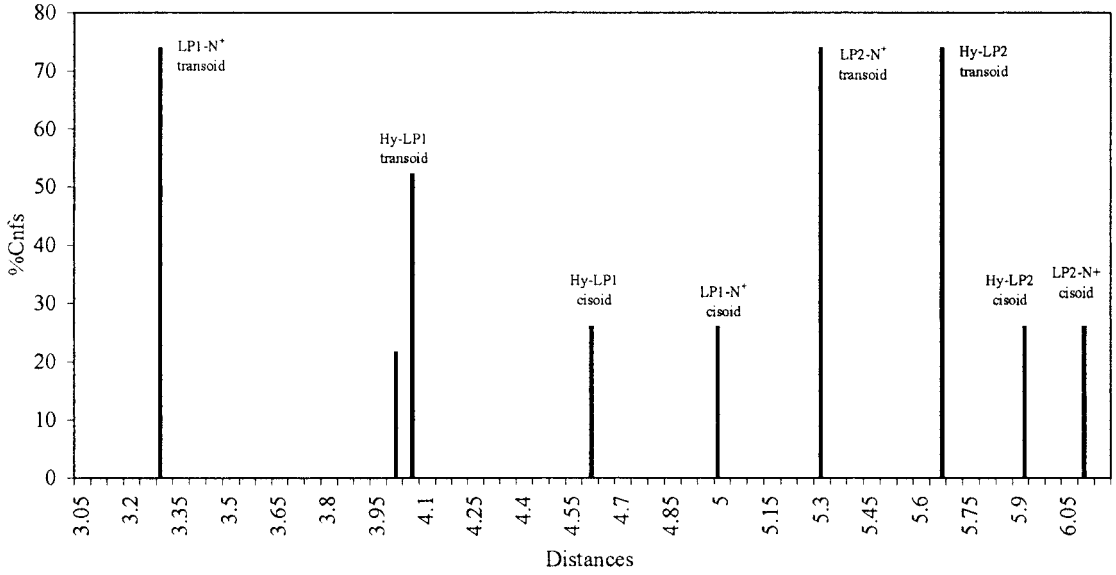

Figure 2. Conformers population (%) and distance (Å) distribution of (+)-anatoxin a **10**.

hydrophobic areas were found in all the examined ligands. They were easily located through their mass centre (centroid) defined by all the heavy atoms of tetrahydropyridine, pyrrolidine and azetidine rings and of the nitrogen bicyclic systems.

A careful search for possible HB acceptor sites was carried out by considering the aromatic nitrogen and carbonyl oxygen lone pairs. As for the carbonyl compounds, both the oxygen lone pairs, called LP1 and LP2 to indicate their vicinal or distal position relative to the cationic centre, respectively, were considered as possible HB acceptor sites. The selection of one of the two lone pairs was done on the basis of their geometrical relationship with all the other pharmacophoric elements. Lone pairs were first generated with default geometries in SYBYL [23] and then modified to hydrogen bond extension vectors (HBVs) according to Vedani et al. [24]. The heteroatom-lone pair distances were therefore adjusted to 1.8 Å and the C=O-HBV valence angle to 135°, to closely mimic the ideal HB geometries observed in X-ray crystallographic structures [25].

A preliminary investigation of the spatial arrangements of putative pharmacophoric points was performed by means of a conformational analysis of the flexible ligands. All the torsion angles were taken into account, allowing an increment of 10° for each rotatable bond, including those involved in ring closure. The AMBER* force field [26] and the Powell-Reeves conjugate gradient (PRCG) [27] were used for energy minimization. To examine a large conformational space, conformations within an energy window of 12 kcal/mol were retained. In the conformer population of compounds bearing a rotatable carbonyl group (**10** and **19**), lower values of standard deviation (sd) and coefficient of variation (cv) were recorded when measuring the mean distances to LP2 than to LP1. The total count of the conformers carried out at the end of the conformational search indicated that the transoid state of the two ligands (**10** and **19**) was the most populated, being the transoid-cisoid ratio approximately 7:3 and the energy window ΔE of 8.403 kcal/mol for cmpd **19** (129 conformers) and of 7.469 kcal/mol for cmpd **10** (23 conformers).

The same distances between the supposed pharmacophoric points were then measured for all the conformers generated for each ligand. Interestingly, the mean distances between the distal lone pair LP2 and the other pharmacophore elements always presented values of sd and cv lower than the corresponding distances involving the vicinal lone pair LP1. The results are reported in detail in Table 1, where the conformation population size is under the column head 'cnfs'. The average distances between the hypothesised pharmacophoric elements, i.e., the charged nitrogen (N^+), the pyridyl nitrogen (N), the vicinal (LP1) and distal (LP2) oxygen carbonyl lone pairs, the hydrophobic centroid (Hy) are reported in the other columns (Table 2). The average distance N-N^+ calculated for the

Table 1. Total energies and energy components (kcal/mol) for (+)-anatoxin a **10** minimum energy conformers in chloroform and water (solvent continuum method)

