

A quantitative structure–activity relationship investigation into agonist binding at GABA_C receptors

Deborah L. Crittenden^{a,b}, Mary Chebib^a, Meredith J.T. Jordan^{b,*}

^aFaculty of Pharmacy, University of Sydney, NSW 2006, Australia

^bSchool of Chemistry, University of Sydney, NSW 2006, Australia

Received 16 June 2005; revised 25 July 2005; accepted 26 July 2005

Available online 7 October 2005

Abstract

The quantitative structure–activity relationship (QSAR) model constructed in this work represents the first quantitative investigation into agonist binding at GABA_C receptors. This model is based upon the three-dimensional structures of (γ-aminobutyric acid (GABA) and 12 other biologically active GABA analogues. These structures are obtained by geometry optimization and conformational exploration at MP2/6-31 + G* within the conductor-like screening solvation model (COSMO). The biological activity data are obtained from previous two-electrode voltage clamp electrophysiological studies on recombinant GABA_C ρ1 receptors expressed in *Xenopus laevis* oocytes. A QSAR model to predict GABA_C agonist binding is constructed from molecular superposition data, generated by least-squares superposition of the stable aqueous phase conformations of GABA and its known biologically active analogues to the bioactive conformation of TACA, the most potent GABA_C receptor agonist. A significant relationship is found between the root-mean-squared deviation in atomic position for the stable conformer of each GABA analogue most closely fitting the template TACA structure and the natural log of the normalized biological activity ($R^2=0.91$, $p<0.0001$). On the basis of molecular superposition and QSAR results, a pharmacophore model describing three-dimensional features and key interactions at the GABA_C receptor binding site is proposed. As there exists no direct experimental knowledge of structure of the GABA_C binding site, this approach provides a feasible and reliable alternative to gain insight into its three-dimensional structure.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Quantitative structure activity relationship; GABA; γ-aminobutyric acid; analogues; Ab initio; GABA_C receptor

1. Introduction

γ-Aminobutyric acid (GABA), the major inhibitory neurotransmitter in the mammalian central nervous system (CNS), exerts its biological activity through interaction with three major classes of GABA receptors, termed GABA_A, GABA_B and GABA_C receptors [1]. GABA_A and GABA_C receptors are ionotropic, ligand-gated receptors [2], whereas GABA_B receptors are metabotropic, G-protein coupled receptors [3]. GABA_C receptors can be distinguished from GABA_A receptors on the basis of significant differences in physiological role, pharmacological profile and molecular biology [4]. Physiologically, GABA_C receptors are predominantly implicated in visual processing [3–7], whereas

GABA_A receptors are more ubiquitous and play a role in a wide variety of CNS functions [8]. Pharmacologically, GABA_C receptors can be distinguished from GABA_A receptors on the basis of their selectivity for the agonist (+)-*cis*-2-aminomethylcyclopropane-1-carboxylic acid (CAMP) [9] and the antagonist 1,2,5,6-tetrahydropyridin-4-ylphosphinic acid (TPMPA) [10]. GABA_C receptors also exhibit a longer mean channel open time, less desensitization and a slower activation/deactivation rate than GABA_A receptors [11–13]. Molecular biology and genetics studies [3,14] have revealed that the human ρ1 and ρ2 GABA_C receptor subunits share 74% sequence identity with one another, while only sharing 30–38% sequence identity with their GABA_A receptor counterparts. While expression of only one type of GABA_C receptor subunit is required to form functional homomeric, pentameric receptors, GABA_A receptors require expression of at least three different subunits to form heteromeric, pentameric receptors [2]. Although, GABA_C receptor subunits may coexpress with

* Corresponding author. Tel.: +61 422 817 479.

E-mail address: m.jordan@chem.usyd.edu.au (M.J.T. Jordan).

GABA_A receptor subunits to form functional receptors in some species [15], in general this does not readily occur [4].

Current experimental evidence suggests that drugs acting at GABA_C receptors may be useful in the treatment of visual [16], sleep [17,18] and cognitive disorders [19–21]. As GABA_C receptors are less widespread in the CNS than GABA_A receptors, they have therapeutic potential as a more selective drug target, enabling development of neurologically active drugs without unwanted side effects. For example, there is no evidence for GABA_C receptors being involved in mood or motor function, with no reports of GABA_C selective drugs having anxiolytic, anxiogenic, convulsant or anticonvulsant actions [4]. This provides a major impetus for understanding the structure–activity relationships of drugs that act at GABA_C receptors and for the subsequent design and development of further ligands that are specific to GABA_C receptors.

The $\rho 1$ subtype of the GABA_C receptor has been the subject of the most experimental investigation and possesses the most complete pharmacological profile of all three GABA_C receptor subtypes. This receptor subtype has therefore been chosen as the subject of the current quantitative structure–activity relationship (QSAR) analysis. Although, there have been previous QSAR investigations into the relationship between ligand binding at the agonist/competitive antagonist binding site and biological activity at GABA_A [22,23] and GABA_B [24,25] receptors, this work represents the first QSAR study into agonist binding at GABA_C receptors. Quantitative structure–activity relationship studies of ligand binding at GABA receptors pose a challenging problem for two reasons. Firstly, there currently exists no direct experimental evidence of either the biologically active conformations of

the ligands or the receptor-binding site. Therefore, it is necessary to accurately determine all stable solution-phase structures of the zwitterionic ligands, which in itself is a non-trivial computational problem. In this work, as in previous works [26,27], we put considerable effort into determining the aqueous-phase structures of GABA and its biologically active analogues using quantum chemical methods. Secondly, it can be difficult to obtain reliable biological activity data, and in this work we compile a comprehensive pharmacological profile for $\rho 1$ GABA_C receptors.

