Pergamon

SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry Letters 13 (2003) 3213–3216

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

A Convenient Method for the Computer-Aided Molecular Design of Carborane Containing Compounds

Jayaseharan Johnsamuel,[a,*] Youngjoo Byun,[a] Thomas P. Jones,[b] Yasuyuki Endo[c] and Werner Tjarks[a]

[a]*The Ohio State University, College of Pharmacy, 500 W. 12th Avenue, Columbus, OH 43210, USA*
[b]*Tripos Inc., 1699 South Hanley Road, St. Louis, MO 63144, USA*
[c]*Faculty of Pharmaceutical Sciences, Tohoku Pharmaceutical University, 4-4-1, Komatsushima, Aoba-ku, Sendai 981-8558, Japan*

Received 22 April 2003; revised 19 June 2003; accepted 24 June 2003

Abstract—Computer-aided molecular design (CAMD) of carborane containing compounds is of growing interest for scientists involved in boron neutron capture therapy (BNCT) and other pharmaceutical applications. However, the complex organo-metallic structures of carboranes pose difficulties in modeling and docking of these structures. This is the first report of a new strategy for modeling and docking of carborane containing molecules with the readily available software packages HyperChem, SYBYL and FlexX. It is intended as a guide for boron chemists interested in using CAMD of carborane containing agents for medical applications such as BNCT.

Due to their high boron content and chemically modifiable properties, carborane clusters have been a preferred choice as boron moieties in the design and synthesis of boronated agents for boron neutron capture therapy (BNCT).[1,2] By analogy with fullerenes,[3,4] carboranes have recently also been used as hydrophobic pharmacophores in the design of derivatives of biologically active compounds such as estradiol, retinoic acid, and teleocidin.[5–8] Some of these carboranyl analogues interacted effectively with the corresponding receptor enzymes and exhibited equal or even higher biological activity compared to their endogenous counterparts. These carborane constructs were easily synthesized with few reaction steps, thus, providing feasible synthetic routes to biologically active carboranyl agents.

So far, carborane-based drug design utilizing computer-aided molecular design (CAMD) has been applied only in relatively few cases.[5–9] This is mainly due to the complex structures of carboranes with 6-fold coordinated carbon and boron atoms, and the unavailability of potential energy functions for the boron atom in most of the commercially available software packages.[10]

The earlier reports on CAMD of carboranyl derivatives of estradiol,[6] retinoic acid,[5] and teleocidin[8] used a software package ADAM,[5,6,8,11] which is not readily accessible to scientist interested in CAMD.

In this paper, a general strategy for modeling and docking of carborane containing derivatives using readily available software packages HyperChem. 5.1, SYBYL 6.8 and FlexX is described in detail. This strategy is intended as a guide for synthetic boron chemists interested in CAMD of carborane containing agents for medical applications. We have also addressed the general applicability of CAMD in modeling studies.

General modeling and docking strategies: The protein data bank file (PDB) entitled 1ERE (3.1 Å structure resolution) of the human estrogen receptor protein (hERαLBD)[12] was obtained from the protein data bank [Research Collaboratory for Structural Bioinformatics (RCSB) (http://www.rcsb.org/pdb)]. Compounds **1**–**4** (Fig. 1) were used for docking.

HyperChem. 5.1 (HYPERCUBE, Inc., 419 Phillip Street, Waterloo, Ontario N2L 3X2, Canada) and SYBYL 6.8 (TRIPOS Inc., 1699 South Hanley Rd., St. Louis, MO, 63144, USA) were used for modeling. Estradiol (**1**) was built and minimized using SYBYL.

