Predicting CYP Reactivity and Sites of Metabolism in ADMET Predictor®

Pankaj R. Daga

Please note: this presentation, including questions from the audience, is being recorded and may be made available.

Drug Metabolism in Early Drug Discovery

- Drug discovery is time-consuming, expensive and labor-intensive
- Up to 25% of compounds withdrawn due to metabolic, pharmacokinetic, or toxic problems
- Metabolic liability can lead to several diverse issues

Which parts of molecule are subject to metabolic reactions

- Modify portions of new molecule to modulate its metabolism to improve its safety and efficacy
- Crucial for rational drug design to mitigate ADME/toxicity issues

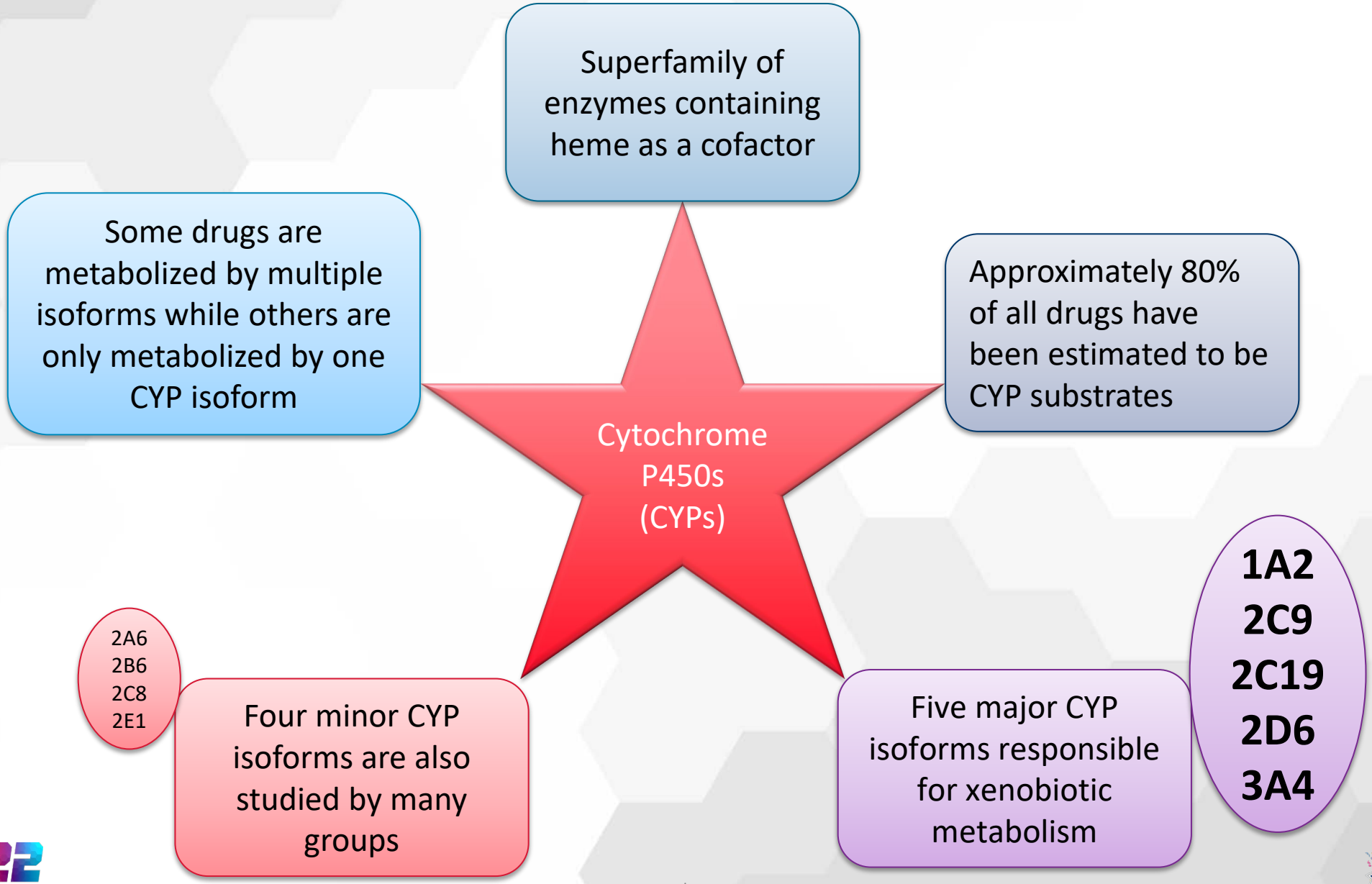
Which enzymes can metabolize a newly designed molecule

- Drug-Drug Interactions predicted and modified to improve drug's applicability

Metabolized into a Toxic/Non-toxic product

- Will have a low therapeutic index
- Need to be administered carefully

Cytochrome P450s (CYPs)



Computational CYP Models

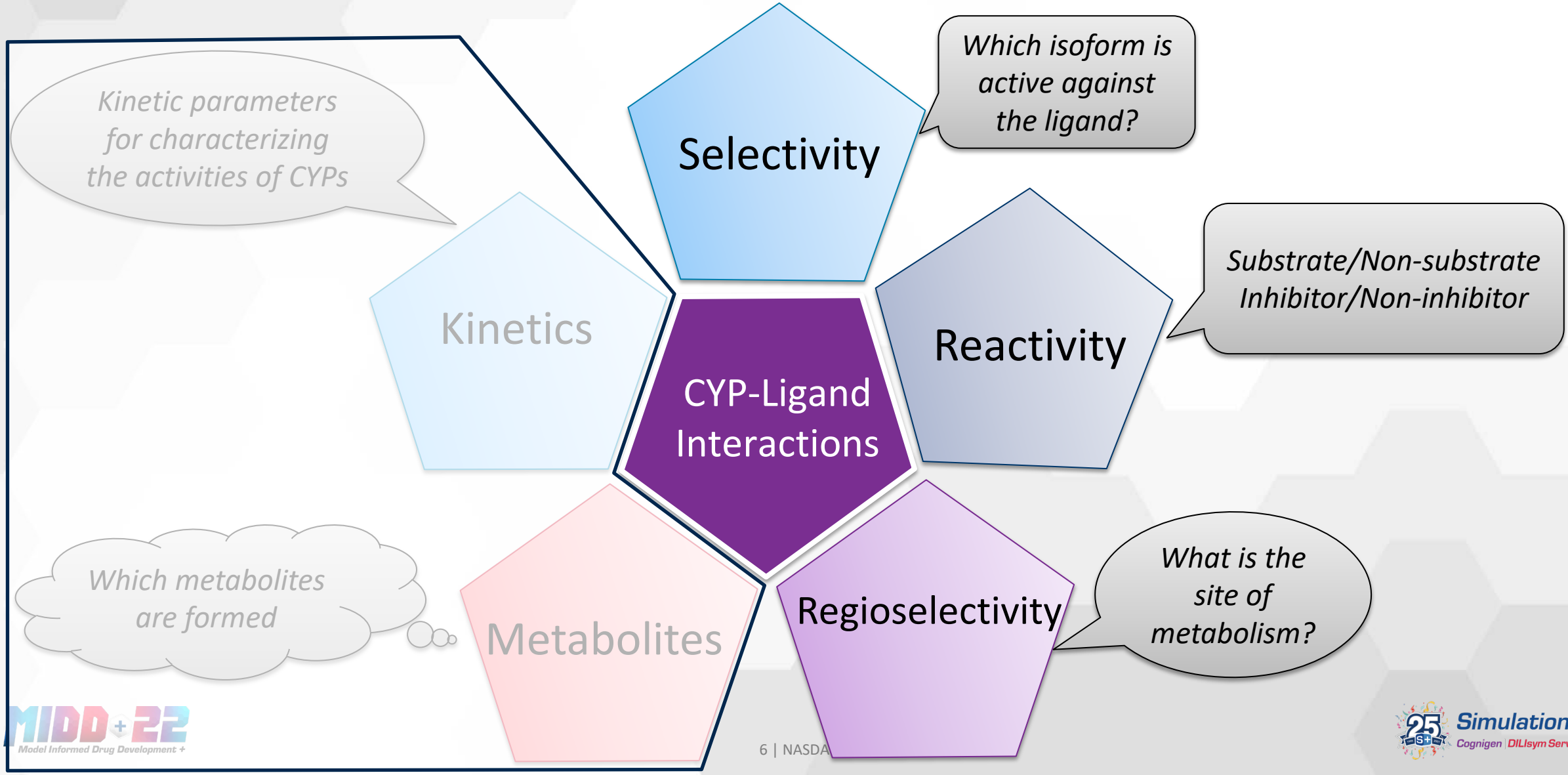
- Experimental identification of SoMs or metabolites
  - expensive and time-consuming task
- *In silico* prediction of SoMs and metabolites
  - Reduce costs and time needed

