Validating property and metabolite predictions for some novel antimalarial compounds

Robert D. Clark

Simulations Plus, Inc.

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Is That All There Is?

|< < Prev Random Next > >|

Permanent link to this comic: https://xkcd.com/1838/
Image URL (for hotlinking/embedding): https://imgs.xkcd.com/comics/machine_learning.png

A Prospective Validation Drug Design Project …

Simulations Plus Completes Molecule Design Phase,
Issues RFQs for Molecule Synthesis

Company Uses Proprietary Software Tools to Design
New Drug-like Molecules for Resistant Malaria Parasite

LANCASTER, CA, September 13, 2011

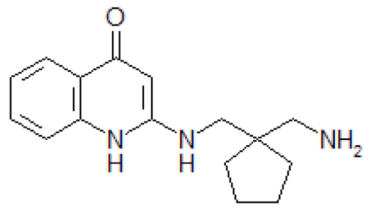
Simulations Plus, Inc. (NASDAQ: SLP), a leading provider of consulting services and software for pharmaceutical discovery and development, today announced that it has completed the molecule design phase of its malaria NCE (new chemical entity) project, and has issued requests for quotes to synthesize up to twelve compounds for testing…

…that Worked Pretty Well

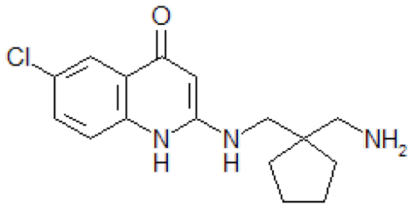
Compound	Predicted K_i (μM)[a]	XC_{50} (μM)[b,c] 3d7(-)	XC_{50} (μM)[b,c] Dd2(+)	Resistance Ratio
SLP0007	0.049	10.0	46	4.6
SLP0008	0.051	1.61	6.4	3.9
SLP0004	0.023	0.55	2.3	4.1
SLP0010	0.037	0.37	1.78	4.8
SLP0005	0.037	0.30	1.47	5.0
SLP0003	0.025	0.106	0.21	2.0
SLP0006	0.038	0.037	0.24	6.6
CID 44534046[d]	0.112	0.89	4.6	5.2
CID 44535189[d]	0.077	0.85	8.6	10.1

[a] Predicted inhibition constant for soluble dihydroorotate dehydrogenase (DHODH) from *Plasmodium falciparum*.
[b] Concentration required to reduce parasite growth rate by 50%.
[c] (-) and (+) denote chloroquine-susceptible and -resistant strains, respectively.
[d] Pubchem IDs for the most active analogs from the GSK data set.

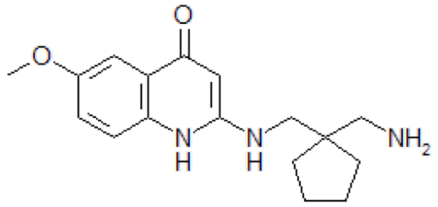
Structures for the Seven Leads Synthesized



