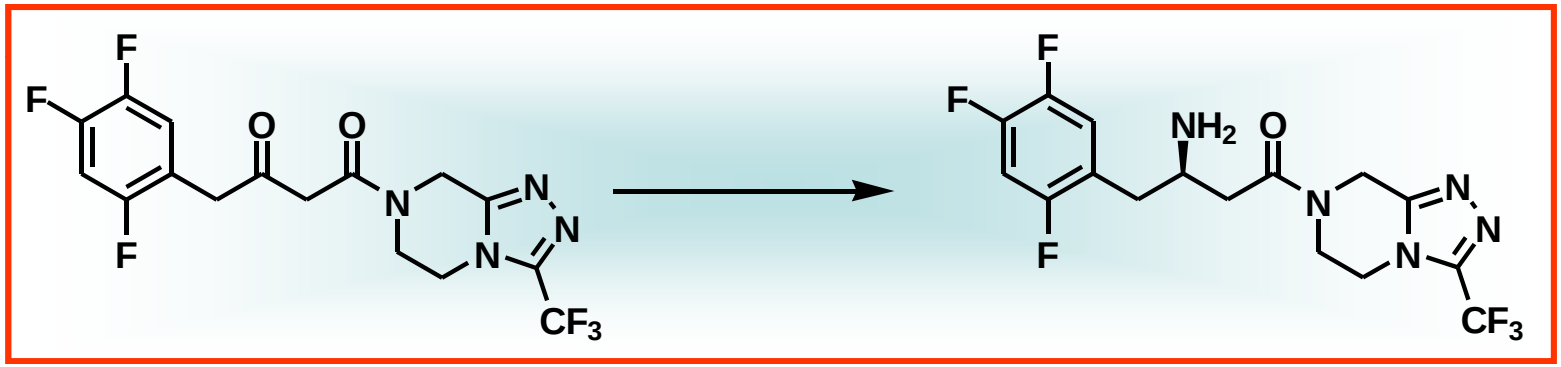
A Biocatalytic Manufacturing Route for Januvia®



Challenges in Catalysis III

Royal Society of Chemistry

Nov 2, 2011

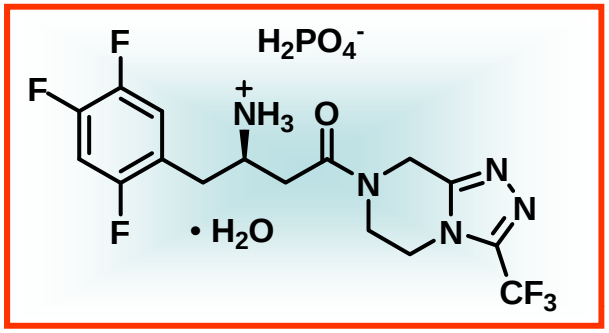
Jake Janey

Bristol-Myers Squibb

CODEXIS

MERCK

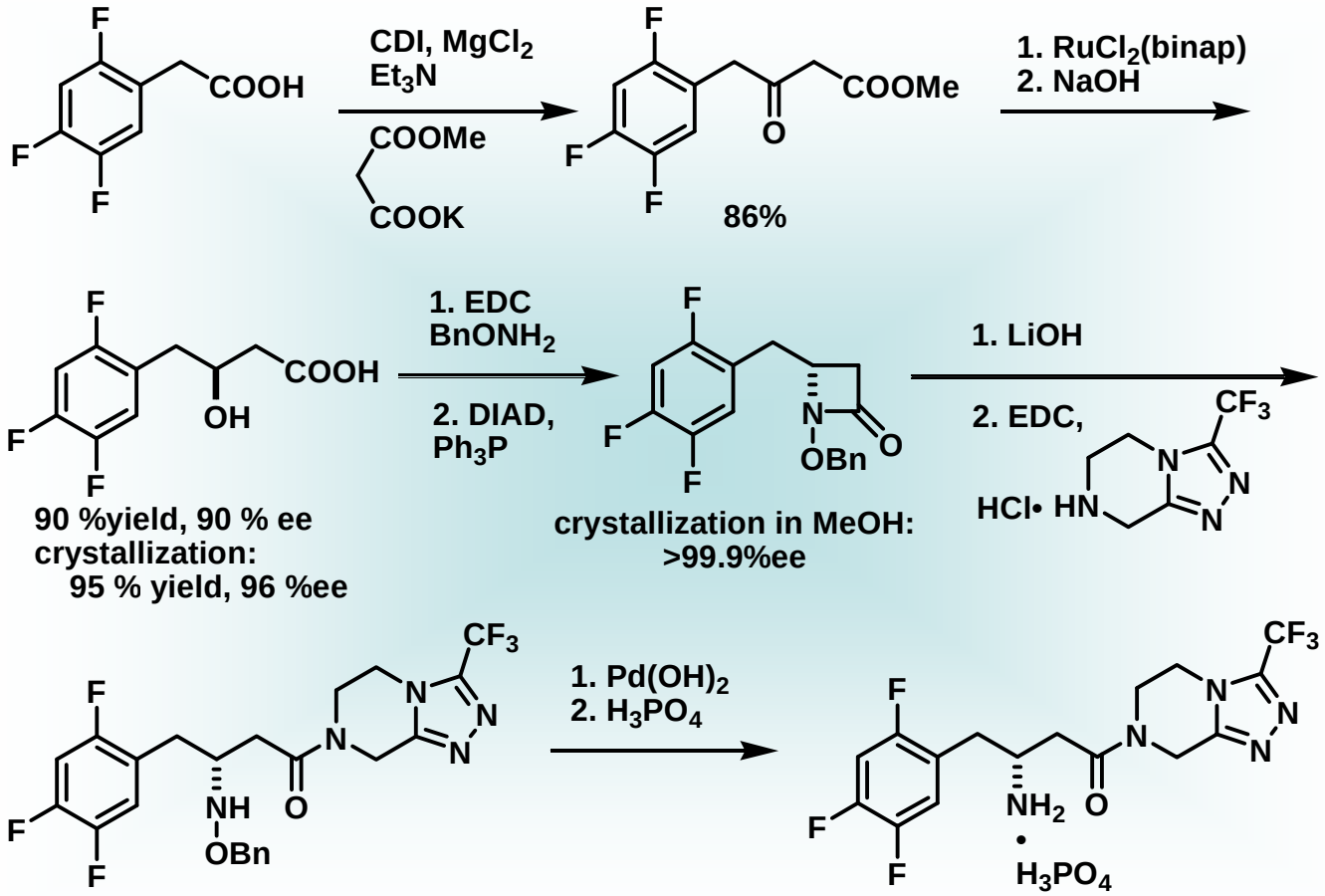
Sitagliptin: A DPPIV inhibitor



Active Ingredient in
Januvia®
Janumet®
Juvisync®

- A first in class dipeptidyl peptidase-4 inhibitor (DDP-4)
- Novel Mechanism for treatment of Type II diabetes
- Major advantages:
  - Oral rather than injectable
  - Unlikely to cause hypoglycemia

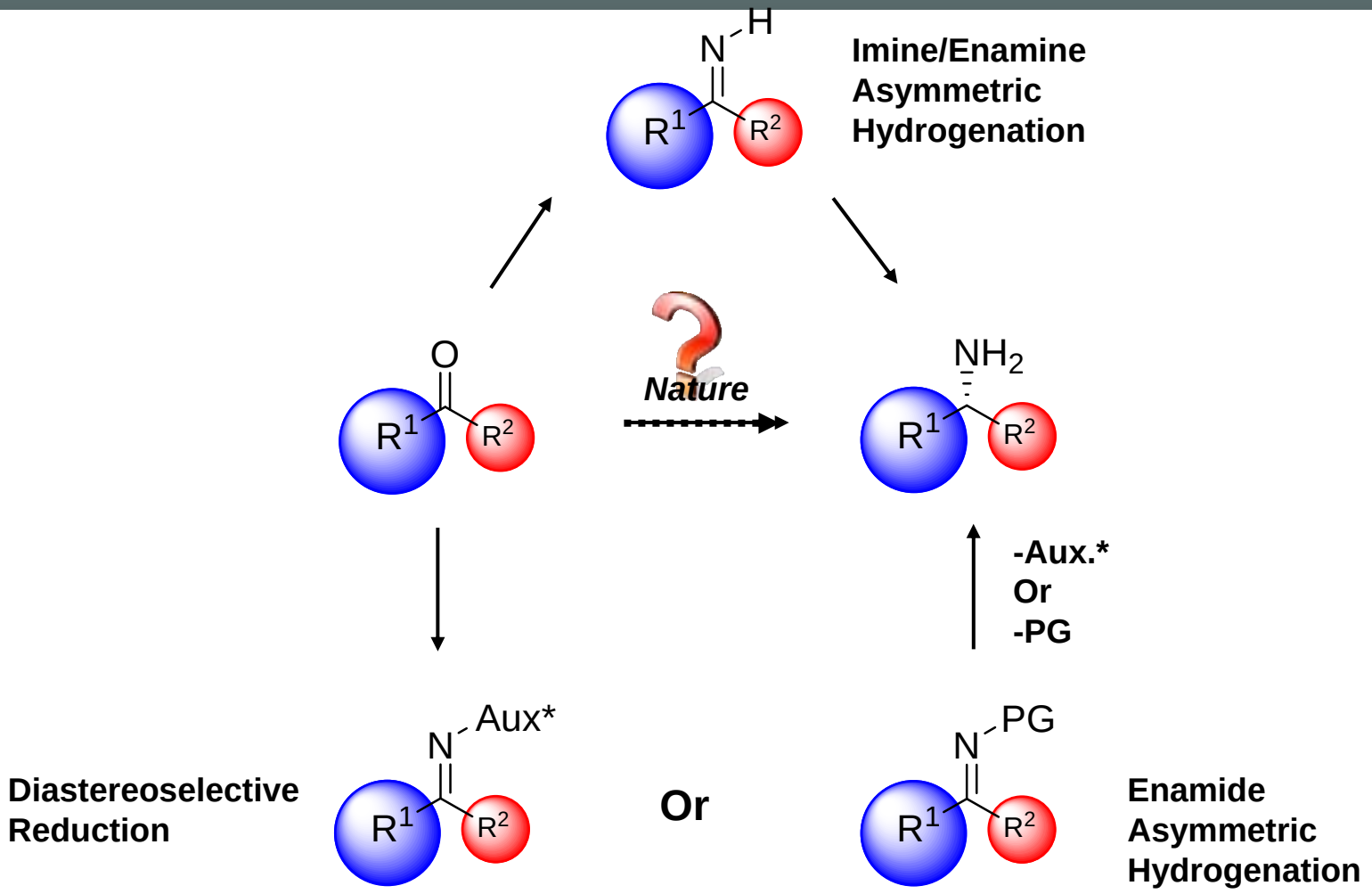
Sitagliptin: 1st Generation Route



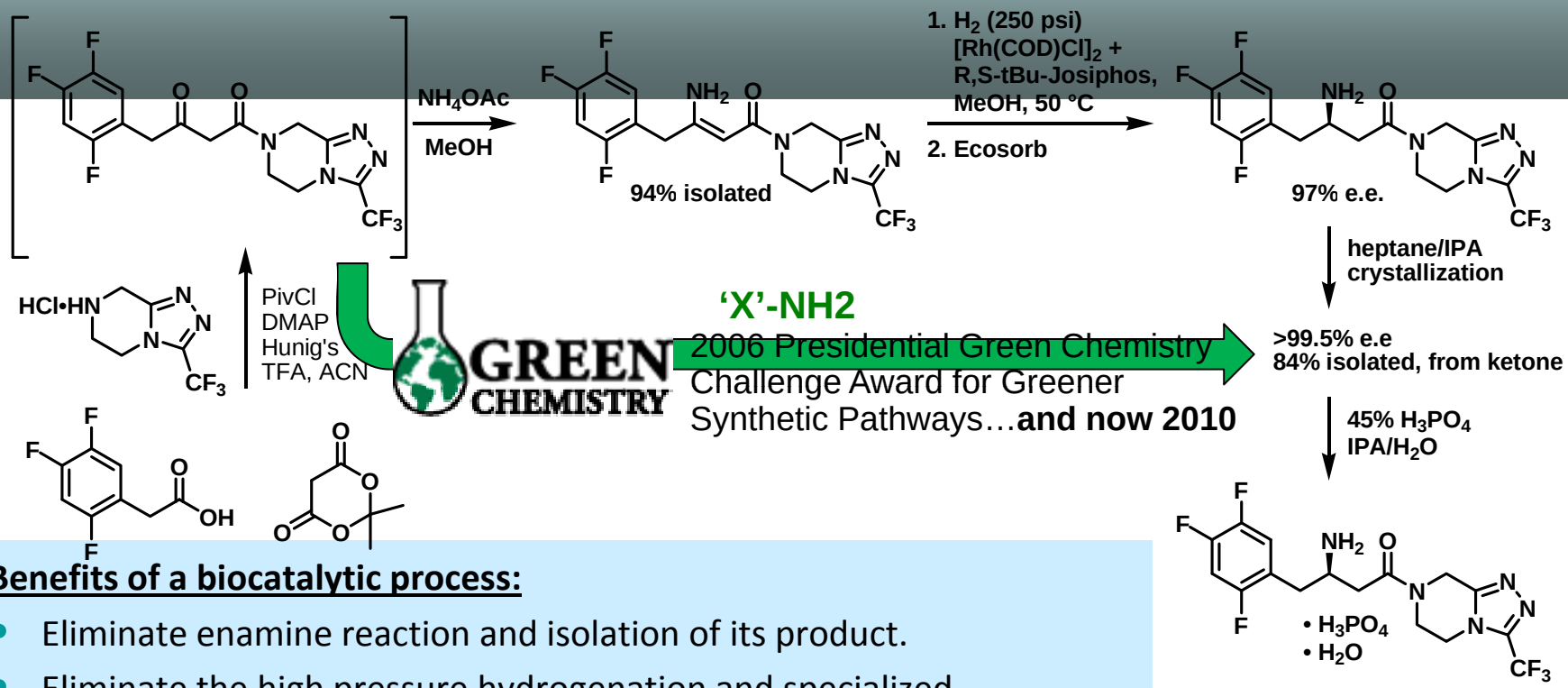
- **9 Steps, 52% overall yield, >100Kg of sitagliptin prepared**
- **Two recrystallizations required for ee upgrade, lengthy, expensive**

Hansen, K. B.,et al. *Org. Proc. Res. Dev.* **2005,** *9* (5), 634

General Asymmetric Aminations



The Aspirational Process



GREEN CHEMISTRY

'X'-NH2

2006 Presidential Green Chemistry Challenge Award for Greener Synthetic Pathways…**and now 2010**

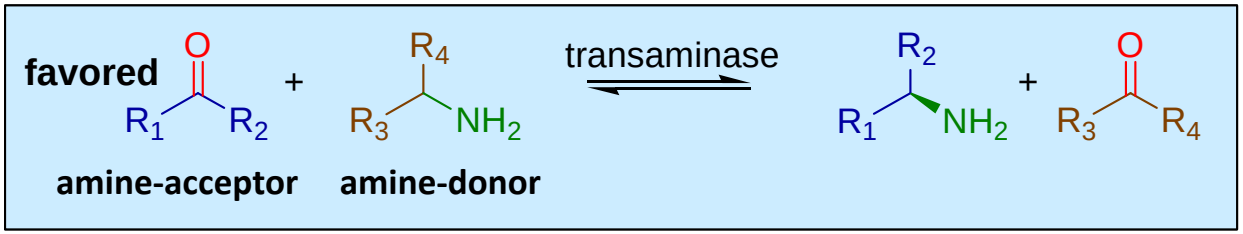
Benefits of a biocatalytic process:

- Eliminate enamine reaction and isolation of its product.
- Eliminate the high pressure hydrogenation and specialized equipment.
- Eliminate heavy metal and carbon treatment to remove it.
  Rh: $760-$10,000 per ounce, $4,061 5 year average
- Provide higher enantioselectivity to eliminate upgrade crystallization with yield loss.
- ***Economics of biocatalytic process need to be better than those of current process.***

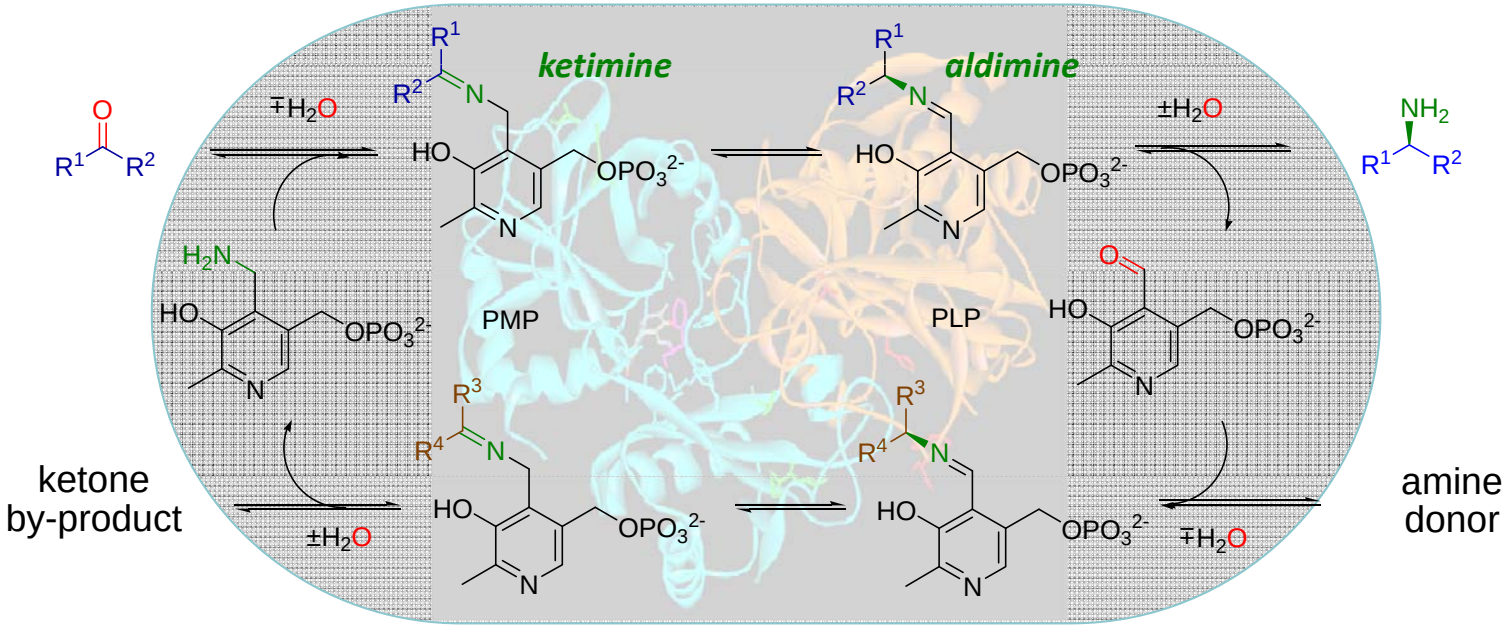
- 3 isolations, ~74% overall yield (77% from enamine).
- Insufficient enantioselectivity requires yield-reducing crystallization to upgrade.

