Supporting Information: Modeling Reactivity to Biological Macromolecules with a Deep Multitask Network

Tyler B. Hughes,*,† Na Le Dang,*,† Grover P. Miller,*,‡ and S. Joshua Swamidass*,†

†*Department of Pathology and Immunology, Washington University School of Medicine, Campus Box 8118, 660 S. Euclid Ave., St. Louis, Missouri 63110, United States*

‡*Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Sciences, Little Rock, Arkansas 72205, United States*

Contents

Data

AMD Training Data

The reactions used in the training data set are included in a comma-separated values text file: *TrainingAMDRegistryNumbers.csv*. This file includes the reaction registry and molecule registry numbers from the December 2014 Accelrys Metabolite Database (AMD) release. Four binary indicator columns designate molecules labeled reactive or nonreactive to cyanide, DNA, glutathione (GSH), and protein. Molecules with unknown reactivity have missing values for that target(s).

Labeling of Site of Reactivity Training Data

Each reaction in the AMD contains the structures of both the starting molecule and the product. The product structures enabled sites of reactivity to be labeled on the starting molecules. For example, the structure of the cyanide conjugate enabled the site of reactivity to be labeled on a nefazodone metabolite (Figure 1). First, the maximum common substructure was calculated between the cyanide conjugate and the nefazodone metabolite. Second, the connectivity distance matrix was calculated for the nefazodone metabolite. Third, for the cyanide conjugate, the atom within the maximum common substructure closest (by connectivity) to the cyanide group was identified. Fourth, this atom was mapped back to the respective atom within the maximum common substructure in the nefazodone metabolite. Similar algorithms were developed for labeling DNA, GSH, and protein sites of reactivity.

Unreactive Epoxide Training Data

While epoxides are generally quite reactive, they can be stabilized in certain cases, such as by the presence of electron-donating groups on epoxide carbons.[1] A shortcoming of our previous method for predicting GSH reactivity was that it predicted all epoxides reactive.[2] To address this, we supplemented our training data with 63 naturally occurring[3] and known nonreactive[1] epoxides. We hypothesized that epoxides found in nature are less likely to be reactive, for if they were very unstable, then they would likely not be observed. The SMILES strings for all 63 epoxides are included in the text file *TrainingNegativeEpoxides.smi*.

AMD External Test Data

After assembling the training data, a new version of the AMD was released with several new reactions, which was tapped for possible new molecules as a test data set. This yielded 14 new molecules that reacted with GSH. The reactions used in this external test set are included in a comma-separated values text file: *TestAMDRegistryNumbers.csv*. This file includes the reaction registry and molecule registry numbers from the June 2015 Accelrys Metabolite Database (AMD) release.

To quantify how similar the external test set was to the training set, we first calculated path based fingerprints (with depth 8) for each of the external test set and training set molecules. Second, we calculated the MinMax similarity (a variant of Tanimoto suitable for

count fingerprints) of each fingerprint of the external test set and each fingerprint of the training set.[4] Third, for each molecule in the external test set, we recorded its maximum similarity score to any molecule in the training set. Across all 14 external test set molecules, this procedure yielded an average maximum similarity of 0.396 ± 0.235. A similar procedure to measure the internal diversity of the training set was also performed, yielding an average maximum similarity of 0.701 ± 0.224. This result implies that the test set is somewhat dissimilar from the training set, which is encouraging because the model successfully generalized to these new molecules (Figure 2 and Figure S7).

Descriptors

The reactivity model used 15 molecule-level (Table S1) and 194 atom-level (Table S2) topological descriptors, each of which describes a chemical property of an atom or the molecule containing an atom. This descriptor set was based on a previous set of topological descriptors, with several expansions.[2,5,6] New descriptors included AlphaBetaUnsat, AmideNitrogen, AromaticNeighbors, ATOMS, BN_*b*_*d*, CarboxylOxygen, CorrectedBondRad, CorrectedVdwRad, ElectronAffinity, ElectronNeg, Epoxide, HbondAcceptor, HbondDonor, HYDROGENS, Ionization, LonePair_*d*, Mass, MaxBonds, MichaelAcceptor, MichaelAcceptorSubstituted, NA_*d*_*e* where 1) $d = 4$ and e = C, N, O, P, S, F, Cl, Br, or I, or 2) d = 0–3 and e = F, Cl, Br, or I, *hyb*_*d*_*e*, NitroOxygen, PA_*d*_*e* where 1) $d = 4$ and e = C, N, O, P, S, F, Cl, Br, or I, or 2) d = 0–3 and e = F, Cl, Br, or I, PartialCharges, PhosphateOxygen, RINGS, SulfateOxygen, and TotalBondOrder

Table S1: Molecule-level topological descriptors used by the XenoSite Reactivity Model.

Label	Definition
ATOMS	number of heavy atoms
B_A	number of aromatic bonds
B_D	number of double bonds
B_S	number of single bonds
B_T	number of triple bonds
BONDS	total number of bonds
HBA_1	number of hydrogen bond acceptors Pybel SMARTS string 1
HBA_2	number of hydrogen bond acceptors Pybel SMARTS string 2
HBD	number of hydrogen bond donors
HYDROGENS	number of hydrogens
logP	octanol/water partition coefficient
MR	molar refractivity
MW	molecular weight
RINGS	number of rings
TPSA	topological polar surface area

Table S2: Atom-level topological descriptors used by the XenoSite Reactivity Model.

