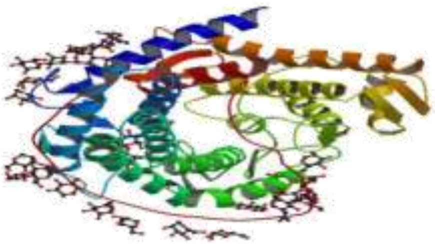
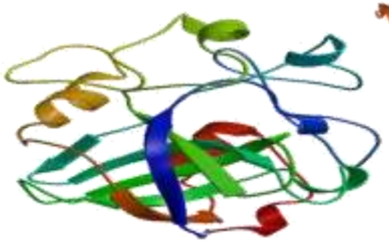
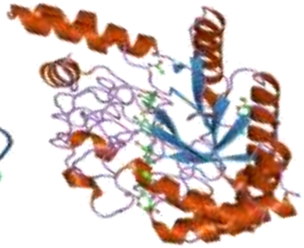
Engineering Cellulase Families for Biomass-to-Biofuel Conversion Processes



Prof. Pete Heinzelman
University of Oklahoma

United States Fossil Fuel Dependence

2010 United States gasoline consumption - 140 billion gallons[1]

1- http://www.eia.doe.gov/basics/quickoil.html

Environmental

Economic

Geopolitical

Biofuel Mandate

Federal government mandates
21 billion gallons of transportation
fuel from inedible biomass by 2022

Biofuel Production Gap

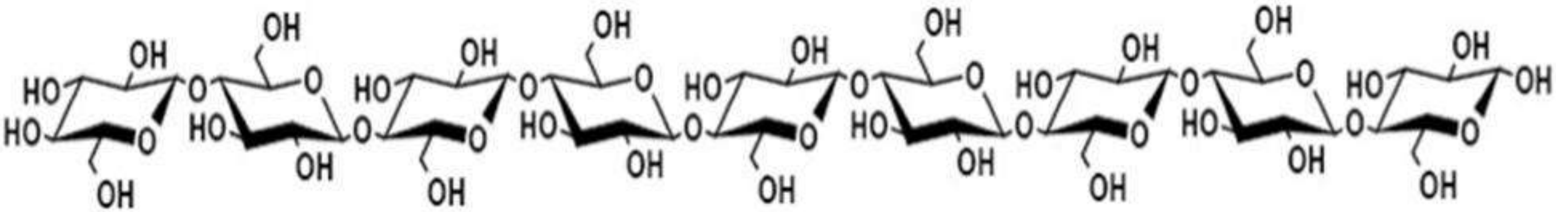
Biofuel from biomass
2010 production-
<100 million gallons[2]

2-http://en.wikipedia.org/wiki/Cellulosic_ethanol_commercialization

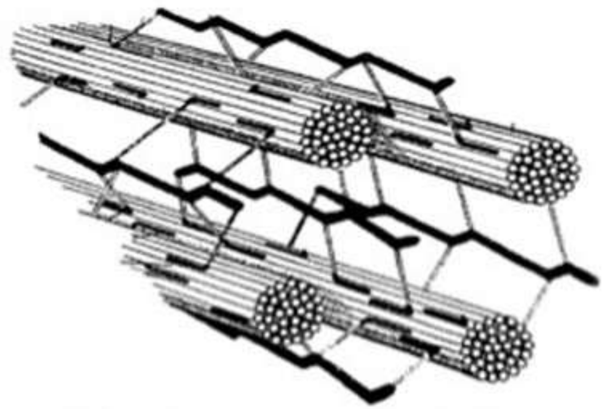
Cellulosic Biomass

Waste material or "energy crops"

Cellulose Structure

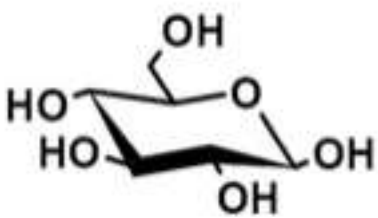


Cellulose: Chains contain ~10,000 β-1,4 linked glucose subunits



Cellulose chains associate into bundles

Cross-linked by hemicellulose (5C & 6C sugars) and lignin (phenols)

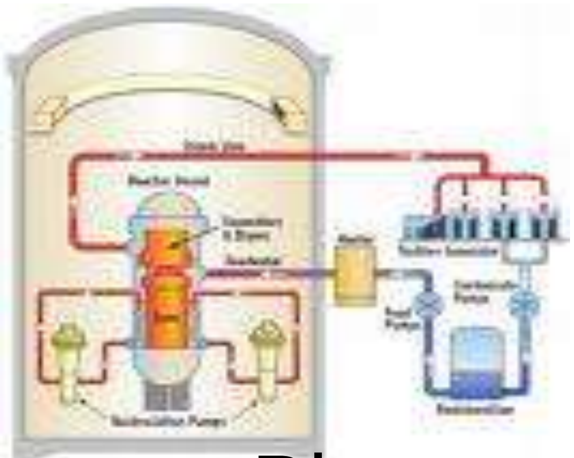


Glucose - Feedstock for biofuel production

http://www.cchem.berkeley.edu/mmargrp/research/Cellulase/Cellulose.jpg

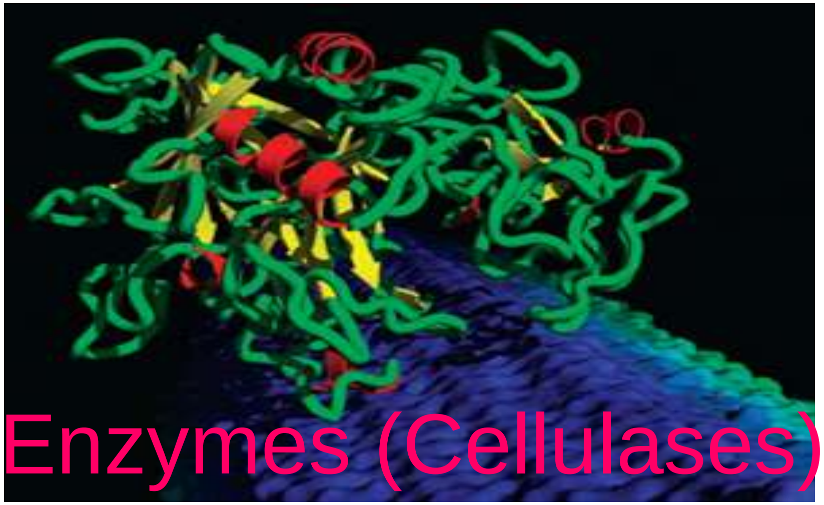
Biomass To Biofuel

Cellulosic Biomass



Pretreat Biomass -

Heat, acid/base, pressure



Glucose

Cellulase Cost Challenge

High cost of cellulases
is a major limiting factor

Estimated enzyme cost - 50 cents/gallon[3]

DOE target cost - 10 cents/gallon

Five-fold cellulase cost reduction

3- Wilson D, 2009

Cellulase Production

Hypocrea jecorina

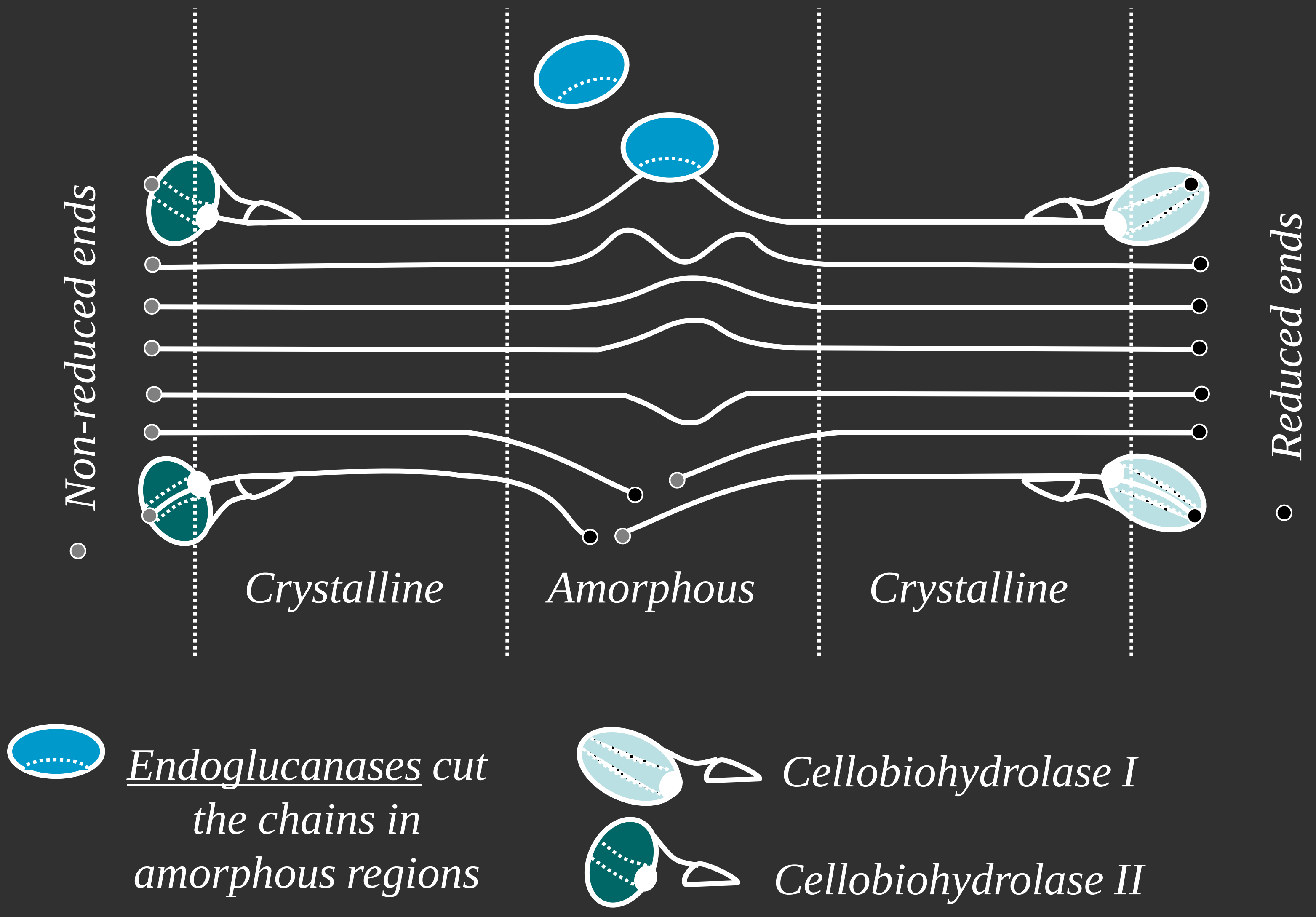
Filamentous, mesophilic fungus

Derives sugar from plant matter

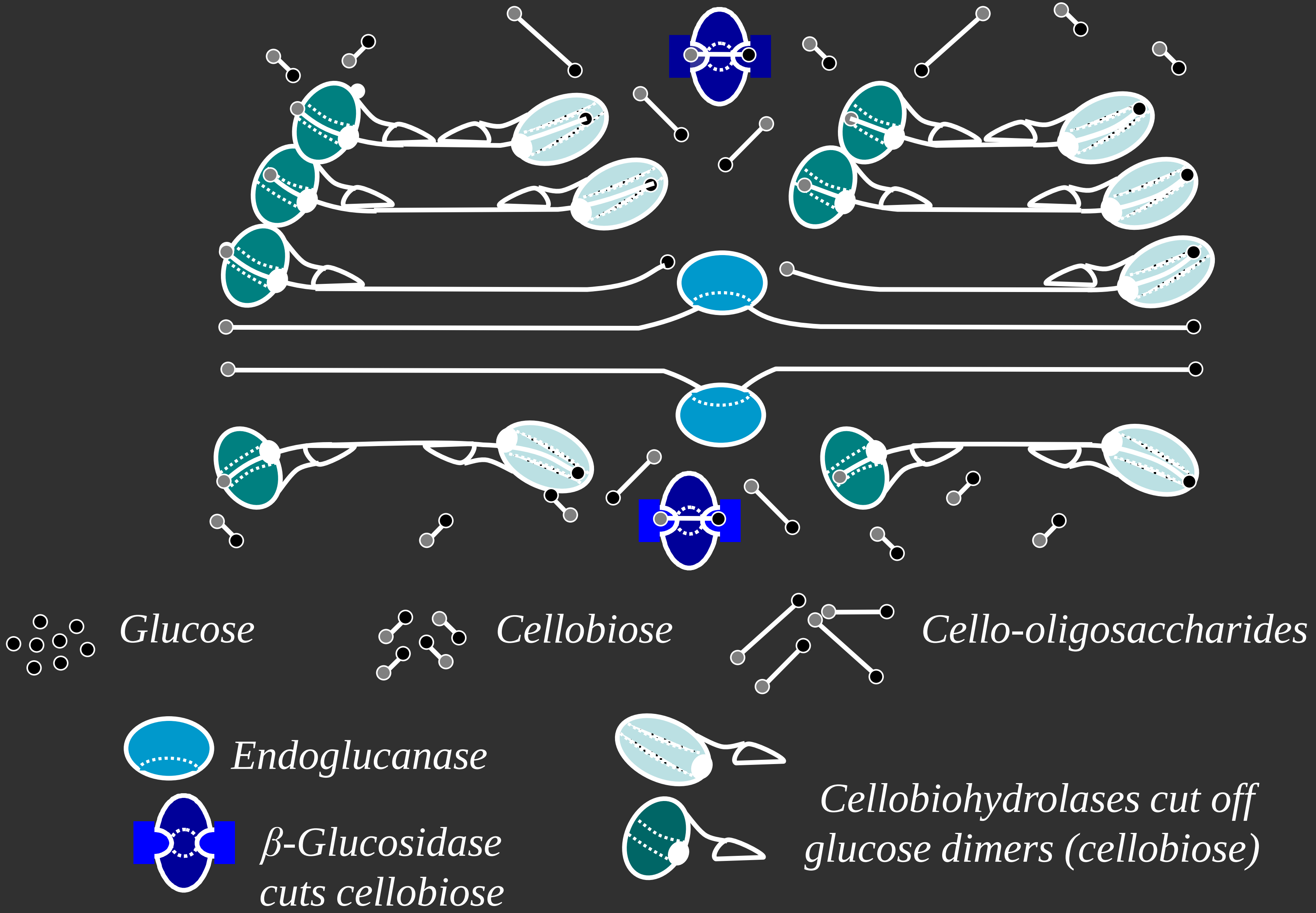
Cultured to secrete high cellulase titers (100g/L)