	Etot	Eele	Evdw	Estr	Ebend	Etors	Esolv
Transoid state							
Boat ($CHCl_3$)	−17.282	0.208	2.811	0.564	6.176	14.134	−41.195
Boat (water)	−42.559	0.466	2.748	0.599	6.078	14.209	−66.718
Chair ($CHCl_3$)	−18.659	−0.129	2.247	0.490	6.097	12.937	−40.370
Chair (water)	−44.438	0.186	2.442	0.602	6.025	12.696	−66.436
Cisoid state							
Boat ($CHCl_3$)	−14.144	1.498	2.464	0.514	5.982	17.717	−42.380
Boat (water)	−40.521	1.632	2.567	0.576	5.954	17.598	−68.893
Chair ($CHCl_3$)	−14.923	2.089	1.305	0.378	5.860	17.361	−40.029
Chair (water)	−42.082	2.342	1.685	0.490	5.927	16.814	−69.395

Table 2. Average (avg) distances among the indicated pharmacophore features, as calculated from the analysis of the selected conformers (cnfs), and associated statistical parameters sd (standard deviation) and cv (coefficient of variation)

cmps	cnfs	N-N^+	LP1-N^+	LP2-N^+	Hy-LP1	Hy-LP2	Hy-N^+
13	36	4.99	6.12		6.32		1.34
14	28	4.43	5.61		6.19		1.34
15	835	5.96	7.17		7.50		1.26
8	1	4.89	5.13	6.54	4.30	5.77	1.77
10	23	3.87	3.77	5.52	4.20	5.72	1.49
16	16	4.59	5.91		6.06		0.99
7	16	4.54	5.85		6.09		1.28
12	1	4.53	5.94		5.76		1.56
17	46	4.46	5.99		6.17		1.28
19	129	4.41	4.46	6.15	4.07	5.45	1.43
18	12	4.99	6.04		5.73		1.27
avg		4.70	5.64	6.08	5.67	6.07	1.36
sd		0.52	0.91	0.45	1.06	0.54	0.20
cv		11.18	16.09	7.42	18.72	8.93	14.65

selected conformer population of each ligand is nearly 4.70 Å, a value close to that found by Beers and Reich.

In a susbequent step, further studies were addressed to the identification of the bioactive orientation of the enone group of (+)-anatoxin a **10** and homoarecolone **19**. Previous conformational analyses, carried out by molecular mechanics (COSMIC force field) by Thompson and coll. [28] on (+)-anatoxin a **10**, pointed out four minimum energy conformers, deriving from the combination of chair and boat conformations with the *s-cis* and *s-trans* enone geometries. The chair conformations were found to be more stable than the boat ones, with the *s-trans* chair conformer showing the lowest energy. The energy difference between the two chair conformations resulted almost entirely from the greater stabilising electrostatic interaction between the oxygen carbonyl and the protonated nitrogen that were closer in the chair-trans (3.68 Å) than in the chair-cis (4.27 Å) conformations. A similar behaviour was observed for the boat conformers. It is worth noting that ^{1}H-NMR studies in $CDCl_3$ had shown that the *s-cis* and *s-trans* forms of the chair conformers were almost equally populated, most likely due to a possible energy leveling effect of the solvent. To verify whether this effect increases as the dielectric constant of the solvent increases, we

carried out a further conformational study, using the *solvent continuum* method (chloroform and water) of MACROMODEL [29] and OPLS [30] force field for energy calculations. The results summarized in Table 3 confirm that the chair conformers are more stable than the boat ones and that as the solvent polarity increases the energy difference decreases from 3.14 to 2.04 and from 3.74 to 2.35 kcal/mole for the boat and chair conformers, respectively. These data suggest that in the physiological medium the enone system could indifferently assume *s-cis* and *s-trans* arrangements as the bioactive conformations.

QXP (Quick EXPlore) studies

QXP (Quick EXPlore) is a module of Flo98, a software package for structure-based drug design including fast and efficient algorithms for flexible docking and template fitting [31, 32]. For template fitting QXP uses a procedure (TFIT) based on a full conformational search and similarity match (fitting) of two or more molecules simultaneously. QXP assigns automatically short-range attractive forces between similar atoms (polarity, charge and hydrogen bond ability) in different molecules. In the template fitting procedure (molecular superposition), the normal intramolecular non-bonded energies are replaced by the superposition energies (E_{sup}), while internal energies (E_{int}) are calculated by the normal force field ignoring non bonded energies. The combined minimisation of these two energies (E_{sup} and E_{int}) yields structures with optimal superimposition and relatively low internal energy. Within a given energy range, the program finds different solutions ranked according to their total energy ($E_{tot} = E_{sup} - E_{int}$).

The template fitting module of QXP was applied to the analysis of our set of nicotinic agonists, the most active ligand, deschloroepibatidine **13**, being taken as the reference compound. Fifty solutions were found within an energy window of 3.59 kcal/mol. The best fitting model showed E_{tot} and E_{sup} values of -817.7 and -834.6 kcal/mol, respectively. It is worth noting that such a model pointed out the same pharmacophore features found by us when initialising the nAChR pharmacophore model. In particular, the distal lone pair of the carbonyl groups (LP2, coloured in magenta in Figure 3) pointed toward the same region of the pyridyl lone pairs. The fact that the 'longer' pyridyl ether ligand **15**, compared to other ligands, is slightly displaced would suggest a different binding mode or a certain flexibility of the receptor binding site

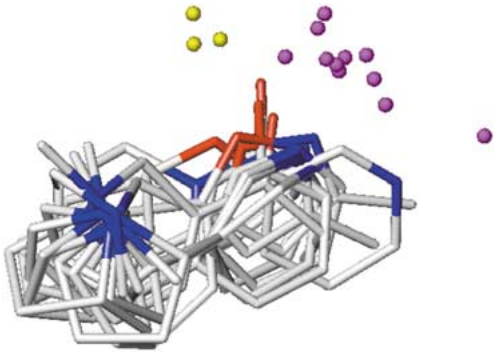

Figure 3. Overlay of the pharmacophore set of ligands as determined by QXP using deschloroepibatidine **13** as the reference compound. Distal lone pairs of carbonyl oxygens (LP2) and pyridyl nitrogens are shown as magenta balls. Vicinal lone pairs of carbonyl oxygens (LP1) are represented as yellow balls.

