HOMOLOGY MODEL RECAP

Donovan Blatz

Carlie A. LaLone, Ph.D.

Office of Research and Development
Center for Computational Toxicology and Exposure, Great Lakes Toxicology and Ecology Division

EPA United States Environmental Protection Agency

Overview

- Look into N214A/F amino acid change with DUET protein stability and Autodock Vina for both Bayer and ITASSER protein models
- Overview ITASSER process and scoring for homology modeling
- Compare ITASSER created models visually to Bayer created models
- Dock flupyradifurone and thiacloprid to models of interest

SeqAPASS: Sequence-based results

Total Match | Partial Match | Not a Match | Susceptible Yes | Susceptible No

Common Name	Similar Susceptibility	Amino Acid 1	Amino Acid 2	Amino Acid 3	Amino Acid 4	Amino Acid 5
Apis mellifera CYP9Q3	Y	111G	214N	310S	371V	372L
Dufourea novaeangliae CYP9DL4	N	111N	214N	307T	368A	369A
Apis mellifera CYP9Q1	Y	111G	211Q	308T	370I	371L
Apis mellifera CYP9Q2	N	113S	216T	310T	371I	372A
Apis cerana CYP9Q3	Y	111G	214N	310S	371I	372L
Apis dorsata CYP9Q3	Y	111G	214N	310S	371I	372L
Apis mellifera CYP9Q3	Y	111G	214N	310S	371V	372L
Bombus impatiens CYP9Q4	N	112E	215T	308A	369V	370A
Bombus impatiens CYP9Q5	N	112E	215T	308T	369V	370A
Bombus impatiens CYP9Q6	N	111D	214F	306S	367I	368T
Eufriesa mexicana CYP9Q7	N	111D	214M	307S	368I	369A
Eufriesa mexicana CYP9Q8	N	111E	214M	307T	368I	369A
Habropoda laboriosa CYP9Q9	N	109R	212L	305T	366I	367V
Apis florea CYP9Q-like-1	N	111N	214T	306T	367I	368A
Tetragonila carbonaria CYP9Q-like-1	N	111K	215S	309S	370I	371A
Apis florea CYP9Q-like-2	N	113S	216T	309T	370V	371V
Tetragonila carbonaria CYP9Q-like-2	N	111K	214S	307S	368I	369A
Andrena vaga CYP9Q-like	N	111A	211F	304T	365I	366G
Andrena haemorrhoa CYP9Q-like	N	111A	211F	304T	365I	366G
Colletes cunicularius CYP9Q-like	N	111E	214N	307T	368I	369A
E nigrescens CYP9Q-like	N	111E	214S	307T	368I	369V
Lasioglossum xanthopus CYP9Q-like	N	111D	215N	308S	369V	370A
Macropis fulvipes CYP9Q-like	N	108K	211T	301S	362V	363A
Melitta haemorrhoidalis CYP9Q-like	N	111E	214L	306T	367I	368A
N lathburiana CYP9Q-like	N	111E	214A	307S	368I	369A
N melanderi CYP9Q-like	N	111D	214D	307S	368S	369P
Xylocopa violacea CYP9Q-like	N	111D	214T	307S	368I	369V
Melipona quadrifasciata CYP9Q-like	N	111K	214S	307T	368I	369A

Amino Acid info

ID	Name	Side Chain	Size
A	Alanine	Aliphatic	89.094
C	Cysteine	Sulfur-Containing	121.154
D	Aspartic Acid	Acidic	133.104
E	Glutamic Acid	Acidic	147.131
F	Phenylalanine	Aromatic	165.192
G	Glycine	Aliphatic	75.067
H	Histidine	Basic	155.156
I	Isoleucine	Aliphatic	131.175
K	Lysine	Basic	146.189
L	Leucine	Aliphatic	131.175
M	Methionine	Sulfur-Containing	149.208
N	Asparagine	Amidic	132.119
P	Proline	Aliphatic	115.132
Q	Glutamine	Amidic	146.146
R	Arginine	Basic	174.203
S	Serine	Hydroxylic	105.093
T	Threonine	Hydroxylic	119.119
U	Seleno-cysteine	Sulfur-Containing	168.064
V	Valine	Aliphatic	117.148
W	Tryptophan	Aromatic	204.228
X	Unknown	Unknown	-
Y	Tyrosine	Aromatic	181.191

N214F – Amidic to Aromatic AND MW diff. 33.073 g/mol
N214A – Amidic to Aliphatic AND MW diff. 43.025 g/mol

Question: Are mutations in structure destabilizing and consistent with sequence-based predictions?