Label	Definition	Depth/Range	Total Number
AlphaBetaUnsat	binary value indicating whether a neighboring atom has an unsaturated bond to a third atom	0–1	1
AmideNitrogen	binary value indicating whether a nitrogen is part of an amide	0–1	1
Aromatic	binary value indicating whether atom is aromatic	0–1	1
AromaticNeighbors	number of aromatic neighbors for atom	0–3	1
BN_*b*_*d*	number of bonds of type b depth d bonds away	b = single, aromatic, double, triple d = 1–2	8

CarboxylOxygen	binary value indicating whether atom is an oxygen in a carboxyl group	0–1	1
CorrectedBondRad	scales the covalent radius by 0.95 for $sp2^2$ and 0.90 for sp hybrids	0.57–1.39	1
CorrectedVdwRad	scales the van der Waals radius by 0.95 for sp^2 and 0.90 for sp hybrids	1.40–2.05	1
ElectronAffinity	the electron affinity of this element	-0.07–3.61	1
ElectronNeg	the Pauling electronegativity (eV) of this element	2.0–3.98	1
Epoxide	binary value indicating whether atom is in a epoxide	0–1	1
HbondAcceptor	binary value indicating whether atom is a hydrogen-bond acceptor	0–1	1
HbondDonor	binary value indicating whether atom is a hydrogen-bond donor	0–1	1
*hyb*_*d*	total number of atoms of hybridization *hyb* depth *d* bonds away	*hyb* = sp, sp^2, or sp^3 *d* = 0–4	15
*hyb*_*d*_*e*	number of atoms of hybridization *hyb* depth *d* bonds away of type element *e*	*hyb* = sp, sp^2, or sp^3 *d* = 0–4 *e* = C, N, and O or S if *hyb* ≠ sp	50
Hydrogens	number of hydrogens bound to atom	0–3	1
Ionization	the ionization potential (eV) for this element	8.30–17.42	1
LonePair_*d*	number of lone pairs depth *d* bonds away	*d* = 0–4	5
Mass	the average atomic mass for this element	10.81–200.59	1
MaxBonds	the maximum expected number of bonds for this element	1–6	1
MaxInvRingSize	maximum inverted ring size	0–0.33	1
MichaelAcceptor	binary value indicating whether atom is in a Michael acceptor	0–1	1
MichaelAcceptorSubstituted	binary value indicating whether atom is in a substituted Michael acceptor	0–1	1
NA_*d*_*e*	number of atoms depth *d* bonds away of type element type *e*	*d* = 0–4 *e* = C, N, O, P, S, F, Cl, Br, or I	45
NitroOxygen	binary value indicating whether a neighboring atom has an unsaturated bond to a third atom	0–1	1
NRings	number of rings containing atom	0–4	1
PA_*d*_*e*	percentage of atoms depth *d* bonds away of element type *e*	*d* = 1–4 *e* = C, N, O, P, S, F, Cl, Br, or I	36
PartialCharges	Gasteiger partial charge of atom	-0.78–0.58	1
PhosphateOxygen	binary value indicating whether atom id an oxygen in a phosphate group	0–1	1
$Ring_n$	within ring of size *n*	*n* = 3–8	6
Rotors	number of rotatable bonds for atom	0–4	1
Span	(maximum path length from current atom) / (maximum path length from all atoms within the molecule)	0–16	1
SpanInverted	1/(1 + Span)	0.058–1	1
SpanNormalized	Span/(maximum span within molecule)	0–1	1
SulfateOxygen	binary value indicating whether atom is an oxygen in a sulfate group	0–1	1
TotalBondOrder	sum of all bond orders for bonds connected to atom	1–6	1

Results

Several experiments were performed during model construction, optimization, and analysis, the results of which are here presented. First, we used an expanded set of topological molecule-level (Table S1) and 194 atom-level (Table S2) descriptors, which significantly improved over previously used topological descriptors (Figure S1). Second, we jointly modeled several types of reactivity in a multitask learning format, which we hypothesized would improve predictions on the smaller data sets through transfer learning.[7,8] Indeed, the multitask model outperformed the individual modeling approach, especially at predicting protein sites of reactivity (Figure S2). Third, a modular input layer (Figure 5) was used to group related

descriptors, rather than a traditional three-layer neural network structure. The modular structure enabled reduction of the total number of weights in the model by 50%, while retaining the same performance (Figure S3). This weight-reduced, modularly-structured model outperformed a traditionally-structured model with the same number of weights. Fourth, we tested whether supplementing the topological descriptors with quantum chemical descriptors improved performance. We found that a topology-only model matched the performance of a model that included both topological and quantum chemical descriptors (Figure S4), so we did not include the quantum chemical descriptors in the final model. Fifth, to analyze the inner workings of our final model, we performed a permutation sensitivity analysis, a method we have previously used to analyze similar networks (Figure S5).[2,6,9] Sixth, we found that including nonreactive epoxides in training improved the model's ability to distinguish between reactive and nonreactive epoxides (Figure S6).