Rule-based	Structure-based	Ligand-based
• Derive likely metabolites and SoMs by applying a dictionary of biotransformation rules • Compiled by human experts	• Focus on the substrate–protein interaction • Requires high quality crystal structures and very time-consuming	• Employ machine learning and atom-level descriptors • Rely heavily on available experimental data on SoMs and metabolites

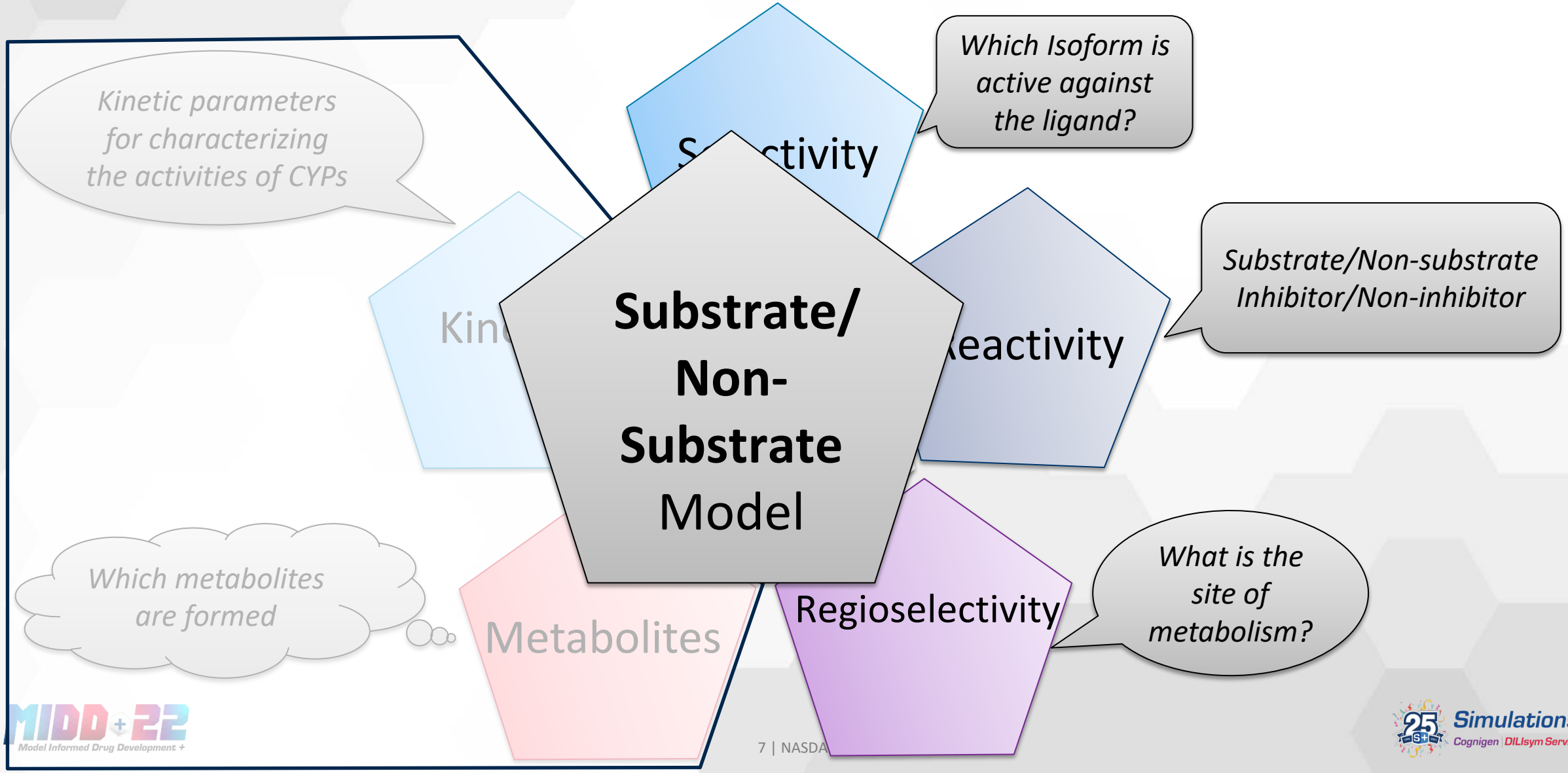
Šícho, *et al, J. Chem. Inf. Model.* **2017**, *57*, 1832−1846

CYP-LIGAND INTERACTIONS IN ADMET PREDICTOR®

What does ADMET Predictor® offer?