Transaminase Reactions

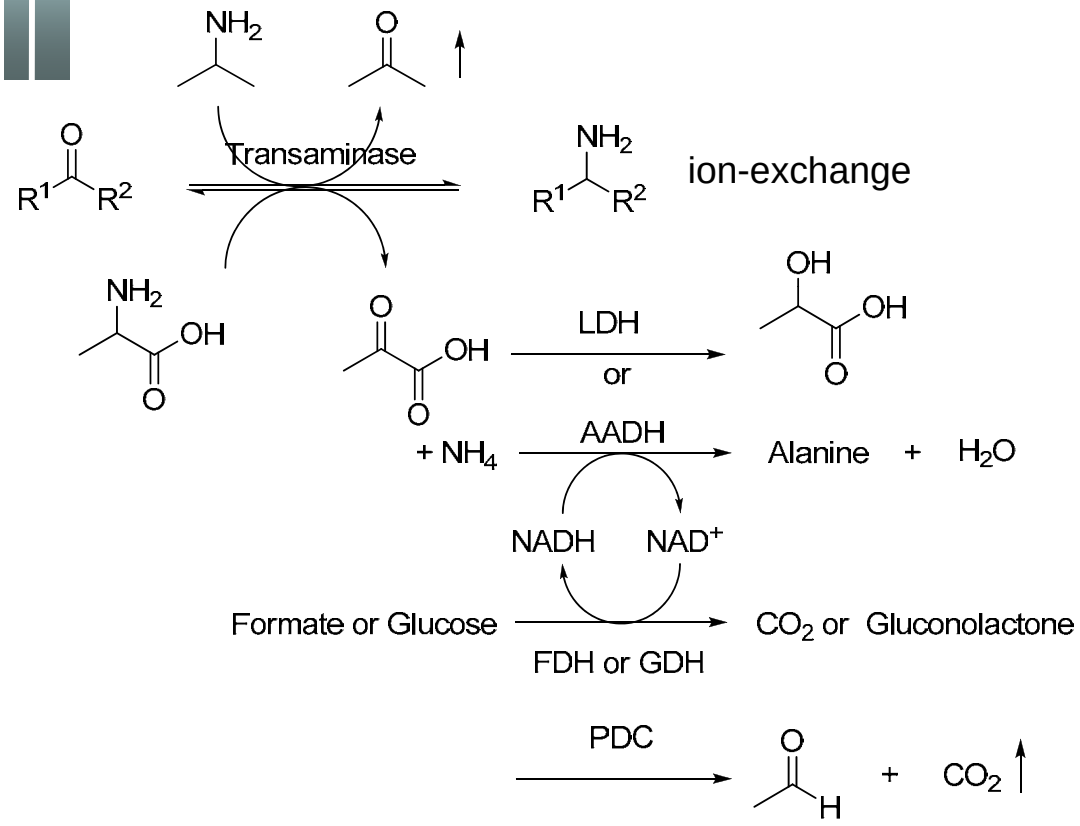
Equilibrium reaction that usually favors the ketone



Reaction mechanism:



Transaminase Processes



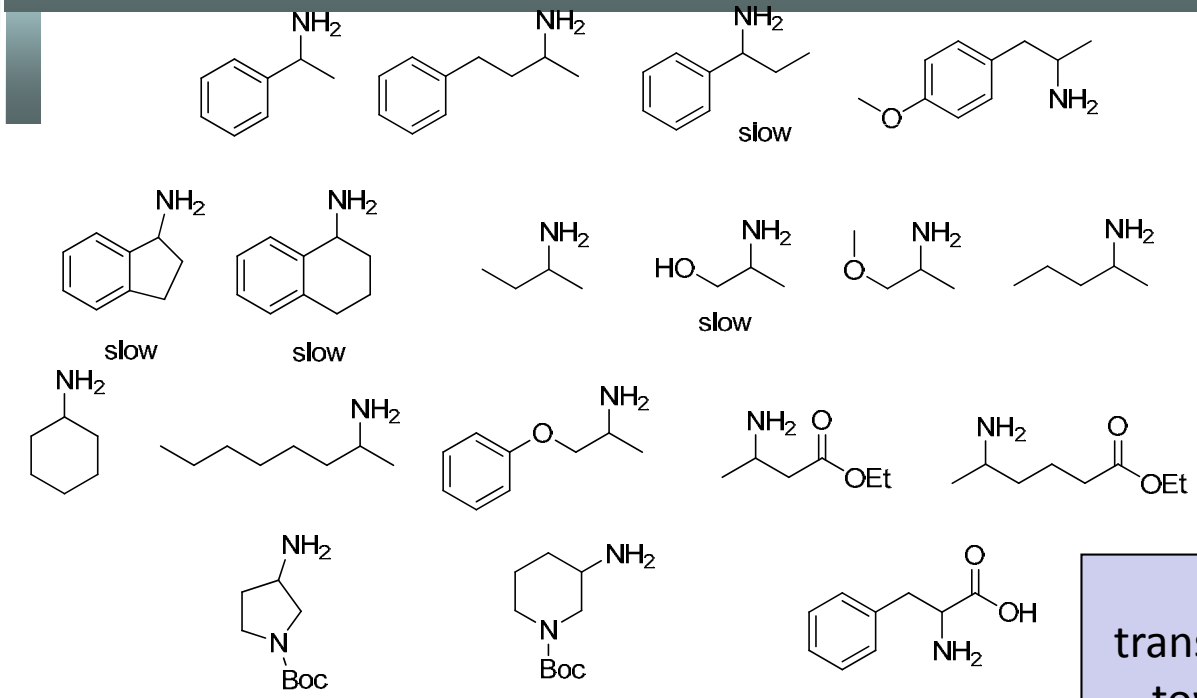
Equilibrium tends to favor the ketone; much effort on shifting equilibrium toward product amine reported:

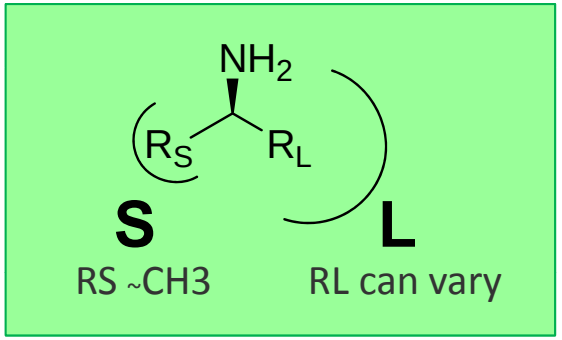
- Pyruvate to lactate using LDH,
- Pyruvate recycled back to alanine using AADH,

➢ These require co-factor recycling (GDH or FDH).

- Pyruvate to acetaldehyde using PDC,
- Amine product removed via resin – also prevent product inhibition.
- **Isopropylamine to acetone – removal by distillation**

Fundamental Problem… No Activity in Nature

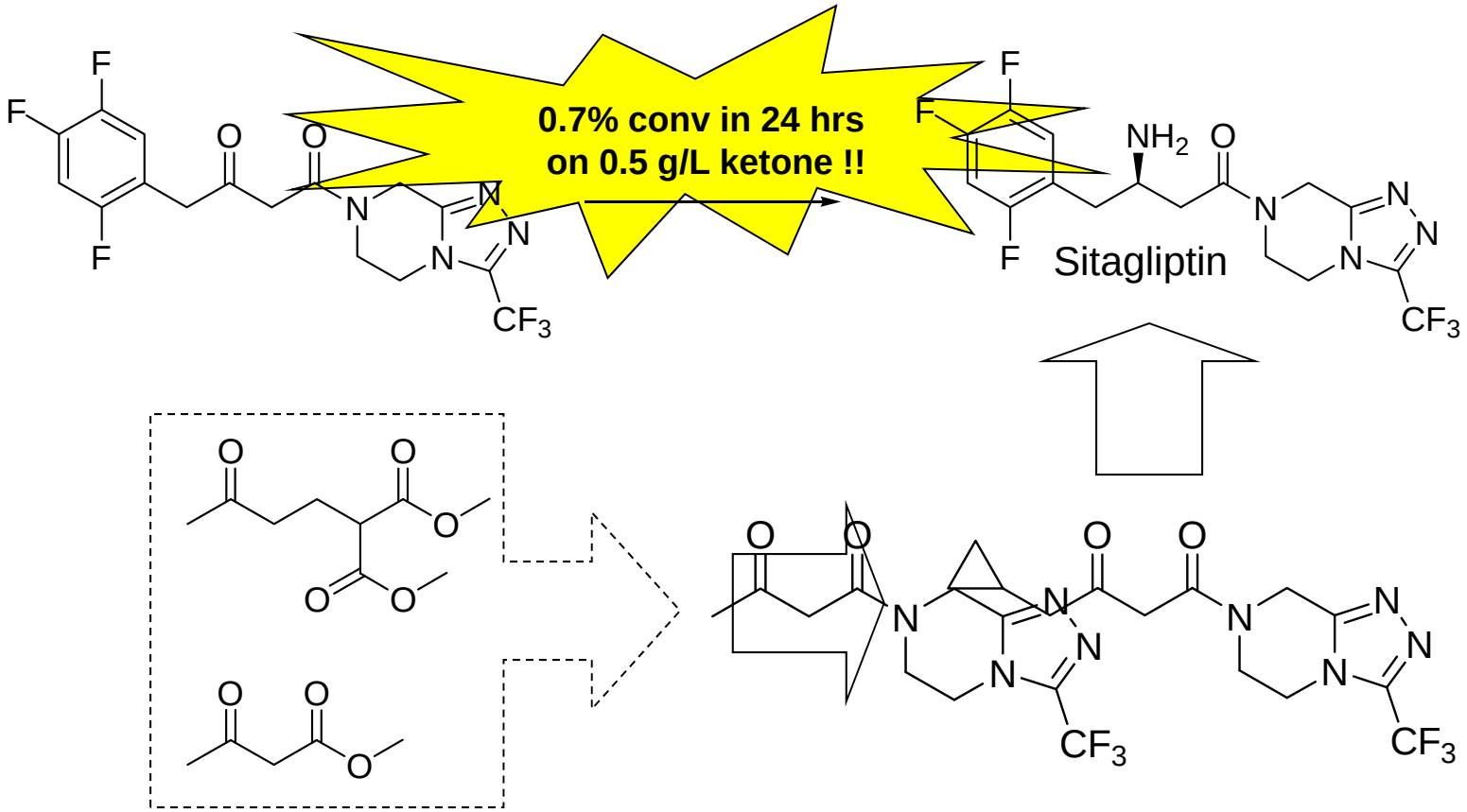




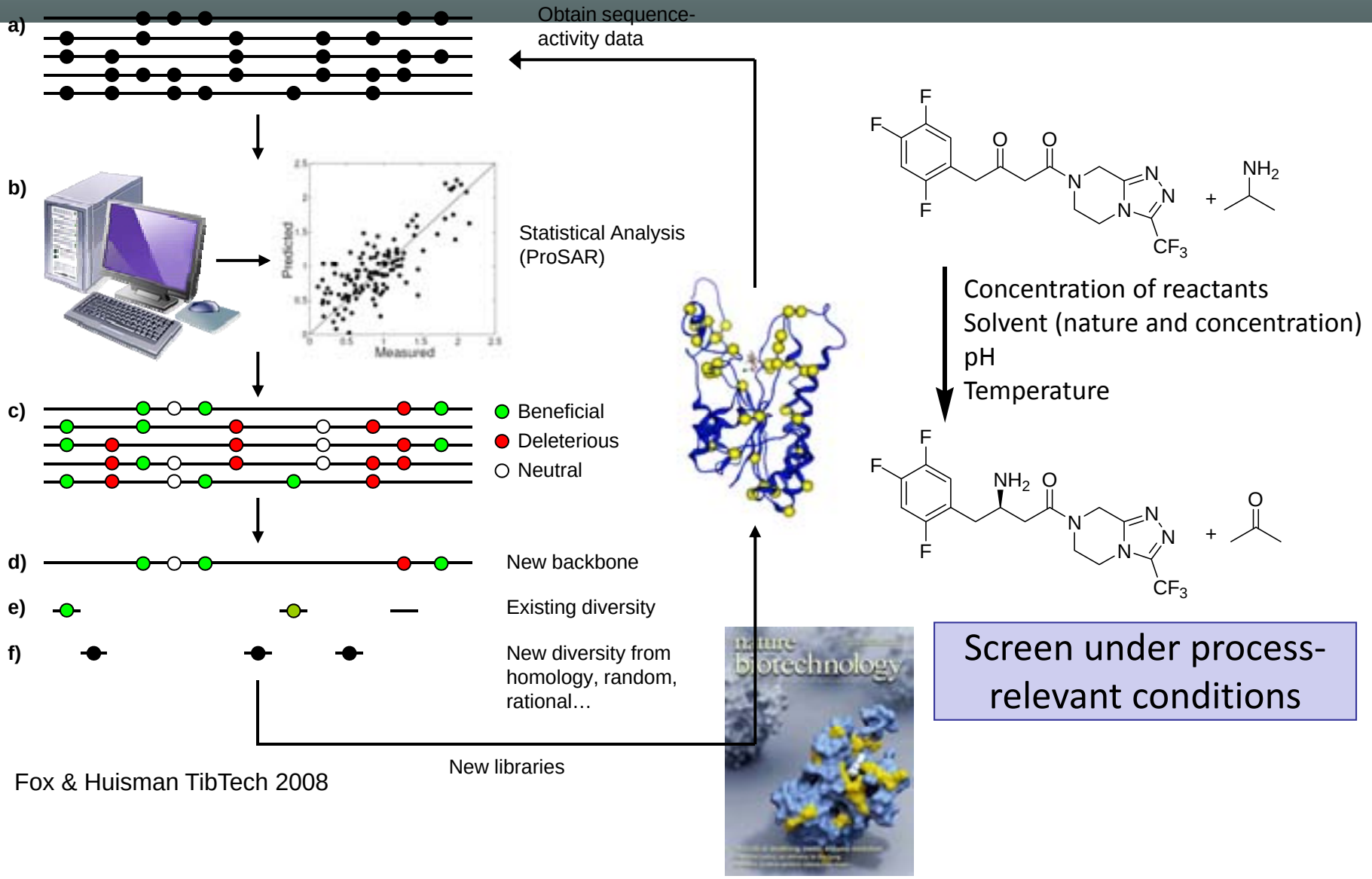
Commercially available transaminases showed no activity toward pro-sitagliptin ketone!

- Many examples of multi-Kg deliveries on methyl ketones and cyclic at Merck
- ATA-103 and ATA-113 broadest *S*-selective transaminases,
- ATA-117 broad *R*-selective transaminase,
- Accept variety of substrates, both donor and acceptor, but <u>small substituent cannot be larger than methyl group</u>.

Substrate Walking and Directed Evolution for Sitagliptin

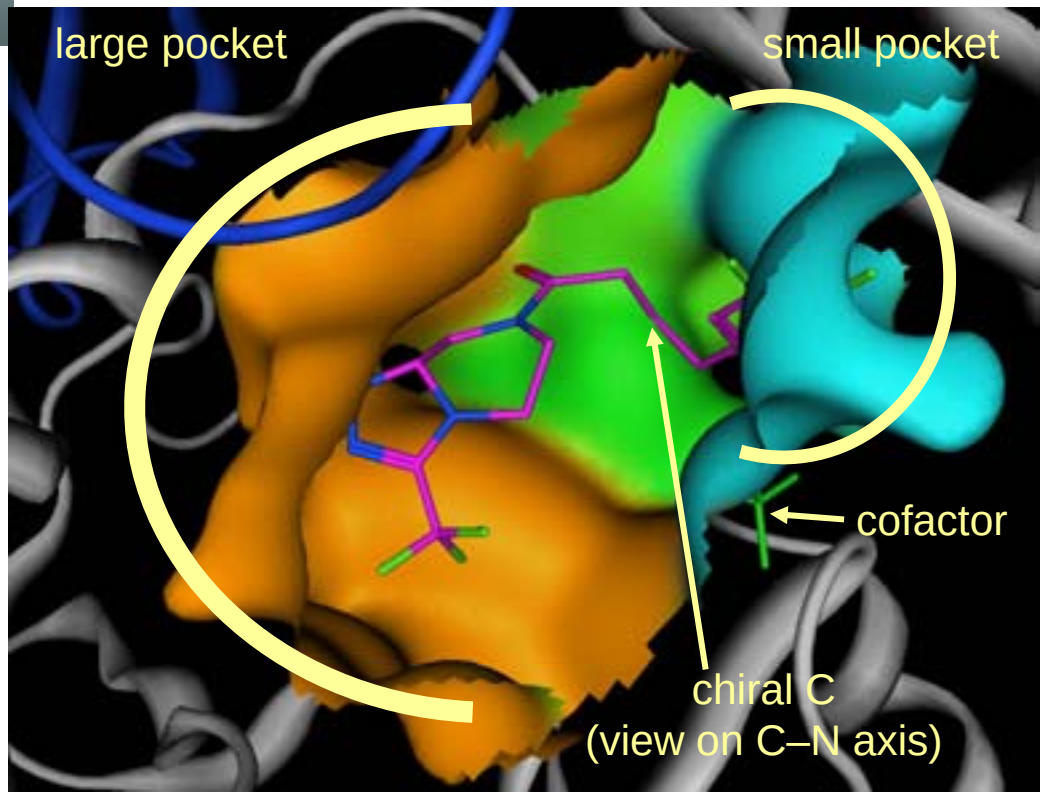


State-of-the-Art Enzyme Evolution



Fox & Huisman TibTech 2008

Active Site Homology Model



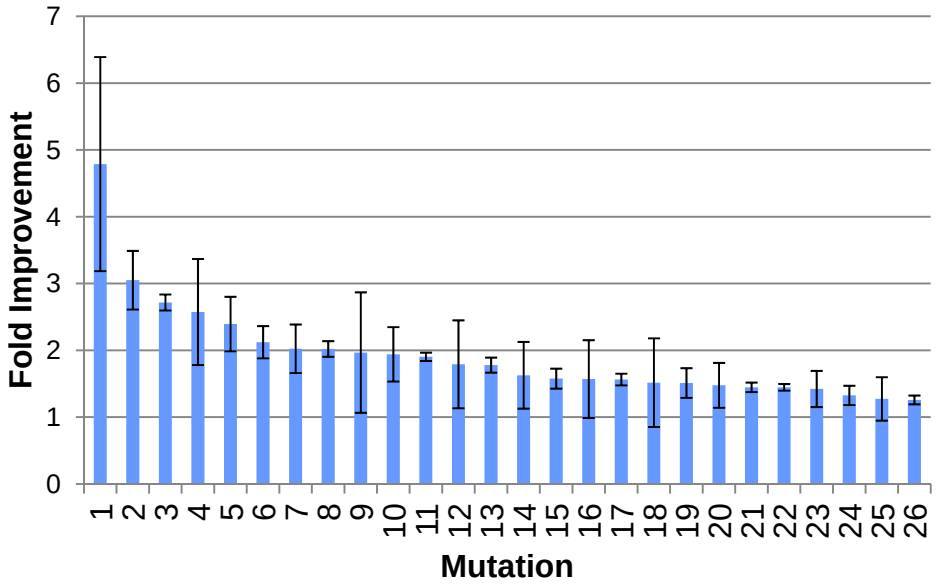
- The large group fits in the large binding pocket, but not optimally.
- The small group (F_3-phenyl) does not fit in the small binding pocket.

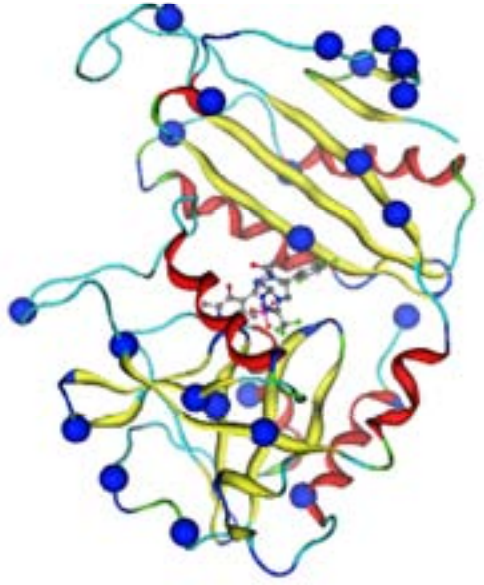
homology model subunits
large pocket
small pocket
PLP and catalytic residues

Directed Evolution Tools: Error prone PCR, site directed mutagenesis, Gene Shuffling, Modeling

Transaminase Diversity from CAPS Libs

- **576 mutations from 7 homologs**
- Screen 18 plates
- **Sequence hits**
- **26 mutations giving 1.25X-5X improvments.**
- **10 times higher hit rate than random mutagenesis.**
- **Synthesize new libraries**





Evolution for Process Fitness

- Evolution rounds 3-9 focused on increasing enzyme activity and in-process stability.
  - Generated, sorted (ProSAR), and recombined mutations across the whole protein.
- **Successive rounds were screened under increasingly challenging conditions: substrate loading, iPM concentration, co-solvent, pH and temperature**

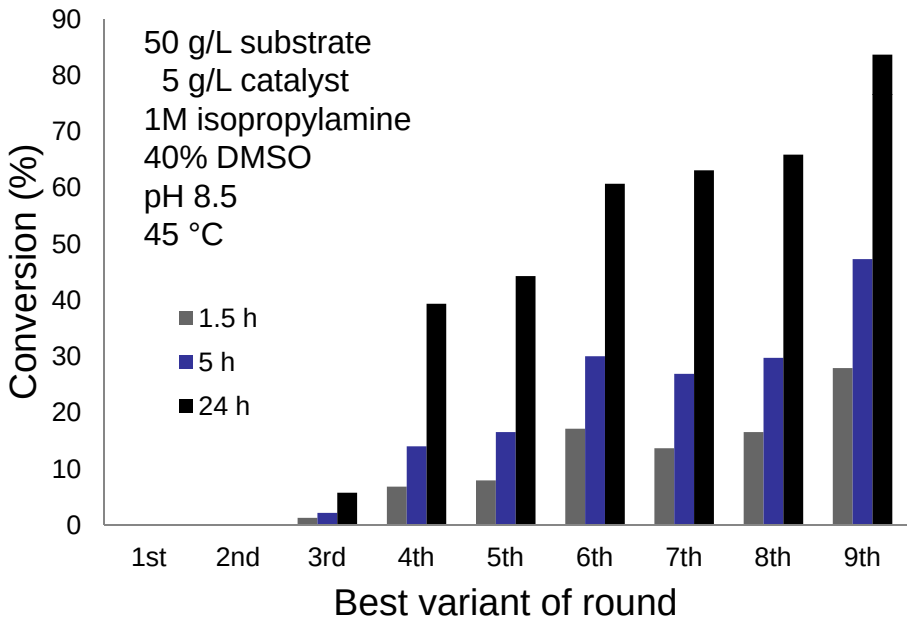
Round #	1 and 2	3	4	5	6	7-9	10-11
substrate g/L	2	5	10	40	100	100	200-275
[iPM], M	0.5	0.5	1.0	1.0	1.0	1.0	1.0
cosolvent	5% DMSO	5% MeOH	5% MeOH	10% MeOH	20% DMSO	30-40% DMSO	50% DMSO
pH	7.5	7.5	8.5	8.5	8.5	8.5	10
temp, °C	22	30	30	45	45	45	45

- Merck process development began with best round 5 variant.
  - Biocatalyst evolution and process development then proceeded in parallel.
  - Merck's process changes were reflected in biocatalyst screening conditions.