Performances of Previous Topological Descriptors and New Topological Descriptors

From the structure of each molecule, an in-house python script calculated 15 molecule-level (Table S1) and 194 atom-level topological descriptors (Table S2), each of which describes a chemical property. These descriptors were used as input to the reactivity model, which learned a mapping between the descriptors and sites of reactivity or molecule reactivity. The majority of these topological descriptors have been shown to be useful for the XenoSite metabolism,[5] GSH reactivity,[2] and epoxidation models.[6] This study used an expanded set of topological descriptors, which in 10-fold cross-validated experiments improved performance over the previous set of descriptors, for cyanide, DNA, GSH, and protein (Figure S1). Two metrics were used to assess performance. First, the "average site AUC" was computed by calculating the area under the ROC curve (AUC) for each molecule and averaging the AUCs for each molecule in the data set.[2,6] Second, the "top-two metric" was computed, which is standard for assessing sites of metabolism predictions.[2,5,10,11] This metric considers a molecule as correctly predicted if any of its sites of reactivity are predicted in the first or second rank positions.

Individual vs Multitarget Training

A single model was built to collectively predict reactivity to cyanide, DNA, GSH, and protein, instead of modeling each type of reactivity individually. We hypothesized that by modeling several times of reactivity in concert, we would improve predictions on the smaller data sets through transfer learning. Indeed, the multitask model outperformed the individual modeling approach, especially at predicting protein sites of reactivity (Figure S2).

Effect of Modular Neural Network

A biologically-inspired modular network architecture was used[12–16] (Figure 5). Its modular input layer allowed incorporation of *a priori* knowledge, guiding discovery of meaningful interactions among input descriptors. Each module was a fully connected network between a group of input descriptors and a subset of nodes in the first hidden layer. If there were strong

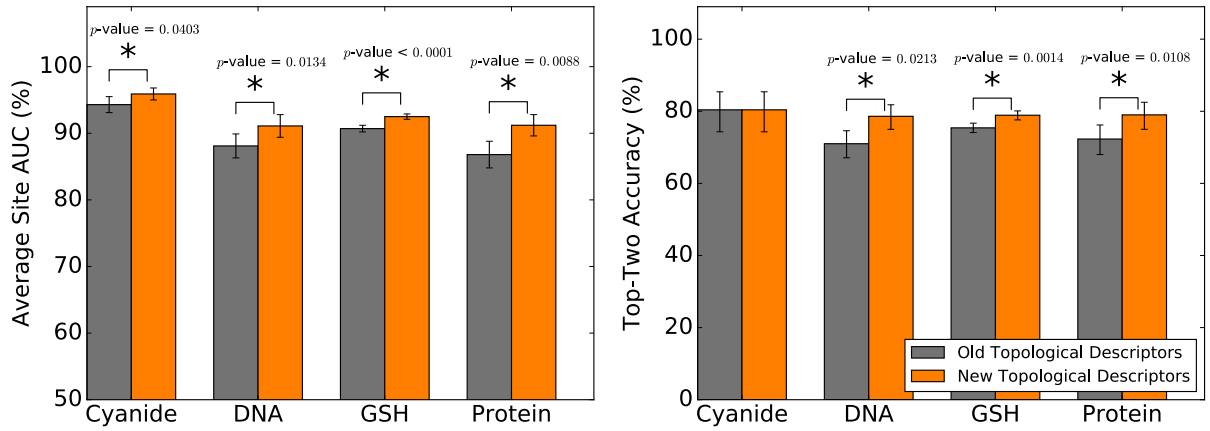

Figure S1: An expanded set of topological descriptors improved over previously used topological descriptors. The reactivity model used 15 molecule-level (Table S1) and 194 atom-level topological descriptors (Table S2), each of which describes a chemical property. An in-house python script calculated these descriptors from the structure of each molecule. The majority of our topological descriptors have been shown to be useful for the XenoSite metabolism, reactivity, and epoxidation models.[2,5,6] To evaluate the value of these new descriptors, we used them as input to a traditionally-structured neural network with one hidden layer, and compared the 10-fold cross-validated performances to those of the old set of topological descriptors inputted to the same network structure. Measured by the average site AUC metric, the models built with the new set of topological descriptors significantly outperformed the old descriptors across all four nucleophiles. By the top-two metric, the performances of the new descriptor set were significantly higher than the performances of the old descriptors for DNA, glutathione (GSH), and protein. The average site AUC metric measures how often reactive atoms were ranked above nonreactive atoms within reactive molecules, and the top-two metric represents the percentage of time a reactive atom is found in the top-two rank positions within a molecule. The models with the new topological descriptors had the same total number of weights as the models with the old topological descriptors, which was achieved by decreasing the number of hidden nodes in the new topological descriptor model to offset the greater number of input descriptors. Results significantly different by paired t-tests are indicated by asterisks.

