



Efficient Molecular Docking of Drug Molecules into DNA Targets and their Enrichment by Cutting-Edge Technologies

S. A. Kumar, M. Petersen, M. G. B. Drew, M. M.-L. Grage,
C. J. Cardin, Y. Gan

published in

*From Computational Biophysics to Systems Biology (CBSB08),
Proceedings of the NIC Workshop 2008,*
Ulrich H. E. Hansmann, Jan H. Meinke, Sandipan Mohanty,
Walter Nadler, Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **40**, ISBN 978-3-9810843-6-8, pp. 281-284, 2008.

© 2008 by John von Neumann Institute for Computing
Permission to make digital or hard copies of portions of this work for
personal or classroom use is granted provided that the copies are not
made or distributed for profit or commercial advantage and that copies
bear this notice and the full citation on the first page. To copy otherwise
requires prior specific permission by the publisher mentioned above.

<http://www.fz-juelich.de/nic-series/volume40>

Efficient Molecular Docking of Drug Molecules into DNA Targets and their Enrichment by Cutting-Edge Technologies

S. A. Kumar^{1,2}, M. Petersen², M. G. B. Drew¹, M. M.-L. Grage²,
C. J. Cardin¹, and Y. Gan¹

¹ School of Chemistry, University of Reading, Whiteknights, Reading RG6 6AD, Berkshire, UK
E-mail: scr04sa@rdg.ac.uk

² Nucleic Acids Center, Department of Physics and Chemistry,
University of Southern Denmark, DK-5230 Odense C, Denmark

A novel DNA mediated anti-cancer mechanism is aimed at stabilizing DNA G-Quadruplex structures. G-tetrads display high polymorphism and are found in the telomeric ends of chromosomes. In normal cells they exist as G-tetrads, producing knots that prevent the access of chromosomal DNA to the replication machinery. During cell division, these G-tetrad knots are resolved and the G-rich DNA is exposed to the replication machinery as single strands. Small molecules stabilizing G-Quadruplex DNA structures are thought of as potential novel anti-cancer agents. Several novel DACA analogues have been designed, synthesized and crystallized in reaction with the G-Quadruplex DNA. Regrettably they failed to diffract even synchrotron light. In the present investigation these analogues were studied for their binding interactions and stabilisation potential of a G-Quadruplex DNA (TGGGGT)⁴ using molecular mechanics and combined QM/MM molecular docking techniques.

1 Introduction

The premise for the present study is that stable drug bound DACA analogue complexes with minimal bound free energy show potential to evolve as anti-cancer agents. A variety of docking algorithms have been tested for docking a number of DACA analogues in the G-Quadruplex DNA. The resulting binding poses have been compared to the crystal structure of a similar drug molecule, daunomycin¹ (PDB ID:1O0K). It turned out that the results of spherical polar coordinate shape complementarity based HEX² docking, and fragment growth based Glide-XP³ (using the OPLS 2005 force field) resulted in realistic binding poses. The results were enriched by describing the polarization of the charge field in the drug - DNA interface by QM/MM single point calculations. The QM method was density functional theory (DFT) B3LYP and calculations were carried out with both 3-21G and 6-31G* basis sets through an iterative Quantum Polarized Docking Workflow⁴. In the following the results of the different molecular docking techniques will be discussed with an emphasis on the differences between force field based and combined QM/MM based predictions.

2 Motivation

Molecular docking is a means of computationally investigating ligand binding to a receptor in order to reduce the laboratory work, as well as justify chemical/physical explanations

for observed phenomena. Docking programs employ an empirical energy function to determine and optimize the interaction energy between the drug candidate and the active site. Structure-based docking methods automatically sample ligand conformations and protein-ligand interactions with a specified region of the protein surface. MM docking results are reported as the lowest energy/highest scoring pose(s) for each ligand, typically one pose for a large compound collection. Hybrid force field electronic structure docking algorithms divide the energy function into a quantum mechanical part describing the ligand, a force field based part describing (most of the) receptor and an interaction part taking care of the embedding of the two descriptions in the area where they interact. In the current case, we see the application of molecular docking as a trial and error process and seek to validate the predictions based on comparison with similar structures solved crystallographically or by NMR studies. Rigid docking turned out not to be successful in producing the drug binding poses observed in the crystal structure. This served as a motivation to examine higher accuracy methods.