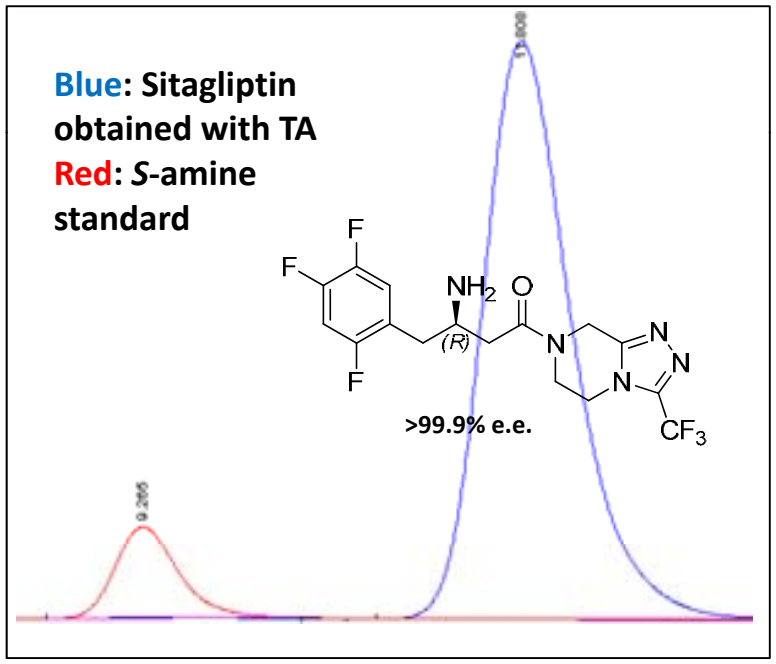
Minor Enantiomer Never Detected

Head-to-Head Comparison of TA Progeny

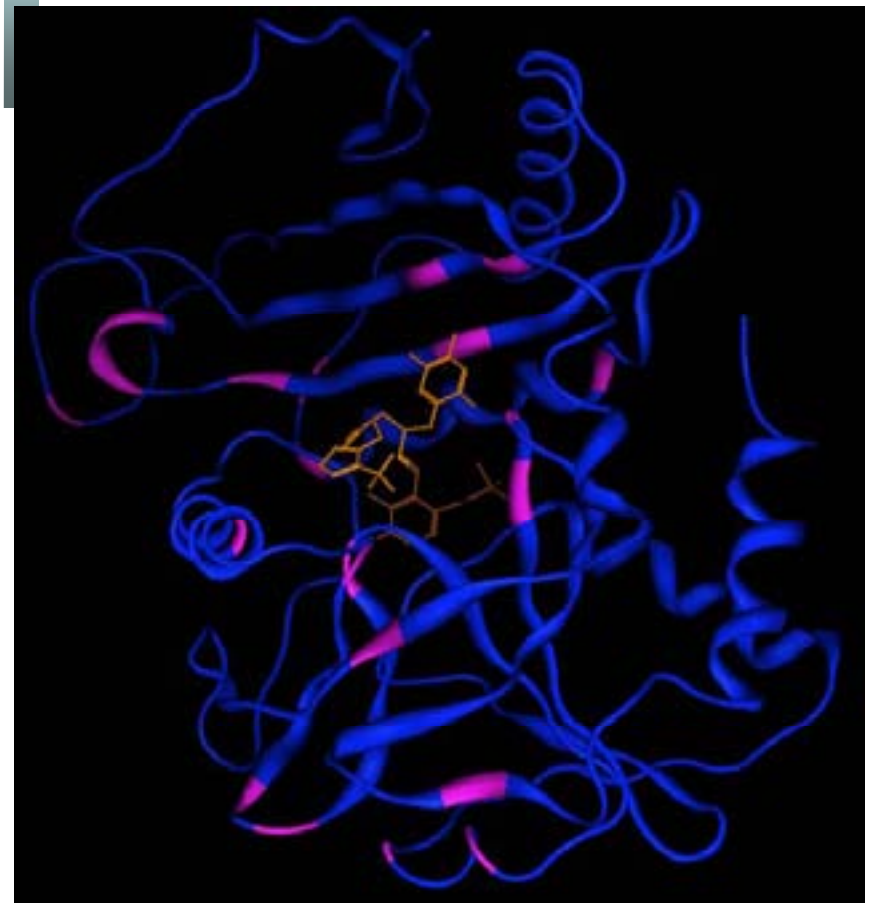
Top variant of each round was tested under identical conditions (vial reactions):



Chiral purity of reaction product:



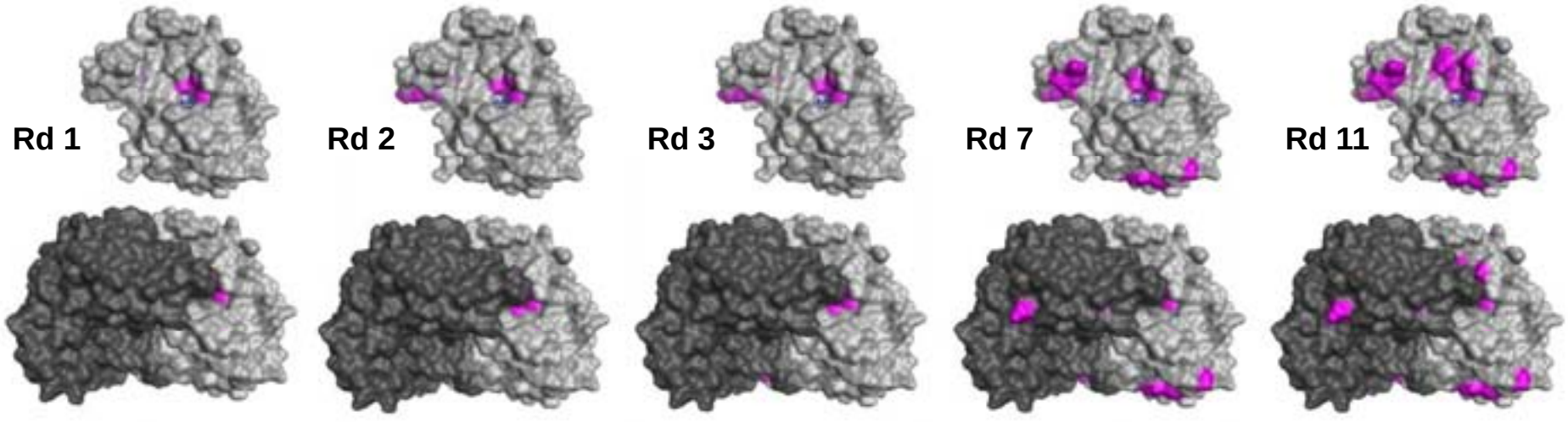
Summary and Current "Final" Biocatalyst



- A transaminase-based process for sitagliptin synthesis was:
  - first enabled by substrate walking and structure-guided directed evolution generating an activity that previously did not exist,
  - then improved 4-to-5-orders of magnitude by state-of-the-art directed evolution technology in parallel with process development.
- Final catalyst contains 25 mutations.
- Of the 16 amino acid residues predicted to be interacting with the substrate:
  - 2 are catalytically essential,
  - 7 were mutated in this catalyst (50%)

Global Map of Mutations

Monomer View



Active Dimer View

Accumulated mutations highlighted in purple

From NO Hits to Industrial Enzyme

www.sciencemag.org SCIENCE VOL 329 16 JULY 2010

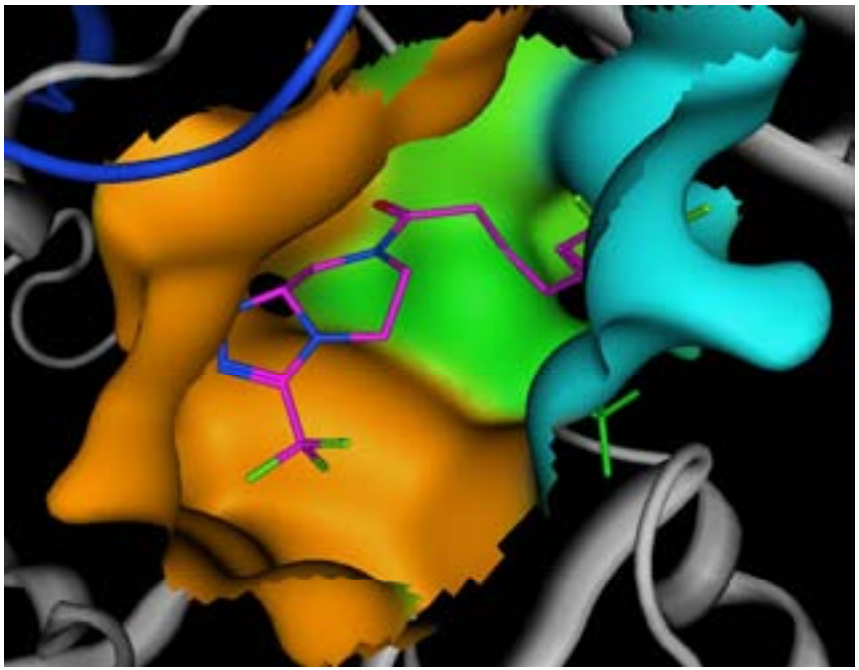
Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture

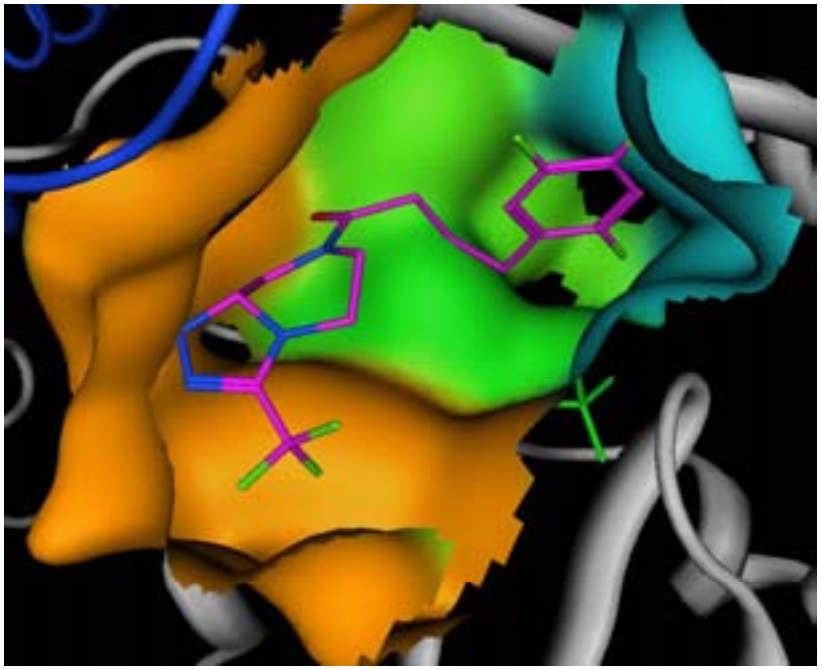
Christopher K. Savile,[1]* Jacob M. Janey,[2]* Emily C. Mundorff,[1] Jeffrey C. Moore,[2] Sarena Tam,[1] William R. Jarvis,[3] Jeffrey C. Colbeck,[1] Anke Krebber,[1] Fred J. Fleitz,[2] Jos Brands,[2] Paul N. Devine,[2] Gjalt W. Huisman,[1] Gregory J. Hughes[2]

2010 GREEN CHEMISTRY

Januvia™ (sitagliptin phosphate)

Janumet™ (sitagliptin/metformin HCl)

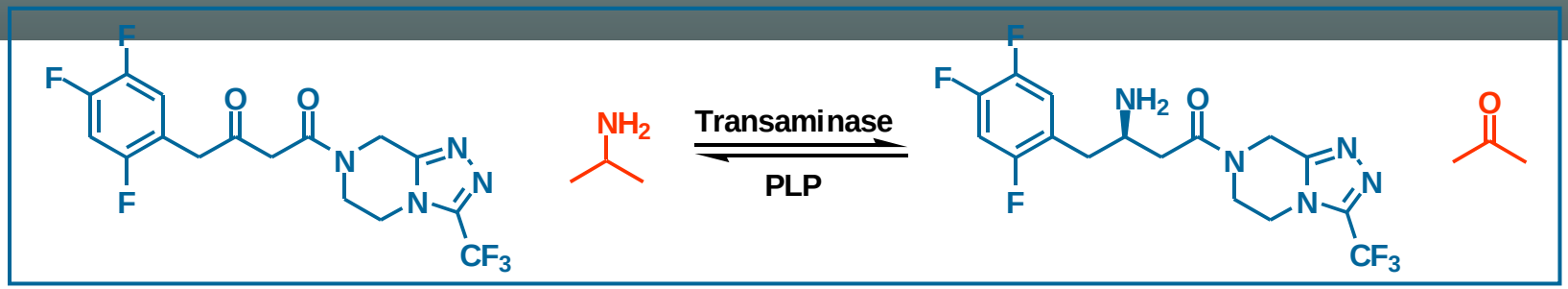




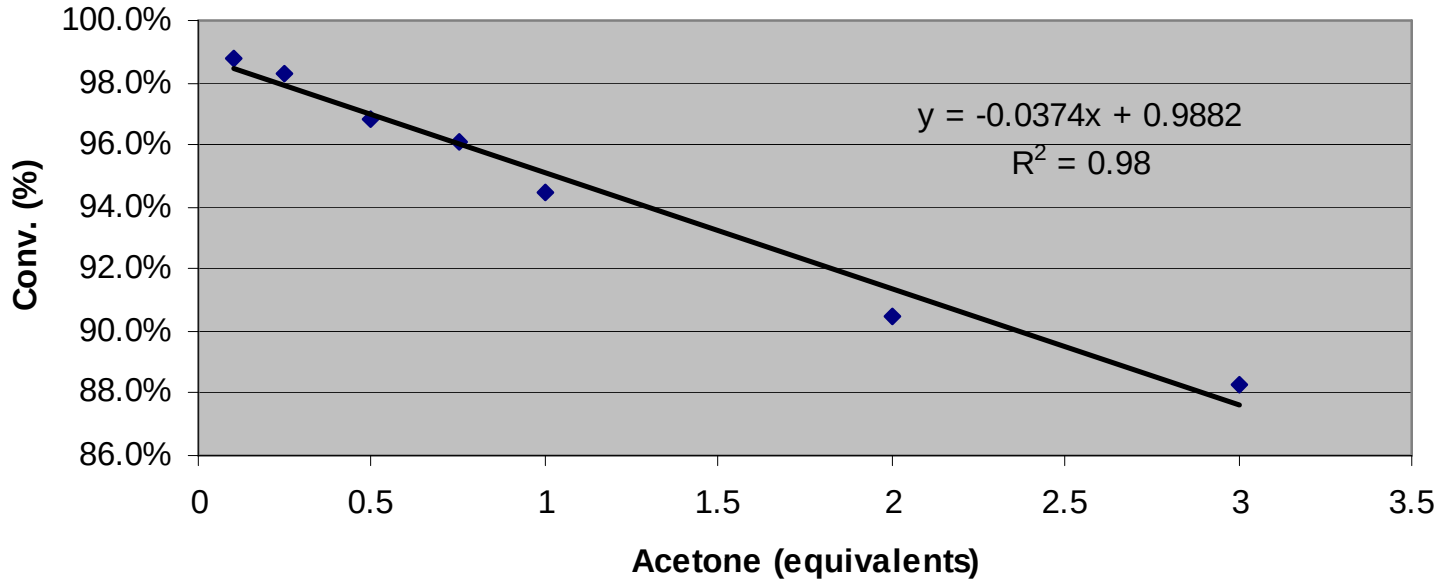
Transamination
Process Challenges

- **Ketoamide solubility <1g/L in water and only 9-10g/L in DMSO/water**
- **Second phasing: free base, imine dimer, and ketoamide**
  - **Feed in ketoamide as DMSO solution…to make 50%v/v in water**
- **pH control: pKa product < isopropyl amine (+ loss to vapor phase)**
- **pH probe fouling**
  - **Feed in 4M iPr-amine in water with feedback loop**
  - **On pilot scale…eliminate pH cart and use set charge rate 4M iPr-amine**
- **Acetone removal needed to drive equilibrium**
  - **Use vacuum and nitrogen sweep**
  - **Monitor acetone and conversion with ReactIR (calibrated)**
- **Enzyme removal during work-up (emulsion issues and regulatory)**
  - **Ppt. enzyme with HCl then filter (kilo scale)**
  - **Extract away enzyme with IPAc/IPA (pilot/factory scale)**

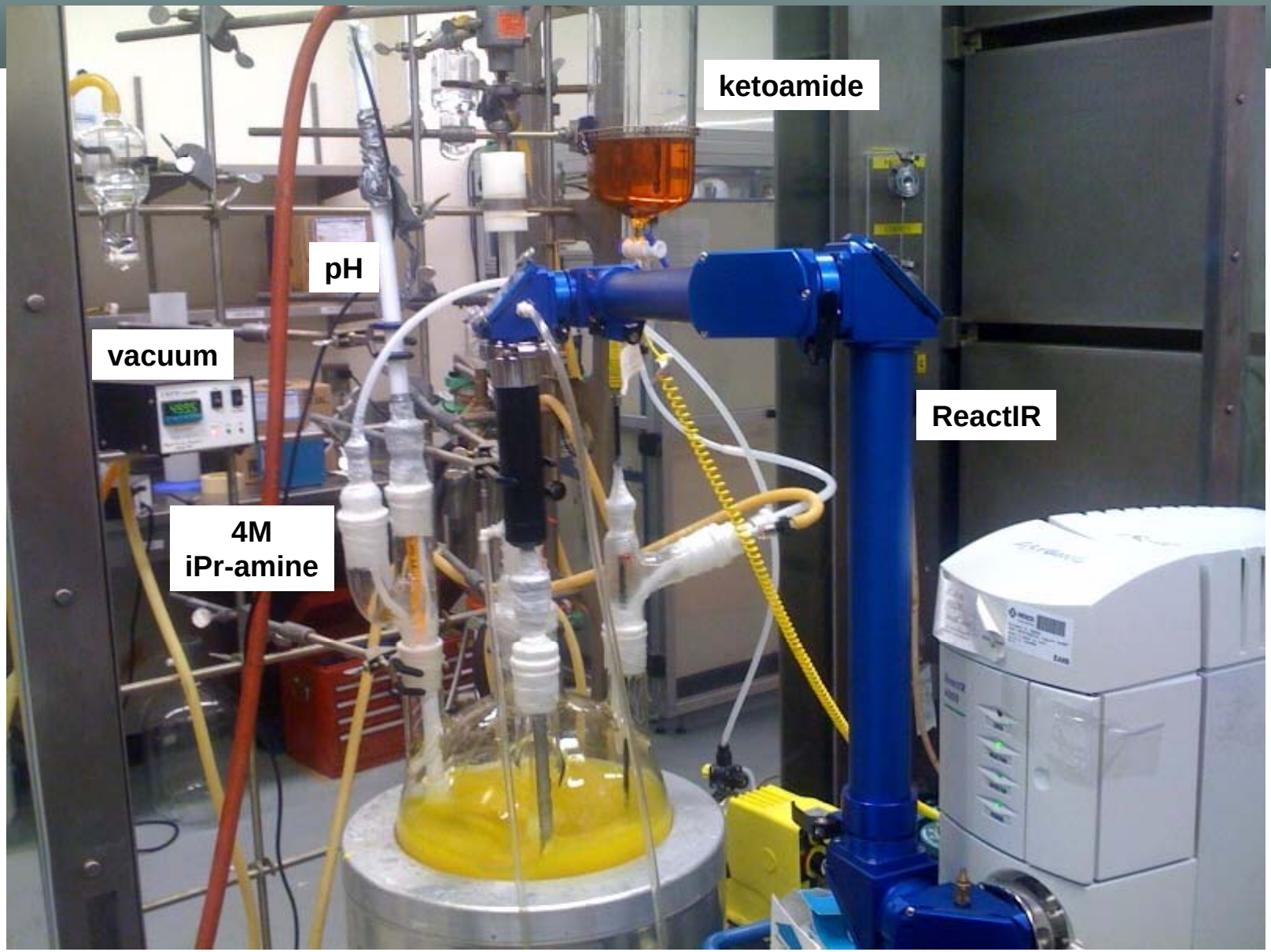
Acetone Equilibrium Study



Acetone Equilibrium, pH 10



Kilo Scale Set-up



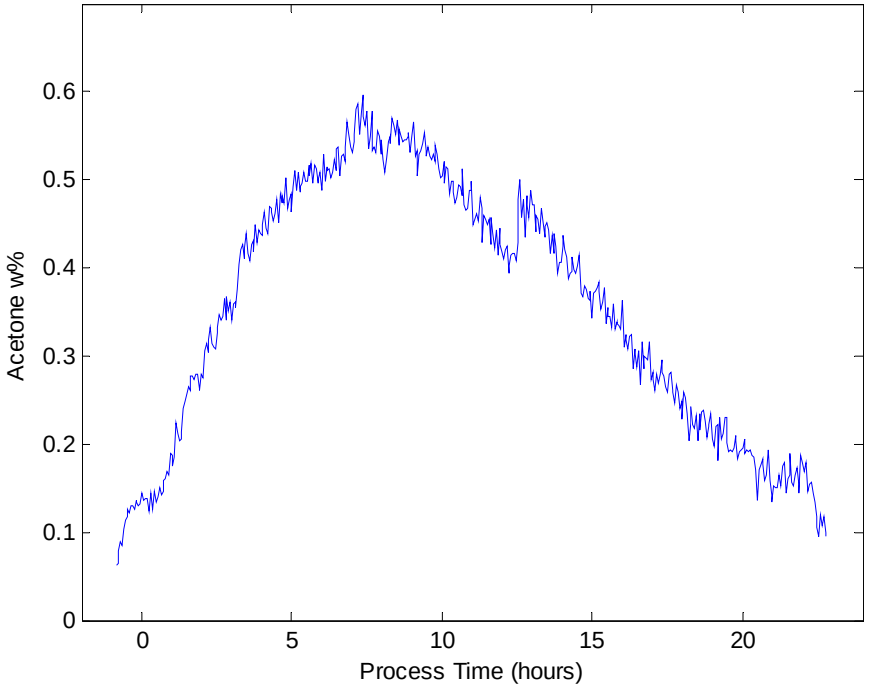
React IR Acetone Monitoring PAT

Pilot Plant Batches (30 kg or 90 kg)

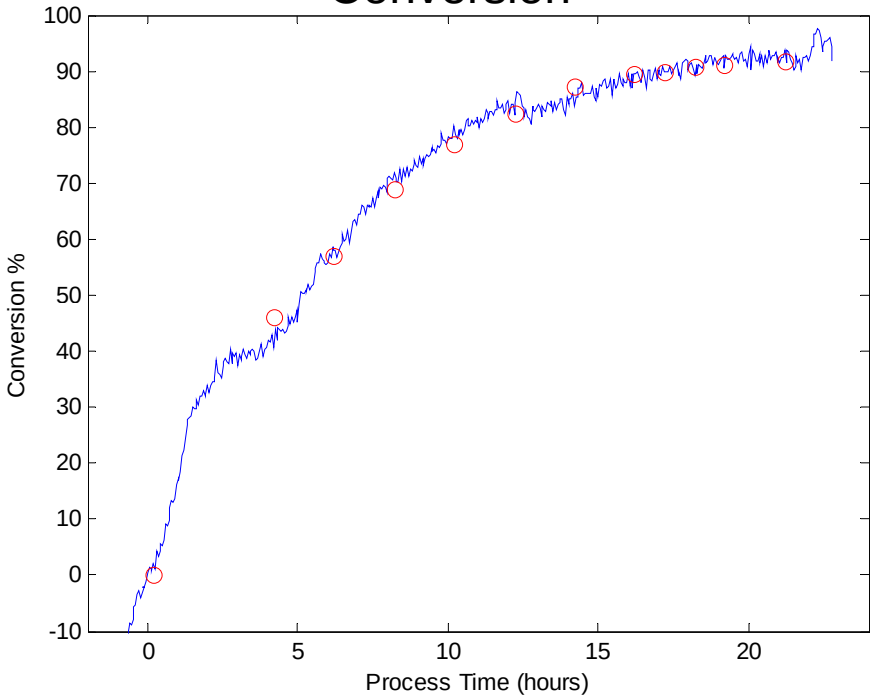
- 150-375 torr
- 2-10 fps nitrogen sweep of headspace
- 2-4 m/s tip speed
- End of Reaction acetone concentration = <0.2%
- 96% Conversion

Highly scale dependant

Acetone



Conversion



Many Factors to Optimize

- **FDA filings requires "Quality by Design": A way to allow process changes within a defined "operating space" without having to re-file**

Transaminase Factors

- **DMSO/Water Ratio and split**
- **Agitation**
- **Vacuum**
- **Nitrogen Sweep**
- **pH range**
- **Temperature**
- **Order of operations**
- **Buffer strength**
- **Enzyme charge**
- **PLP charge**
- **iPr-amine initial charge**
- **Ketoamide addition rate**

How do we identify an operating range for this many parameters that may/may not interact?