2. Methods

Biological activity data for the compounds listed in Table 1 and illustrated in Fig. 1 were taken from the literature [28–39]. These data were generated using two-electrode voltage clamp electrophysiological techniques on recombinant GABA_C $\rho 1$ receptors expressed in *Xenopus laevis* oocytes. The set of structural data for the compounds listed in Table 1 was assembled either by obtaining stable solution-phase geometries directly from the literature [26, 27] or by using literature methods to generate the required structural data [26,27]. In general terms, the literature method for identifying stable solution-phase structures of solvated biological zwitterions involves performing a geometry optimization using ab initio molecular orbital methods to calculate molecular energies and a dielectric continuum solvation model to model solvent effects. Specifically, the molecular energies were calculated using second-order Moller-Plesset Perturbation Theory (MP2) [40–43], with a double-zeta Pople basis set augmented with

Table 1

Full chemical names and abbreviations of GABA analogues with activity at GABA_C receptors

	Abbreviation	Full chemical name	Activity
1	GABA	4-Aminobutanoic acid	Agonist
2	CACA	<i>cis</i> -4-Aminobut-2-enoic acid	Agonist
3	TACA	<i>trans</i> -4-Aminobut-2-enoic acid	Agonist
4	(+)-CAMP	(+)- <i>cis</i> -2-Aminomethylcyclopropane-1-carboxylic acid	Agonist
5	(+)-TAMP	(+)- <i>trans</i> -2-Aminomethylcyclopropane-1-carboxylic acid	Agonist
6	(-)-TAMP	(-)- <i>trans</i> -2-Aminomethylcyclopropane-1-carboxylic acid	Agonist
7	(+)-4-ACP-2-E	(1R,4S)-4-Aminocyclopent-2-ene-1-carboxylic acid	Agonist
8	Muscimol	5-Aminomethyl-3-hydroxy-isaxazole	Agonist
9	Homohypotaurine	4-Aminosulfonic acid	Agonist
10	Isoguvacine	4-Aminocyclohex-1-ene-1-carboxylic acid	Agonist
11	2-F-TACA	<i>trans</i> -4-Amino-2-fluorobutanoic acid	Agonist
12	I4AA	Indole-4-acetic acid	Both
13	A2MBA	4-Amino-2-methylbutanoic acid	Both
14	(-)-CAMP	(-)- <i>cis</i> -2-Aminomethylcyclopropane-1-carboxylic acid	Antagonist
15	(-)-4-ACP-2-E	(1S,4R)-4-Aminocyclopent-2-ene-1-carboxylic acid	Antagonist
16	THIP	4,5,6,7-Tetrahydroisoxazolo[5,4-c]pyridin-3-ol	Antagonist
17	Isonipectic acid	Piperidin-4-ylcarboxylic acid	Antagonist
18	2-Me-TACA	<i>trans</i> -4-Amino-2-methylbutanoic acid	Antagonist
19	DFP	4-Aminomethyl-2,6-difluorophenol	Antagonist
20	(S)-4-ACP-1-E	(S)-4-Aminomethylcyclopent-1-ene-1-carboxylic acid	Antagonist
21	(R)-4-ACP-1-E	(R)-4-Aminomethylcyclopent-1-ene-1-carboxylic acid	Inactive

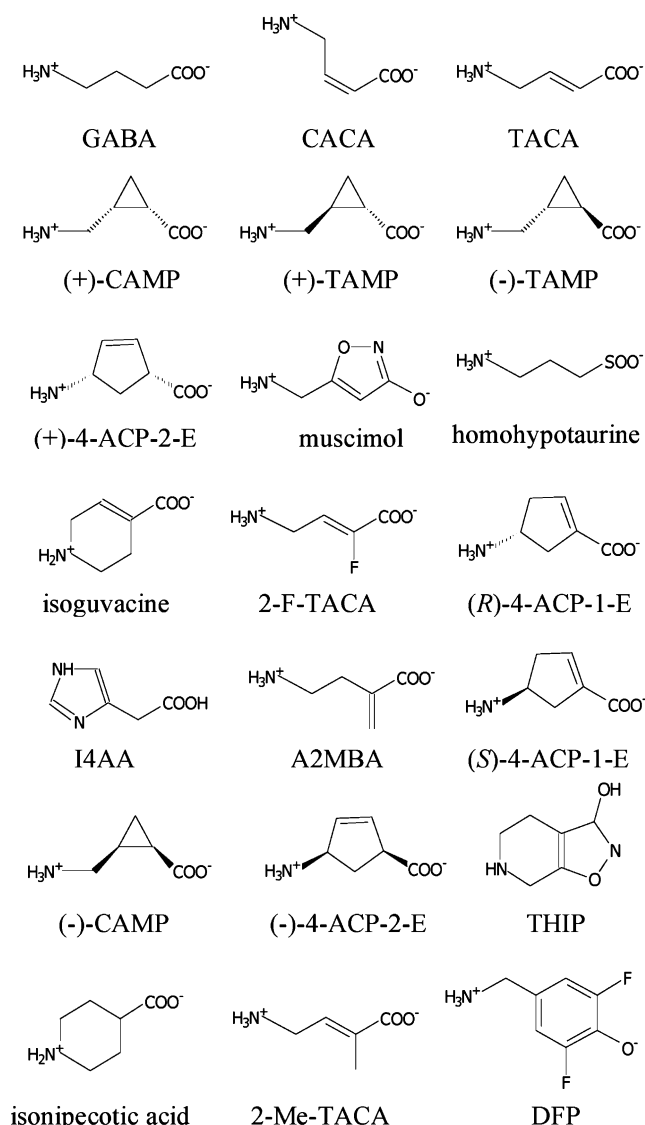


Fig. 1. Line structures of GABA analogues that have activity at GABA_C receptors.

diffuse and polarization functions ($6-31+G^*$) [44–48] used to describe the electronic wavefunction. These calculations were carried out within a dielectric continuum, as described by the conductor-like screening solvation model (COSMO) [49,50]. Stable rotamers were identified by systematically searching the conformational space of each molecule. This was accomplished by rotation about all rotatable bonds of the optimized structures and subsequent reoptimization from any other minima thus identified.

A common biologically active substructure was identified by considering the conformationally restricted GABA analogues TACA, (*S*)-4-ACP-1-E and (*R*)-4-ACP-1-E. These analogues were chosen either because they have a high affinity for the GABA_C receptor (TACA) or limited conformational flexibility ((*R*)-4-ACP-1-E) or both ((*S*)-4-ACP-1-E), where the affinity for the GABA_C receptor is measured in terms of the EC₅₀ or IC₅₀ value of the

compound. Although, (*S*)-4-ACP-1-E is a potent antagonist at GABA_C receptors, it is suitable for use as a template to determine the biologically active agonist substructure because it is a *competitive* antagonist. This indicates that it binds at the same site as agonist molecules, although, it produces an opposite biological effect. (*R*)-4-ACP-1-E, on the other hand, is a weak agonist at GABA_C receptors. Both (*R*)- and (*S*)-4-ACP-1-E are particularly useful as template molecules due to their complete conformational rigidity.

All stable TACA conformers were superposed upon the (*R*)-4-ACP-1-E and (*S*)-4-ACP-1-E template structures by least squares fitting of all heteroatoms corresponding to the GABA backbone. In total, seven atoms were fit—two oxygen atoms, one nitrogen atom and four carbon atoms. This atom–atom superposition scheme was chosen over a pharmacophore superposition scheme, as the extra information contained in the conformation of the hydrocarbon backbone is necessary to distinguish between the various ligands. The implication of this choice of superposition scheme is that steric interactions between the hydrocarbon backbone and the binding site are approximately as important in determining ligand affinity as the electrostatic and dipole–dipole interactions between the pharmacophoric groups and the binding site. This implies that the $\rho 1$ subtype of the GABA_C receptor possesses a relatively constricted binding site. This, in turn, is consistent with the known pharmacological profile of this receptor, which indicates that this receptor can only be activated by small molecules, with the largest known ligand being THIP (see Fig. 1). Furthermore, previous investigations [28,30,33–36] have indicated that ligands that can attain a flat or slightly buckled conformation have the highest affinity for the receptor, suggesting that the binding site encapsulates a narrow cavity. Hence, the root-mean-squared deviation is atomic position obtained from the atom–atom superposition approach adopted here is expected to be a suitable parameter for quantifying differences in molecular structure.

The TACA conformer illustrated in Fig. 2 was identified as the common substructure required for optimal biological activity, as it fits closely to the potent antagonist (*S*)-4-ACP-1-E while the fit to the weak agonist (*R*)-4-ACP-1-E shows a much larger root-mean-squared deviation in atomic position. This is confirmed by superposition of this structure upon other conformationally restricted GABA analogues. In particular, this substructure is identified among the stable solution-phase conformers of both the potent agonist muscimol and the moderately potent competitive antagonist THIP. This TACA conformer, which is characterized by

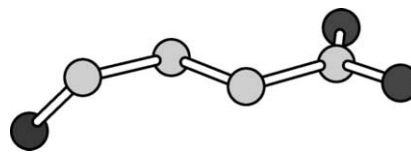


Fig. 2. Putative TACA structure required for optimal affinity at the GABA_C $\rho 1$ receptor.

a N–C–C–C dihedral angle of -118° , a C–C–C–C dihedral angle of -169° and a C–C–C–O dihedral angle of -178° , was used as the template molecule in subsequent structural superpositions for all compounds listed in Table 1.

A QSAR model for predicting agonist binding at GABA_C receptors was constructed by examining the correlation between the observed biological activity and root-mean-squared deviation (RMSD) in atomic position between template TACA conformer and the optimized solution phase geometries of the other GABA analogues. These RMSD values were obtained by least squares fitting of seven backbone heteroatoms of each of the GABA analogues to the corresponding seven backbone heteroatoms of the template TACA molecule. For most of the compounds studied here, this involved directly comparing the terminal nitrogen, four connecting carbon atoms and the two oxygen atoms of the carboxylate group. The exceptions were the compounds that contained other heteroatoms or which contained heteroatoms incorporated by a heterocyclic ring structure. The superposition procedure for each of these compounds—muscimol, homohypotaurine, isoguvacine, isonipecotic acid, I4AA, THIP and DFP—will be discussed separately below.

Muscimol was superposed upon the template TACA conformer by comparison of the hydroxyl oxygen and isoxazole nitrogen of muscimol with the carboxylate group of TACA. The four carbon atoms and terminal amino group of muscimol were fit to the four carbon atoms and terminal amino group of TACA. The fitting process for homohypotaurine was the same as for the compounds where a direct comparison was possible, with the exception that the sulfonic sulfur atom was superposed onto the carboxylate carbon. For isoguvacine and isonipecotic acid, it was possible to superpose the carbon atoms in the template TACA molecule to the carbon atoms on either side of the cyclopentene or cyclopentane ring. In these cases, both superpositions were carried out. I4AA possesses two nitrogen atoms, incorporated in its indole ring. Of these two nitrogen atoms, only the nitrogen atom in the γ position was used in the superposition. Like muscimol, THIP possess an isoxazole ring, and the molecular superposition was carried out by fitting the hydroxyl oxygen and isoxazole nitrogen to the carboxylate oxygens of TACA. However, like isoguvacine, THIP also possesses a cyclohexene ring, and the carbon atoms in the template TACA molecule can be superposed upon the carbon atoms on either side of the cyclohexene ring. Again, both superpositions were carried out. DFP is structurally unique amongst the GABA analogues considered here, possessing a fluorine-substituted phenol ring in place of a carboxylic acid or isoxazole group. In this case, the carboxylate oxygen atoms of TACA was fit to one of the fluorine substituents and the phenolic oxygen, and the carbon atoms between the terminal nitrogen group and the fluorine substitution were superposed upon the carbon atoms in the template TACA molecule.

In the cases where there was more than one stable conformer or more than one way to perform the molecular superposition or both, all possible superpositions were carried out, giving a set of RMSD values. The RMSD values used in the QSAR analysis were chosen by considering only the conformer most closely matching the template TACA molecule, except where the superposed geometry had a substituent that occupied an excluded volume, as discussed below. In this case, the RMSD value for the conformer with the best fit where the substituent did not occupy an excluded volume was used instead.

Linear regression analysis was performed with the natural log of the normalized biological activity as the response variable and the root-mean-squared deviation in atomic position as the dependent variable for the GABA_C receptor agonists. This analysis was only carried out for the analogues with agonist activity, as more biological activity data was available for agonists than antagonists. Accordingly, the relationship between biological activity and similarly to the bioactive template TACA molecule was easier to observe. The statistical significance of the model was assessed by one-way Analysis of Variance (ANOVA). Errors in the model were quantified by performing leave-one-out cross validation testing and examining the residual difference between the predicted and actual biological activity data. All statistical calculations were carried out in Microsoft Excel.

3. Results and discussion

Energetic and selected geometric data for the stable solution-phase conformers of homohypotaurine, isoguvacine, 2-F-TACA, I4AA, A2MBA, THIP, isonipecotic acid, 2-Me-TACA and DFP are presented in Table 2. Analogous data for the other compounds studied here may be obtained from previous work [27]. The RMSD data for the superposition of the bioactive TACA conformer on the GABA_C receptor agonists most closely matching the TACA template are presented in Table 3. The structures (+)-CAMP, (–)-CAMP, (+)-TAMP and (–)-TAMP superposed upon the template TACA conformer are illustrated in Fig. 3. The results of the superposition for the other GABA analogues are not illustrated here, but the molecular coordinates and two-dimensional illustrations of the superposed geometries may be obtained as Supplementary information. The results of the linear regression analysis are presented in Table 3 and illustrated in Fig. 4. The results of leave-one-out cross validation testing are present in Table 4. The correlation between biological activity and RMS deviation in atomic position was found to obey the relationship:

$$\ln \left[\frac{EC_{50}}{EC_{50}(TACA)} \right] = 5.4541 \times RMS$$

Table 2

Energetic and selected geometric data for the stable solution-phase conformers of homohypotaurine, isoguvacine, 2-F-TACA, I4AA, A2MBA, THIP, isonipecotic acid, 2-Me-TACA and DFP, as calculated at MP2/6-31+G* within the COSMO solvation model

	N–C–C–C	C–C–C–C	C–C–C–O	C–C–C–O	(ΔE (kJ/mol))
Isoguvacine (1)	48.9	75.3	8.2	–174.0	221.2
	–66.6	–54.5	149.0	–33.3	
Isoguvacine (2)	–13.4	175.8	15.3	–164.1	0.0
	48.5	166.0	–167.7	12.9	
Isonipecotic acid (1)	–55.4	–74.5	70.3	–108.9	7.5
	53.6	74.4	–54.0	126.8	
Isonipecotic acid (2)	–55.5	178.4	–57.2	122.9	0.0
	55.7	–178.9	65.0	–114.9	
THIP (1)	–47.2	–163.4	0.1	179.7	0.4
	13.9	–179.9	179.8	–0.5	
THIP (2)	46.4	164.7	–0.2	–179.8	0.0
	–14.6	179.4	–179.8	0.6	
2-F-TACA (1)	0.7	–179.6	–3.1	176.8	9.6
2-F-TACA (2)	–115.1	–178.6	–1.6	178.3	0.0
2-F-TACA (3)	115.8	179.4	–2.9	176.8	0.3
2-Me-TACA (1)	121.1	–178.6	–29.6	149.7	0.2
2-Me-TACA (2)	–120.1	–175.4	–37.5	143.7	0.0
A2MBA (1)	–44.6	–48.3	77.1	–102.2	7.1
A2MBA (2)	176.4	–67.4	–18.0	162.4	18.8
A2MBA (3)	78.0	–81.2	30.9	–149.6	0.0
A2MBA (4)	44.6	48.1	–76.4	103.0	7.2
A2MBA (5)	–176.5	67.9	16.4	–163.9	18.8
A2MBA (6)	–78.2	80.8	–30.0	150.4	0.0
	N–C–C–C	C–C–C–S	C–C–C–O	C–C–C–O	
Homohypotaurine (1)	–64.5	–173.2	–177.6	–63.6	23.8
Homohypotaurine (2)	179.9	179.6	57.5	–57.0	22.8
Homohypotaurine (3)	66.1	175.2	–63.5	–177.6	24.2
Homohypotaurine (4)	–177.0	–68.4	–62.0	–176.1	22.9
Homohypotaurine (5)	75.6	–84.2	64.0	176.8	0.0
Homohypotaurine (6)	–42.0	–49.9	72.7	–174.6	7.3
Homohypotaurine (7)	–75.7	84.5	–63.7	–176.5	0.1
Homohypotaurine (8)	176.9	70.0	62.0	176.0	23.0
Homohypotaurine (9)	42.8	48.9	–72.7	174.6	7.3
	N–C–C–C	C–C–C–C	C–C–C–O	C–C–C–OH	
I4AA (1)	–179.2	–128.2	–120.4	59.9	2.6
I4AA (2)	–179.8	120.7	120.7	–59.9	0.0
I4AA (3)	–179.2	–128.2	59.9	–120.4	2.6
I4AA (4)	–179.8	120.7	–59.9	120.7	0.0
	N–C–C–C	C–C–C–C	C–C–C–O	C–C–C–F	
DFP	89.8	–179.0	–0.4	–179.6	0.0
	–89.6	179.0	0.5	179.6	

The F -statistic of 464.1 returned by one-way ANOVA corresponds to a probability of 2×10^{-10} that the correlation between the natural log of the normalized biological activity data and the RMS deviation in atomic position data is due to chance. This is equivalent to a probability of less than 0.0000001% for a chance correlation, indicating that a significant linear relationship

exists between the two data sets. This is consistent with the R^2 value of 0.9063, which also indicates the existence of a statistically significant relationship.

The univariate regression analysis performed here has two types of associated error. Firstly, there is an uncertainty associated with estimating the regression coefficient. Secondly, there is the error introduced by using the model

Table 3

Root-mean-squared deviations in heteroatom position from the template TACA molecule, actual and predicted biological activity data for GABA analogues with agonist activity at GABA_C receptors. The references for the biological activity data are given in the final column

Compound name	EC ₅₀ (μM)	RMS fit (Å)	Predicted EC ₅₀ (μM)	Residuals (μM)	Ref.
GABA	1.01	0.345	3.47	2.46	[36]
CACA	37.4	0.737	29.5	−7.9	[28]
TACA	0.53	0.000	0.53	0.00	[33]
(+)-CAMP	39.7	0.772	35.7	−4.0	[35]
(+)-TAMP	59.1	0.802	42.2	−16.9	[35]
(−)-TAMP	8.90	0.383	4.29	−4.61	[35]
(+)-4-ACP-2-E	135	0.985	114	−21	[39]
Muscimol	1.48	0.351	3.6	2.11	[37]
Homohypotaurine	4.59	0.374	4.1	−0.52	[33]
Isoguvacine	99.0	0.959	99.2	0.2	[30]
2-F-TACA	2.43	0.080	0.82	−1.61	[33]
I4AA	7.57	0.467	6.75	−0.82	[32]
A2MBA	182	1.138	264	82	[33]

to predict the biological activities. Fortunately, both types of error are easy to quantify, facilitating assessment of the reliability of the predicted biological activities.

The error in the regression coefficient can be estimated either directly as the standard error returned by the regression analysis or through leave-one-out cross validation testing. Leave-one-out cross validation testing is particularly useful for identifying which, if any, outlying data points unduly influence the value of the regression coefficient obtained. If any single data point dominates the regression analysis, the results obtained excluding that point will be significantly different from all other results. The value of the regression coefficient using all available data was $5.4541 (\text{Å})^{-1}$. The standard error in the estimate of the regression coefficient was returned directly from the regression analysis as a value of $0.2532 (\text{Å})^{-1}$. The median value of the regression coefficient averaged over all leave-one-out cross validation testing runs was $5.4545 (\text{Å})^{-1}$. These two estimates of the regression coefficient can be considered the same, within the given error tolerance. The maximum deviation in the leave-one-out estimates of the regression coefficient was $0.0949 (\text{Å})^{-1}$ and a relatively narrow spread of values can be observed from the results shown in Table 4. This indicates that no one data point is dominating the regression analysis.

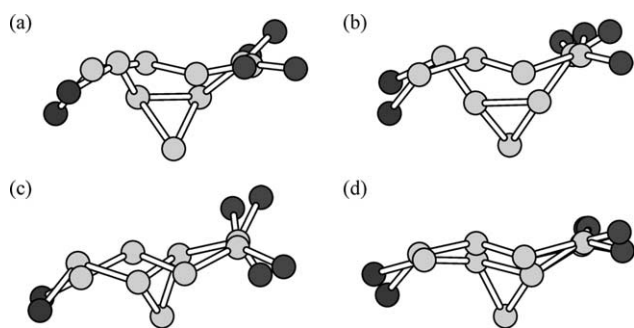


Fig. 3. (a) (+)-CAMP, (b) (−)-CAMP, (c) (+)-TAMP and (d) (−)-TAMP superposed on the template TACA molecule.

Errors in the predicted biological activities can be quantified either by examining the residual difference between the predicted and actual biological activity data or by performing leave-one-out cross validation testing. The average root-mean-squared residual difference returned directly from analysis of the linear regression data was found to be $24.0 \mu\text{M}$, whereas the average root-mean-squared difference from leave-one-out cross validation testing was $32.4 \mu\text{M}$. Leave-one-out cross validation testing provides a more rigorous upper bound on the error estimate, as it involves making biological activity predictions for data points not included in the training set, whereas the residual errors returned by linear regression analysis are optimized to the data points included in the training set. Thus, $32.4 \mu\text{M}$ can be considered the average uncertainty in a biological activity prediction made by this linear regression model. This implies that our linear regression model is not sufficiently accurate to discriminate between compounds with similar biological activities, but may be useful in identifying and categorizing the difference between high affinity and low affinity ligands.

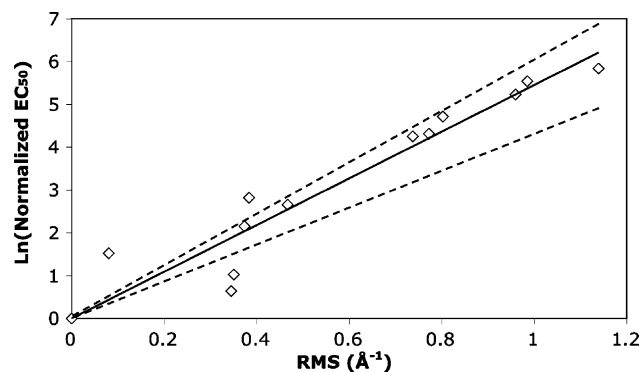


Fig. 4. Linear regression plot illustrating the correlation between biological activity and root-mean-squared deviation in the fitting of all heteroatoms to the template TACA molecule illustrated in Fig. 2. The EC₅₀ values have been normalized to the EC₅₀ of TACA. The regression equation is shown as the solid line, while the dashed lines represent the 95% confidence interval.

Table 4
Results of leave-one-out cross validation analysis of the relationship between the normalized biological activity and the root-mean-squared fit of each GABA analogue to a template TACA conformer

Compound name	EC ₅₀ (μM)	Parameter value	Predicted EC ₅₀ (μM)	Residuals (μM)
GABA	1.01	5.530	3.57	2.56
CACA	37.4	5.420	28.8	−8.6
TACA	0.53	5.454	0.53	0.00
(+)-CAMP	39.7	5.438	35.3	−4.4
(+)-TAMP	59.1	5.401	40.4	−18.7
(−)-TAMP	8.90	5.404	4.20	−4.70
(+)-4-ACP-2-E	135	5.419	110	−25
Muscimol	1.48	5.510	3.66	2.18
Homohypotaurine	4.59	5.446	4.05	−0.54
Isoguvacine	99.0	5.454	99.2	0.2
2-F-TACA	2.43	5.439	0.82	−1.61
I4AA	7.57	5.444	6.72	−0.85
A2MBA	182	5.549	294	112

From the residual error data presented in Table 3 and the leave-one-out cross validation data presented in Table 4, it is also observed that there is significant heterogeneity in the predicted errors for each data point. The compounds with the highest EC₅₀ values tend also to have residual errors that are larger in absolute magnitude, although, the relative errors are highest for compounds with higher biological activities. In particular, leave-one-out cross validation analysis predicts 2.4- and 1.4-fold errors relative to actual biological activity for GABA and muscimol, respectively.

Linear regression analysis is the most simple of all statistical models for finding relationships between correlated data. The simplicity of linear models has both advantages and disadvantages. The major disadvantage is that they do not tend to perform as well in pure predictive ability as nonlinear models such as neural networks [51]. However, linear regression models also possess some advantages over nonlinear models. They do not suffer from the over-fitting problems encountered in training neural networks, and thus are applicable even to relatively small data sets [51]. Linear regression models are also less computationally intensive than nonlinear models. However, the major advantage of linear regression models is that their

Table 5
Biological activity data for GABA analogues with antagonist activity at GABA_C receptors

Compound name	IC ₅₀ (μM)	Ref.
(−)-CAMP	890	[35]
(S)-4-ACP-1-E	6	[39]
(−)-4-ACP-2-E	> 300	[39]
THIP	30.3	[37]
Isonipetric acid	> 300	[29]
DFP	75.5	[34]
2-Me-TACA	31	[33]
I4AA	1.9	[32]
A2MBA	28.8	[33]

References for the biological activity data are given in the final column.

simplicity allows for straightforward physical interpretation of the results [51].

The qualitative picture of binding at the GABA_C receptor produced by the QSAR and molecular superposition results can be summarized according to two principles. Firstly, the efficacy of agonist binding is related to the similarity of the agonist structure to the biologically active TACA structure shown in Fig. 2; the greater the similarity, the lower the EC₅₀ and higher the agonist potency. Secondly, on the basis of the superposed structures, it is postulated that steric interactions with the binding site determine the nature of the biological activity; that is, whether a compound has agonist or antagonist activity, or both. For reference, the relative affinities of all known GABA_C receptor antagonists are given in Table 5. This represents a qualitative model for distinguishing between agonist and antagonist activity. The effects of steric interactions can be summarized by considering the biologically active TACA structure. Two planes are defined with respect to this structure in Fig. 5. Fig. 5(a) illustrates the plane containing the carboxylate group and double bond and Fig. 5(b) illustrates the plane perpendicular to this. The location of substituents or steric bulk are defined with respect to the TACA structure and these two planes. Substituents can be ‘above’ or ‘below’ the plane containing the double bond and carboxylate group, as shown in Fig. 5(a). Alternatively, they can be ‘behind’ or ‘in front’ of the perpendicular plane illustrated in Fig. 5(b).

In general, substitutions ‘in front of’ the double bond are tolerated but lead to antagonist activity, for example compounds 2-Me-TACA and THIP [33,37]. The exception to this rule is the case of 2-F-TACA, which is an agonist [33]. However, fluorine is a small substituent that is approximately isosteric with hydrogen, and thus may not be subject to the steric interactions that are postulated to occur for larger substituents [33]. Substitutions ‘behind’ of the double bond, on the other hand, result in agonist activity, as exemplified by muscimol [37]. The exception to this rule is the case of DFP, which is an antagonist [34] although, its steric bulk is projected behind the TACA double bond.

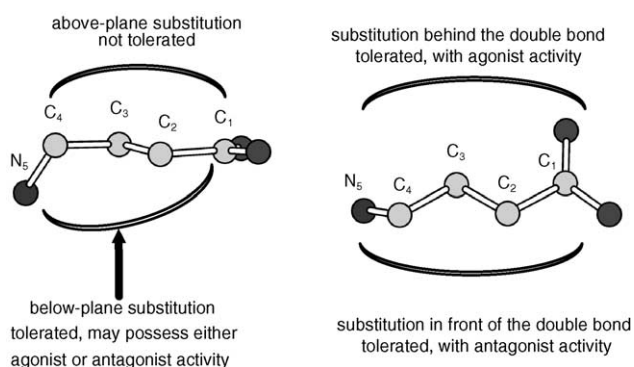


Fig. 5. Schematic diagram of the proposed pharmacophore defining agonist and antagonist activity at GABA_C receptors, using (a) a side-on view of the bioactive TACA conformation and (b) a top-down view of the bioactive TACA conformation.

However, DFP has very different steric and electrostatic properties to the other GABA_C receptor ligands considered here. Further investigation to elucidate the reason for this deviation may be useful in terms of designing novel agents with biological activity at GABA_C receptors.

The trends for above- and below-plane substitution are slightly different. It is postulated that substitution between C2 and N5 is tolerated if the substitution leads to steric bulk below the plane of the double bond but not tolerated if that substitution leads to steric bulk directed above the plane. In this case, the molecule will flex and/or reorient to avoid such interactions. This hypothesis explains both the difference in biological activity of (+)- and (–)-CAMP and the difference in potency of (+)- and (–)-TAMP. The optimal fit of (+)-CAMP to the template TACA molecule results in orientation of the cyclopropane ring substituent below and behind the TACA double bond, as shown in Fig. 3(a). Conversely, the optimal fit of (–)-CAMP to the template TACA molecule results in orientation of the cyclopropane ring above and behind the TACA double bond. However, as this arrangement is sterically unfavourable, the (–)-CAMP molecule will reorient and flex into an alternative orientation with the cyclopropane ring orientated below and in front of the TACA double bond, which represents only a marginally suboptimal fit. This superposition is illustrated in Fig. 3(b). Therefore, (+)-CAMP is an agonist at the GABA_C receptor while (–)-CAMP is an antagonist, according to the principles for in-plane substitution outlined above. The difference in biological activity of (+)-4-ACP-2-E and (–)-4-ACP-2-E is analogous to the difference in biological activity of (+)- and (–)-CAMP.

The optimal fit of (–)-TAMP to the template TACA molecule results in orientation of the cyclopropane ring directly below the TACA double bond (Fig. 3(c)), while the optimal fit of (+)-TAMP results in orientation of the cyclopropane ring directly above the TACA double bond. Again, this arrangement is not tolerated on the grounds of adverse steric interactions, and (+)-TAMP is hypothesized to reorient and flex to remove above-plane steric interactions. This results in the superposition of a different minimum energy structure in an orientation with a suboptimal fit to the template TACA molecule, but with the cyclopropane group orientated below the plane of the double bond. This superposition is shown in Fig. 3(d). As this fit does not result in steric bulk directed in front of the double bond, (+)-TAMP is still an agonist at GABA_C receptors, albeit less potent than (–)-TAMP.

Analogues with substitutions between the C3 and N5 positions exhibit similar behaviour. For example, isoguvacine, whose steric bulk falls behind the double bond, is an agonist at GABA_C receptors. Conversely, isonipicotic acid and THIP, whose steric bulk falls in front of the double bond, are antagonists. This also explains why I4AA and A2MBA possess both agonist and antagonist activity. For example, I4AA possesses two minimum energy conformers, which both have a similar RMSD value for their fit to

the template TACA molecule. Superposition of conformer 1 results in the molecule attaining the antagonist conformation while superposition of conformer 2 results in the molecule attaining the agonist conformation. A similar situation exists for A2MBA, although, A2MBA possesses six stable conformers. The torsional flexibility of these molecules means that they can access both the agonist and antagonist conformations.

The key features of the GABA_C receptor-binding site can be inferred from Fig. 5. The binding site must possess both a positively charged or polarized amino acid residue and a negatively charged, negatively polarized or aromatic residue, separated by a cavity which will optimally accommodate the TACA conformer shown in Fig. 2, with a C–N distance of 4.775 Å. This cavity must also possess a region of extruding steric bulk above the plane of the TACA double bond, creating an excluded volume. Further, although the cavity must accommodate steric bulk both in front of and behind the double bond of TACA, the interactions between the substituents and the protein must be significantly different on either side, with substitution behind the double bond leading to agonist activity and substitution in front of the double bond leading to antagonist activity.

4. Conclusions

In this work, we have constructed the first quantitative model for predicting biological activity at GABA_C receptors. According to this model, the agonist activities of GABA analogues at GABA_C receptors are correlated with the three-dimensional fit of the bioactive TACA structure to the most similar analogue conformer. In general, the closer the fit to the biologically active TACA conformer, the higher the affinity for the receptor and lower the EC₅₀ value. There is, however, an average root-mean-squared uncertainty of 32.4 μM in the predictions made by the linear regression model constructed in this work. Hence, this model is not expected to be sufficiently accurate to discriminate between compounds with similar biological activities but may facilitate discrimination between high and low affinity GABA_C receptor ligands. On the basis of this QSAR model we have, however, proposed a qualitative model describing key interactions at the GABA_C receptor-binding site that is consistent with all experimental data. This model has been used to gain insight into the structural and chemical features influencing the binding affinity and biological activity of 19 compounds experimentally identified as being active at GABA_C receptors. According to this model, agonist and antagonist activity can be determined by examining steric interactions at the binding site. Knowledge of the factors that determine biological activity and optimal receptor binding is expected to aid in the subsequent classification and development of GABA_C receptor ligands.

Acknowledgements

DLC acknowledges the financial support of an Australian Postgraduate Award.

References

- [1] H.P. Rang, M.M. Dale, J.M. Ritter, *Pharmacology*, fourth ed., Churchill Livingstone, Edinburgh, 1999, pp.478–480.
- [2] For a review of GABA_A and GABA_C receptors, see: M. Chebib, G.A.R. Johnston, *J. Med Chem.* 43 (2000) 1427.
- [3] For a review of GABA_B receptors, see: N.G. Bowery, S.J. Enna, *J. Pharmacol. Exp. Ther.* 292 (2000) 2.
- [4] G.A.R. Johnston, M. Chebib, J.R. Hanrahan, K.N. Mewett, *Curr. Drug Targets - CNS & Neuro. Disorders* 2 (2003) 260.
- [5] G.R. Cutting, L. Lu, B.F. O'Hara, L.M. Kasch, C. Montrose-Rafizadeh, P.M. Donovan, S. Shimada, S.E. Antonarakis, W.B. Guggino, G.R. Uhl, H.H. Kazazian, *Proc. Natl. Acad. Sci. USA* 88 (1991) 2673.
- [6] K. Wegelius, M. Pasternack, J.O. Hiltunen, C. Rivera, K. Kaila, M. Saarna, M. Reeben, *Eur. J. Neurosci.* 10 (1998) 350.
- [7] M.A. McCall, P.D. Lukasiewicz, R.G. Gregg, N.S. Peachey, *J. Neurosci.* 22 (2002) 4163.
- [8] For a review of GABA_A receptors see: P. Krogsgaard-Larsen, B. Frølund, F.S. Jørgensen, A. Schousboe, *J. Med. Chem.* 37 (1994) 2489.
- [9] R.K. Duke, R.D. Allan, M. Chebib, J.R. Greenwood, G.A.R. Johnston, *Tetrahedron Asymm.* 9 (1998) 2533.
- [10] D. Ragozzino, R.M. Woodward, Y. Murata, F. Eusebi, L.E. Overman, R. Mileli, *Mol. Pharmacol.* 50 (1996) 1024.
- [11] A. Feigenspan, H. Wassle, J. Bormann, *Nature* 361 (1993) 159.
- [12] A. Feigenspan, J. Bormann, *Eur. J. Pharmacol.* 288 (1994) 97.
- [13] J. Amin, D.S. Weiss, *Receptors Channels* 2 (1994) 227.
- [14] G.R. Cutting, S. Curristin, H. Zoghbi, B. O'Hara, M.F. Seldin, G.R. Uhl, *Genomics* 12 (1992) 801.
- [15] H.H. Qian, Y. Pan, *Mol. Brain Res.* 103 (2002) 62.
- [16] M. Bailey, B.E. Albrecht, K.J. Johnson, M.G. Darlison, *Biochim. Biophys. Acta Gene Struct. Express.* 1447 (1999) 307.
- [17] C. Arnaud, P. Gauthier, C. Gottesmann, *Psychopharmacology* 154 (2001) 415.
- [18] C. Gottesmann, *Neuroscience* 111 (2002) 231.
- [19] G.A.R. Johnston, P.M. Burden, K.N. Mewett, M. Chebib, *PCT International Application* (1998) Patent Number WO 98/58939.
- [20] J.R. Hanrahan, K.N. Mewett, M. Chebib, P.M. Burden, G.A.R. Johnston, *J. Chem. Soc., Perkin Trans. 1* (2001) 2389.
- [21] R.K. Wurtmann, R.K.K. Lee, *PCT International Application* (2002) Patent Number WO 98/58939 A2.
- [22] M.L. Lorenzini, L. Bruno-Blanch, G.L. Estiú, *J. Mol. Struct. (Theochem.)* 454 (1998) 1.
- [23] E. Gálvez-Ruano, I. Iriepa, A. Morreale, D.B. Boyd, *J. Mol. Graphics Model.* 20 (2001) 183.
- [24] M.L. Lorenzini, L. Bruno-Blanch, G.L. Estiú, *Int. J. Quant. Chem.* 70 (1998) 1195.
- [25] G. Costantino, A. Macchiarulo, A.E. Guadix, R. Pellicciari, *J. Med. Chem.* 44 (2001) 1827.
- [26] D.L. Crittenden, M. Chebib, M.J.T. Jordan, *J. Phys. Chem. A* 108 (2004) 203.
- [27] D.L. Crittenden, M. Chebib, M.J.T. Jordan, *J. Phys. Chem. A* 109 (2005) 4195.
- [28] M. Chebib, R.J. Vandenberg, W. Froestl, G.A.R. Johnston, *Eur. J. Pharmacol.* 329 (1997) 223.
- [29] Y. Murata, R.M. Woodward, R. Mileli, L.E. Overman, *Bioorg. Med. Chem. Lett.* 6 (1996) 2071.
- [30] M. Chebib, K.N. Mewett, G.A.R. Johnston, *Eur. J. Pharmacol.* 357 (1998) 227.
- [31] D. Krehan, B. Frølund, P. Krogsgaard-Larsen, J. Kehler, G.A.R. Johnston, M. Chebib, *Neurochem. Int.* 42 (2003) 561.
- [32] Y. Chang, D.F. Covey, D.S. Weiss, *Mol. Pharmacol.* 58 (2000) 1375.
- [33] M. Chebib, R.J. Vandenberg, G.A.R. Johnston, *British J. Pharmacol.* 122 (1997) 1551.
- [34] M. Chebib, G.A.R. Johnston, J.P. Mattsson, K. Rydström, K. Nilsson, J. Qiu, S.H. Stevenson, R.B. Silverman, *Bioorg. Med. Chem. Lett.* 9 (1999) 3093.
- [35] R.K. Duke, M. Chebib, V.J. Balcar, R.D. Allan, K.N. Mewett, G.A.R. Johnston, *J. Neurochem.* 75 (2000) 2602.
- [36] M. Chebib, R.K. Duke, R.D. Allan, G.A.R. Johnston, *Eur. J. Pharmacol.* 430 (2001) 185.
- [37] R.M. Woodward, L. Polenzani, R. Mileli, *Mol. Pharmacol.* 43 (1993) 609.
- [38] T. Kusama, C.E. Spivak, P. Whiting, V.L. Dawson, J.C. Schaeffer, G.R. Uhl, *British J. Pharmacol.* 109 (1993) 200.
- [39] D.L. Crittenden, A. Park, J. Qiu, R.B. Silverman, R.K. Duke, G.A.R. Johnston, M.J.T. Jordan, M. Chebib, *J. Neurochem.* submitted for publication.
- [40] J.A. Pople, J.S. Binkley, R. Seeger, *Int. J. Quant. Chem., Quant. Chem. Symp.* 10 (1976) 1.
- [41] R. Krishnan, J.A. Pople, *Int. J. Quant. Chem.* 14 (1978) 91.
- [42] R.J. Bartlett, D.M. Silver, *J. Chem. Phys.* 62 (1975) 3258.
- [43] R.J. Bartlett, G.D. Purvis, *Int. J. Quant. Chem.* 14 (1978) 561.
- [44] P.C. Hariharan, J.A. Pople, *Theoret. Chim. Acta* 28 (1973) 213.
- [45] M.M. Francl, W.J. Pietro, W.J. Hehre, J.S. Binkley, M.S. Gordon, D.J. DeFrees, J.A. Pople, *J. Chem. Phys.* 77 (1982) 3654.
- [46] T. Clark, J. Chandrasekhar, P.v.R. Schleyer, *J. Comput. Chem.* 4 (1983) 294.
- [47] R. Krishnan, J.S. Binkley, R. Seeger, J.A. Pople, *J. Chem. Phys.* 72 (1980) 650.
- [48] P.M.W. Gill, B.G. Johnson, J.A. Pople, M.J. Frisch, *Chem. Phys. Lett.* 197 (1992) 499.
- [49] A. Klamt, G. Schuurmann, *J. Chem. Soc., Perkin Trans. 2* (1993) 799.
- [50] V. Barone, M. Cossi, J. Tomasi, *J. Comput. Chem.* 19 (1998) 404.
- [51] R. Guha, P.C. Jurs, *J. Chem. Inf. Comput. Sci.* 44 (2004) 1440.