Table 3. Average distances between the pharmacophore features and corresponding statistical parameters as calculated after QXP runs

	$N\text{-}N^+$	$LP2\text{-}N^+$	Hy-LP2	$Hy\text{-}N^+$
avg	4.81	6.31	6.08	1.36
sd	0.70	0.78	0.72	0.21
cv	14.49	12.36	11.84	14.62

in accommodating ligands of different size. A certain dispersion of the points indicating the positions of the unshared lone pairs may be ascribed, at least in part, to the fact that QXP looks for the better superposition of similar atom types and not of their lone pair vectors. As for the ligands bearing an enone system, the QXP search algorithm selected the transoid over the cisoid geometry as the best models, because of its slightly lower energy.

It must be observed that the mean distances between the cationic centre and the HB acceptor nitrogen atom ($N\text{-}N^+$) in the model is 4.81 Å, a distance superimposable to that of the Beers and Reich pharmacophore.

DISCO (DIStance COmparison) studies

The eleven ligand conformers selected for pharmacophore development were subjected to a further comparative analysis by the software DISCO [33] (an acronym for Distance Comparison), a powerful tool to derive pharmacophores. The best matches among all ligands is achieved through the clique detection algorithm of Brint and Willett [34]. The pharmacophore

features are taken from both the ligand and the receptor. In order to derive a putative pharmacophore for our set of molecules we followed the classical DISCO steps: (i) setting up a molecular spreadsheet containing the eleven active ligands; (ii) multisearch analysis, using default options, to produce a set of conformers (up to 25) for each ligand; (iii) feature assignments to each conformer; (iv) selection of the reference compound; (v) progressive increasing of feature size; (vi) model computations; (vii) statistical and features analysis of all models found; (viii) selection and refinement of the best model allowing the mismatch of at most one ligand and (ix) limiting the conformer selection for each ligand within lower energy window.

The most active ligand in our set, deschloroepibatidine (**13**), was once again taken as the reference compound. Firstly, we looked for the lowest number of common features for identifying quickly a pharmacophore hypothesis. Several useless models showing good statistics but lacking physicochemical meaning were found. By increasing the feature-size DISCO derived a good model (*Model 169* in Table 4), with four pharmacophore features, namely a donor atom centred on the positive nitrogen and a corresponding acceptor site on the receptor, an acceptor atom consisting of an aromatic nitrogen or a carbonyl oxygen and the respective donor site on the receptor. In order to find out other chemical features possibly responsible for high binding affinity, we gradually increased the feature size from five to eight. DISCO selected *Model 233* as the best one showing an improved definition of the potential pharmacophore site points even if the statistical parameters are not as good as we would have liked. This model added to the previous one a new feature, i.e. a hydrophobic centroid representing a cyclic aliphatic moiety. To improve the statistical results of *Model 233*, we allowed DISCO to mismatch at least one ligand. DISCO mismatched ligand **16**, that is the only one having an azetidine ring, and found *Model 337* as the best one. This model presents improved statistics when compared to *Model 233*, while keeping the same pharmacophore features. As the last refinement, we allowed DISCO to select conformers for each ligand with an energy penalty up to 2.5 kcal/molÅ over the absolute minimum energy conformer. DISCO derived *Model 341*, which is still fully compatible with our starting models. Five essential features were found by DISCO: a hydrophobic centre (Hy), a donor atom (DA) and an acceptor atom (AA) on the ligand and, for the last two, the corresponding acceptor (AS) and donor (DS) sites on the receptor.

Table 4. DISCO models (**13** as the reference compound) and parameters

Model	Size	ndropped	tol max	d mean
169	4	none	1.5	6.46
233	5	none	3.5	5.80
337	5	**16**	1.5	5.70
341	5	**16**	1.5	5.11

Table 5. DISCO parameters for *Model 341* (see text)

Cmps	Cnfs	Fit[a]	Overlap[b]	ΔE[c]	Tolerance[d]
13	1	0	146.5	0	0
14	5	0.38	116.62	0	0.58
15	2	1.14	69.5	1	0.58
8	1	0.75	113.12	0	0.75
10	1	0.60	85.88	0	0.48
7	10	0.53	87.5	1.37	0.47
12	1	0.70	86.75	0	0.48
17	5	1.38	59	2.43	0.8
19	3	0.91	81.12	0.90	0.95
18	24	0.85	70.12	1.08	0.48

[a]The fit is expressed as the r.m.s. deviation of feature coordinates from those of reference **13**.
[b]The overlap volume represents the highest intersection volume of the selected conformer with **13**.
[c]ΔE is the energy difference (kcal/mol) between the selected conformer and the lowest energy conformer found by a Multisearch analysis.
[d]The tolerance is a measure of the deviation of the distances between the features in the ligand and in the reference compound **13**.

