Molecular Modeling of mono- and bis-Quaternary Ammonium Salts as Ligands at the $\alpha 4\beta 2^*$ Nicotinic Acetylcholine Receptor Subtype Using Nonlinear Techniques

Submitted: March 7, 2005; Accepted: May 30, 2005; Published: October 25, 2005

Joshua T. Ayers,^{1,2} Aaron Clauset,³ Jeffrey D. Schmitt,⁴ Linda P. Dwoskin,¹ and Peter A. Crooks¹

¹College of Pharmacy, University of Kentucky, Lexington, KY 40536-0082

²Current address: AstraZeneca Pharmaceuticals LP, 1800 Concord Pike, Wilmington, DE 19850-5437

³Department of Computer Science, University of New Mexico, Albuquerque, NM 87131

⁴Targacept Inc, 200 East First Street, Winston-Salem, NC 27101-4165

ABSTRACT

The neuronal nicotinic acetylcholine receptor (nAChR) has been a target for drug development studies for over a decade. A series of *mono*- and *bis*-quaternary ammonium salts, known to be antagonists at nAChRs, were separated into 3 structural classes and evaluated using both self-organizing map (SOM) and genetic functional approximation (GFA) algorithm models. Descriptors from these compounds were used to create several nonlinear quantitative structure-activity relationships (QSARs). The SOM methodology was effective in appropriately grouping these compounds with diverse structures and activities. The GFA models were also able to predict the activities of these molecules. Charge distribution and the hydrophobic free energies were found to be important indicators of bioactivity for this particular class of molecules. These QSAR approaches may be a useful to screen and select *in silico* new drug candidates from larger compound libraries to be further evaluated in *in vitro* biological assays.

KEYWORDS: self-organizing map, genetic functional approximation, neuronal nicotinic acetylcholine receptor

INTRODUCTION

Over the past decade, the neuronal nicotinic acetylcholine receptor (nAChR) has become a target for drug development. The most prevalent nAChR subtype in the mammalian brain is the $\alpha 4\beta 2^*$ receptor subtype, which is the primary high affinity [³H]nicotine binding site. Ligands that specifically interact with this nAChR subtype are of interest in several pathologies, including tobacco dependence and drug addiction, Alzheimer's disease, Parkinson's disease, Tourette's syndrome, schizophrenia, and pain.¹⁻⁶ Nicotine and other nAChR agonists (eg, epibatidine) are not selective

for $\alpha 4\beta 2^*$ nAChR subtypes and produce significant side effects, making these compounds less attractive as drug candidates. Consequently, previous studies have generated quantitative structure-activity relationships (QSAR) in an attempt to understand the interactions of these agonists with their recognition sites on nAChRs.⁷⁻¹¹

To date, several of these QSAR studies have used linear regression analysis of ligand-receptor binding or function in order to elucidate the nature of the ligand interactions within several nAChR subtypes.⁸⁻¹⁰ Of note, QSAR studies have been used to investigate ligand interactions at the $\alpha 4\beta 2^*$ nAChR subtype.^{9,12,13} Though generalities for the ligand interactions were drawn from the SARs generated, insufficient model validation limited the ability of the previously used models to recognize the training set of compounds. As such, generation of a nAChR pharmacophore has proven difficult. In this respect the earlier pharmacophore model generated by Sheridan et al, despite being based on crude pharmacological data, still constitutes an approach worthy of further exploration.¹⁴⁻¹⁶

Another model surmises that a water molecule may mediate binding of a ligand to the receptor, perhaps through an interaction with the protonated nitrogen atom of the ligand or the hydrogen bonding acceptor moiety in the ligand molecule.¹⁶ The crystallization of an acetylcholine binding protein (AChBP) from snail glia has also provided insight into the binding domain of the nAChR. Brejc et al have determined that the loop C (Changeux's nomenclature) of the acetylcholine binding site was crystallographically less well resolved than the rest of the binding site,¹⁷ leading to the possibility that part of the binding site may take on different conformations, depending on ligand shape and size. This finding represents a plausible challenge to the notion of a static pharmacophore. Owing to the complexity of the nAChR molecular recognition process and the difficulty in obtaining high-resolution (ie, microscopic kinetics) data for specific receptor subtypes, unambiguous models defining specific nAChR subtype pharmacophores do not yet exist.