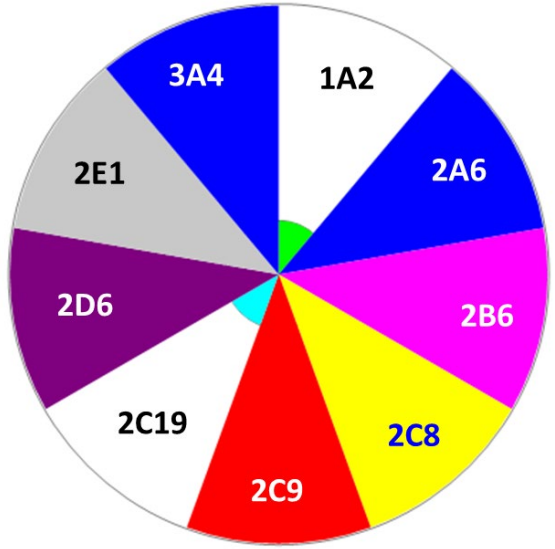


Selectivity : Reactivity Combined



CYP Substrate Models: The Data

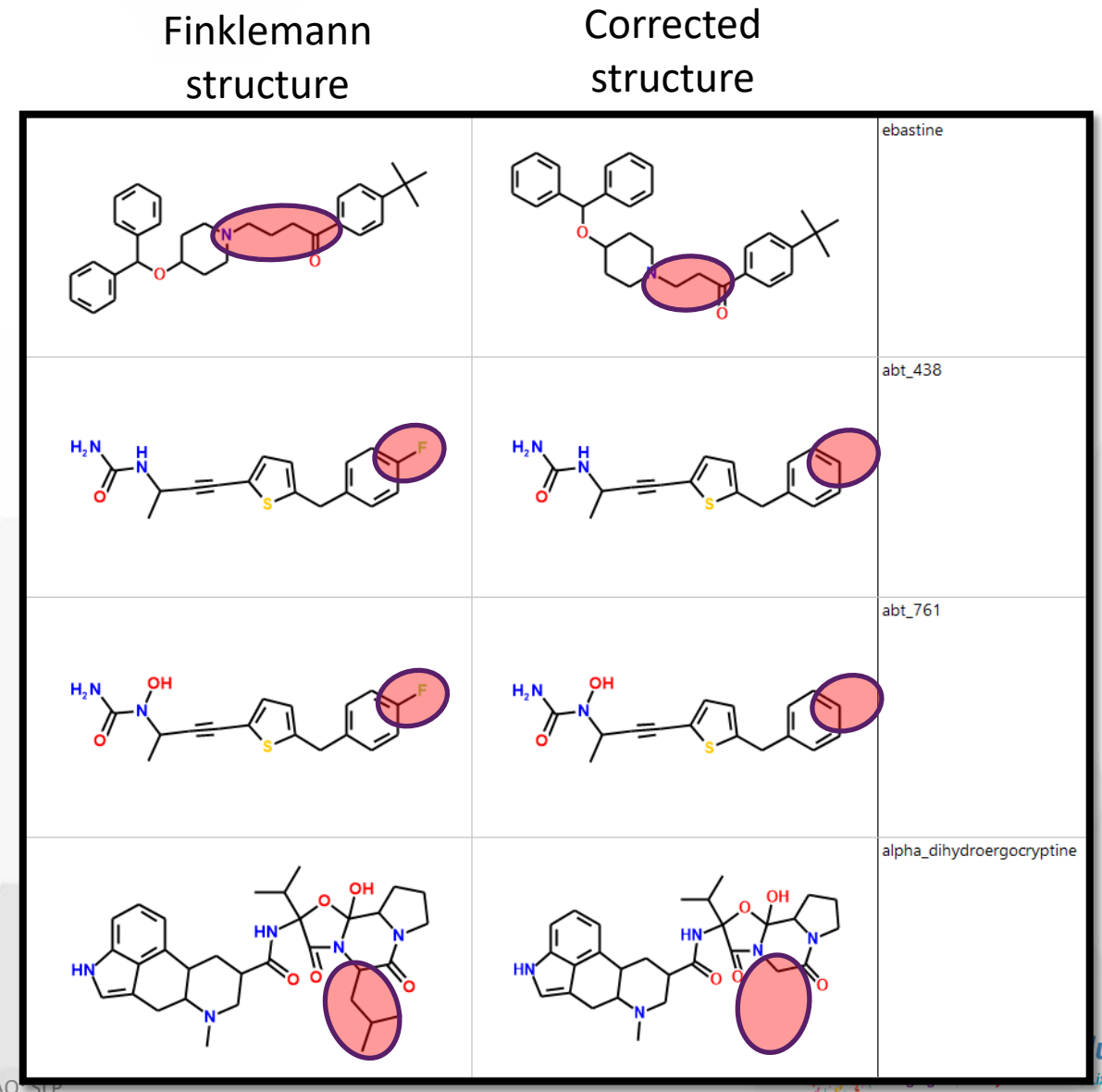
- CYP Substrate classification models for nine CYP isoforms
- Models built using data from various sources
  - Compiled datasets from literature (Sheridan et al, Zaretzki et al)
  - Commercial and academic databases
  - FDA review submissions
  - Literature
- Data included specific metabolites as well as cases where no metabolite is reported



Enzyme	Object	Object Metabolite
CYP2D6	(S)-gallopamil	(S)-gallopamil disappearance
CYP3A4	ebastine	ebastine disappearance
CYP2C19	tacrolimus	tacrolimus disappearance
CYP3A4	E2101	E2101 metabolite M4
CYP3A4	E2101	E2101 metabolite M1
CYP2C19	siponimod	siponimod disappearance
CYP2D6	(S)-fluoxetine	norfluoxetine
CYP3A4	dydrogester...	dydrogesterone disappearance
CYP2C19	sildenafil	N-desmethyl sildenafil (UK-103 320)

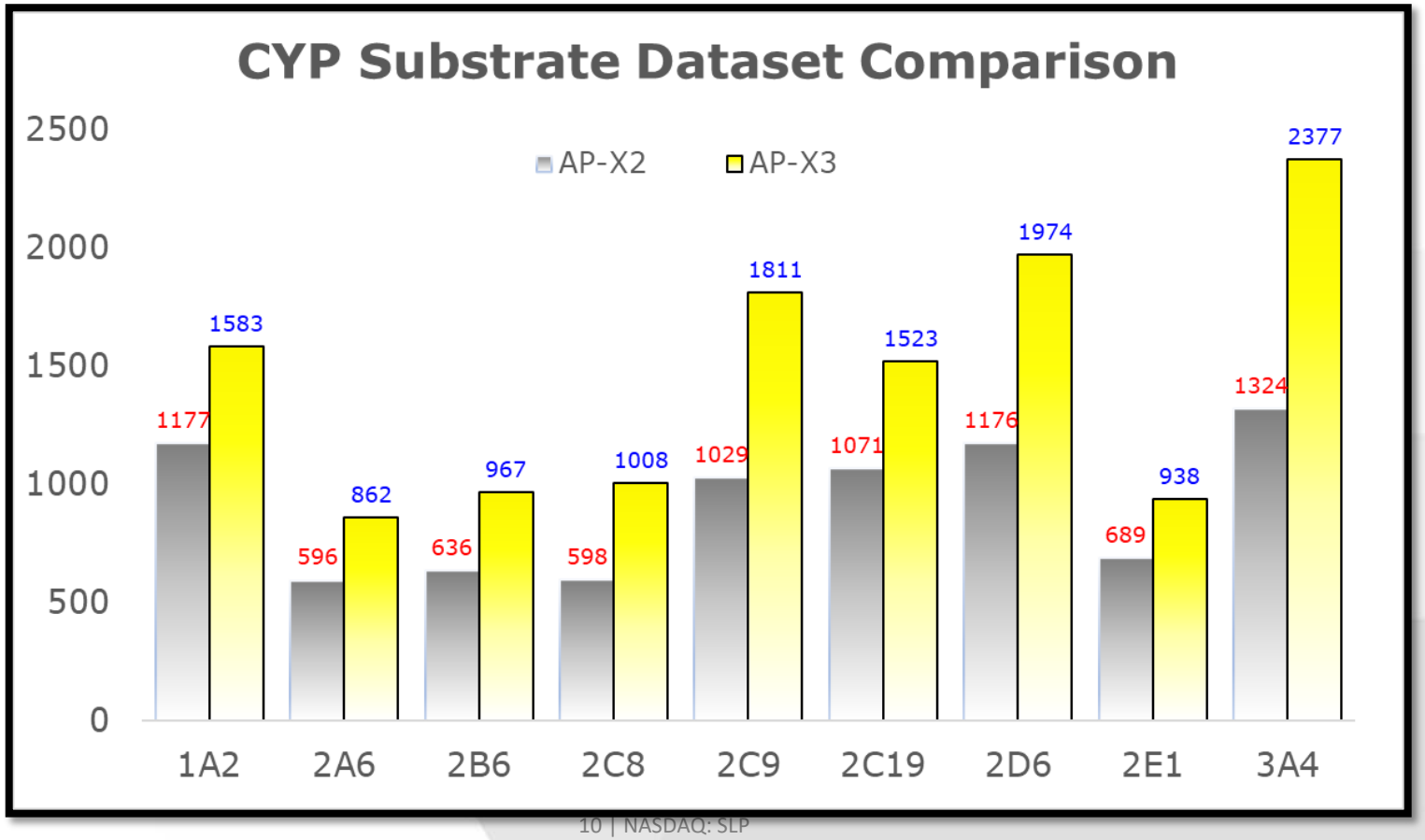
CYP Substrate Models: Data Curation

- Correcting chemical structures
- Updating correct site of metabolism
- Merge datasets from different sources
  - Remove identical duplicates
  - In case of extra SoM available, combine sites
- Exclude epoxidation sites



CYP Substrate Models: Final Datasets

- 2931 cmpds in CYP Substrate/Non-Substrate dataset
  - At least ~35% larger dataset compared to previous SLP dataset
  - Average increment: ~55%

