CHEMMEDCHEM

DOI: 10.1002/cmdc.200800198

Discovery of NMDA Glycine Site Inhibitors from the Chemical Universe Database GDB

Kong Thong Nguyen,[a] Salahuddin Syed,[a] Stephan Urwyler,[b] Sonia Bertrand,[c] Daniel Bertrand,[c] and Jean-Louis Reymond*[a]

Drug development faces the ever-increasing challenge of discovering new bioactive small molecules that have not already been investigated in the more than 150-year history of medicinal chemistry.[1] This problem is particularly critical for drug targets that require very small organic ligands, for which the range of potential structures is limited. One possible way to overcome this difficulty would be to search through the entire chemical space of these small molecules using virtual screening tools and to identify promising ligands for synthesis and testing. Herein we report the first example of such an approach for ligands of the *N*-methyl-D-aspartic acid (NMDA) receptor glycine site. This receptor is an important drug target implicated in synaptic plasticity, neuronal development, learning, and memory. Inhibiting the NMDA receptor may help prevent neuronal cell death caused by glutamate excitotoxicity in acute and chronic neurodegenerative disorders such as stroke, epilepsy, Huntington's, and Alzheimer's disease.[2] Starting with our recently reported chemical universe database (GDB) that lists all compounds of C, N, O, and F up to 11 atoms obeying simple stability and synthetic feasibility rules,[3] we show that virtual screening followed by synthesis and testing leads to several new NMDA glycine site ligands. Both ligand binding assays and functional investigations reveal that the identified ligands interact with the glycine binding site of the NDMA receptors, inhibiting receptor function by direct competition with glycine.

The crystal structure of the NMDA receptor glycine site was recently reported.[4] Glycine is bound through a series of hydrogen bonds at the bottom of a narrow channel, leaving limited free space. The known NMDA glycine site ligands are indeed very small analogues of glycine, such as D-alanine and D-serine,[5] D-cycloserine,[6] and small cyclic amino acids,[7] suggesting a restricted structural range to explore (Figure 1). Despite these constraints, we hypothesized that new ligands for this receptor might be found by screening our chemical universe database (GDB), which is rich in very small organic molecules. Screening was undertaken in three stages, starting with a substructure-based selection of promising ligands from GDB, followed by high-throughput docking to select the best virtual hits, and finally synthesis and testing of a few selected compounds for their capacity to displace the binding of a glycine site radioligand, and for their action on the functional properties of NMDA receptors expressed in *Xenopus* oocytes.

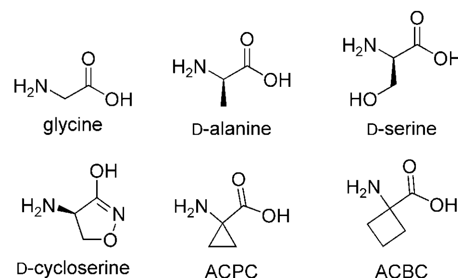

Figure 1. Known NMDA glycine site ligands.

The first stage of selection was realized by using a 2D chemical substructure fragment-based Bayesian classifier, which determines a bioactivity probability score for any compound from the product of the relative frequency of occurrence of all its substructures in known active versus inactive compounds.[8] As the active set we collected 61 ligands reported to interact with the NMDA receptor. This selection included ligands that interact with the glycine site, the glutamate site, or with an undetermined site of action, thus covering relatively diverse structural types as opposed to the much narrower selection that would have resulted by considering only confirmed glycine site ligands, which are almost exclusively α-amino acids. The inactive set was composed of a selection of 190 371 small organic molecules from ChemACX.[9] The classifier was trained and applied to screen the non-small-ring version of GDB (14.7 million structures), which returned 15 061 hits. These hits were then expanded to 69 367 stereoisomers using CORINA,[10] and docked into the NMDA glycine site using AutoDock 3.0.5.[11] The docking program positioned glycine in the crystallographically observed position, confirming the accuracy of conformational sampling as usually observed with such programs.[12] A typical Gaussian distribution of docking energies was observed across the library, with glycine appearing among the best 1.03 % of the ligands (Figure 2). The known reference ligands (Figure 1) also docked stronger than glycine, suggesting that the scoring and ranking of the docked conformations could be used for selecting ligands.