SLP0001: 17 µM

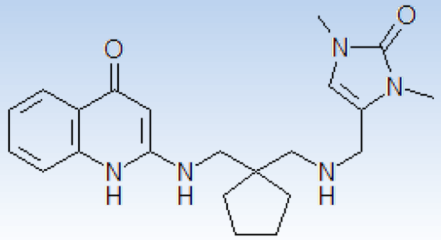


SLP0002: 0.73 µM

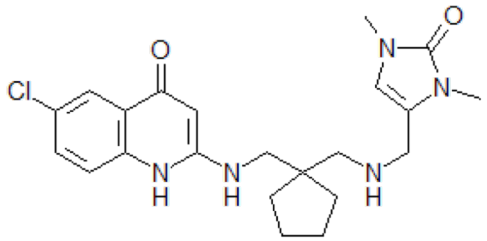


SLP0009: 2.3 µM

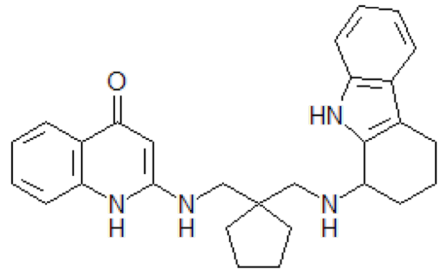
intermediates



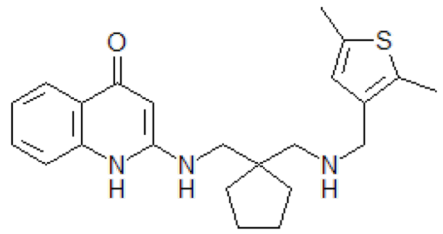
SLP0007: 10.0 µM



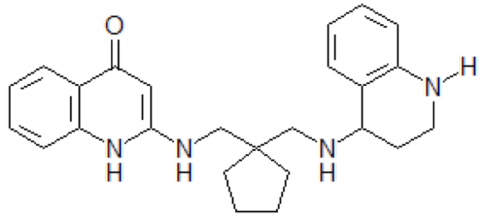
SLP0008: 1.61 µM



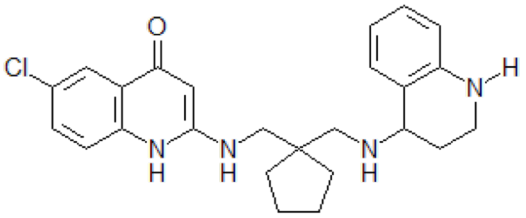
SLP0003: 0.106 µM



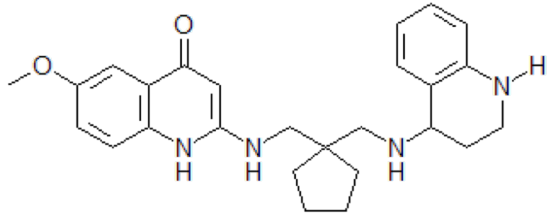
SLP0005: 0.30 µM



SLP0004: 0.55 µM



SLP0006: 0.037 µM



SLP0010: 0.37 µM

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

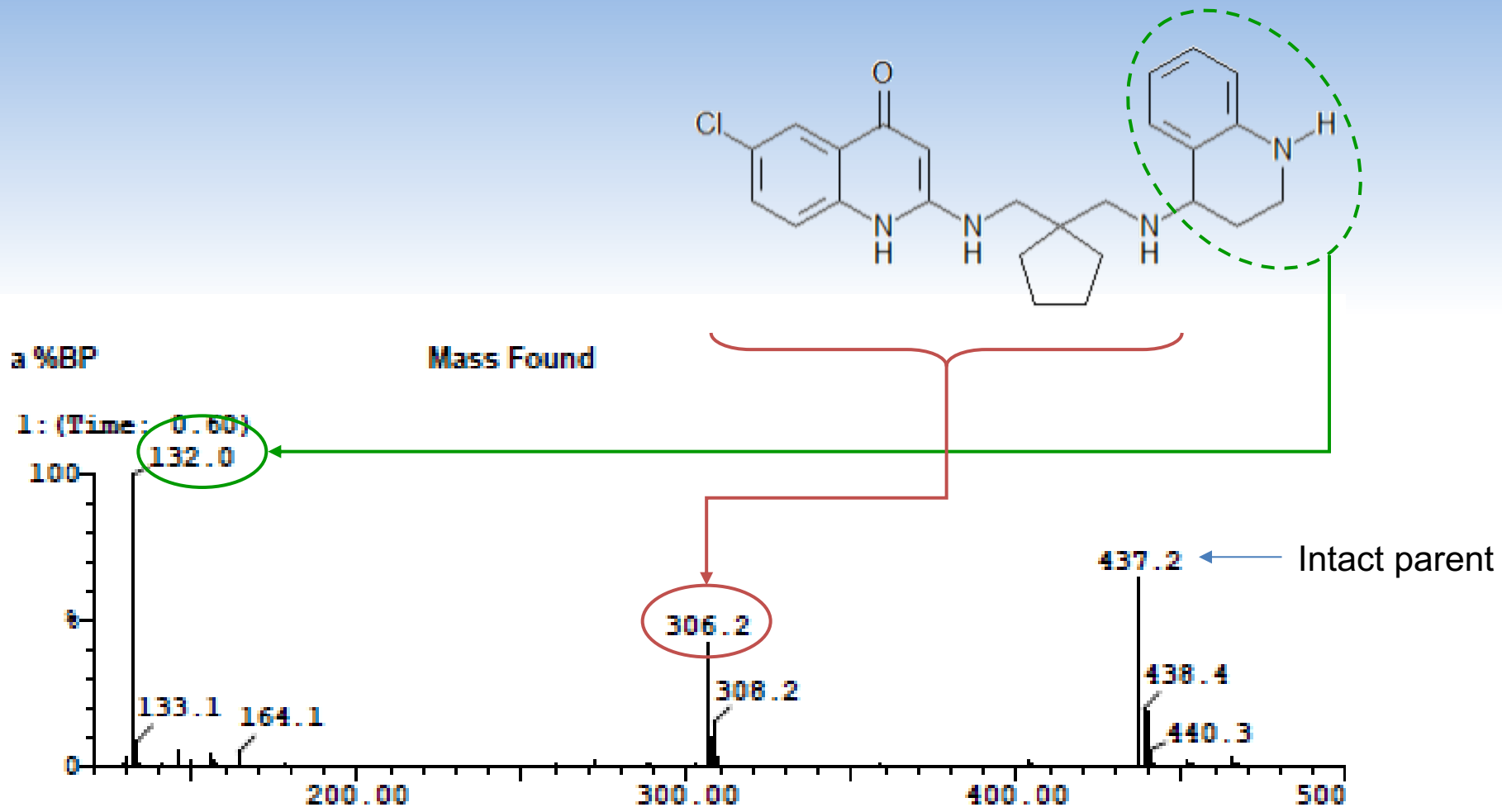
Rates of CYP Metabolism *in vitro*

CYP	Property	SLP0005	SLP0003	SLP0004	SLP0006	
1A2	Substrate?	No (60%)[a]	Yes (48%)	No (41%)	Yes (43%)	
	$CL_{1\mu M}$	(6.9) / 10[b,c,d]	4.9 / 9.3	(5.5) / 7.3	10.7 / ND	← RMSE: 2-fold
2C9	Substrate?	No (89%)	No (77%)	No (96%)	No (76%)	
	$CL_{1\mu M}$	(137) / 4.1	(231) / ND	(214) / 1.8	(245) / ND	
2C19	Substrate?	No (94%)	No (93%)	No (87%)	No (97%)	
	$CL_{1\mu M}$	(532) / 14.6	(115) / 3.8	(642) / 3.5	(571) / 2.2	
2D6	Substrate?	Yes (82%)	Yes (63%)	Yes (63%)	Yes (82%)	
	$CL_{1\mu M}$	330 / **31**[e]	381 / **10.6**[e]	420 / **30**[e]	447 / **18**[e]	
3A4	Substrate?	Yes (92%)	Yes (98%)	Yes (98%)	Yes (98%)	
	$CL_{1\mu M}$	41 / 256	135 / 179	155 / 298	84 / 263	← RMSE: 3-fold
HLM	CL	171 / 220	138 / 120	117 / 230	250 / 620	← RMSE: 2-fold

[a] Prediction (confidence)
[b] Predicted / observed clearance at 1 µM
[c] µL/min/mg HLM protein
[d] Parentheses indicate predicted clearance *were the compound to be a substrate*
[e] **Autoinhibition expected**

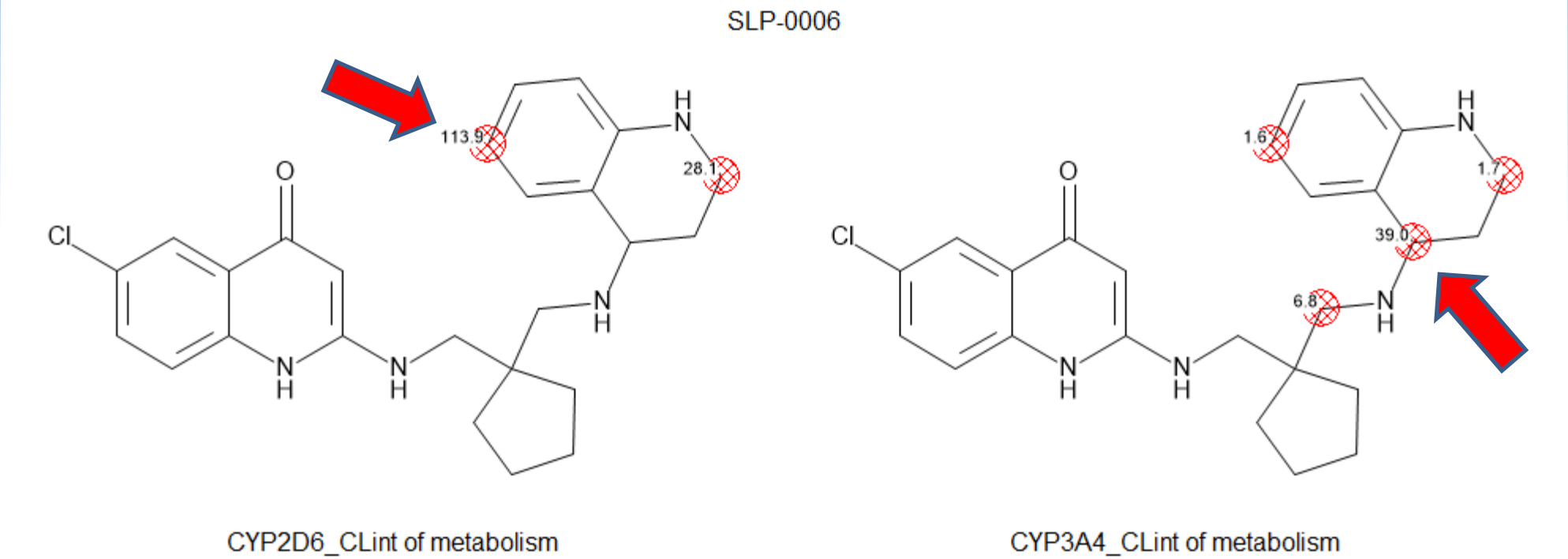
SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Parent Fragmentation: Mass Spectrum for SLP-0006