*Corresponding author. Tel.: +1-614-688-3149; fax: +1-614-292-2435; e-mail: johnsamuel.1@osu.edu

0960-894X/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0960-894X(03)00674-7

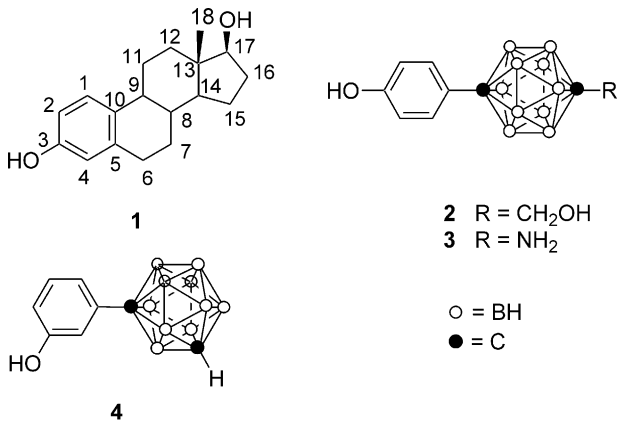

Figure 1. Structures of ligands **1–4**.

The atomic point charges were calculated using the Gasteiger–Hückel method. The molecule was minimized using the Maximin2 minimizer and the TRIPOS force field/parameters until an energy gradient of 0.005 *K*cal/mol was reached.

The carboranyl estradiol derivatives **2–4** (Fig. 1) and *o-*, *m-*, and *p*-carborane were constructed on the HyperChem platform and were minimized by the semi-empirical AM1 method to an energy gradient of 0.005 *K*cal/mol. The atomic point charges of *o-*, *m-*, and *p*-carborane as well as compounds **2–4** were re-calculated using the MOPAC interface of SYBYL applying the semi-empirical AM1 method after import of their PDB files generated by Hyperchem.

Docking of ligands **1–4** into the active site of estrogen receptor was performed using FlexX (TRIPOS Inc., 1699 South Hanley Road, St. Louis, MO 63144, USA). Docking produced 30 possible docked conformations for each of the ligands **1–4** and the Cscore™ program of SYBYL scored each conformation. Cscore™ scoring functions include rmsd values,[13] ChemScore,[14] Dock_score,[15] G-score,[16] FlexX_score,[17] and PMF_score.[18] Among the 30 conformational solutions of ligands 1–4, the ones with the best FlexX_score (rank 1) were chosen as the optimal conformational poses[13] in all docking experiment. The rank 1 conformations showed better binding interactions compared to other solutions. SYBYL was used to generate dynamic hydrogen bonds between the best-docked conformational pose of a ligand and the amino acid residues in the active site of the protein. The same software package was used to visualize the binding mode of the docked protein–ligand complexes by generating a Connolly-type MOLCAD surfaces for estradiol with a probe sphere diameter of 1.4 Å based on the X-ray structures from the PDB file, 1ERE. The MOLCAD surface of estradiol was superimposed on the best-docked conformations of ligands 1–4 to visualize the binding of 1–4 within the active sites.

Results and Discussion

HyperChem, but not Insight II or SYBYL, provided a convenient way for generating and minimizing *o-*, *m-*, and *p*-carborane as well as compounds **2–4** using the Allow Ions option and geometry optimization with the semi-empirical AM1 method. The obtained geometries of *o-*, *m-*, and *p*-carborane correlated sufficiently with those obtained from ab initio calculations at the Hartree–Fock/6-31G* level[19] and experimental electron diffraction data.[20]

Unfortunately, HyperChem does not provide any platform for docking operations. Thus, modeling of **2–4** and their docking into the active site of estrogen receptor had to be carried out using SYBYL and FlexX, respectively. PDB files of compounds **2–4**, generated by HyperChem, were imported into SYBYL, partially reconstructed, and saved as 'mol2' files. All boron atoms 'B' in ligands **2–4** were subsequently changed to carbon atoms 'C.3' using the Built/Edit option in SYBYL. The geometries, bond lengths, bond angles, dihedral angles, and atomic point charges of the structures, obtained from semi-empirical AM1 calculations based on carboranyl structures, did not change during this operation. After docking, boron atoms were re-colored to green for easier identification (Fig. 2). This simple but effective modification circumvents the fact that both SYBYL and FlexX do not contain the empirical potential energy functions for the boron atom.[21] We did not notice any limitation of this modification on docking processes and we are not aware of any limitation on extending this methodology to other modeling studies.