The top docking compounds were almost all acyclic, in agreement with the narrow binding channel available at the glycine site. Carboxylic acids were strongly enriched through the docking selection owing to energetically favorable salt-bridge interactions with the side chain of Arg 131(A) at the

[a] K. T. Nguyen, Dr. S. Syed, Prof. Dr. J.-L. Reymond
Department of Chemistry and Biochemistry, University of Bern
Freiestrasse 3, 3012 Bern (Switzerland)
Fax: (+41) 31-631-8057
E-mail: jean-louis.reymond@ioc.unibe.ch

[b] Dr. S. Urwyler
Novartis Institutes for BioMedical Research, Neuroscience
4002 Basel (Switzerland)

[c] S. Bertrand, Prof. Dr. D. Bertrand
Department of Neuroscience, Medical Faculty, University of Geneva
1 rue Michel Servet, 1211 Geneva 4 (Switzerland)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cmdc.200800198.

 WILEY InterScience DISCOVER SOMETHING GREAT

© 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

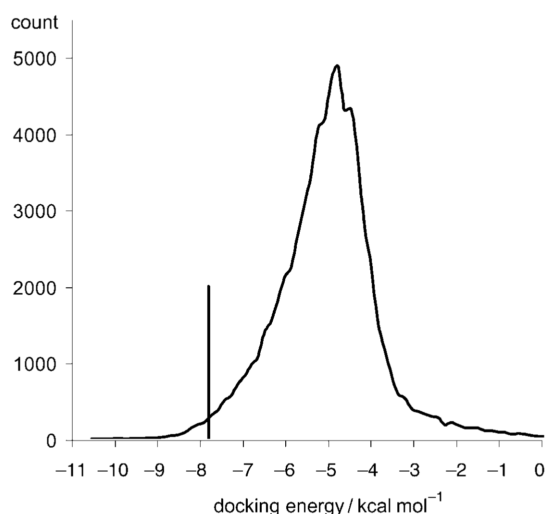

Figure 2. Distribution of docking energies for the 69367 docked stereoisomers. Stereoisomers were generated from SMILES using CORINA and docked into the glycine binding site of the NMDA receptor (PDB code: 1PB7) using AutoDock 3.0.5 for 10 cycles. In each case the energy of the most favorable pose was used as the docking energy (DE). Glycine (DE $=-7.82$ kcal mol^{-1}, vertical line) docked into the crystallographically observed position.

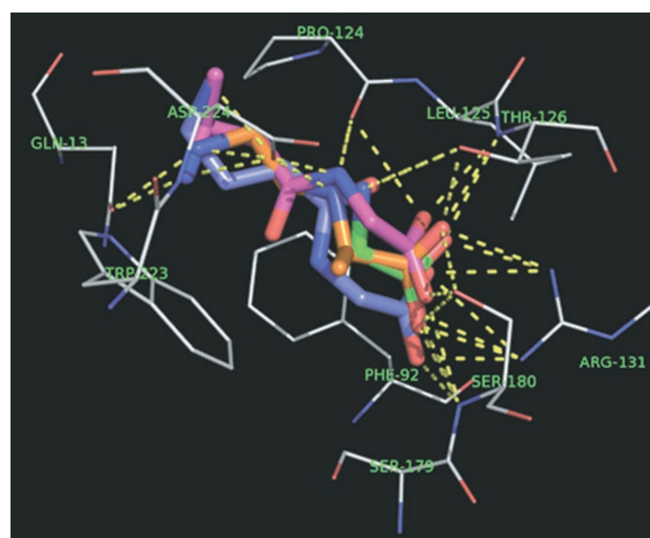

Figure 3. Binding modes within the NMDA glycine site (PDB code: 1PB7). Glycine (green); virtual hit **1** (blue); virtual hit **2** (mauve); virtual hit **7** (orange). Residue numbering according to PDB code: 1PB7. Arg 131 corresponds to Arg 523 in Ref. [4].

bottom of the cleft (Table 1, Figure 3).[4,13] Twenty-three of the top-docking ligands were selected for closer investigation, including six commercially available compounds and 17 additional molecules, which were synthesized (Figure 4). Although most of these compounds were known, none of them had been reported for NMDA glycine site binding.

