From Structure to Screen:

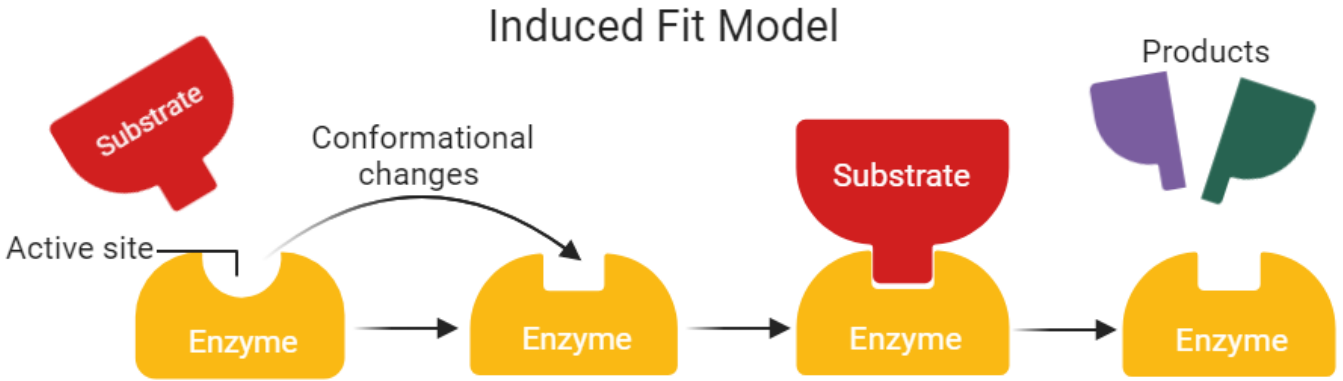
MD-Ready Enzyme-Substrate Pairs

Maximilian Faissner

Main Supervisor: **Christoph Flamm** (Institute for Theoretical Chemistry)

In collaboration with:
Shelley Copley, Walter Fontana, Lynn Kamerlin

The textbook lie: enzymes are *perfectly* specific

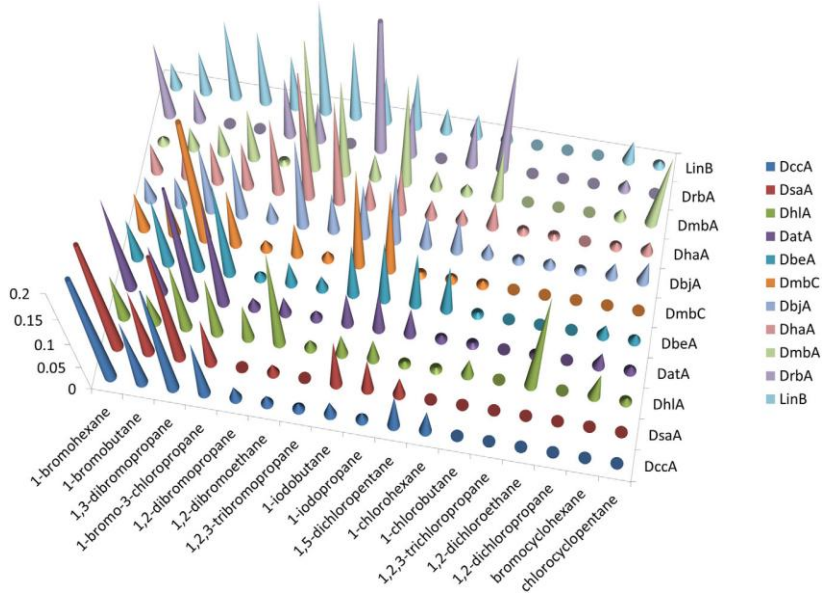


Reality: many enzymes are **promiscuous**:
accept multiple substrates / side reactions

Many enzymes have broad specificity profiles

Enzyme specificity Profile (ex: Haloalkane Dehalogenases)

L . Carlucci et al. (2015)



activity on many substrates per enzyme

If promiscuity is everywhere, **predicting** it becomes crucial.

Supervision / MATOMIC (Mathematical Modelling for Microbial Community Induced Metabolic)