Large decreases in cellulase production cost unlikely

Cellulases Act Synergistically



Cellulases Act Synergistically

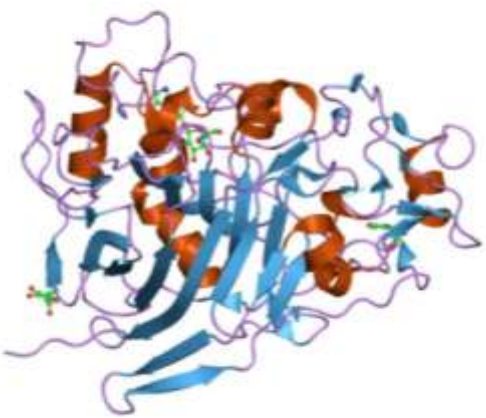


From J.McBride

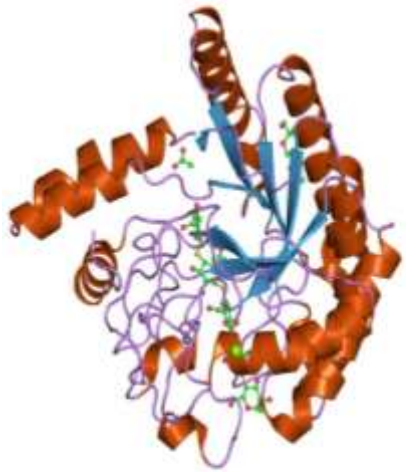
Multidimensional Challenge

What synergies are most important?

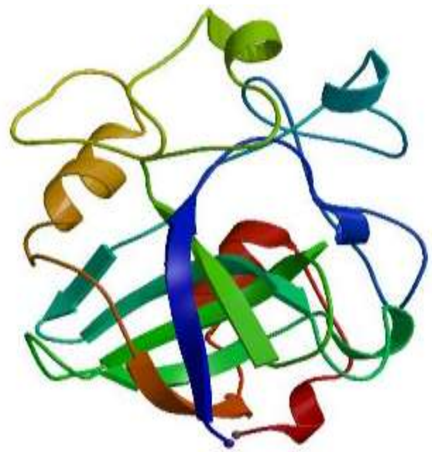
What enzyme properties limit cellulose hydrolysis?



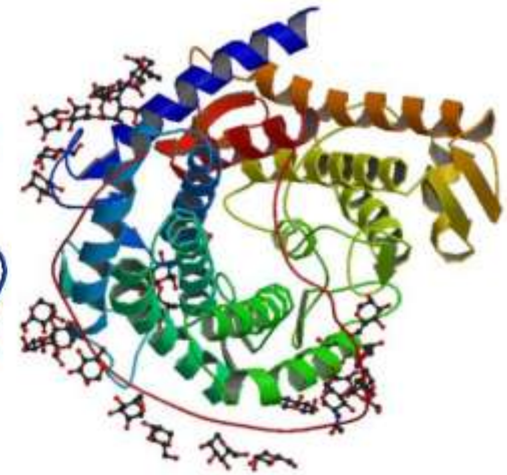
Cellobiohydrolase Class I



Cellobiohydrolase Class II



Endoglucanase



β-Glucosidase

Need holistic approach to cellulase improvement

Biomass Variability

Different feedstocks & processing configurations

Need to optimize cellulases for every process

Cellulases Are Slow

Hydrolyze cellulose on order 1 bond/sec at 50 ℃

Hydrolysis processes run for days

Increase hydrolysis temperature - Arrhenius effect

Thermostability limits hydrolysis temperature (~65 ℃)

Increase cellulase stability-

Leverage Arrhenius effect

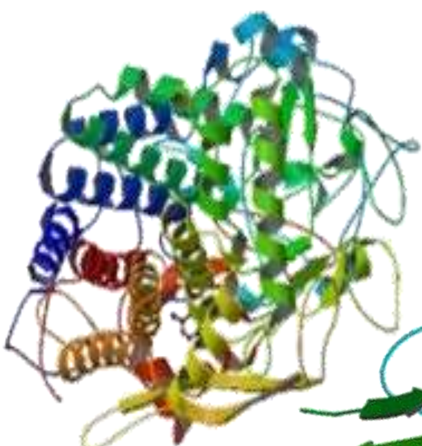
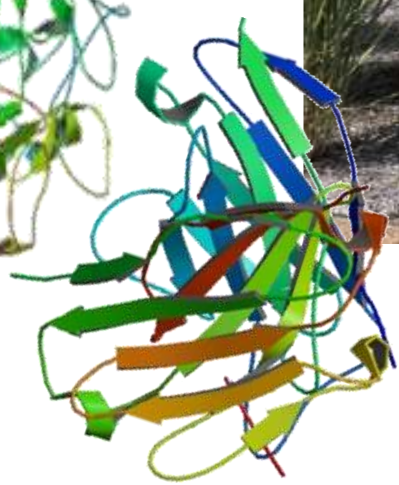
Longer enzyme lifetime at higher temperature

Cellulase Challenges

Different classes of cellulases act synergistically

Biomass feedstock & process variability

Cellulase thermostability

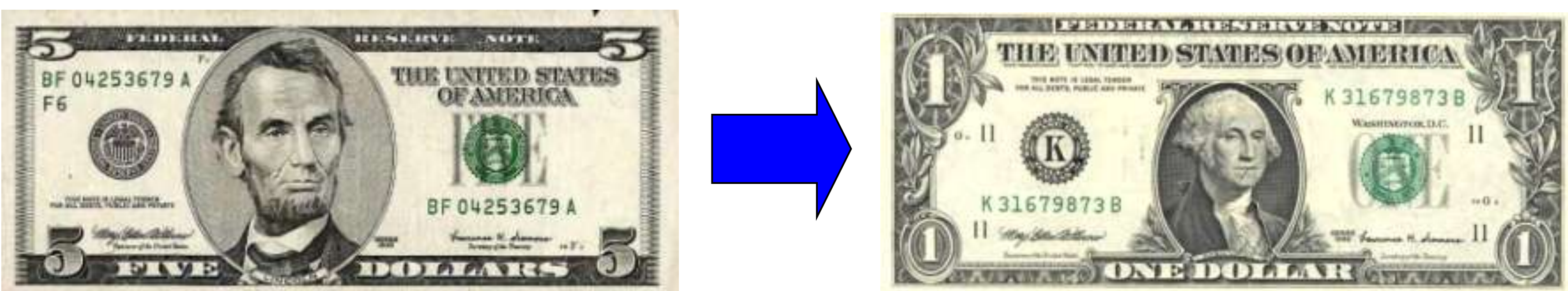



Motivation For Cellulase Families

Families containing thousands of thermostable cellulases with different functional properties -

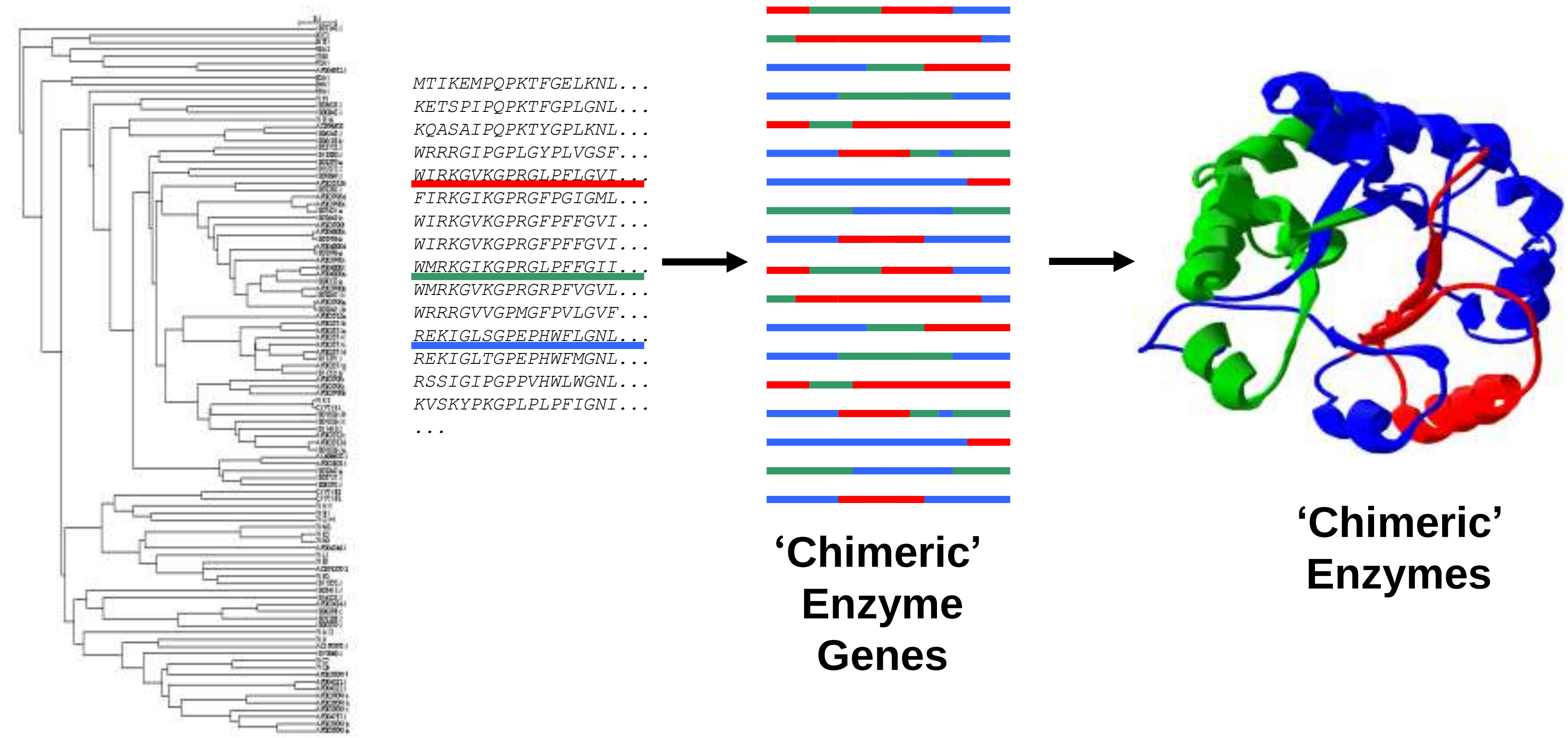
Application-specific, optimized cellulase mixtures

Target: Five-fold cellulase cost reduction



Structure-Guided Recombination: Enzyme Chimera Families

Recombine “blocks” of amino acids from related enzyme genes - Enzyme families containing thousands of active chimeras