The estrogen receptor protein (hERαLBD) contains six ligand binding domains.[12] Each domain contains one receptor site and the amino acid residues, sequences, and active sites are conserved in all domains. In order to simplify the docking process, five of the six domains of the estrogen receptor protein[22–24] were truncated using SYBYL and the resulting monomeric ligand binding domain was used for docking of ligands **1–4**. The minimized structure of estradiol (**1**) was docked first with the estrogen receptor site and the structure of the resulting protein-ligand complex was compared with that of the X-ray crystallograpically determined structure of the same complex. The X-ray structure showed hydrogen-bonding interactions of the hydroxyl groups at positions **3** and **17** of estradiol with the amino group of the arginine residue (R394) and the imino group of the histidine residue (H524), respectively. The optimal conformational pose corresponded to the highest FlexX_score value (rank 1) and the lowest rmsd value (0.80 Å). The docked estradiol–estrogen receptor protein complex reproduced native binding interactions as shown in Figure 2A. A rmsd value of ≤ 2 Å for the docked ligand–protein structure also suggests that estradiol was appropriately docked within the active site of the estrogen receptor.

FlexX docking of ligands **2–4** was carried out using the Run One Ligand option of FlexX. The structures of optimal conformational poses of ligands **2–4** (Table 1) within the active site of the estrogen receptor and the superimposed MOLCAD surface of the X-ray geometry of estradiol were visually examined for binding interactions.

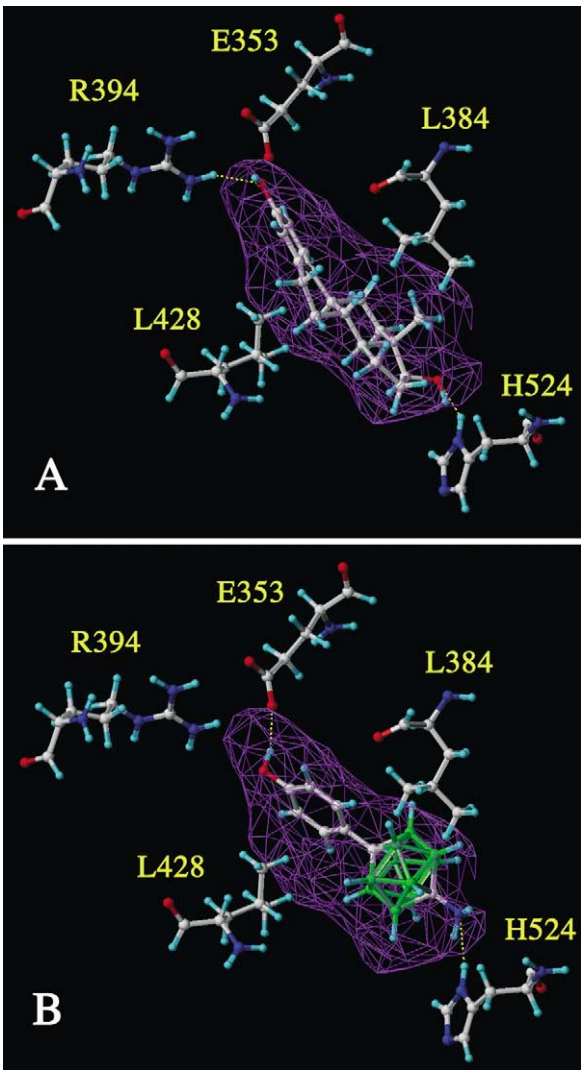

Figure 2. Binding modes of the optimal conformational poses of ligands, 1 (A) and 3 (B), with the active site of the estrogen receptor. The yellow line represents hydrogen bond ($\leq$3.3 Å). The pink grid represents the MOLCAD surface of X-ray geometry of estradiol.

All poses of the ligands **2–4** fitted within the MOLCAD surface area of the X-ray geometry of estradiol as demonstrated by the example of compounds **1** and **3** (Fig. 2A and B). The ligands showed hydrogen bondings between either the alcoholic group (**2**) or carboranyl amino group (**3**), and the imino group of the histidine residue (H524) while the phenolic hydroxyl group in these structures interacted with either the amino group of the arginine residue (R394) (**2–3**) or the carboxylic function of the glutamate residue (E353) (**4**). The bulky carborane clusters of ligands **2–4** were positioned in the active site similar to the decahydronaphthalene portion of estradiol and showed hydrophobic interactions with the isobutyl group of the leucine residues (L384 and L428) as illustrated in Figure 2B.