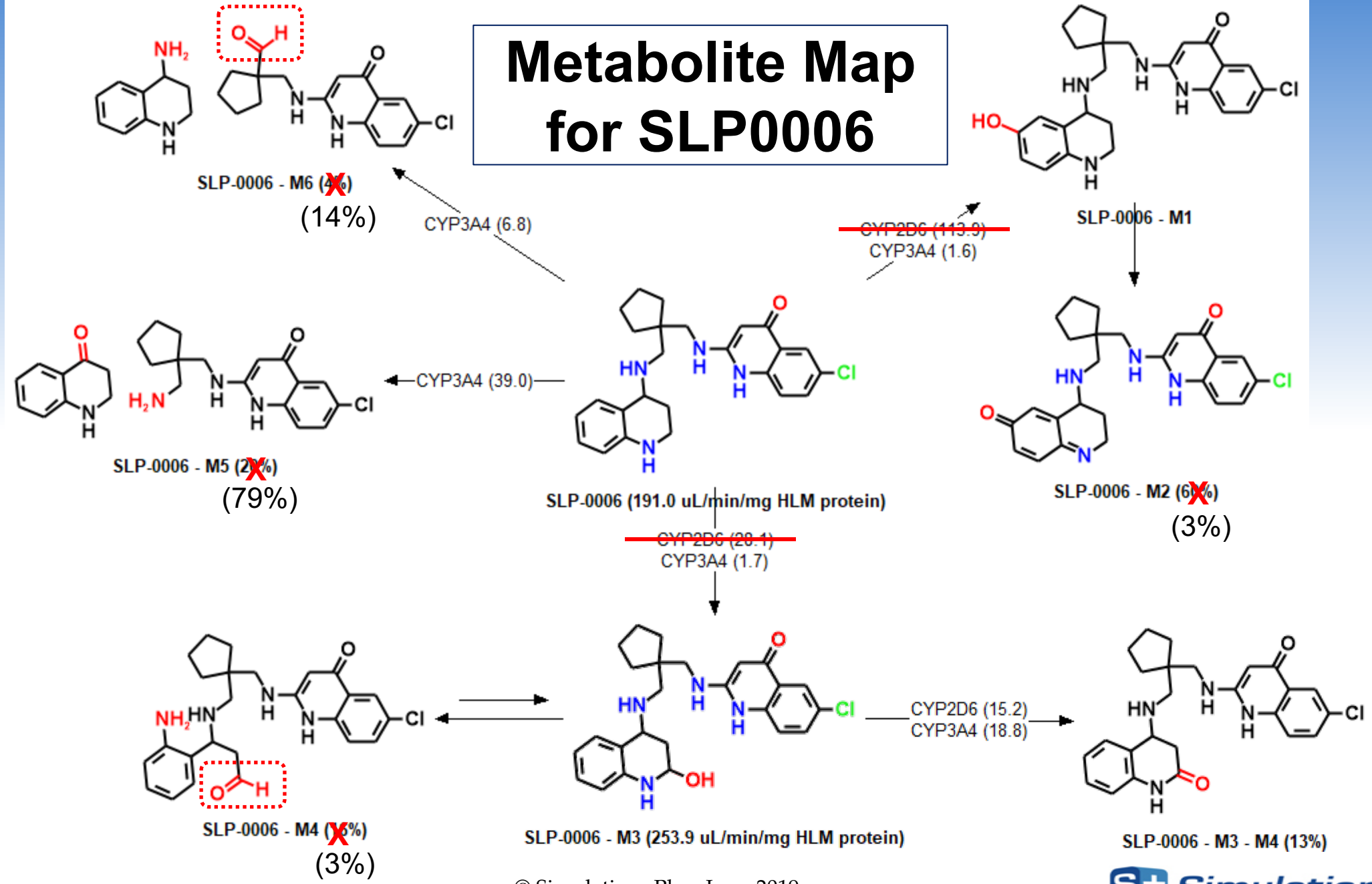
SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Predicted Metabolic Hot Spots for SLP0006



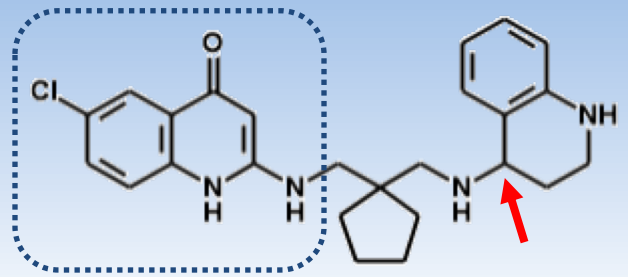
SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Metabolite Map for SLP0006



SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

HPLC-MS Data for SLP0006

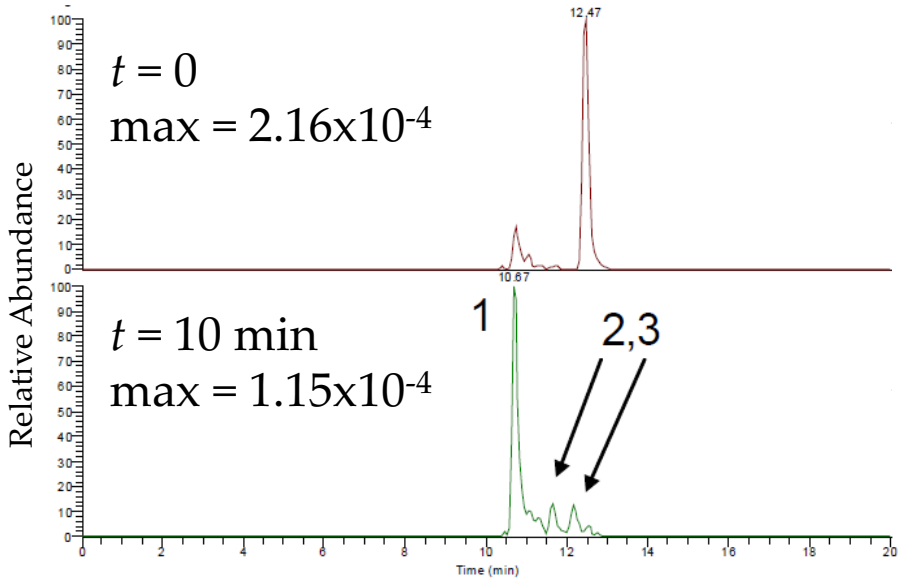


α

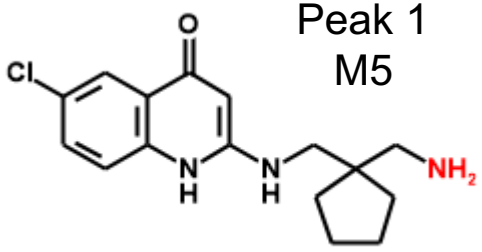
195.6

Table I. Results of detection of metabolites of test compound KXN-4783

Analyte	exact mass (M+H)	Detected peaks at RT (min)	MS/MS data collected**
KXN-4783	437.2103	12.47	NO
Parent+O	453.2052	NF	NO
Parent+2O	469.2001	NF	NO
Parent+2O-H2O	451.1895	NF	NO
Met1	306.1368	10.67*; 11.7; 12.5*	YES
Met3	259.1805	NF	NO
Met3+O	275.1754	NF	NO
alpha	195.0320	10.67*	NO