Use DUET to explore mutations in protein structure

- **Need structural models**
  - **PDB**
  - **Point mutation**

Protein stability of N214A/F of Apis mellifera CYP9Q3

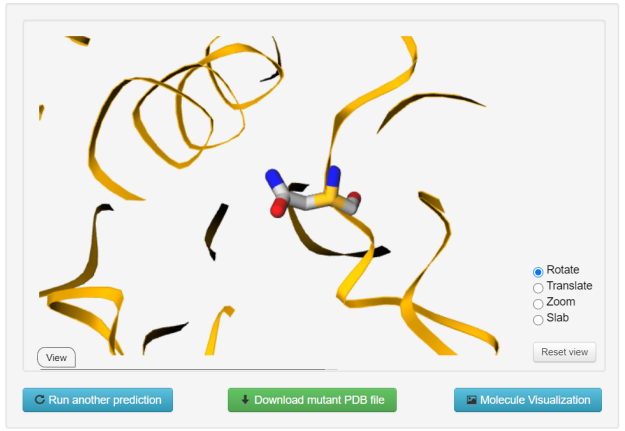
I-TASSER
Protein Structure & Function Predictions

mCSM Predicted Stability Change (ΔΔG):
-0.643 Kcal/mol (*Destabilizing*)

SDM Predicted Stability Change (ΔΔG):
0.92 Kcal/mol (*Stabilizing*)

DUET Predicted Stability Change (ΔΔG):
-0.181 Kcal/mol (*Destabilizing*)

Mutation:
Wild-type: **ASN**
Position: **214**
Mutant-type: **ALA**
Chain: **A**
Secondary structure: **Loop or irregular**



mCSM Predicted Stability Change (ΔΔG):
-0.761 Kcal/mol (*Destabilizing*)

SDM Predicted Stability Change (ΔΔG):
0.74 Kcal/mol (*Stabilizing*)

DUET Predicted Stability Change (ΔΔG):
-0.5 Kcal/mol (*Destabilizing*)

Mutation:
Wild-type: **ASN**
Position: **214**
Mutant-type: **PHE**
Chain: **A**
Secondary structure: **Loop or irregular**

BAYER

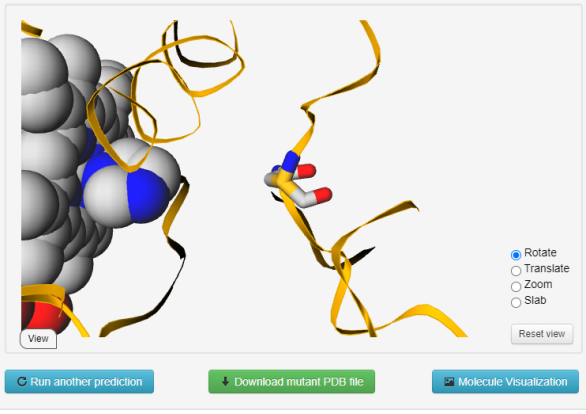
Both AM CYP9Q3 raw model

mCSM Predicted Stability Change (ΔΔG):
-0.259 kcal/mol (*Destabilizing*)

SDM Predicted Stability Change (ΔΔG):
1.29 kcal/mol (*Stabilizing*)

DUET Predicted Stability Change (ΔΔG):
0.293 kcal/mol (*Stabilizing*)

Mutation:
Wild-type: **ASN**
Position: **214**
Mutant-type: **ALA**
Chain: **A**
Secondary structure: **Loop or irregular**

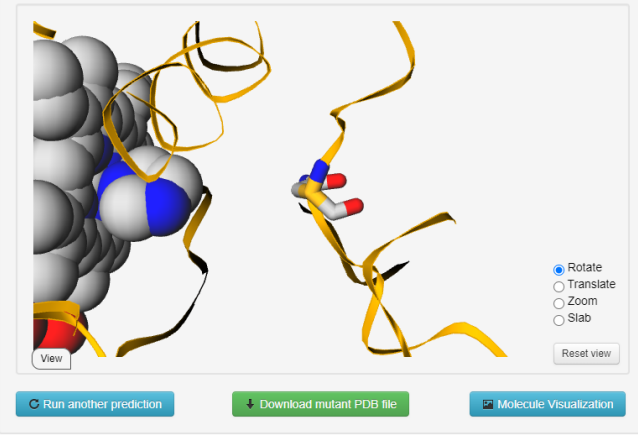


mCSM Predicted Stability Change (ΔΔG):
-0.625 kcal/mol (*Destabilizing*)

