Homology modeling and molecular docking

recap

Donovan J. Blatz Carlie A. LaLone Ph.D

Office of Research and Development Center for Computational Toxicology and Exposure, Great Lakes Toxicology and Ecology Division The views expressed in this poster are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA.

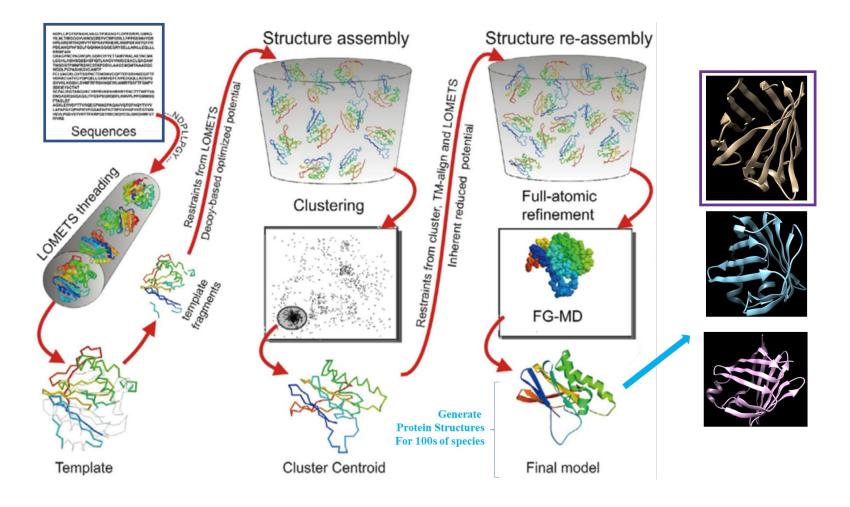


Overview

- Created homology models for 10 high priority proteins
 - Look at model scores
- Docked thiacloprid and flupyradifurone to models
 - Look at docking scores across species and pesticides
- Compared scores between both docking areas



Th Creating homology models with the Iterative Threading ASSEmbly Refinement (ITASSER) tool





Models created and scores

Protein	C-Score	TM-Score	RMSD	decoys	density
CYP6AS11	1.32	0.90+-0.06	4.5+-3.0	7707	0.571
CYP6AS13	1.21	0.88+-0.07	4.7+-3.1	8571	0.533
CYP6BD1	0.68	0.81+-0.09	5.9+-3.7	4028	0.314
CYP9DL5	1.58	0.94+-0.06	4.1+-2.7	6294	0.761
CYP9DM1	1.42	0.91+-0.06	4.3+-2.9	8639	0.65
CYP9DM2	1.23	0.88+-0.07	4.8+-3.1	9784	0.555
CYP9FT2	1.75	0.96+-0.05	3.7+-2.5	5596	0.902
CYP9FU3	1.14	0.87+-0.07	5.0+-3.2	7131	0.486
CYP9FZ2	1.17	0.87+-0.07	4.9+-3.2	7127	0.506
CYP9Q10	1.36	0.9+-0.06	4.5+-3.0	7430	0.605

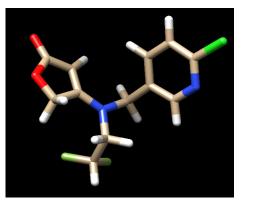
Project	Protein	Template	Species	Model #	Cscore	TM-Score	RMSD	Decoys	Density
CYP9Q	CYP3A4	1TQN	Human	1 of 5	1.09	0.86+-0.07	5.0+-3.2	3316	0.556
CYP9Q	CYP9Q3	1TQN	Apis mellifera	1 of 5	1.01	0.85+-0.08	5.2+-3.3	4559	0.428
CYP9Q	CYP9Q4	1TQN	Bombus terrestris	1 of 5	0.83	0.83+-0.08	5.6+-3.5	2088	0.365
CYP9Q	CYP9BU2	1TQN	Osmia bicornis	1 of 5	0.57	0.79+-0.09	6.1+-3.8	1579	0.28
CYP9Q	CYP9Q3	None	Apis mellifera	1 of 5	1.07	0.86+-0.07	5.1+-3.3	4515	0.456
CYP9Q	CYP9Q4	None	Bombus terrestris	1 of 5	0.72	0.80+-0.09	5.9+-3.7	3153	0.326
CYP9Q	CYP9BU2	None	Osmia bicornis	1 of 5	0.54	0.79+-0.09	6.2+-3.8	1042	0.272