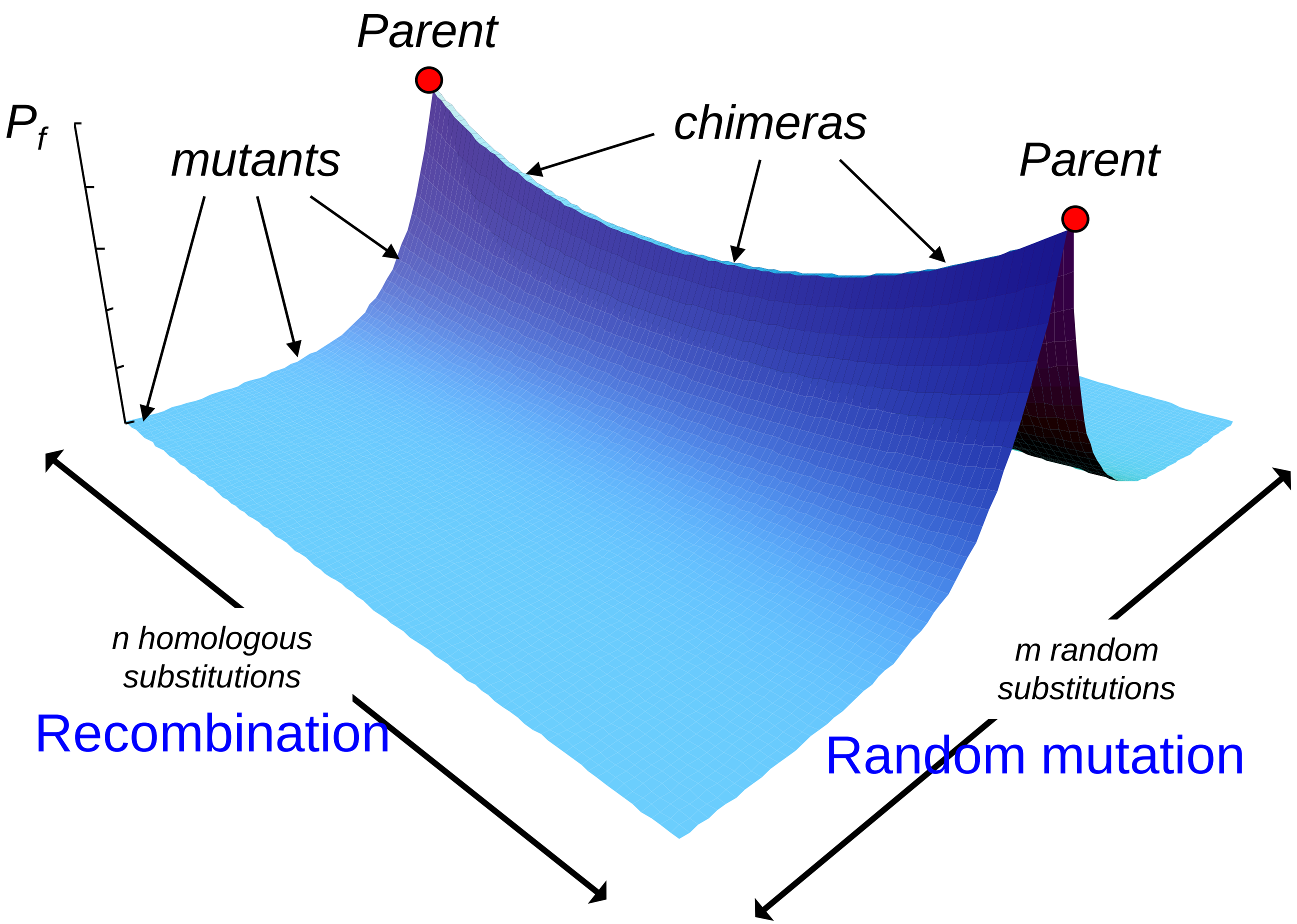
Chimeric enzymes contain dozens of mutations - Property diversity

Natural Mutations

Recombine natural enzymes - Mutations "approved" by nature

Enzyme chimeras with dozens of mutations retain activity

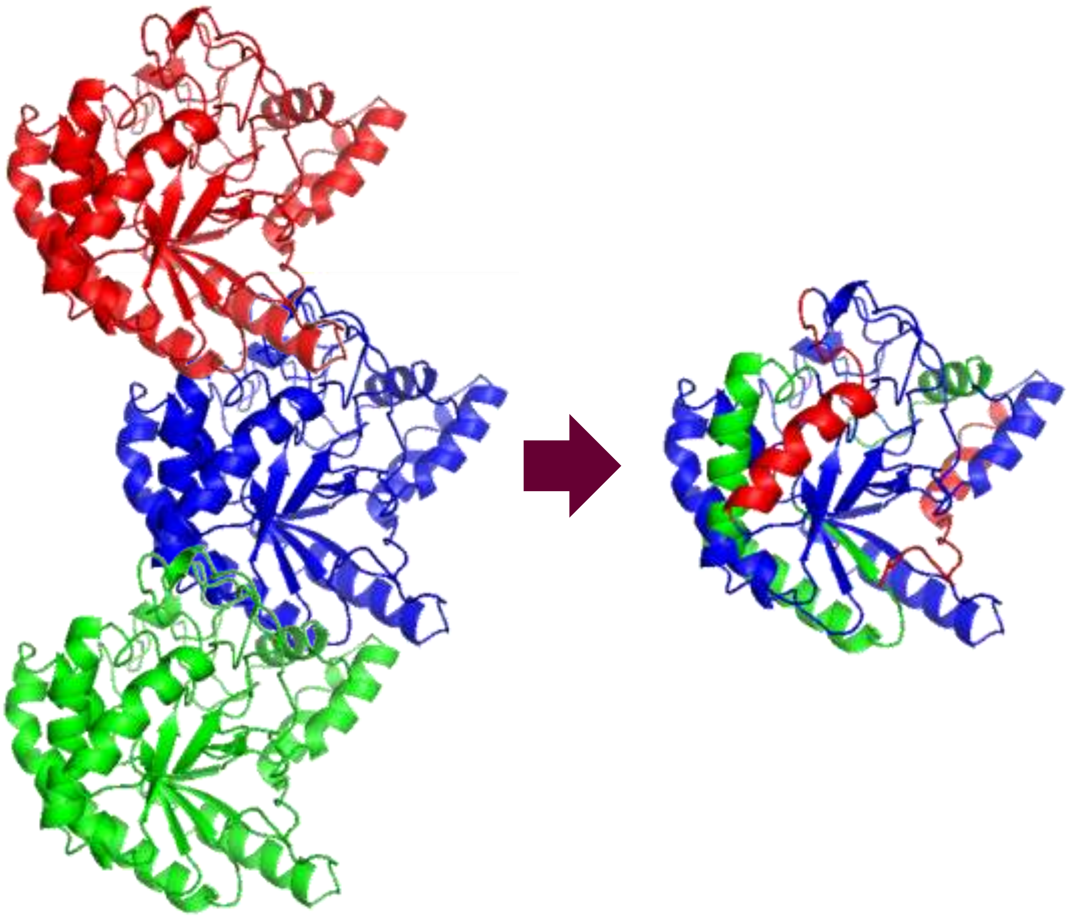
Increased number of mutations - Increased property diversity

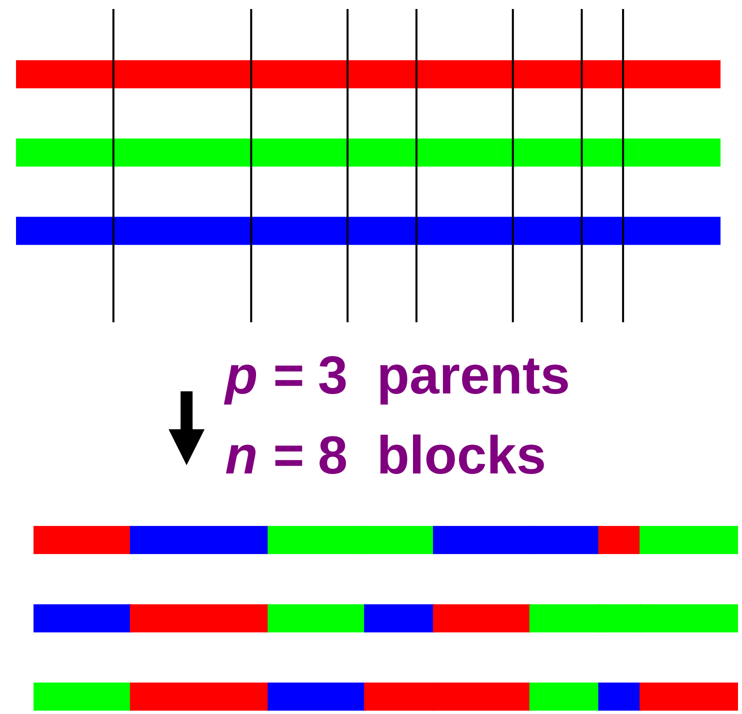


Drummond et al., PNAS 2005

Chimera Family Design

Recombine 8 blocks from 3 parents to yield 6,561 chimeras



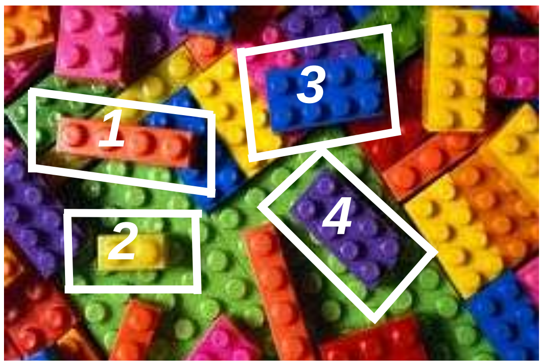


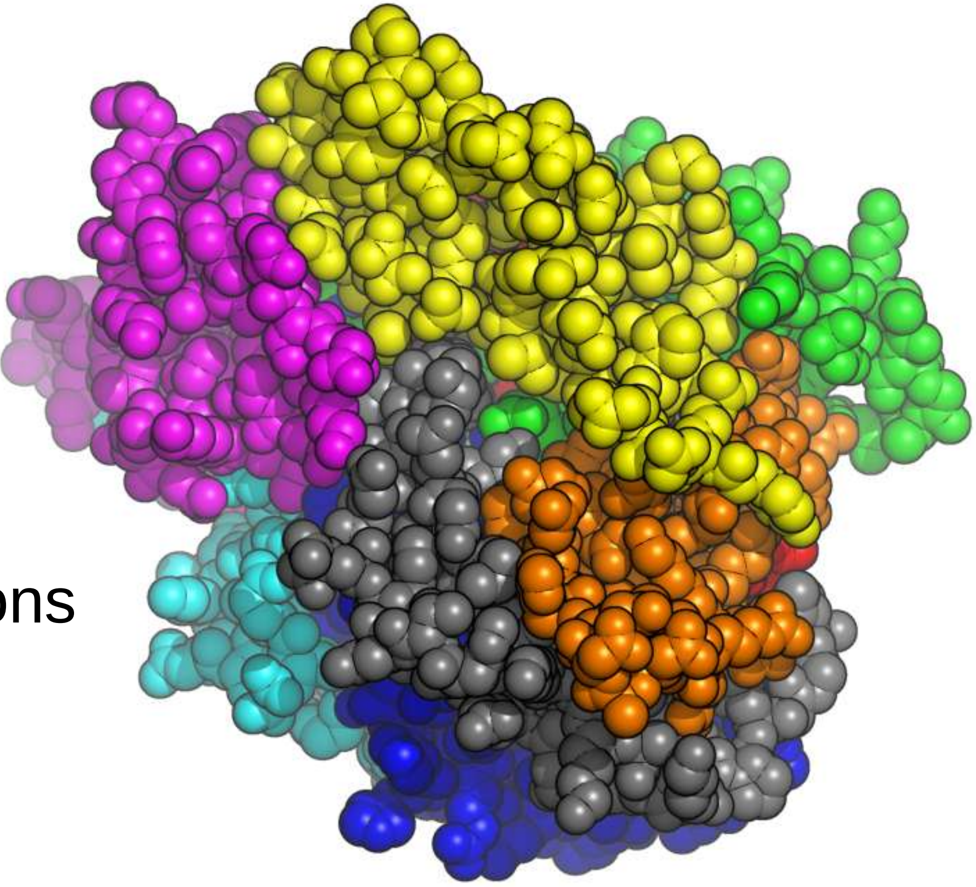
p^n = 6,561 *chimeras*

Otey et al. PLoS Biology 2006, Voigt et al. Nat. Struct. Biol. 2002

SCHEMA
Structure-Guided Recombination

Use crystal structures (or models) to identify swappable "blocks" of protein sequence amenable to recombination





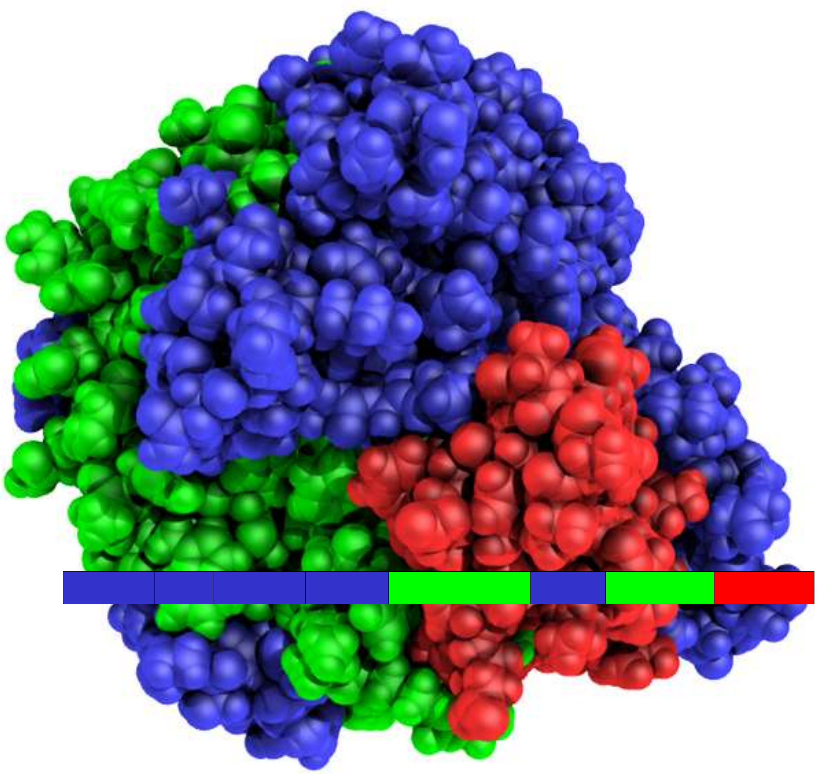
Minimize number of interactions broken upon recombination -