The 23 selected compounds were tested by radioactive ligand displacement assays for the NMDA glycine site, the NMDA glutamate site, and the AMPA receptor (Table 2). In these assays, the positive controls glycine and L-glutamate, as well as D- and L-alanine displayed the expected binding affinities. Five ligands (22% of the tested compounds) showed detectable binding to the NMDA glycine site, including the strongest docking virtual hit **7**, D-diaminopropionic acid, amino isobutyric acid, and the dipeptides **1** and **2**. Although binding was relatively weak (IC_{50}~80 μM), the interaction was selective for the NMDA glycine site, with the exception of compound **2**, which also bound to the NMDA glutamate site. The compounds showed no detectable binding at the AMPA receptor.

Dipeptides **1** and **2**, which showed the best affinities of the tested compounds and suggested a possible lead structure, were selected for further study. Five dipeptides, **18**–**22**, featuring variations of the amino acid components, were prepared by coupling Boc-γ-aminobutyric acid or Boc-β-alanine with the

Table 1. Numbers and percentages of compound types in the various libraries during virtual screening for NMDA glycine site ligands.

Library	GDB	GDB No Small Rings[a]	Known Actives[b]	Inactives (ACX)[c]	Bayesian Hits[d]	Top Docking[e]	Tested
Total of library	26434571	14763172	61	190371	15061	712	23
Monoamines[f]	28.8%	20.5%	31.1%	7.0%	21.7%	20.8%	21.7% (5)
Diamines[f]	7.2%	3.6%	1.6%	0.8%	35.0%	25.4%	13.0% (3)
Amino acids[g]	0.29%	0.26%	23.0%	1.5%	7.3%	22.9%	39.1% (9)
Diamino acids[h]	0.05%	0.033%	1.6%	0.08%	6.3%	18.1%	21.7% (5)
Amino diacids[g]	0.0014%	0.002%	11.5%	0.20%	0.37%	1.8%	4.4% (1)
Triamines[f]	0.95%	0.380%	3.3%	0.072%	17.5%	4.2%	0
Triamino acids[i]	0.0028%	0.0019%	0%	0.006%	1.06%	1.7%	0
Others[j]	62.6%	75.2%	27.9%	90.3%	10.8%	5.1%	0
Acyclics	4074634	4074634	18	18432	4255	614	20
Cyclics	22359937	10688538	43	171939	10806	98	3

[a] All structures of GDB that do not contain any 3- or 4-membered rings. [b] 61 ligands reported to be active on the NMDA receptor were collected; these include ligands for the glycine site, the glutamate site, and ligands with no established site of action. [c] Organic compounds extracted from the ChemACX compound collection (CambridgeSoft version 8) with $M_r<500$ and containing only C, N, O, H, and halogen (F, Cl, Br, I) atoms, excluding the 61 known actives. [d] Ligands with a Bayesian score higher than the highest-scoring compounds from the inactive set. [e] Ligands with a more negative estimated docking energy than glycine as estimated by AutoDock 3.0.5 (see Figure 1). [f] Structures without carboxylic groups and for which the number of nitrogen atoms with only H or saturated C atom neighbors is exactly one (monoamines), two (diamines), or three (triamines). [g] Monoamines with exactly one (amino acids) or two (amino diacids) carboxylic groups (CO_2H). [h] Diamines with exactly one carboxylic group. [i] Triamines with exactly one carboxyl group; the presence of other functional groups, including amides interrupting the carbon chain, are not considered in this classification. [j] Guanidines are particularly frequent in the known actives, Bayesian hits, and top docking lists.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