All the DISCO models identified: a) the distal lone pair, along the axis AA-DS, as possible HB acceptor site; b) the positive nitrogen as a HB donor or coulombic interaction site and c) the mass centroid, indicative of hydrophobic interactions. *Model 341* was then selected as the most significant and informative pharmacophore model. Its main features are represented in a graphical form in Figure 4. The enone conformations selected by DISCO for ligands **19** and **10** are transoid and cisoid respectively.

In summary, *Model 341* derived by DISCO showed the following characteristics: 5 features, tolmax = 1.5, Dmean = 5.11 and one mismatch (**16**). Tolmax is a measure of the deviation of the distances between the features in the ligand and in the reference compound **13** and Dmean is the average interpoint distance.

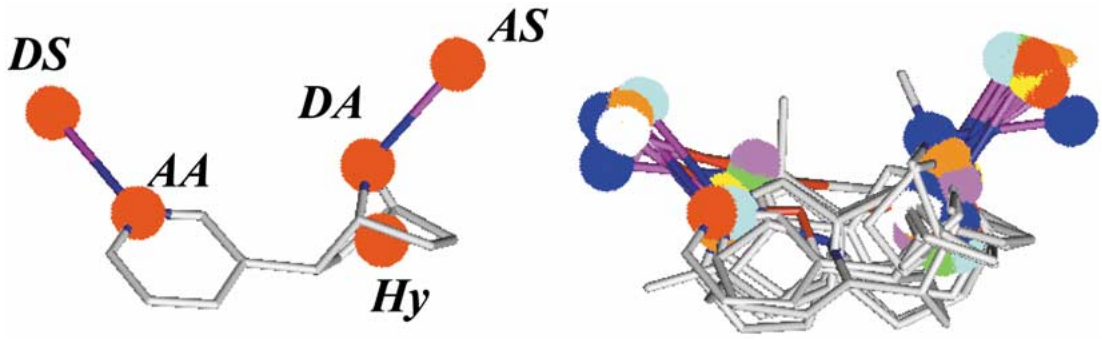

Figure 4. Left: representation of the pharmacophore points identified by DISCO on deschloroepibatidine **13** selected as reference compound; DS, DA, AS, AA and HY stand for donor site, donor atom, acceptor site, acceptor atom and hydrophobic centre respectively. The pharmacophore ligands, as superimposed by DISCO, are reported on the right-hand side. Different colours are associated with the spheres representing the pharmacophore features of the different nAChR agonists.

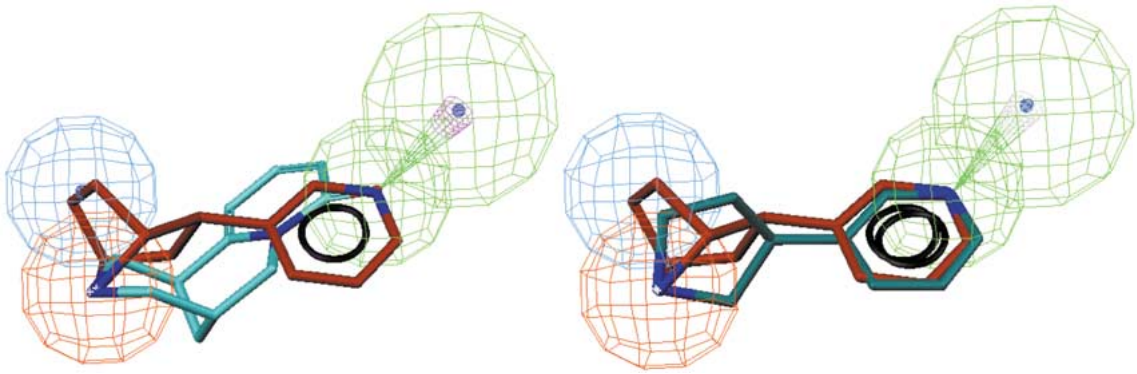

Figure 5. Superimposition of **8** (coloured in cyan, on the left) and **18** (coloured in green, on the right) onto **13** (coloured in red). The grid-sphere contours, coloured in red, blue and green, indicate ionic, hydrophobic and hydrogen bond interaction areas, respectively. The projecting cones show the HB directionality and the small spheres HB donor sites on the receptor.

Catalyst/HipHop studies

The Catalyst/HipHop program [35, 36] is a powerful tool in pharmacophore generation. With the aim of exploring the conformational space available to each molecule, Catalyst collects conformers using a poling technique. As far as the feature detection is concerned, HipHop contains, by default, definitions for HB donors and acceptors, negative and positive charge centres, aliphatic accessible surface and aromatic hydrophobic regions. HB donors and acceptors are coded as vectors with two ends: the hydrogen bonded atom and the projected point on a complementary site atom whose position depend on the HB directionality. In our study, HipHop was applied in three steps: (i) generation of a conformational model for each agonist structure in the pharmacophore set; (ii) feature detection and (iii) identification of the three dimensional configuration of chemical features shared by all molecules in the set. Conformational models were generated by the module Confirm, a 'quasi-exhaustive' search module based on a modified CHARMM force field provided with a poling algorithm for covering the conformational space [37]. The poling algorithm adds penalty functions to the force field, based on the chemical features (termed as poles). By using the 'best conformer analysis' implemented in Catalyst, a representative collection of conformers was generated within a 20 kcal/mol range of the computed minimum energy conformer for each molecule of our set [38].