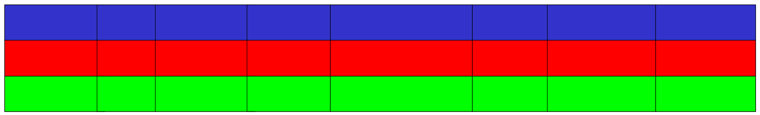
Maximize number of active chimeras

Voigt et al., Nat. Struct. Biol. 2002

Recombination of Fungal Cellobiohydrolase II (CBH II) Cellulase Enzymes

S. cerevisiae used as recombinant secretion host





Parent 1: *H. insolens*

Parent 2: *H. jecorina*

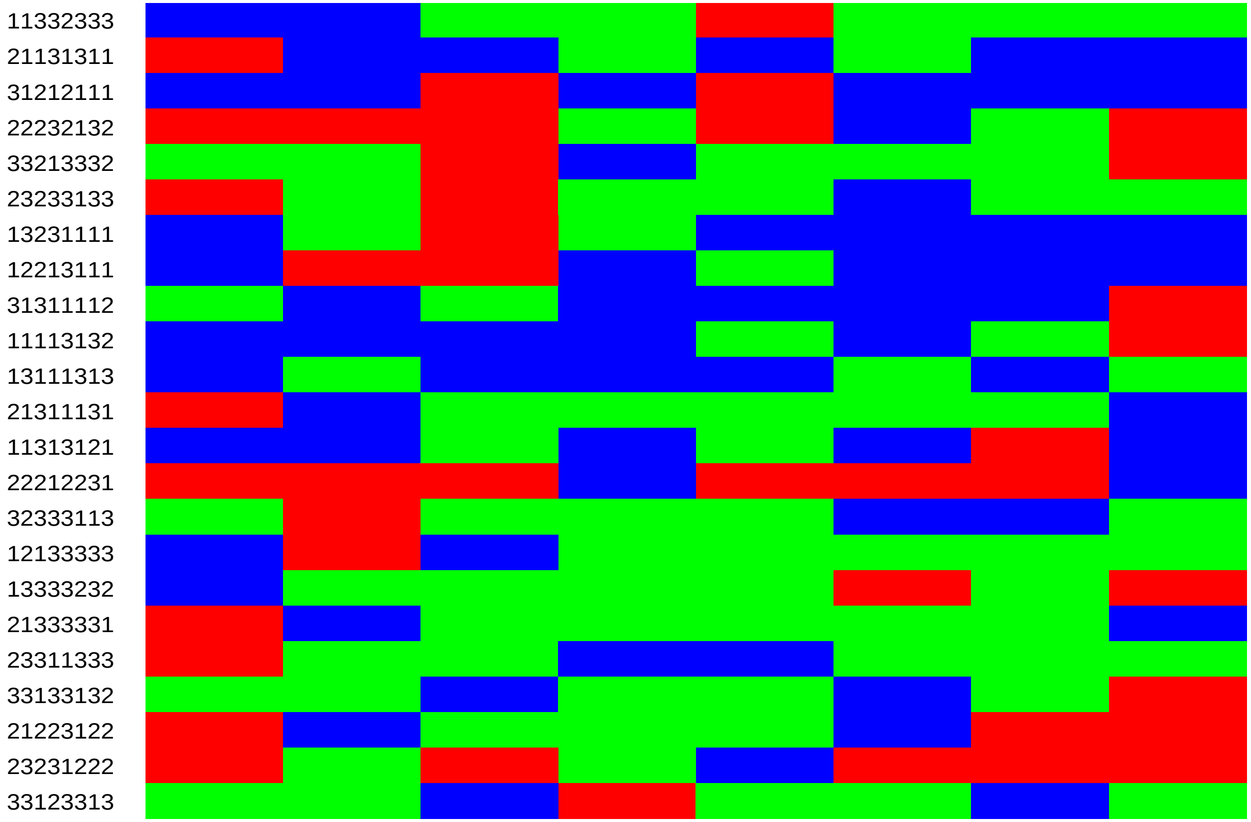
Parent 3: *C. thermophilum*

3^8 = 6,561 chimeras

Heinzelman et al. PNAS 2009

CBHII Chimera Sample Set

Parent 1: *H. insolens*
Parent 2: *H. jecorina*
Parent 3: *C. thermophilum*

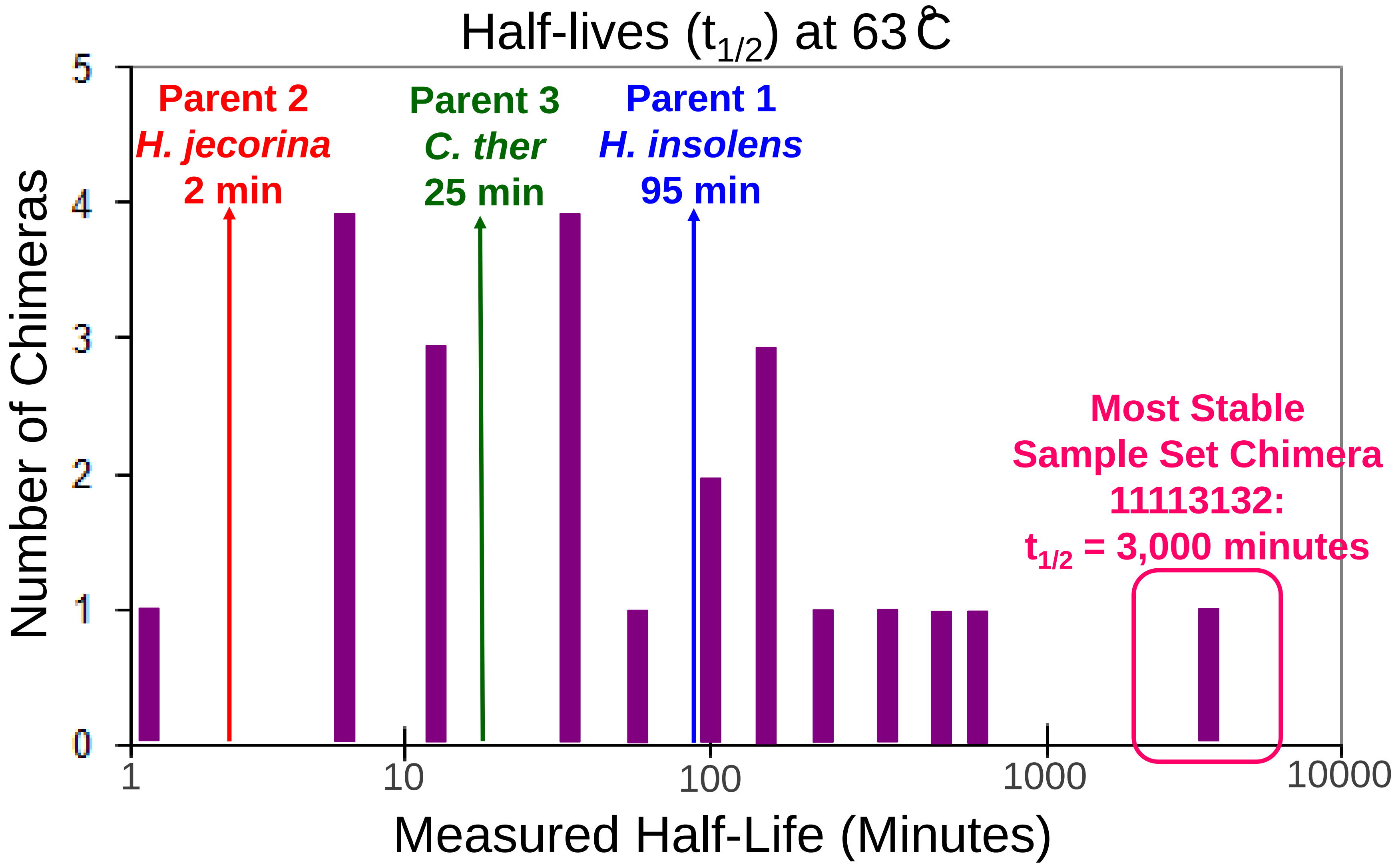


48 genes synthesized, 23 secreted chimeras

Average of 36 mutations from closest parent

H. jecorina underrepresented in active chimeras (Block 4)

