Bioorganic & Medicinal Chemistry Letters 19 (2009) 3832–3835

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

3-(Aminomethyl)piperazine-2,5-dione as a novel NMDA glycine site inhibitor from the chemical universe database GDB

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

[a] Department of Chemistry and Biochemistry, University of Berne, Freiestrasse 3, 3012 Berne, Switzerland
[b] Department of Neuroscience, Medical Faculty, 1, rue Michel Servet CH-1211 Geneva 4, Switzerland

ARTICLE INFO

Article history:
Received 18 February 2009
Revised 1 April 2009
Accepted 3 April 2009
Available online 10 April 2009

Keywords:
Virtual screening
Glutamate receptors
Diketopiperazine
Neurochemistry
Docking

ABSTRACT

Docking of randomly selected compounds from the chemical universe database GDB-11, which contains all organic molecules up to 11 atoms of C, N, O, F possible under consideration of simple chemical stability and synthetic feasibility rules, into the NMDA receptor glycine site (1pb7.pdb) lead to the identification of 3-(aminomethyl)piperazine-2,5-dione **3** and its close analog 5-(aminomethyl)piperazine-2,3-dione **4** as possible new ligands for this drug target, which is implicated in synaptic plasticity, neuronal development, learning and memory. Synthesis of these compounds in 4 and 6 steps, respectively, and testing by radioligand displacement assays and electrophysiological measurements in *Xenopus* oocytes show that while **4** is inactive, **3** is indeed an inhibitor of glycine, with an estimated K_D of 50 μM.

In small molecule drug discovery one attempts to find the best possible ligands for a given biological target.[1] Ideally it should be possible to test the ensemble of all possible organic molecules to find such ligands. In our efforts towards this goal, we recently assembled the chemical universe database GDB-11, which contains 26.4 million molecules up to 11 atoms of C, N, O, and F possible under consideration of simple chemical stability and synthetic feasibility rules.[2] We have shown that active compounds can be readily found in GDB, as in the example of dipeptides **1** and **2** identified as inhibitors of the NMDA receptor glycine site by virtual screening, synthesis and testing.[3] This receptor is an important drug target implicated in synaptic plasticity, neuronal development, learning and memory for which inhibitors might have clinical application.[4]

The discovery of **1** and **2** provided an important proof-of-principle for using GDB in drug discovery. However the ligands were very similar to known inhibitors of the NMDA glycine site such as D-alanine and D-serine,[5] D-cycloserine,[6] and small cyclic amino acids,[7] which was to be expected since compound selection had been guided by structural analogy to these molecules. Herein we report the identification of 3-(aminomethyl)piperazine-2,5-dione **3** and its close analog 5-(aminomethyl)piperazine-2,3-dione **4** by direct virtual screening of GDB-11 using docking to the NMDA receptor glycine site (Fig. 1). In these compounds the key carboxylic group present in amino acid inhibitors and interacting with Arg131 is replaced by the diketopiperazine pharmacophore. Synthesis and activity testing by radioligand displacement assays and functional testing on the NMDA receptor expressed in *Xenopus* oocytes shows that **3** indeed binds to the NMDA glycine site and inhibits glycine with an estimated K_D of 50 μM. This experiment shows for the first time that ligand discovery from GDB-11 is possible on the basis of the target protein structure only without prior knowledge of other small molecule ligands.

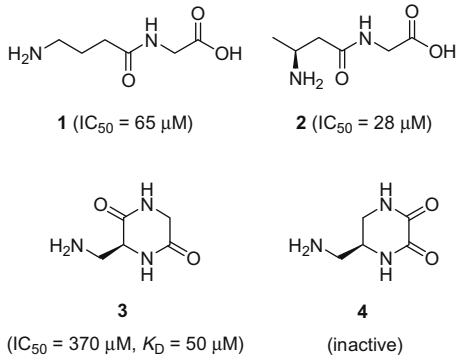

Figure 1. NMDA receptor glycine site virtual hits **1**–**2** (Ref. 3) and **3**–**4** (this work) identified by virtual screening of GDB-11. **1**, **2** and **3** bind and inhibit the receptor at the glycine site, **4** is not active.

* Corresponding author. Fax: +41 31 631 80 57.
E-mail address: jean-louis.reymond@ioc.unibe.ch (J.-L. Reymond).

0960-894X/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2009.04.021

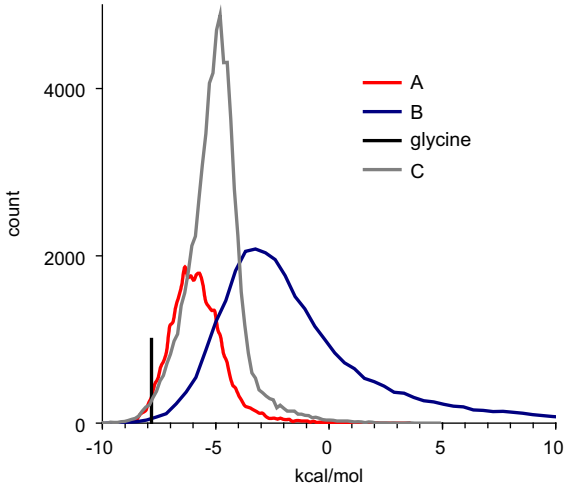

Figure 2. Distribution of docking energies of (A, red line) the 31,121 stereoisomers generated from 8000 randomly selected GDB structures; (B, blue line) the 81,877 stereoisomers generated from the 35,810 virtual hits from the Bayesian classifier trained with the docking results of A; (C, grey line) the 69,367 stereoiosmers generated from the 15,061 virtual hits from the Bayesian classifier trained with known NMDA-receptor inhibitors, as described in Ref. 3. The black vertical line indicates the DE of glycine (−7.82 kcal/mol). Stereoisomers were generated from SMILES using CORINA and docked into the glycine binding site of the NMDA-receptor (pdb: 1PB7)[20] using AUTODOCK3.0.5 for 10 cycles. In each case the energy of the most favourable pose was used for scoring.