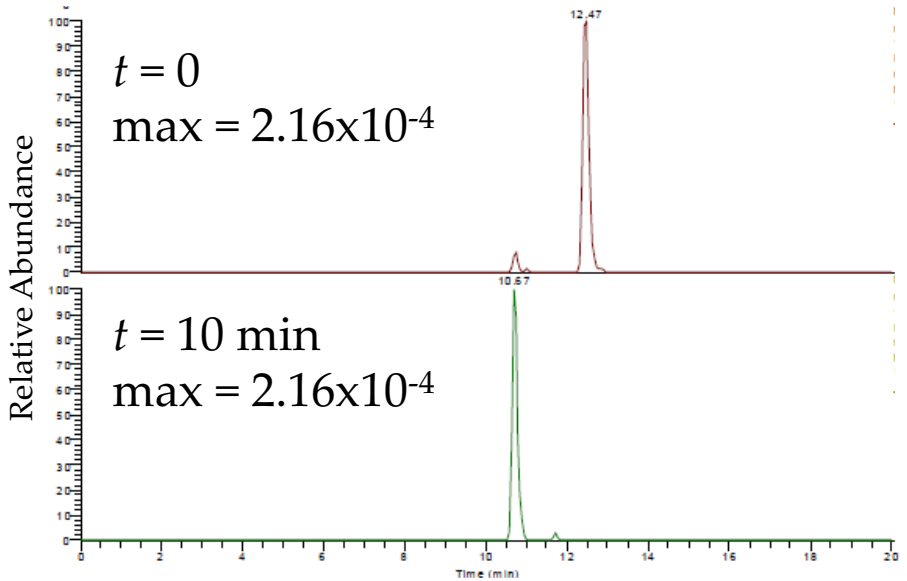


detection @ mass 306.14



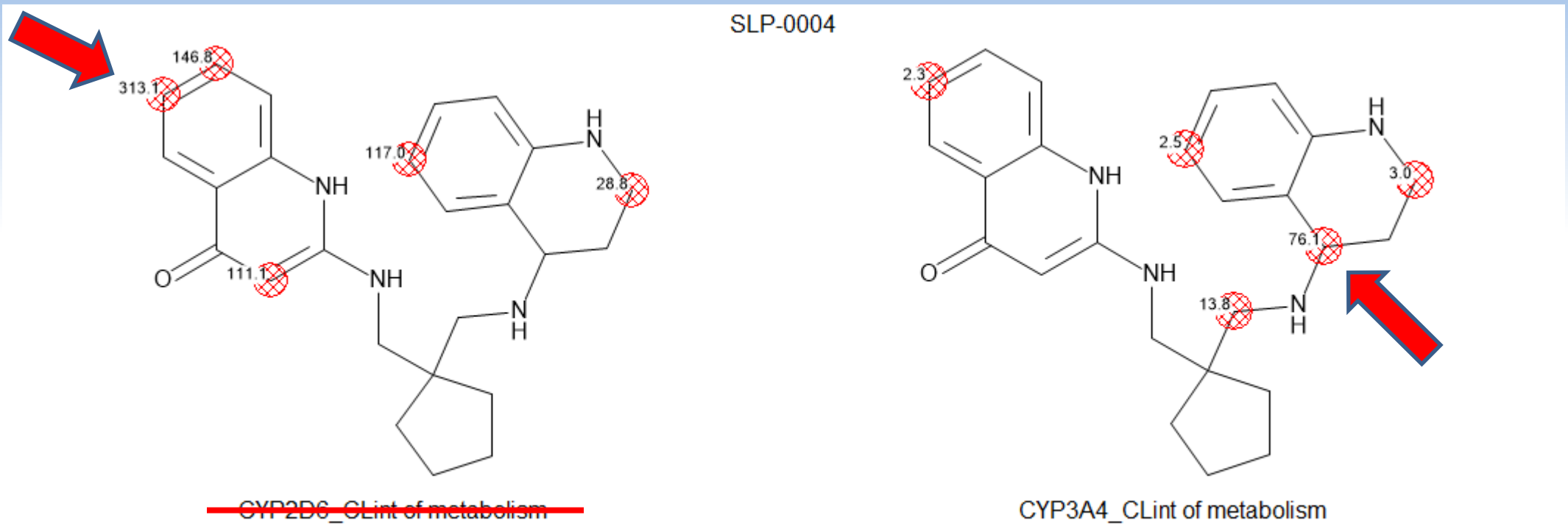
Peak 1
M5

Predicted: 79%
Observed: 63%



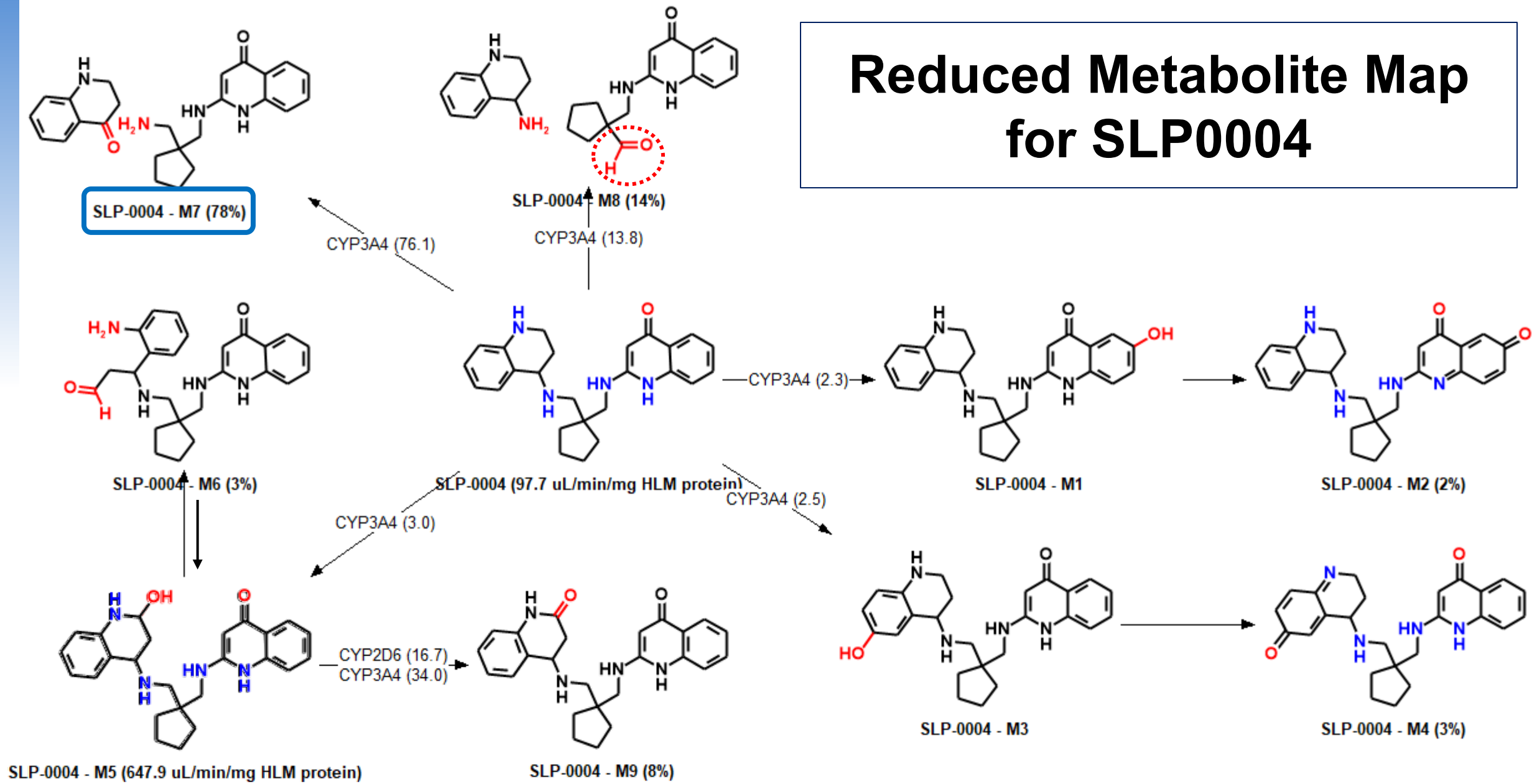
detection @ mass 195.03

Predicted Metabolic Hot Spots for SLP0004



SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

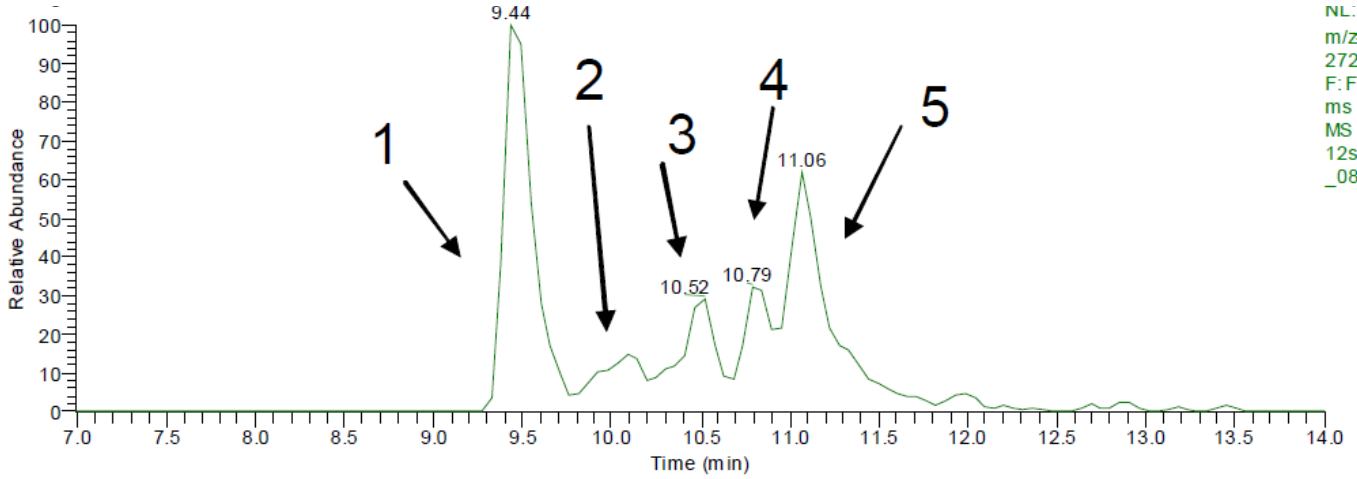
Reduced Metabolite Map for SLP0004

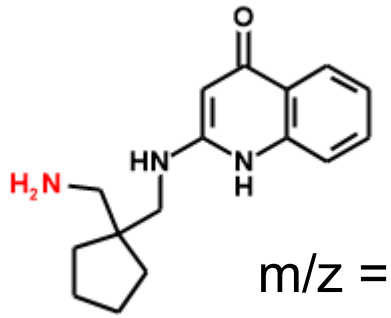


SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

In vitro Metabolite Yield for SLP-0004

RT(min)	Mass Peaks	Tentative ID	Obs. Yield	Pred. Yield
9.44*	272, 161	(1) SLP0002: M7	22%	59%
10.1	419, 272	(2) aldehyde: M6	8%	
10.5	419, 272	(3) 2'-hydroxyl: M5	8%	8%
10.8	417, 272	(4) lactam (sec.): M9	4%	
11.05*	403, 272, 161	(5) parent SLP0004	25%	----
			67%	