The following features were selected by Catalyst as the starting point for generating the HipHop hypothesis: (a) a HB acceptor (HBA), (b) a hydrophobic aliphatic site (HY), (c) a positive charge (PC) [39]. The whole ligand pharmacophore set, together with its conformational population models, was submitted to HipHop with the purpose of identifying a common feature hypothesis of Catalyst. Among the several hypotheses generated, one emerged as consistent for

eight molecules of our set. Three molecules (namely **12**, **16** and **18**) were left out when deriving such a hypothesis since they, unexpectedly, were not satisfactorily mapped by Catalyst. However, even if these three compounds were not used in the hypothesis development, we were able to prove that ligands **16** and **18** matched very well the model, as can be seen on the right-hand side of Figure 5, where **18** is superimposed on **13**. On the other hand, the conformer population of **12** presented low energy hits that quite well fit the pharmacophore hypothesis. In other words, the mismatch of ligands **12, 16** and **18** by catalyst is surprising indeed and deserves further investigations.

Due to the relatively small size of some of our molecules, the 'minimum permitted interfeature spacing' was reduced to 0.5 Å from the default value of 3 Å, the latter being more suitable for large sized molecules. As can be seen in Table 6, the most active ligand in our set (**13**), was chosen as the reference compound setting a *Principal* value of 2 and a *MaxOmitFeat* value of 0. A *MaxOmitFeat* value of 0 associated with the reference compound forces it to map all the features of each pharmacophore hypothesis whereas a *MaxOmitFeat* value of 1 means that the compound can lack, at most, one feature only. When *Principal* is set to 2, all the pharmacophore features of the compound are used to define the initial set of hypotheses, whereas compounds are allowed to map only the pharmacophore hypothesis when *Principal* is set to 1.

In a second step, a further refinement of our pharmacophore was sought by improving the feature mapping. A double feature-weight was assigned to HBA compared to HY and to PC, being HY and PC too close in space. In doing so, a more valuable definition of the feature mapping was worked out. The HipHop outcomes confirmed the validity and robustness of our pharmacophore model. In fact, according to our previous results, PC, HY and HBA were identified as relevant points of interaction. Ammonium sites represent critical points for ionic interactions and were well identified by HipHop. Moreover, in full agreement with our expectations, the hydrophobic region is better defined in bicyclic systems than in molecules having only one aliphatic ring. Molecules bearing aromatic rings accommodated HBA on aromatic nitrogens whereas compounds with enone systems showed, as potential HBA, the carbonyl oxygen. The HBA is located close to the LP2 in the case of (-)-cytisine **8**, as shown on the left hand side of Figure 5, and along the axis of the C=O bond when considering compounds **19** and **10**. As for ligand **8**, such a result indicates the distal lone pair LP2 as an accessible point for HB interaction and this perfectly fulfills our expectations. With regard to ligands **19** and **10**, the lack of a decisive preference for one carbonyl lone pair resulted in a poorer HB directionality.

Generating nAChR pharmacophore model through MIPSIM studies

The main results obtained by applying different methods to the pharmacophore development are summarised in Table 7. Our interaction model for agonist nicotinic ligands is constituted by three main features placed at relatively constant distances and angles as assessed by the three diverse approaches used (DISCO, QXP and ours).

Besides the geometrical-based investigation of the pharmacophore features, we attempted to generate a property-based pharmacophore, by assessing the molecular interaction potential (MIP) similarity between the deschloroepibatidine **13**, chosen as the reference, and the other nAChR agonists in the dataset. MIPs have been extensively used for the comparison of series of compounds in structure-activity relationship studies [40]. In this study, we wanted to prove whether the best MIP superposition resulted also in a common feature alignment over the whole series of nAChR agonists, as it should be the case when a reliable pharmacophore model does exist.

The molecular charge distribution and geometry of the eleven nAChR agonists was optimised by means of *ab initio* calculations based on the STO-3G basis set, using the software GAUSSIAN94 [41]. The molecular models were then submitted to MIPSIM (Molecular Interaction Potentials SIMilarity analysis) [42], a powerful software tool for the analysis and comparison of MIPs. MIPSIM includes modules that perform analogous tasks to those performed by the previous software MEPSIM [43], i.e. MIPMIN to find MIP minima and MIPCOMP to compare MIP distributions of pairs of molecules and to search for the alignments that maximise similarity. Nevertheless, MIPSIM has interesting new features in comparison to MEPSIM: the program transparently integrated with other programs, like GAMESS [41] or GRID [44], which provide one with the computation of the potentials to be analysed or compared, or with programs that allow one the statistical analysis of the results, as is the case for GOLPE [45].

The molecular interaction potentials of each nAChR agonist was measured by scanning its sur-

Table 6. Catalyst/HipHop parameters used for the hypothesis generation.