SDM Predicted Stability Change (ΔΔG):
0.85 kcal/mol (*Stabilizing*)

DUET Predicted Stability Change (ΔΔG):
-0.346 kcal/mol (*Destabilizing*)

Mutation:
Wild-type: **ASN**
Position: **214**
Mutant-type: **PHE**
Chain: **A**
Secondary structure: **Loop or irregular**



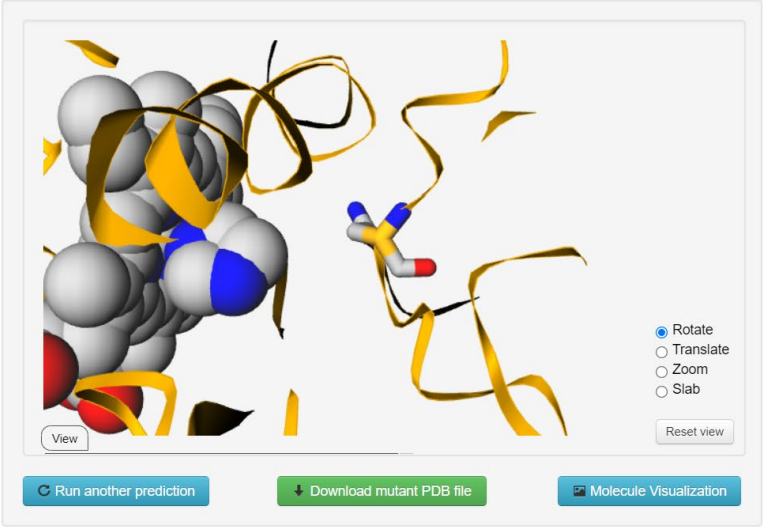
Protein Stability N214A/F of Apis mellifera CYP9Q3 cont.

mCSM Predicted Stability Change (ΔΔG):
-1.265 kcal/mol (*Destabilizing*)

SDM Predicted Stability Change (ΔΔG):
0.19 kcal/mol (*Stabilizing*)

DUET Predicted Stability Change (ΔΔG):
-1.034 kcal/mol (*Destabilizing*)

Mutation:
Wild-type: **ASN**
Position: **214**
Mutant-type: **ALA**
Chain: **A**
Secondary structure: **Loop or irregular**

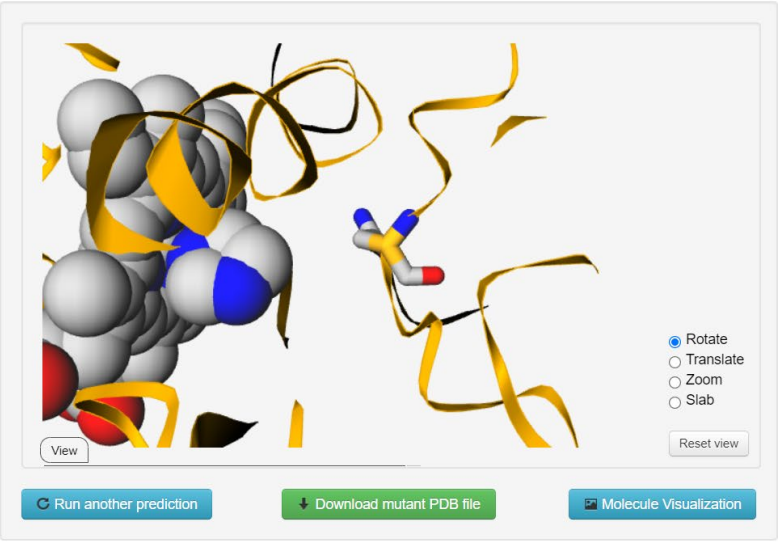


mCSM Predicted Stability Change (ΔΔG):
-1.15 kcal/mol (*Destabilizing*)

SDM Predicted Stability Change (ΔΔG):
-0.48 kcal/mol (*Destabilizing*)

DUET Predicted Stability Change (ΔΔG):
-1.267 kcal/mol (*Destabilizing*)