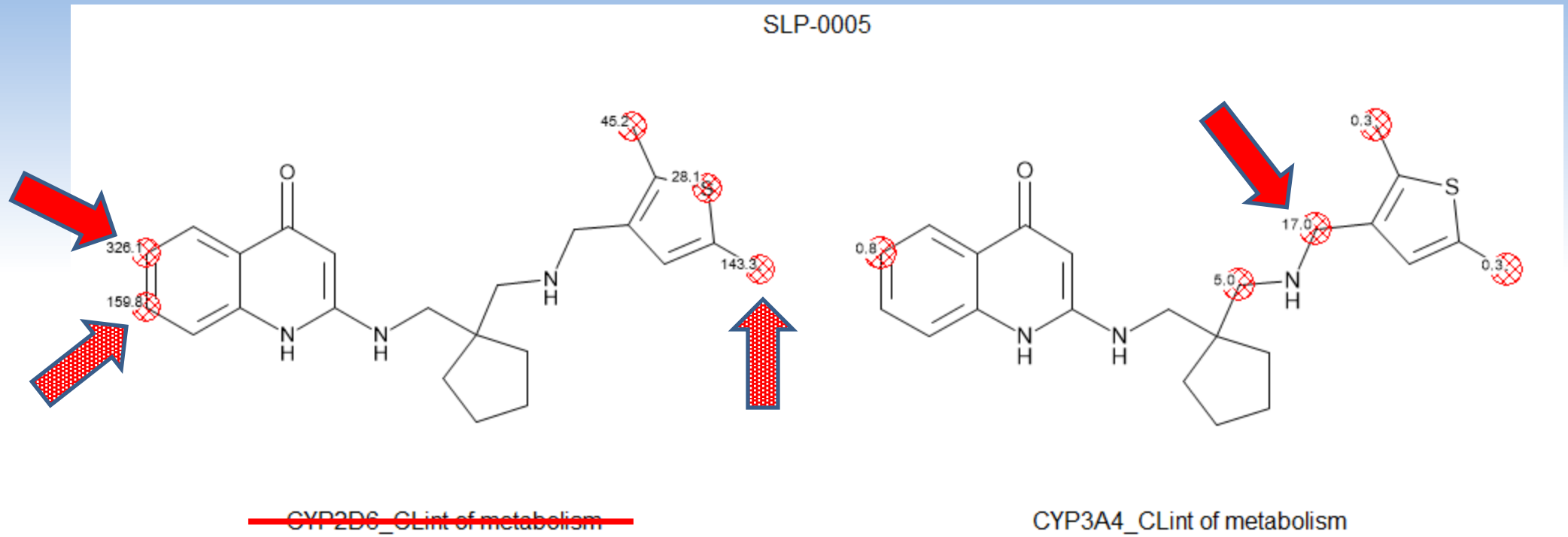
m/z = 272.1

SLP-0004 - M7 (78%)

Observed: 52% of metabolites

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Predicted Metabolic Hot Spots for SLP0005