Cmps	Principal[a]	MaxOmitFeat[a]	Cnfs[a]	Features/Cnfs[a]
13	2	0	12	3.00
14	1	1	14	4.93
15	1	1	56	3.80
8	1	1	8	4.75
7	1	1	12	3.00
17	1	1	17	4.53
10	1	1	14	5.36
19	1	1	28	5.14

[a]Principal, MaxOmitFeat and Features/Cnfs values define the weight of compounds when an hypothesis is generated, the number of features that can be missed and the total number of features divided by the number of conformers, respectively (see text).

Table 7. Comparison among pharmacophores for nAChR agonists developed in this study

Method	Distances				Angles		
	N-N^+	LP2-N^+	Hy-LP2	Hy-N^+	^LP2	^Hy	^N^+
Ours	4.70	6.08	6.07	1.36	12.09	82.96	84.94
DISCO[a]	4.54	5.86	5.79	1.39	13.16	85.71	81.13
QXP	4.81	6.31	6.08	1.36	12.28	92.62	75.10

[a]2-(3-pyridyl)azetidine was excluded from the DISCO analysis (see text).

rounding space with diverse chemical probes, available in GRID, to simulate hydrophobic (DRY), HB acceptor (carbonyl oxygen), HB donor (hydroxyl oxygen) and ionic (carboxylate) interactions.

The MIP similarity was maximised by means of a molecular alignment optimisation process where deschloroepibatidine **13**, chosen as the reference, is kept fixed whereas the other nAChR agonists are free to translate and rotate. The MIP similarity was measured by the Pearson correlation coefficient.

The similarity values referred to the best MIP superpositions are illustrated in Figure 6, showing that higher similarity is achieved when using probes accounting for polar-like interactions rather than when considering the hydrophobic DRY probe. Interestingly, all the alignments resulting from MIPCOMP calculations were in agreement with our pharmacophore hypothesis. In fact, whatever probe was used, the resulting maximum MIP similarity alignments generate the maximum overlap of the previously described pharmacophore features over the whole nAChR agonist set. In other words, maximising the MIP similarity by means of rotations and translations provided a significant overlap of the common pharmacophore features. As can be observed in Figure 7, the MIP superpositions are in agreement with those suggested by the pharmacophore features.

Conclusions

The results obtained using different computational methods, namely DISCO, QXP, Catalyst/HipHop and MIPSIM, allowed us to develop a physicochemically well-defined pharmacophore model for nAChR agonists, which shows three key features, spatially arranged at proper distances and angles: (i) a cationic centre, involved in a coulombic or HB interaction, (ii) a lone pair of a pyridyl nitrogen or of a carbonyl oxygen as HB acceptor site, and (iii) the centre of a hydrophobic region.

The main improvements of our model over the existing pharmacophores could be summarised as follows: (i) the model has been generated starting from a properly selected set of structurally varying ligands; (ii) the three proposed key pharmacophore features are geometrically unrelated; (iii) the hydrogen bond directionality has been well determined; (iv) the centre of a hydropohobic moiety has been established as a new parmacophoric feature.

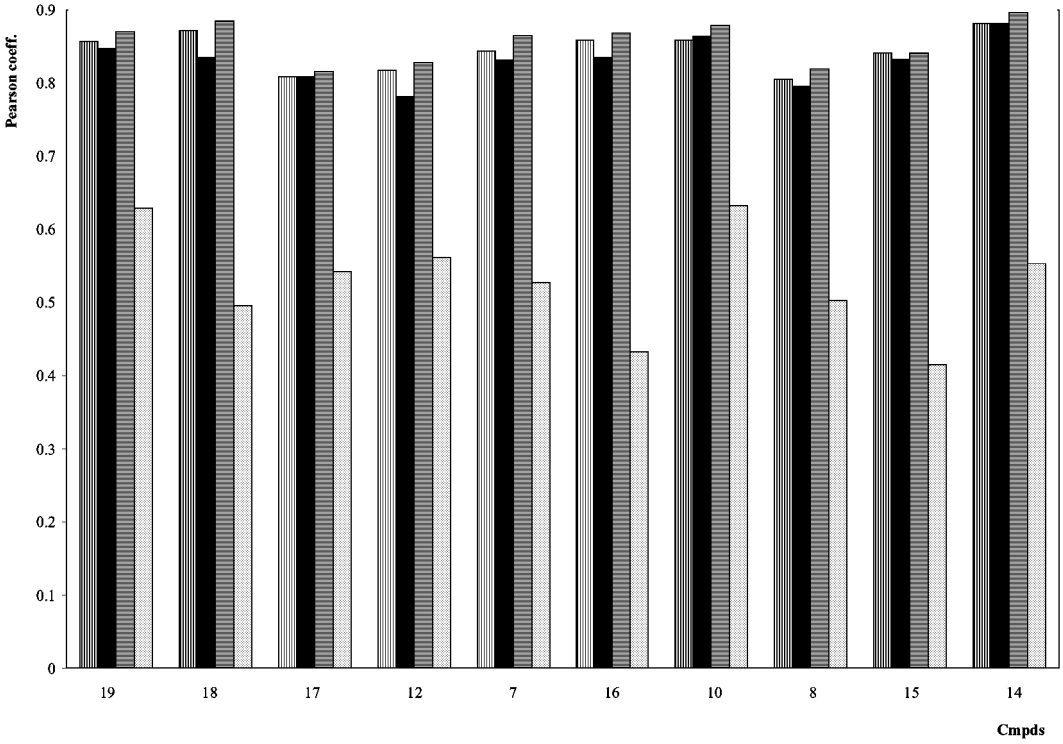

Figure 6. Bar plot of similarity indices (Pearson coefficient, deschloroepibatidine **13** as reference compound) for the GRID probes: ionic (vertical lines), HB donor (solid black), HB acceptor (horizontal lines), and hydrophobic (dotted lines).