QSAR descriptor data have been previously established as having nonlinear characteristics.¹⁸⁻²⁰ Therefore, alternative

Corresponding Author: Peter A. Crooks, College of Pharmacy, University of Kentucky, Lexington, KY 40536-0082. Tel: (859) 257-1718; Fax: (859) 257-1718; E-mail: pcrooks@email.uky.edu

approaches such as self-organizing maps (SOMs) and the genetic functional approximation (GFA) algorithm have been used to generate QSAR for several biological systems²¹⁻²⁵ and may be better suited for analysis of ligand interactions with $\alpha 4\beta 2^*$ nAChRs. Regression analysis typically employed in QSAR, such as partial least squares or multiple least squares are limited to relating linear descriptor trends to a given activity or property. Considering the complexity in the modeling of the *mono*- and *bis*-quaternary ammonium salt interactions with nAChRs described in this study, we elected to use SOMs and GFAs in the QSAR analysis described below as one valid modeling approach.

EXPERIMENTAL

Biological Data

Biological data were obtained from the [³H]nicotine binding assay using rat striatal membranes, a well-established assay reported to identify the $\alpha 4\beta 2^*$ subtype.²⁶⁻²⁹ The data set consisted of 90 *mono*- and *bis*-substituted quaternary ammonium compounds together with other related compounds developed in the laboratory of Dr Peter Crooks (University of Kentucky College of Pharmacy, Lexington, KY). The chemical structures are presented in Figure 1. Activities (ie, $\alpha 4\beta 2^*$ receptor affinities) are expressed as experimentally determined K_i values.

Figure 1. N-Alkylated quaternary ammonium salts and their affinities at $\alpha 4\beta 2^*$ nAChR subtypes.

Descriptor Generation

Descriptors were generated using the chemical modeling program Cerius2 (Accelrys Inc, San Diego, CA) and QSARIS (MDL Information Systems Inc, San Liandro, CA). Descriptors were divided into one of several physical property categories: conformational, electronic, informational, molecular shape analysis (MSA), quantum mechanical, receptor, spatial, structural, thermodynamic, or topological. The same descriptors were used for the generation of both the SOMs and GFAs. (Information on the 47 descriptors used in this study can be obtained by contacting the corresponding author.)

Self-organizing Maps With Descriptor Selection

In order to nonlinearly relate bioactivity and molecular features, a set of computational models were developed using a custom-built stepwise descriptor selection algorithm coupled to a supervised SOM.³⁰ The algorithm was developed using Matlab 6.1 with the SOM Toolbox 2.0 (<http://www.cis.hut.fi/projects/somtoolbox/>) (MathWorks, Inc, Natick, MA).

Data Sets

Three data sets were created: (1) both *mono*- and *bis*-salts (90 molecules; the “ALL” data set); (2) *mono*-quaternary ammonium salts (67 molecules; the “MONO” data set); and (3) *bis*-quaternary ammonium salts (23 molecules; the “BIS” data set). A binning scheme was created to convert the numerical bioactivity into categorical data since our implementation of SOMs uses categorical data. The scheme assigned roughly the same number of molecules to each of 3 bins. In using the supervised SOM learning algorithm provided in the toolbox, this binning scheme allowed for a more accurate clustering of the data sets and for extracting the underlying commonalities of the molecules.

Bin 1: $0 < x \leq 5$

Bin 2: $5 < x \leq 50$

Bin 3: $50 < x \leq 100$

The data set being modeled was partitioned into a set of test molecules (20%) and training molecules (80%). Membership in the test or training set was random, with the one requirement that the proportion of molecular population in each bin be identical in both test and training sets. The descriptor blocks for each data set were normalized to unit variance and mapped to a logistic function. Descriptors with zero variance were removed, as were pairs with linear correlation above 0.90, retaining the most representative of the pair.

The SOM-Stepwise Feature Reduction Algorithm

Finally, a SOM-oriented variation of a forward entry stepwise regression algorithm was employed to select a small

number of independent variables from which the SOM model was constructed. The algorithm used is identical to the forward entry stepwise method with the following variations:

1. the entry statistic was defined to be the percentage of training set molecules correctly classified by the model, any increase of which resulted in the inclusion of the descriptors being tested;
2. the entry statistic was calculated by constructing a SOM using a supervised learning algorithm; and
3. the descriptors to be entered into the model were chosen in pairs, 1 from each of 2 randomized lists of the full set of descriptors.