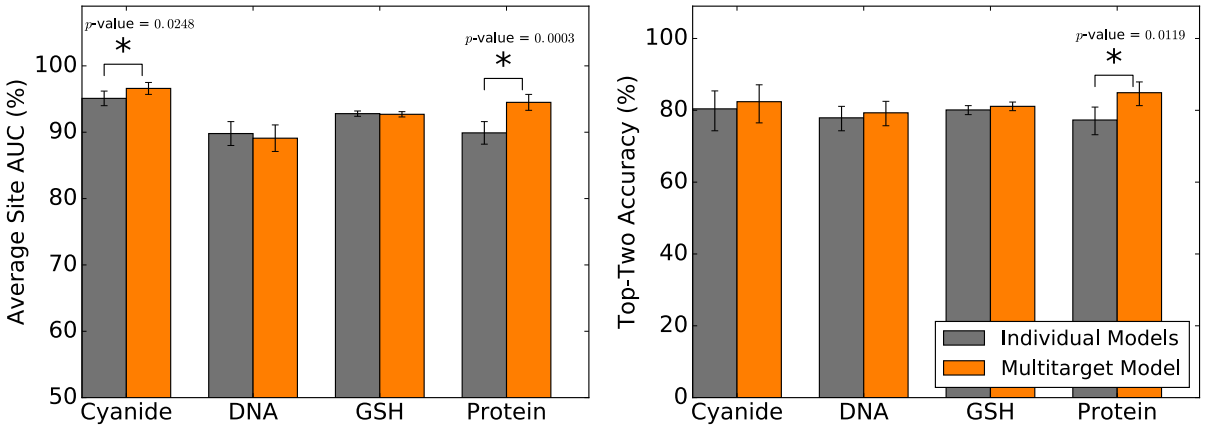

Figure S2: Multitask learning improved over individual models. Jointly modeling several types of reactivity in a multitask learning format improved predictions through transfer learning,[7,8] especially at predicting protein sites of reactivity. The predictive performances of 10-fold cross-validated scores produced by training individual models for each nucleophile are compared to those of a single model that outputs predictions for all four nucleophiles. This multitask model had the same total number of weights as the individual models combined. Results significantly different by paired t-tests are indicated by asterisks.

correlations between the grouped descriptors, the connected nodes in the first hidden layer were able to capture the information in a more parsimonious manner than a traditionally-structured three-layer model. From our previous work, we observed that the contribution of each descriptor to the final model depended not only on what kind of physical, chemical or quantum mechanic characteristic it represented, but also on the location of the atom or bond it depicted.[2,5,6] We hypothesized that the impact of neighboring chemical entities (atoms, bonds, or hybridization systems) could be described by a small number of core features. With this in mind, we combined descriptors by their distances to the atom of interest into identity and hybridization neighborhood groups (Table S3). Compared to a traditionally structured model, the modular input layer enabled reduction of the number of parameters of the model by 50%, while retaining accuracy (Figure S3). This is advantageous, because using the simplest possible model—the one with the fewest total number of weights—improves generalization potential by reducing the chances of overfitting.

Table S3: Descriptor groups in the modular multi-target neural network.

Group	Input Nodes	Number of ^1H nodes
Atom Element	NA_d_e with $d = 0$	3
Atom Hybridization	hyb_d_e with $d = 0$	3
Atoms One Bond Away	NA_d_e and PA_d_e with $d = 1$	3
Atom Hybridization One Bond Away	hyb_d_e with $d = 1$	3
Atoms Two Bonds Away	NA_d_e and PA_d_e with $d = 2$	3
Atom Hybridization Two Bonds Away	hyb_d_e with $d = 2$	3
Atoms Three Bonds Away	NA_d_e and PA_d_e with $d = 3$	3
Atom Hybridization Three Bonds Away	hyb_d_e with $d = 3$	3
Atoms Four Bonds Away	NA_d_e and PA_d_e with $d = 4$	3
Atom Hybridization Four Bonds Away	hyb_d_e with $d = 4$	3