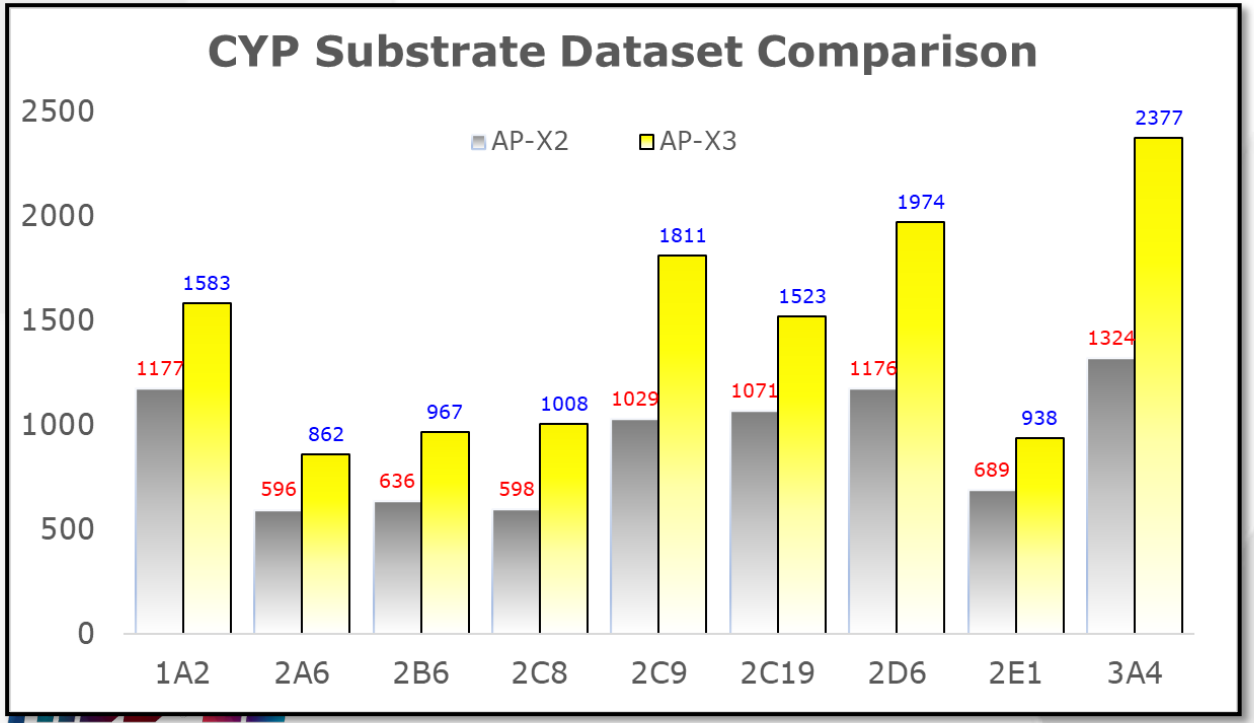


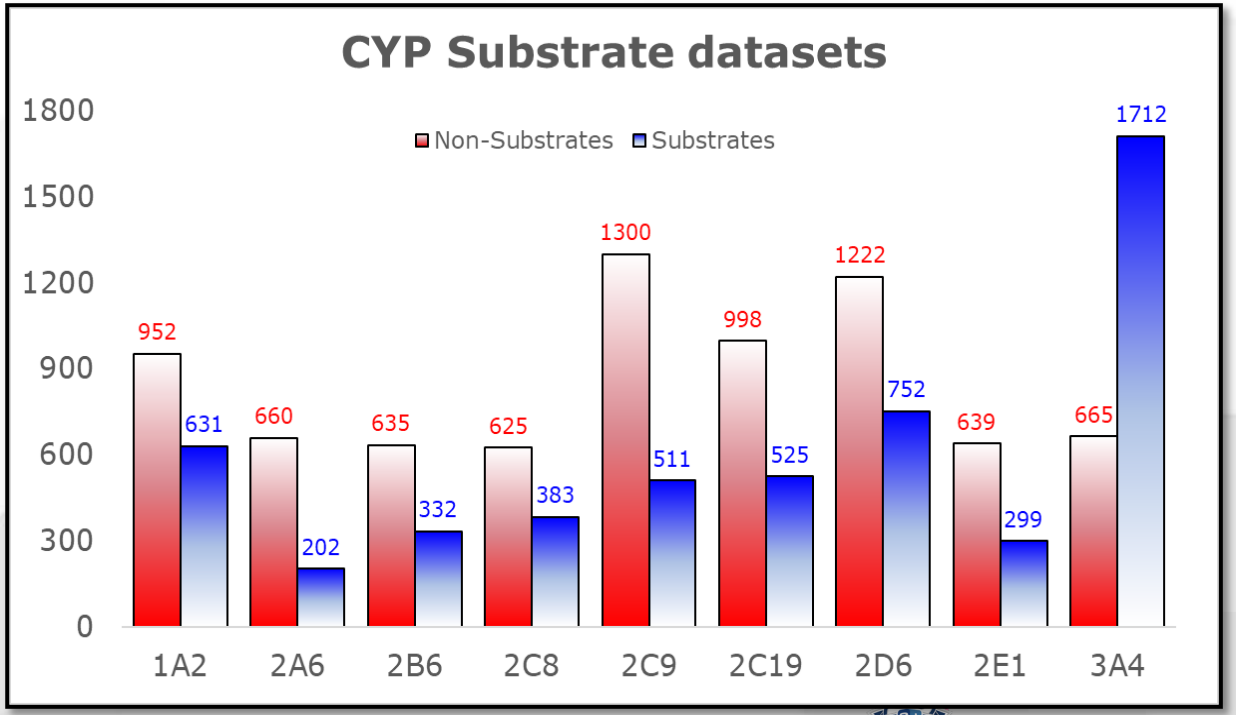
CYP Substrate Models: Final Datasets

- Comparison with our previous dataset
  - At least ~35% larger dataset compared to previous SLP dataset
  - Average increment: ~55%