Algorithm training was conducted iteratively until some convergence criteria were met.

Experimental Design

For each of the 3 data sets, 10 independent runs were conducted, each with distinct training set/test set samplings. At the end of each training process, the test set molecules were predicted with the trained SOM model, providing the benchmark statistic (ie, percentage correctly classified). Averaging these statistics over the 10 runs produced an estimate of the algorithm's success in both learning the training data and generalizing to the test data. Conducting multiple runs to some extent minimizes sampling bias, as well as bias arising from the stochastic SOM initialization process. The best model for each of the 3 data sets was further examined using custom visualization tools, as shown in Figures 2, 3, and 4.

Genetic Functional Activity Algorithms

GFAs as described by Rogers and Hopfinger²⁰ were implemented using Cerius² software (Version 4.6, Accelrys Inc, San Diego, CA). GFAs were conducted using static 100 member populations; linear, quadratic, and offset quadratic basis functions; the mutation probability for both additions and deletions was set at 0.2%. Lack-of-fit criterion was set at 1.0, and initial equation length was set to 3. Models were evolved over 50 000 to 80 000 generations using standard convergence criterion.

RESULTS AND DISCUSSION

Self-organizing Maps

U-matrix maps represent the descriptor-based relationships of the training set of molecules by the SOM and are constructed by projecting all selected descriptor's contributions to the mapping.

D-matrix maps indicate the classification mapping of the SOM, as learned from the corresponding training sets. All 3 maps illustrate strong clustering.

Figure 2. (A) D-matrix generated from the ALL data set (red numbers indicate the classification mapping of each node); (B) U-matrix SOM generated from the ALL data set (red indicates class 1; green, class 2; blue, class 3). Inter-node distances are shown in gray-scale coloring, the closer the distance the darker the shading.

The color level in all the maps (see Figures 2-4) describes the relative clustering in the data; the darker areas indicate small distances, while the lighter areas indicate larger distances. The colored hexagons indicate the points on the map corresponding to training molecules and are colored by the affinity classification and labeled with the corresponding training molecule names.

SOM Models of the ALL Data Set

The 10 models learned the training set with high classification accuracy ($93\% \pm 0.02\%$) and were able to correctly classify the test compounds with a percentage accuracy of $68\% \pm 0.08\%$. The average number of descriptors chosen via the stepwise SOM was 11, selected from a total number

Figure 3. (A) D-matrix generated from the MONO data set; (B) the U-matrix representation.

of 47 descriptors. False negative values, where the model placed the compound 2 bins removed from its actual value, were low, accounting for less than 2% of the failures generated by the model. There were no false positives reported for any of the models. Figure 2 is a graphical representation of the best of the 10 models generated from the ALL data set. The descriptors selected by the best model were as follows: S_dO, S_aaCH, IC, S_sCH3, S_sssNH, S_tCH, and PHI (see below for more detail).

The D-matrix (Figure 2A) shows that there are 3 general clusters for the full complement of compounds in the ALL data set. The descriptor, IC (informational content), is an information-theoretic topological index and may be identifying steric qualities of the molecules. The descriptors used by the model mostly correspond to the electrotopological (E-Top) indices. S_dO describes the sum of the energies of

Figure 4. (A) D-matrix generated from the BIS data set; (B) the U-matrix representation.

carbonyl moieties in the set from the *N-n*-alkylnicotinamide series. As the carbonyl-containing compounds of this family are inactive, *S*_{dO} may be acting largely as a binary classifier. In addition, the best model used descriptors for the energy states of all CH₃, aromatic CH, total CH, and NH bonds (ie, *S*_{sCH₃}, *S*_{aaCH}, *S*_{tCH}, and *S*_{sssNH}, respectively). These descriptors may differentiate the C3-functionality of the pyridine ring; the *S*_{dO} and *S*_{sssNH} descriptors together likely map the SAR relative to the inclusion of the amide functionality at the C3-position of the pyridine ring, leading to a decrease in affinity. Another descriptor that may demonstrate the importance of the C3-position functionality is the Kier flexibility index, PHI. The compounds in the *N-n*-alkylpicolinium series, which contain a methyl group at the C3-position of the pyridine ring, were also inactive and grouped with the *N-n*-alkylnicotinamide salts. The active *N-n*-alkylnicotinium salts, with a 2-*N*-methylpyrrolidino functionality at the C3-position of the pyridine ring, however, were separated from the inactive compounds. The rigidity of the methyl and amide function-