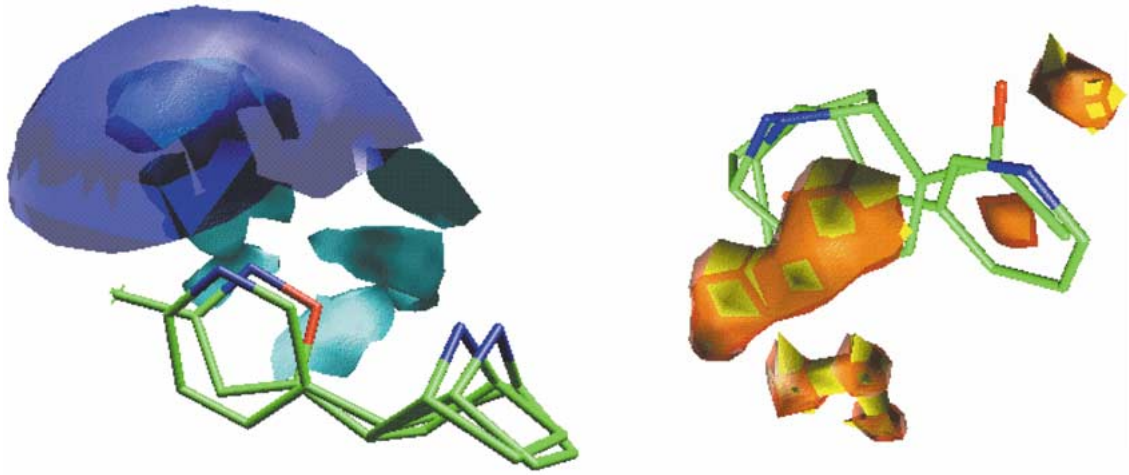

Figure 7. Right hand side: MIP shapes of (+)-anatoxin a **10** (−0.65 kcal/mol, in brown) superimposed on those of deschloroepibatidine **13** (−0.7 kcal/mol, in yellow) on the basis of GRID distributions using the DRY probe and MIPCOM calculations (similarity coeff. = 0.632). Shapes indicate common hydrophobic interaction regions. Left hand side: MIP shapes of epiboxidine **14** (−2.2 kcal/mol, in cyan) superimposed on those of deschloroepibatidine **13** (−2.2 kcal/mol, in blue) on the basis of GRID distributions using the carbonyl oxygen probe and MIPCOM calculations (similarity coeff. = 0.881), shapes indicate shared HB interaction regions.

In general, a pharmacophore is considered as a mere working hypothesis [46], often the only rational tool available to the medicinal chemist when the receptor structure is not known. A challenge is still how to render a pharmacophore something more than *ad hoc* rationalization of structure-activity relations. We believe that the results obtained from the comparison among different computational methods presented herein, one of them, MIPSIM, being a property-based method, represent a successful attempt to generate a pharmacophore of a general utility for nicotinic agonist binding to $\alpha_4\beta_2$ nAChRs. The reliability of our model both in terms of calculated physicochemical properties, such as the molecular interaction potential, and parameters of fitting, such as indices of similarity among properties, make us quite confident that it could be profitably used, together with predictive 2D and 3D QSAR models recently developed in our group [19], to drive database screening and to design new potentially active nAChR agonists.

Acknowledgements

We are very grateful to Professor Peter Goodford (University of Cambridge), Professor Roberto Todeschini (University 'La Bicocca', Milan), Professor Maurizio Botta (University of Siena), Dr. Colin McMartin and Professor Luisa Mosti (University of Genoa) for providing a free access to Grid, Dragon 1.0, Catalyst/HipHop, QXP and DISCO softwares, respectively. Financial support by MURST (Rome) and CNR (Rome) are gratefully acknowledged.