Composition of each dataset

Substrate Vs Non-Substrates





CYP Substrate Models: Approach

Artificial Neural Network Ensemble

Constitutional Descriptors

Topological Indices

Electrotopological State Indices

Charge-based Descriptors

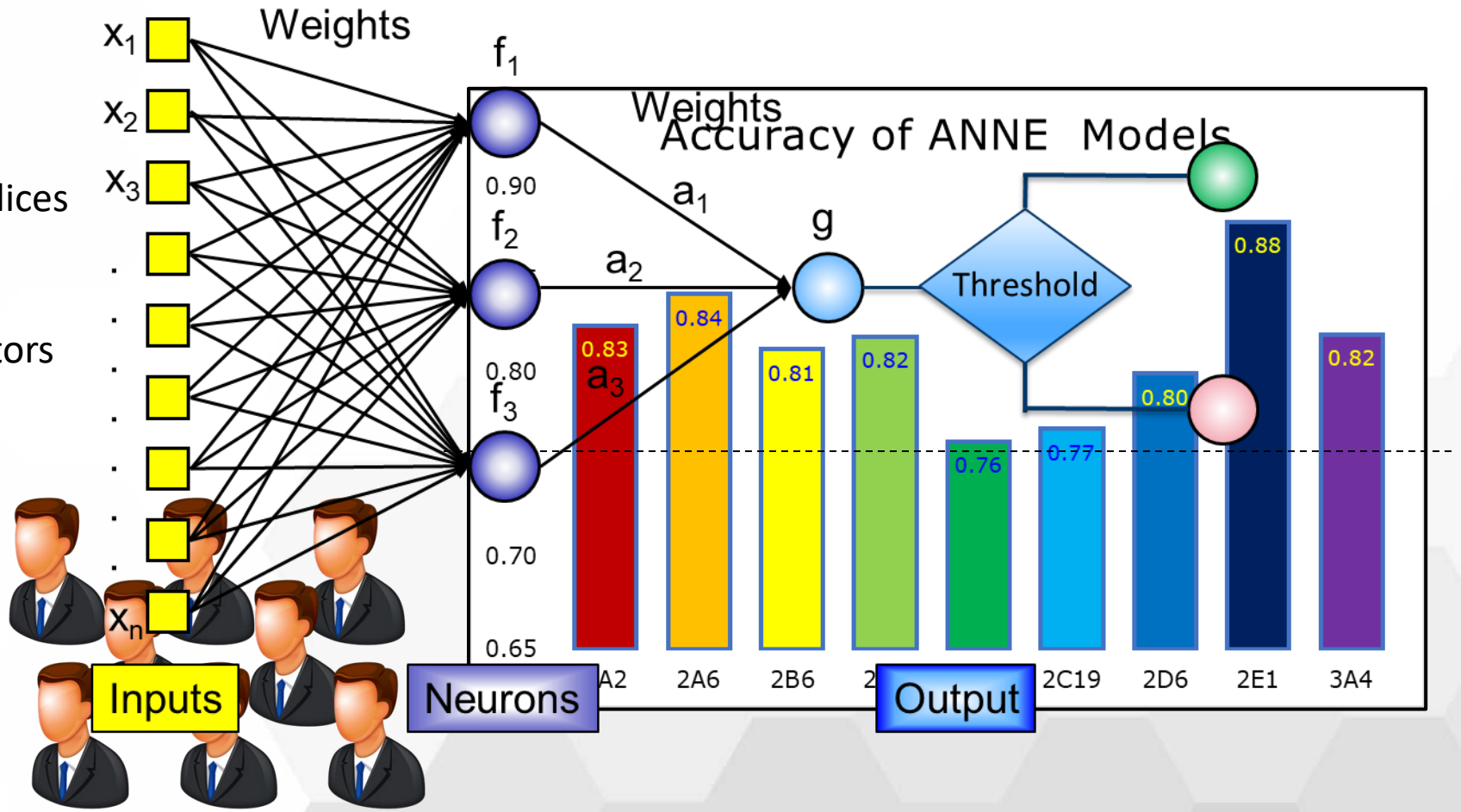
Hydrogen Bonding Descriptors

Ionization Descriptors

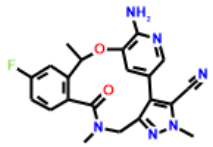
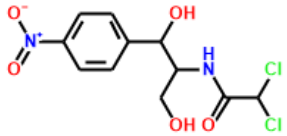
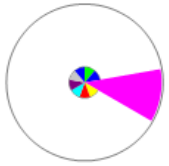
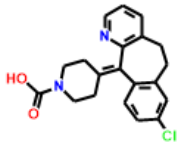
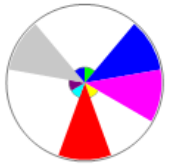
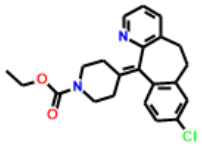
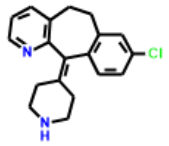
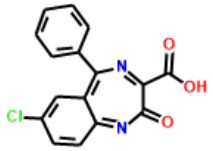
Functional Groups

Moriguchi Descriptors

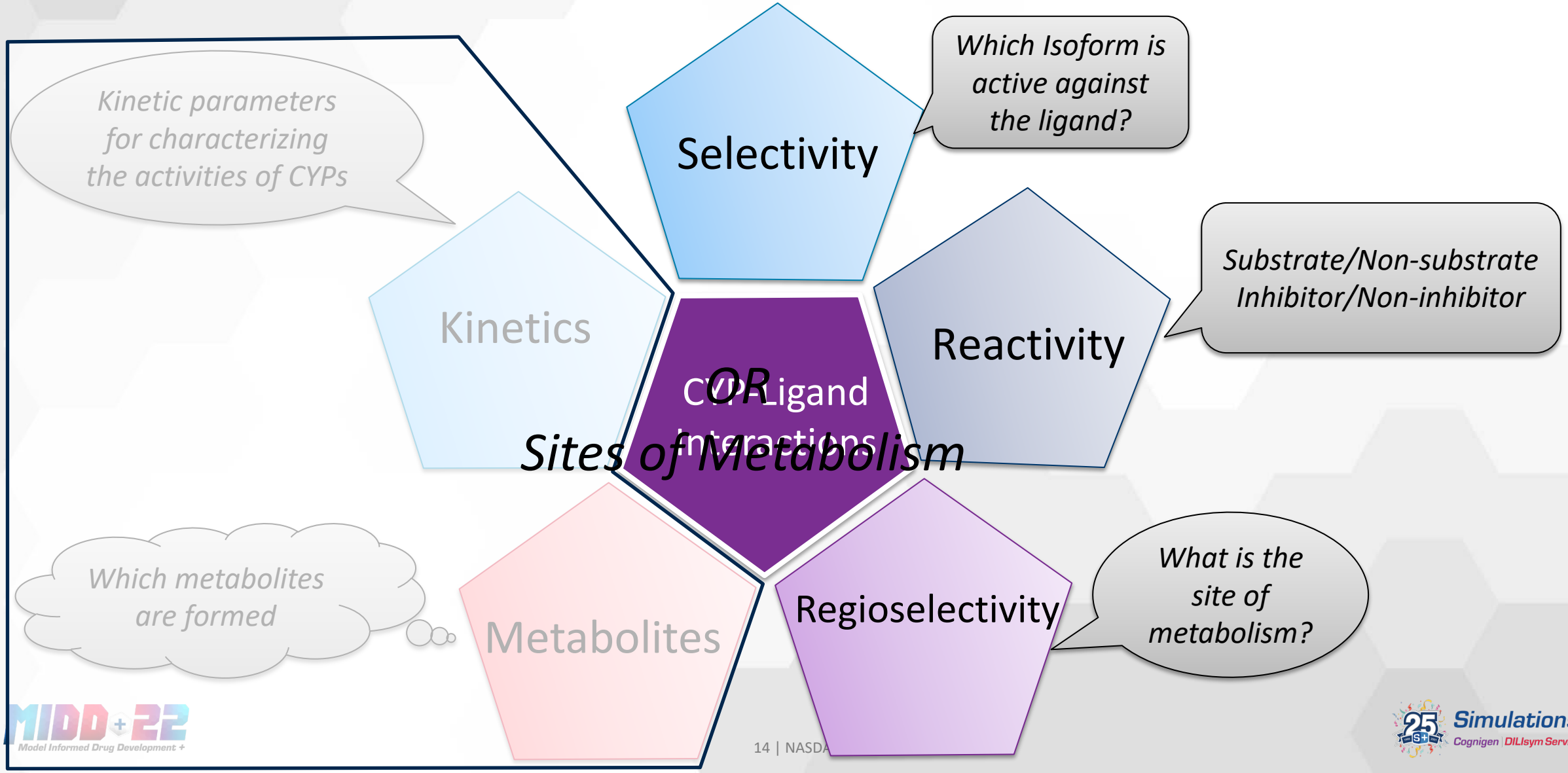
Vs



Final Output

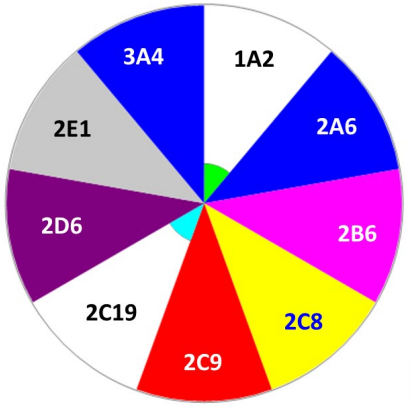
Structure	Identifier	*CYP Substr*	CYP1A2_Substr	CYP2A6_Substr	CYP2B6_Substr	CYP2C8_Substr	CYP2C9_Substr	CYP2C19_Substr	CYP2D6_Substr	CYP2E1_Substr	CYP3A4_Substr
	Lorlatinib		Yes (66%)	No (98%)	No (89%)	Yes (56%)	No (65%)	No (65%)	No (95%)	No (98%)	Yes (98%)
	Chloramphenicol		No (54%)	No (84%)	No (98%)	No (54%)	No (98%)	Yes (67%)	No (95%)	No (98%)	No (44%)
	O-deethyl-Loratid...		No (75%)	No (98%)	No (63%)	Yes (91%)	Yes (66%)	Yes (57%)	No (66%)	No (89%)	Yes (89%)
	loratadine		Yes (79%)	No (94%)	Yes (89%)	Yes (91%)	Yes (66%)	Yes (67%)	Yes (56%)	No (67%)	Yes (98%)
	desloratadine		Yes (70%)	No (98%)	Yes (71%)	Yes (83%)	No (65%)	Yes (67%)	Yes (87%)	No (87%)	Yes (98%)
	Clorazepic acid		No (81%)	No (98%)	No (83%)	Yes (60%)	No (76%)	No (80%)	No (95%)	No (98%)	No (33%)