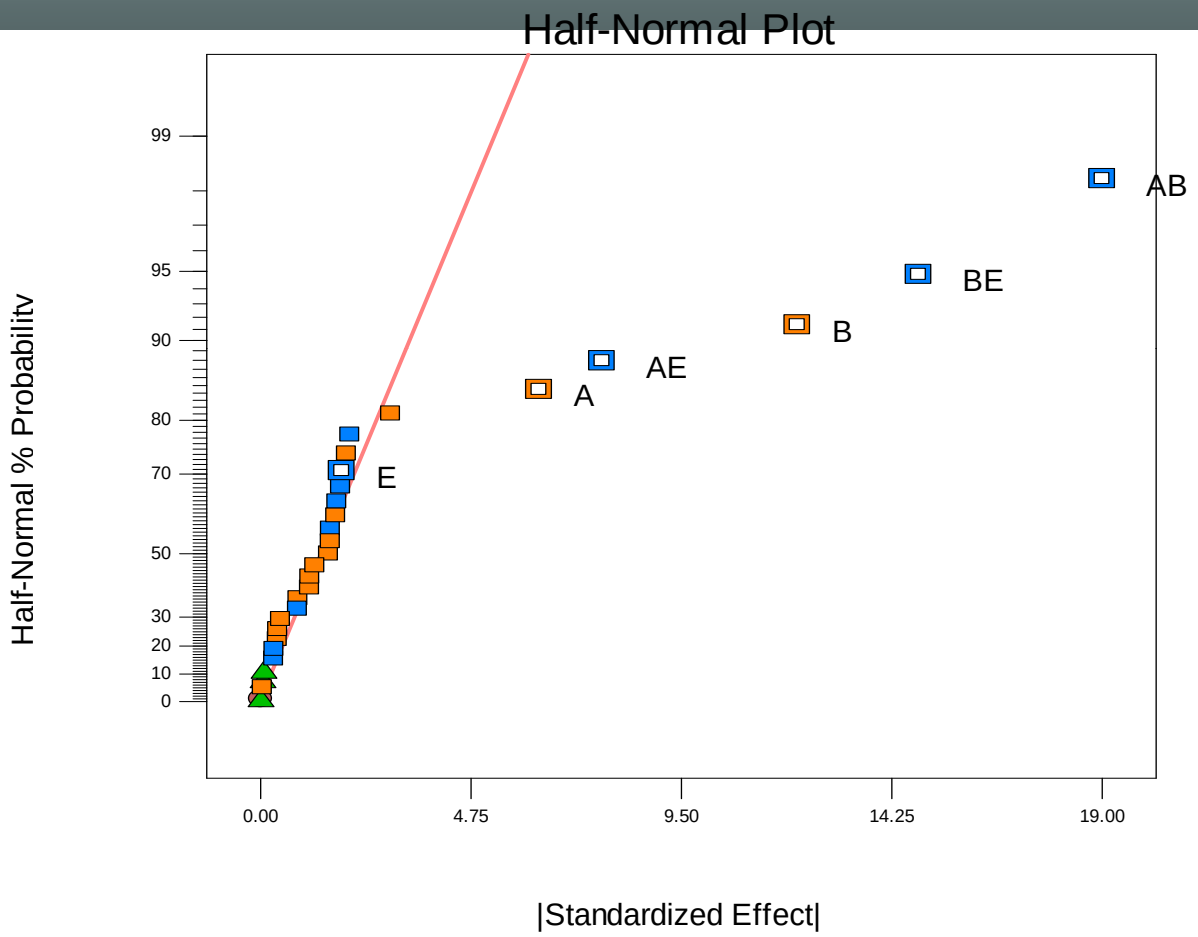
QbD via DOE

Transaminase Results: DOE 1

Design-Expert® Software
Assay Yield

▲ Error from replicates

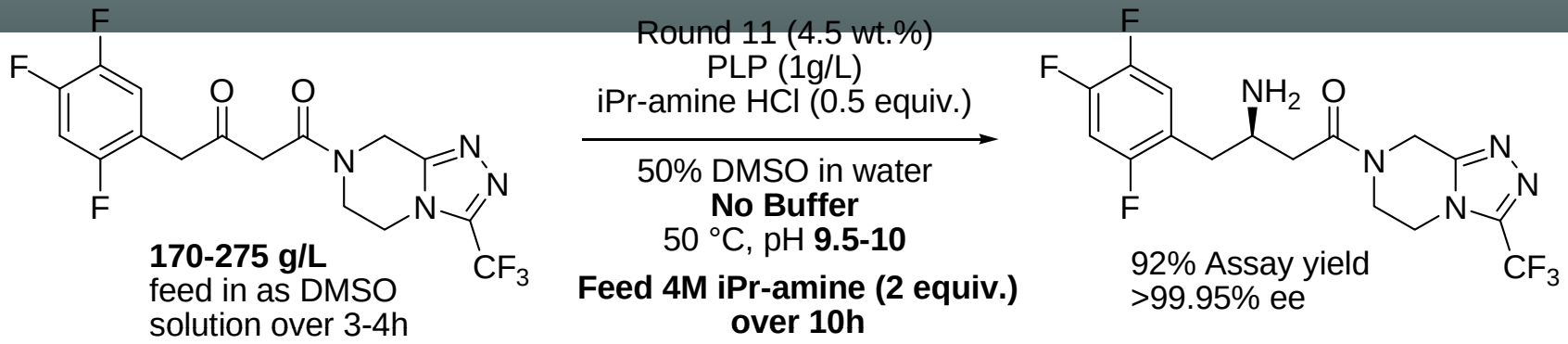
Shapiro-Wilk test
W-value = 0.956
p-value = 0.463
A: Rxn Temp
B: pH
C: DMSO in Disso
D: DMSO in Reactor
E: IPA.HCl Stock Soln
F: Enzyme
G: PLP
■ Positive Effects
■ Negative Effects



- **Fractional design: 16 exp. + repeats, 4 center points**
- **Yield in a set time was used as output**

QbD

Transaminase Results: Final (after 2nd DOE)



170-275 g/L
feed in as DMSO solution over 3-4h

92% Assay yield
>99.95% ee

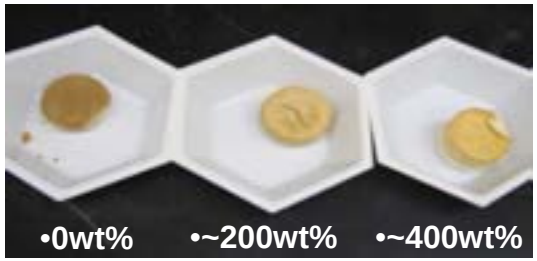
EOR: 97% sitagliptin, 3% ketoamide , 0.1% dimer, ~1% olefin

Key Features:

- No pH control needed…continuous 4M iPr-amine feed
- No buffering needed
- Reaction time cut to 12h from 21h (better stirring and higher pH)
- Vacuum (150 torr), nitrogen sweep (10 fps), and 4 m/s tip speed gives fast acetone removal
- **Enzyme remains ~85% active at EOR…insensitive from pH 7-12 and up to 50 °C after aging 3 days in 50% DMSO.**

Work-Up QbD

Filtration Removal Process



Replaced filtration with extractive process

Extractive Removal Process

IPA / IPAc

Increasing IPA

- Enzyme Gel in downstream extractions – affected settling time and removal of enzyme
- Requires polishing filters
- **Extremely poor flux**

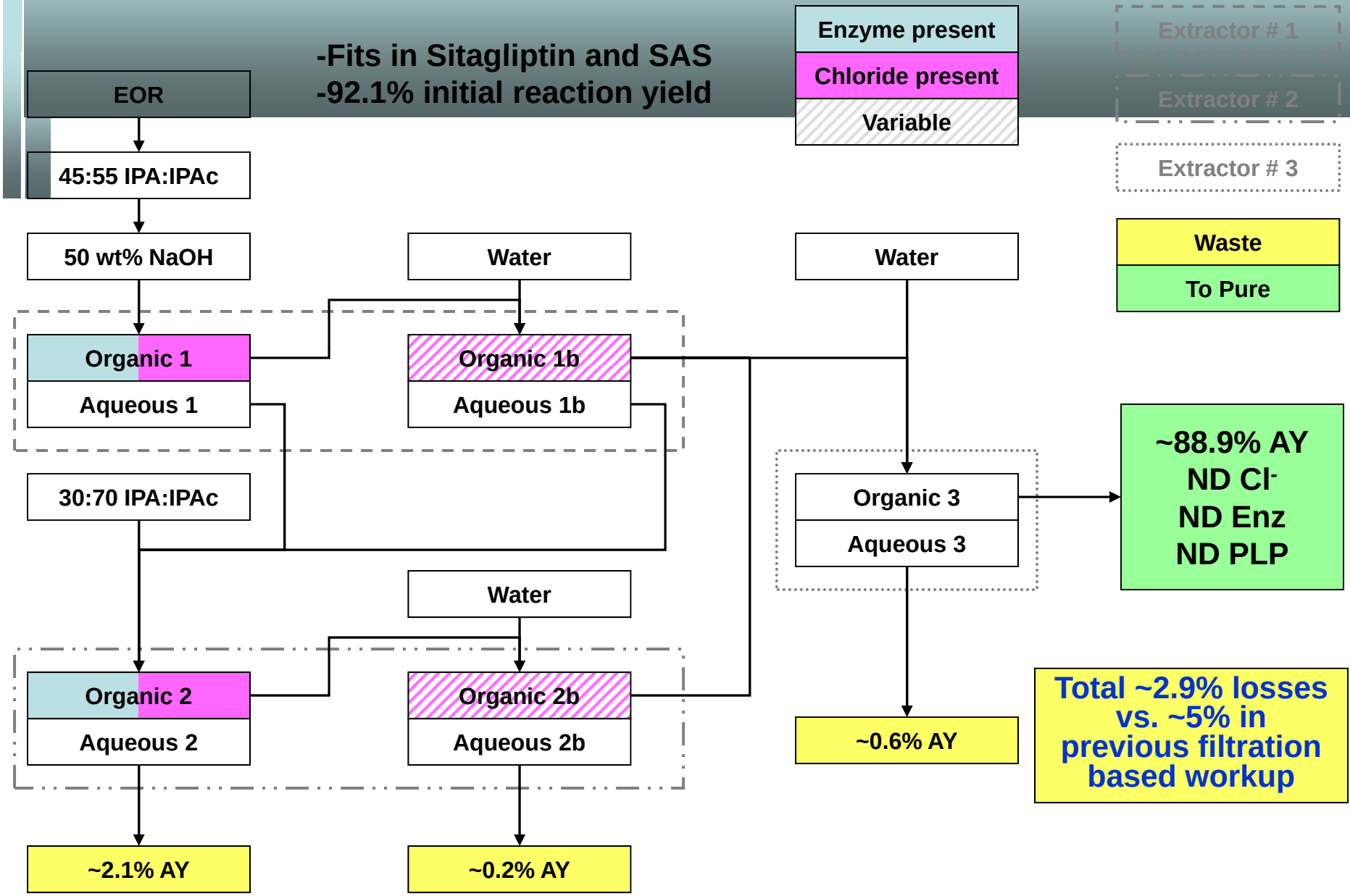
- Eliminated long filtration times
- More portable (less capital dependence for enzyme removal)
- Cleaner extraction interfaces throughout workup
- **General for all enzymes**

• Simple: Add 0.5 reaction volume IPA then 0.5 IPAc (or any other alcohol/organic combo)

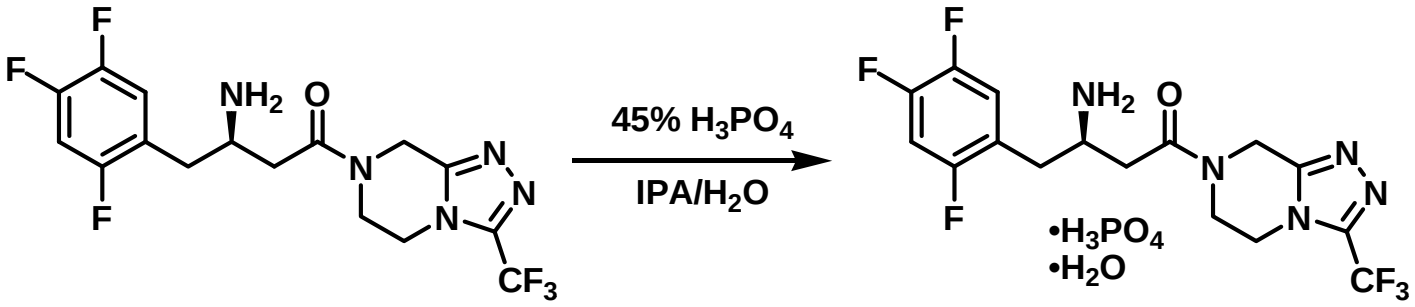
MRL & MMD

Extractive Workup Summary

-Fits in Sitagliptin and SAS
-92.1% initial reaction yield



Final Pure Step



Januvia
(Sitagliptin Phosphate Monohydrate)

Recall: Current route charges crystalline free base
New route use IPA solution

98% yield (vs. 96% in current)
Same physical properties
Acceptable purity profile

- Higher yields and purity…can increase crystallization yields (less need to reject impurities)
- Crude IPA stream has only ketoamide, IPAc, and DMSO as new impurities
- Pure step tolerates up to 13% ketoamide (complete rejection). DMSO/IPAc rejected

<10 ppm protein by size exclusion HPLC and fluorescent detector

QbD
Results

Original

Enamine Amide Hydrogenation → Free Base Crystallization → API Pure Crystallization

Solid

84% yield

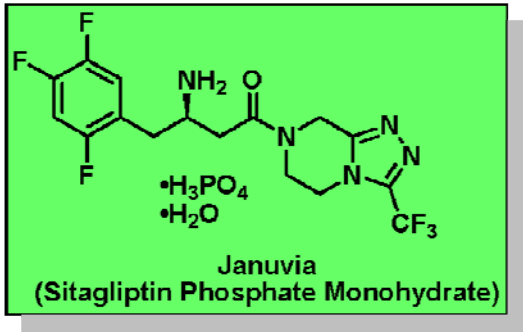
96% yield

New

Ketoamide Transamination → Work-up (Enzyme Removal) → API Pure Crystallization

90% yield

98% yield

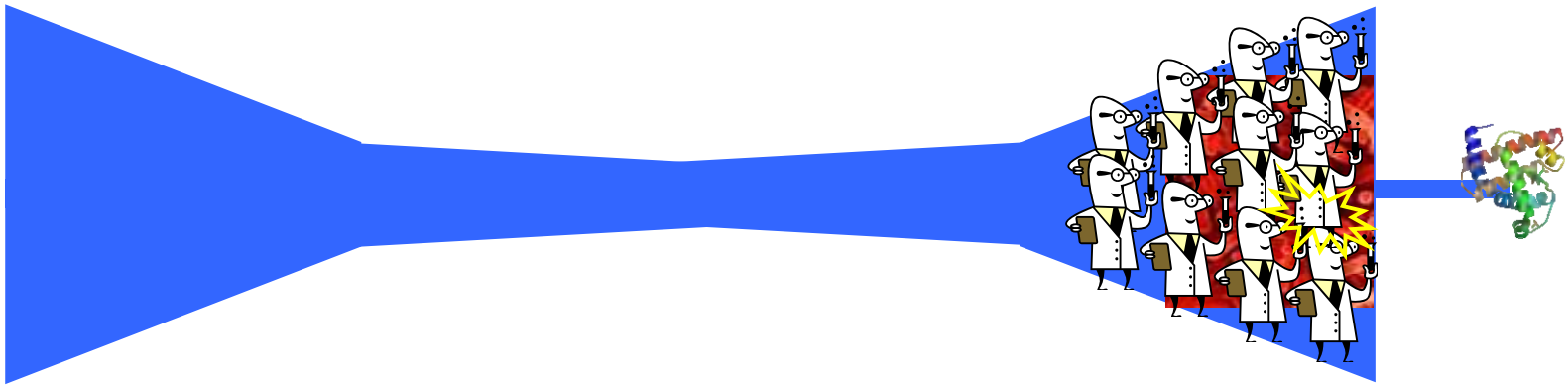


Januvia
(Sitagliptin Phosphate Monohydrate)

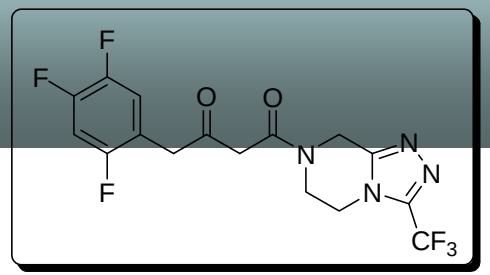
- Through Process, 2 Steps
- 88% Yield from Ketoamide
- 7% Increase in overall yield
- ~30% more productivity
- PMI from 38 down to 31
- No special equipment needed
- No metals
- No high pressure hydrogen
- Reneweable catalyst vs. mined

Develop the Best Process …or the Best Solution?

Screening → **Optimization** → **Final Process Definition**

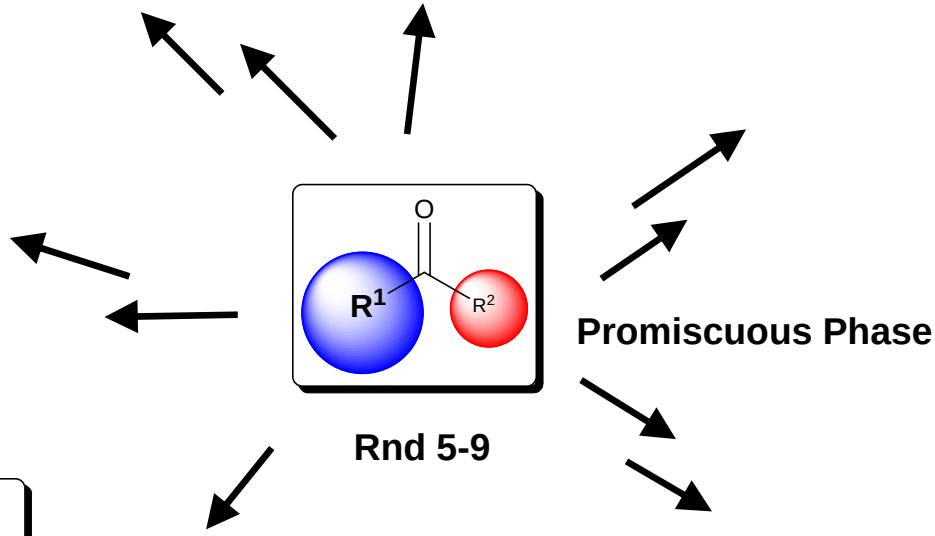


Applications Beyond Januvia….