References

1. McDonald, I.A., Cosford, N. and Vernier, J.M., Ann. Rep. in Med. Chem., 5 (1995) 41.
2. Holladay, M.W., Lebold, S.A. and Lin, N.H., Drug Dev. Res., 35 (1995) 191.
3. Holladay, M.W., Dart, M.J. and Lynch., J.K., J. Med. Chem., 40 (1997) 4169.
4. Arneric, P. and Brioni, J.D. (ed.), Neuronal Nicotinic Receptors, Wiley-Liss, New York, NY, (1999).
5. Levin, E.D., Drug Dev. Res., 38 (1996) 188.
6. Gualtieri, F., Pharm. Acta Helvet., 74 (2000) 85.
7. Newhouse, P.A. and Kelton, M., Pharm. Acta Helvet., 74 (2000) 91.
8. Shacka, J.J. and Robinson, S.E., Med. Chem. Res., 6 (1996) 444.
9. Miyazawa, A., Fujiyoshi, Y., Stowell, M. and Unwin, N., J. Mol. Biol., 288 (1999) 765.
10. Dart, M.J., Wasicak, K.B., Ryther, K.B., Schrimpf, M.R., Kim, K.H., Anderson, D.J., Sullivan, J.P. and Meyer, M.D., Pharm. Acta Helvet., 74 (2000) 115.
11. Beers, W.H. and Reich, E., Nature, 225 (1970) 917.
12. Sheridan, R.P., Nilakantan, R., Dixon, J.S. and Venkataraghavan, R.J., J. Med. Chem, 29 (1986) 899.
13. Barlow, R.B. and Johnson, O., Br. J. Pharmacol., 98 (1989) 799.
14. Manallack, D.T., Gallagher, T. and Livingstone D.J., In Deviller, James (Ed.) Neural networks in QSAR and Drug Design, Rillieux la Pape, France, 1996.
15. Barreiro, E.J., Barreiro, G., Guimarães, C.R.W. and Bicca de Alencastro, R., J. Mol. Str: Theochem., 532 (2000) 11.
16. Tonder, J.E., Olesen, P., Hansen, J.B., Begtrup, M. and Petterson, I., J. Comput. Aid. Mol. Des., 15 (2001) 247.
17. van de Waterbeemd, Han (ed.), Chemometric Methods in Molecular Design, Wiley-VCH, Weinheim ,1995.
18. Damaj, M.I., Glassco, W., Dukat, M., May, L.M., Glennon, R.A. and Martin, B.R. Drug Dev. Res., 38 (1996) 177.
19. Nicolotti, O., Pellegrini-Calace, M., Altomare, C., Carotti, A. and Sanz, F., Curr. Med Chem., in press.
20. DRAGON1.0, Talete srl, Todeschini, R. and Consonni, V. (2000).
21. Allen, F.H. and Kennard O., *Chem. Des. Autom News*, 8 (1993) 31.
22. Stewart J.J.P. MOPAC: a semiempirical orbital program J. Comput Aid. Mol. Des. 4. (1990), 1.
23. SYBYL, TRIPOS Inc., St. Louis, MO, USA.
24. Vedani, A. and Duniz, J.D., J. Am. Chem. Soc., 107 (1985) 7653.
25. Ippolito, J.A., Alexander, R.S. and Christianson, D.W., J. Mol. Biol., 215 (1990) 457.
26. Kollman, P.A., Case, D., Singh, U.C., Alagona, G., Profeta, S. and Weiner, P. J. Am. Chem. Soc., 106 (1984) 765.
27. Polak, E. and Ribiere, G., Rev. Fran. Infor.Rech. Oper., 16 (1969) 35.
28. Thompson, P.E., Manallack, D.T.,. Blaney, F.E. and Gallagher, T. J. Comput. Aid. Mol. Des., 6 (1992) 287.
29. MacroModel 6.0; Copyright Columbia University 1986–1997.
30. Jorgensen, W.L. and Tirado-Rives, J., J. Am. Chem. Soc., 110 (1988) 1657.
31. McMartin, C. and Bohacek R.S., J Comput. Aid. Mol. Des., 9 (1995) 237.
32. McMartin, C. and Bohacek R.S., J Comput. Aid. Mol. Des., 11 (1997) 333.
33. Martin, Y.C., Bures, M., Dahaner, E., DeLazzer, J., Lico, I. and Pavlik, P., J. Comput. Aid. Mol. Des. 7 (1993) 83.
34. Brint, A. and Willett, P., J. Chem. Inf. Comput. Sci., 27 (1987) 152.
35. Molecular Simulations Inc, 9865, Scranton Road, San Diego, CA 92121, USA.
36. Barnum, D., Greene, J., Smellie, A. and Sprague, P., J. Chem. Inf. Comput. Sci. 36 (1996) 563.
37. Smellie, A., Teig, S. and Towbin, P., J. Comp. Chem., 16 (1995) 171.
38. Smellie, A., Kahn, S.D. and Teig, S., J. Chem. Inf. Comput. Sci., 35 (1995) 295.
39. Greene, J., Kahn, S., Savoj, H., Sprague, P. and Teig, S., J. Chem. Inf. Comput. Sci., 34 (1994) 1297.
40. Orozco, M. and Luque F.J., In: Murray, J.S. and Sen, K. (Eds.), Molecular Electrostatic Potentials: Concepts and Applications. Theoretical and Computational Chemistry, Vol. 3, Elsevier, Amsterdam, 1996. pp. 181–218.
41. Schmidt, M.W., Baldridge, K.K., Boatz, J.A., Elbert, S.T., Gordon, M.S., Jensen, J.H., Koseki, S., Matsunaga, N., Nguyen, K.A., Su, S.J., Windus, T.L., Dupuis, M. and Montgomery J.A., J. Comput. Chem., 14 (1993) 1347.

42. De Càceres, M., Villà, J., Lozano, J.J. and Sanz, F. Bioinformatics, 16 (2000) 568.
43. Sanz F., Manaut F., Rodriguez J., Lozoya E. and Lopez-de-Brinas E., J. Comput. Aid. Mol. Des., 7 (1993) 337.
44. Goodford, P.J., Med. Chem., 28 (1985) 849.
45. Baroni, M., Costantino, G., Cruciani, G., Riganelli, D., Vagli, R. and Clementi, S., Quant. Struct.-Act. Rel., 12 (1993) 9.
46. Wermouth, C.G., Ganellin, C.R., Lindberg, P. and Mitscher, L.A., Ann. Rep. Med. Chem., 36 (1998) 385.