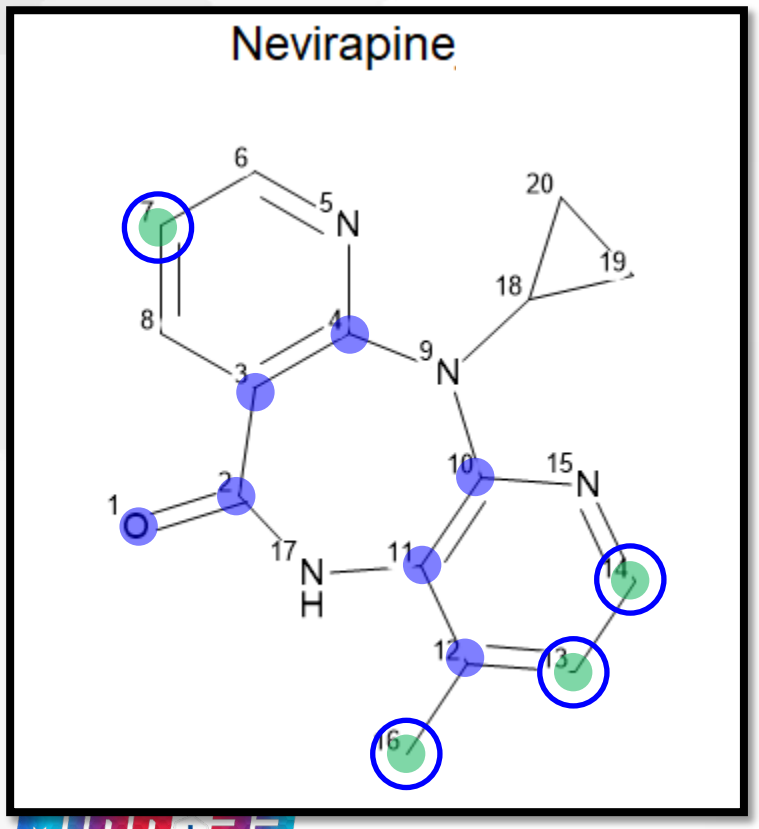
Regioselectivity Models in ADMET Predictor®



SoM Models : The Data

- Site of metabolism (SOM) models for nine CYP isoforms
- 1440 unique compounds
- >5500 unique metabolic sites





CYP	# Molecules	Total Atoms	Excluded Atoms	Included Atoms	Negatives	Positives
	Nevirapine	**20**	**7**	**13**	**9**	**4**
1A2	435	8804	3769	5035	4277	758
2A6	123	1958	756	1202	999	203
2B6	201	3841	1444	2397	2064	333
2C8	208	5304	2162	3142	2776	366

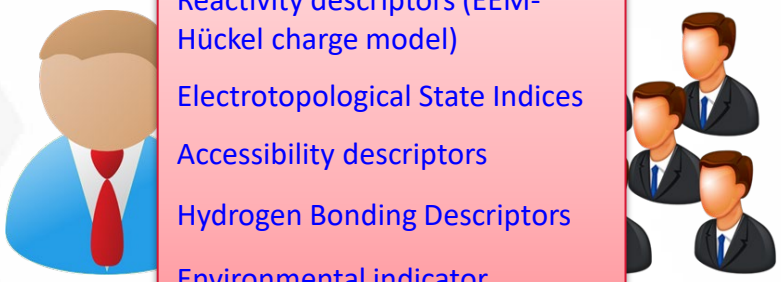
SoM Models: Approach

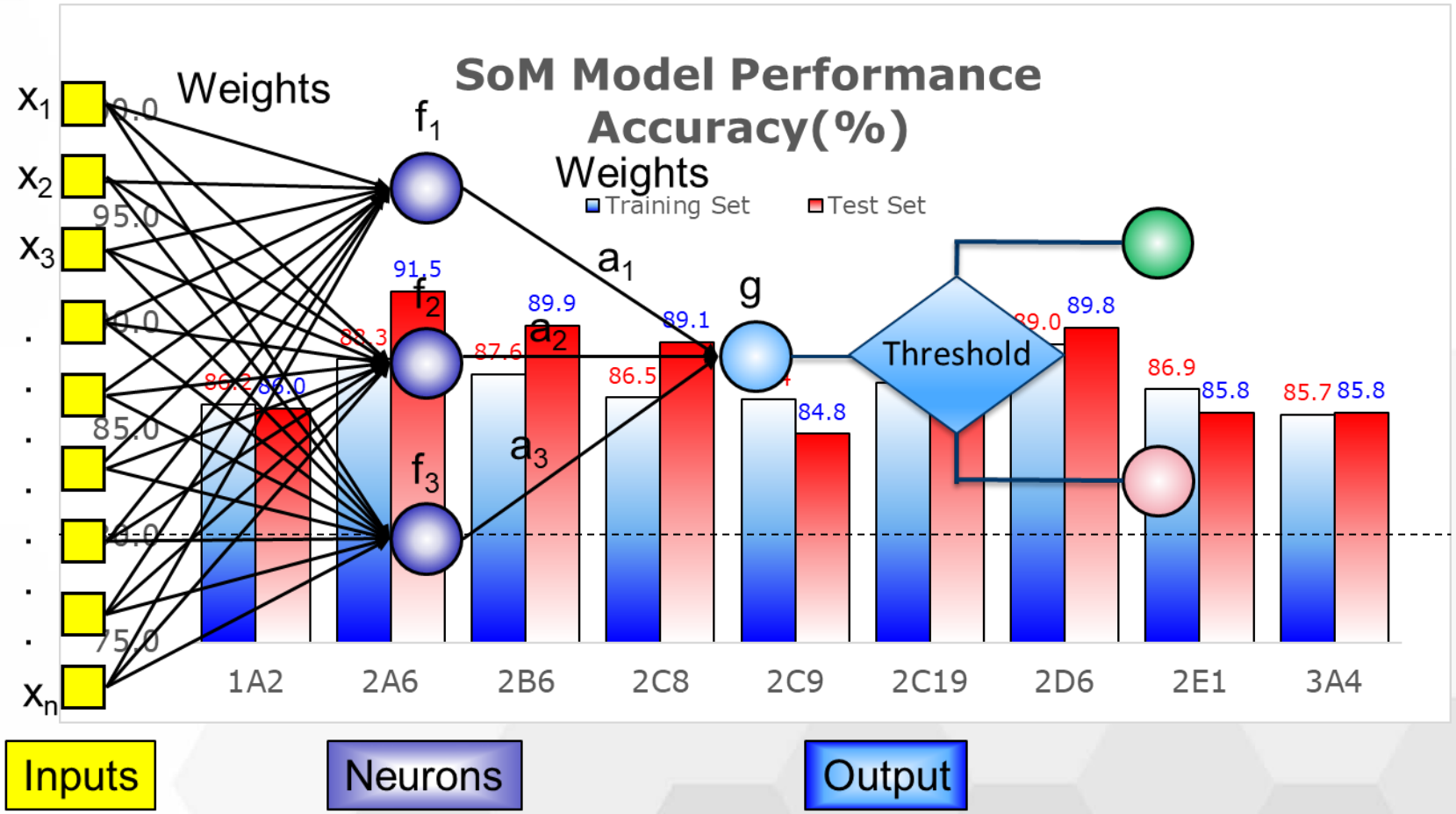
Molecular descriptors

- Constitutional Descriptors
- Topological Indices
- Electrotopological State Indices
- Charge-based Descriptors
- Hydrogen Bonding Descriptors
- Ionization Descriptors
- Functional Groups
- Moriguchi Descriptors