alities when compared with the more flexible pyrrolidine ring may also be important with respect to the observed affinity of these compounds at the $\alpha 4\beta 2^*$ nAChR.

The U-matrix (Figure 2B) reflects chemically intuitive trends; the longer saturated *N-n*-alkylnicotinium salts were clustered together. In addition, the unsaturated *N-n*-alkylnicotinium compounds, NONB-3y, NDNB-4t, and NDNB-3y, were similar in bioactivity and clustered near the saturated *N-n*-alkylnicotinium compound NNNI. However, NONI, the less active 8-carbon analog, was clustered among the active compounds, pointing to a limitation in the SOM or the descriptors. Shorter carbon chain *N-n*-alkylnicotinium compounds and *N-n*-alkylpyridinium compounds were likewise clustered together. Of interest, rotationally restricted analogs were not found in the freely rotating *N-n*-alkylnicotinium compound cluster. This reflects a bias toward their different affinities, as the model does not recognize the 2 classes as structurally similar but actually classifies the restricted rotation analogs as structurally similar to the longer chain *N-n*-alkylpyridinium analogs (NEcPB and NDDPI).

A strong distinction is observed between the different structural classes of *bis*-substituted quaternary ammonium compounds. The active *bis*-quaternary ammonium salts are closely associated with the *N-n*-alkylnicotinium compounds, as well as the classical muscle nAChR antagonist, decamethonium bromide (DEC). The *bis*-quaternary ammonium compounds clustered nearer to the *mono*-substituted *N-n*-alkylpyridinium salts. This clustering may indicate that the *bis*-quaternary ammonium compounds bind to the receptor in a fashion similar to the *N-n*-alkylnicotinium compounds.

SOM Models of the MONO Data Set

The 10 models learned the training set with high accuracy ($95\% \pm 0.03\%$) and were able to predict test compound affinities with $72\% \pm 0.10\%$ accuracy. The average number of descriptors chosen via the stepwise SOM regression method was 7.8 descriptors, selected from a total number of 42. A 7.2 percentile of the missed assignments by the model were false positives. False negatives accounted for less than 1% of the failures generated by the model. Figure 3 is a graphical demonstration of the best model generated from the *mono*-substituted series of compounds. From the model generated, the descriptors used to create the model were *S*_{aaCH}, Jurs-RASA, Jurs-PNSA-1, Foct, *S*_{dCH2}, Fh2o, and LogP.

The *mono*-substituted quaternary ammonium salt model based on the MONO data set also demonstrated 3 distinct areas of clustering (Figure 3A) and used different descriptors than those used in the ALL compound set, with only 1 similar descriptor (ie, *S*_{aaCH}). The model did use a similar descriptor, *S*_{dCH2}, to describe the importance of the energy states of all of the CH₂ groups in the molecule. One of the new descriptors, Jurs-PNSA-1 describes the partial

negative surface area (ie, the sum of the solvent-accessible surface areas of all negatively charged atoms). Also included in the model is the related Jurs-RASA descriptor, which indicates the relative hydrophobic surface area, defined as the total hydrophobic surface area divided by the total molecular solvent-accessible surface area. The hydrophobic character of compounds in this series appears to be a major determining factor for their affinity and is supported by the 3 additional related descriptors (eg, Foct, Fh2o, and LogP). These descriptors are thermodynamic descriptors that estimate the desolvation free energy for octanol, water, and the octanol/water partition coefficient, respectively.