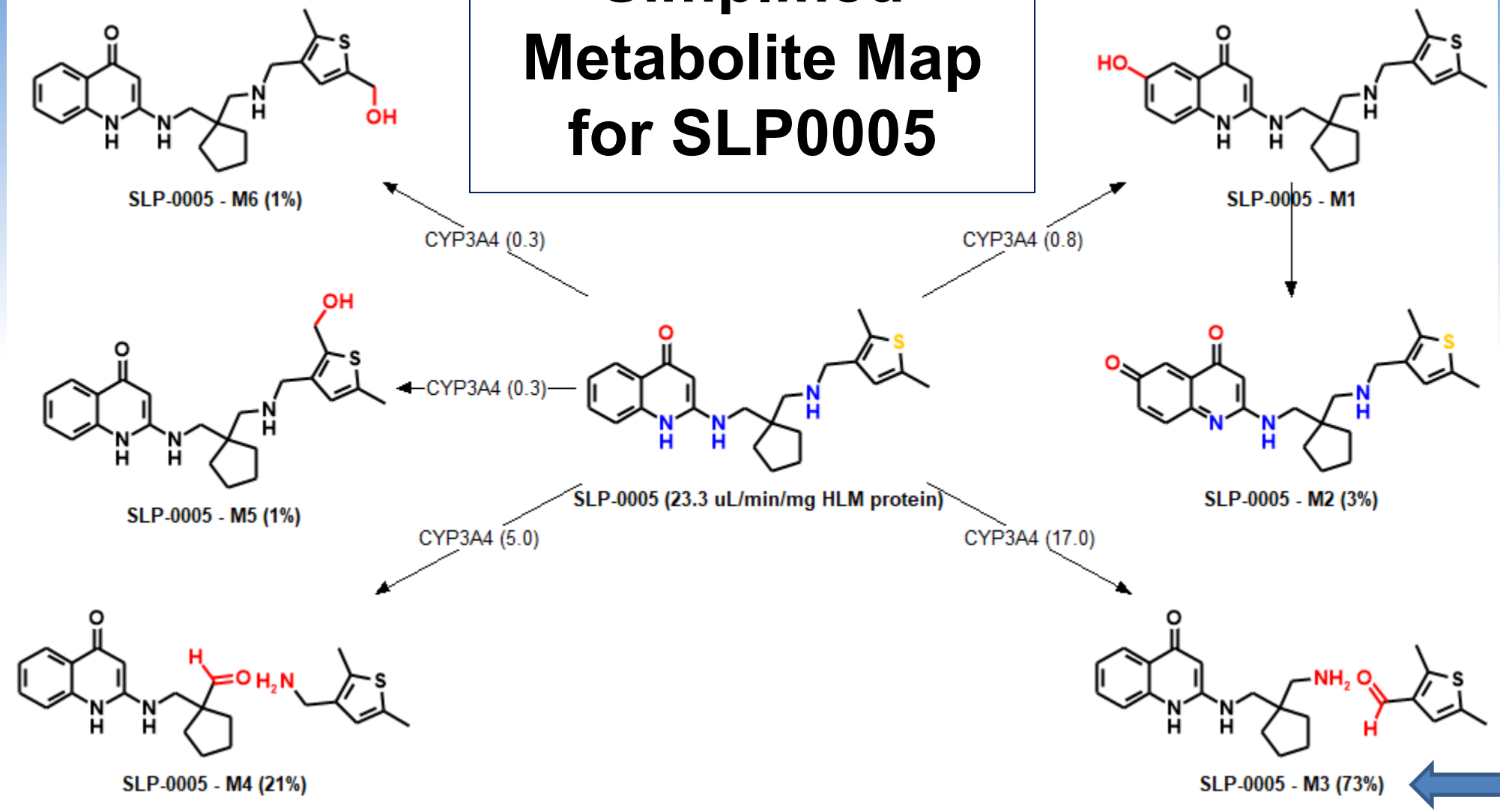


CYP2D6_CLint of metabolism

CYP3A4_CLint of metabolism

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

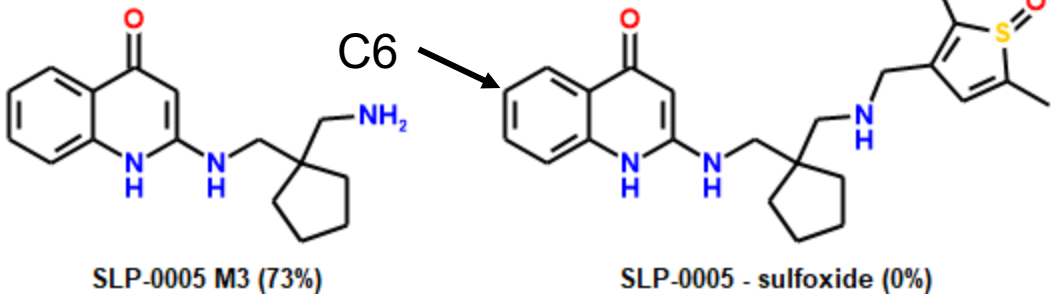
Simplified Metabolite Map for SLP0005



SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

In vitro Metabolite Yields for SLP-0005

RT (min)	Mass Peaks	Tentative ID	Obs. Yield	Pred. Yield	S+logP
9.6	272	SLP0001: M3	<4%	48%	2.26
9.8	428	6-OH sulfoxide	1%	----	2.40
10.5	412, 272	sulfoxide	18%	----	2.95
11.0	272	2'-CH2OH: M5	13%	0.7%	3.57
11.7	412, 272	5'-CH2OH: M6	≥5%	0.7%	3.57
12.2	396, 272	SLP0005	35%	----	4.71

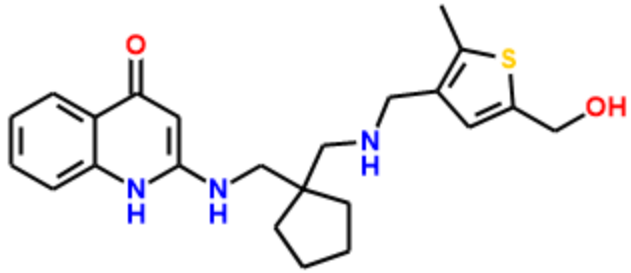


SLP-0005 M3 (73%)

Obs: <6% of metabolites

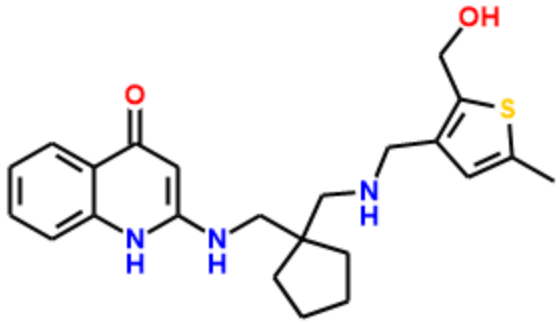
SLP-0005 - sulfoxide (0%)

Obs: > 28% of metabolites



SLP-0005 - M6 (1%)

Obs: > 8% of metabolites

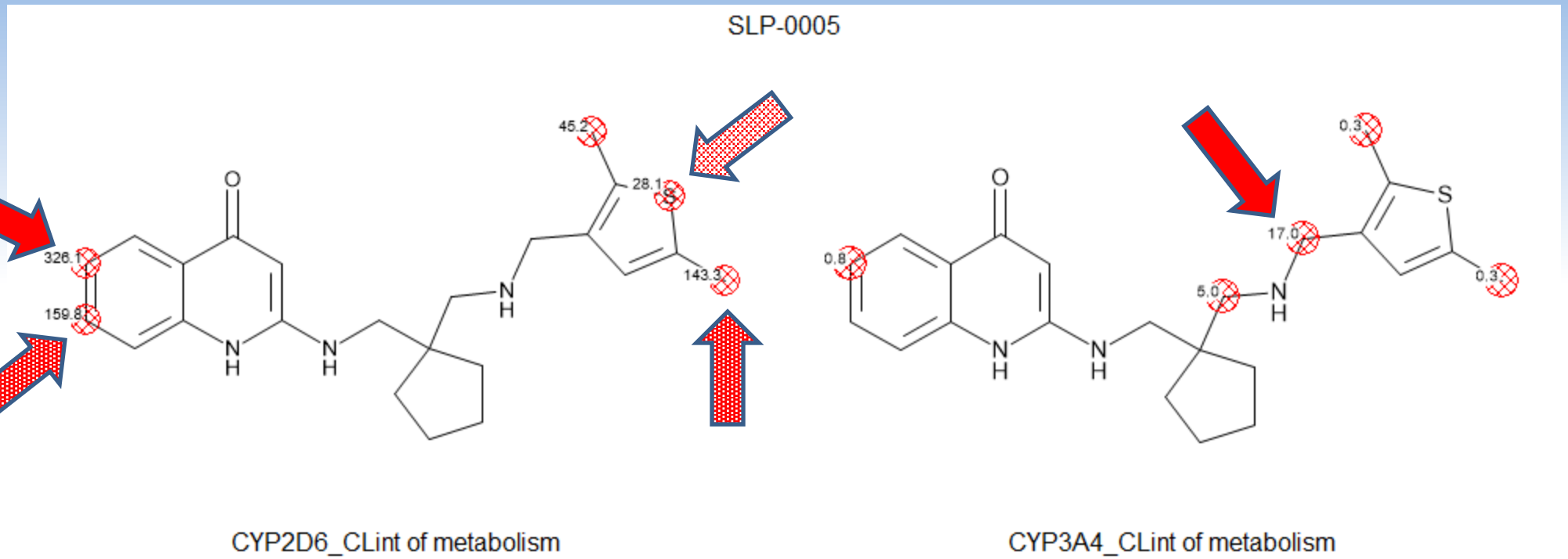


SLP-0005 - M5 (1%)

Obs: 20% of metabolites

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Predicted Metabolic Hot Spots for SLP0005