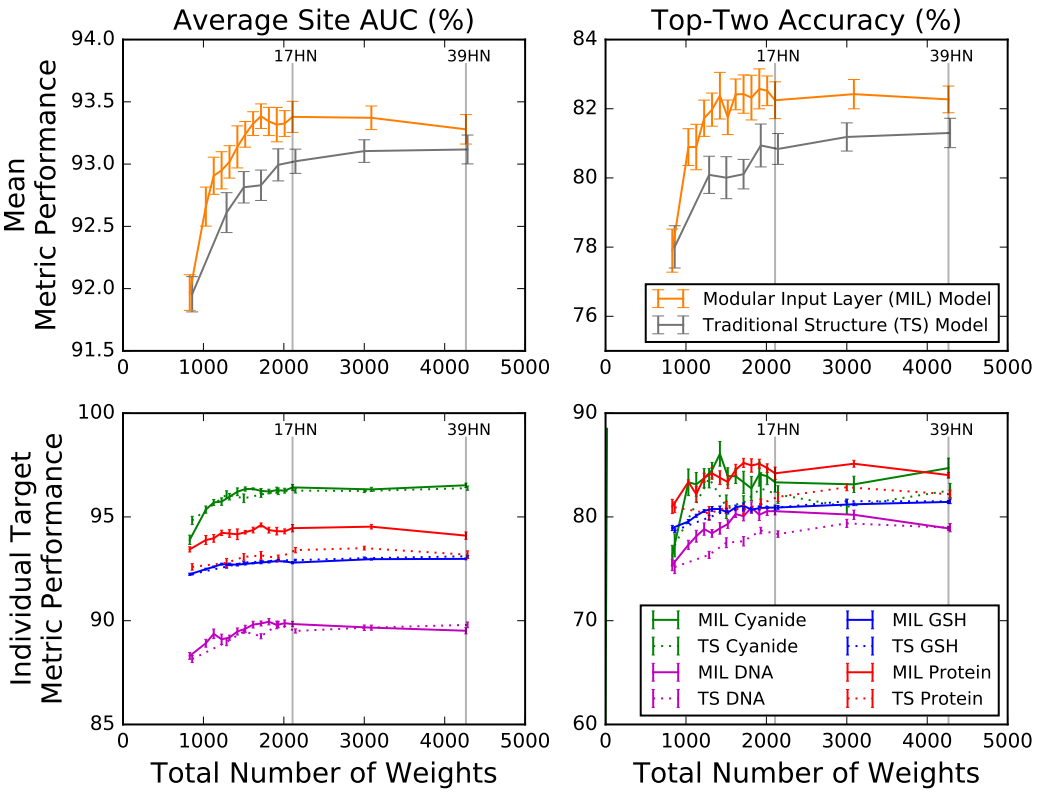

Figure S3: Use of a modular input layer enabled reduction of the total number of weights by 50%**, while retaining accuracy.** By varying the number of hidden nodes in each model, these charts plot model performance as a function of the total number of weights. Top, the cross-validated performances of neural network models built with either a modular input layer (MIL) or a traditional structure (TS) are evaluated by both the average site AUC and top-two metrics. Each data point represents the mean cross-validated performance of all four nucleophilic targets, with 10 iterations of 10-fold cross validation. Originally, the TS model used 20 hidden nodes, which ensured weight-controlled experiments in Figure S2. Due to its different structure, the MIL model required 39 hidden nodes to have the nearest possible total number of weights (4264) to that of the TS model with 20 hidden nodes (4284 total weights). For this number of parameters, the MIL model only slightly outperformed the TS model. However, the MIL model allowed greater reduction of weights than the TS model. For example, a MIL model with 17 hidden nodes matched the performances of the MIL model with 39 hidden nodes, despite only having about half as many weights (2108). In contrast, the TS model performances decreased steadily as the number of weights were reduced. For all future experiments, the MIL model with 17 nodes was selected as the model that optimizes performance versus simplicity. Bottom row, performances are broken down for all four nucleophilic targets individually for both the MIL and TS models. On each graph, the models with 17 hidden nodes (17HN) and 39 hidden nodes (39HN) are indicated by vertical lines.

Effect of Adding Quantum Chemical Descriptors

In our previous reactivity study, we noted that our GSH reactivity model primarily relied upon topological descriptors instead of quantum chemical descriptors.[2] Consequently, this study began by only using topological descriptors, and we later tested whether adding the quantum chemical descriptors from the previous study boosted model performance (Figure S4).[2] The same total number of weights were used in both models, so that any difference in their performances was solely due to the different descriptor sets. As measured by both the average site AUC and top-two metrics, there was no significant difference in performance between the two descriptor sets. Therefore, we concluded that there was no need to include quantum chemical descriptors, and proceeded with a topology-only model. Eliminating quantum chemical descriptors offered several advantages for the final model. First, they are significantly quicker to compute. Second, our model no longer depends on the exact 3D conformation of the molecule. Most commonly, chemical information is disseminated in SMILES format, which only contains 2D information. By eliminating the need to calculate 3D coordinates, using topology-only models removes one layer of extrapolation required to make predictions. Third, the same SMILES string or molecule will always receive the exact same predictions by the topological model, in contrast to the previous GSH model that included quantum chemical descriptors, and therefore could produce slightly different predictions depending on the exact conformation of an input molecule.

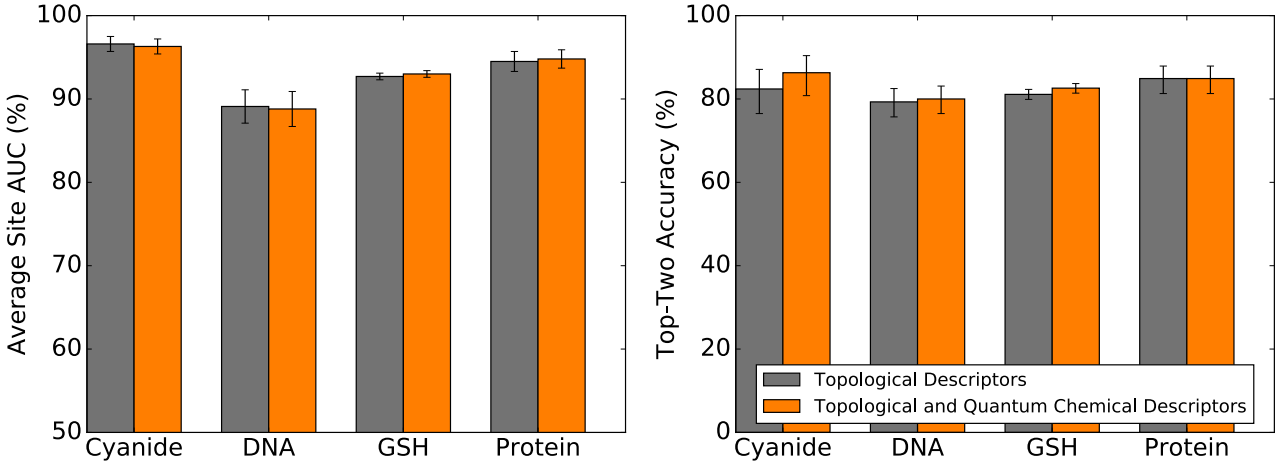

Figure S4: A topology-only model matched the performance of a model supplemented with quantum chemical descriptors. The cross-validated predictions of a model exclusively trained on topological descriptors are compared to those of a model trained with both topological and quantum chemical descriptors. The inclusion of quantum chemical descriptors did not improve performance compared to only using topological descriptors. The same total number of weights were used for both models. No results were significantly different by paired t-tests. For example, the p-values of the cyanide average site AUC and top-two results were 0.58 and 0.32, respectively.