The long chain *N-n*-alkylnicotinium salts and the unsaturated *N-n*-alkenylnicotinium salts were clustered together in the SOM model (Figure 3B). These compounds share similar structures and $\alpha 4\beta 2^*$ nAChR affinities. The unsaturated and saturated *N*-alkylpyridinium salts were grouped together with the shorter chained *N-n*-alkylnicotinium compounds, which indicates, according to the descriptors chosen, that both the alkyl chain and the pyridine C3-position substituent are important for determining $\alpha 4\beta 2^*$ nAChR affinity. The longer chain *N-n*-alkylnicotinium salts exhibited the highest affinities. This finding correlates with a larger number of lipophilic groups (ie, CH₂ groups in the *N*-alkyl chain and in the *N*-methylpyrrolidine ring), whereas the longer chain *N-n*-alkylpyridinium salts and the shorter chain *N-n*-alkylnicotinium salts have fewer numbers of CH₂ groups and, therefore, lower affinities for the $\alpha 4\beta 2^*$ nAChR receptor.

The third cluster contains 2 interesting compound families: *N-n*-alkylnicotinamide salts, which were inactive in the [³H]nicotine binding assay, and the conformationally restricted *N-n*-alkylnicotinium analogs, which were also inactive. The *N-n*-alkylnicotinamide salts possess a hydrophilic moiety at the 3-position of the pyridinium ring. Similar to many nAChR ligands, affinity at the $\alpha 4\beta 2^*$ nAChR subtype is positively influenced by a hydrophobic interaction, whereas a hydrophilic moiety at the 3-position of the pyridinium ring serves to decrease affinity. The low-affinity, conformationally-restricted *N-n*-alkylnicotinium analogs have a high degree of hydrophobicity, and the model was able to differentiate them from other nicotinium molecules. Although it is unclear exactly which descriptor or descriptors may map this phenomenon, it is clear that the SOMs can differentiate in a manner that takes into account both distributed (hydropathy) and localized (functional groups) determinants of the observed SAR.

SOM Models of the BIS Data Set

The 10 models learned the training set with high accuracy ($98\% \pm 0.03\%$) and predicted the test compounds with a percentage accuracy of $58\% \pm 0.15\%$. The average number of descriptors chosen via the stepwise SOM was 4.8 selected

from a total number of 30 descriptors. False positive and false negative values were not observed in this model. Figure 4 is a graphical representation of the best model generated from the *bis*-substituted series of compounds. From the model generated, the descriptors used to create the model were S_{ssCH2}, CIC, S_{aasC}, Jurs-RPCS, and Jurs-WNSA-3.

The model developed for the *bis*-quaternary ammonium salts based on the BIS data set also shows 3 distinct clustering areas (Figure 4A). This model contains the descriptors S_{ssCH2} (energy states of the CH₂ bonds) and S_{aasC} (energy states of the C bonds), again stressing the importance of the lipophilic nature of these compounds. The model also contained the descriptor, CIC (complementary informational content), which is an information-theoretic topological index similar to IC and may be identifying steric interactions of the molecules. The *bis*-quaternary ammonium salt model used the Jurs spatial descriptors to generate the model. However, unlike the *mono-N-n*-alkylated quaternary ammonium salt model, which uses a hydrophobic interaction descriptor, the *bis*-quaternary ammonium salt SOM used a positive charge surface area descriptor, Jurs-RPCS. This descriptor represents the most positively charged atom and its relative accessible surface area. Since these *bis*-molecules possess 2 positive charges, the selected descriptors may indicate an increased importance of charge distribution within the series. The model also uses the Jurs-WNSA-3 descriptor, which is similar to the Jurs-PNSA-1 descriptor used in the *mono-N-n*-alkylated quaternary ammonium salt model, except that it represents only a partial weighting of the surface accessible portions of the molecule. The highest affinity members of this data set, the *bis*-quaternary ammonium salts, were clustered together. DEC was also clustered in the high affinity bin, whereas the classical ganglionic nAChR antagonist, hexamethonium bromide (HEX), was clustered with the lower affinity compounds, viz, the *bis*-quinolinium and the *bis*-pyridinium salts. In addition, the *bis*-picolinium compounds were grouped with 2 of the *bis*-quinolinium salts in the weakest affinity set. Therefore, the presence of the 2-*N*-methylpyrrolidine ring of the *bis*-nicotinium salts plays an important role in determining the affinities of this series of compounds for the $\alpha 4\beta 2^*$ nAChR receptor. It has been demonstrated that a pK_a shift occurs in the pyrrolidino nitrogen when the pyridino functionality is quaternized.³¹ Therefore, the importance of the 2-*N*-methylpyrrolidine may be owing to increased lipophilicity and/or to the possibility that the less basic nitrogen may act as a hydrogen bond acceptor.