Various scoring values of ligands **2–4** are listed in Table 1 by comparison with the estrogenic activity of these compounds[6] in order to evaluate the effectiveness of the Cscore^{TM} scoring functions for the estrogen receptor protein–ligand complexes. Among the Cscore^{TM} scoring functions, FlexX_score[17] (based on empirical functions), PMF_score[18] (based on statistical ligand–receptor atom-pair interaction potentials), D_score[15] (based on both electrostatic and hydrophobic contributions to the binding energy), and Chemscore[14] (based on a diverse training set of 82 receptor–ligand complexes) produced values for ligands **1–4** that showed some correlation with their respective biological activities. Ligand **4**, with the lowest biological activity of all carboranyl estradiol derivatives, had also the lowest scoring values (Table 1). For G_score,[16] which computes accurately scoring values for ligand–receptor complexes having many polar interactions, no obvious correlation between the scoring values for ligands **1–4** and respective biological activities could be observed. A possible explanation could be the presence of the carborane cluster in ligands **2–4**, which interacts strongly with the receptor in a hydrophobic fashion. Similar sub-optimal correlation of values from Cscore^{TM} scoring functions with biological activities has been reported previously in different experimental/computational settings.[13,25–27]

In conclusion, the combination of HyperChem and SYBYL generated geometries of the *o-*, *m-*, and *p*-carborane clusters resembled closely those obtained previously using different theoretical and experimental methods.[19,20,28] FlexX docked ligands 2–4 into the active site of the estrogen receptor within the MOLCAD surface of estradiol and all ligands reproduced native estradiol binding interactions. Similar results were obtained using the software package ADAM for docking compounds **2–4** into the active site of the estrogen receptor. Computed values from Cscore^{TM} scoring functions, except G_score, showed some correlation with the biological activities.

Based on our studies, it appears that computed scoring functions from Cscore,^{TM} visual comparison between

Table 1. Cscore^{TM} scoring values for complexes of the estrogen receptor with optimal conformational poses of ligands **1–4**

Ligand	FlexX	G_score	PMF_score	D_score	Chemscore	Hydrogen bonds	Biological activity[a]	Interaction energy (*K*cal)[b]
1	−19	−229	−46	−117	−40	E353, R394, H524	xxx	−54.67
2	−19	−164	−42	−208	−39	R394, H524	xxxx	−51.60
3	−18	−177	−36	−175	−36	E353, G521, H524	xxxx	—
4	−14	−199	−27	−153	−35	E353	x	−45.34

[a]Biological activity is defined as the capacity of compounds **1–4** to induce transcriptional activation of COS-1 cells and was estimated from results obtained by Endo et al.[3] xxxx = very high, xxx = high, xx = medium, x = low.
[b]Interaction energies between estrogen receptor protein and ligands **1–4**, generated by ADAM, were taken from ref 6.

geometries of docked ligands and MOLCAD surfaces, and comparison between the binding interactions of docked ligands with those within ligand–receptor protein X-ray structures can be valuable tools for CAMD of carboranyl derivatives.

The lack of CAMD approaches involving carboranes is a major draw back in BNCT compound development. The described strategy for modeling molecules containing carborane clusters with the accessible software packages SYBYL, FlexX, and HyperChem should be of value for synthetic chemists involved in BNCT compound development.

Acknowledgements

This work was funded by US Department of Energy grant DE-FG02-90ER60972. The authors thank Professor Emeritus Albert H. Soloway, College of Pharmacy, The Ohio State University, Professor Christopher M. Hadad, Department of Chemistry, The Ohio State University, and John C. Hackett, graduate student, College of Pharmacy, The Ohio State University, for fruitful comments and discussions on CAMD.