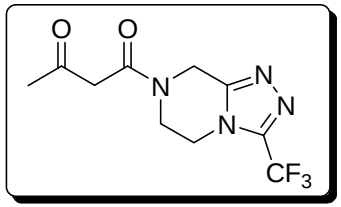
Rnd 11

Activity

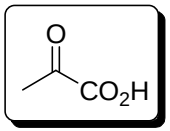


Promiscuous Phase

Rnd 5-9

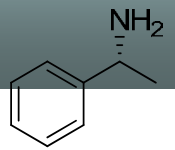


Rnd 1

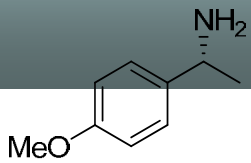


WT

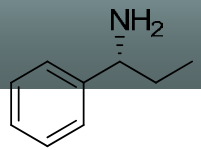
Ketone Survey



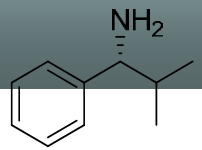
ATA-117: 72% conv.
Rd 11: TA: 72% conv.
>99% ee



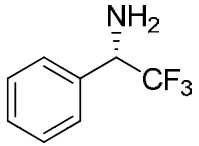
ATA-117: 4% conv.
Rd 11 TA: 28% conv.



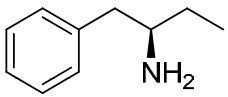
ATA-117: 0% conv.
Rd 11 TA: 80% conv.



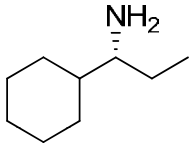
ATA-117: 0% conv.
Rd 11 TA: 56% conv.



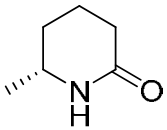
ATA-117: 0% conv.
Rd 11 TA: 99% conv.,
67% yield, 99% ee



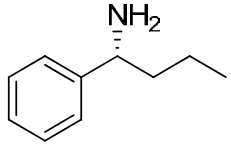
ATA-117: n.d.
Rd 11 TA: 90% conv.



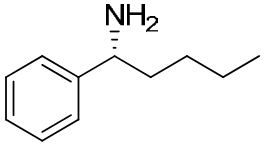
ATA-117: n.d.
Rd 11 TA: 60% conv.



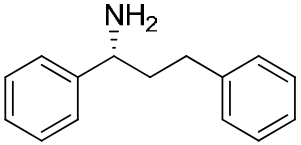
ATA-117: 100% conv.
Rd 11 TA: 100% conv.



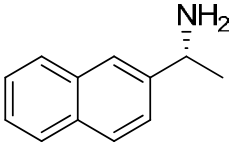
ATA-117: 0% conv.
Rd 11 TA: 56% conv.



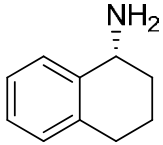
ATA-117: 0% conv.
Rd 11 TA: 31% conv.



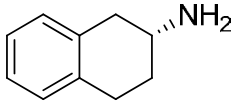
ATA-117: n.d.
Rd 11 TA: 9% conv.



ATA-117: 17% conv.
Rd 11 TA: 95% conv.,
95% yield, 99% ee



ATA-117: 10% conv.
Rd 11 TA: 15% conv.



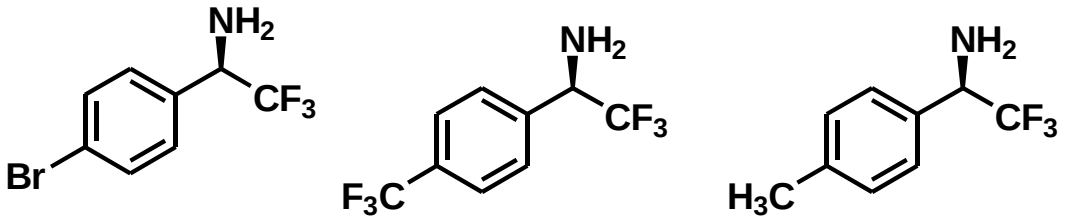
ATA-117: 17% conv.
Rd 11 TA: 100% conv.

n.d. = not determined

Transamination of Trifluoromethyl Ketones

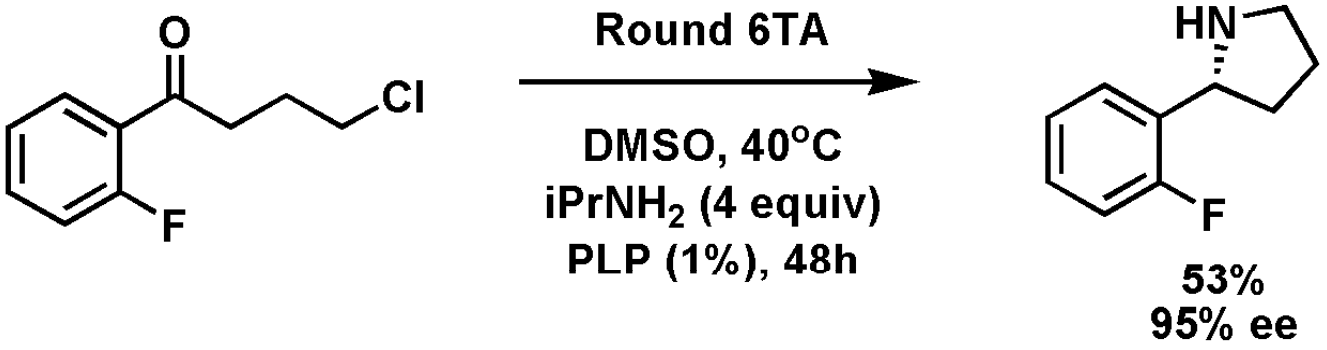
100% Conv
99% ee

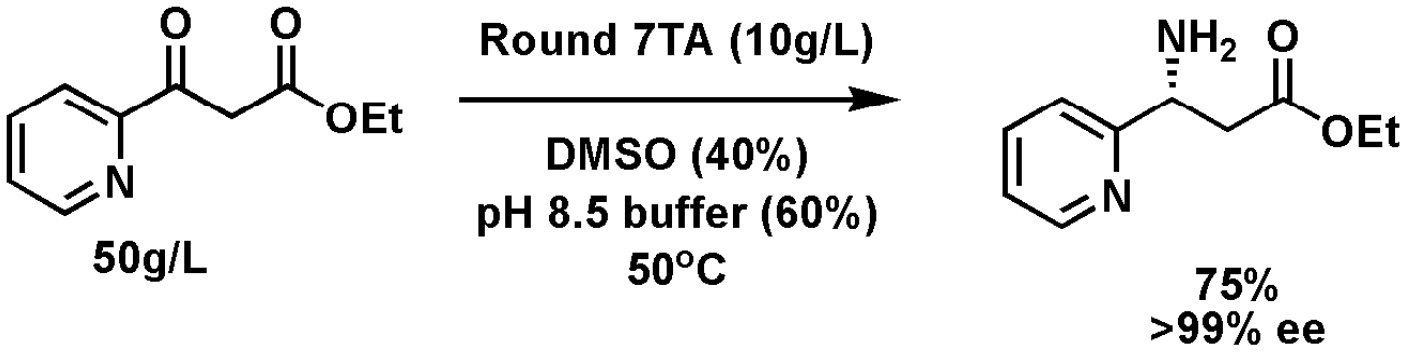
- Sterically/ Electronically unfavourable



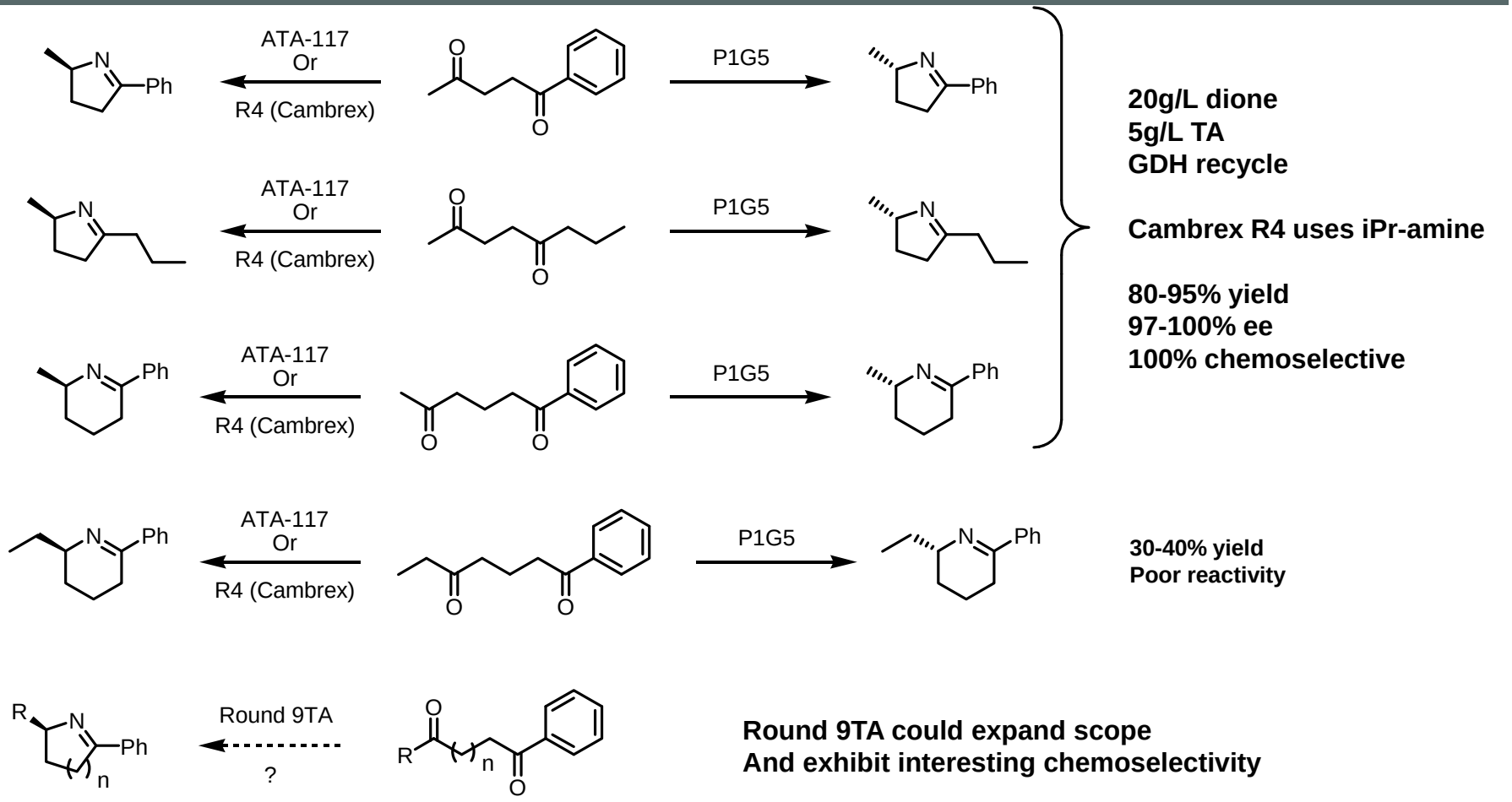
100% Conversion, >98% ee

Transamination of Non-Sitagliptin Ketones





Transamination of Diones



Broad Impact

Sitagliptin Transaminase enzyme is a general tool to convert most ketones to enantiopure R-amine…
A unique transaminase with wide application

Prior to 2008: Ketone to amine required 2-4 steps and ~2 weeks
Post 2008 (sitagliptin transaminase): Ketone to amine requires 1 step and 2 days

- **Accelerated Lead Op…Fast prep of >99%ee amines from easy to access ketones**
- **High success rate for making R-amines, but S-amines limited to methyl and cyclic ketones**
- **Predictable stereochemical outcome (high degree of confidence on absolute stereochem.)**
- **Large late stage impact…multiple programs use transaminases**

2008 top 200 branded drugs: 48/200 (24%) contain R-transaminase stereocenters
12/200 (6%) contain S-transaminase stereocenters
Combined: 60/200 (30%) could use a transamination

2008 top 200 generic drugs: 49/200 (24.5%) contain R-transaminase stereocenters
9/200 (4.5%) contain S-transaminase stereocenters
Combined: 58/200 (29%) could use a transamination

Acknowledgements

Merck Biocatalysis
Jeffrey Moore
Paul Devine
Greg Hughes
Birgit Kosjek
Brendan Grau
Krista Morley
Fred Flietz

Analytical Support
Naijun Wu
Kate Vogel
Yadan Chen
Frank Bernardoni
Catherine Lancaster (WP)
Gabe Graffius

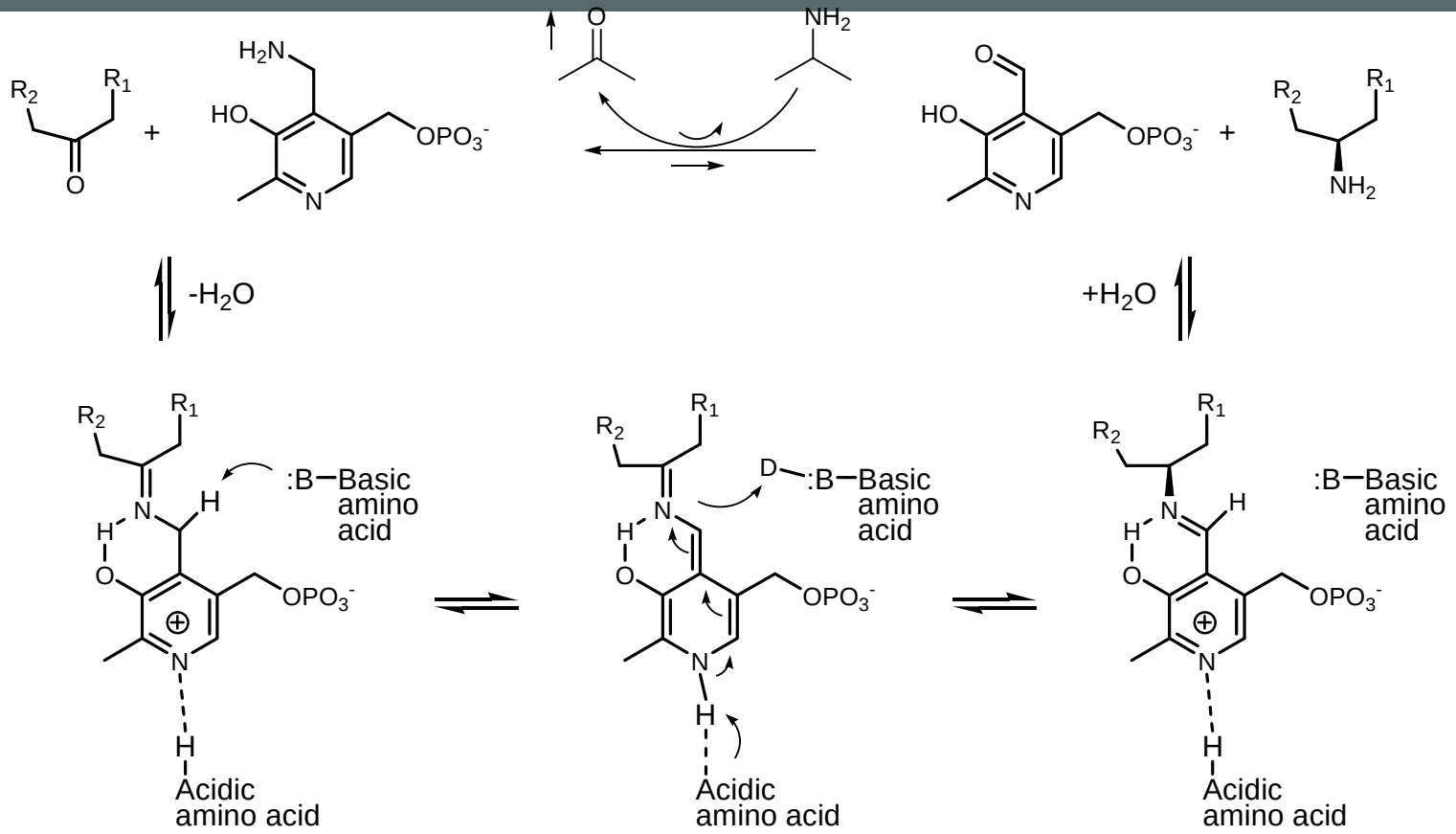
Management
Jos Brands
Skip Volante
Mahmoud Kaba (GTO)

Engineering support
Marguerite Mohan
Paul Fernandez
Sean Harrington
Alexei Kalinin
Jon Jurica
Chuck Orella
Zhihao Lin
Tseng-En Hu
George Zhou
Shane Grosser

Codexis R&D
Chris Savile
Emily Mundorff
Will Jarvis
Sarena Tam
Jeff Colbeck
John Munger
Anke Krebber
Lisa Moore
Gjalt Huisman

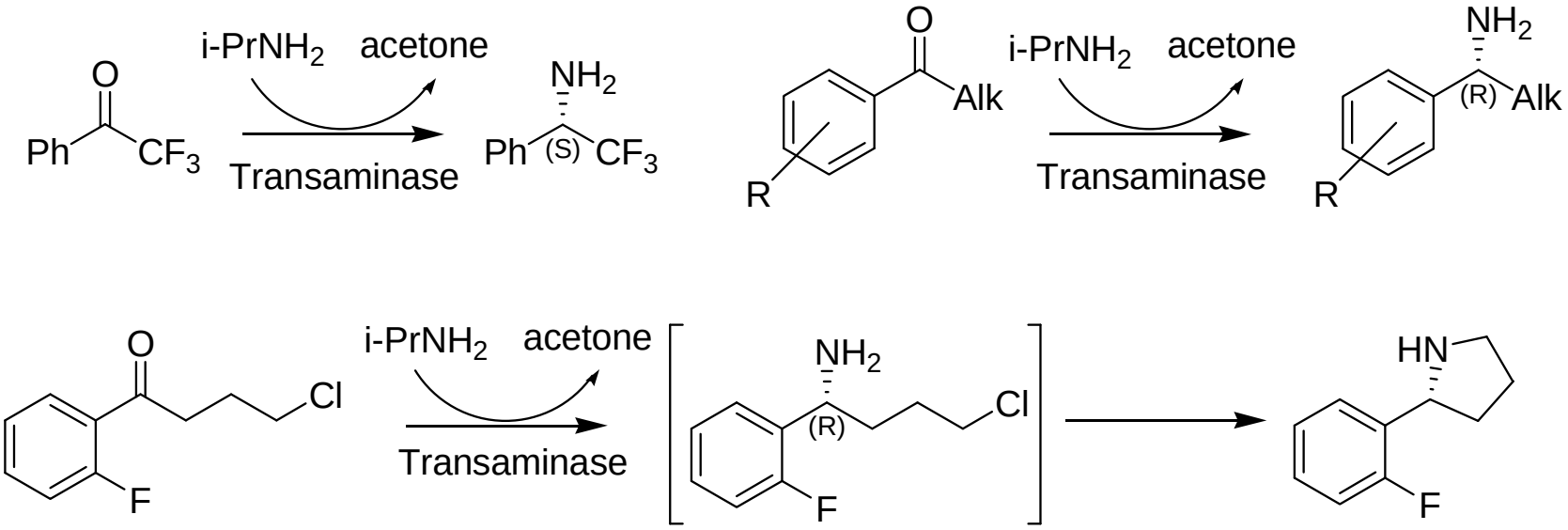
Codexis Operations
Carrine Ng
David Standish
Jon Postlethwaite
Bob Sato
Peter Seufer-Wasserthal
David Gray
Kevin Bishop

Mechanism

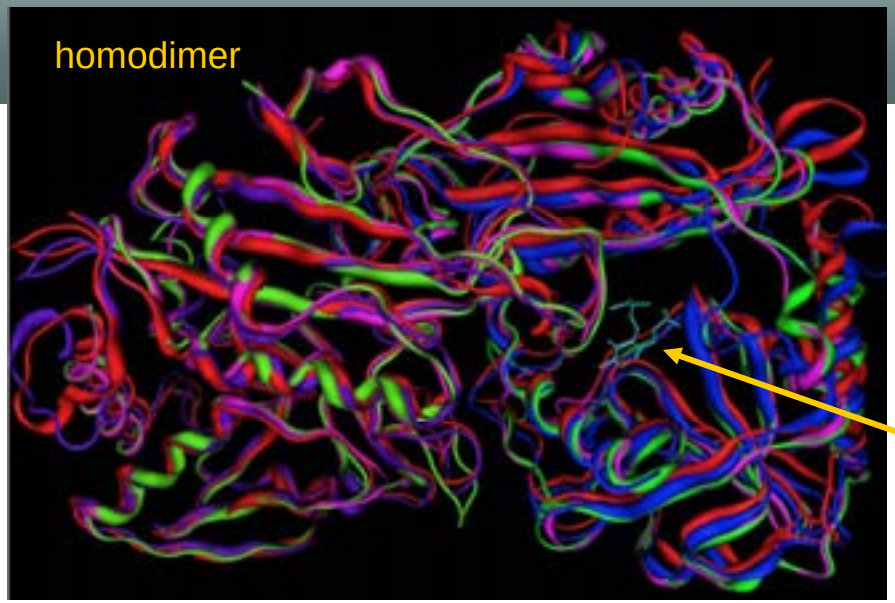

R2
R1
O
+
H2N
HO
OPO3-
N
O
NH2
HO
OPO3-
N
+
R2
R1
NH2
-H2O
+H2O
:B–Basic amino acid
H
Acidic amino acid
D
:B–Basic amino acid
H
Acidic amino acid
:B–Basic amino acid
H
Acidic amino acid