Genetic Functional Activity Algorithms

Genetic functional algorithm models

N-Alkylpyridinium and *N*-Alkyl-3-substituted pyridinium salts

The resulting equation was derived using the GFA methods:

$$K_i = 161.582 + 0.004244 (\text{Area})^2 - 2.7\text{e-}5 (\text{E-DIST-mag})^2 - 5.12623 (\text{Vm}) + 36.4862 (\text{SC-3_P}) + 0.001581 (\text{Jurs-WPSA-1})^2, \quad (1)$$

where r^2 is 0.835734 and q^2 is 0.720.

The GFA methodology was applied to generate 3 equations for the prediction of the quaternary ammonium compound's affinity at the $\alpha 4\beta 2^*$ nAChR subtype. Equation 1 was generated to correlate the descriptors to the bioactivity of all of the *N-n*-alkylpyridinium salts, *N-n*-alkylpicolinium salts, and *N-n*-alkylnicotinamide salts (MONO data set). The equation reveals a positive correlation with increased molecular volume (Vm) and area (Area), while the molecular connectivity index of longer chain lengths (SC-3_P) has a negative correlation. Of interesting, the compounds in the *N-n*-alkylnicotinamide series, which have larger molecular volumes than their corresponding pyridinium cousins, possess less affinity toward this nAChR subtype. In addition, the affinities of the *N-n*-alkylnicotinamide salts were also decreased when compared with the simpler *N-n*-alkylpyridinium salts. The GFA also had a negative correlation with the E-DIST-mag descriptor, which describes the entropy of the matrix. The Jurs-WPSA-1 descriptor describes the surface-weighted charged partial surface areas; an indication that the positive charge distribution plays a role in receptor recognition (note that the converse descriptor, Jurs-PNSA-1, was used in the corresponding SOM model).

N-Alkylnicotinium salts

The resulting equation was derived using the GFA methods:

$$K_i = 43.5762 + 1.0\text{e-}6 (\text{E-DIST-mag})^2 - 3.03532 (\text{Jurs-RPCS})^2 - 6862.5 (\text{Jurs-FNSA-3})^2 + 0.000602 (\text{MW})^2, \quad (2)$$

where r^2 is 0.592152 and q^2 is 0.369.

Equation 2 models the SAR of the *N-n*-alkylnicotinium salts only. As can be seen in Equation 1, the E-DIST-mag descriptor was used; however, a positive correlation was found with this set of data. Jurs-RPCS, a measurement of the relative positive surface charge, and Jurs-FNSA-3, fractional negatively charged partial surface areas, positively correlate with affinity for the $\alpha 4\beta 2^*$ nAChR binding site, with the latter having the biggest contribution. Of interest, molecular weight (MW) is also negatively correlated. Since the highest affinity analogs NDDNI and NDNI (C_{12} and the C_{10} *N-n*-alkylnicotinium compounds, respectively) possess the highest molecular weight, the spatial partial charge interactions described by the Jurs descriptors must be more heavily weighted in this case. In addition, this data set did not include NONI, the C_8 *N-n*-alkylnicotinium analog. The GFA

was unable to generate a reliable model with NONI included, indicating that the descriptor set was likely insufficient to fully describe the SAR of this series.

bis-quaternary ammonium salts

The resulting equation was derived using the GFA methods:

$$K_i = -141.363 + 1449.55 (\text{SIC})^2 + 0.000602 (\text{Jurs-WNSA-2})^2 - 47.5526 (\text{IC})^2 + 2.43134 (\log Z), \quad (3)$$

where r^2 is 0.810622 and q^2 is 0.657.

The affinity of the *bis*-quaternary ammonium compounds is modeled by Equation 3. Similar to the *N-n*-alkylnicotinium salts and the SOM maps, the *bis*-quaternary ammonium salts used a surface-weighted, partial-charge descriptor. Again, the information-theoretic topological indices are found in the model equation. IC and SIC (structural informational content) are present and may be identifying steric qualities of the molecules and Log Z (logarithm of the Hosoya index), which deals with connectivity of the molecule. It also appears that the Jurs descriptors are effective in describing the SAR of the quaternary ammonium salts. The GFA selected a surface-weighted, partial negative charge descriptor, Jurs-WNSA, which is most likely related to a portion of the molecule, possibly the aromatic pyridinium ring.