Mutation:
Wild-type: **ASN**
Position: **214**
Mutant-type: **PHE**
Chain: **A**
Secondary structure: **Loop or irregular**



AM CYP9Q3 Bayer's minimized model

Apis mellifera CYP9Q3 Dock prepped

mCSM Predicted Stability Change (ΔΔG):

-0.643 kcal/mol (*Destabilizing*)

SDM Predicted Stability Change (ΔΔG):

0.92 kcal/mol (*Stabilizing*)

DUET Predicted Stability Change (ΔΔG):

-0.181 kcal/mol (*Destabilizing*)

Mutation:

Wild-type: **ASN**
Position: **214**
Mutant-type: **ALA**
Chain: **A**
Secondary structure: **Loop or irregular**

mCSM Predicted Stability Change (ΔΔG):

-0.761 kcal/mol (*Destabilizing*)

SDM Predicted Stability Change (ΔΔG):

0.74 kcal/mol (*Stabilizing*)

DUET Predicted Stability Change (ΔΔG):

-0.5 kcal/mol (*Destabilizing*)

Mutation:

Wild-type: **ASN**
Position: **214**
Mutant-type: **PHE**
Chain: **A**
Secondary structure: **Loop or irregular**

ITASSER min (100 grad) and overall scoring

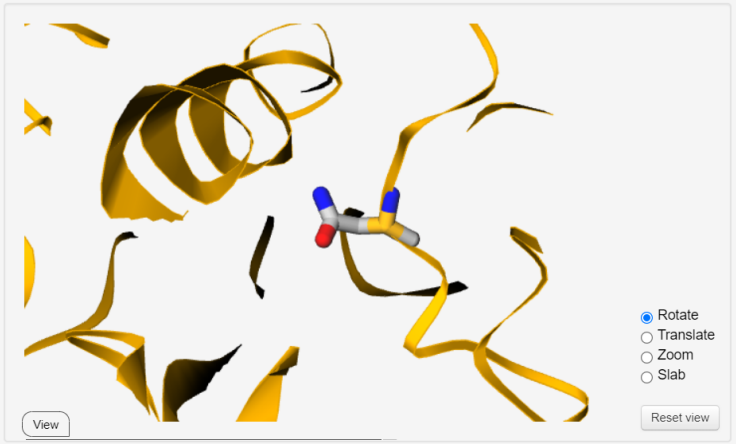
I-TASSER
Protein Structure & Function Predictions

mCSM Predicted Stability Change (ΔΔG):
-0.659 kcal/mol (*Destabilizing*)

SDM Predicted Stability Change (ΔΔG):
0.92 kcal/mol (*Stabilizing*)

DUET Predicted Stability Change (ΔΔG):
-0.197 kcal/mol (*Destabilizing*)

Mutation:
Wild-type: **ASN**
Position: **214**
Mutant-type: **ALA**
Chain: **A**
Secondary structure: **Loop or irregular**

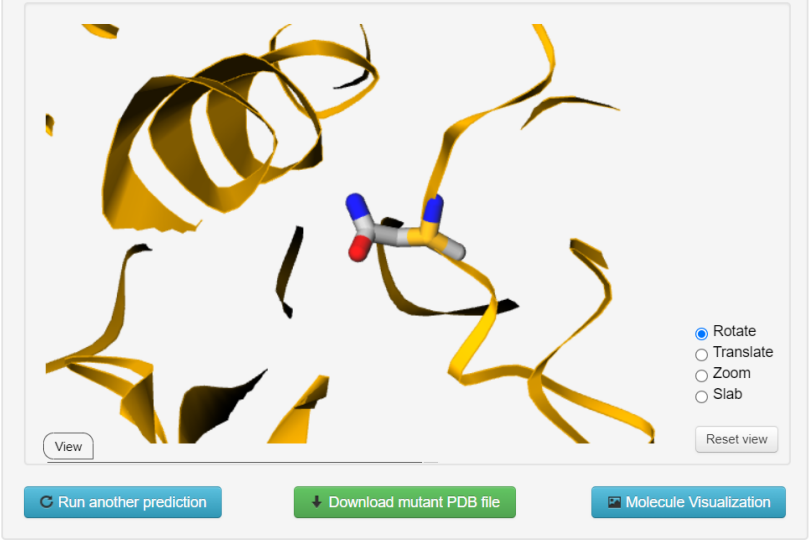


mCSM Predicted Stability Change (ΔΔG):
-0.7 kcal/mol (*Destabilizing*)