Other Examples

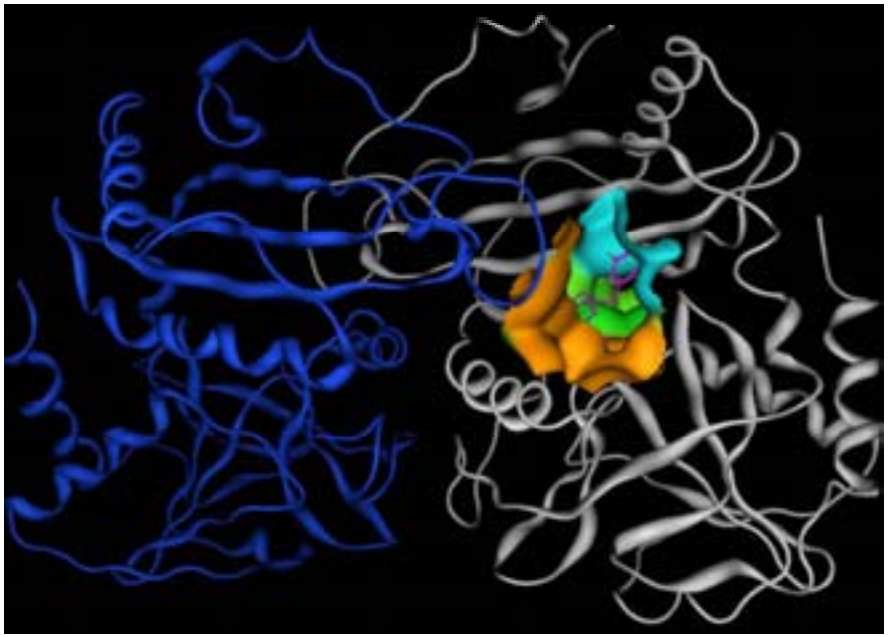
Previously Unreactive



Homology Model for ATA-117



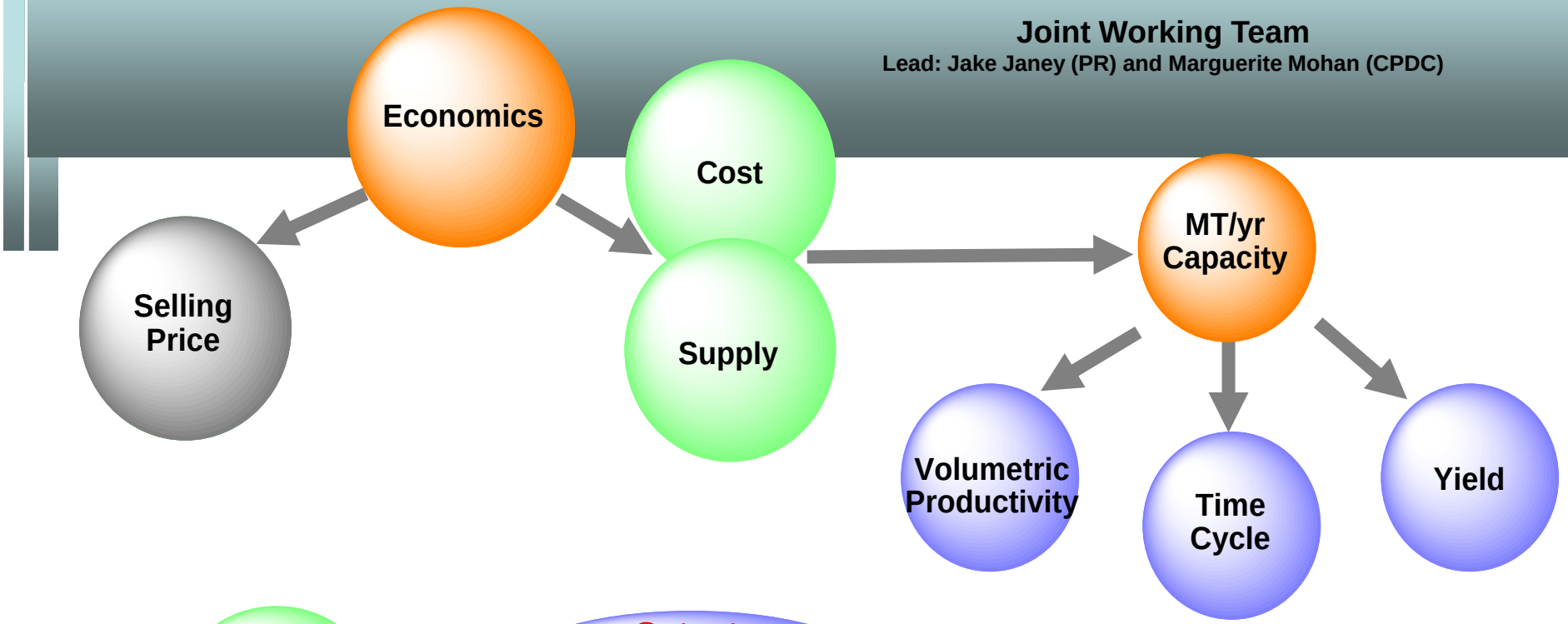
- Overlay alignment of backbone ribbon structures of homologs and homology model.
  - homolog 1
  - homolog 2
  - homolog 3
  - homology model

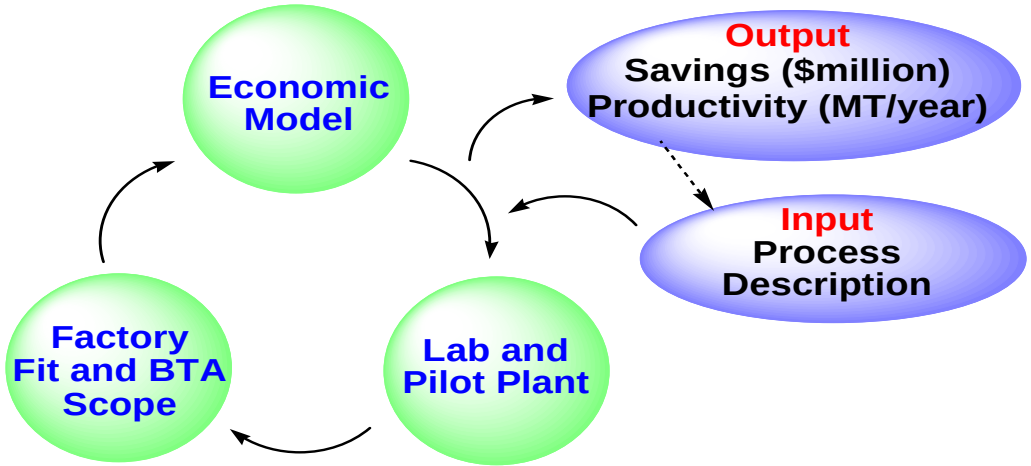


- Homology model ribbon structure showing surface of one binding site
  - Active sites are at subunit interfaces.
  - homology model subunits
  - large pocket
  - small pocket
  - PLP and catalytic residues

Co-Evolution of a Process and an Enzyme

Joint Working Team
Lead: Jake Janey (PR) and Marguerite Mohan (CPDC)



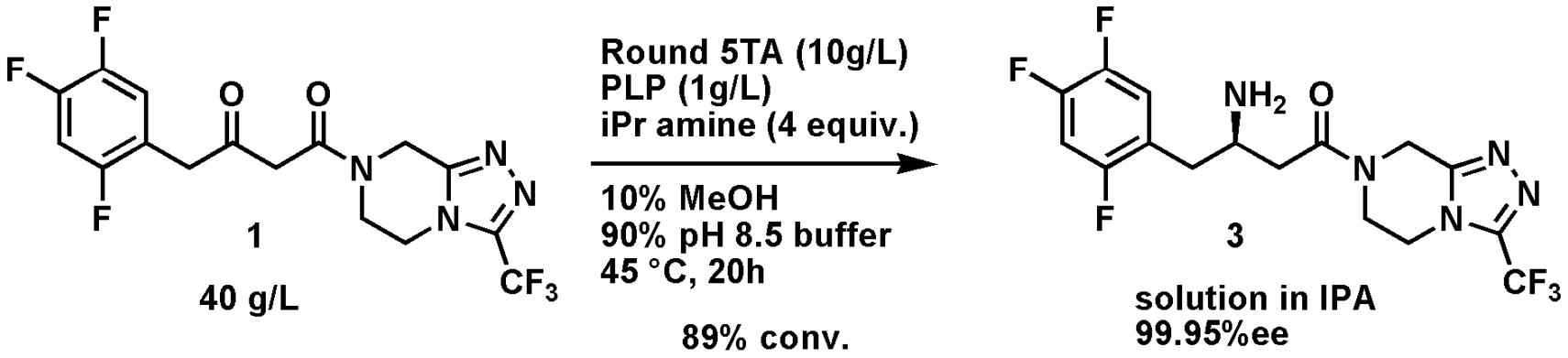


Basis for use of Process Engineering Model to Optimize Development by:

PC, CPDC, ADC/PAT, GPC Ops, GTO, EM, Procurement, Supply Chain, Financial, Process Eng, Codexis

Transamination
Initial Process Development

- **Process developments (Rounds 5-7):**



Codexis: 40g/L ketoamide **1** in 10% MeOH with 10 g/L round 5TA

Challenges	**Solutions**
• Oiling out of ketoamide **1**	• 50% DMSO rather than 10% MeOH
• Imine formation (**1+3**)	• Add ketoamide/DMSO solution over 3h
• pH control	• Control pH to 8.5 with 4M isopropyl amine
• loss of isopropyl amine	
• Conversion (89%)	• **Acetone removal**

QbD
Challenges

Original

Enamine Amide Hydrogenation → **Free Base Crystallization** → **API Pure Crystallization**

2006: First FDA approved QbD Filing

Assay Yield
% ee

Control Point:
Isolated Solid
- Impurities

Isolated Solid
- Purity
- Particle Size
- Yield

- Already Marketed…Need to match or exceed filed API Specs
- No new capital allowed…make it fit into the factory
- Must Increase factory productivity (MT/year)…not only quality, but must have better yield
- Through Process…complicates QbD as steps are interrelated

New

Ketoamide Transamination → **Work-up (Enzyme Removal)** → **API Pure Crystallization**

Through Process Reliant on Analytical and PAT

Assay Yield
React IR
pH

- Assay Yield
- Enzyme
- Inorganics (NaCl)
- Crude Stream Solvent (IPAc/IPA/DMSO/Water)

Isolated Solid
- Purity
- Particle Size
- Yield
- Residual Enzyme (N/D)

QbD
Transaminase Approach

- **Very large operating space identified in Pilot Plant and Lab**
  85% conversion in 24h gave passing quality API

- **Used the following productivity constraints as outputs in DOE to identify ideal operating space**
  - **>93.0% conversion (correlated to assay yield)**
  - **<12h reaction time**
  - **<0.1% on any new impurities**

- **Acetone removal drives reaction kinetics and is scale dependant (S/V, agitation, vacuum, sweep)**
  - **Ran DOE reactions “sealed” on multimax with condenser to be internally consistent**

- **Ran all experiments using same RM’s and one same 4-reactor multimax with continuous pH control (4M iPr-amine feed)…set feed rate in factory**

Lessons Learned

- **Initial Activity is not needed to evolve a manufacturing ready enzyme**
- **Design the enzyme to fit the process needs (or do it in parallel), i.e. evolve the catalyst to fit the chemistry needs (or factory fit)**
- **Productivity: volumetric efficiency (>0.4M) and time cycles (<20h)**
- **Use extractions: Enzyme partitions based on density of organic layer so use alcohols (IPA) to adjust organic density…a general work-up**
- **DOE works: Going from pH 8.5 to pH 9.8 cut time cycle in half (10h)**
- **Acetone removal is easy with vacuum and sweep, but you lose iPr-amine as well**
- **Fixed charges of base can eliminate pH probes and continuous feedback**

New Chemistry
Cross Organizational Challenges

Technical (Process Chemistry)

- Ketoamide solubility <1g/L in water and only 9-10g/L in DMSO/water
- Second phasing: free base, imine dimer, and ketoamide
- pH control: pKa product < isopropyl amine (+ loss to vapor phase)
- pH probe fouling
- Acetone removal needed to drive equilibrium
- Enzyme removal during work-up (emulsion issues and regulatory)

No Precendence Or Procedure for New Chemistry On Marketed Product

Bioanalytical

- Enzyme tracking and removal
- Enzyme supply, storage, "assay"
- Enzyme characterization

Logistical

- Supply chain
- Validation strategy
- PVE's
- Procurement
- Capacity/Demand planning

Engineering

- Scalability
- Work-up (filter vs. extract)
- Acetone removal (S/V vs. pressure and sweep)
- Fit, time cycles, yields…productivity
- Through process QbD
- Capital Constraints

Analytical

- Reaction sampling
- Rapid enzyme assay
- Acetone monitoring (PAT)
- pH monitoring (PAT)
- Crude into pure stream (solvents)

Regulatory

- Enzyme Spec?
- Enzyme in "Final Step"
- API characterization and stability
- Drug Product
- Filing Strategy

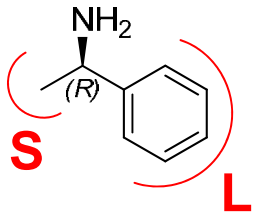
Lead Team from: Factory Site, Procurement, Regulatory, Analytical, Engineering (co-lead), Supply Chain

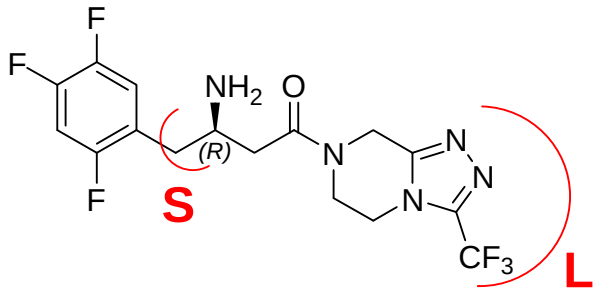
Problem and Approach

- Problem: **No** commercially accessible transaminase (*R* or *S*-selective), **nor** any Codexis in-house transaminase showed any detectable activity (LC/MS/MS) on the pro-sitagliptin ketone.
  - *Detectable initial activity is a prerequisite to directed evolution*

Problem and Approach

- Problem: **No** commercially accessible transaminase (*R* or S-selective), **nor** any Codexis in-house transaminase showed any detectable activity (LC/MS/MS) on the pro-sitagliptin ketone.
  - *Detectable initial activity is a prerequisite to directed evolution*
- Sitagliptin can be mapped to an *R*-selective transaminase (ATA-117):

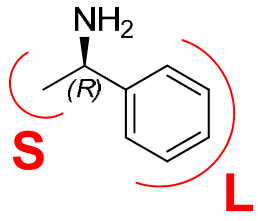


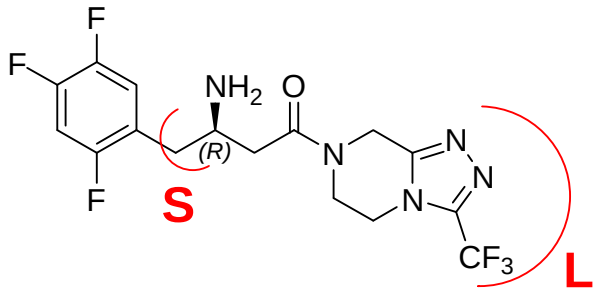


Problem and Approach

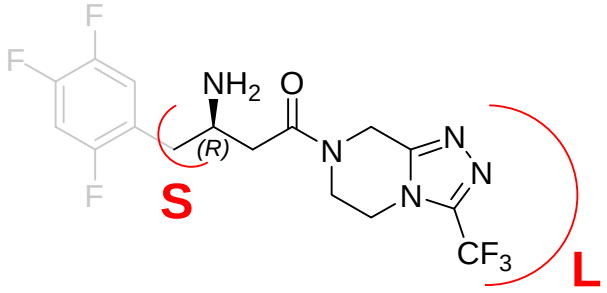
- Problem: **No** commercially accessible transaminase (*R* or *S*-selective), **nor** any Codexis in-house transaminase showed any detectable activity (LC/MS/MS) on the pro-sitagliptin ketone.
  - *Detectable initial activity is a prerequisite to directed evolution*

- Sitagliptin can be mapped to an *R*-selective transaminase (ATA-117):





- Approach: Evolve an *R*-selective transaminase on a truncated substrate that maps to established methyl ketone substrates, followed by further expansion of small pocket to accommodate F_3-phenyl :



Structure Guidance for ATA-117 Evolution

- To generate initial activity on pro-sitagliptin ketone, substantial reengineering of the ATA-117 active site would be needed.
- No tertiary structure for this enzyme, nor any close sequence homolog was available to identify binding site residues.

Structure Guidance for ATA-117 Evolution

- To generate initial activity on pro-sitagliptin ketone, substantial reengineering of the ATA-117 active site would be needed.
- No tertiary structure for this enzyme, nor any close sequence homolog was available to identify binding site residues.
- Structures of three distant homologous transaminases were reported.
  - Between 24-30% sequence identity to ATA-117.
  - Closest sequence identity among the three is <50%.

% homology	ATA-117	homolog 1	homolog 2	homolog 3
ATA-117	100	29.6	24.0	25.6
homolog 1		100	25.3	25.6
homolog 2			100	47.1
homolog 3				100

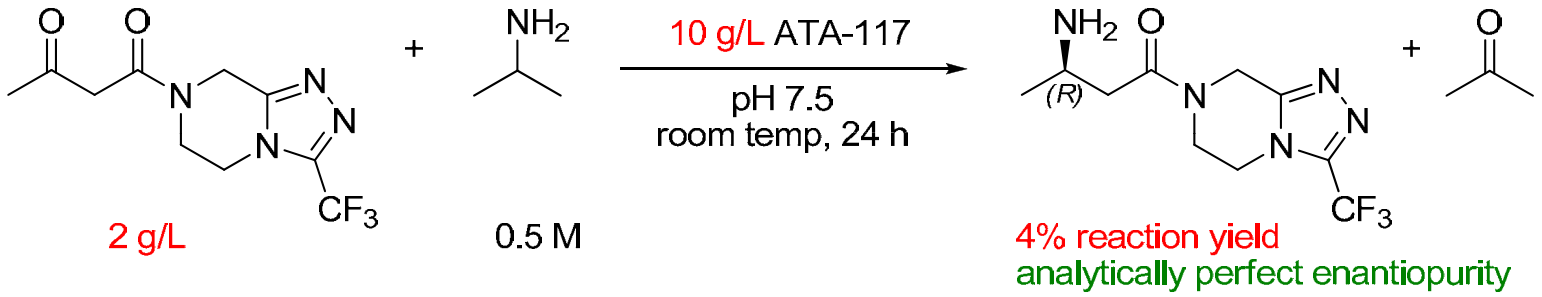
Structure Guidance for ATA-117 Evolution

- To generate initial activity on pro-sitagliptin ketone, substantial reengineering of the ATA-117 active site would be needed.
- No tertiary structure for this enzyme, nor any close sequence homolog was available to identify binding site residues.
- Structures of three distant homologous transaminases were reported.
  - Between 24-30% sequence identity to ATA-117.
  - Closest sequence identity among the three is <50%.

% homology	ATA-117	homolog 1	homolog 2	homolog 3
ATA-117	100	29.6	24.0	25.6
homolog 1		100	25.3	25.6
homolog 2			100	47.1
homolog 3				100

- The three reported tertiary structures closely overlap, allowing a predictive homology model to be built:

Initial Activity on Methyl Ketone Surrogate



- ATA-117 is a Codexis catalogue product that was used for this experiment.
- It is an unnatural, close homolog of a wild-type R-selective transaminase.

Establishing Activity on the Pro-Sitagliptin Ketone

- Evolution round 1a:
  - Site saturation libraries of large pocket mutations screened on the methyl ketone.
  - Identified multiple single mutants with improved activity.
  - A key single mutation gave *11-fold* greater activity over ATA-117.

Establishing Activity on the Pro-Sitagliptin Ketone

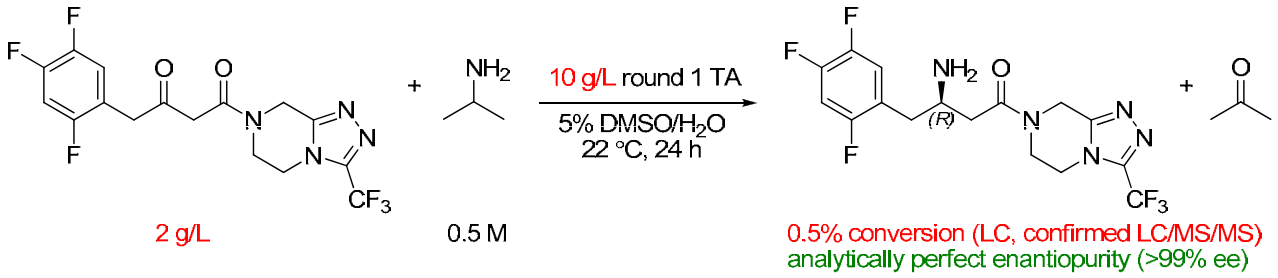
- Evolution round 1a:
  - Site saturation libraries of large pocket mutations screened on the methyl ketone.
  - Identified multiple single mutants with improved activity.
  - A key single mutation gave *11-fold* greater activity over ATA-117.

- Evolution round 1b:
  - Combinatorial libraries of small pocket mutations with smaller amino acids, in combination with the key large pocket mutation of round 1a were designed and generated.

Establishing Activity on the Pro-Sitagliptin Ketone

- Evolution round 1a:
  - Site saturation libraries of large pocket mutations screened on the methyl ketone.
  - Identified multiple single mutants with improved activity.
  - A key single mutation gave *11-fold* greater activity over ATA-117.