Descriptors Driving Prediction Performance

To investigate the inner workings of the reactivity model, we used a permutation sensitivity analysis.[2,6,9] This procedure identifies the descriptors heavily relied upon by the model. First, using the full training data set, a baseline model is built, and its performance calculated on this same training data. Performance is measured by the average site AUC metric, but a top-two analysis was also performed with similar results (data not shown). Secondly, the importance of groups of descriptors (Table S4), and individual descriptors, was quantified by calculating the fall in the model's performance on training data after shuffling the descriptor values randomly. This procedure was performed 10 times for each descriptor category, and the mean performance drop reported (Figure S5). Permutation sensitivity analysis was performed both for the final topology-only model and the model that included quantum chemical descriptors. The relative importance of descriptors for both models were fairly similar, and none of the quantum chemical descriptors were heavily weighted by the model that included them in training.

Effect of Adding Unreactive Epoxides

Our previous method for predicting GSH reactivity predicted all epoxides reactive.[2] While epoxides are generally quite reactive, they can be stabilized due to specific molecular context. To improve the model's ability to differentially predict epoxide reactivity, we supplemented our AMD training data with 63 naturally occurring[3] and known nonreactive[1] epoxides. We hypothesized that epoxides found in nature are less likely to be reactive, for if they were very unstable, then they would likely not be observed. We included these nonreactive epoxides in the training data, and compared the cross-validated scores on all epoxide carbons to the equivalent predictions of a second model, trained without the nonreactive epoxides scores. Performance was measured by Epoxide AUC, which reflects how often reactive epoxide carbons are ranked above nonreactive epoxide carbons. We found that including nonreactive epoxides in training improved the model's ability to distinguish between reactive and nonreactive epoxides (Figure S6).

Table S4: Descriptor groups used for sensitivity analysis.

Atom Element	NA_d_e with $d = 0$
Atom Hybridization	hyb_d_e with $d = 0$
Atoms One Bond Away	NA_d_e and PA_d_e with $d = 1$
Atom Hybridization One Bond Away	hyb_d_e with $d = 1$
Atoms Two Bonds Away	NA_d_e and PA_d_e with $d = 2$
Atom Hybridization Two Bonds Away	hyb_d_e with $d = 2$
Atoms Three Bonds Away	NA_d_e and PA_d_e with $d = 3$
Atom Hybridization Three Bonds Away	hyb_d_e with $d = 3$
Atoms Four Bonds Away	NA_d_e and PA_d_e with $d = 4$
Atom Hybridization Four Bonds Away	hyb_d_e with $d = 4$
Bonds	B_A, B_D, B_S, B_T, and BONDS
HydrogenBonds	HBA_1, HBA_2, and HBD
Ring Information	MaxInvRingSize, NRings, and $Ring_n$
Span	Span, SpanInverted, and SpanNormalized

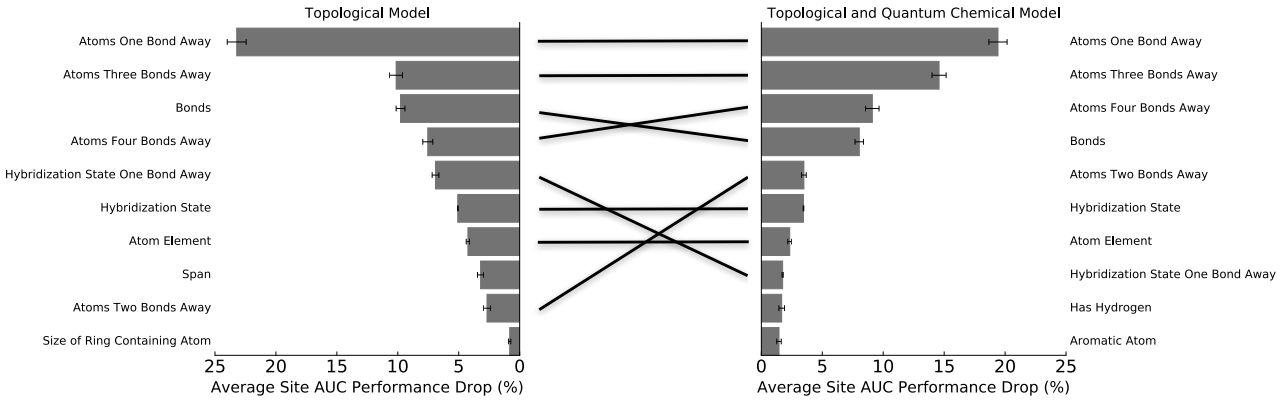

Figure S5: The importance of specific descriptors to the final topology-only model (left) and the topological and a model that included quantum chemical descriptors (right). A permutation sensitivity analysis measured the importance of descriptors for each model.[2,6,9] This procedure identified the descriptors heavily relied upon by the model. First, using the full training data set, a baseline model was built, and its performance calculated on this same training data. Performance was measured by the average site AUC metric. Second, the importance of groups of descriptors (Table S4), and individual descriptors, was quantified by calculating the fall in the model's performance on training data after shuffling the respective descriptor values randomly. Each chart displays the 10 most important descriptor or descriptor groups, decreasing in order of importance from top to bottom. Each data point represents the average model performance drop after 3 iterations of permuting the relevant descriptors.