SDM Predicted Stability Change (ΔΔG):
0.74 kcal/mol (*Stabilizing*)

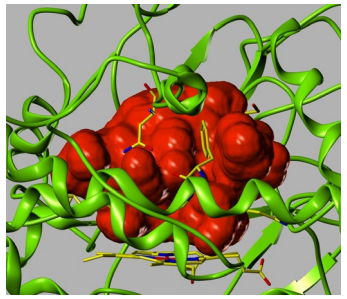
DUET Predicted Stability Change (ΔΔG):
-0.447 kcal/mol (*Destabilizing*)

Mutation:
Wild-type: **ASN**
Position: **214**
Mutant-type: **PHE**
Chain: **A**
Secondary structure: **Loop or irregular**

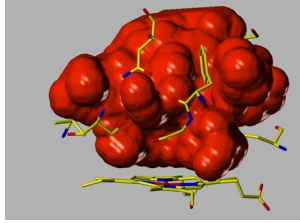


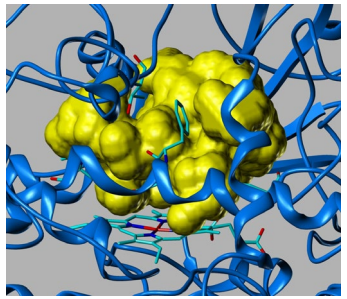
		ITASSER	100	Bayer	2500
		N214A	N214F	N214A	N214F
Duet predicted stability change	Raw model	-0.181	-0.5	0.293	-0.346
	Min model	-0.197	-0.447	-1.034	-1.267

A robust computational modeling approach predicting P450-mediated detoxification *in silico* largely fails due to the lack of respective bee P450 crystal structures