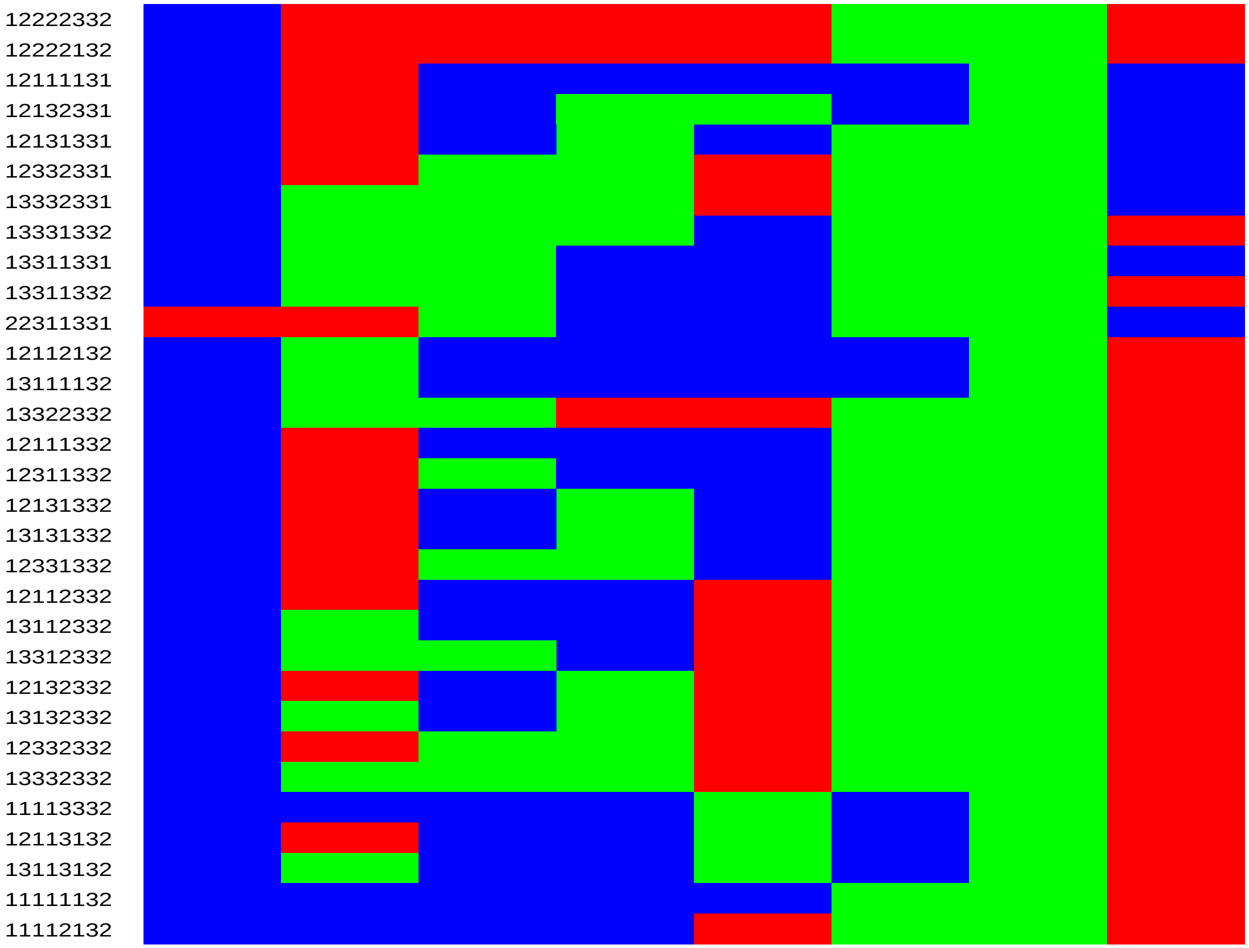
Broad Range of Chimera Stabilities



Stable CBHII Sequence Prediction

Qualitative block classification - stabilizing, destabilizing or neutral

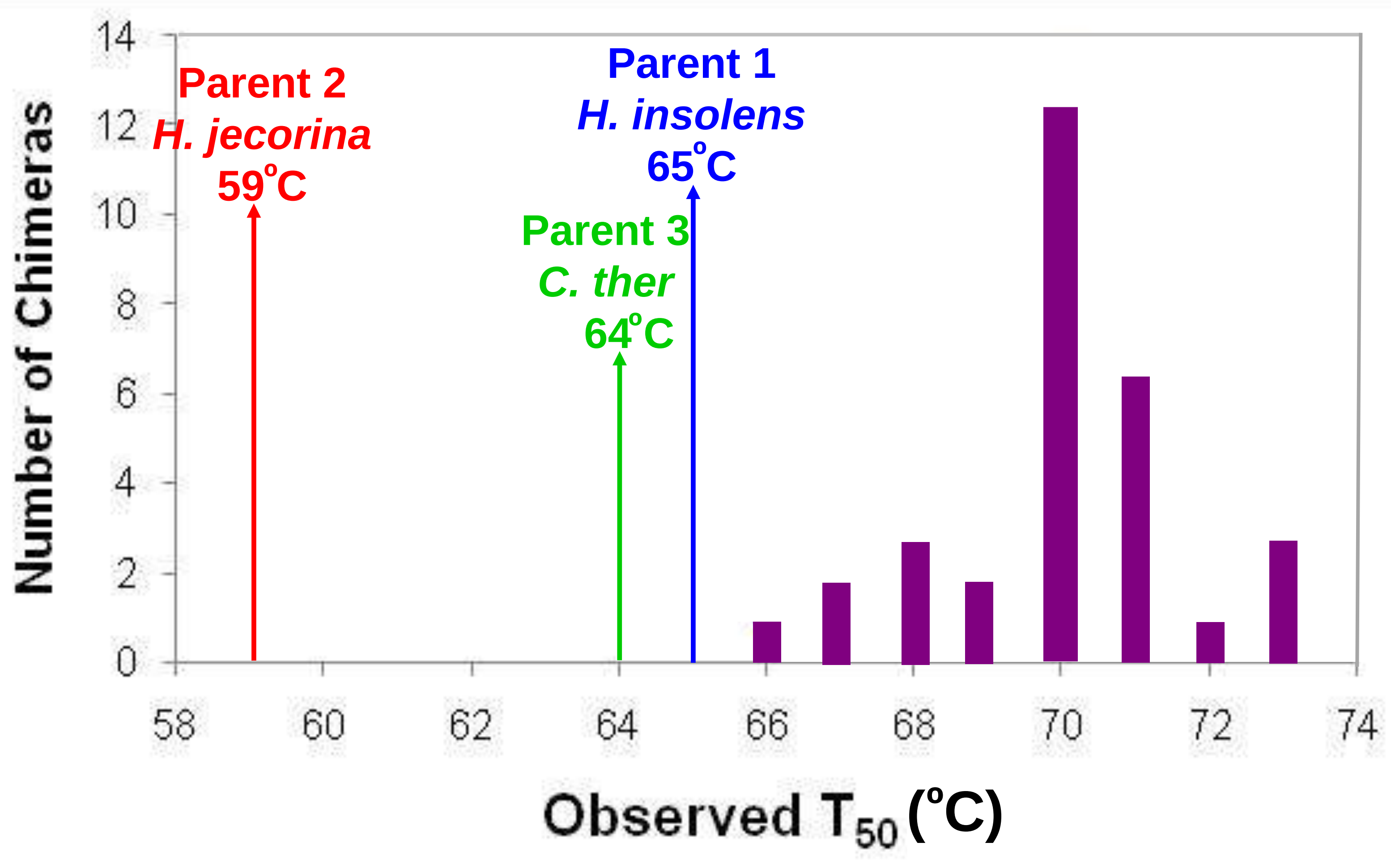
Synthesized 41 chimeras enriched in stabilizing & neutral blocks



Stability Predictions Validated

31 of 41 predicted stable chimeras secreted

All 31 are more stable than the parents!!

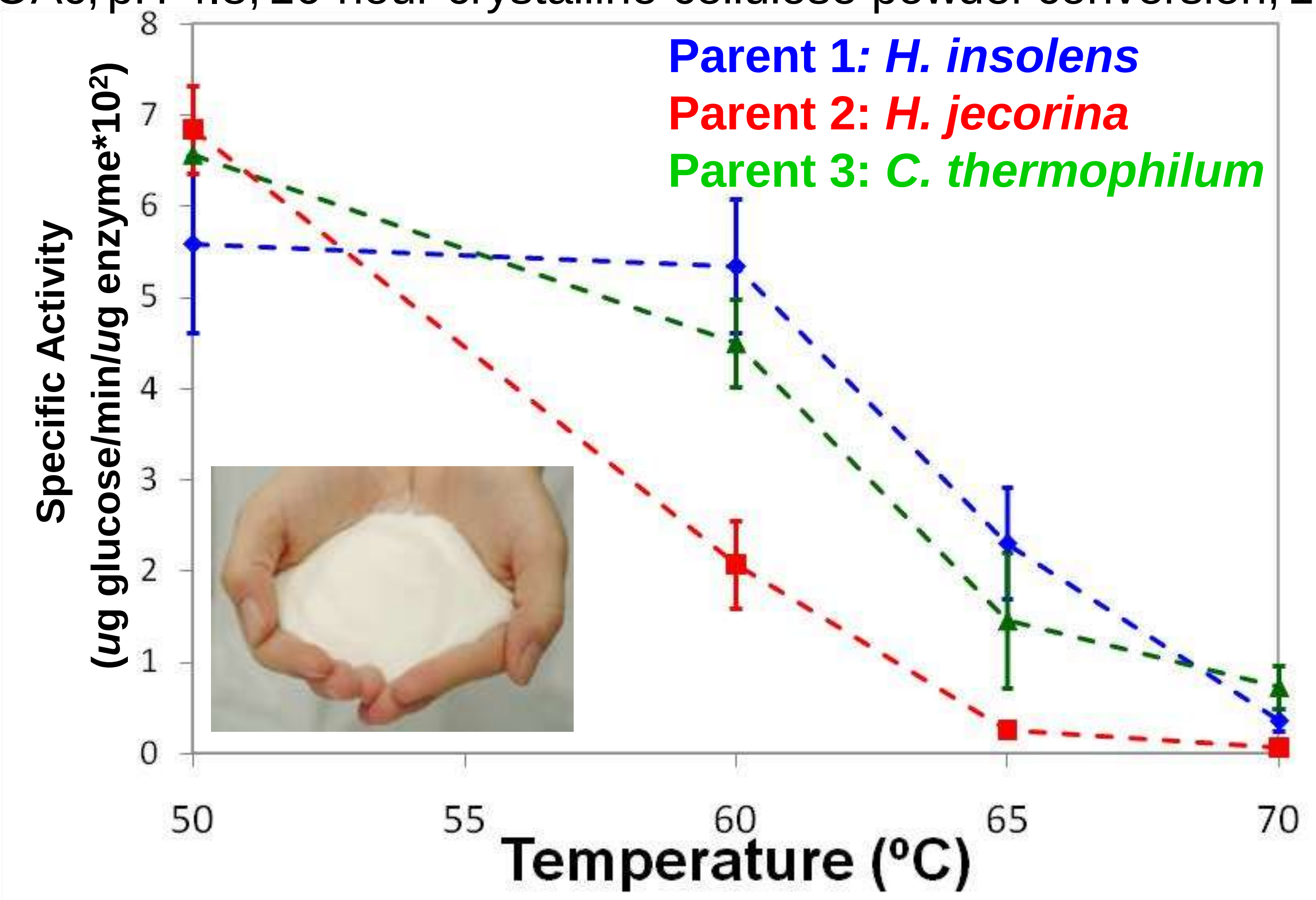


CBHII Parent Biomass Conversion

Decreasing cellulose conversion from 50°C to 70°C

Parent CBHIIs inactive at 70°C

50 mM NaOAc, pH 4.8, 16-hour crystalline cellulose powder conversion, 1 *u*g/mL CBH II