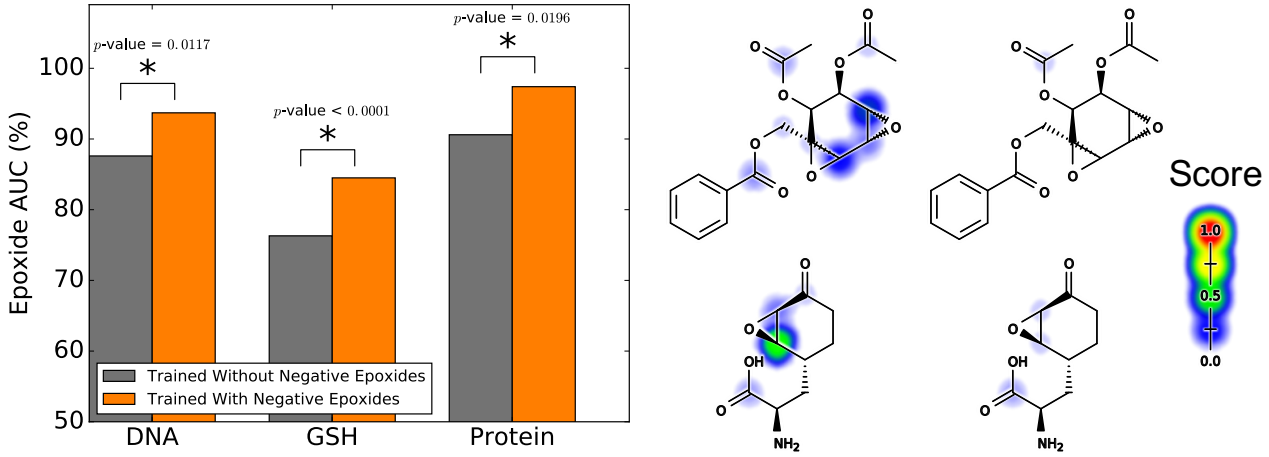

Figure S6: Including negative epoxides in training improves the model's ability to distinguish between reactive and nonreactive epoxides. Left, training without negatives (in grey) and predicting on nonreactive epoxides is compared to training with several negative epoxides (in orange) in the cross-validation procedure. Performance is measured by epoxide AUC, which reflects how often reactive epoxide carbons are ranked above nonreactive epoxide carbons. The improved epoxide AUC for the model trained with the nonreactive epoxides demonstrates that the inclusion of this additional data significantly improved the model's ability to determine whether epoxides are reactive. Right, two example epoxides are visualized with the protein reactivity predictions of the model trained without negative epoxides (left column) and with negative epoxides (right column). Top row: crotepoxide, bottom row: anticapsin.[3] The model trained without examples of nonreactive epoxides failed to recognize that these epoxides are unreactive. Results significantly different by paired t-tests are indicated by asterisks.