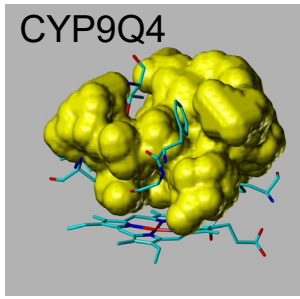


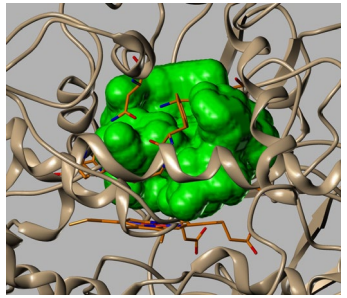
CYP9Q3



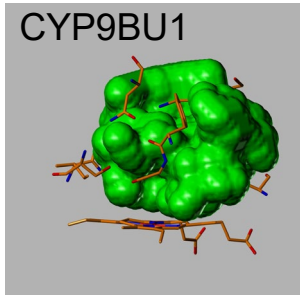


CYP9Q4

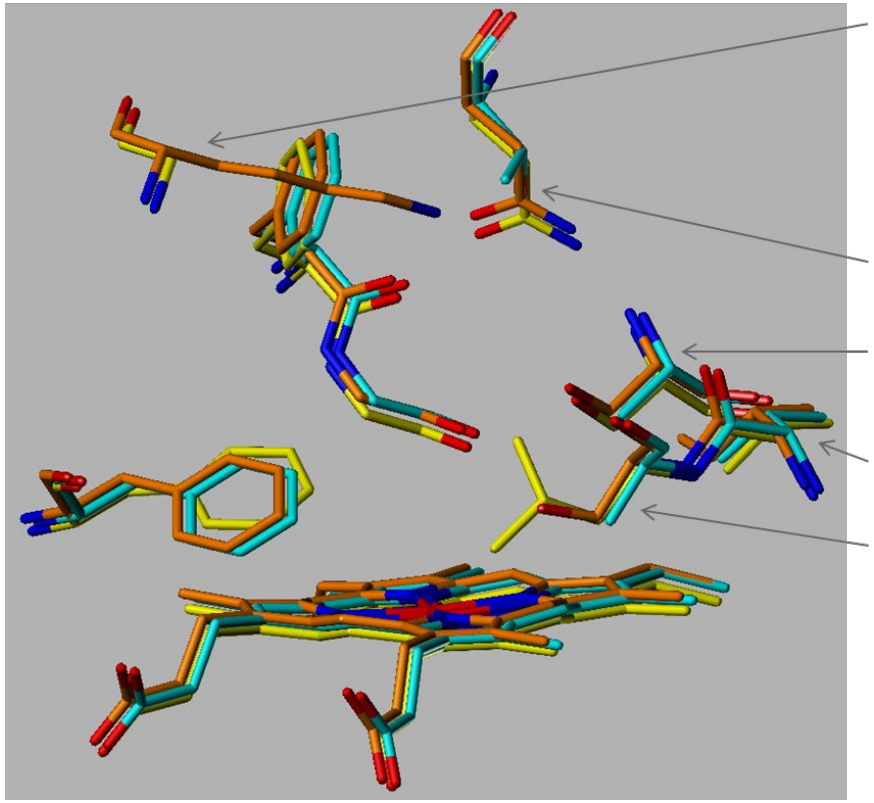




CYP9BU1



Q3 / Q4 / BU1



G111 / - / L111

The backbone loops were predicted significantly different in this region and so the equivalent residue is too far away from the catalytic oxygen

