



# Virtual Screening for Endocrine Disrupting Compounds

Melissa A. Pasquinelli, Stephen Little, Susan Laws\*, and James Rabinowitz

U.S. EPA, ORD, Computational Toxicology Research Program

research & development

IIC-2

## Science Question

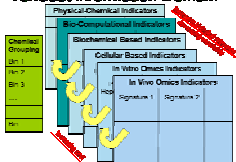
Which chemicals are most likely to bind to steroid hormone receptors and disrupt the endocrine system?

## Research Goals

Utilize quantum chemistry and other molecular modeling methods in order to develop computational predictive tools that can be used to screen for endocrine disrupting effects.

Interface the virtual screening tools with complementary experimental and computational approaches such as the ToxCast initiative in order to aid the Agency in prioritizing data requirements for risk assessment.

## ToxCast Information Domain



ToxCast is a multi-level information domain scheme for studying chemical toxicity, an NCCT initiative led by Robert Kavlock, David Dix, and Keith Houck.

First phase: Perform a comprehensive study on a set of diverse chemicals across all information domains.

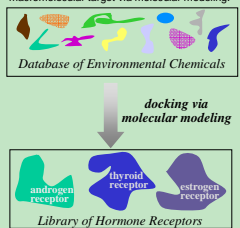
Second phase: The information that is gathered will be used to extract relationships among the information domains.

Third phase: Utilize these relationships to inform computer models in order to improve their abilities to suggest priorities for future bioassays and other regulatory purposes.

## Methods

### What is Virtual Screening?

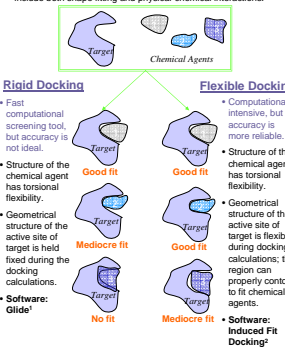
Each chemical agent in the database is "docked" into the binding region of each macromolecular target via molecular modeling.



1. Devise a set of compounds that will provide a rigorous test of the model, in collaboration with ORD scientists in NHEERL. Build a database of these chemicals that are properly prepared for the virtual screening process (such as the addition of hydrogen atoms, proper definition of bond orders, stereochemistry, etc.).
2. Develop a diverse library of nuclear receptor targets, focusing first on human steroid hormone receptors and then including other environmentally relevant nuclear receptor targets that have experimental information available of sufficient quality.
3. Utilize the virtual screening tool to screen the database of chemical agents with the library of nuclear receptors. Optimize the virtual screening tool in order to minimize the false negative rate.
4. Once optimized, use the virtual screening tool to predict the endocrine disrupting potential of unknown chemicals. Each chemical-target pair predicted to interact by the virtual screening tool will undergo advanced computational and/or experimental testing.

### What is Docking?

The "best fit" for each chemical agent is calculated in the binding region of the macromolecular target. The best fit docking poses include both shape fitting and physical-chemical interactions.



#### Rigid Docking

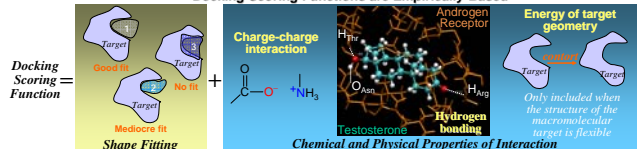
- Fast computational screening tool, but accuracy is not ideal.
- Structure of the chemical agent has torsional flexibility.
- Geometrical structure of the active site of target is held fixed during the docking calculations.
- Software: Glide<sup>1</sup>

#### Flexible Docking

- Computationally intensive, but accuracy is more reliable.
- Structure of the chemical agent has torsional flexibility.
- Geometrical structure of the active site of target is flexible during docking calculations; the region can properly contort to fit chemical agents.
- Software: Induced Fit Docking<sup>2</sup>

5. Dissect the information obtained from the virtual screening for particular characteristics of endocrine disrupting chemicals, such as across receptor classes or species.
  - Integrate the knowledge with complementary experimental and computational approaches, such as the ToxCast initiative.
  - This phase of the project could benefit from collaborations with researchers from the STAR Bioinformatics Research Centers.