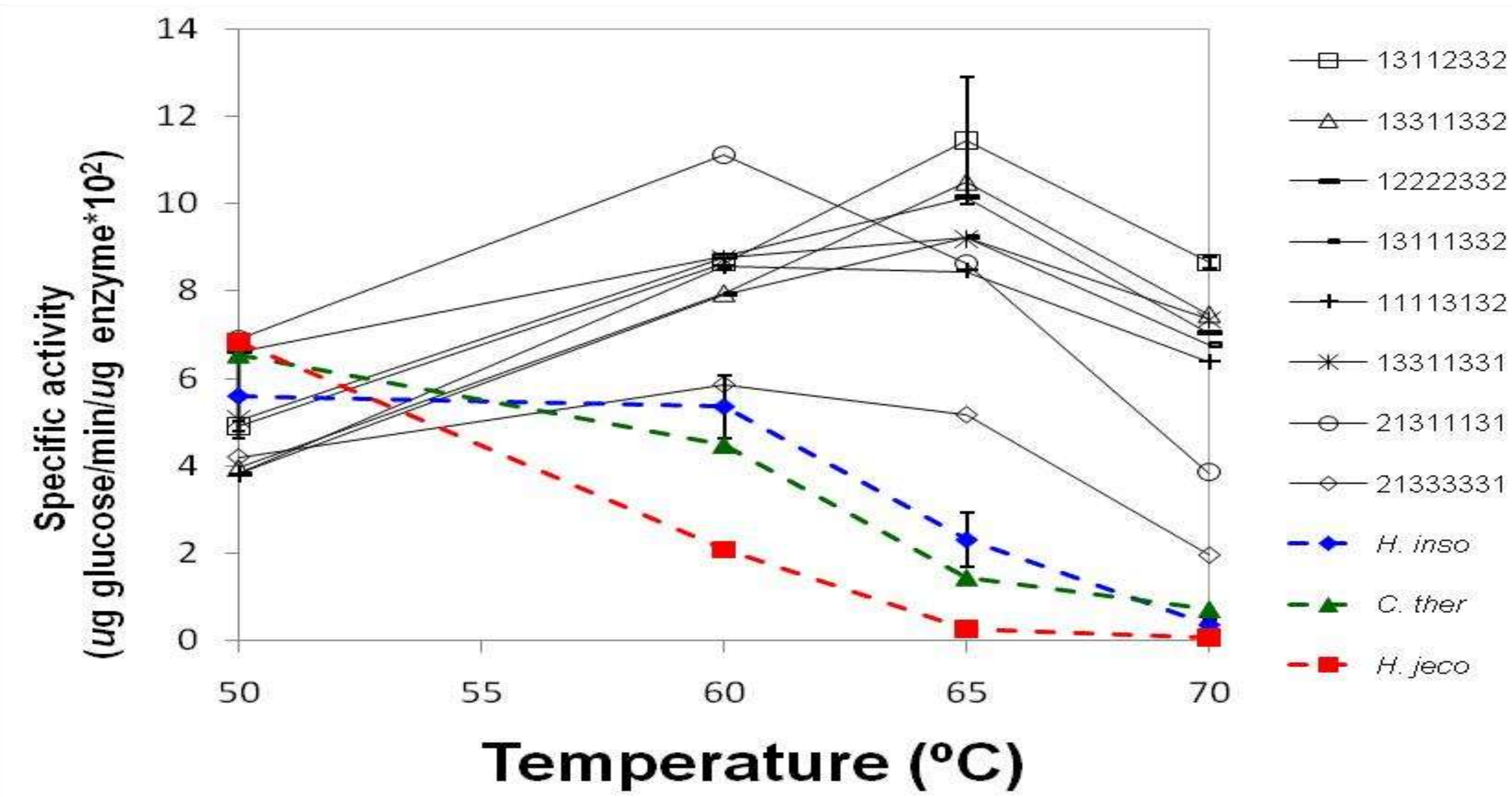


Chimeras Hydrolyze More Biomass

CBHII chimera conversion increases at high temperature

Seven of eight chimeras outperform parent CBHIIs

All eight chimeras remain active at 70°C



Three Parent SCHEMA Recombination

Identified stabilizing blocks - Chimeras more stable than parents

Three Parent SCHEMA Recombination

Identified stabilizing blocks - Chimeras more stable than parents

Predicted stable chimera sequences with 100% accuracy

Three Parent SCHEMA Recombination

Identified stabilizing blocks - Chimeras more stable than parents

Predicted stable chimera sequences with 100% accuracy

Chimeras have improved long-time cellulose hydrolysis activity

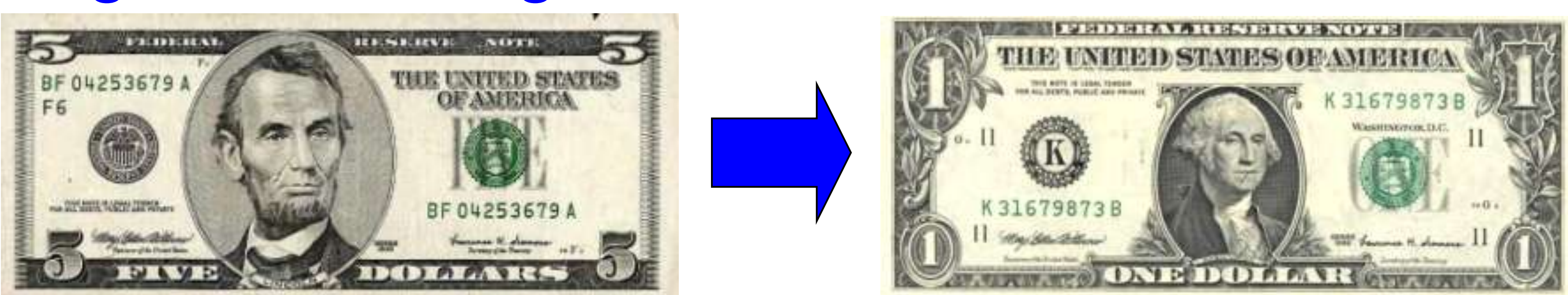
Three Parent SCHEMA Recombination

Identified stabilizing blocks - Chimeras more stable than parents

Predicted stable chimera sequences with 100% accuracy

Chimeras have improved long-time cellulose hydrolysis activity

Progress toward goal of 5X cellulase cost reduction



Three Parent SCHEMA Recombination

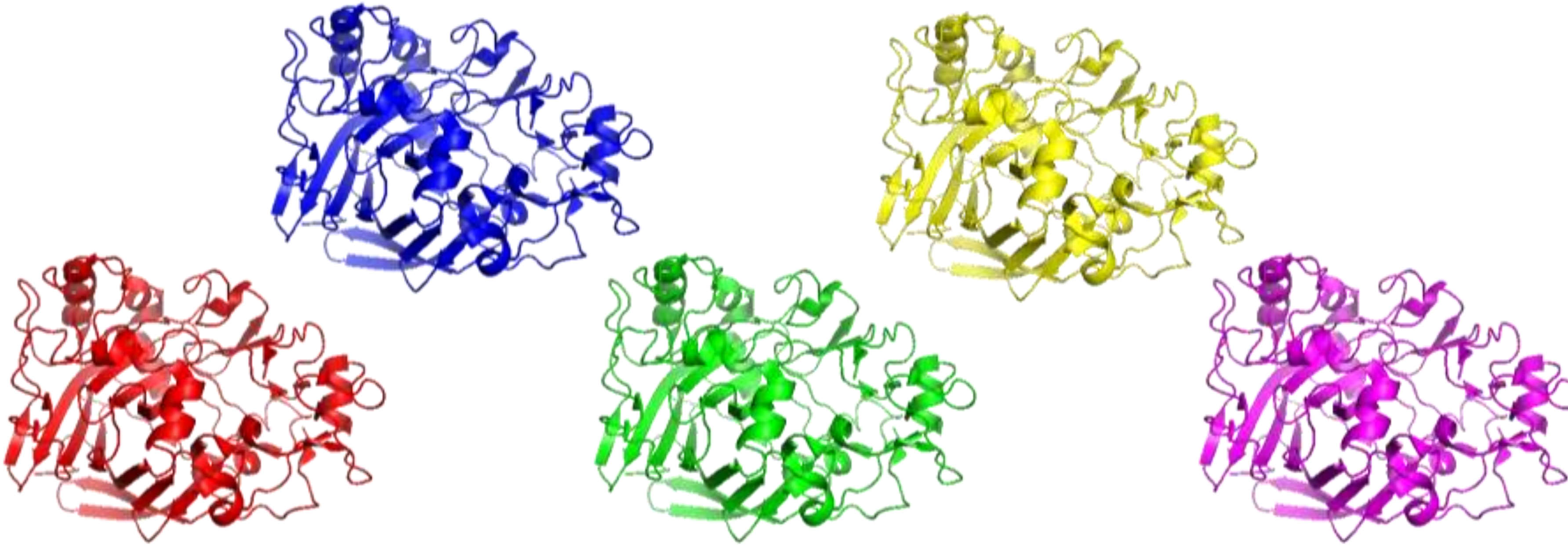
Identified stabilizing blocks - Chimeras more stable than parents

Predicted stable chimera sequences with 100% accuracy

Chimeras have improved long-time cellulose hydrolysis activity

Could we have done better?

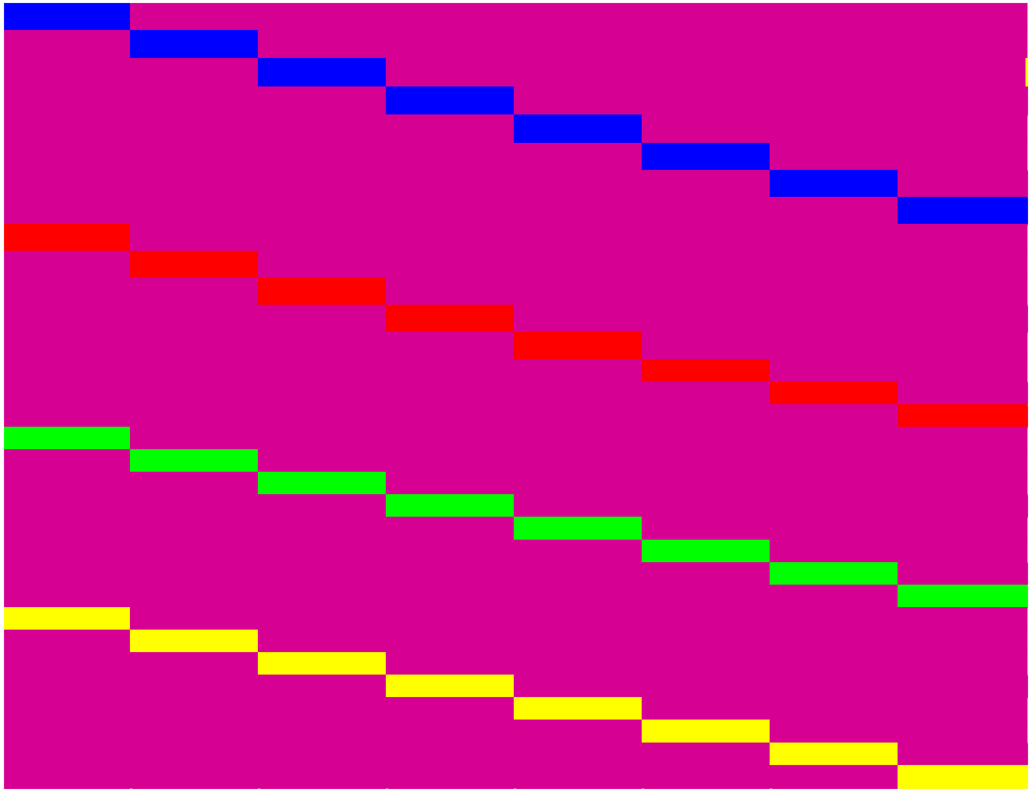
Expanded SCHEMA Families: “p” Parent Recombination



Increase probability of identifying desirable blocks increasing with number of parents (p)

Design sample set chimeras to maximize fraction of sample set chimeras that are secreted

Single Block Substitution "Monomeras"



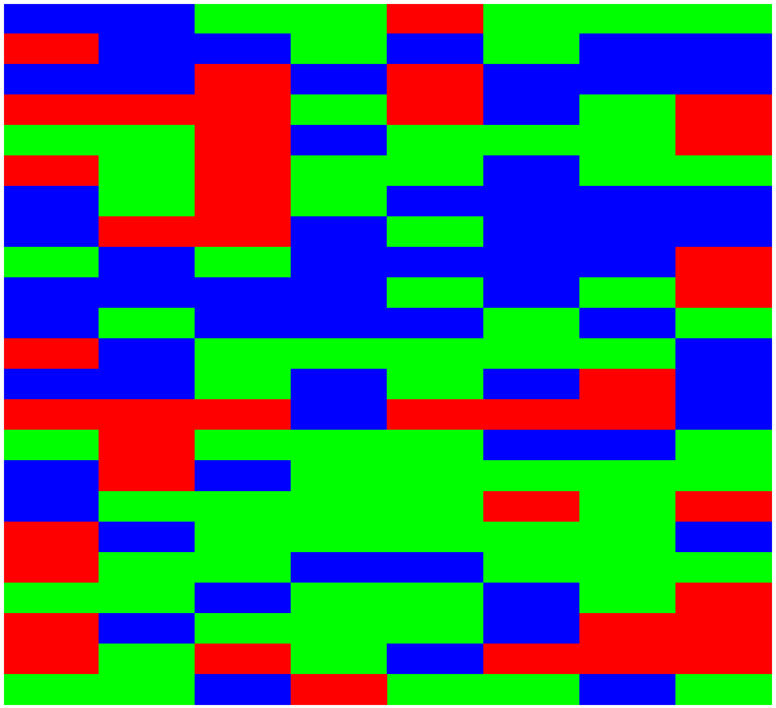
Choose one **highly secreted** parent

Substitute other parents 1 block at a time

Screen (8*p) blocks from "p" parents with 8*(p-1) monomeras

Prevents poorly secreted blocks from "polluting" sample set

More Blocks From Smaller Sample Set

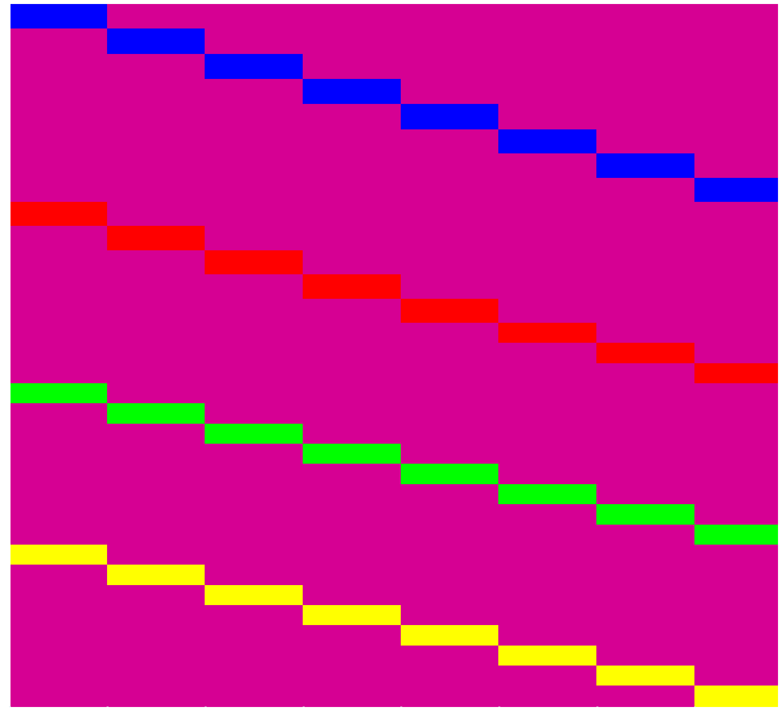


3-Parent Chimera Screen

48 sample set chimeras

3*8 = 24 blocks

$3^8 = 6{,}561$ chimeras



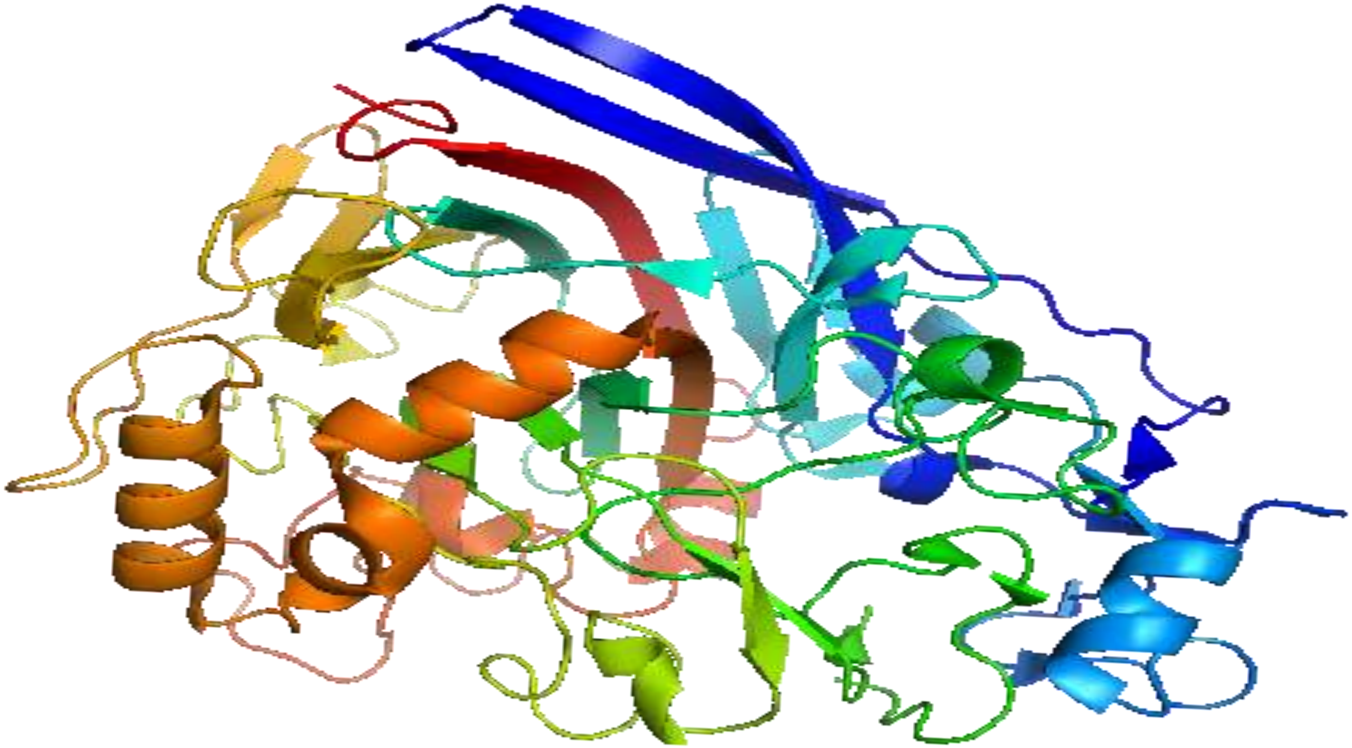
5-Parent Monomera Screen

32 monomeras

5*8 = 40 blocks

$5^8 = 390{,}625$ chimeras

Monomera Screening Method: Fungal Cellobiohydrolase Class I (CBH I)