…

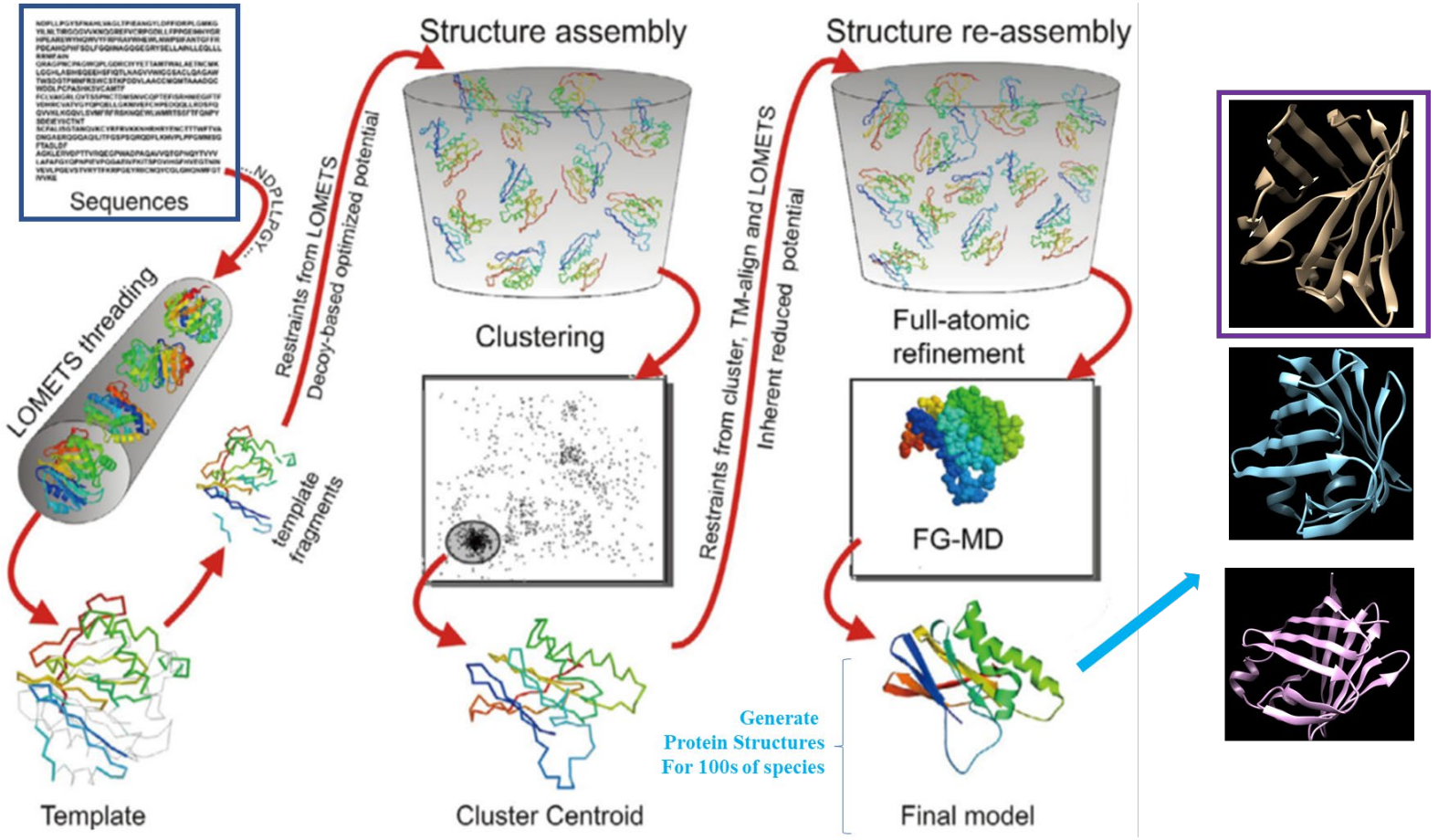
N214 / T218 / N214

S310 / A308 / S306

V371/ V369/ I367

L372 / A370 / S368

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



Models created for each species

- Apis mellifera
  - CYP9Q3 with restraint 1TQN (ITASSER)
  - CYP9Q3 with no restraint (ITASSER)
  - CYP9Q3 with no restraint (Phyre2)
- Bombus terrestris
  - CYP9Q4 with restraint 1TQN (ITASSER)
  - CYP9Q4 with no restraint (ITASSER)
  - CYP9Q4 with no restraint (Phyre2)
- Osmia bicornis
  - CYP9BU1 with restraint 1TQN (ITASSER)
  - CYP9BU1 with no restraint (ITASSER)
  - CYP9BU1 with no restraint (Phyre2)