It is important to note that the SOM and GFA methodologies used the same set of descriptors at the outset of training and, not surprisingly, models resulting from the 2 methods used different subsets of these descriptors. A combination of 3 factors is thought to underlie this observation: (1) GFAs and SOMs model nonlinearity in fundamentally different ways; (2) each method approaches descriptor selection in fundamentally different ways; and (3) they both have different susceptibilities to getting stuck in local minima.

CONCLUSION

Two different computational modeling methodologies were used to model the SAR of the interaction of quaternary ammonium compounds with the $\alpha 4\beta 2^*$ nAChR subtype. The SOM methodology was effective in appropriately grouping compounds with diverse structures and activities. GFAs were also able to predict compound activities from a set of physical descriptions. Taken together, the charge distribution and the hydrophobic free energies seem to be important indicators of bioactivity for this particular class of molecules. However, both methodologies exhibited limitations when compounds with similar physical properties but vastly different biological activities were evaluated. Nevertheless, the current study demonstrates that nonlinear *in silico* methodologies can be used to screen large series of compounds, and successful models can be generated to aide

in the selection of active compounds from larger compound libraries to be further evaluated.

ACKNOWLEDGMENTS

The authors acknowledge generous funding of this research from the National Institutes of Health (NIH) (Bethesda, MD) grant DA017548.