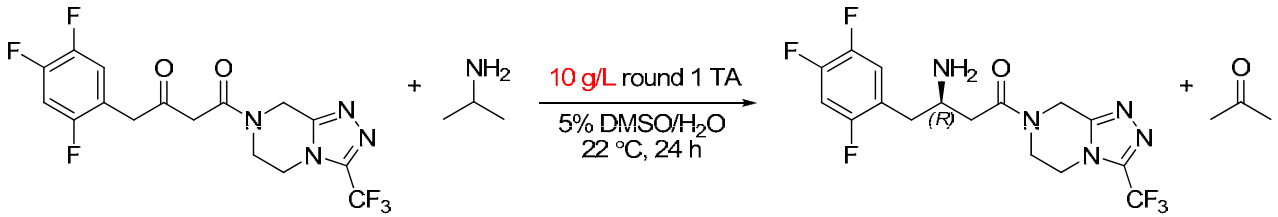
- Evolution round 1b:
  - Combinatorial libraries of small pocket mutations with smaller amino acids, in combination with the key large pocket mutation of round 1a were designed and generated.
  - A combination of three small pocket mutations, in combination with the key large pocket mutation provided the first detectable activity on the pro-sitagliptin substrate:



Establishing Activity on the Pro-Sitagliptin Ketone

- Evolution round 1a:
  - Site saturation libraries of large pocket mutations screened on the methyl ketone.
  - Identified multiple single mutants with improved activity.
  - A key single mutation gave *11-fold* greater activity over ATA-117.

- Evolution round 1b:
  - Combinatorial libraries of small pocket mutations with smaller amino acids, in combination with the key large pocket mutation of round 1a were designed and generated.
  - A combination of three small pocket mutations, in combination with the key large pocket mutation provided the first detectable activity on the pro-sitagliptin substrate.



  - No activity detected with the small pocket mutations in the absence of the key large pocket mutation.

Evolution for Process Fitness

- Evolution rounds 3-9 focused on increasing enzyme activity and in-process stability.
  - Generated, sorted (ProSAR), and recombined mutations across the whole protein.
- Successive rounds were screened under increasingly challenging conditions: substrate loading, iPM concentration, co-solvent, pH and temperature.

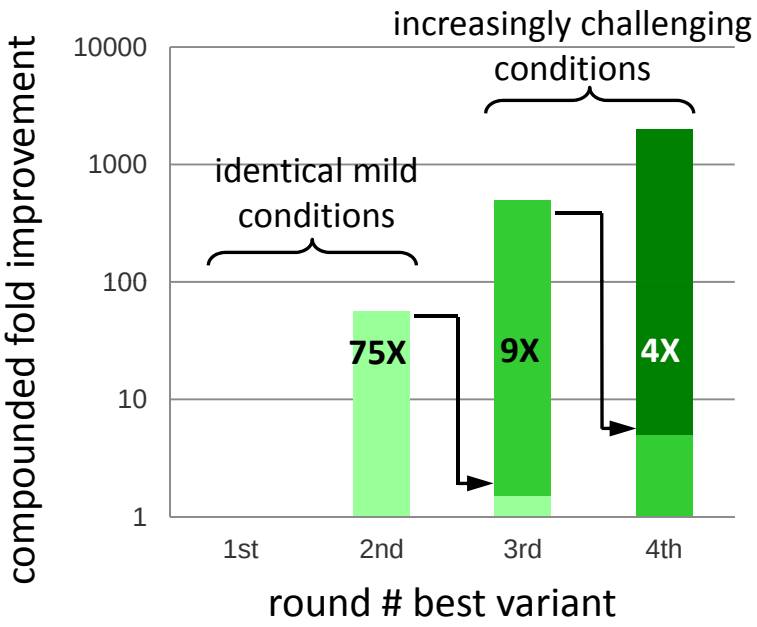
Evolution for Process Fitness

- Evolution rounds 3-9 focused on increasing enzyme activity and in-process stability.
  - Generated, sorted (ProSAR), and recombined mutations across the whole protein.
- Successive rounds were screened under increasingly challenging conditions: substrate loading, iPM concentration, co-solvent, pH and temperature

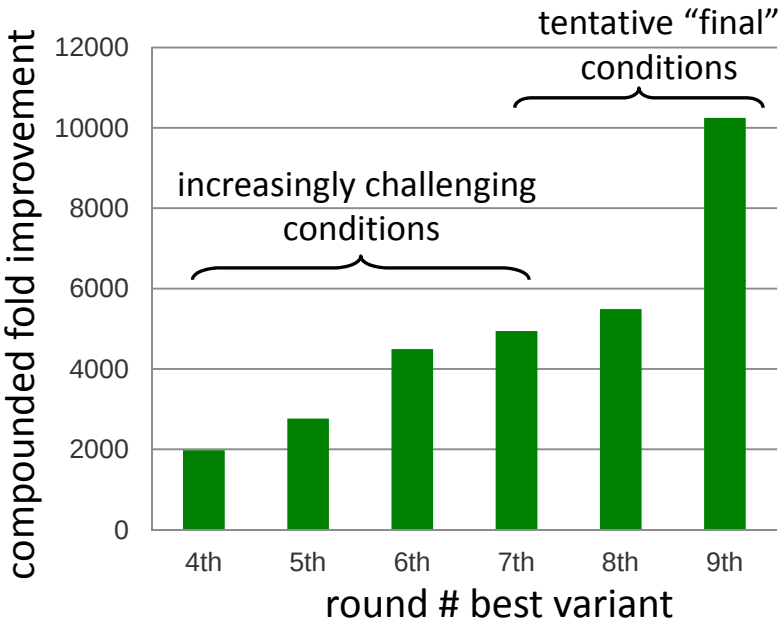
Round #	1 and 2	3	4	5	6	7-9
substrate g/L	2	5	10	40	100	100
[iPM], M	0.5	0.5	1.0	1.0	1.0	1.0
cosolvent	5% DMSO	5% MeOH	5% MeOH	10% MeOH	20% MeOH	25% DMSO
pH	7.5	7.5	8.5	8.5	8.5	8.5
temp, °C	22	30	30	45	45	45

Compounded Fold Improvements

rounds 1-4, log scale

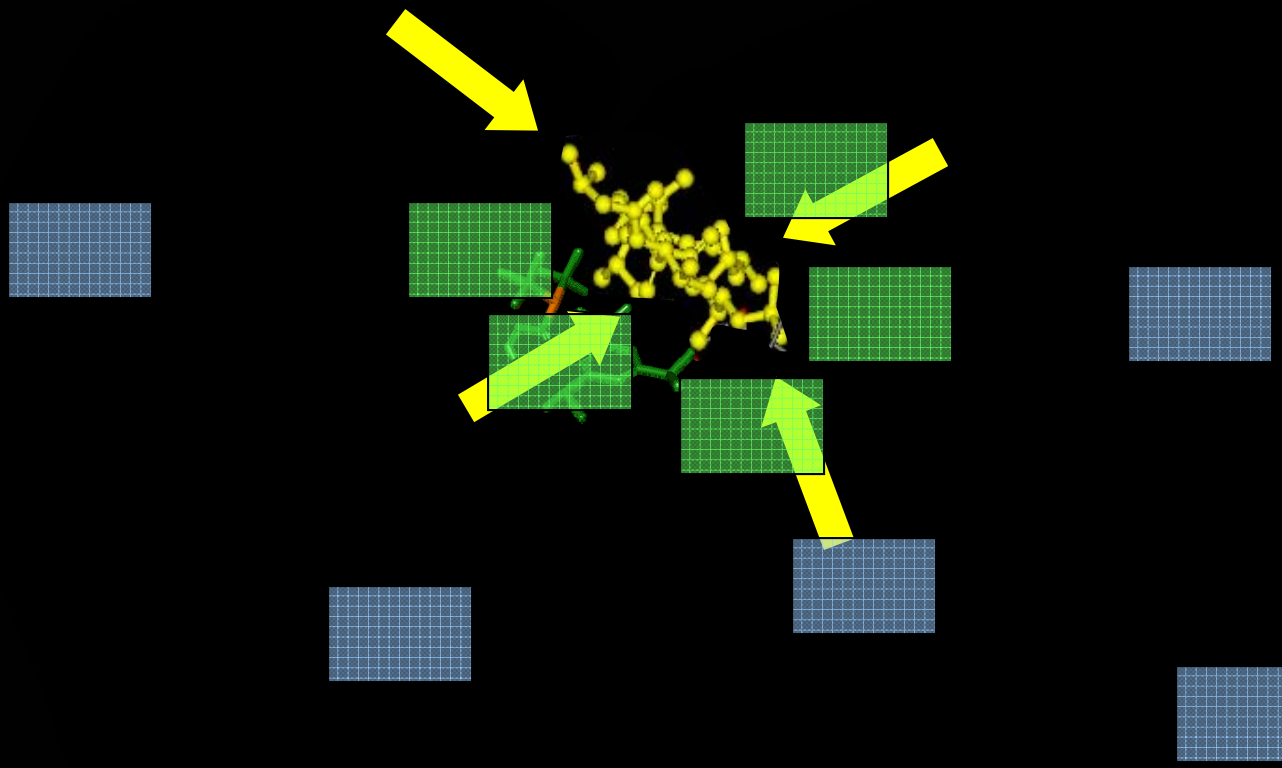


rounds 4-9, linear scale



- Rapid, exponential catalyst improvement initially.
- Evolutionary pressure was modified as understanding of the system improved.
- Improvements due to improved substrate binding, gene expression, thermostability, *in process* stability, and presumably other unknown factors.

Enzyme Catalysts



Renewable resources

Evolution

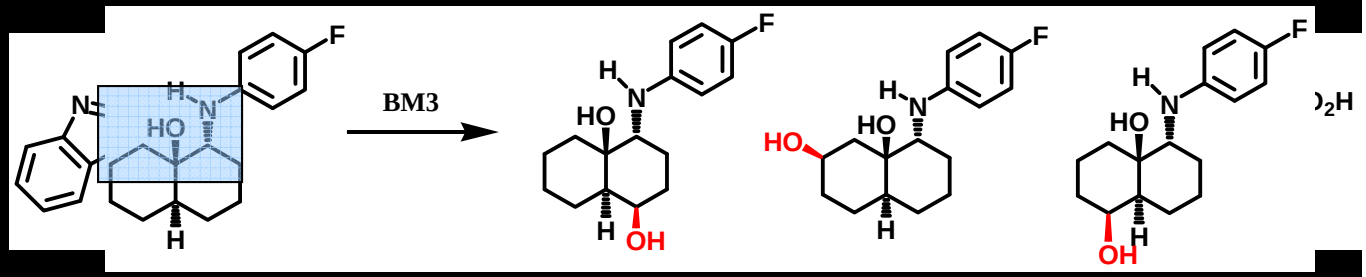
Increase stability

Biodegradable

Generate Activity

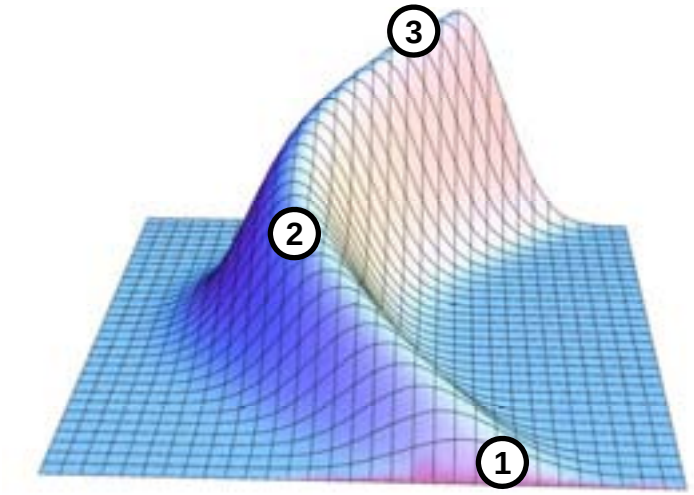
Mild Conditions

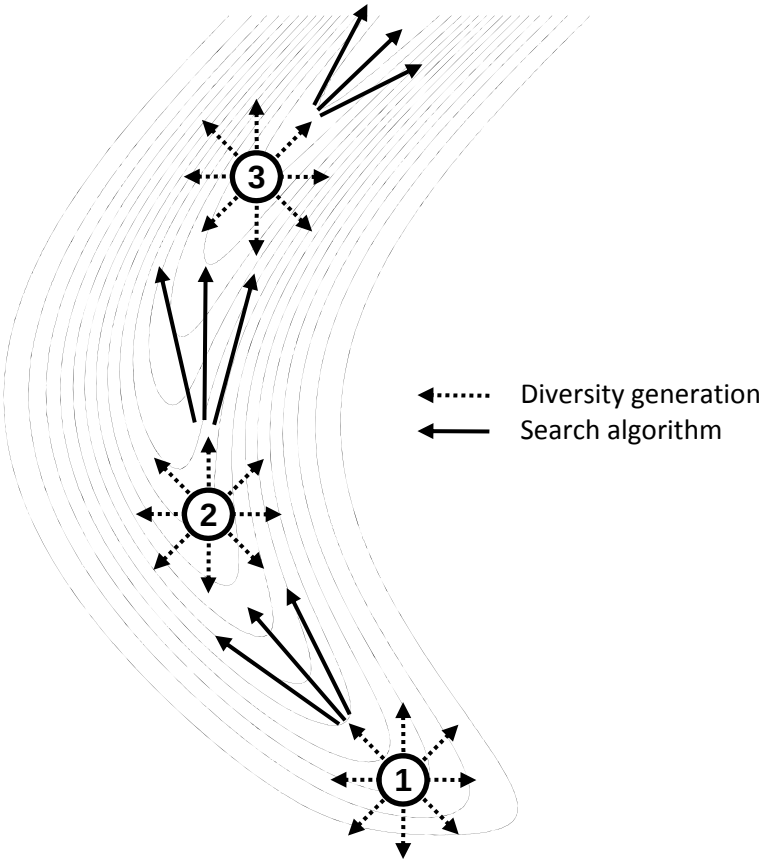
Enhance Selectivity



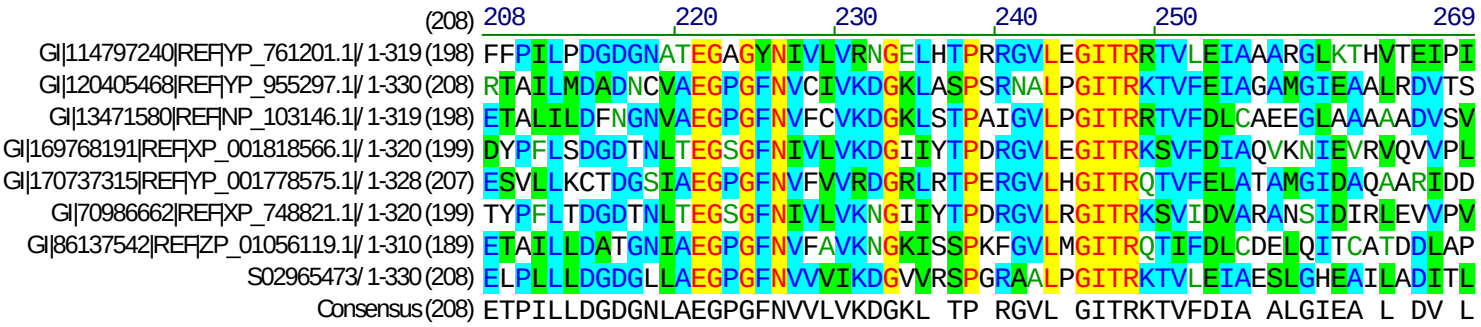
Guiding Principles for Enzyme Optimization

- Fitness function
- Diversity generation
- Search algorithm





Diversity Generation via CAPS

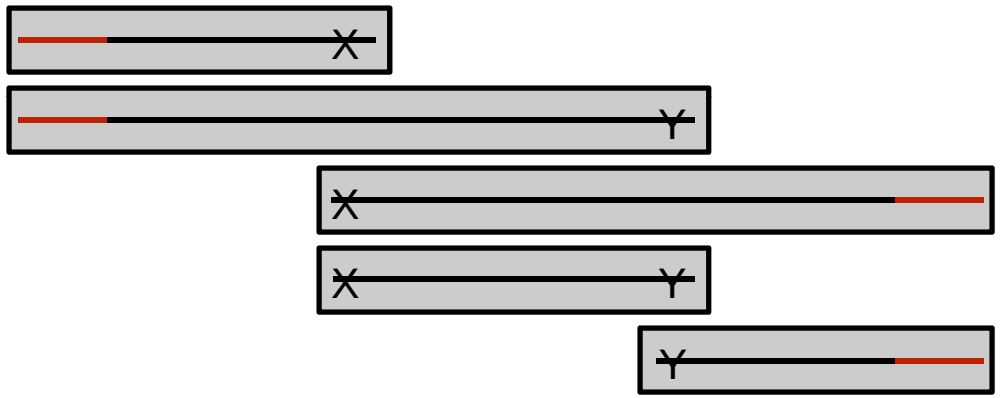


7 homologs, 36%-48% identity → Conservative diversity filter → 576 mutations, 6 plates

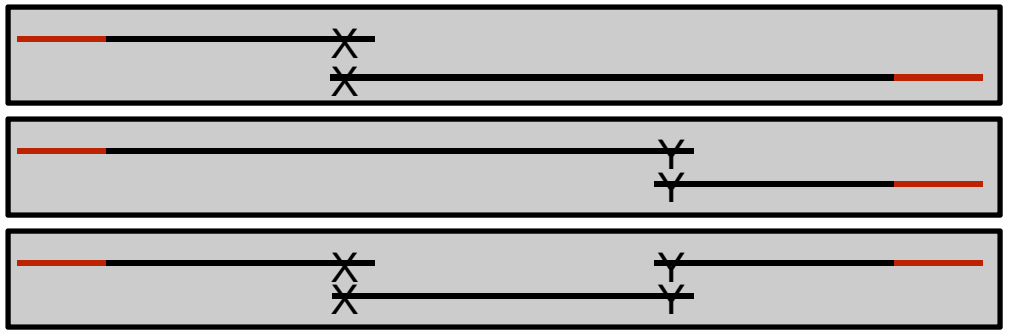
Automated Parallel SOEing (APS)

APS software automatically generates the location and sequence of primers. X and Y correspond to mutagenic primers.

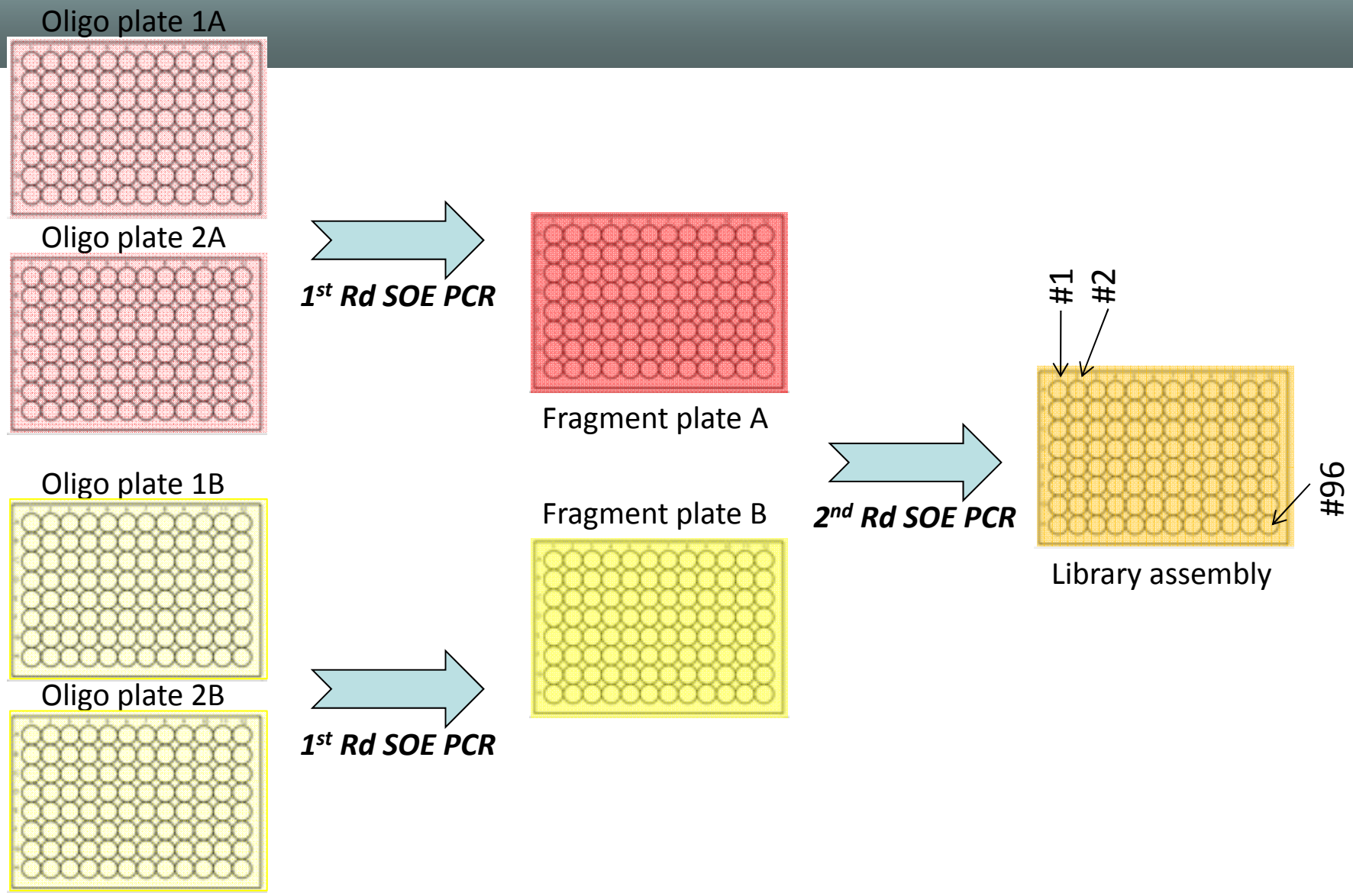
A script is written to dilute/mix primers, template and reactions for a 1st round of Splicing by Overlap Extension (SOE) PCRs to generate the necessary fragments.