CBH I Background Parent Selection

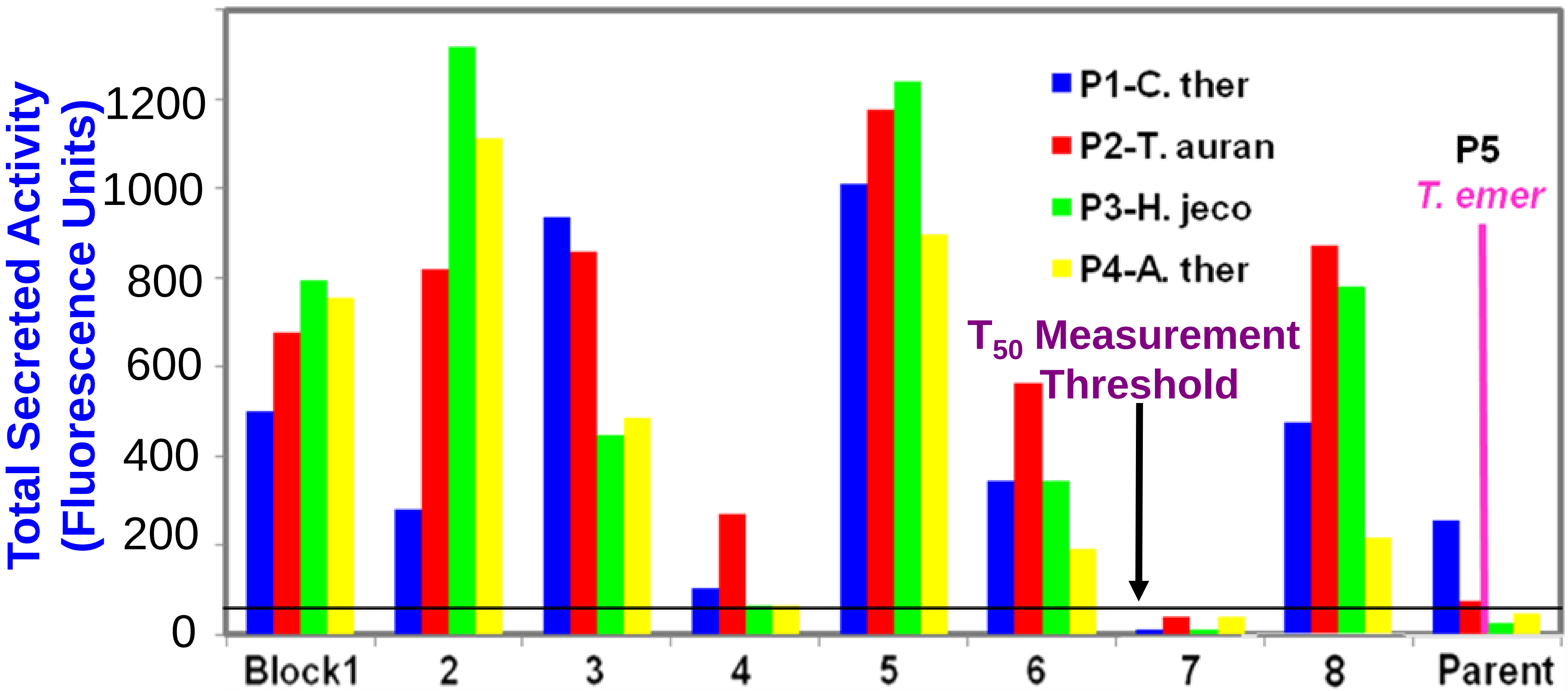
CBH I Parent	Total Secreted Activity (Fluorescence Units)	T_{50} (C)
P1 - C. ther	**400**	**59.9 +/- 0.5**
P2 - T. auran	**70**	**62.2 +/- 0.4**
P3 - H. jeco	**20**	**ND**
P4 - A. ther	**40**	**ND**
P5 - T. emer	**1000 +/- 100**	**62.9 +/- 0.3**

Five CBH I parents - 62-81% sequence identity

T. emersonii most highly expressed - **Background parent**

His_6-isolated Parent 1, 2, & 5 specific activities approximately equal

88%Monomeras Secreted



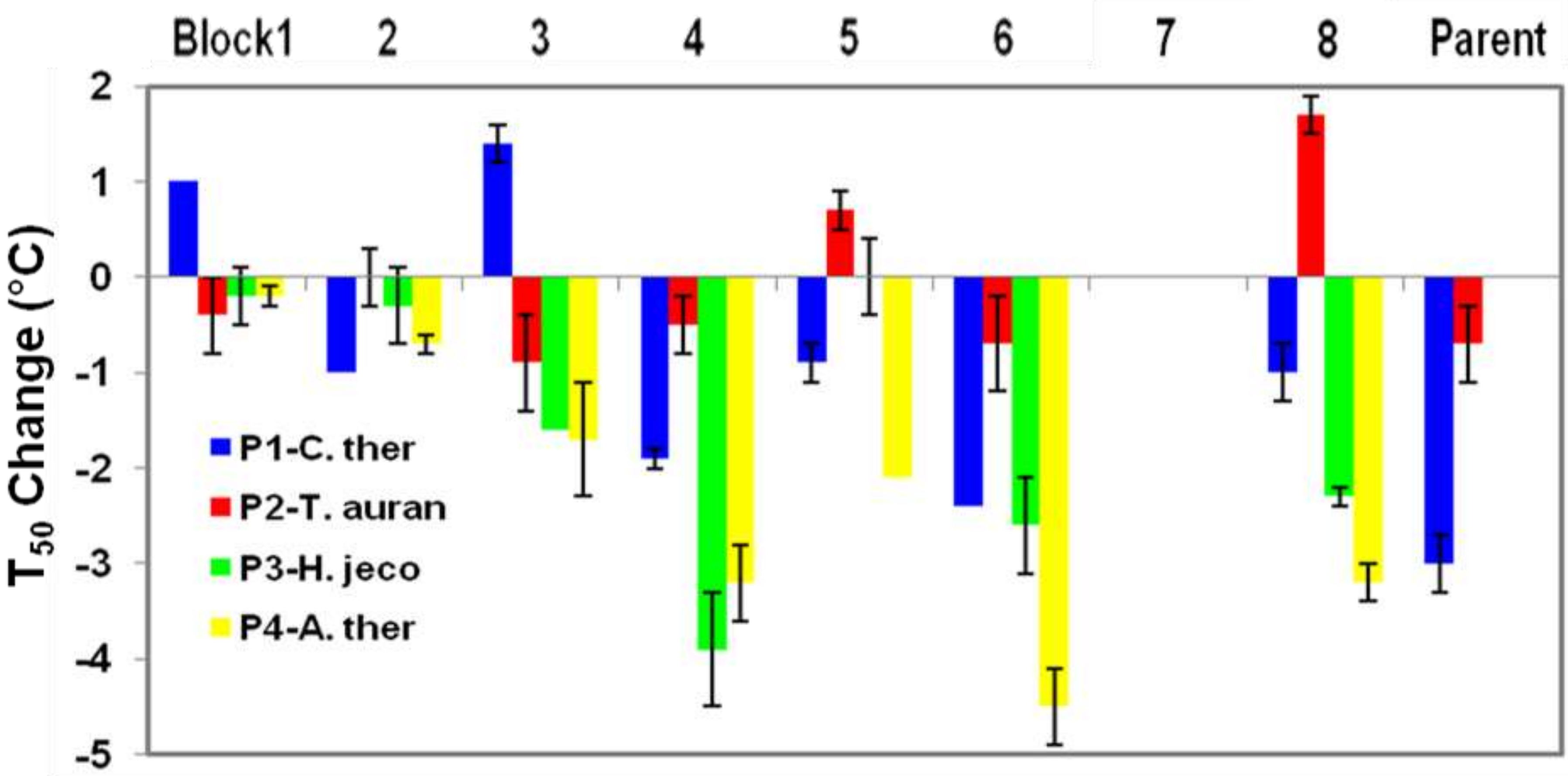
Secretion for 28 of 32 monomeras adequate for T_{50} measurement

Block 7 substitutions not tolerated

Block 4 substitutions mildly tolerated

Mixed block sample set would have had few active chimeras

Stabilizing Blocks Identified



Four stabilizing blocks - B1P1, B3P1, B5P2 & B8P2

Five neutral blocks - B1P2, B1P3, B1P4, B2P1 & B5P3

Blocks 4 and 7 from parent 5 “not destabilizing”

100% Stability Prediction Accuracy

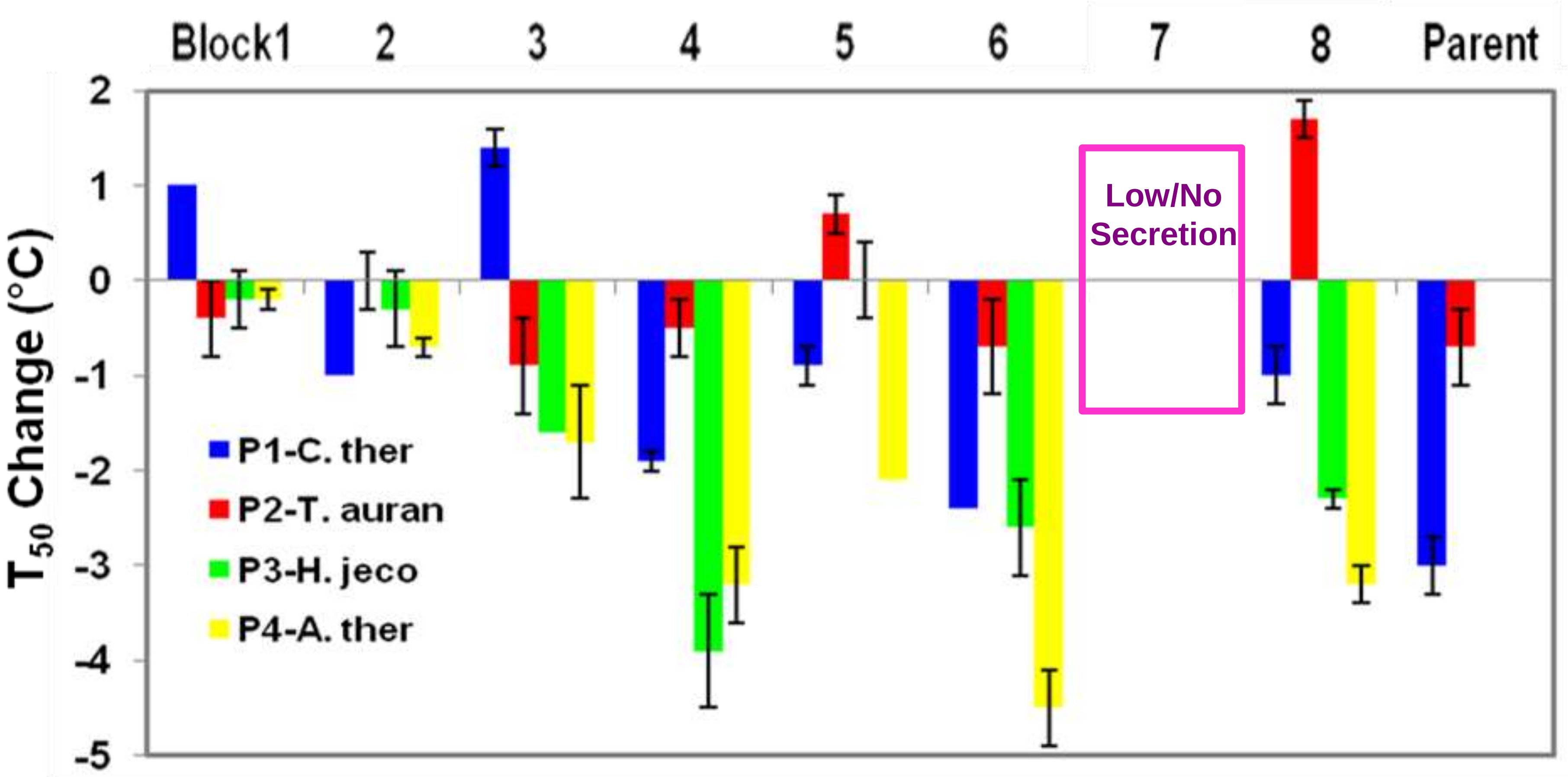
CBH I Sequence	CBH I Parent Represented At Each Block Position	T_{50} (°C)	Secreted Activity
11111111		59.9 +/- 0.5	330
22222222		62.2 +/- 0.4	70
33333333		ND	20
44444444		ND	40
55555555		62.9 +/- 0.3	1000
34152252		64.0 +/- 0.1	980
55153552		64.3 +/- 0.0	1440
32153252		64.3 +/- 0.2	440
55155552		64.4 +/- 0.7	950
22153252		64.4 +/- 0.2	560
52152552		64.5 +/- 0.0	1500
12153252		64.7 +/- 0.2	280
45153252		64.8 +/- 0.2	1100
12153552		64.9 +/- 0.3	470
25152252		65.0 +/- 0.1	970
13152552		65.0 +/- 0.0	1510
12152252		65.3 +/- 0.1	440
55153252		65.3 +/- 0.2	870
55552252		65.6 +/- 0.7	800
55152552		65.7 +/- 0.1	1280
55152252		66.3 +/- 1.0	850

Recombined stabilizing & neutral blocks

16 of 16 chimeras with T_{50} > All 3 secreted parents

Average of 37 mutations per chimera

Subblock Recombination



All block 7 substitutions markedly reduce secretion

Block 7 is largest block - 116 of 437 CBH I amino acids

Subdivide block 7 to create desirable "subblocks"