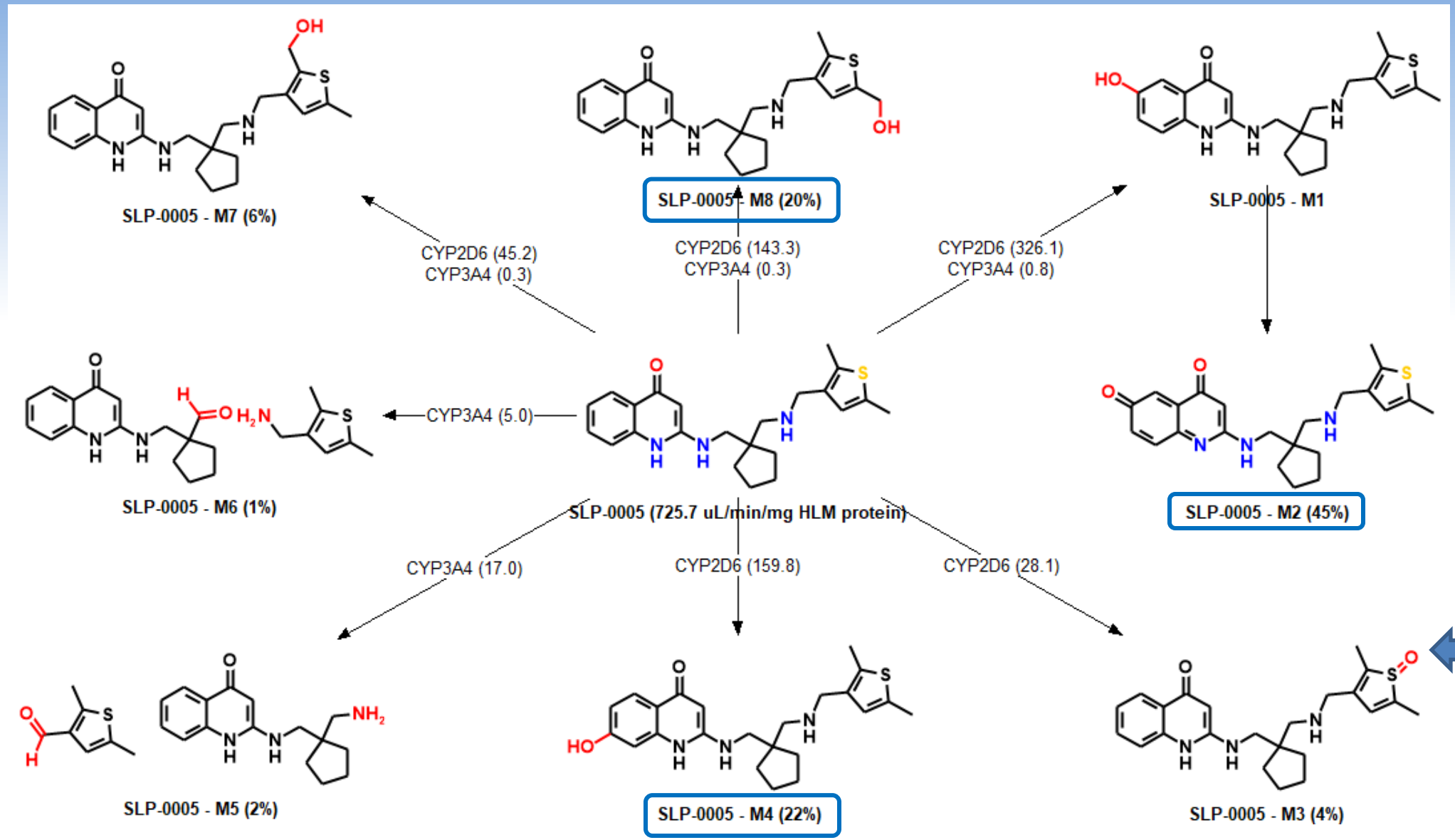


CYP2D6_CLint of metabolism

CYP3A4_CLint of metabolism

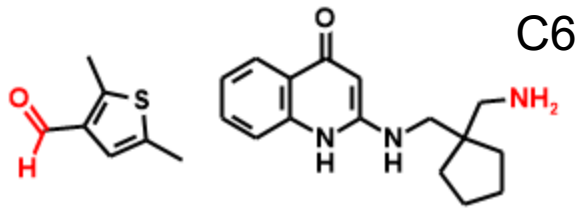
SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Full Metabolite Map for SLP0005



In vitro Metabolite Yields for SLP-0005

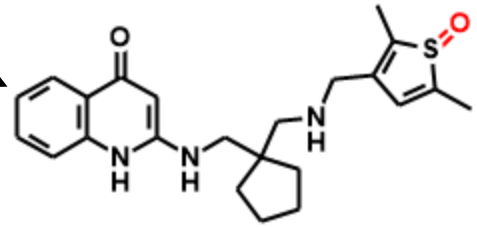
RT (min)	Mass Peaks	Tentative ID	Obs. Yield	Pred. Yield	S+logP
9.6	272	SLP0001: M5	<4%	1.3%	2.26
9.8	428	6-OH sulfoxide: M1	1%	3%	2.40
10.5	412, 272	sulfoxide: M3	18%		2.95
11.0	272	5'-CH_2OH: M8	13%	13%	3.57
11.7	412, 272	2'-CH_2OH: M7	≥5%	5%	3.57
12.2	396, 272	SLP0005	35%	----	4.71



SLP-0005 - M5 (2%)

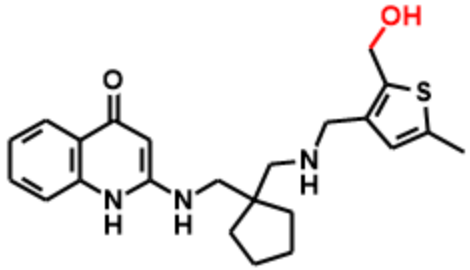
Obs: <6% of metabolites

C6



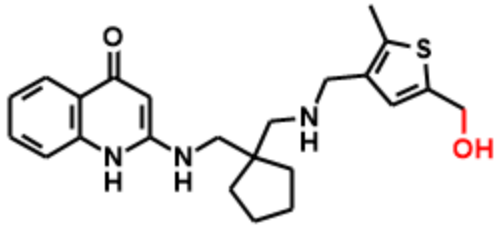
SLP-0005 - M3 (4%)

Obs: >32% of metabolites



SLP-0005 - M7 (6%)

Obs: >7.5% of metabolites

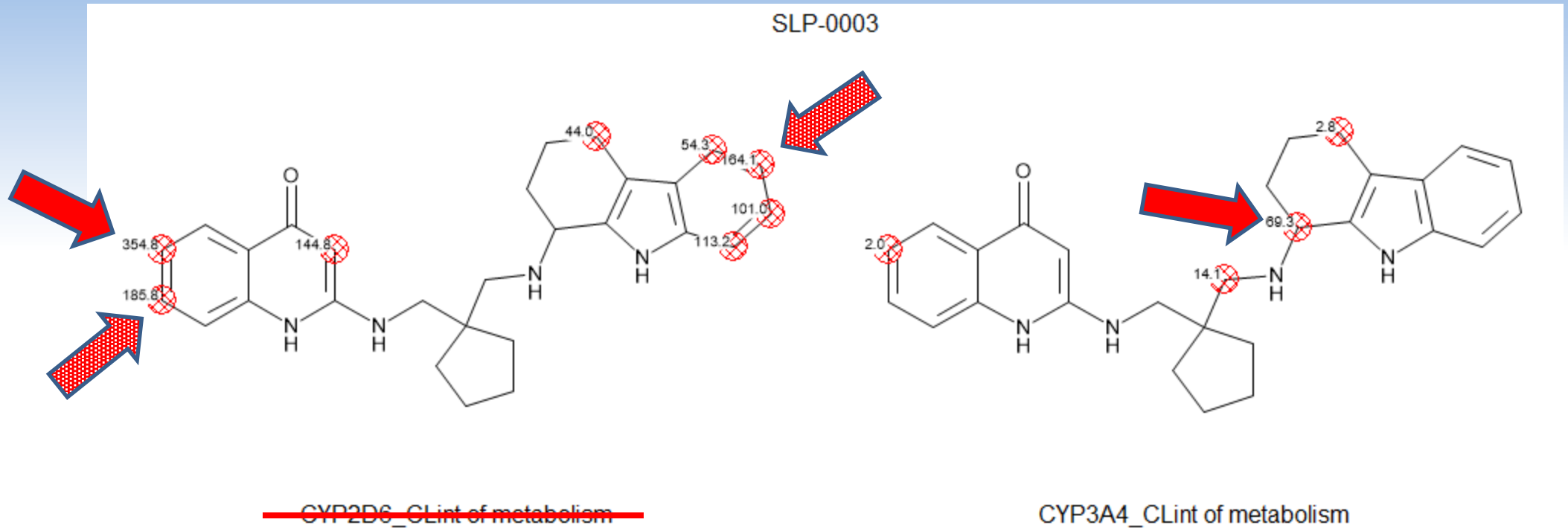


SLP-0005 - M8 (20%)