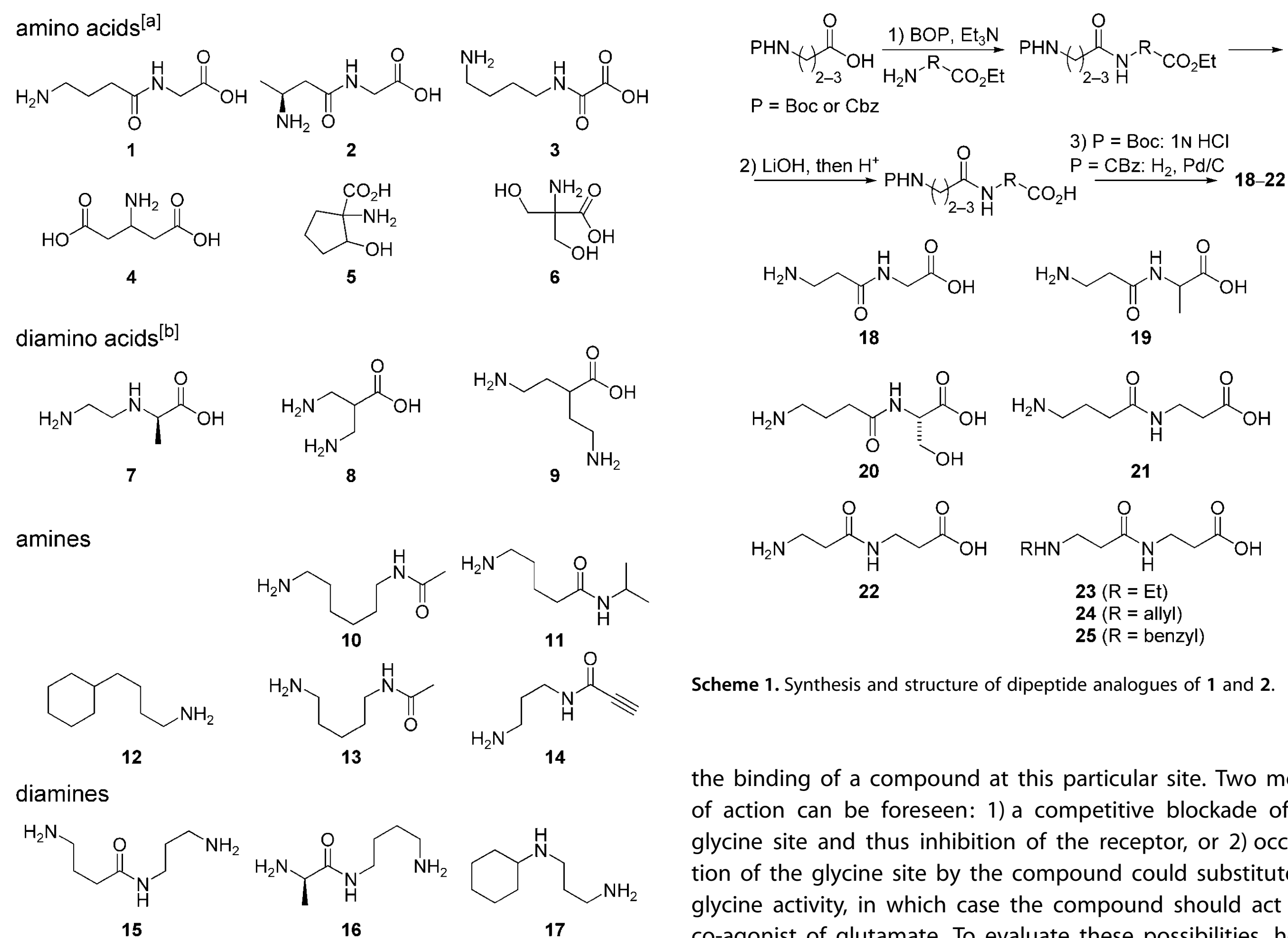

Figure 4. GDB compounds selected for synthesis and testing. (See Supporting Information for synthesis procedures; docking poses are shown in Figures S1–S4). [a] The presence of other functional groups, including amides that interrupt the carbon chain, is not considered for classification. The following virtual hits were purchased: amino acids: D-2,3-diaminopropanoic acid (D-Dap), β-alanine, GABA, aminoisobutyric acid (Aib), 2-aminobutyric acid. [b] Diamino acids: L-Lysine. The virtual hits D- and L-alanine of known activity were also included as additional controls. Compound **8** was prepared as previously described.[14]

Scheme 1. Synthesis and structure of dipeptide analogues of **1** and **2**.

ethyl esters of glycine, β-alanine, L-serine, or D,L-alanine followed by saponification of the ethyl ester and acidic removal of the Boc group (Scheme 1). Whereas most compounds obtained by this route were pure, compound **20** contained small amounts of serine (~2% as detected by amino acid analysis) formed during the final Boc deprotection step, as had been observed for a synthesis of **1** via the Boc route (0.5% glycine after Boc cleavage).[15] Compound **20** was therefore prepared similarly to **1** starting with CBz-protected γ-aminobutyric acid with final deprotection by hydrogenation, giving the pure product (see Supporting Information). To our delight, compounds **18**–**22** also bound to the NMDA glycine site in the radioactive ligand displacement assays, with affinities similar to those of the initial virtual hits **1** and **2**, confirming the activity observed with this type of compound (Table 2).