Parent Two Subblocks Moved Into Parent Five

Subblocks from parent 2

Constant residues from parent 5

Number of mutations in subblock

B7P2 has highest identity (73%) to B7P5

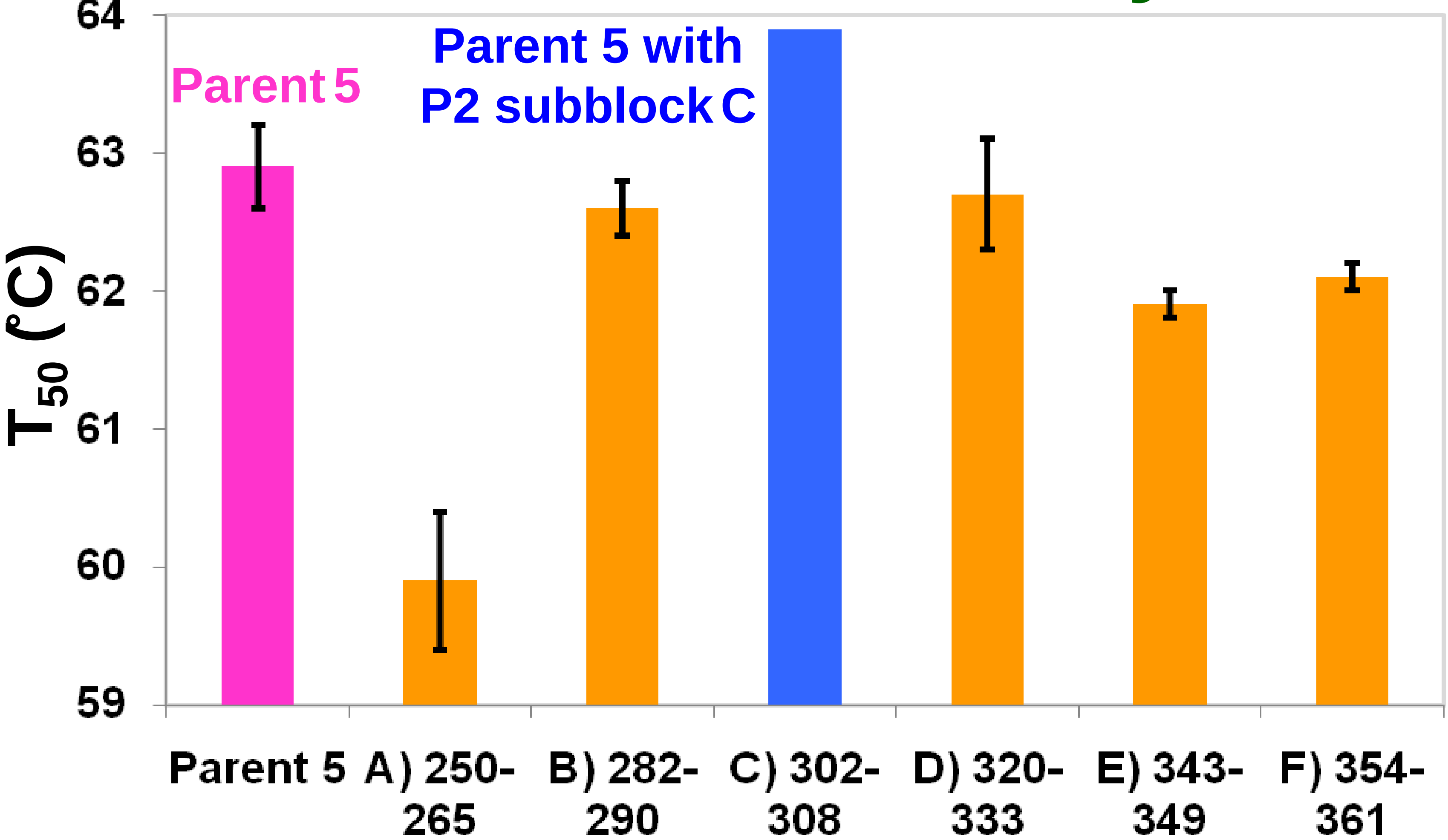
Group proximal mutations into subblocks for cloning convenience

Subblocks contain between 2 & 7 mutations each

Move P2 subblocks into P5 one at a time

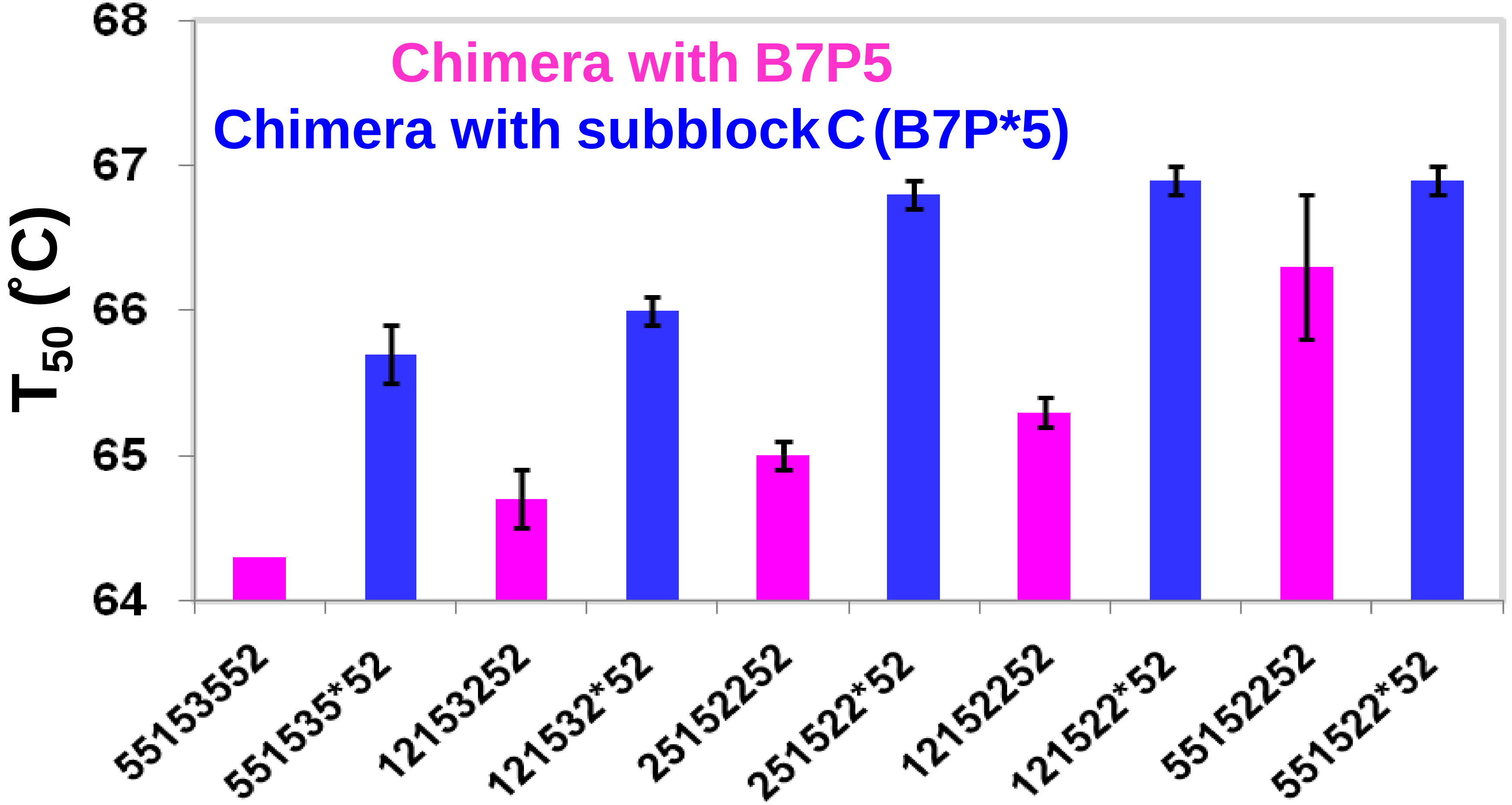
Identify P2 subblocks that improve stability

Subblock “C” From Parent Two Increases Stability



Parent 2 subblock “C” (residues 302-308) increases T_{50}

B7P*5 Subblock Chimeras: Increased Stability

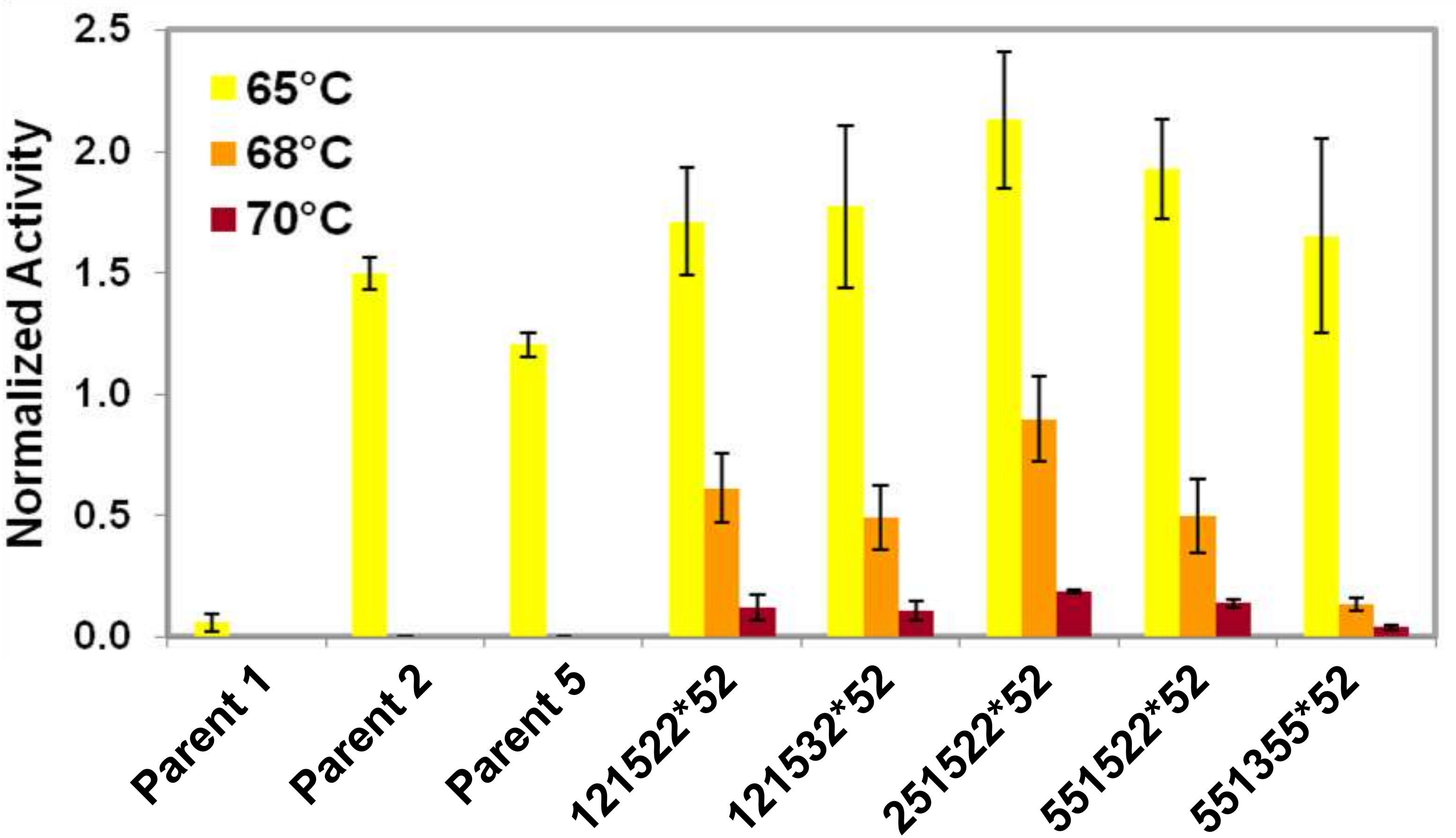


B7P*5 block (contains subblock "C")
increases T_{50} values for 5 of 5 chimeras

Subblock "C" Chimeras Convert Biomass at Elevated Temperature

Long-Time Solid Cellulose Conversion

15 *u*g/mL CBH I, 60 mg/mL Lattice NT cellulose, 50mM pH 4.8 NaOAc, 16-hour Rx



5 of 5 P7B*5 chimeras convert biomass at up to 70°C

Parents completely inactive above 65°C

“p”-Parent CBHI SCHEMA Recombination & Block Screening

Monomera screening approach identified stabilizing blocks

“p”-Parent CBHI SCHEMA Recombination & Block Screening

Monomera screening approach identified stabilizing blocks

Screening of only 32 monomeras elucidated chimeras among most stable of 400,000 possible chimera sequences

“p”-Parent CBHI SCHEMA Recombination & Block Screening

Monomera screening approach identified stabilizing blocks

Screening of only 32 monomeras elucidated chimeras among most stable of 400,000 possible chimera sequences

Stabilizing subblock from nonstabilizing recombination block

“p”-Parent CBHI SCHEMA Recombination & Block Screening

Monomera screening approach identified stabilizing blocks

Screening of only 32 monomeras elucidated chimeras among most stable of 400,000 possible chimera sequences

Stabilizing subblock from nonstabilizing recombination block

Chimeras catalyze biomass conversion at higher T than parents

“p”-Parent CBHI SCHEMA Recombination & Block Screening

Monomera screening approach identified stabilizing blocks

Screening of only 32 monomeras elucidated chimeras among most stable of 400,000 possible chimera sequences

Stabilizing subblock from nonstabilizing recombination block

Chimeras catalyze biomass conversion at higher T than parents

Validated monomera screening as GENERAL approach - Extrapolate to other enzymes