References and Notes

1. Soloway, A. H.; Tjarks, W.; Barnum, B. A.; Rong, F.-G.; Barth, R. F.; Codogni, I. M.; Wilson, J. G. *Chem. Rev.* **1998**, *98*, 2389.
2. Valliant, J. F.; Guenther, K. J.; King, A. S.; Morel, P.; Schaffer, P.; Sogbein, O. O.; Stephenson, K. A. *Coordin. Chem. Rev.* **2002**, *232*, 173.
3. Marcorin, G. L.; Da Ros, T.; Castellano, S.; Stefancich, G.; Bonin, I.; Miertus, S.; Prato, M. *Org. Lett.* **2000**, *2*, 3955.
4. Friedman, S. H.; Ganapathi, P. S.; Rubin, Y.; Kenyon, G. L. *J. Med. Chem.* **1998**, *41*, 2424.
5. Endo, Y.; Iijima, T.; Yaguchi, K.; Kawachi, E.; Inoue, N.; Kagechika, H.; Kubo, A.; Itai, A. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1307.
6. Endo, Y.; Iijima, T.; Yamakoshi, Y.; Fukasawa, H.; Miyaura, C.; Inada, M.; Kubo, A.; Itai, A. *Chem. Biol.* **2001**, *8*, 341.
7. Endo, Y.; Yaguchi, K.; Tsuji, M.; Yamaguchi, K.; Shudo, K. *Chem. Pharm. Bull.* **1999**, *47*, 699.
8. Endo, Y.; Yoshimi, T.; Kimura, K.; Itai, A. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2561.
9. Tsuji, M.; Koiso, Y.; Takahashi, H.; Hashimoto, Y.; Endo, Y. *Biol. Pharm. Bull.* **2000**, *23*, 513.
10. Tao, P.; Lai, L. *J. Comput. Aid. Mol. Des.* **2001**, *15*, 429.
11. Mizutani, M. Y.; Tomioka, N.; Itai, A. *J. Mol. Biol.* **1994**, *243*, 310.
12. Brzozowski, A. M.; Pike, A. C.; Dauter, Z.; Hubbard, R. E.; Bonn, T.; Engstrom, O.; Ohman, L.; Greene, G. L.; Gustafsson, J. A.; Carlquist, M. *Nature* **1997**, *389*, 753.
13. Gohlke, H.; Hendlich, M.; Klebe, G. *J. Mol. Biol.* **2000**, *295*, 337.
14. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. *Comput. Aided Mol. Des.* **1997**, *11*, 425.
15. Ewing, T. J. A.; Kuntz, I. D. *J. Comput. Chem.* **1997**, *18*, 1175.
16. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727.
17. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, *261*, 470.
18. Muegge, I.; Martin, Y. C. *J. Med. Chem.* **1999**, *42*, 791.
19. Hermansson, K.; Wojcik, M.; Sjoeberg, S. *Inorg. Chem.* **1999**, *38*, 6039.
20. Hnyk, D.; Rankin, D. W. H.; Robertson, H. E.; Hofmann, M.; Schleyer, P. V.; Buhl, M. *Inorg. Chem.* **1994**, *33*, 4781.
21. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
22. Tedesco, R.; Thomas, J. A.; Katzenellenbogen, B. S.; Katzenellenbogen, J. A. *Chem. Biol.* **2001**, *8*, 277.
23. Sippl, W. *Bioorg. Med. Chem.* **2002**, *10*, 3741.
24. Mortensen, D. S.; Rodriguez, A. L.; Carlson, K. E.; Sun, J.; Katzenellenbogen, B. S.; Katzenellenbogen, J. A. *J. Med. Chem.* **2001**, *44*, 3838.
25. Rognan, D.; Bissantz, C.; Dedier, S.; Logean, A.; Reinelt, S. *Chimia* **2000**, *54*, 658.
26. Bissantz, C.; Folkers, G.; Rognan, D. *J. Med. Chem.* **2000**, *43*, 4759.
27. Clark, R. D.; Strizhev, A.; Leonard, J. M.; Blake, J. F.; Matthew, J. B. *J. Mol. Graph. Model.* **2002**, *20*, 281.
28. Davidson, M. G.; Hibbert, T. G.; Howard, J. A. K.; Mackinnon, A.; Wade, K. *J. Chem. Soc., Chem. Commun.* **1996**, *19*, 2285.