3 Materials and Methods

Different DACA analogues were sketched and minimized in gas phase using the UFF force field to prepare an ensemble of starting structures of drug molecules with no atomic clashes in their geometries. Several different docking algorithms were used at varying levels of accuracy, i.e. pure molecular mechanics and combined QM/MM methods. Of the algorithms tested, the HEX and the Schrodinger GLIDE methods produced good binding predictions which agreed with the binding clues learnt from a similar drug bound crystal structure. All docking experiments were conducted by blind docking, i.e. without specifying the binding site of the drug molecule in the receptor.

Glide XP is a force field based fragment docking method. QM/MM docking is achieved in combination with the IMPACT, Q-SITE and JAGUAR modules of Schrodinger suite, using primarily the OPLS 2005 force field. The Glide XP scoring function is designed to improve the results compared to the SP scoring function. This is achieved by adding terms to the function, the additional terms include: i) Coulomb energy of interacting atoms, ii) Van der Waal's energy of atoms, iii) Terms to favour binding interactions, iv) Terms to hinder binding interactions.

The terms that favour binding interactions include terms with parameters to evaluate the hydrophobic enclosures, hydrogen bonding possibilities between neutral-neutral H-bonding motifs, hydrogen bonding possibilities between charged-charged motifs, pi stacking interactions and pi-cation interactions. The terms that hinder binding interactions include terms with parameters to score desolvation in the binding site, and a term to calculate the inter-molecular strain energy based on proximity distances of ligand heavy atoms. On the other hand, the Hex docking algorithm is purely a structure complementarity matching algorithm based on polar spherical coordinates. The molecular surface is represented by several radial expansions of spherical harmonic basis functions. Surfaces are generated for both receptor and ligand molecules and the shape complementarity is scored to arrive at binding predictions.

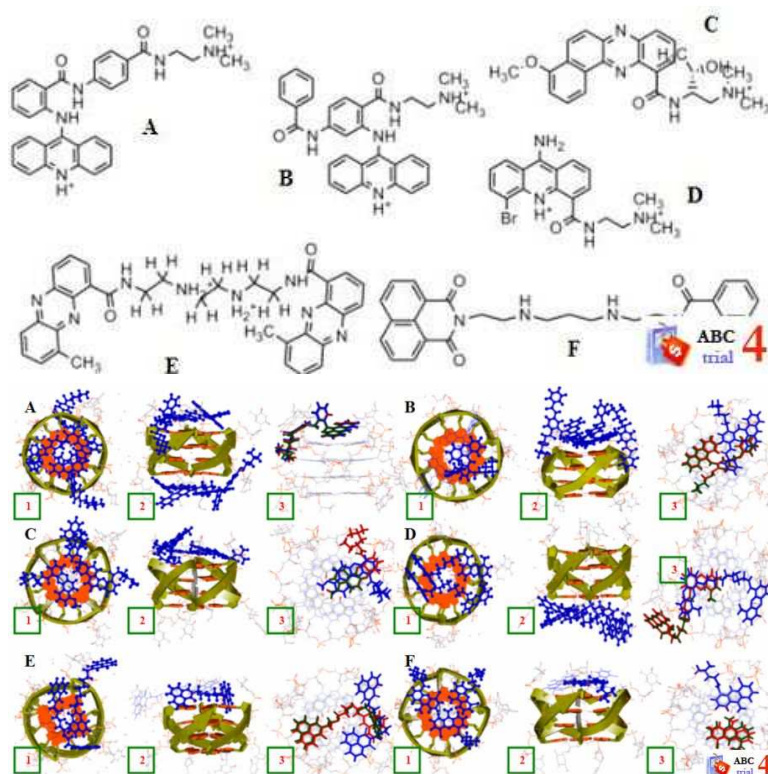


Figure 1. A = Drug 1, B = Drug 2, C = Drug 3, D = Drug 4, E = Drug 5, F = Drug 6. 1 & 2 = Hex docking poses, 3 = Glide docking poses. Colour Codes: Blue = Glide XP OPLS 2005, Red = QM/MM B3LYP with 3-21G basis sets, Green = QM/MM with 6-31G* basis sets.