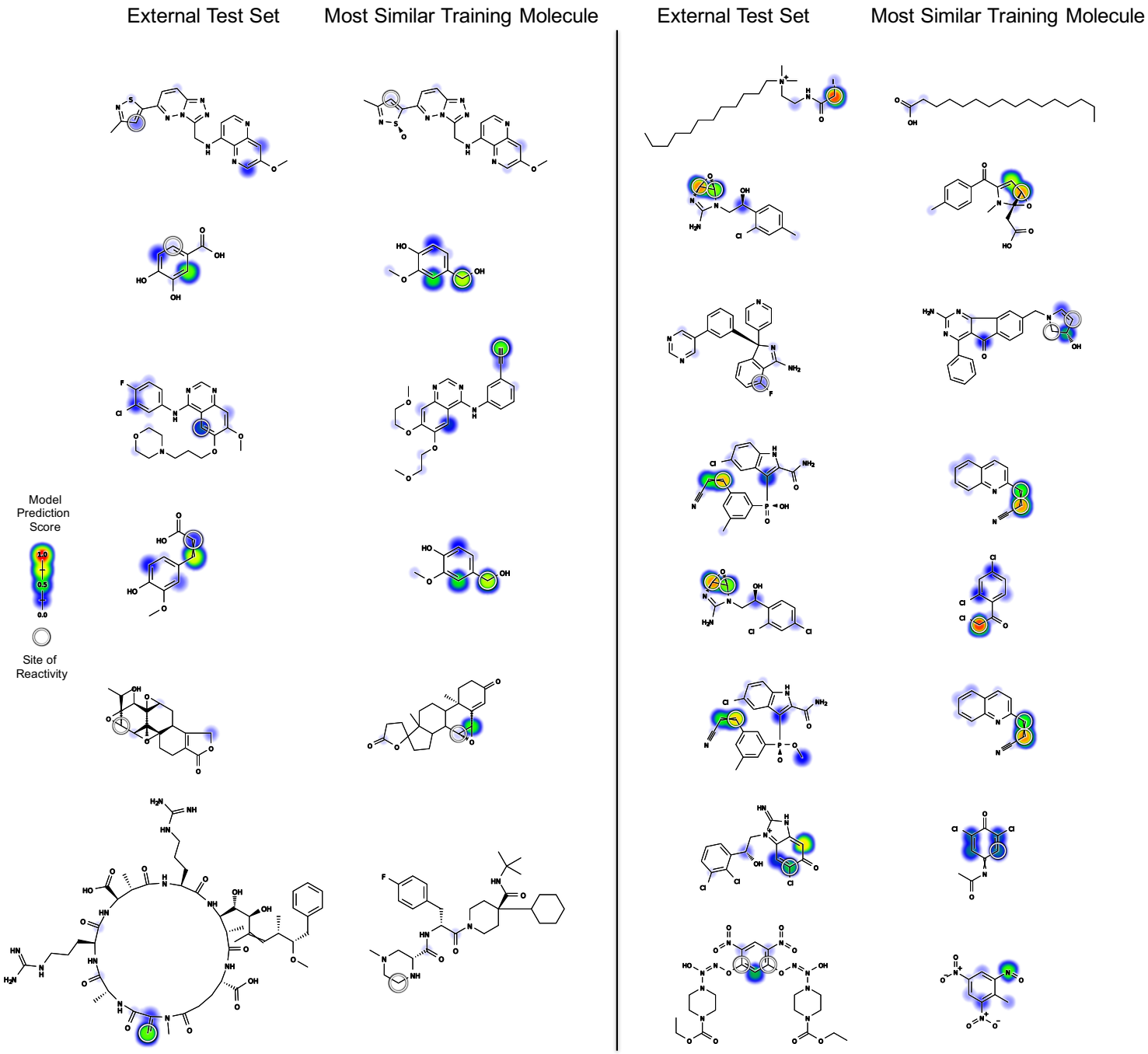

Figure S7: The model generalized to an external test set. Each molecule of the external test set is visualized alongside its most similar neighbor from the training data set. Similarity was measured using the MinMax similarity (a variant of Tanimoto suitable for count fingerprints) on path based fingerprints.[4] The colored shading indicates GSH atom reactivity scores. Experimentally-known sites of reactivity are labeled with white circles.

References

(1) Delaine, T.; Ponting, D. J.; Niklasson, I. B.; Emter, R.; Hagvall, L.; Norrby, P.-O.; Natsch, A.; Luthman, K.; Karlberg, A.-T. Epoxyalcohols: Bioactivation and Conjugation Required for Skin Sensitization. *Chem. Res. Toxicol.* **2014**, *27*, 1860–1870.

(2) Hughes, T. B.; Miller, G. P.; Swamidass, S. J. Site of Reactivity Models Predict Molecular Reactivity of Diverse Chemicals with Glutathione. *Chem. Res. Toxicol.* **2015**, *28*, 797–809.

(3) Marco-Contelles, J.; Molina, M. T.; Anjum, S. Naturally occurring cyclohexane epoxides: sources, biological activities, and synthesis. *Chem. Rev.* **2004**, *104*, 2857–2900.

(4) Ralaivola, L.; Swamidass, S. J.; Saigo, H.; Baldi, P. Graph kernels for chemical informatics. *Neural networks* **2005**, *18*, 1093–1110.

(5) Zaretzki, J.; Matlock, M.; Swamidass, S. J. XenoSite: Accurately predicting CYP-mediated sites of metabolism with neural networks. *J. Chem. Inf. Model.* **2013**, *53*, 3373–3383.

(6) Hughes, T. B.; Miller, G. P.; Swamidass, S. J. Modeling epoxidation of drug-like molecules with a deep machine learning network. *ACS Cent. Sci.* **2015**, *1*, 168–180.

(7) Collobert, R.; Weston, J. A unified architecture for natural language processing: Deep neural networks with multitask learning. Proceedings of the 25th international conference on Machine learning. 2008; pp 160–167.

(8) Pan, S. J.; Yang, Q. A survey on transfer learning. *Knowledge and Data Engineering, IEEE Transactions on* **2010**, *22*, 1345–1359.

(9) Hunter, A.; Kennedy, L.; Henry, J.; Ferguson, I. Application of neural networks and sensitivity analysis to improved prediction of trauma survival. *Comp. Meth. Progr. Biom.* **2000**, *62*, 11–19.

(10) Cruciani, G.; Baroni, M.; Benedetti, P.; Goracci, L.; Fortuna, C. G. Exposition and reactivity optimization to predict sites of metabolism in chemicals. *Drug Discovery Today: Technol.* **2013**, *10*, e155–e165.

(11) Zamora, I.; Afzelius, L.; Cruciani, G. Predicting drug metabolism: a site of metabolism prediction tool applied to the cytochrome P450 2C9. *J. Med. Chem.* **2003**, *46*, 2313–2324.

(12) Azam, F. Biologically inspired modular neural networks. Ph.D. thesis, Citeseer, 2000.

(13) Anand, R.; Mehrotra, K.; Mohan, C. K.; Ranka, S. Efficient classification for multiclass problems using modular neural networks. *Neural Networks, IEEE Transactions on* **1995**, *6*, 117–124.

(14) Lu, B.-L.; Ito, M. *Biological and Artificial Computation: From Neuroscience to Technology*; Springer: Berlin, 1997; pp 330–339.

(15) Auda, G.; Kamel, M. Modular neural network classifiers: A comparative study. *J. Intell. Rob. Syst.* **1998**, *21*, 117–129.

(16) Balakin, K. V.; Ekins, S.; Bugrim, A.; Ivanenkov, Y. A.; Korolev, D.; Nikolsky, Y. V.; Ivashchenko, A. A.; Savchuk, N. P.; Nikolskaya, T. Quantitative structure-metabolism relationship modeling of metabolic N-dealkylation reaction rates. *Drug Metab. Dispos.* **2004**, *32*, 1111–1120.