REFERENCES

- Lindstrom J. Nicotinic acetylcholine receptors in health and disease. *Mol Neurobiol*. 1997;15:193-222.
- Dani JA, Biasi M, Liang Y, et al. Potential applications of nicotinic ligands in the laboratory and clinic. *Bioorg Med Chem Lett*. 2004;14:1837-1839.
- Lloyd GK, Williams M. Neuronal nicotinic acetylcholine receptors as novel drug targets. *J Pharmacol Exp Ther*. 2000;292:461-467.
- Hogg RC, Bertrand D. Nicotinic acetylcholine receptors as drug targets. *Curr Drug Target CNS Neurol Disord*. 2004;3:123-130.
- Dwoskin LP, Xu R, Ayers JT, Crooks PA. Recent developments in neuronal nicotinic acetylcholine receptor antagonists. *Curr Opin Ther Patents*. 2000;10:1561-1581.
- Dwoskin LP, Crooks PA. Competitive neuronal nicotinic receptor antagonists: a new direction for drug discovery. *J Pharmacol Exp Ther*. 2001;298:395-402.
- Nicolotti O, Pellegrini-Calace M, Carrieri A, et al. Neuronal nicotinic receptor agonists: a multi-approach development of the pharmacophore. *J Comput Aided Mol Des*. 2001;15:859-872.
- Nielsen SF, Nielsen EO, Olsen GM, Liljefors T, Peters D. Novel potent ligands for the central nicotinic acetylcholine receptor: synthesis, receptor binding, and 3D-QSAR analysis. *J Med Chem*. 2000;43:2217-2226.
- Dukat M, Fiedler W, Dumas D, et al. Pyrrolidine-modified and 6-substituted analogs of nicotine: a structure-affinity investigation. *Eur J Med Chem*. 1996;31:875-888.
- Barlow RB, McLeod LJ. Some studies on cytosine and its methylated derivatives. *Br J Pharmacol*. 1969;35:161-174.
- Barlow RB, Johnson O. Relations between structure and nicotine-like activity: X-ray crystal structure analysis of (-)-cytosine and (-)-lobeline hydrochloride and a comparison with (-)-nicotine and other nicotine-like compounds. *Br J Pharmacol*. 1989;98:799-808.
- Tonder JE, Olesen PH, Hansen JB, Begtrup M, Pettersson I. An improved nicotinic pharmacophore and a stereoselective CoMFA-model for nicotinic agonists acting at the central nicotinic acetylcholine receptors labeled by [³H]-N-methylcarbamylcholine. *J Comput Aided Mol Des*. 2001;15:247-258.
- Tonder JE, Olesen PH. Agonists at the alpha4beta2 nicotinic acetylcholine receptors: structure-activity relationships and molecular modeling. *Curr Med Chem*. 2001;8:651-674.
- Schmitt JD. Exploring the nature of molecular recognition in nicotinic acetylcholine receptors. *Curr Med Chem*. 2000;7:749-800.
- Sheridan RP, Nilakantan R, Dixon JS, Venkataraghavan R. The ensemble approach to distance geometry: application to the nicotinic pharmacophore. *J Med Chem*. 1986;29:899-906.
- Glennon RA, Dukat M, Liao L. Musings on alpha4beta2 nicotinic acetylcholine (nACh) receptor pharmacophore models. *Curr Top Med Chem*. 2004;4:631-644.
- Brejck K, van Dijk WJ, Klaassen R, et al. Crystal structure of AChBP reveals the ligand-binding domain of nicotinic receptors. *Nature*. 2001;411:269-276.
- Zupan J, Gasteiger J. *Neuronal Networks for Chemists*. New York, NY: VCH; 1993.
- Kohonen T. *Self-organizing Maps*. Berlin, Germany: Springer; 2001.
- Rogers D, Hopfinger AJ. Application of genetic function approximation to quantitative structure-activity relationships and quantitative structure-property relationships. *J Chem Inf Comput Sci*. 1994;34:854-866.
- von Korff M, Steger M. GPCR-tailored pharmacophore pattern recognition of small molecular ligands. *J Chem Info & Comp Sci*. 2004;44:1137-1147.
- Schneider G, Nettekoven M. Ligand-based combinatorial design of selective purinergic receptor (A2A) antagonists using self-organizing maps. *J Comb Chem*. 2003;5:233-237.
- So S, van Helden S, van Geerestein V, Karplus M. Quantitative structure-activity relationship studies of progesterone receptor binding steroids. *J Chem Info & Comp Sci*. 2000;40:762-772.
- Yuan H, Parrill A. QSAR studies of HIV-1 integrase inhibition. *Bioorg Med Chem*. 2002;10:4169-4183.
- Gokhale V, Kulkarni V. Understanding the antifungal activity of terbinafine analogues using structure-activity relationship (QSAR) models. *Bioorg Med Chem*. 2000;8:2487-2499.
- Wilkins LHJr, Grinevich VP, Ayers JT, Crooks PA, Dwoskin LP. N-n-Alkylpyridinium analogs, a novel class of nicotinic receptor antagonists: interaction with alpha4beta2* and alpha7* neuronal nicotinic receptors. *J Pharmacol Exp Ther*. 2003;304:400-410.
- Xu R, Dwoskin LP, Grinevich VP, Deaciuc G, Crooks PA. Neuronal nicotinic acetylcholine receptor binding affinities of boron-containing nicotine analogues. *Bioorg Med Chem Lett*. 2001;11:1245-1248.
- Ayers JT, Dwoskin LP, Deaciuc AG, Grinevich VP, Zhu J, Crooks PA. bis-Azaaromatic quaternary ammonium analogues: ligands for alpha4beta2* and alpha7* subtypes of neuronal nicotinic receptors. *Bioorg Med Chem Lett*. 2002;12:3067-3071.
- Grinevich VP, Crooks PA, Sumithran P, Haubner AJ, Ayers JT, Dwoskin LP. N-n-Alkylpyridinium analogs, a novel class of nicotinic receptor antagonists: selective inhibition of nicotine-evoked [3H]dopamine overflow from superfused rat striatal. *J Pharmacol Exp Ther*. 2003;306:1011-1020.
- Bayram E, Santago P, Harris R, Xiao Y-D, Clauset AJ, Schmitt JD. Genetic algorithms and self-organizing maps: a powerful combination for modeling complex QSAR and QSPR problems. *J Comput Aided Mol Des*. 2004;187-9:483-493.
- Crooks PA, Ravard A, Wilkins LH, Teng L, Buxton ST, Dwoskin LP. Inhibition of nicotine-evoked [3H]dopamine release by pyridino N-substituted nicotine analogues: a new class of nicotinic antagonist. *Drug Dev Res*. 1995;36:91-102.