ITASSER scoring text

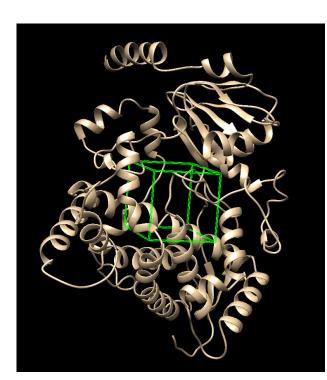
- Correlation coefficient of Cscore of the first model with TM-score to the native structure is 0.91, while the coefficient of C-score with RMSD to the native structure is 0.75
- TM-score >0.5 indicates a model of correct topology and a TM-score<0.17 means a random similarity
- A higher cluster density means the structure occurs more often in the simulation trajectory and therefore signifies a better-quality model



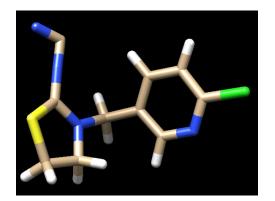
Docking process



Flupyradifurone



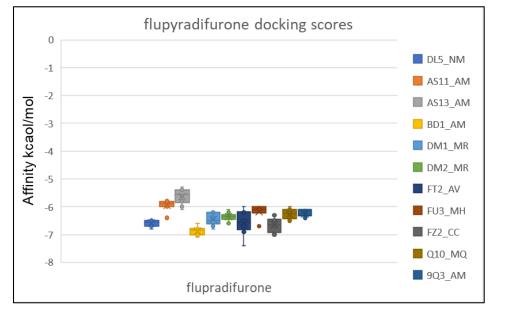
Docking box areas
EPA:
-17.098,-22.338,-11.492
20,15,15
Bayer:
19.7332,25.7236,-4.67613
16.0055,17.7187,23.9048
(Used negative x,y,z coordinates due to location of ITASSER models)



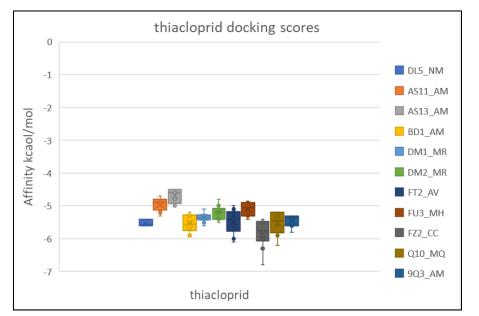
Thiacloprid



Docking scores between models and boxes



		Average	docking sco			
	AS11_AM	AS13_AN	1 BD1_AM	DM1_MR	DM2_MR	
flupyradifurone	-5.94	-5.6	3 -6.87	-6.45	-6.36	EPA Box
паругаанагоне	-6	-5.6	8 -6.52	-6.54	-6.24	Bayer Box
this also wid	-4.96	-4.6	9 -5.53	-4.29	-5.19	EPA Box
thiaclorpid	-5.3	-4.9	7 -5.4	-5.32	-4.91	Bayer Box
	flupy	thia				
Reference	-6.25	-5.4	4 EPA Box			
CYP9Q3_AM	_AM -6.61 -5.6		4 Bayer Bo	x		



"It includes 5 of the CYP9Q-orthologs we tested and have thiacloprid and flupyradifurone as their confirmed substrate and 5 enzymes (3 from the honey bee and two from Megachile rotundata), where metabolism of those insecticides is not present *in-vitro.*"



Previously made SeqAPASS heat map

Total Match Partial Match Not a Match Susceptible No									
NCBI Accession	Scientific Name	Protein Name	Common Name	Similar Susceptibility	Amino Acid 1	Amino Acid 2	Amino Acid 3	Amino Acid 4	Amino Acid 5
XP_006562363.1 Ap	pis mellifera	cytochrome P450 9e2	Honey bee	Y	111G	214N	310S	371V	372L
XP_016922294.2 Ap	pis cerana	LOW QUALITY PROTEIN: cytochrome P450 9e2-like	Asiatic honeybee	Y	111G	214N	310S	3711	372L
XP_017758640.1 Eu	ufriesea mexicana	PREDICTED: cytochrome P450 9e2-like	Orchid bees	Ν	110K	214M	307T	3681	369A
XP_031775226.1 Ap	pis florea	cytochrome P450 9e2-like	Little honeybee	Ν	112K	216T	309T	370V	371V
XP_017794730.1 Ha	abropoda laboriosa	PREDICTED: cytochrome P450 9e2-like	Digger bees	N	108K	212L	305T	3661	367V
XP_031837097.1 No	omia melanderi	cytochrome P450 9e2-like	Alkali bee	Ν	110K	214D	307S	368S	369P



Conclusions

- Models created are "quality models" based off scoring metrics
- Sequence based predictions differ from structural based predictions
- SeqAPASS predictions can be difficult when comparing sequences that don't align with template protein