Obs: 20% of metabolites

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

Predicted Metabolic Hot Spots for SLP0003

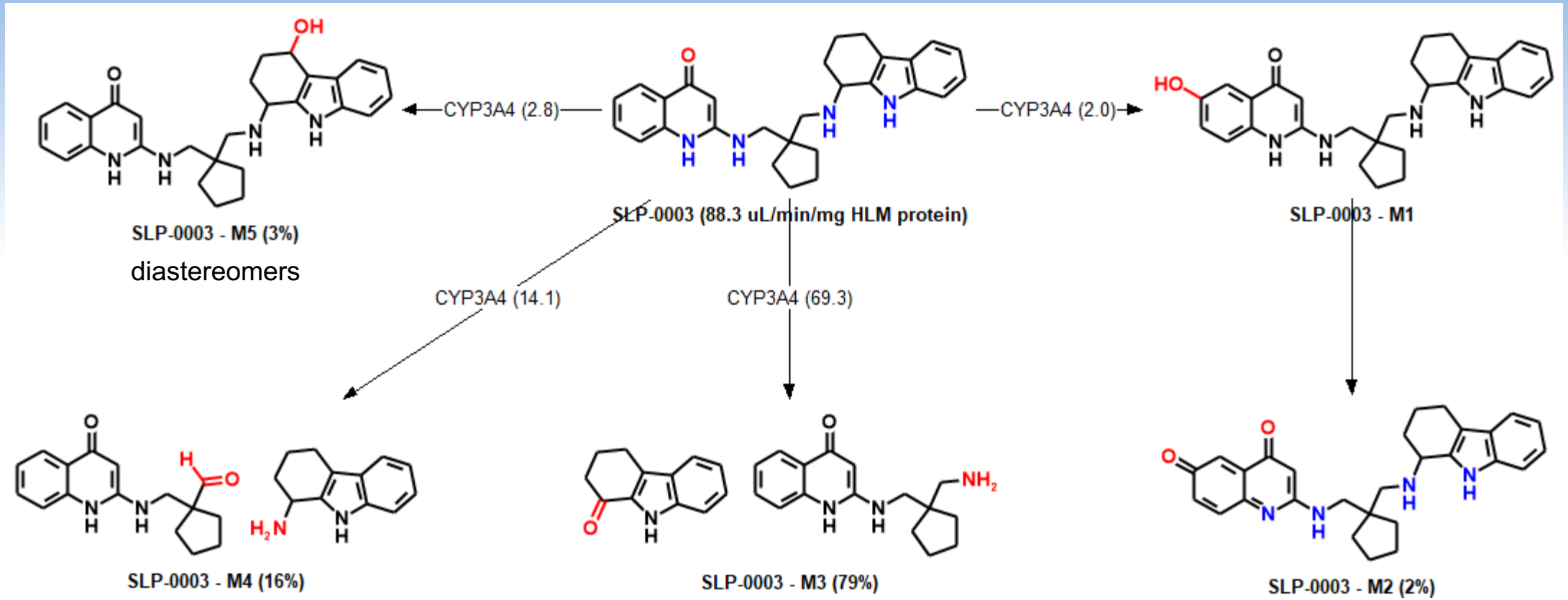


CYP2D6_CLint of metabolism

CYP3A4_CLint of metabolism

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

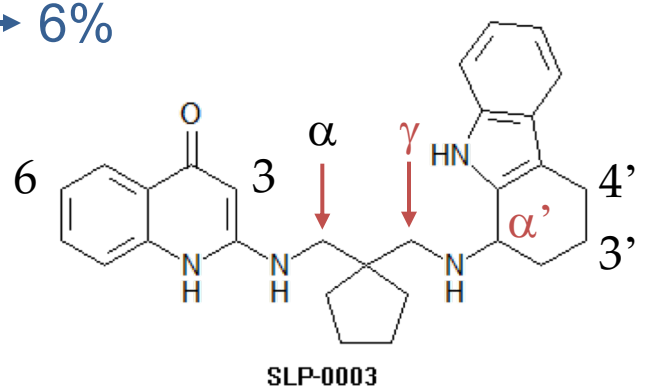
Simplified Metabolite Map for SLP0003



SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS

In vitro Metabolite Yields for SLP-0003*

RT(min)	Mass Peaks	Tentative ID	Observed	Predicted
9.5	272	SLP0001: M3	7%	42% (32% of metabolites)
10.4	473	3,4'-dihydroxy[ab]	2%	
10.9	473	6,4'-dihydroxy[ab]	4%	
11.2	473	6,4'-dihydroxy[ab]	2%	
11.5	457, 272	4'-hydroxy[a]: M5	4%	9%[c]
11.7	271	γ cleavage: M4	6%	6%
11.9	457	4'-hydroxy[a]: M5	3%	
12.2	457	4'-ketone[b]	3.5%	
13.3	272	3',4'-dehydro[b]	4.5%	
14.2	441, 272	SLP0003 (parent)	47%	



SLP-0003

* Based on a simplified metabolite map that only takes CYP3A4 oxidation into account

[a] Suspected diastereomers [b] Secondary metabolites of M5 [c] Sum of M5 and its metabolites

Metabolite Assignment Considerations & Caveats

- One compound can (and often does) give rise to two or more mass spectral (MS) peaks - *but* there should only be one HPLC peak per compound,
- Some parent ions fragment too quickly to show up in the MS
- Yield calculations are based on fragment ion abundances and assume that all fragments have been accounted for.
- Some structurally similar metabolites may not be separated under the chromatographic conditions used
- Characteristic fragments convey information about site(s) of oxidation by their presence *or* absence, but absence of evidence is not *proof* of absence
- Several isomer assignments are based on the presumed monotonic relationship between HPLC retention time and S+logP

Conclusions

- Metabolite maps can get very complicated very quickly.
- The potential for autoinhibition can complicate things.
  - the combination of individual K_m predictions and inhibitor classification models can help anticipate such effects.
- One of the values of models lies in giving an indication of when things are more complicated than one might expect.
  - And how to resolve ambiguities, i.e., by retesting at different concentrations.
- logP estimates can help rationalize HPLC profiles.
- Quantitating unusual oxidations – e.g., sulfur in thiophene – is hard

⇒The commercial substrate, inhibitor and kinetic models in ADMET Predictor 9.5™ were able to produce metabolic maps for four novel compounds that were *consistent with most of* the observed *in vitro* experimental results, especially in straightforward cases.

Cast of Supporting Characters

Simulations Plus, Inc.
- Michael Lawless
- Walter S. Woltosz
- Marvin Waldman
- David Miller
- Aleksandra Mikosz
- Robert Fraczkiewicz
- Dechuan Zhuang
- Adam Lee
- Jinhua Zhang
- Viera Lukacova
- Michael B. Bolger

Cognigen
- Denise Morris
- Ted Grasela

Kalexsyn
- Bob Gadwood
- Gary Chinigo

UC Riverside
- Karine Le Roch
- Jacques Prudhomme

Tres Cantos (GSK)
- Francisco Javier Gamo-Benito
- Maria Jose Lafuente-Monasterio
- Santiago Ferrer-Bazaga

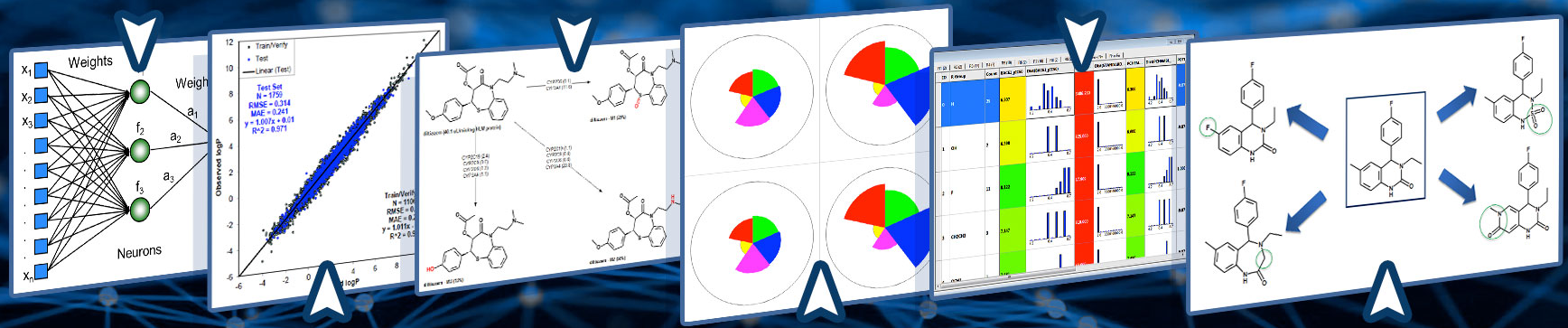
Medicines for Malaria Venture (MMV)
- Jeremy Burrows
- Xavier Ding

And thank you!

ADMET Predictor® 9.5

ADMET Property Estimation & Model Building

QSAR Model Building **CYP Metabolite Prediction** **R-Table Generation/Analysis**



>140 Predicted Properties **Data Visualization** ***De novo* Design**

bob@simulations-plus.com

SimulationsPlus
SCIENCE + SOFTWARE = SUCCESS