Digging Deeper into Deep Data: Molecular Docking as a Hypothesis-driven Biophysical Interrogation System in Computational Toxicology

Michael Goldsmith, Stephen Little, David Reif, James Rabinowitz • U.S. EPA National Center for Computational Toxicology (NCCT)

MOLECULAR MODELS

Background

• Developing and evaluating predictive strategies to elucidate the mode of biological activity of environmental chemicals is a major objective of the concerted efforts of the US-EPA's computational toxicology program. Aligning these strategies with the Agency's ongoing chemically-specific risk-assessment needs will provide additional molecular-level insight for decision-making purposes.

• Often, data required for extrapolations inherent in human risk assessment are unavailable. In silico methods can be used to provide molecular-level information surrogates that are vital for toxicological mechanistic insight.

• Employing a virtual screening approach, a diverse set of chemicals were computationally docked into multiple macromolecular targets (nuclear receptors) using an extensive docking algorithm. The individual chemical-target poses, scores, and the chemical-protein contexts generated by this approach afforded a virtual affinity fingerprint matrix that provides mechanistic molecular-level insight. Knowledge gained from quantitative and visual analyses (clustering/heat maps, and linkage networks) of these virtual screens demonstrate the utility of these approaches and their ability to resolve differences in ligand panagonism, receptor promiscuity.

• These virtual affinity fingerprint matrices, coupled to tissue-specific receptor distribution data and inference mapping of downstream signal transduction elements, provide a molecular level of accountability that complements experimental high-throughput screening and toxicogenomic endeavors. [This work was reviewed by EPA and approved for publication but does not necessarily reflect official Agency policy.]

Method: Molecular Docking

Docking of both EZ-Guggletaronic geometric isomers against multiple crystal structure derived human NR targets in their agonist-associated (active) conformation (from www.pdb.org) and MSMPR’s optimized ligand set geometries with AM1-BCC charges assigned from MOE (CDG Canada) as found in KiBank (Aizawa 2004), curated from the original publication on guggletaronic polyporpharmacy (Burnis, 2005). In the computational toxicology framework we may also pose this question in terms of polytoxicology (or pan-agonism associated with an adverse rather than therapeutic effect). All performed in eHiTS on ‘fast’ screening mode (fewer match-poses generated) (Zatzakis et al 2006) against the diverse set of targets. The docked structure of EZ isomers are shown docked within the binding pocket from NR (mineralocorticoid receptor) one of the top hits for both isomers. The structural formulae is shown overlaid on the experiment/theory rank ordered bar graphs (magnitude – normalized binding affinity (Kd) to largest value, so large bars = high affinity).

Validation: Target Fishing for Promiscuous Ligands

Results: Docking environmental chemicals into 18 Nuclear Receptors

(A) Docking of 409 diverse compounds from the ToxCast proof-of-concept chemical list against multiple NRs. All docking calculations performed in eHiTS on ‘fast’ screening mode on a dual processor Athlon Opteron 64 bit server in order 24 hours (fewer match-poses generated). Heat map/hierarchical cluster performed in R, with two key groups identified (green-box) = the weakest binders and (red-box) the strong promiscuous NR binders. The actual structures are shown in (C) and (D).

(B) Docking pose slide show of several tight binding ligands to their respective target.

(C) Weaker NR binding cluster of chemicals from in silico screening in A (i)

(D) Strong promiscuous NR binding cluster of chemicals from in silico screening in A (ii)

(E) Chemical (functional group) feature histograms generated in LeadScope for the (left) weak and (right) strong clusters. Adjacent to this feature histogram are the AlacGP profiles of both sets, illustrating major differences in these clusters.

(F) The higher affinity (i.e. top 50) chemicals from (A) with logP < 4 (132 chemicals/noises/index and 285 linkages) were subsequently plotted as a linkage map in Cytoscape. These edges represent the binding of specific ligands to specific targets based on molecular docking (a biosophysical computational model of protein ligand interactions).

Highlights & Future Directions

• The guggletaronic screen qualitatively and quantitatively agree with experiment, a use of tool for screening or ‘fishing’ for putative targets.

• Docking studies demonstrate the practicality of an approach that identifies and clusters both (a) chemicals in a given target-space, and (b) targets in a given chemical space; this experiment has identified compounds that show (i) high NR promiscuity and affinity as well as (ii) high NR specificity with varying degrees of affinity. The top structure in this class is strikingly similar (structurally analogous) to Aizuthin, a known environmental obesogen. Target-space clustering in the context of these ligands suggests weaker binders are smaller than tighter promiscuous binders, have a higher heterosol (D.N.S) count (light have higher halogen count) and have greater degrees of freedom.

• Will consider additional targets (see schema below) such as human serum albumin (shown below in 3D) and lipid binding proteins required to translocate chemicals from the cytoplasm to the nucleus.

Method: Molecular Docking

Validation: Target Fishing for Promiscuous Ligands

Docking of both EZ-Guggletaronic geometric isomers against multiple crystal structure derived human NR targets in their agonist-associated (active) conformation (from www.pdb.org) and MSMPR’s optimized ligand set geometries with AM1-BCC charges assigned from MOE (CDG Canada) as found in KiBank (Aizawa 2004), curated from the original publication on guggletaronic polyporpharmacy (Burnis, 2005). In the computational toxicology framework we may also pose this question in terms of polytoxicology (or pan-agonism associated with an adverse rather than therapeutic effect). All performed in eHiTS on ‘fast’ screening mode (fewer match-poses generated) (Zatzakis et al 2006) against the diverse set of targets. The docked structure of EZ isomers are shown docked within the binding pocket from NR (mineralocorticoid receptor) one of the top hits for both isomers. The structural formulae is shown overlaid on the experiment/theory rank ordered bar graphs (magnitude – normalized binding affinity (Kd) to largest value, so large bars = high affinity).

(F) The higher affinity (i.e. top 50) chemicals from (A) with logP < 4 (132 chemicals/noises/index and 285 linkages) were subsequently plotted as a linkage map in Cytoscape. These edges represent the binding of specific ligands to specific targets based on molecular docking (a biosophysical computational model of protein ligand interactions).

References

• Will also perform analogous screen on rodent targets (mouse and rat) for which sufficient in vivo and in vivo data exists, although this may require homology modeling sparse target sets (most crystallized forms of targets shown in this study are protein sequences from humans expressed in a secondary system).
• More efficient identification and enumeration of biologically/environmentally relevant permutations and progeny of the chemical structures in question would be highly desirable. These include stereoisomers, tautomers, protonation states, metabolites and degradation products)

• vHTS studies of parent compounds provide valuable molecular-level detail in the toxicant-target paradigm. These details, along with additional experimental information, may be used for hypothesis generation and are complementary to hypothesis-driven toxicogenomic inquiry.

Acknowledgements

This work was reviewed by EPA and approved for publication but does not necessarily reflect official Agency policy.™

MRG was Supported in part by a HHS/NIH/DESI Cooperative Training in Environmental Research, EPA 5R8323601. Special thanks to Dr. T. Transeau and Dr. M. Wolf (both Lockheed Martin) for assistance.

International Science Forum on Computational Toxicology
May 21–23, 2007 • Research Triangle Park, North Carolina