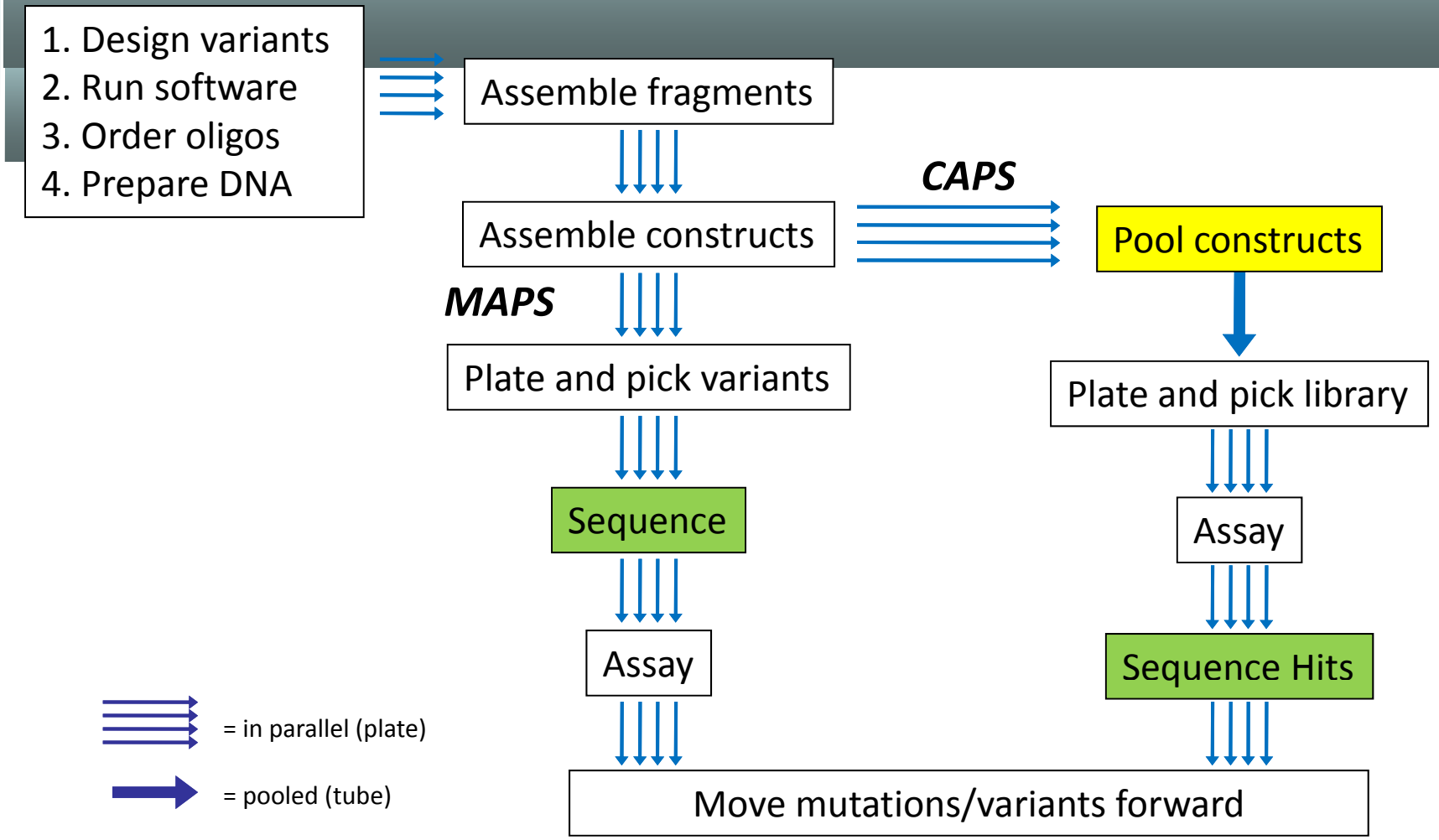
A second script is written to dilute/mix fragments for a 2nd round of SOE PCRs to generate the full length constructs.



CAPS Fragment Assembly



Combined APS (CAPS)



A Non-Standard Mode of Process Development

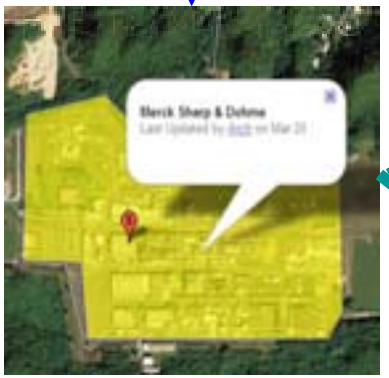
Standard approach used for Proof of Concept:

Lab

Pilot Plant

Factory

Revised approach used for 2nd Generation Process – where capital constraints are set:



Barceloneta Factory

PR and CPDC Labs

Rahway Pilot Plants

Process Engineering Model

QbD
Work-up Approach, Design, Results

- Through process dictates multiple responses in crude stream going into API

Critical Outputs (identified during scale-ups)

- Residual NaCl content…impacts API purity
- DMSO content…causes high ML loses in pure
- Enzyme residue…detectable down to 0.001wt.% unknown regulatory impact

Pure is run in IPA/water with phosphoric acid, hence need to know IPA/water content as well

Inputs

Factors	Units	Low	High
A: Ext. 1 IPA:IPAc Ratio	v:v%	40.0	50.0
B: Ext. 1 Org. Volume	L/kg	4.0	7.0
C: Ext. 2 IPA:IPAc Ratio	v:v%	25.0	35.0
D: Ext. 2 Org. Volume	L/kg	3.0	6.0
E: NaOH Charge Amount	Eq. rel. to HCl	0.0	1.2
F: Reaction Concentration	g/L	160	210

Response

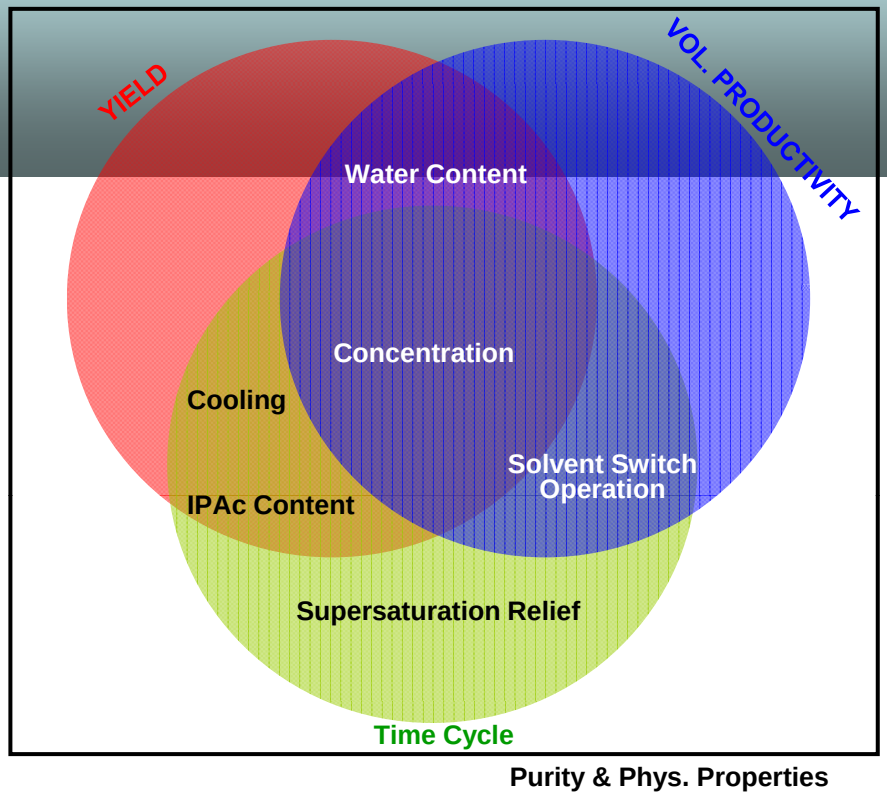
Responses	Units	Low	High
Na^+	m.eq.	0.0006	0.02
Cl^-	m.eq.	0.0006	0.06
DMSO	m.eq.	0.44	5.06
Enzyme	wt%	N.D.	N.Q.

Operational

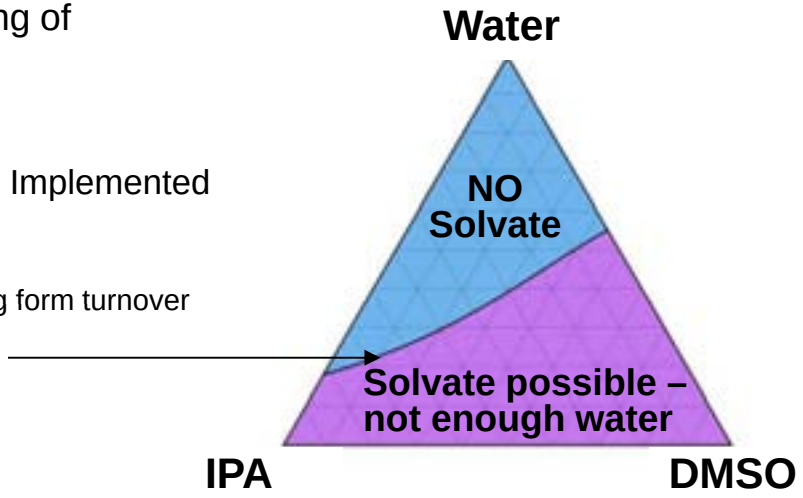
Factors Investigated	Recast Parameters	New Units	Low	High
A: Ext. 1 IPA:IPAc Ratio	Same	wt%	37.5	47.4
B: Ext. 1 Org. Volume	Ext. 1 IPA Charge	kg/kg KA	1.26	2.75
C: Ext. 2 IPA:IPAc Ratio	Same	wt%	23.1	32.7
D: Ext. 2 Org. Volume	Ext. 2 IPA Charge	kg/kg KA	0.59	1.65
E: NaOH Charge Amount	Same	Eq. rel. to HCl	0.0	1.14
F: Reaction Concentration	See Rxn document	-	-	-

Pure DOE Key Findings

- Purity Response
  - No statistical dependence on investigated factors
  - No further investigation planned
- Yield Response
  - Completely dominated by final supernatant KF
- Mv Response
  - PSD impacted by operational (vs. compositional) parameters only.
  - Investigated further by additional DOE experimentation
- Disso Temp Response
  - Strong function of solution composition
  - Investigated further by solubility map with the help of PAT group
- Form Response
  - No statistical dependence on investigated factors
  - Investigated further by thermodynamic modeling of water activity due to introduction of DMSO



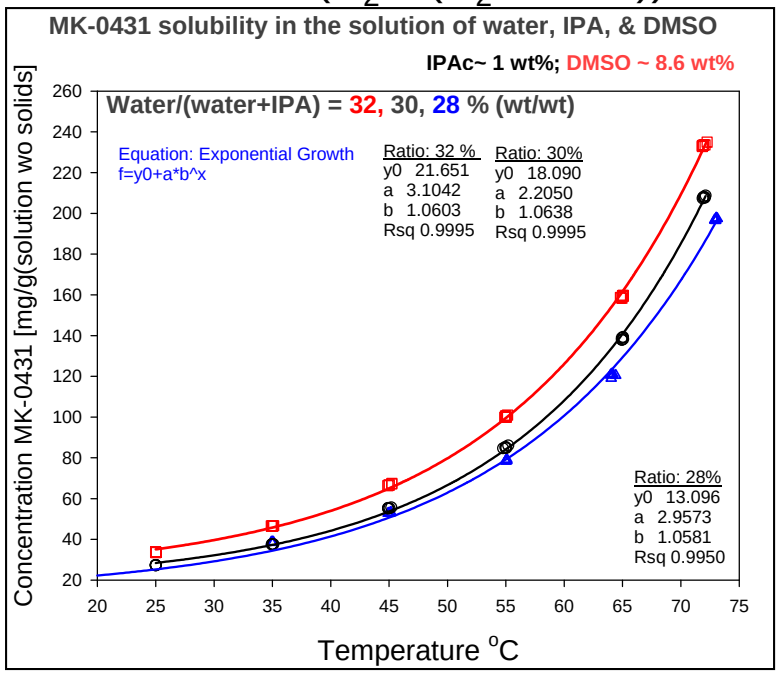
- Thermodynamic modeling of water activity
  - Unifac activity coefficient model - Implemented in Mathematica or excel
- Critical water activity determined by CMSE using form turnover studies (0.784@ 75ºC)



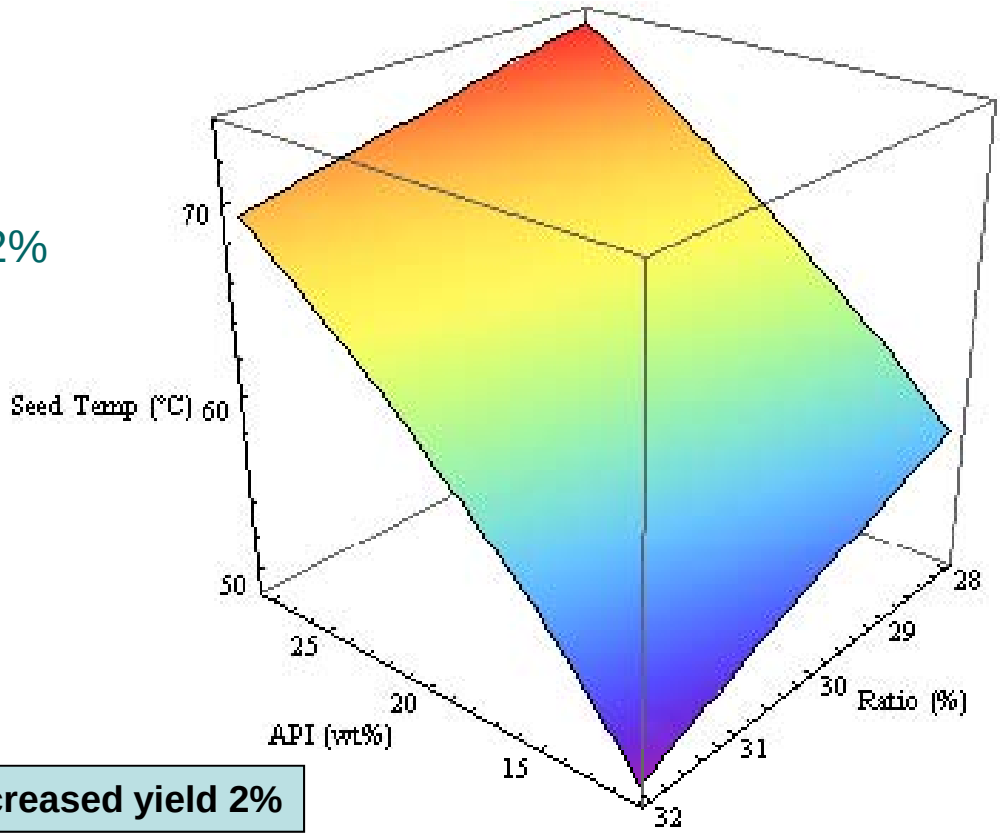
Empirical Model

$$\text{Seed Temp}(^{o}\text{C}) = 29.38 + \sqrt{2378.76 + 70.70\,(\text{wt\%}_{\text{API}}) - 89.39\left(\text{Ratio}_{\text{Water}/(\text{Water+IPA})}\right) + 16.61\,(\text{wt\%}_{\text{DMSO}})}$$

- Empirical model to predict disso temp:
  - Developed via regression of solubility data
  - Applicable Range-
    - **API =** 11.2 – 28.4 wt%
    - **DMSO** = 0 – 3.05 eq. DMSO
    - **IPAc =** 0 – 0.50 eq. IPAc
    - **Ratio** (H_2O/(H_2O +IPA)) = 28 – 32%



- Model Error analysis indicates prediction is ±1.2ºC.



Increased yield 2%

DMSO = 5 wt%

Januvia Transaminase Attributes

Ranges

- **Ketone Loading: Up to 275g/L in <24h**
- **Co-solvent: 0 - 60vol% DMSO (or MeOH)**
- **Temperature: 30 - 60 °C**
- **Enzyme Loading: 3 - 4.5 wt.% (~0.01 mol%)**
- **pH Range: 9 – 11 (no buffer)**

Operating

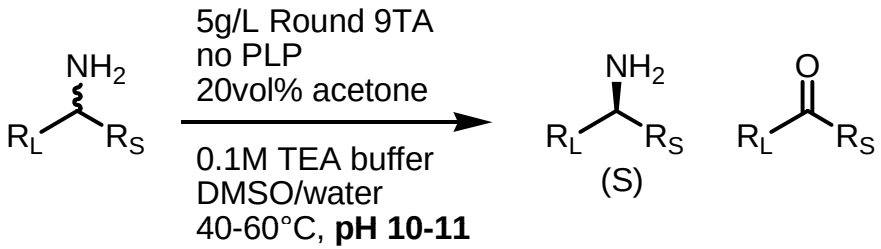
- **Ketone Loading: 250 g/L in 10h**
- **Co-solvent: 50vol% DMSO**
- **Temperature: 50 °C**
- **Enzyme Loading: 4 wt.% (~0.01 mol%)**
- **pH Range: 9.7 – 10 (no buffer)**

Kinetics

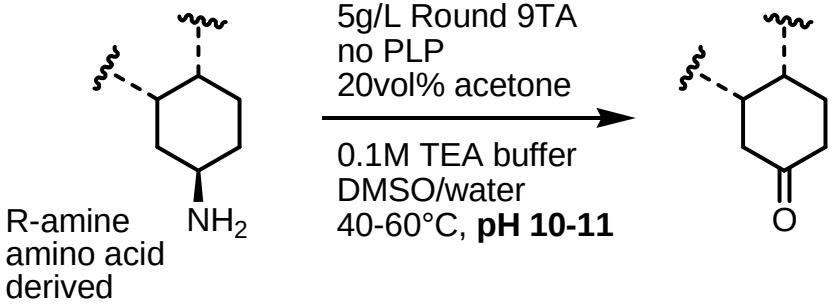
- **25 g L^{-1} h^{-1} productivity**
- **Ketone ~5 g/L solubility…mass transfer limited kinetics**
- **Likely diffusion limited (dilute runs)**

De-aminations

Resolution: *An alternative to S-transaminase*



For ketone preparation



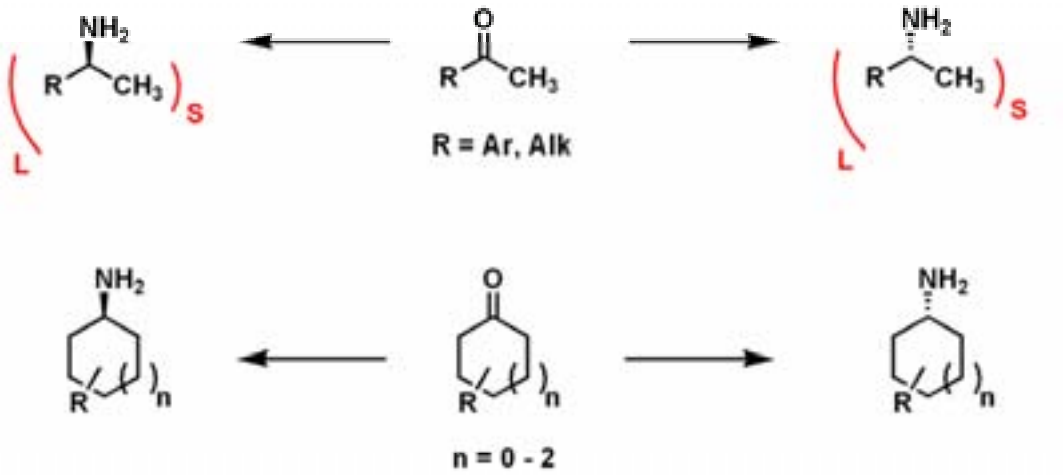
Performance vs. Process Targets

Parameter	Target	ATA-117	1st active variant	Round 9 variant	Final Round 11
substrate load, g/L	100	2	2	110	275
biocatalyst, g/L	≤5	10	10	5.5	6
reaction time, h	≤24	72	24	24	16
conversion, %	>98	0	0.5	90	95
enantioselectivity, e.e.	>99%	--	>99%	>99.9%	>99.9%
Productivity (g/L.hr.g_{TA})		0	$4.2*10^{-5}$	4	5-6

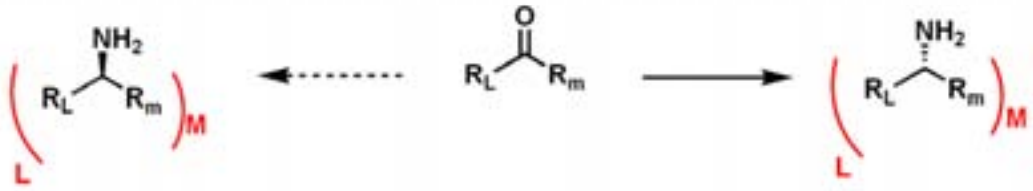
- **25,000-fold improvement from 1st to 11th round best variant.**
- **Infinite improvement from starting enzyme, ATA-117.**

Summary of New Transamination Capabilities

- **Previous Transaminase capabiliies**



- **New Transaminase capabiliies**



- **Most substrates are converted at higher substrate/ lower catalyst loadings**
- **Broader S-selective transaminase capabilities are currently under development**