Shelley Copley

University of Colorado Boulder

Walter Fontana

Harvard Medical School

Lynn Kamerlin

Georgia Tech

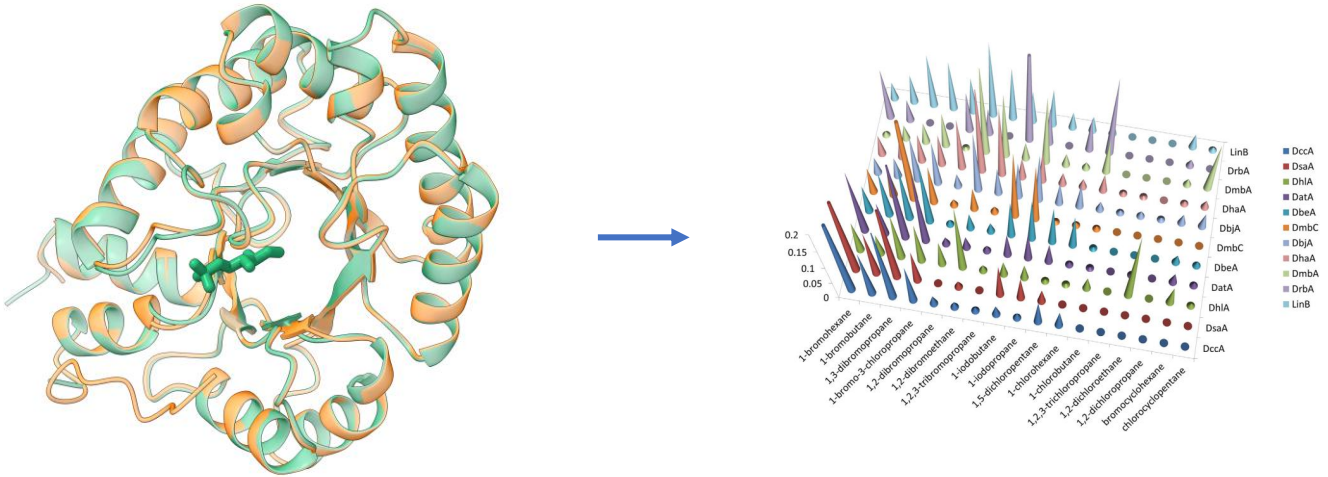
Christoph Flamm

MATOMIC: Focus on Modelling of Metabolic Networks

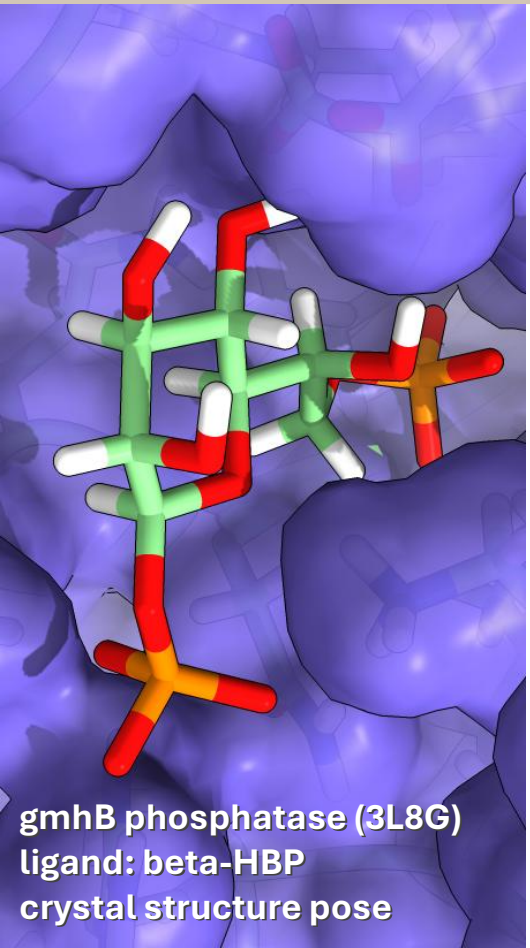
Copenhagen Meeting 04/2025

https://www.sdu.dk/en/forskning/matomic

Project Goal: (Semi-) quantitative promiscuity prediction



Thermodynamic Context of Enzyme Catalysis



Today

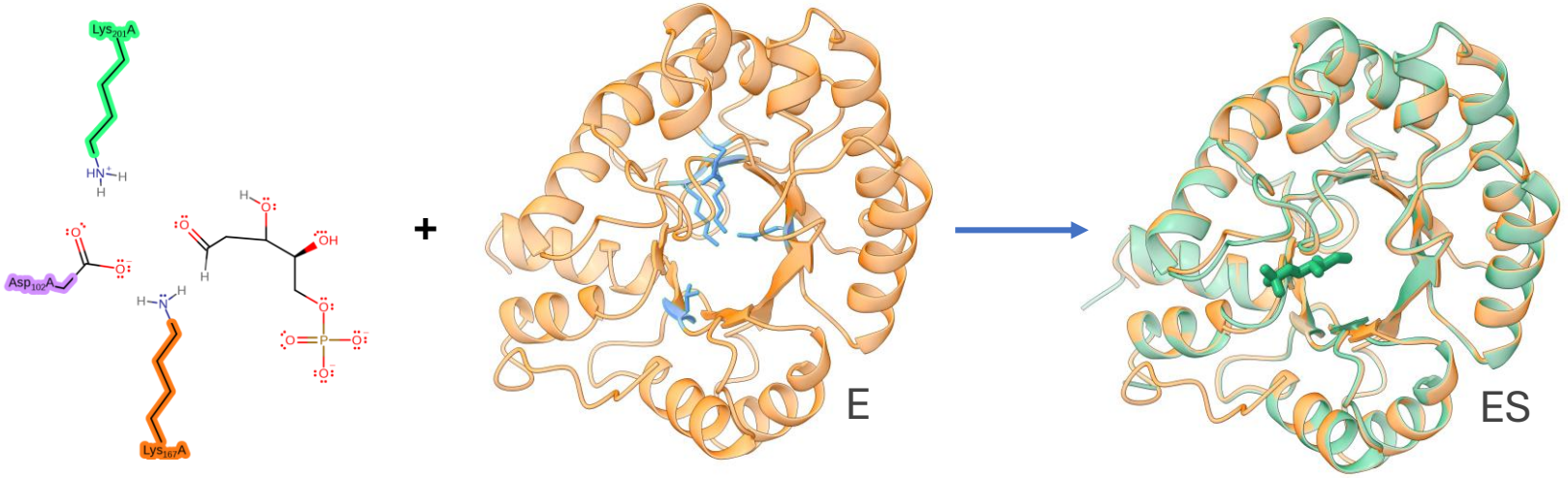
- Define NAC (near attack conformation)
- NACTrace (filtering, NAC populations in MD)
- Curated enzyme/substrate MD dataset

Not today

- MD details (system parameterization)
- Ligand binding affinity estimations (MM-PBSA, …)

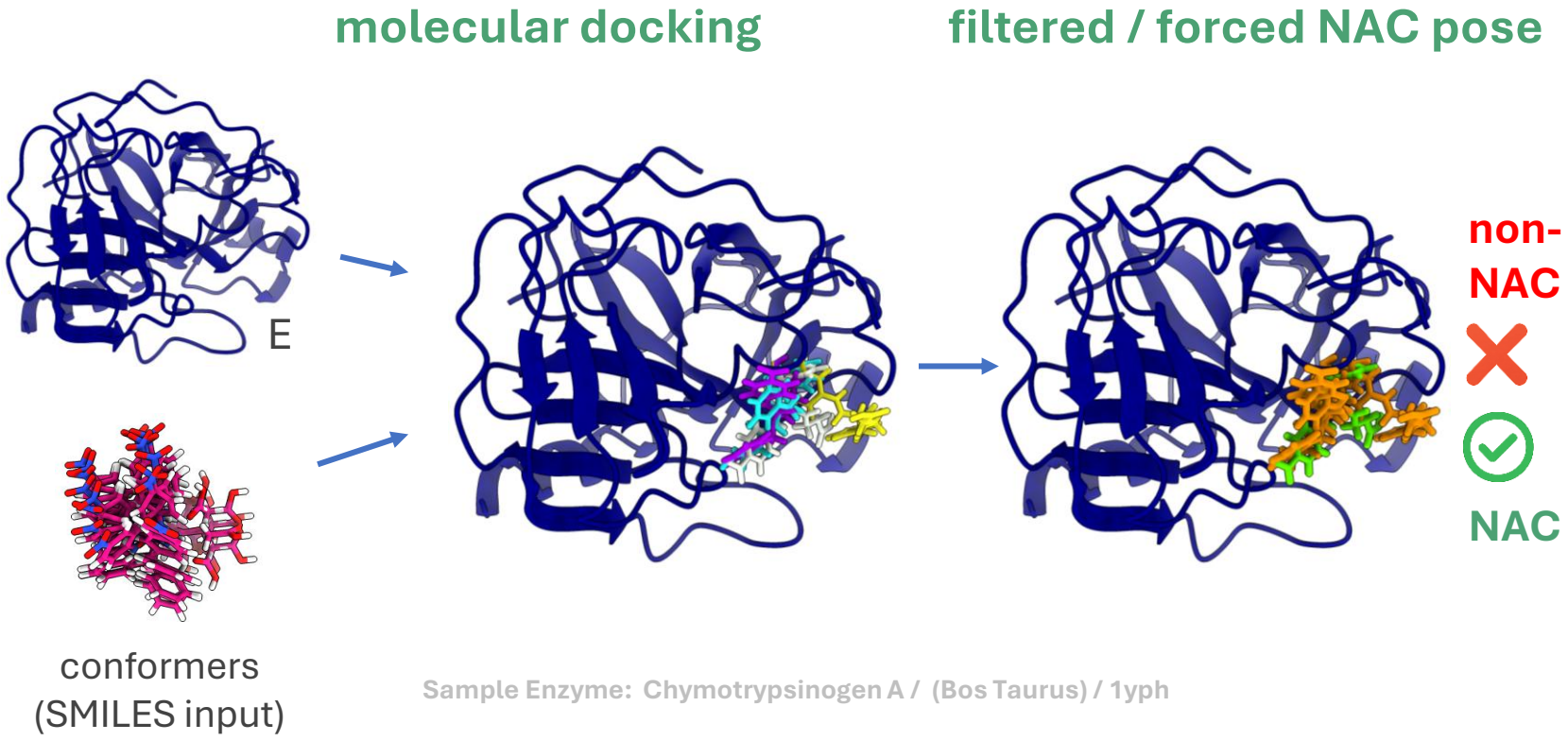
Requirement: Reaction Annotations & Enzyme Structures

We need a well-annotated structures linked to chemistry



Sample Enzyme: Q7MHL9 / Ribose-5-phosphate isomerase / rpiA
(aldose-ketose isomerization)

NACtraces: Catalytic poses from human-readable specs



NACtraces: Catalytic poses from human-readable specs

One NAC definition for

- **pdb files (static)**
- **Trajectories**

Use cases:

- **filter docking poses**
- **measure NAC population in MD**

Output:

- **rank / pass–fail**
- **summary stats (per pose, per trajectory)**

Input 1: YAML reaction spec

Input 2: pdb / trajectory

↓

map site + reactive atoms

↓

compute geometry

↓

NAC stats

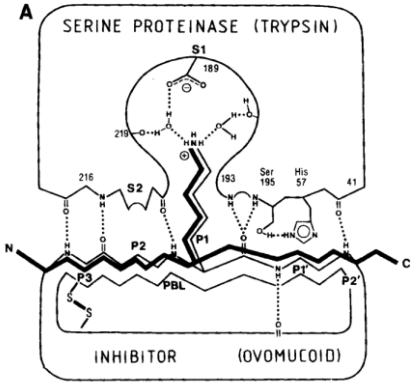
NACtraces: MD trajectory analysis (substrate comparison chymotrypsin)

Chymotrypsin NAC constraints:

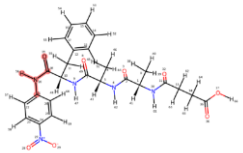
Distance:
- Nucleophile-electrophile ~3.0–3.5 Å.

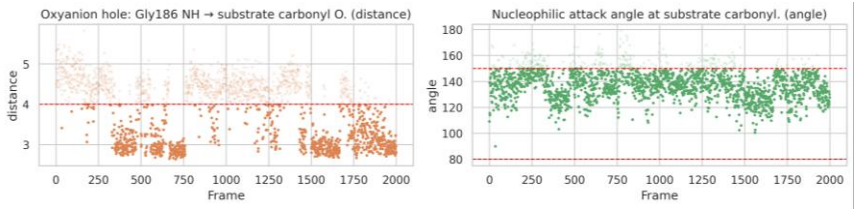
Angle (Bürgi–Dunitz):
- ~105° ± 5–10°

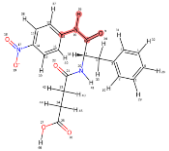
Oxyanion alignment:
- Correct carbonyl oxygen orientation

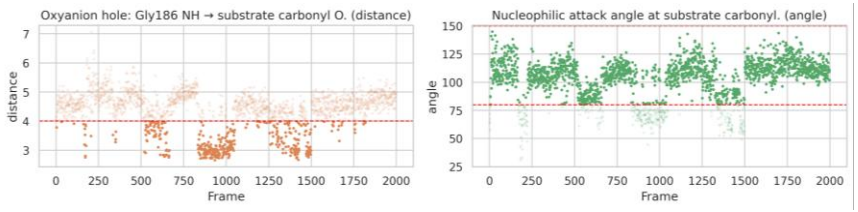


SAAPpNA (N-succinyl-l-alanyl-l-alanyl-l-phenylalanine-p-nitroanilide)









SPpNA (N-Succinyl-l-phenylalanine-p-nitroanilide)

Ligand placement w/ molecular docking: The 2-body problem

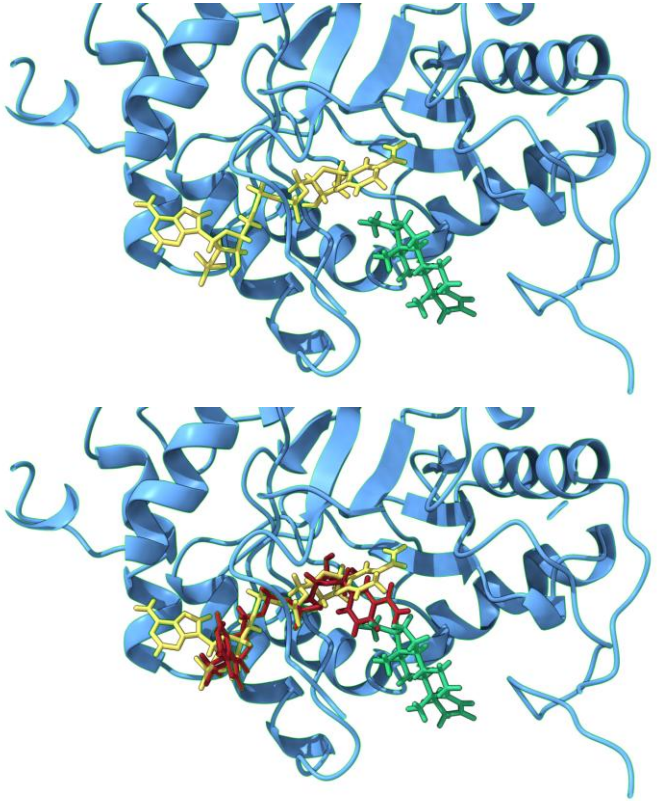
Molecular docking is designed for a 1-body problem (Protein + 1 Ligand)

- Sequential Bias:
  - docking ligand 1 (NADPH) first
  - algorithm optimizes position in empty space
  - likely will shift towards ligand 2
  - Error Propagation

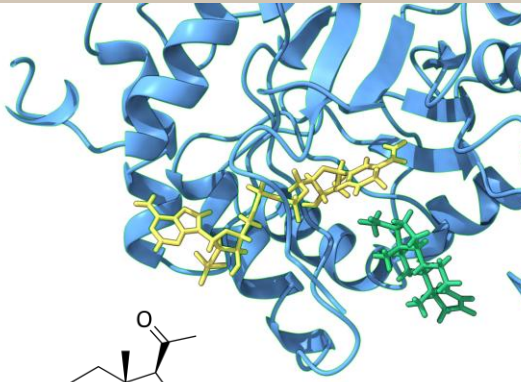
Solutions for common cofactors

- Homology-based templates (via Foldseek)
- AI models (AlphaFold 3, …)

⏳ **WIP …**



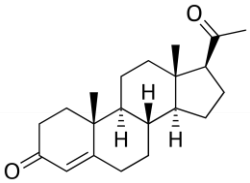
Case Study: AKR1C1 (NADPH-d. oxidoreductase)



Primary activity: 20α-hydroxysteroid dehydrogenase: NADPH-dependent reduction of **C20 ketosteroids**

Secondary activities: Broad ketosteroid reductase activity toward **3- and 17-ketosteroids**

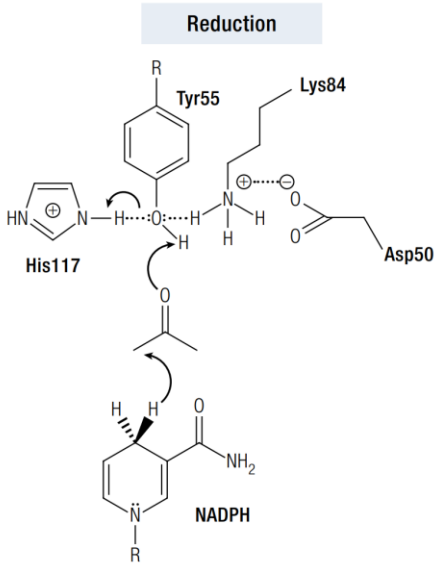
- dependence on substrate orientation
- dependence on active-site interactions



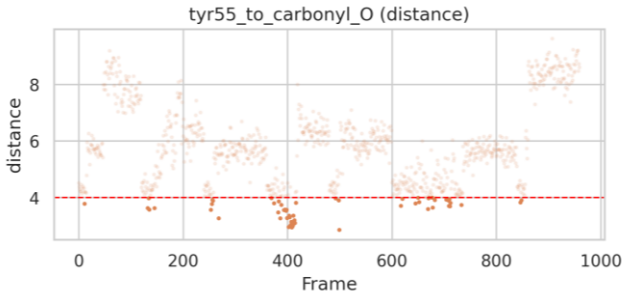
Progesterone

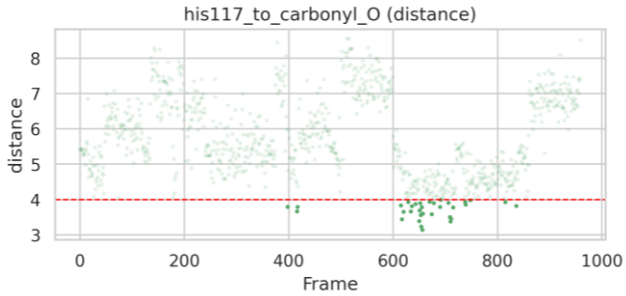
Substrate	Product	Cofactor	Reduction	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m (min^{-1} mM^{-1})
AKR1C1						
Progesterone^R	20α-HydroxyP	NADPH	20-Ketosteroid	5.7	0.93	210[a]
Androsterone	5α-Androstane-3α,17β-diol	NADPH	17-Ketosteroid	21	0.18	9[b]
5α-DHT^R	5α-Androstane-3α,17β-diol	NADPH	3-Ketosteroid	81	0.66	8[b]

NACtraces: MD trajectory analysis (AKR1C1)

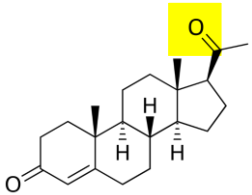


Reaction Mechanism Alignment

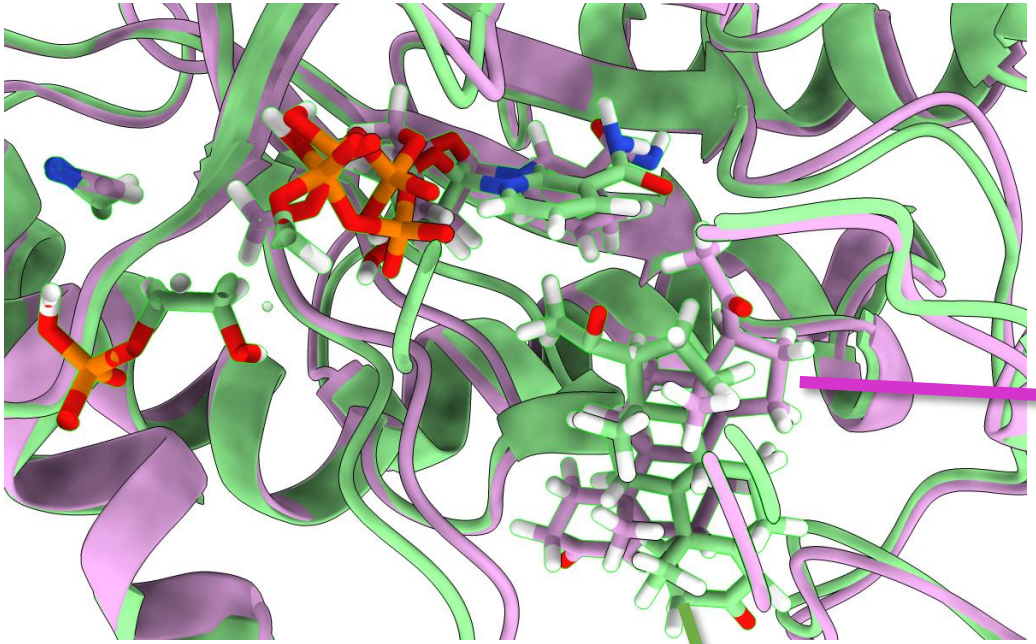




Progesterone



Case Study: AKR1C1 Ternary Complex (NADPH, Progesterone)



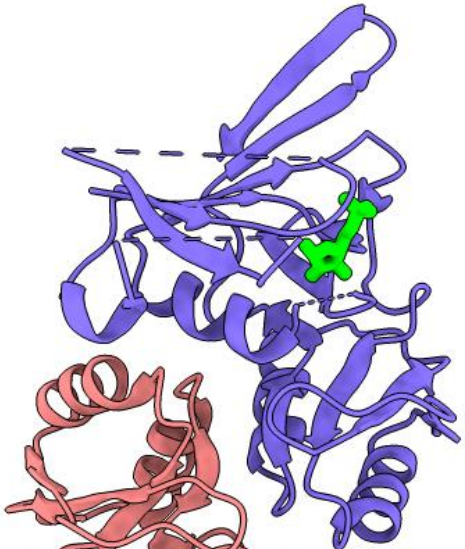
Progesterone,
after 20 ns MD

Progesterone, NAC starting pose

From MD-ready to MD-hard

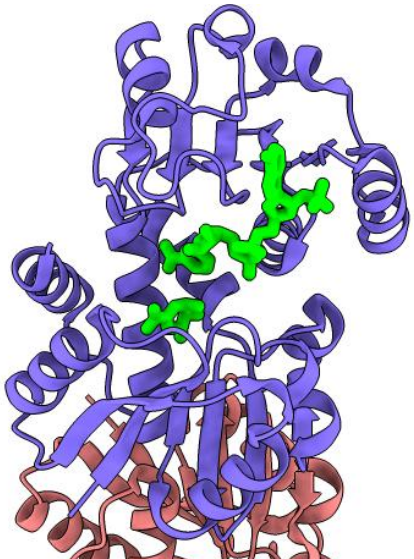
Easy: RpiA (1O8B)

- **Ribose-5-phosphate isomerase**
- **1 ligand**



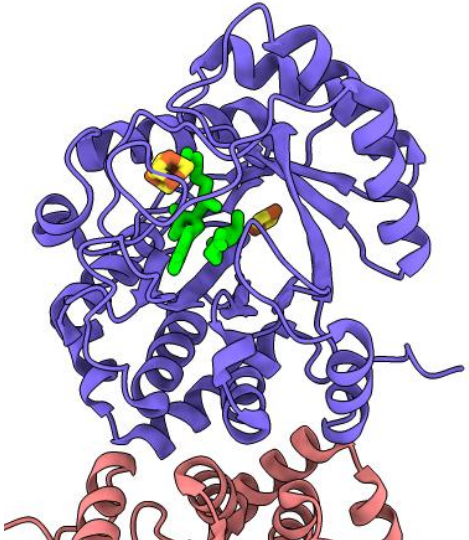
Mid: AroE (2EV9)

- **Shikimate dehydrogenase**
- **2 ligands**



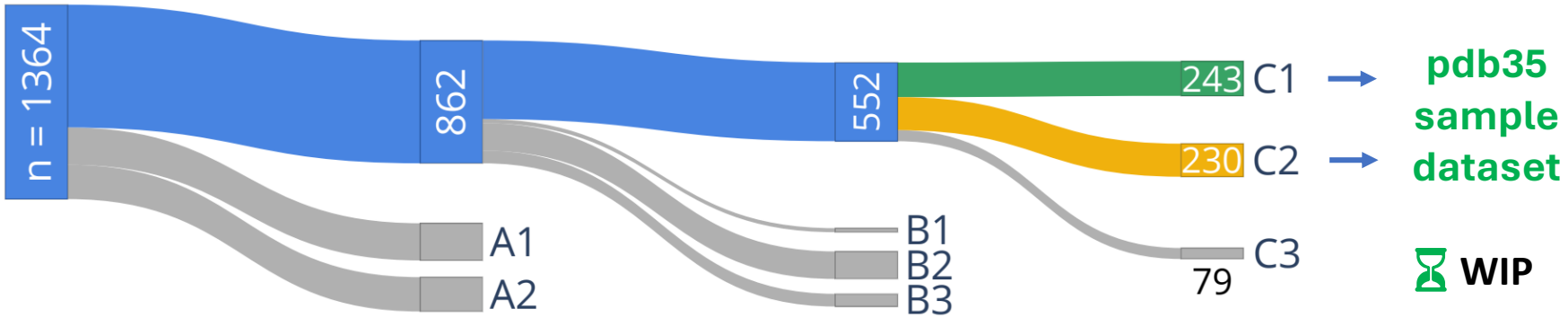
Hard: BioB (1R30)

- **Biotin synthase**
- **Fe–S clusters, radical SAM cofactor, redox states**



Enzyme-Filter project: Genome-scale MD-readiness

Gene filtering pipeline: iEC1364_W

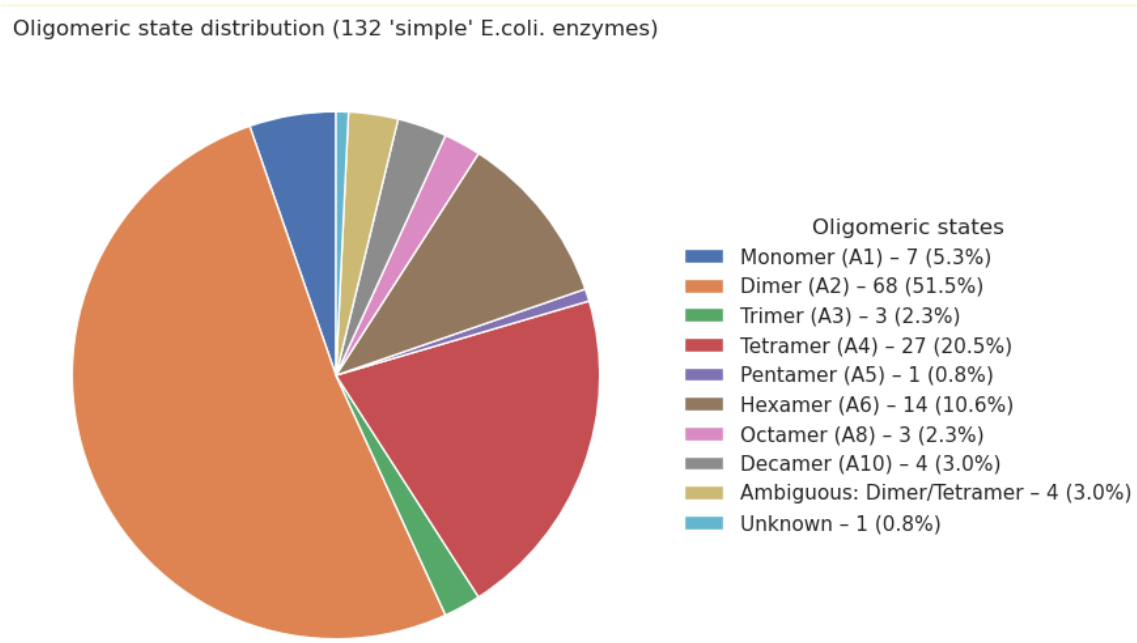


Stage 1: Annotations
A1: Rhea-linked reactions
A2: Simple genetics (GPR)

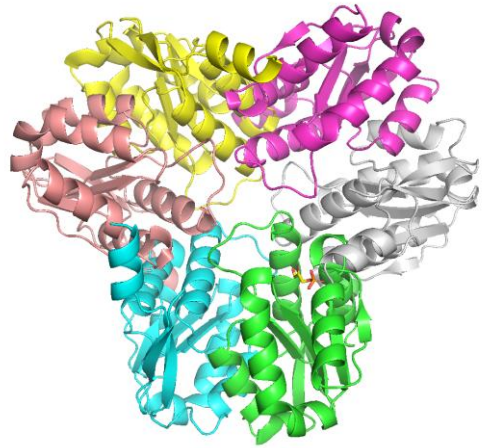
Stage 2: Size limits
B1: Protein size
B2: Cofactor complexity
B3: Substrate complexity

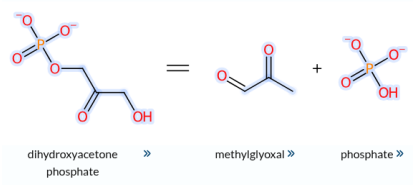
Stage 2: Modeling
C1: Simple MD Setup
C2: Moderate Setup Effort
C3: Cofactor/Redox Heavy

MD-readiness: Oligomeric States

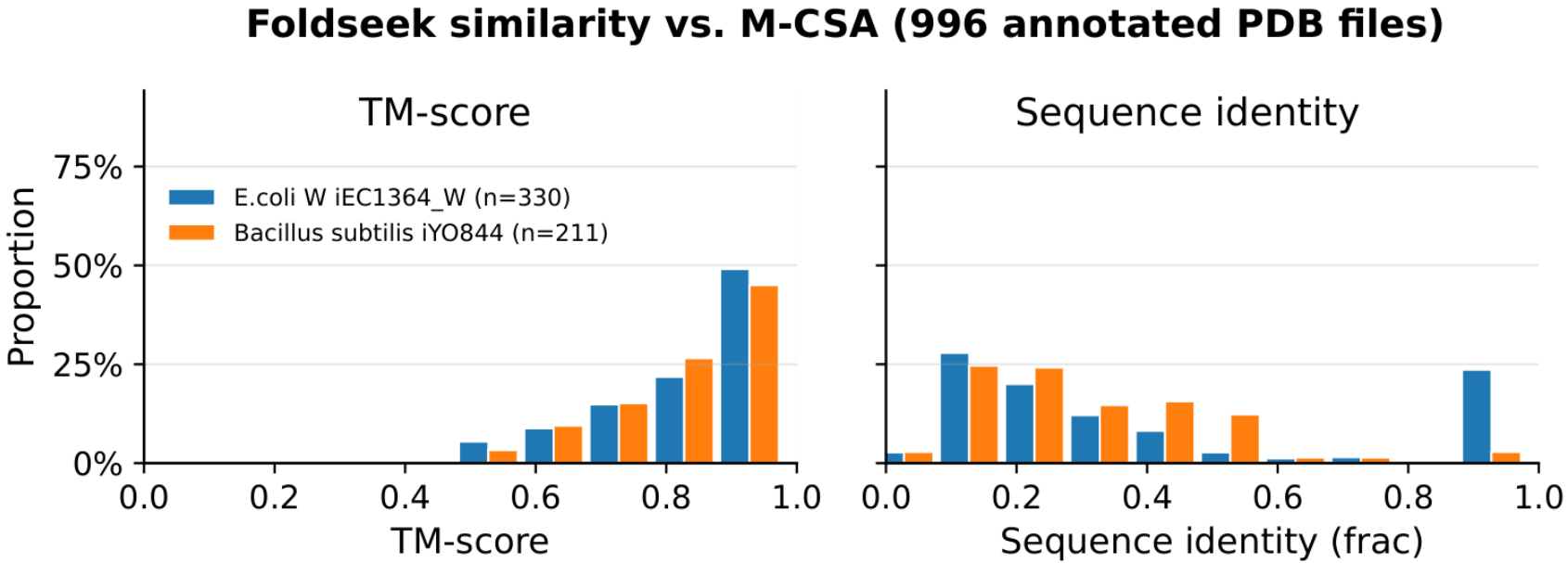


MgsA, EC 4.2.3.3 / 1EGH
Methylglyoxal synthase





MD-readiness: *E.coli / B. subtilis* structural availability

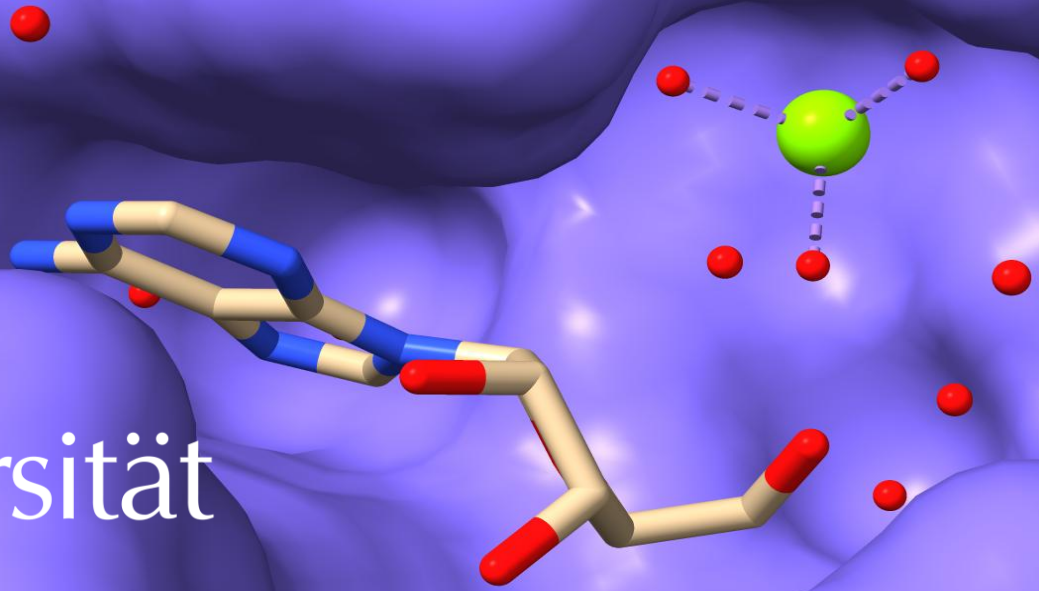


Project Status

- ⏳ Pdb35 sample dataset: WIP
  - characterize native substrates first
  - kinases?
  - dehydrogenases?
- ⏳ (Automated-) MD parameterization setup: WIP

- ☑ NACTraces project
- ☑ Enzyme-Filter project

Thanks!



universität wien

tbi MATOMIC novo nordisk fonden

RCSB ID: 1RMT
AphA phosphatase
complexed with adenosine, Mg^{2+}