4 Results and Discussions

There is an overall agreement between the results of Hex and Glide XP docking predictions *fig.1*, although minor variations exist in i.e. side chain conformations etc. In general Glide XP placed the conformations better than the Hex poses which might be due to the bias of the pure shape complementarity principle of Hex. This result in the deviations of chemical geometry which are implied better in various force fields.

The important binding features deduced from the crystal structure are:

1) Planar aromatic rings formed stacking interactions with the DNA bases, 2) Flexible side chains were favored in the electrostatically stable groove regions of the DNA, 3) The G-Quadruplex terminal surface was enough only to accommodate two molecules of daunomycin in fully stacked manner.

Daunomycin is made up of a planar four-ring chromophore and a sugar moiety, while the DACA chromophore is a planar three-ring system with a more flexible side chain and no sugar moiety. As the DACA analogues contain aliphatic side chains in contrast to the heterocyclic sugar moiety in daunomycin, a more compact packing of the sidechains in the

DNA groove regions is expected.

DACA analogues; drug1, drug3 and drug4 are more simple structures containing comparatively shorter length aliphatic side chains connected to the aromatic ring system and thus the Hex and Glide predictions turned out to agree well. In the case of drug2, an aromatic ring branches out from the planar pharmacophore, which is further bifurcated with another aromatic ring on one side and an aliphatic side chain on the other side. Thus drug2 is quite complex in nature, something which might require knowledge of the physical/chemical parameters for a better prediction. The Hex binding pose exhibits an electrostatically less favorable state where the bifurcated groups are dumped in the same region. This is in opposition to the Glide output in which the bifurcation maintains the placement of the aromatic and the aliphatic side chain bulks on opposite sides. In the current case, it can also be suggested that a combination of geometrical and force field methods can lead to a robust prediction mechanism.

DACA analogues; drug5 and drug6 are structures which can typically be considered as tougher molecules for docking. The complexity is due to their long, highly flexible aliphatic linkers. It evinces that approximations in force field parameters and/or negligence of force field parameters lead to highly deviating and less favorable binding predictions when compared to predictions made by combined QM/MM models. The inclusion of QM derived charge polarization clearly has exhibited a better performance that agrees with hints taken from the crystal structure. The Hex predictions place the linkers passing through the terminal surface of G-Quadruplex with the chromophores kept away from stacking interactions. Glide XP heavily under-performed the drug5 binding pose with many fewer interactions between the drug and DNA, which is rectified by redocking with included QM polarized charges.

In conclusion the charge polarization plays a very important role in drug - receptor recognition and binding and can be included by combining traditional MM and QM methods. Furthermore our experiments show that simple algorithms considering mainly the shape complementarity perform nearly as well as the force field parameters based molecular docking methods. In overall, molecular docking can be better achieved by combining several different methods which utilizes a combination of chemical/physical information, shape complementarity and induced effects in the drug-receptor interface.

Acknowledgments

We are very thankful to all our academic and industrial collaborators. Special thanks to Nucleic Acids Center at SDU, MC FP6 Project: Nucleic Acid Based Drug Design and the Danish Center for Scientific Computing.

References

1. Clark, G. R., Pytel, P. D., Squire, C. J., Neidle, S., *J. Am. Chem. Soc.*, **125**, 4066–4067, 2003.
2. Ritchie, D. W., Kemp, G. J. L., *J. Comp. Chem.* **20**(4), 383–395, 1999.
3. GLIDE, version 4.5, (Schrodinger, LLC: Portland, OR, 2000).
4. Cho, A. E., Gullar, V., Berne, B. J., Friesner, R., *J. Comp. Chem.* **26**, 915–931, 2005.