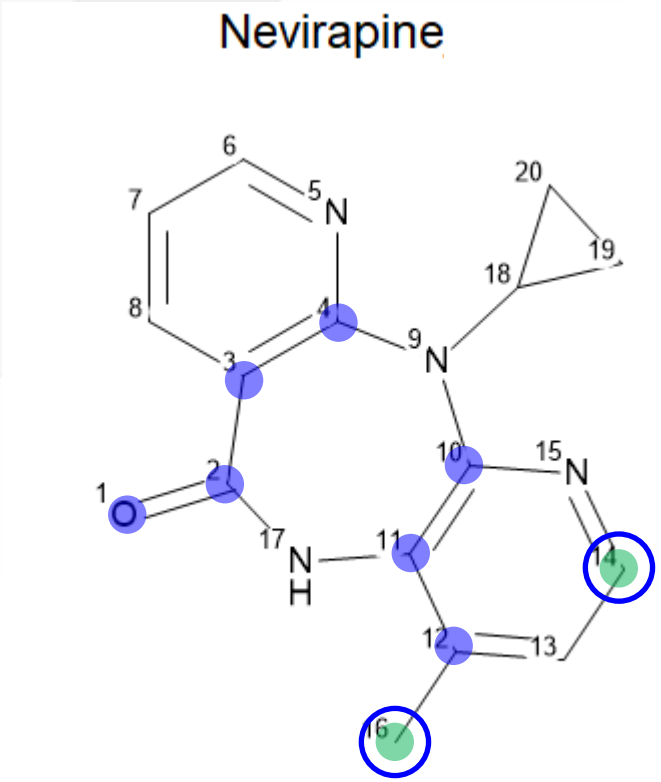
Atomic descriptors

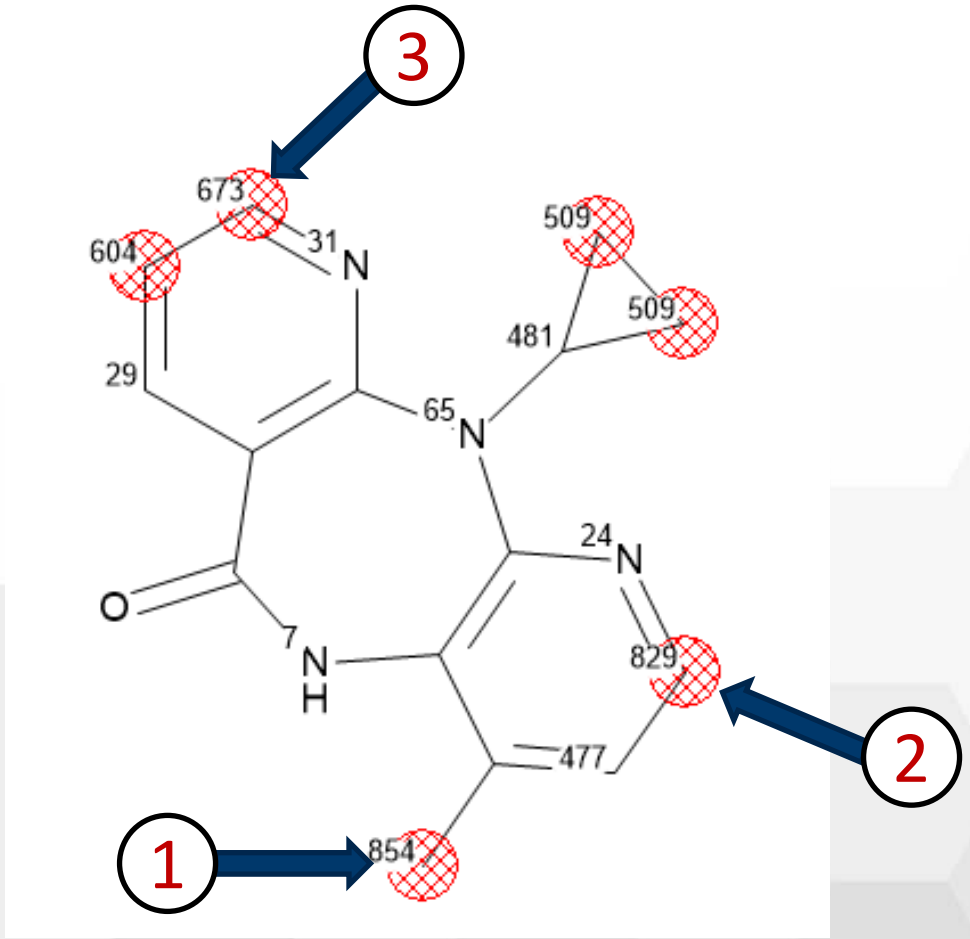
- Constitutional Descriptors
- Reactivity descriptors (EEM-Hückel charge model)
- Electrotopological State Indices
- Accessibility descriptors
- Hydrogen Bonding Descriptors
- Environmental indicator
- Topological Indices





Measuring Performance: Top Scoring Sites





Top-two Metric to Evaluate Performance

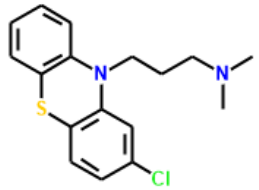
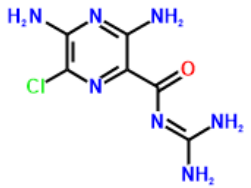
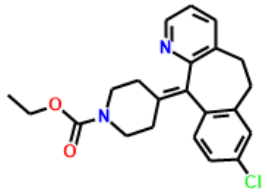
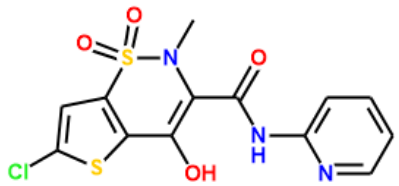
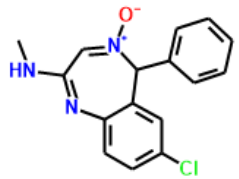
Table 2. XenoSite Is the Most Accurate Method for Predicting the Metabolism of the Majority of Curated CYP Substrates[a,b]

isozyme	1A2	2A6	2B6	2C8	2C9	2C19	2D6	2E1	3A4	HLM	
number of substrates	271	105	151	142	226	218	270	145	475	680	average
XenoSite[c]	87.1	85.7	83.4[(1)]	88.7[(2)]	86.7	89.0[(3)]	88.5	83.5	87.6	89.4[(4)]	87.0
RS-Predictor[d]	83.4	85.7[(4)]	82.1	83.8	84.5	86.2	85.9	82.8	82.3	86.2[(4)]	84.3
SMARTCyp	80.0	86.0	77.0	83.0	84.0	86.0	83.0	82.0	78.0		82.1
StarDrop					78.0		75.3		74.1		75.8
Schrödinger					72.1		68.1		76.4		72.2
Fingerprint[e]	66.1	63.8	64.2	65.5	68.1	69.3	74.4	64.1	71.2	75.3	68.2
random model	26.0	31.9	24.8	22.6	22.2	20.2	21.1	36.5	21.0	26.3	25.3

	435	**123**	**201**	**208**	**345**	**342**	**455**	**166**	**960**		
ADMET Predictor®	**89.2**	**88.9**	**87.1**	**96.6**	**81.0**	**81.6**	**85.5**	**92.0**	**84.0**		**87.3**