Comparing model scores

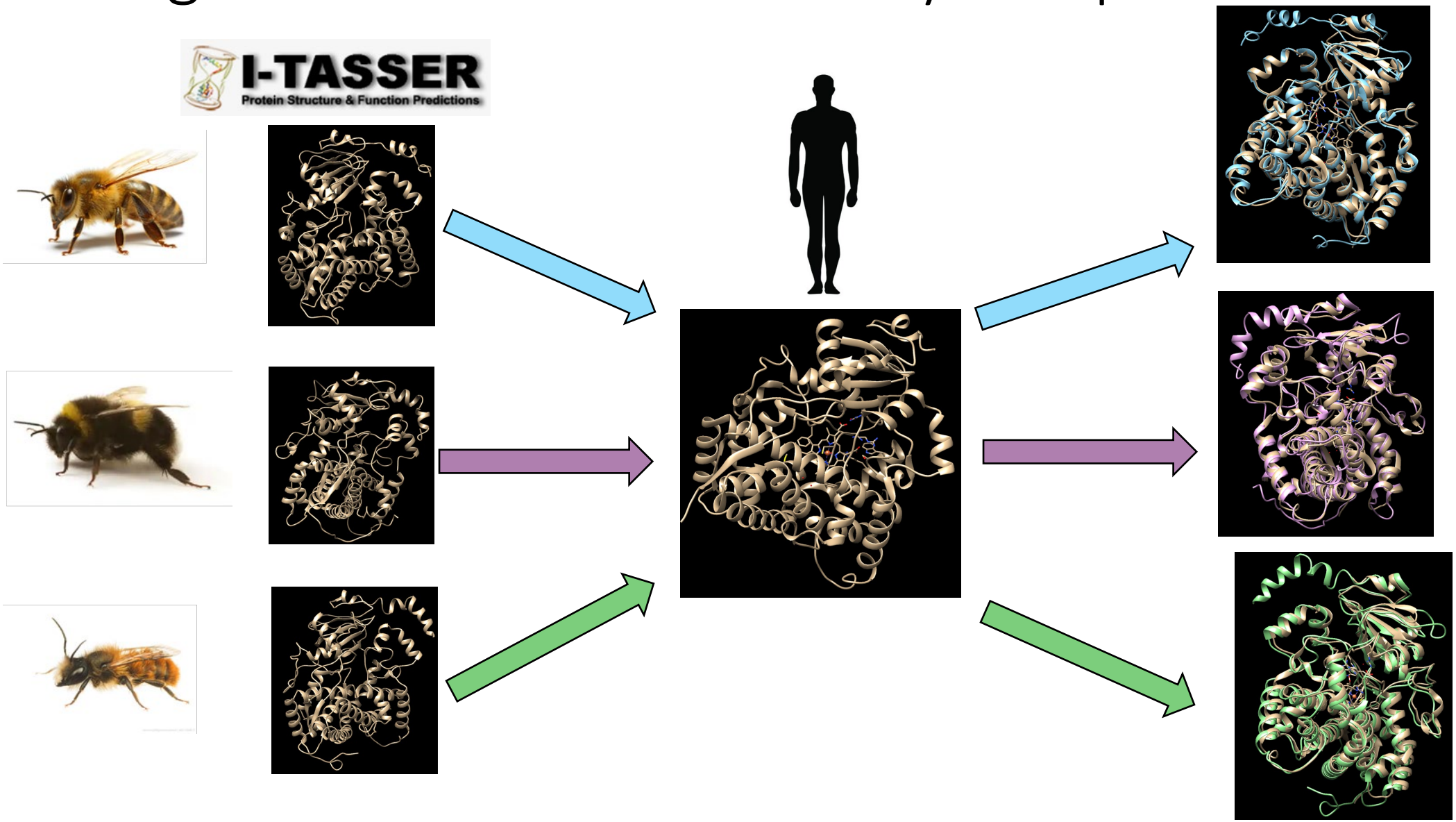
ITASSER Scoring text

- Correlation coefficient of C-score 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

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	CYP9BU1	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	CYP9BU1	None	Osmia bicornis	1 of 5	0.54	0.79+-0.09	6.2+-3.8	1042	0.272

United States Environmental Protection Agency

Utilizing UCSF Chimera to visually compare models

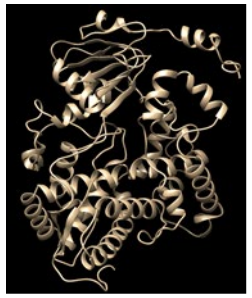


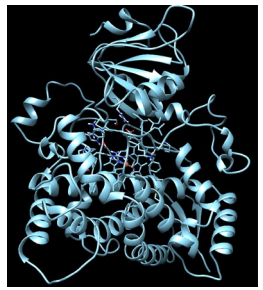
EPA United States Environmental Protection Agency

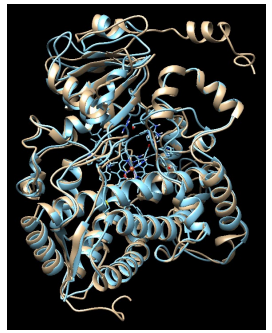
Utilizing UCSF Chimera to visually compare models

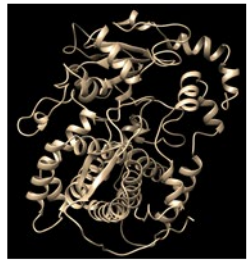
I-TASSER Protein Structure & Function Predictions

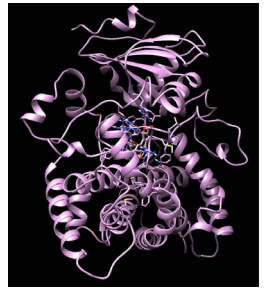
BAYER

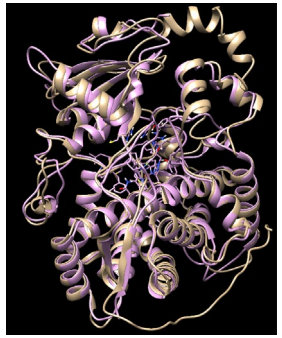


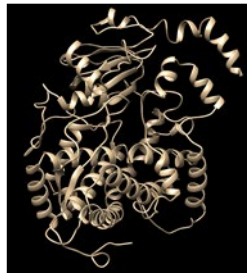


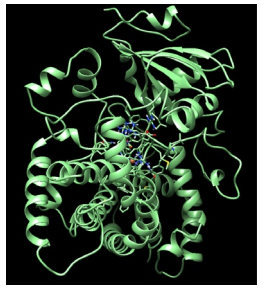


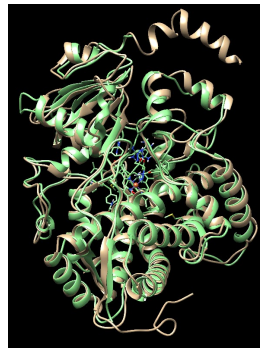












EPA United States Environmental Protection Agency

Completed work recap and next possible steps…

- Minimize ITASSER structures to match Bayer .min structures
- Look at superimposed positions for critical amino acids
- Dock flupyradifurone and thiacloprid to structures and determine docking scores
- Point mutate Bayer models and analyze docking scores of flupyradifurone and thiacloprid
- Create homology models for other bee species of interest
  - Possible step