## Docking Scoring Functions are Empirically-Based



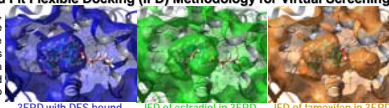
## Chemical and Physical Properties of Interaction

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

## Results

### Test of Induced Fit Flexible Docking (IFD) Methodology for Virtual Screening

As an initial test of the IFD methodology<sup>2</sup>, three different PDB<sup>3</sup> crystal structures of the human estrogen receptor alpha (hERα) were used to cross-dock their crystallized ligands with IFD: estradiol (1GWR<sup>4</sup>), tamoxifen (3ERT)<sup>5</sup>, and diethylstilbestrol (3ERD)<sup>6</sup>. Rigid docking predicts no binding of tamoxifen to the 3ERD structure, or a false negative.



The pictures to the left show how the binding pocket for 3ERD contorts in order to fit the larger ligand molecules, estradiol (green) and tamoxifen (orange).

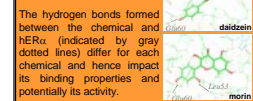
### Using Induced Fit Docking Methodology for Virtual Screening of Endocrine Disrupting Compounds

In collaboration with Susan Laws and other ORD scientists in NHEERL<sup>4</sup>, the 3ERD structure for hERα was then used to screen a set of chemicals that represent a wide range of biological activity with ER as determined from experimental measurements. These chemicals were also screened with the human Androgen Receptor (hAR)<sup>7</sup> and hERβ<sup>8</sup>. The results indicate a need for a better docking scoring function that can differentiate strong, weak, and no-binders.

Chemical Agent	Ki (nM) ERα	ERα 3ERD	ERβ 1QKM	AR 1T65
[R] EM652	na	-59.2	-46.8	-63.6
[S] EM652	na	-61.3	-52.3	-60.6
diethylstilbestrol	0.04	-39.7	-42.1	-46.7
hydroxytamoxifen	0.1	-59.3	-56.9	-59.6
17β-estradiol	0.1	-49.7	-50.1	-50.4
mesa-hexestrol	0.2	-41.0	-42.3	-46.0
RS-hexestrol	0.2	-43.8	-48.2	-45.2
SS-hexestrol	0.2	-45.1	-47.8	-48.8
coumestrol	42	-42.9	-46.1	-49.1
5,11-diethyl-5,6,11,12-tetrahydrochrysen-2,8-diol	na	-42.5	-47.0	-51.6
2-ethylhexyl 4-hydroxybenzoate	480	-36.8	-39.3	-42.1
4-nonylphenol	636	-35.2	-35.8	-38.4
diadzein	6000	-43.4	-45.6	-45.7
4,4',4''-ethane-1,1,1-triyltriphenol	8250	-46.9	-46.6	-48.6
4-(1-methyl-1-phenylethyl)phenol	16480	-34.2	-34.8	-35.9
4,4'-propane-2,2'-diylidiphenol	24800	-36.0	-40.4	-40.0
4-octylphenol	39520	-32.4	-34.1	-34.6
decan-1-ol	74310	-25.4	-22.0	-28.4
6,6'-dithiodi(2-naphthol)	80630	-51.7	-52.1	-52.9
4,4'-sulfonyldiphenol	82450	-37.8	-39.1	-43.1
morin	100000	-44.8	-54.8	-53.0
phenyl(2,3,4-trihydroxyphenyl)methanone	102000	-40.6	-42.7	-45.0
2,2-dimethoxy-1,2-diphenylethaneone	136000	-38.7	-41.1	-41.0
ethyl-4-hydroxybenzoate	330000	-31.6	-31.6	-31.5
(3Z)-hex-3-en-1-yl salicylate	No bind	-36.6	-35.5	-36.8
(4-methylphenyl)(phenyl)methanone	No bind	-33.3	-34.6	-32.9
3,4-dimethoxybenzaldehyde	No bind	-30.9	-29.5	-28.5
4,4'-propane-2,2'-diylbis(2-tert-butylphenol)	No bind	-43.1	-44.3	-49.2
4,4'-thiodibenzene-1,3-diol	No bind	-38.7	-39.4	-45.0
5-chloro-2-hydroxy-N-phenylbenzamide	No bind	-41.3	-40.5	-40.8
diphenyl isophthalate	No bind	-50.8	-53.4	-50.3

weak estrogenic binders (from experiments)

The pictures to the right show calculated poses of hERα with 17β-estradiol, its natural ligand, and diadzein and morin, environmental compounds that are weakly estrogenic.