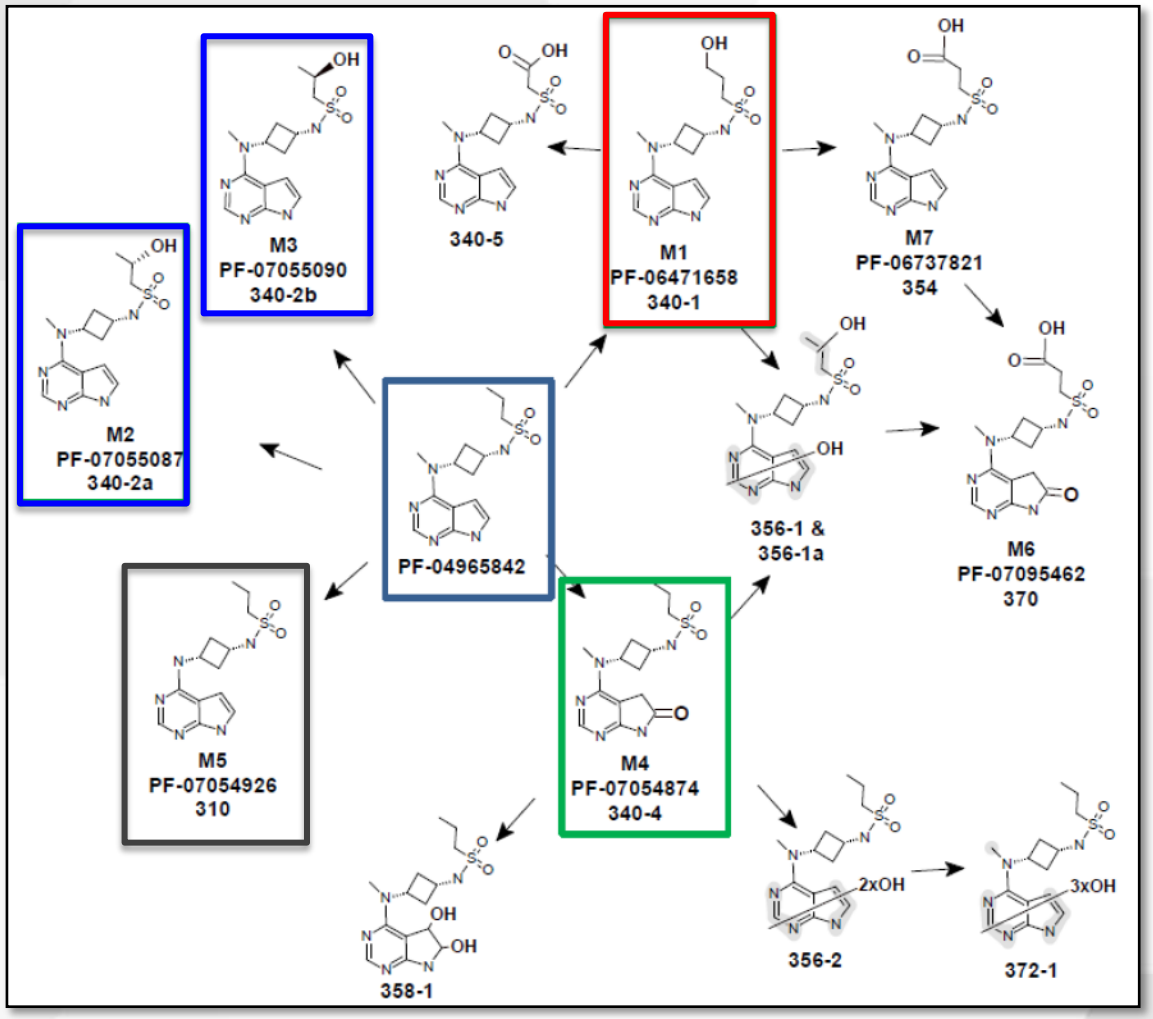
Zaretzki, *et al*, *J. Chem. Inf. Model.* **2013**, *53*, 3373−3383

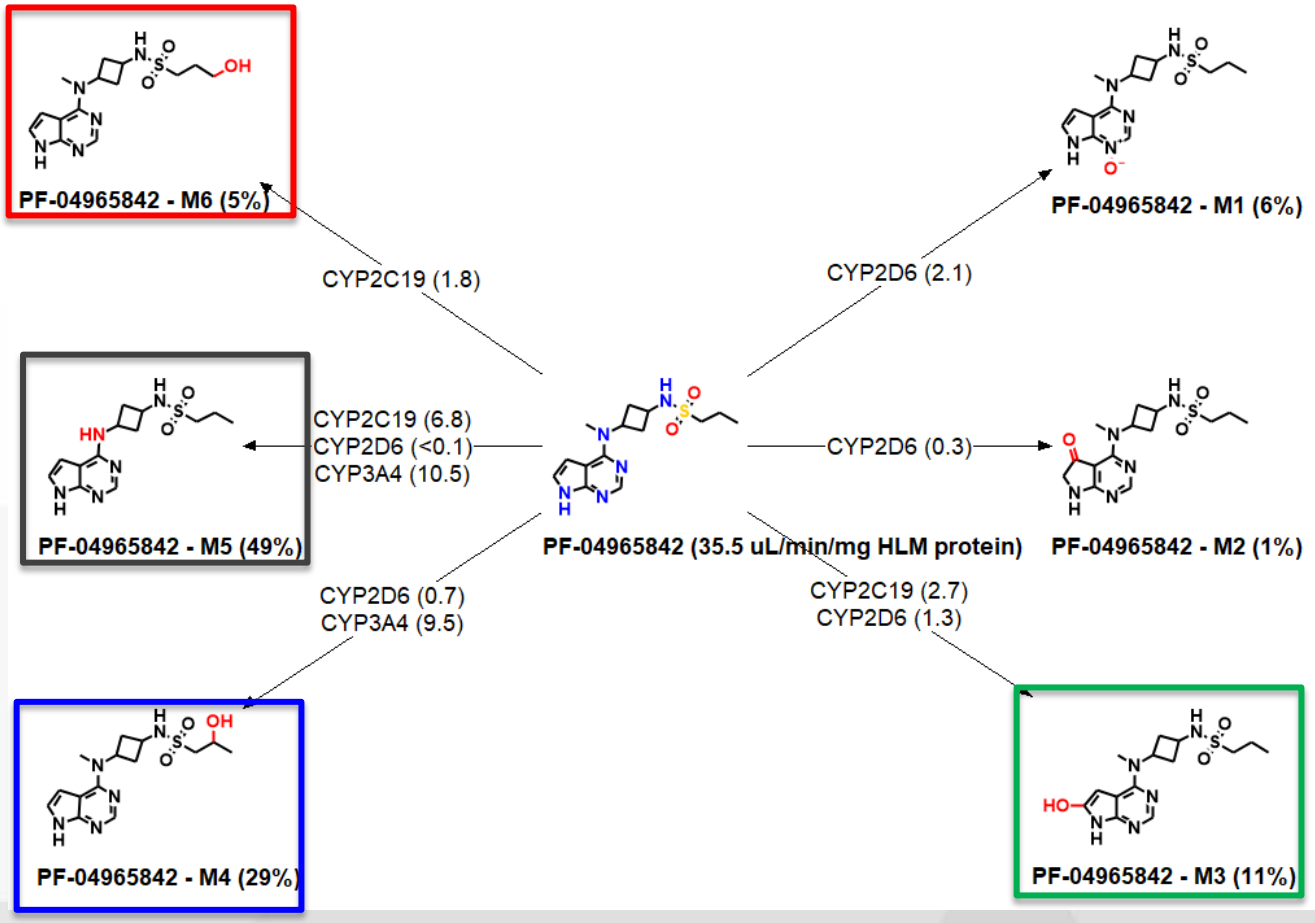
CYP Substrate Dependent Models in ADMET Predictor®

Structure	Identifier	CYP3A4_Substr	CYP3A4_Sites	CYP3A4_Km	CYP3A4_Vmax	CYP3A4_CLint
	Chlorpromazine	Yes (92%)	C20(972); C21(972); S8(962)	93.039	3.936	4.696
	Amiloride	No (84%)	NonSubstrate	NonSubstrate	NonSubstrate	NonSubstrate
	Loratadine	Yes (98%)	C4(923); C17(581); C18(543)	49.191	35.132	79.277
	Lornoxicam	No (40%)	NonSubstrate	NonSubstrate	NonSubstrate	NonSubstrate
	Chlordiazepoxide	Yes (98%)	C9(996)	43.507	6.859	17.500

Case Study: Abrocitinib

Recently approved by FDA To treat refractory, moderate-to-severe atopic dermatitis (Jan 14, 2022)





https://www.accessdata.fda.gov/drugsatfda_docs/nda/2022/213871Orig1s000MultidisciplineR.pdf

Conclusions

- We recently expanded chemical applicability of CYP models
  - Nine CYP Substrate models
  - Nine CYP regioselectivity models (Sites of Metabolism)
- The new models showed higher accuracy, over 80%
- Models are applicable to a very broad range of molecules
- When combined with pre-defined set of biotransformation rules, can be successfully used to predict plausible metabolites of small molecules

Acknowledgements

- Michael Lawless
- Phyo Phyo Zin
- David Miller
- Marvin Waldman
- Bob Clark

SimulationsPlus

MIDD+22

Model Informed Drug Development

Q&A

Questions & Answers

pankaj.daga@simulations-plus.com