In our search for NMDA glycine site inhibitors from GDB-11, we hypothesized that innovative ligands might be found if virtual screening would be guided from the protein structure only without using information from existing ligands. With docking as our main tool,[8] we first screened a randomly selected subset of 8000 GDB structures. The structures were expanded to 31,121 stereoisomers using CORINA[9] and evaluated by AUTODOCK3.0.5,[10] resulting in a typical gaussian curve for the distribution of docking energies (Fig. 2). There were 702 molecules docking better than glycine, among which amino acids were significantly enriched, showing the natural tendency of the binding pocket to select for that class of compounds.

A Bayesian classifier was applied next to extract further ligands from GDB. Bayesian classifiers determine 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, and are computationally very fast and therefore applicable to screen large databases.[11] 141 of the top docking hits were selected as actives to train the classifier, considering structures that would be relatively straightforward to prepare, in particular amides, while complex combinations of polyenes, amidines and hydrazones were rejected. All compounds docking weaker than glycine were used as inactives. The classifier returned 35,810 additional compounds from GDB scoring higher than the lowest Bayesian score for the active set, from which 182 compounds docked stronger than glycine.

Analysis of functional group enrichment showed that docking did not select for monoamines, but strongly enriched amino acids and amides, which is similar to our previously reported study[3] (Table 1). Amides were very frequent in the final hit set because we chose this functional group for ease of synthesis in the selection of top-docking compounds from the first round of docking. Docking also enriched hydrogen bonding donor (HBD) and acceptor sites (HBA). The random GDB selection used in this study had fewer HDB (2.08) than the compounds previously selected on the basis of known actives (5.45), but significantly more HBA (3.17 vs 2.13). The higher HBA in the GDB random set compared to the Bayesian hit sets might explain its better mean docking energy reflecting high-scoring interactions with donor sites on the proteins.

As in our previous study[3] virtual screening favoured acyclic compounds, with many virtual hits showing structures reminiscent of **1** and **2**. On the other hand the hit set also contained a few six-membered ring heterocycles, in particular the intriguing diketopiperazine ligands (*S*)-**3** and (*S*)-**4** at rank #58 and #20, and docking energies of −8.36 kcal/mol and −8.81 kcal/mol, respectively. In their best docking pose, the key carboxylate pharmacophore of known amino acid inhibitors was replaced by the diketopiperazine group, and the primary amino group, although placed quite differently from that of glycine, interacted with anionic residues similarly to the amino group of dipeptides **1** and **2** (Fig. 3). Since **3** and **4** had never been reported previously, we set out to synthesize and test these ligands against the NMDA glycine site.

Diketopiperazine **3** was prepared as the (*S*) enantiomer in four steps and 33% overall yield (Scheme 1). Peptide coupling between CBz-protected L-asparagine (**5**) and glycine methyl ester gave dipeptide **6**. Oxidative Hofmann degradation[12] of the asparagine carboxamide side-chain using *I,I*-bis(trifluoroacetoxy)iodobenzene[13] and protection of the primary amine gave the Boc-derivative **7**.[14] Hydrogenation of the CBz group of the terminal α-amino group and intramolecular cyclization under mild basic

Table 1
Enrichment statistics during virtual screening

Library	GDB-11 Random set	1st Round[a] top docking	Selected[b]	Bayesian hits[c]	2nd round[a] top docking	Ref. 3 Bayesian hits[d]	Ref. 3 top docking[a]
Size	8000	702	141	35,810	182	15,061	712
Monoamines[e] %	25.8	28.4	42.6	33.8	24.7	21.7	20.8
Aminoacids[f] %	0.36 (29)[e]	1.00 (7)	2.1 (3)	0.42	18.1 (150)	7.3 (1100)	22.9 (163)
Amides[g] %	9.2	16.0	24.8	23.3	41.8	3.6	10.3
HBD average[h]	2.08	3.32	1.59	2.29	2.82	5.45	5.25
HBA average[i]	3.17	3.38	2.84	2.35	4.94	2.13	2.63
Acyclics %	27.5	71.8	55.3	55.9	91.2	28.3	86.2

[a] Molecules of the random set (1st round) or from the Bayesian hits in this work (2nd round) or from Ref. 3 docking stronger than the crystallographic ligand glycine (DE = −7.83 kcal/mol) in the NMDA glycine site binding pocket as estimated by AUTODOCK 3.0.5.
[b] Subset of 1st round top docking with simple functional groups.
[c] All compounds from GDB-11 scoring better than the worst of the selected compounds in a Bayesian classifier trained with the 141 selected compounds as actives and all structures with DE >−7.82 kcal/mol as inactives.
[d] Bayesian hits from Ref. 3, in this case the classifier was trained with known actives and an ACX random set as inactives.
[e] Monoamines are molecules without carboxylic groups and exactly one nitrogen atom with only H or saturated C-atoms neighbours.
[f] Amino acids are monoamines with exactly one carboxylic acid. The absolute number of molecules is indicated in parentheses.
[g] Amides are molecules with at least one amide function.
[h] HBD is the hydrogen bond donor site count, that is, the total number of NH and OH bonds.
[i] HBA is the hydrogen bond acceptor site count, that is, the total number of lone pairs.