The hydrogen bonds formed between the chemical and hERα (indicated by gray dotted lines) differ for each chemical and hence impact its binding properties and potentially its activity.

## Conclusions

- Including protein flexibility during the simulation of target-toxicant interactions is significant.
- A docking "scoring function" is needed that differentiates between weak binding chemicals and non-binders in order to minimize the false negative rate.

## Future Directions

- Repeat virtual screening with the same set of chemicals using a protein that is not a nuclear receptor, such as a signaling protein, and other nuclear receptors.
- Add more chemicals to the test set, including chemicals with known androgen and other endocrine activity. Collaboration with scientists from ORD/NERL and ORD/NHEERL.
- Since flexible docking is not computationally feasible for high-throughput screening, develop a database of target poses for each receptor that can be used with rigid docking methods.
- Since scoring functions are currently optimized for pharmaceutical discovery, a better scoring function needs to be developed that can differentiate among strong, weak, and no binding compounds. One potential idea is to develop a knowledge-based scoring function that includes the experimental data. This phase of the project could interface with the STAR Bioinformatics Research Centers.

## Impact and Outcomes

This research supports the Agency's goals in the multi-year plans for Human Health and Endocrine Disruption. It addresses the significant Agency need for predictive models for hazard identification in the sub-areas of (1) QSAR and other computational approaches, and (2) high throughput screening.

This project supports two of the objectives of the NCCT as stated in the Computational Toxicology Framework: (2) develop predictive models for categorizing and prioritizing chemicals in the environment, and (3) improve quantitative risk assessments. Molecular modeling and other tools derived from computational chemistry can be used in conjunction with available experimental data to develop methods for screening environmental chemicals for toxicological effects on human health and ecosystems. This virtual screening tool will lead to a more efficient risk assessment process.

In collaboration with Dr. Laws and other scientists in ORD/NHEERL, a scheme for prioritizing bioassay requirements for identifying potentially endocrine disrupting chemicals is being developed and applied. This scheme provides a rational strategy for the application of bioassay efforts to determine endocrine disrupting chemicals. In collaboration with the ToxCast program of ORD/NCCT, complex computational molecular modeling approaches will become an integral part of chemical categorization and prioritization for more diverse toxicological effects.

## References

1. Glide version 4.0. Schrodinger, LLC, New York, NY, 2005.
2. Schrodinger Suite 2006 Induced Fit Docking protocol: Glide version 4.0. Schrodinger, LLC, New York, NY, 2005.
3. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. *Nucleic Acids Research* 28, 235-242 (2000).
4. A. Warmark, E. Troutler, J.A. Gustafson, R.E. Hubbard, A.M. Brzozowski, A.C.W. Pike. *J. Biol. Chem.* 277 21862 (2002).
5. A.K. Shiau, D. Barnard, P.M. Loria, L. Cheng, P.J. Kushner, D.A. Agard, C.L. Greene. *Cell* 95 921-937 (1998).
6. S.C. Laws, S. Yavanbayk, R.L. Cooper, J.C. Eldridge. Submitted to *Toxicol. Sci.* (2006).
7. PDB entry 1T65. E. Estévez-Perrina, J.M.R. Moore, E. Mai, E. Delgado-Rodriguez, P. Nguyen, J.D. Baxter, B.M. Buxner, P. Webb, R.J. Fletterick, R.K. Guy. *J. Biol. Chem.* 280 8066-8068 (2005).
8. PDB entry 1QKM. A.C.W. Pike, A.M. Brzozowski, R.E. Hubbard, J. Bonn, A.G. Thorsell, O. Engstrom, J. Ljunggren, J.-A. Gustafson, M. Carlquist. *Embo J.* 18 101 (1999).

\* Reproductive Toxicology Division, National Health and Environmental Effects Research Laboratory, Office of Research and Development, U.S. EPA

Long Term Goal II

COMPUTATIONAL TOXICOLOGY

This poster does not necessarily reflect EPA policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.