Because it is well documented that glycine is a co-agonist at the NMDA receptor,[16] it is crucial to evaluate the outcome of the binding of a compound at this particular site. Two modes of action can be foreseen: 1) a competitive blockade of the glycine site and thus inhibition of the receptor, or 2) occupation of the glycine site by the compound could substitute for glycine activity, in which case the compound should act as a co-agonist of glutamate. To evaluate these possibilities, heterologous expression in *Xenopus* oocytes was carried out by the injection of cDNAs encoding the glycine binding subunit NR-1 together with the cDNA encoding the glutamate binding subunits NR-2A or NR-2B. Oocytes expressing the desired NMDA receptor were first challenged with a solution containing both glutamate and glycine that provided the reference agonist evoked current. Effects of compounds were then tested in the presence of glutamate and glycine (same as control) to evaluate the competition with glycine, then glutamate was applied alone followed by exposure of glutamate + test compound to detect any possible co-agonist activity (respectively traces A, B, C, and D in Figure 5a). Typical glutamate-evoked currents recorded with this experimental paradigm are illustrated in Figure 5a and 5b.

The functional assays were performed for virtual hits **1** and **2** and their analogues **18**–**22**. Five of these seven compounds inhibited the signal in the presence of both glycine and glutamate with either one or both receptor subtypes, with IC_{50} values in the range of 100–500 μM (Table 3). The different and somewhat higher IC_{50} values from the oocyte assay relative to those from the radioligand assay probably reflect more stringent competition by glycine than by the radioligand, but might also indicate a different response of the NMDA receptor in the two assays. Functional testing of further derivatives showed that alkylation of the simple β-Ala-β-Ala dipeptide **22** at the terminal amine to provide more hydrophobic and possibly druglike analogues gave additional active ligands **23**–**25**.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 2. Radioactive ligand displacement assays with selected virtual hits from GDB and analogues.[a]

Compound	NMDA receptor glutamate site ([3H]CGP39653)			NMDA receptor glycine site ([3H]MDL-105519)			
	pIC_{50}	IC_{50} [μM]	N[b]	pIC_{50}	IC_{50} [μM]	nH[c]	N[b]
L-Glu	6.81±0.29	0.16[d]	3	<4	>100		2
Glycine	<4.5	>30	3	6.16±0.13	0.69	0.98±0.07	3
D-Ala	<5	>10	2	6.02±0.07	0.96	0.76±0.05	3
L-Ala	<4	>100	2	4.03±0.04	93	1.13±0.1	3
1	<4	>100	2	4.19±0.13	65	1.19±0.04	3
2	4.14±0.06	72[e]	3	4.55±0.06	28	1.05±0.03	3
7	<4	>100	2	3.95±0.04	112	1.14±0.05	3
D-Dap	<4	>100	2	4.14±0.09	72	0.99±0.1	3
AiB	<4	>100	3	<4	(44±5%)[f]		3
18	<4	>100	2	4.08±0.01	83	1.34±0.05	3
19	<4	>100	2	3.86±0.03	138	1.48±0.08	3
20	<4	>100	2	4.62±0.01	24	0.97±0.12	3
21	<4	>100	2	<4	(24±1%)[f]		2
22	<4	>100	2	4.06±0.03	87	1.07±0.23	3

[a] All ligands shown in Figure 4 were tested; only ligands showing activities are listed. The ligands were also tested on the AMPA receptors and found to have no effect (except for L-glutamate). [b] Number of independent measurements. [c] nH is the Hill coefficient. [d] nH = 1±0.15. [e] nH = 1.34±0.47. [f] Percent inhibition at 100 μM. Assays were performed on native rat brain membrane receptor preparations; see Supporting Information for detailed procedures.

Table 3. Electrophysiological data on the NMDA receptor expressed in *Xenopus* oocyte membranes.[a]

Compd	IC_{50} [μM] on NR1+NR2A	IC_{50} [μM] on NR1+NR2B
1	40% at 30 μM[b]	no inh.[c]
2	311±74	97±9
18	470±105	166±4
19	no inh.[c]	498±115
20	no inh.[c]	no inh.[c]
21	no inh.[c]	163±100
22	70% at 30 μM[b]	153±19
23	60±7	70±22
24	168±40	100±35
25	120±10	79±9

[a] cDNAs encoding NR1 and NR2A or NR2B were injected in equal amounts (1:1). Experimental protocol corresponds to that of Figure 5c and 5d in the presence of 100 μM glutamate +10 μM glycine. The data are the average of three independent determinations of inhibition at 0, 30, 100, and 300 μM compound. [b] The signal stabilizes at the indicated level at the higher concentrations. [c] No significant signal inhibition was observed.

The *N*-ethyl derivative **23** in particular, with an IC_{50} value of 60 μM, suggests a possibility for lead optimization (Figure 5c and 5d, and Table 3).

In summary, the above experiments show for the first time drug discovery by virtual screening of an exhaustive representation of chemical space in the form of the chemical universe database (GDB), followed by synthesis and testing. Despite their deceptively simple structures, the identified ligands, in particular the dipeptide inhibitors **1** and **2** and analogues such as **22**, represent a class of previously unknown NMDA glycine site inhibitors. The activity is conserved in the more hydrophobic derivatives **23**–**25**, suggesting that optimization of this lead series might be possible. The application of GDB to the discovery of new ligands for other receptors is ongoing and will be reported in due course.

Acknowledgements

Financial support from the Swiss National Science Foundation and access to the Molinspiration virtual screening toolkit are gratefully acknowledged.

Keywords: computer chemistry · inhibitors · receptors · virtual screening

[1] K. H. Bleicher, H.-J. Böhm, K. Müller, A. I. Alanine, *Nat. Mater. Nat. Rev. Drug Discovery* **2003**, *2*, 369–378.

[2] a) T. G. Smart, *Curr. Opin. Neurobiol.* **1997**, *7*, 358–367; b) D. W. Choi, S. M. Rothman, *Annu. Rev. Neurosci.* **1990**, *13*, 171–182; c) J. W. Olney, *Annu. Rev. Pharmacol. Toxicol.* **1990**, *30*, 47–71; d) J. W. Olney, J. Labruyere, G. Wang, D. F. Wozniak, M. T. Price, M. A. Sesma, *Science* **1991**, *254*, 1515–1518; e) J. A. Kemp, R. M. McKernan, *Nat. Neurosci.* **2002**, *5*, 1039–1042; f) G. P. Franks, *Nat. Rev. Neurosci.* **2008**, *9*, 370–386.

[3] a) T. Fink, H. Bruggesser, J.-L. Reymond, *Angew. Chem.* **2005**, *117*, 1528–1532; *Angew. Chem. Int. Ed.* **2005**, *44*, 1504–1508; b) T. Fink, J.-L. Reymond, *J. Chem. Inf. Model.* **2007**, *47*, 342–353.

[4] H. Furukawa, E. Gouaux, *EMBO J.* **2003**, *22*, 2873–2885.

[5] N. W. Kleckner, R. Dingledine, *Science* **1988**, *241*, 835–837.

[6] W. F. Hood, R. P. Compton, J. B. Monahan, *Neurosci. Lett.* **1989**, *98*, 91–95.

[7] G. B. Watson, T. H. Lanthorn, *Neuropharmacology* **1990**, *29*, 727–730.

[8] a) *Miscreen–Molinspiration Cheminformatics Virtual Screening Toolkit*, 2005.03, Molinspiration Cheminformatics: Slovensky Grob, Slovak Republic, 2005; b) A. Bender, H. Y. Mussa, R. C. Glen, S. Reiling, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 170; c) J. Hert, P. Willett, D. J. Wilton, P. Acklin, K. Azzaoui, E. Jacoby, A. Schuffenhauer, *Org. Biomol. Chem.* **2004**, *2*, 3256-3266; d) P. Japertas, R. Didziapetris, A. Petrauskas, *MiniRev. Med. Chem.* **2003**, *8*, 797- 808.

[9] *ChemACX Ultra*, version 8.0: CambridgeSoft Corporation, Cambridge, MA (USA), 2005.

[10] J. Sadowski, J. Gasteiger, *Chem. Rev.* **1993**, *93*, 2567–2581.

[11] a) G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639–1662; b) S. F. Sousa, P. A. Fernandes, M. J. Ramos, *Proteins Struct. Funct. Gent.* **2006**, *65*, 15–26.

[12] a) G. L. Warren, C. W. Andrews, A.-M. Capelli, B. Clarke, J. LaLonde, M. H. Lambert, M. Lindvall, N. Nevins, S. F. Semus, S. Senger, G. Tedesco, I. D.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

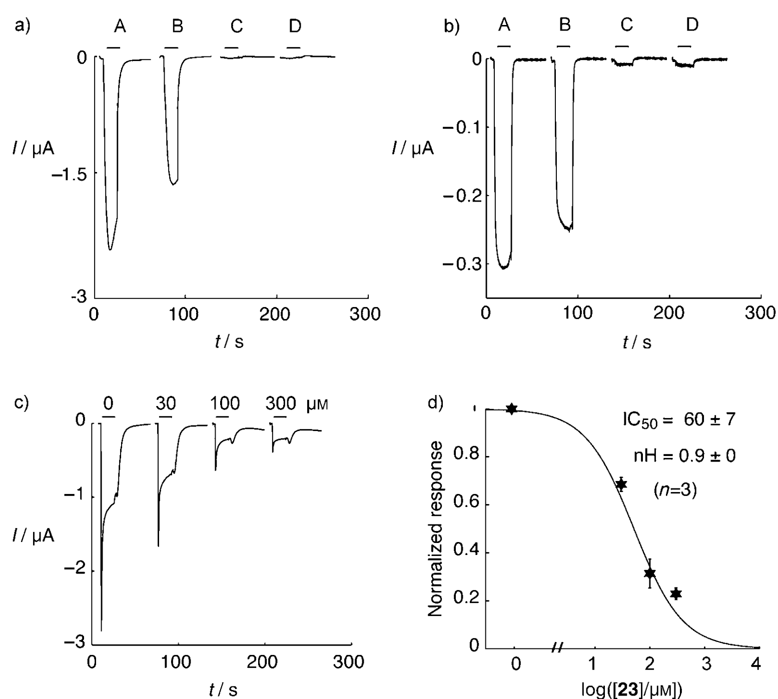

Figure 5. Inhibition of the electrophysiological responses of the NMDA receptor NR1/NR2A in *Xenopus* oocyte membranes. All experiments were carried out with rodent NMDA receptors (see Supporting Information). a) Effects of compound **1**: A: 10 μM glutamate + 1 μM glycine; B: 10 μM glutamate + 1 μM glycine + 10 μM **1**; C: 10 μM glutamate; D: 10 μM glutamate + 10 μM compound **1**. b) Effect of compound **2** (series A–D as for **1**). c) and d) Inhibition by the *N*-ethyl derivative **23** in the presence of 100 μM glutamate + 10 μM glycine.

Wall, J. M: Woolven, C. E. Peishoff, M. S. Head, *J. Med. Chem.* **2006**, *49*, 5912–5931; b) P. C. D. Hawkins, A. G. Skillman, A. Nicholls, *J. Med. Chem.* **2007**, *50*, 74–82; c) N. Moitessier, P. Englebienne, D. Lee, J. Lawandi, C. R. Corbeil, *Br. J. Pharmacol.* **2008**, *153*, S7–26.

[13] C. I. Schneider, S. Urwyler, *Biochem. Pharmacol.* **1992**, *43*, 1693–1699.

[14] D. Lagnoux, E. Delort, C. Douat-Casassus, A. Esposito, J.-L. Reymond, *Chem. Eur. J.* **2004**, *10*, 1215–1226.

[15] Compound **1** prepared by the Boc route and containing 0.5% glycine gave an apparent IC_{50} value of 20 μM in the binding assay and a positive activation signal together with glutamate in the absence of glycine in the electrophysiological assay. The data reported herein are for the product **1** from the CBz route, which is free of any trace of glycine.

[16] a) G. Henderson, J. W. Johnson, P. Ascher, *J. Physiol.* **1990**, *430*, 189–212; b) P. L. Wood, *Life Sci.* **1995**, *57*, 301–310.

Received: June 30, 2008
Revised: July 22, 2008
Published online on September 15, 2008

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim