

Computational(rational) protein engineering

Presented by:

Ladan Mafakher

Pharmaceutical biotechnology of Pasteur institute of medical sciences

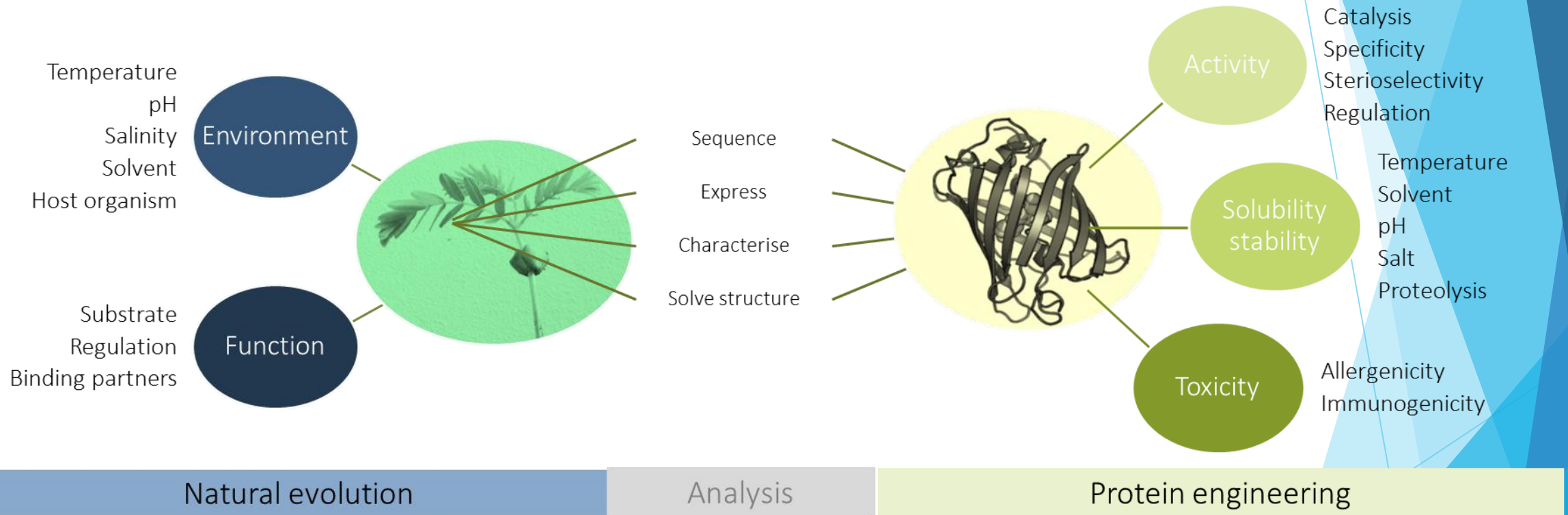
Thalassemia & Hemoglobinopathy Research center, Health research
institute, Ahvaz Jundishapur

Ladan.mafakher@gmail.com

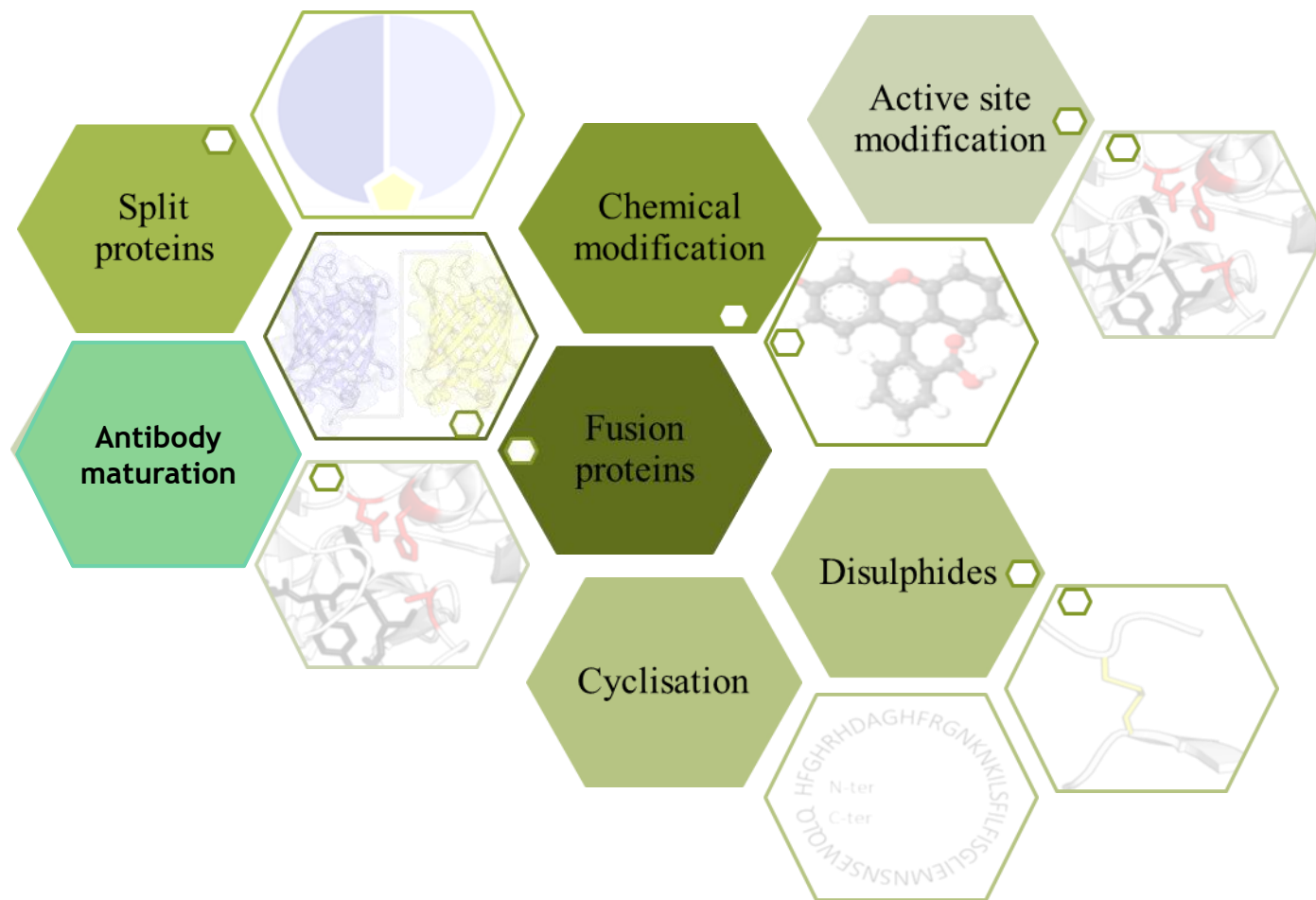
What is protein engineering?

- ▶ Proteins are molecular devices, in the nanometer scale, where biological function is exerted.
- ▶ Proteins are the building blocks of all cells.
- ▶ Nature has sampled only $\sim 10^{12}$ different possible proteins, while a 100 amino acid protein has 20^{100} potential sequence variations.
- ▶ protein engineering is to identify specific changes in the amino acid sequence and to alter such sequence for desired functional properties.
- ▶ construction of new proteins or enzymes with novel or desired functions, through the modification of amino acid sequences using recombinant DNA technology.
- ▶ The sizes and three dimensional conformations of protein molecules are also manipulated by protein engineering.

Aims of protein engineering



Topics



PEG-uricase in the management of treatment-resistant gout and hyperuricemia ☆

Merry R. Sherman ^{a,*}, Mark G.P. Saifer ^a, Fernando Perez-Ruiz ^b

^a Mountain View Pharmaceuticals, Inc., 3475-S Edison Way, Menlo Park, CA 94025, USA

^b Sección de Reumatología, Hospital De Cruces, Pza Cruces, Barakaldo 48903, Spain

Received 22 May 2007; accepted 10 June 2007

Available online 14 August 2007

Abstract

Hyperuricemia results from an imbalance between the rates of production and excretion of uric acid. Longstanding hyperuricemia can lead to gout, which is characterized by the deposition of monosodium urate monohydrate crystals in the joints and periarticular structures. Because such deposits are resolved very slowly by lowering plasma urate with available drugs or other measures, the symptoms of gout may become chronic. Persistent hyperuricemia may also increase the risk of renal and cardiovascular diseases. Unlike most mammals, humans lack the enzyme uricase (urate oxidase) that catalyzes the oxidation of uric acid to a more soluble product. This review describes the development of a poly(ethylene glycol) (PEG) conjugate of recombinant porcine-like uricase with which a substantial and persistent reduction of plasma urate concentrations has been demonstrated in a Phase 2 clinical trial. Two ongoing Phase 3 clinical trials include systematic assessments of gout symptoms, tophus resolution and quality of life, in addition to the primary endpoint of reduced plasma urate concentration.

© 2007 Elsevier B.V. All rights reserved.

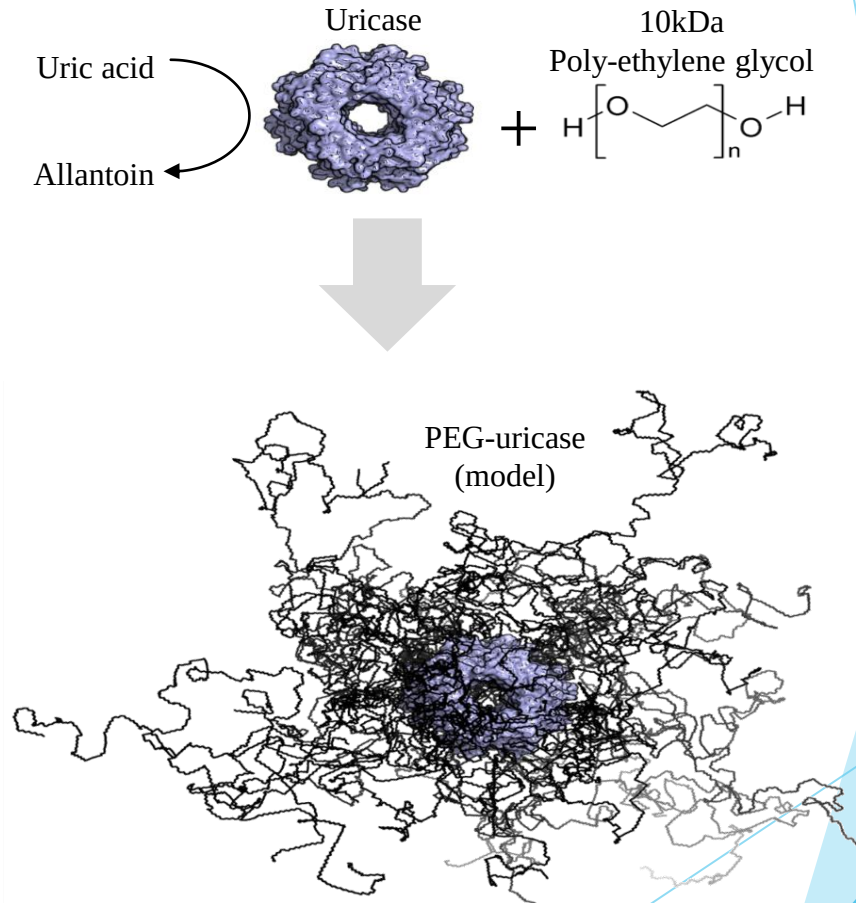
Keywords: Poly(ethylene glycol); Urate oxidase; Therapeutic enzyme; Gout; Hyperuricemia

Uricase

Engineering

Outcome

- ▶ Optimise uricase as gout treatment
- ▶ Reduce immunogenicity
- ▶ Increase serum half-life
 - Attached PEG polymers
 - Lysine coupling
 - Optimised PEG number and length
 - Maximise improvements
 - Avoid destabilisation or activity reduction
- Optimal PEG number and length
 - 10kDA polymers
 - 9 polymers per subunit of the tetramer
- 1000x reduced antigenicity
 - Also improved solubility at neutral pH
 - Also increased serum half-life



Oral Treatment With an Engineered Uricase, ALLN-346, Reduces Hyperuricemia, and Uricosuria in Urate Oxidase-Deficient Mice

Kateryna Pierzynowska^{1,2,3*}, Aditi Deshpande⁴, Nadiia Mosiichuk⁵, Robert Terkeltaub⁶, Paulina Szczurek⁷, Eduardo Salido⁸, Stefan Pierzynowski^{2,3,9} and Danica Grujic^{4*}

OPEN ACCESS

Edited by:

Wei-Jun Qian,
Pacific Northwest National Laboratory
(DOE), United States

Reviewed by:

Ashish K. Solanki,
Medical University of South Carolina,
United States
Hirotaki Yokoi,
Kyoto University Hospital, Japan

*Correspondence:

Kateryna Pierzynowska
kateryna.goncharova@biol.lu.se
Danica Grujic
dgrujic@allonpharma.com

Specialty section:

This article was submitted to
Nephrology,
a section of the journal
Frontiers in Medicine

Received: 02 July 2020

Accepted: 29 September 2020

Published: 24 November 2020

Citation:

Pierzynowska K, Deshpande A,
Mosiichuk N, Terkeltaub R, Szczurek P,
Salido E, Pierzynowski S and Grujic D
(2020) Oral Treatment With an
Engineered Uricase, ALLN-346,
Reduces Hyperuricemia, and
Uricosuria in Urate Oxidase-Deficient
Mice. *Front. Med.* 7:569215.
doi: 10.3389/fmed.2020.569215

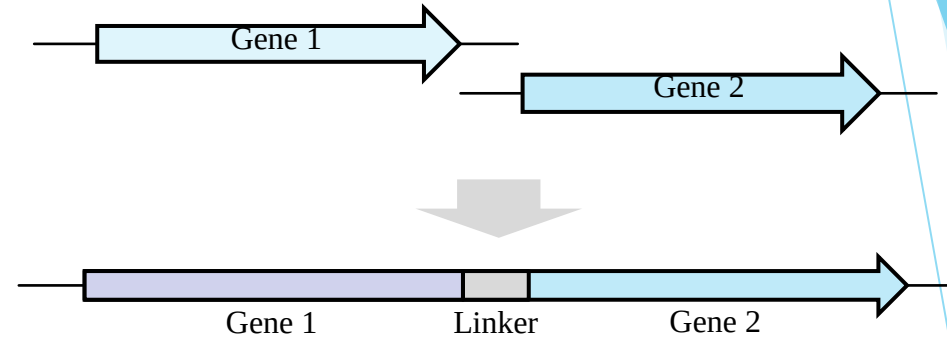
¹ Department of Animal Physiology, Kolanowski Institute of Animal Nutrition and Physiology Polish Academy of Sciences, Jablonna, Poland, ² Department of Biology, Lund University, Lund, Sweden, ³ SGP+Group, Trollhättan, Sweden, ⁴ Allon Pharmaceuticals, Newton, MA, United States, ⁵ Department of Biochemistry and Biotechnology, Vasyk Stolyk Procarpathian National University, Ivano-Frankivsk, Ukraine, ⁶ VA Medical Center, University of California, San Diego, La Jolla, CA, United States, ⁷ Department of Animal Nutrition and Food Sciences, National Research Institute of Animal Production, Balice, Poland, ⁸ Hospital Universitario de Canarias, Universidad La Laguna & Center for Rare Diseases (CIBERER), Tenerife, Spain, ⁹ Department of Biology, Institute Rural Medicine, Lublin, Poland

Limitations in efficacy and/or tolerance of currently available urate-lowering therapies (ULTs), such as oral xanthine oxidase inhibitors, uricosurics, and intravenous uricase agents contribute to the development of refractory gout. Renal excretion is the major route of uric acid elimination, but the intestinal tract plays an increasingly recognized role in urate homeostasis, particularly in chronic kidney disease (CKD) in which the renal elimination of urate is impaired. We targeted intestinal degradation of urate *in vivo* with ALLN-346, an orally administered, engineered urate oxidase, optimized for proteolytic stability, and activity in the gut. We tested ALLN-346 in uricase/urate oxidase deficient mice (URKO mice) with severe hyperuricemia, hyperuricosuria, and uric acid crystalline obstructive nephropathy. A total of 55 male and female URKO mice were used in the two consecutive studies. These seminal, proof-of-concept studies aimed to explore both short- (7-day) and long-term (19-day) effects of ALLN-346 on the reduction of plasma and urine urate. In both the 7- and 19-day studies, ALLN-346 oral therapy resulted in the normalization of urine uric acid excretion and a significant reduction of hyperuricemia by 44 and 28% when therapy was given with food over 24 h or was limited for up to 6 h, respectively. Fractional excretion of uric acid (FEUA) was normalized with ALLN-346 therapy. Oral enzyme therapy with engineered urate oxidase (ALLN-346) designed to degrade urate in the intestinal tract has the potential to reduce hyperuricemia and the renal burden of filtered urate in patients with hyperuricemia and gout with and without CKD.

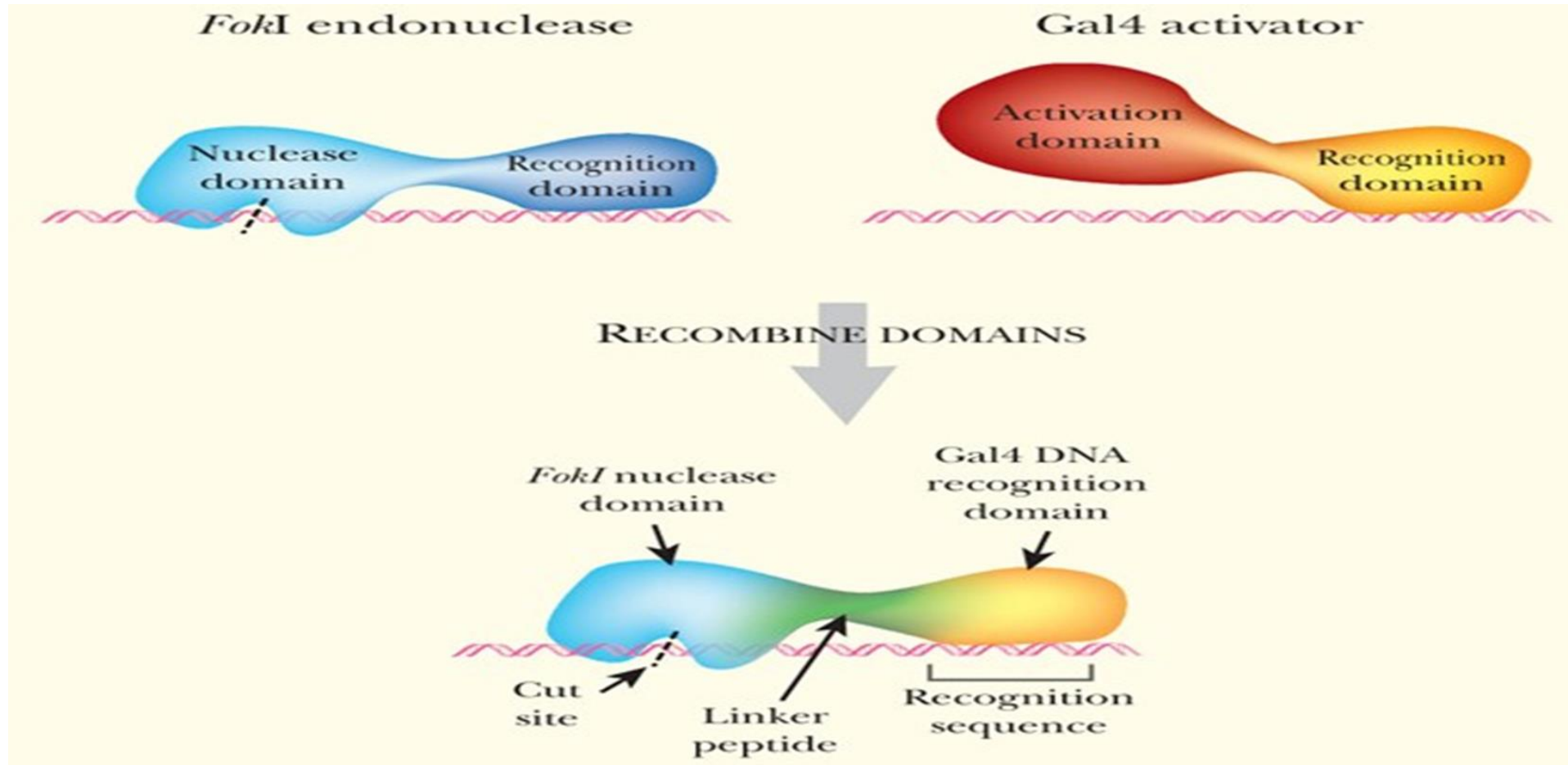
Keywords: gout, CKD, urolithiasis, urate-lowering therapy, ABCG2

Fusion proteins

- ▶ Creation
- ▶ Remove stop codon of first gene
- ▶ Ligate genes together in frame
- ▶ Include linker codons
- ▶ Aims
- ▶ Combine the properties of the components
- ▶ E.g. Addition of antibody Fc fragment to proteins increases their serum half-life
- ▶ Co-localise the components
- ▶ E.g. Set of enzymes that work in a reaction pathway
- ▶ Considerations
- ▶ Linker length and flexibility
- ▶ Ability for proteins rotate relative to each other
- ▶ Distance between protein components
- ▶ Protease resilience
- ▶ Ability for domains to fold



Fusion proteins



Research Article

Computational Design of a DNA- and Fc-Binding Fusion Protein

**Jonas Winkler,¹ Giuliano Armano,² J. Nikolaj Dybowski,¹ Oliver Kuhn,¹
Filippo Ledda,² and Dominik Heider¹**

¹Department of Bioinformatics, Center for Medical Biotechnology, University of Duisburg-Essen, Universitaetsstraße 1-5,
45117 Essen, Germany

²Department of Electrical and Electronic Engineering, University of Cagliari, Piazza d'Armi, 09123 Cagliari, Italy

Correspondence should be addressed to Dominik Heider, dominik.heider@uni-due.de

Received 29 April 2011; Revised 16 June 2011; Accepted 22 June 2011

Academic Editor: Shandar Ahmad

Copyright © 2011 Jonas Winkler et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Computational design of novel proteins with well-defined functions is an ongoing topic in computational biology. In this work, we generated and optimized a new synthetic fusion protein using an evolutionary approach. The optimization was guided by directed evolution based on hydrophobicity scores, molecular weight, and secondary structure predictions. Several methods were used to refine the models built from the resulting sequences. We have successfully combined two unrelated naturally occurring binding sites, the immunoglobulin Fc-binding site of the Z domain and the DNA-binding motif of MyoD bHLH, into a novel stable protein.



Rational affinity maturation of anti-amyloid antibodies with high conformational and sequence specificity

Received for publication, June 18, 2020, and in revised form, February 5, 2021. Published, Papers in Press, March 4, 2021.
<https://doi.org/10.1016/j.jbc.2021.100508>

Alec A. Desai^{1,2}, Matthew D. Smith^{1,2}, Yulei Zhang^{1,2}, Emily K. Makowski^{2,3}, Julia E. Gerson⁴, Edward Ionescu^{1,2}, Charles G. Starr^{1,2,3}, Jennifer M. Zupancic^{1,2}, Shannon J. Moore^{5,6}, Alexandra B. Sutter^{4,7}, Magdalena I. Ivanova^{4,7}, Geoffrey G. Murphy^{5,6}, Henry L. Paulson^{4,5,8}, and Peter M. Tessier^{1,2,3,5,9,*}

From the ¹Department of Chemical Engineering, ²Biointerfaces Institute, ³Department of Pharmaceutical Sciences, ⁴Department of Neurology, ⁵Protein Folding Disease Initiative, ⁶Department of Molecular & Integrative Physiology, ⁷Biophysics Program, ⁸Michigan Alzheimer's Disease Center, ⁹Department of Biomedical Engineering, University of Michigan, Ann Arbor, Michigan, USA

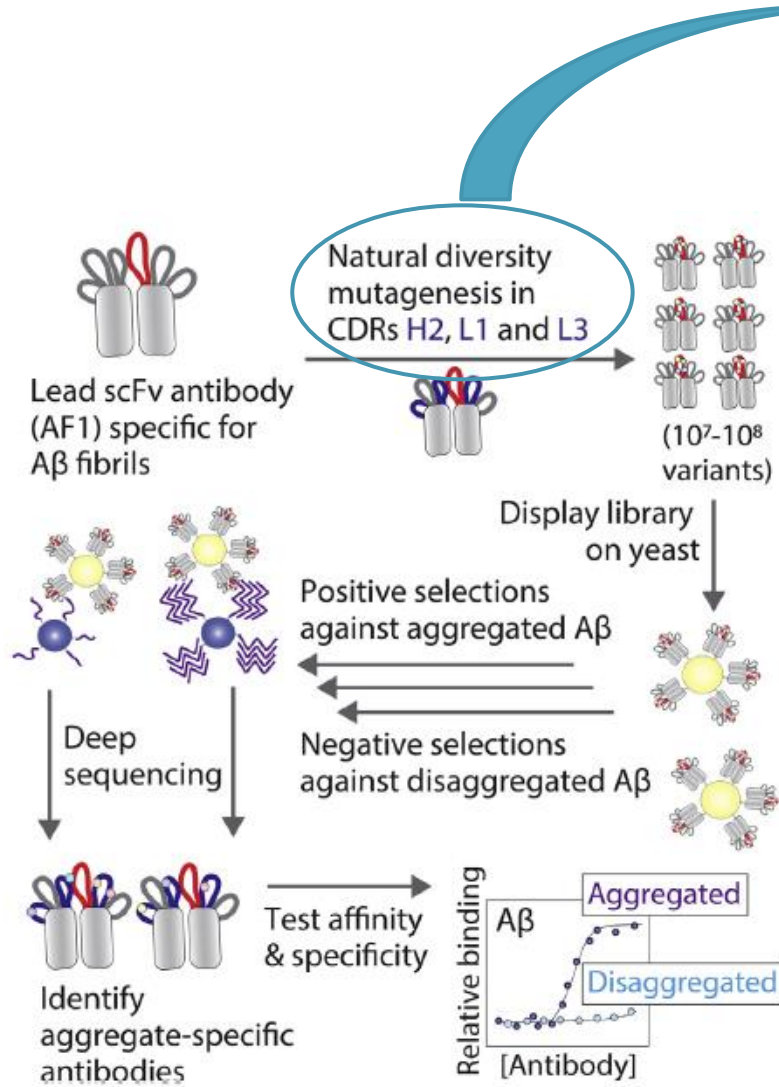
Edited by Paul Fraser

The aggregation of amyloidogenic polypeptides is strongly linked to several neurodegenerative disorders, including Alzheimer's and Parkinson's diseases. Conformational antibodies that selectively recognize protein aggregates are leading therapeutic agents for selectively neutralizing toxic aggregates, diagnostic and imaging agents for detecting disease, and biomedical reagents for elucidating disease mechanisms. Despite their importance, it is challenging to generate high-quality conformational antibodies in a systematic and site-specific manner due to the properties of protein aggregates (hydrophobic, multivalent, and heterogeneous) and limitations of immunization (uncontrolled antigen presentation and immunodominant epitopes). Toward addressing these challenges, we have developed a systematic directed evolution procedure for affinity maturing antibodies against Alzheimer's A β fibrils and selecting variants with strict conformational and

Of the many human disorders facing our society today, neurodegenerative diseases such as Alzheimer's and Parkinson's diseases are arguably the most menacing and least treatable (1, 2). These diseases – which are linked to the formation of toxic prefibrillar oligomers and amyloid fibrils – are particularly concerning because their frequency of occurrence is linked to age and, thus, the number of cases is expected to increase as life expectancy increases in the coming years due to significant advances in treating other human disorders such as cancer and heart diseases.

Conformational antibodies specific for different conformers of amyloid-forming proteins are important for detecting, disrupting, and reversing toxic protein aggregation (3, 4). Several previous reports have demonstrated creative methods for using immunization (4–12), autoantibody screening (5, 13–22), directed evolution (23–26), and rational design methods

Affinity maturation



A

	HCDR2					LCDR1				LCDR3	
Position	52	53	54	56	61	28	30	31	34	92	94
Wild type	Y	T	N	Y	D	D	N	T	A	Y	T
	S	S	G	S	D	I	S	N	A	D	S
	Y	D	S	N	P	G	N	S	S	Y	T
	N	Y	D	T	Q	S	Y	Y	N	N	W
	K	N	N	D	A	V	G	K	H	G	Y
≤50	I	G	F	Y	V	N	K	T	Y	T	F
≤40	D	H	T	G	E	L	R	R	D	S	L
≤30	R	I	K	E	T	D	D	D	C	A	G
≤20	W	T	A	V	L	T	T	H	Q	H	N
≤12	T	W	R	A	R	H	A	Q	G	L	R
≤6	F	R	L	R	G	F	I	I	T	I	A
≤3	A	A	E	I	S	Y	V	A	V	C	I

B

	HCDR2					LCDR1				LCDR3	
Position	52	53	54	56	61	28	30	31	34	92	94
Wild type	Y	T	N	T	D	D	N	T	A	Y	T
Designed library	A	D	D	A	A	D	G	D	D	D	N
	N	N	N	N	G	N	N	N	N	N	G
	T	S	S	S	G	S	S	S	S	S	T
	Y	S	S	S	V	V	Y	Y	Y	Y	Y

Machine learning-aided engineering of hydrolases for PET depolymerization

<https://doi.org/10.1038/s41586-022-04599-z>

Received: 10 October 2021

Accepted: 28 February 2022

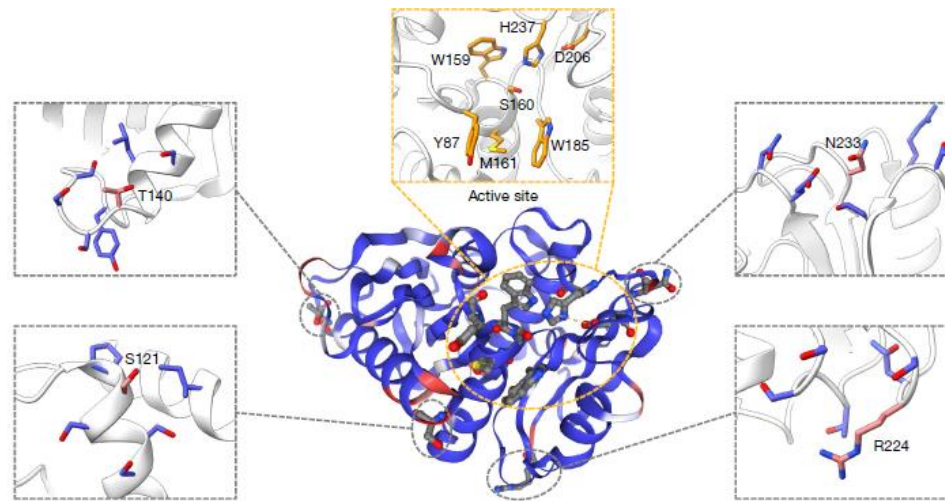
Published online: 27 April 2022



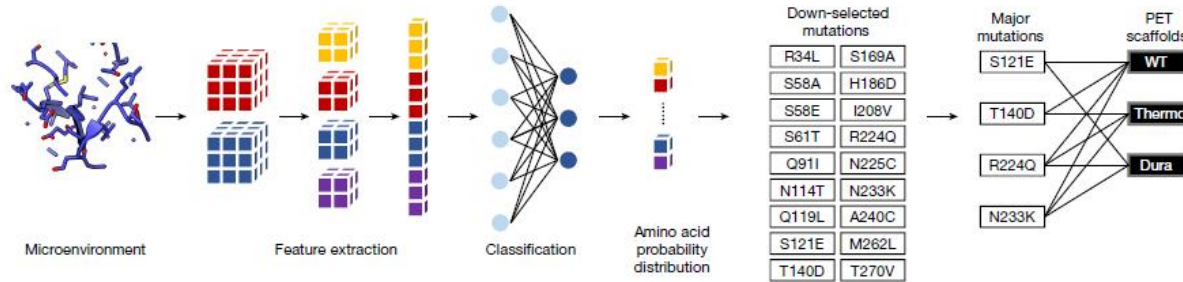
Hongyuan Lu¹, Daniel J. Diaz², Natalie J. Czarnecki¹, Congzhi Zhu¹, Wantae Kim¹, Raghav Shroff^{3,4}, Daniel J. Acosta³, Bradley R. Alexander³, Hannah O. Cole^{1,3}, Yan Zhang³, Nathaniel A. Lynd¹, Andrew D. Ellington³ & Hal S. Alper¹✉

Plastic waste poses an ecological challenge^{1–3} and enzymatic degradation offers one, potentially green and scalable, route for polyesters waste recycling⁴. Poly(ethylene terephthalate) (PET) accounts for 12% of global solid waste⁵, and a circular carbon economy for PET is theoretically attainable through rapid enzymatic depolymerization followed by repolymerization or conversion/valorization into other products^{6–10}. Application of PET hydrolases, however, has been hampered by their lack of robustness to pH and temperature ranges, slow reaction rates and inability to directly use untreated postconsumer plastics¹¹. Here, we use a structure-based, machine learning algorithm to engineer a robust and active PET hydrolase. Our mutant and scaffold combination (FAST-PETase: functional, active, stable and tolerant PETase) contains five mutations compared to wild-type PETase (N233K/R224Q/S121E from prediction and D186H/R280A from scaffold) and shows superior PET-hydrolytic activity relative to both wild-type and engineered alternatives¹² between 30 and 50 °C and a range of pH levels. We demonstrate that untreated, postconsumer-PET from 51 different thermoformed products can all be almost completely degraded by FAST-PETase in 1 week. FAST-PETase can also depolymerize untreated, amorphous portions of a commercial water bottle and an entire thermally pretreated water bottle at 50 °C. Finally, we demonstrate a closed-loop PET recycling process by using FAST-PETase and resynthesizing PET from the recovered monomers. Collectively, our results demonstrate a viable route for enzymatic plastic recycling at the industrial scale.

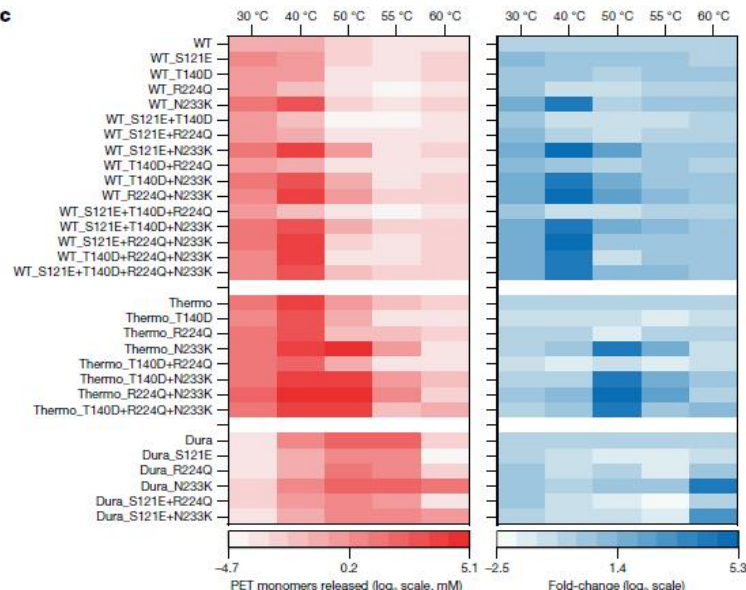
a



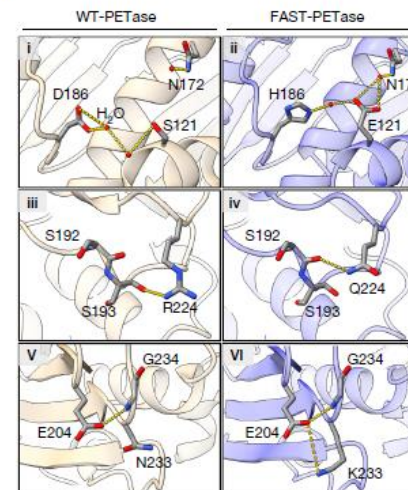
b



c

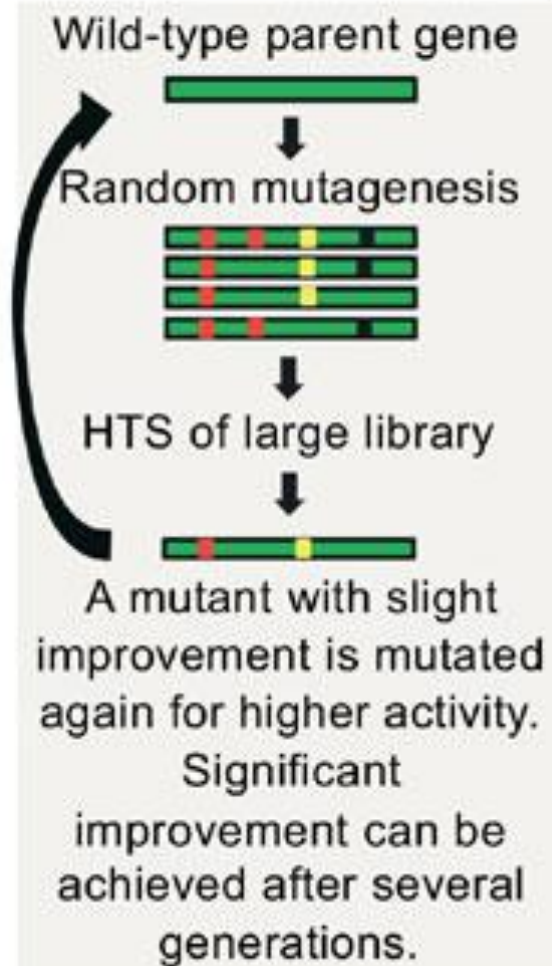


d



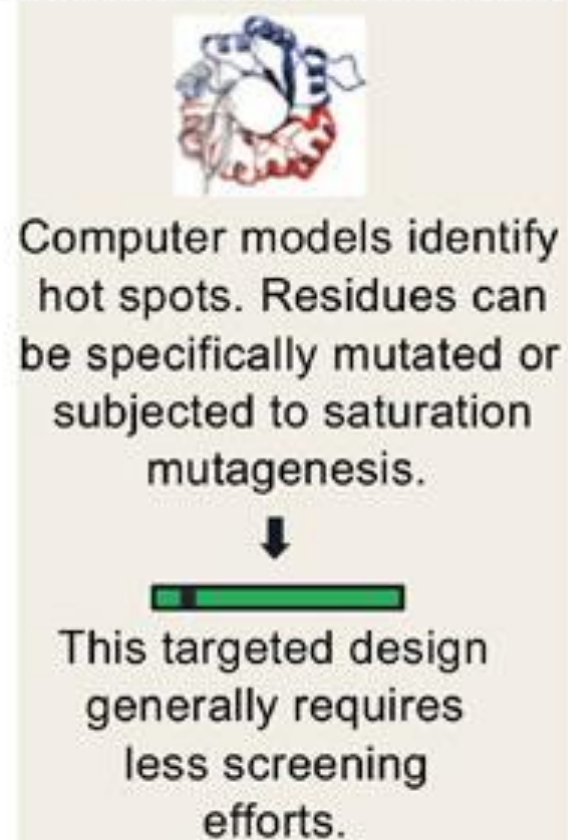
Protein engineering strategy

A. Directed evolution



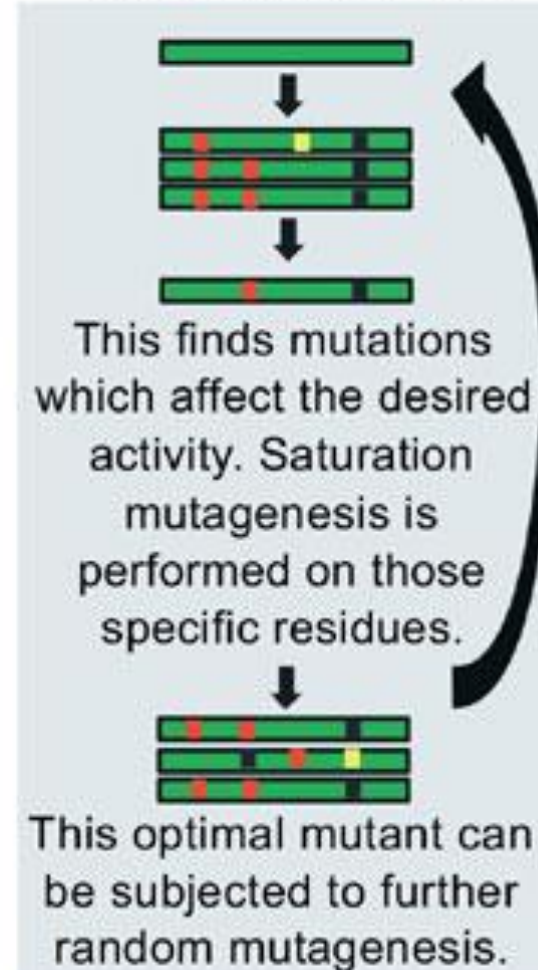
B. Rational design

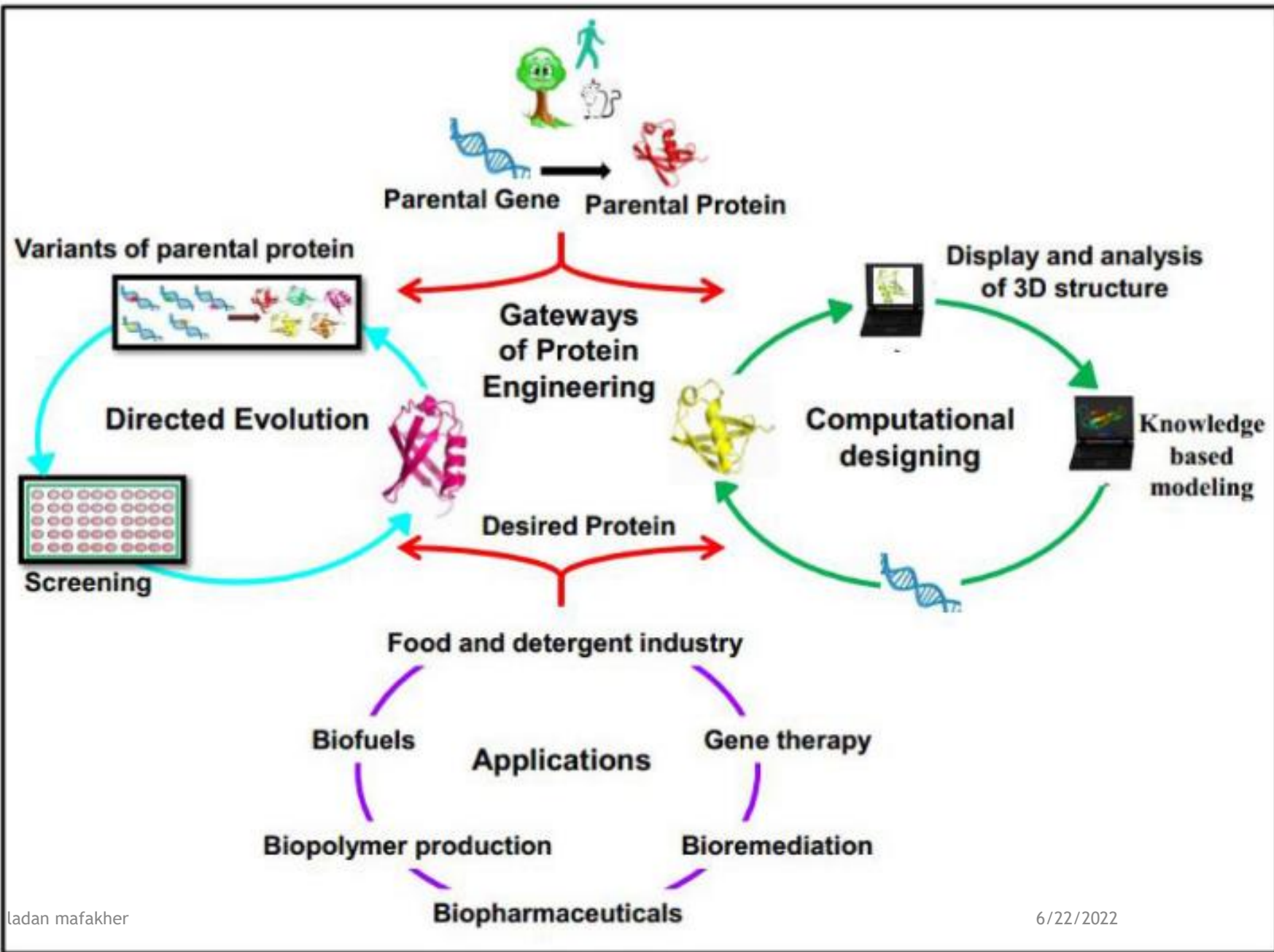
A priori structural and mechanistic information of the protein is required

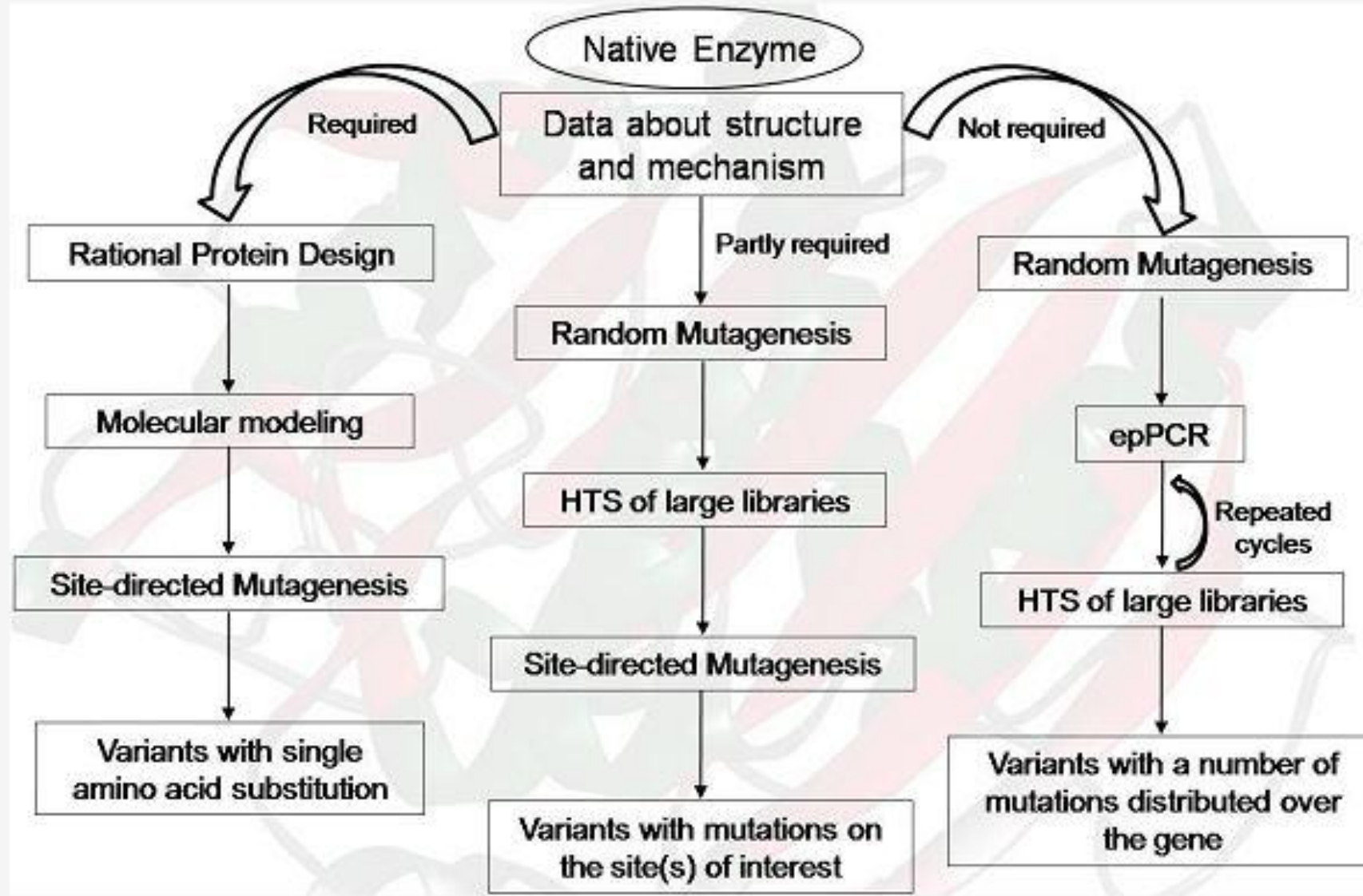


C. Semi-rational design

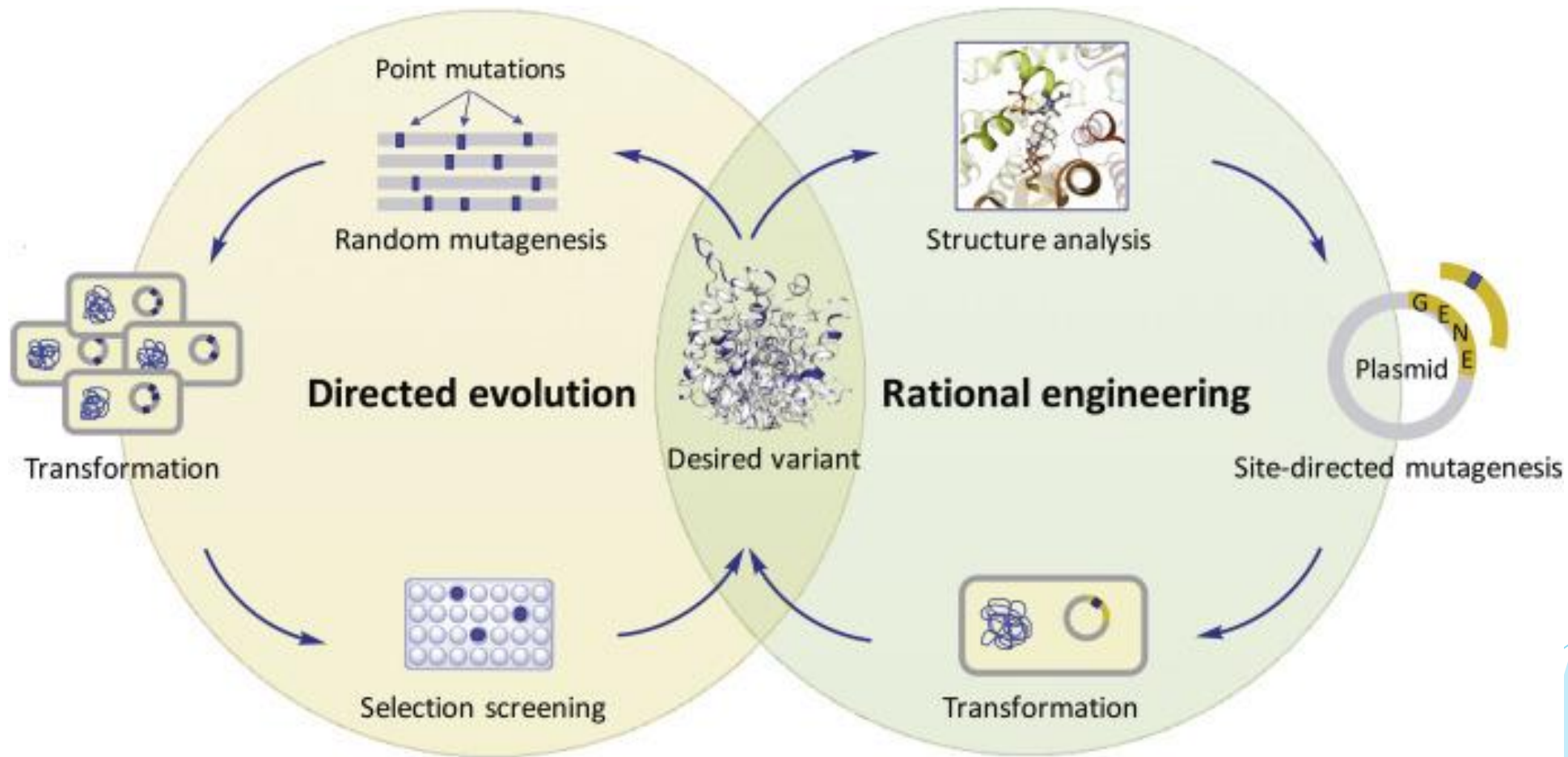
Combination uses directed evolution to create random diversity







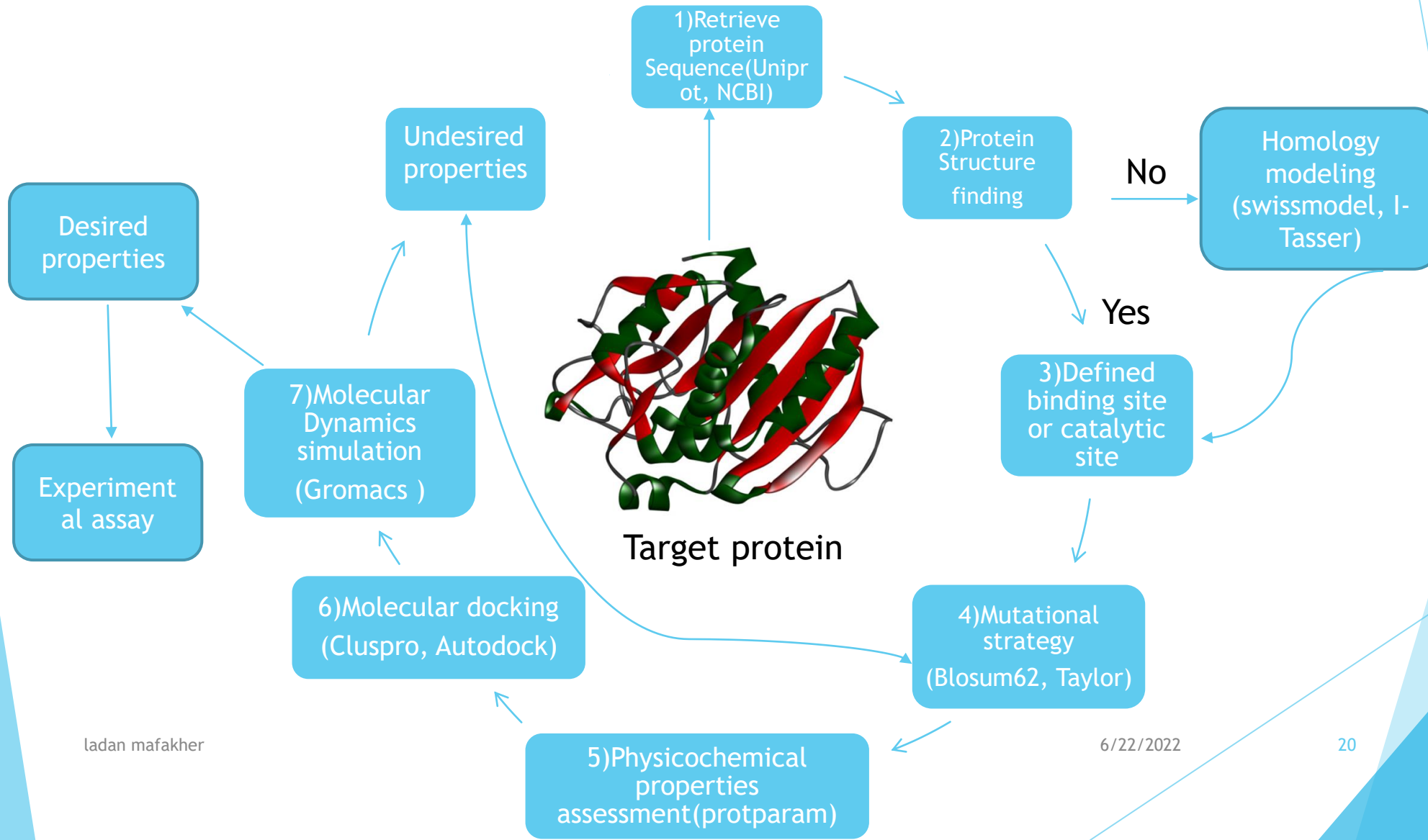
Advantage of rational protein engineering



Workflow of computational protein engineering

- 1 • In silico analysis and design
- 2 • Gene library construct
- 3 • Protein display/expression
- 4 • High throughput screening/selection
- 5 • Validation and characterization
- 6 • Selection the most potential protein

Steps of rational protein engineering



1) Retrieve protein(NCBI)

<https://www.ncbi.nlm.nih.gov/protein>

NCBI Resources How To Sign in to NCBI

Protein Protein human EGFR Search

Create alert Advanced Help

Species

- Animals (2,250)
- Plants (7)
- Fungi (102)
- Bacteria (402)
- Viruses (180)
- Customize ...

Source databases

- PDB (621)
- RefSeq (897)
- UniProtKB / Swiss-Prot (287)
- Customize ...

Sequence length

Custom range...

Molecular weight

Custom range...

Summary 20 per page Sort by Default order

Send to: Filters: Manage Filters

GENE

Was this helpful?  

[EGFR – epidermal growth factor receptor](#)

[Homo sapiens \(human\)](#)

Processed peptides: Epidermal growth factor receptor

Also known as: ERBB, ERBB1, HER1, NISBD2, PIG61, mENA

GeneID: 1956

[RefSeq transcripts \(9\)](#) [RefSeq proteins \(9\)](#) [RefSeqGene \(1\)](#) [PubMed \(5,247\)](#)

Genome Browser

BLAST

Download

Results by taxon

Top Organisms [\[Tree\]](#)

- Homo sapiens (1758)
- Listeria monocytogenes (213)
- Mus musculus (204)
- Human alphaherpesvirus 1 (154)
- Rattus norvegicus (70)
- All other taxa (611)

[More...](#)

NCBI Resources How To Sign in to NCBI

Protein Protein Search

Advanced Help

FASTA

Send to:

Change region shown

epidermal growth factor receptor isoform a precursor [Homo sapiens]

NCBI Reference Sequence: NP_005219.2

[GenPept](#) [Identical Proteins](#) [Graphics](#)

>NP_005219.2 epidermal growth factor receptor isoform a precursor [Homo sapiens]
MRPSTAGAAALLAALCPASRALEKKVCQGTISNKLTLQGTGFDHFLSLQRMFNNEVVLGNLEITYV
QRNYDLSFLKTIQEVAGYVLIALNIVERIPLNLIIRGNMYENSVALAVLSNYDANKTGLKELPMRNL
QELHGAIVRFSNNPACNIVESIQMRDIVSSDFLSNMSMDFQNLHSGCQKCDPSCPNWSCWAGEENCQKL
TKIICAQCCSGRCRGKSPSDCCCHNQCAAGCTGPRESDCLVCRKFRDEATCKDTCPPFLMYNPITYQMDVN
PEGKYSFGATCVKKCPNRYVVDHSGCVACGADSYEMEDGVRKCKKCEGPRKVCNGIGIGEFKDSLS
INATNIKHFNCTISISGLHILPVAFRGDSFTHTPLDPQELDIKTVKEITGFLLIQAWPENRTDLHAF
ENLEIIRGRTKQHGGFSLAVVSLNITSLGLRSLKEISDGDVVISGNKNLCYANTINWKKLFGTSGQKTKI
ISNRGENSKATGQVCHALCSPEGCWGPEDRCVSRNVRGRCVDCNLLGEPEPFVENSECIQCHP
ECLPQAMNITCTGRGPDNCIQAHYIDGPHCVKTCFAGVMGNNITLVNKYADAGHVCHLCHPNTYGTG
PGLEGCPNTNGKIPSIATGMVGAALLLVVALGIGLFMRRRHIVRKRTLRLQLERELVEPLTPSGEAPN
QALLRILKETEFKKIKVLGSGAGFTVYKGLWIPGEKVKIPVAIKELREATSPKANKEILDEAYVMASVD
NPHVCRLLGICLTSTVQLITQLMPFGCLLDYVREHKDNIGSQYLLNWCQVIAKGMNVLDRRLVHRDLAA
RNVLVKTPQHVKITDPLGLAKLLGAEKEYHAEGGKVPKIMMALESILHRIYTHQSDVWSYGVTVWELMTF
GSKPYDGIPIASEISSILEGERLPQFPICITDVYIMVCKWMIDADSRPKFRELIIEFSKMDARPQRYLV
IQGDERMHLPSPTDSNFYRALMDEEDMDVDVDADEYLIPQQGFFSSPSTSRTPLLSSLSATSNSTVACI
DRNLGQSCPIKEDSFLQRYSSDPTGALTSDIDDTFLVPVEYINQSVKRPAGSVQNPVYHNQPLNPAPS
RDPHYQDPHSTAVGNFYLNTVQPTCVNSTFDSPAHWAQKQSHQISLDNPDQDQDFPKEAKPNGIFKGS
TAENAEYLRVAPQSSEFIGA

Analyze this sequence

Run BLAST

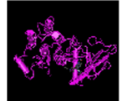
Identify Conserved Domains

Highlight Sequence Features

Find in this Sequence

Show in Genome Data Viewer

Protein 3D Structure

 Crystal Structure of EGFR T790M mutant in complex with naquotinib
PDB: 5Y9T
Source: Homo sapiens
Method: X-ray Diffraction
Resolution: 3.25 Å

[See all 151 structures...](#)

[Articles about the EGFR gene](#)

1) Retrieve protein(UNIPROT)

https://www.uniprot.org/

UniProtKB results

UniProtKB consists of two sections:

- Reviewed (Swiss-Prot) - Manually annotated**
Records with information extracted from literature and curator-evaluated computational analysis.
- Unreviewed (TrEMBL) - Computationally analyzed**
Records that await full manual annotation.

The UniProt Knowledgebase (UniProtKB) is the central hub for the collection of functional information on proteins, with accurate and rich annotation. In addition to capturing the core data mandatory for each UniProtKB entry (mainly, the amino acid sequence, protein name or description, taxonomic data and citation information), as much annotation information as possible is added.

Help UniProtKB help video Other tutorials and videos Downloads

UniProtKB - P00533 (EGFR_HUMAN)

Display

Entry

Protein | Epidermal growth factor receptor

Gene | EGFR

Organism | Homo sapiens (Human)

Status | Reviewed - Annotation score: ★★★★★ - Experimental evidence at protein level¹

Function¹

Receptor tyrosine kinase binding ligands of the EGF family and activating several signaling cascades to convert extracellular cues into appropriate cellular responses (PubMed:2790960, PubMed:10805725, PubMed:27153536). Known ligands include EGF, TGFα/TGF-α, AREG, epigen/EPGN, BTC/betacellulin, epiregulin/EREG and HBEGF/heparin-binding EGF (PubMed:2790960, PubMed:7679104, PubMed:8144591, PubMed:9419975, PubMed:15611079, PubMed:12297049, PubMed:27153536, PubMed:20837704). Ligand binding triggers receptor homo- and/or heterodimerization and autophosphorylation on key cytoplasmic residues. The phosphorylated receptor recruits adapter proteins like GRB2 which in turn activates complex downstream signaling cascades. Activates at least 4 major downstream signaling cascades including the RAS-RAF-MEK-ERK, PI3 kinase-AKT, PLCγ-PKC and STATs modules (PubMed:27153536). May also activate the NF-κB signaling cascade (PubMed:11116146). Also directly phosphorylates other proteins like RGS16, activating its GTPase activity and probably coupling the EGF receptor signaling to the G protein-coupled receptor signaling (PubMed:11602604). Also phosphorylates MUC1 and increases its interaction with SRC and CTNNB1/beta-catenin (PubMed:11403589). Plays a role in enhancing learning and memory performance (By similarity). [By similarity](#) [20 Publications](#)

Isoform 2 may act as an antagonist of EGF action.

(Microbial infection) Acts as a receptor for hepatitis C virus (HCV) in hepatocytes and facilitates its cell entry. Mediates HCV entry by promoting the formation of the CD81-CLDN1 receptor complexes that are essential for HCV entry and by enhancing membrane fusion of cells expressing HCV envelope glycoproteins. [1 Publication](#)

Catalytic activity¹

ATP + L-tyrosyl-[protein] = ADP + H⁺ + O-phospho-L-tyrosyl-[protein] [PROSITE-ProRule annotation](#) [6 Publications](#)

EC:3.7.10.1 [PROSITE-ProRule annotation](#) [6 Publications](#)

Source: Rhea. [Hide](#)

Chemical structures: ATP, L-tyrosyl-[protein], ADP, H⁺, O-phospho-L-tyrosyl-[protein]

UniProtKB - P00533 (EGFR_HUMAN)

Display

Entry

Protein | Epidermal growth factor receptor

Gene | EGFR

Organism | Homo sapiens (Human)

Status | Reviewed - Annotation score: ★★★★★ - Experimental evidence at protein level¹

Function¹

Receptor tyrosine kinase binding ligands of the EGF family and activating several signaling cascades to convert extracellular cues into appropriate cellular responses (PubMed:2790960, PubMed:10805725, PubMed:27153536). Known ligands include EGF, TGFα/TGF-α, AREG, epigen/EPGN, BTC/betacellulin, epiregulin/EREG and HBEGF/heparin-binding EGF (PubMed:2790960, PubMed:7679104, PubMed:8144591, PubMed:9419975, PubMed:15611079, PubMed:12297049, PubMed:27153536, PubMed:20837704). Ligand binding triggers receptor homo- and/or heterodimerization and autophosphorylation on key cytoplasmic residues. The phosphorylated receptor recruits adapter proteins like GRB2 which in turn activates complex downstream signaling cascades. Activates at least 4 major downstream signaling cascades including the RAS-RAF-MEK-ERK, PI3 kinase-AKT, PLCγ-PKC and STATs modules (PubMed:27153536). May also activate the NF-κB signaling cascade (PubMed:11116146). Also directly phosphorylates other proteins like RGS16, activating its GTPase activity and probably coupling the EGF receptor signaling to the G protein-coupled receptor signaling (PubMed:11602604). Also phosphorylates MUC1 and increases its interaction with SRC and CTNNB1/beta-catenin (PubMed:11403589). Plays a role in enhancing learning and memory performance (By similarity). [By similarity](#) [20 Publications](#)

Isoform 2 may act as an antagonist of EGF action.

(Microbial infection) Acts as a receptor for hepatitis C virus (HCV) in hepatocytes and facilitates its cell entry. Mediates HCV entry by promoting the formation of the CD81-CLDN1 receptor complexes that are essential for HCV entry and by enhancing membrane fusion of cells expressing HCV envelope glycoproteins. [1 Publication](#)

Catalytic activity¹

ATP + L-tyrosyl-[protein] = ADP + H⁺ + O-phospho-L-tyrosyl-[protein] [PROSITE-ProRule annotation](#) [6 Publications](#)

Protein Structure Databases

<http://www.rcsb.org/pdb/>

RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB

142015 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

Search by PDB ID, author, macromolecule, sequence, or ligands Go

Advanced Search | Browse by Annotations

FDB-101 PDB EMDataBank Research Data Bank Worldwide Protein Data Bank

f t y d

Welcome

Deposit Search Visualize Analyze Download

A Structural View of Biology

This resource is powered by the Protein Data Bank archive—information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data. The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

Award-Winning Videos on Antibiotic Resistance

2018

July Molecule of the Month

Contact Us

RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB

Display Files Download Files

1P36

T4 LYOSZYME CORE REPACKING MUTANT I100V/TA

DOI: 10.2210/pdb1P36/pdb

Classification: [HYDROLASE](#)

Organism(s): [Enterobacteria phage T4](#)

Expression System: [Escherichia coli](#)

Mutation(s): 3

Deposited: 2003-04-16 Released: 2003-10-07

Deposition Author(s): [Moers, B.H.](#), [Datta, D.](#), [Baase, W.A.](#), [Zollars, E.S.](#), [Mayo, S.L.](#), [Matthews, B.W.](#)

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 1.45 Å

R-Value Free: 0.226

R-Value Work: 0.187

wwPDB Validation

3D Report Full Report

Metric	Percentile Ranks	Value
Rfree		0.229
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		2.9%
RSRZ outliers		6.7%

6/22/2022 23

3D View: Structure | Electron Density | Ligand Interaction

Standalone Viewers Protein Workshop | Ligand Explorer

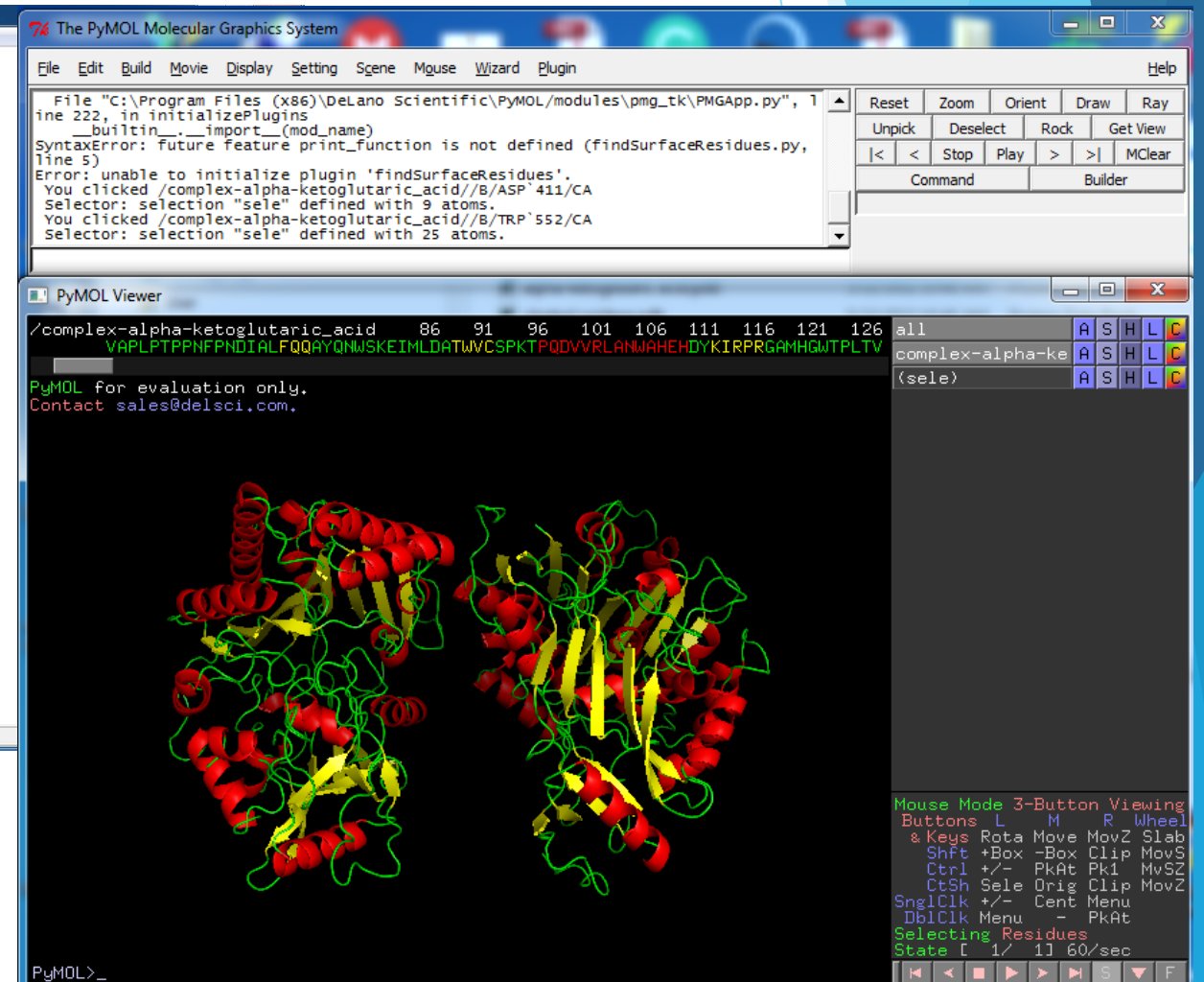
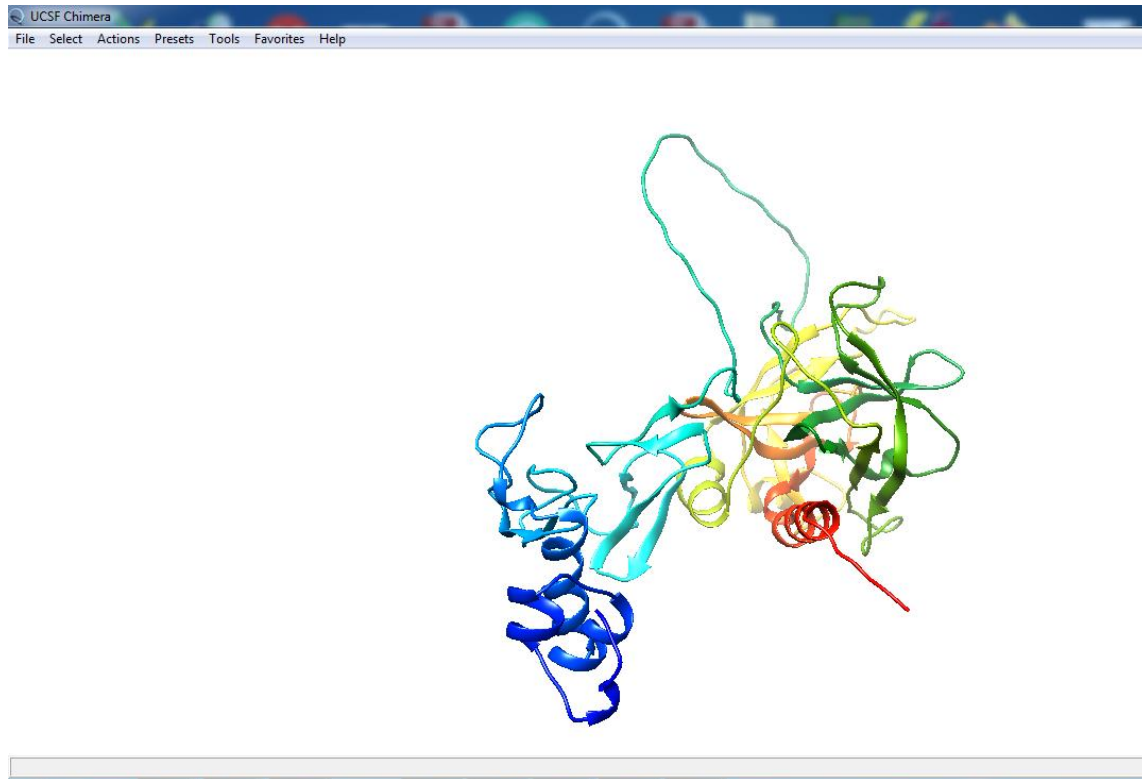
www.rcsb.org/pdb/search/smartSubquery.do?smartSearchSubtype=Structure&idisplay=true&struct_keywords.pdb_keywords.value=HYDROLASE&struct_keywords.pdb_keywords.comparator=contains

Contact Us

How to visualize tertiary protein structure

Pymol(<https://pymol.org/2/#opensource>) ,

Chimera(<https://www.cgl.ucsf.edu/chimera/download.html>)



Homology modeling

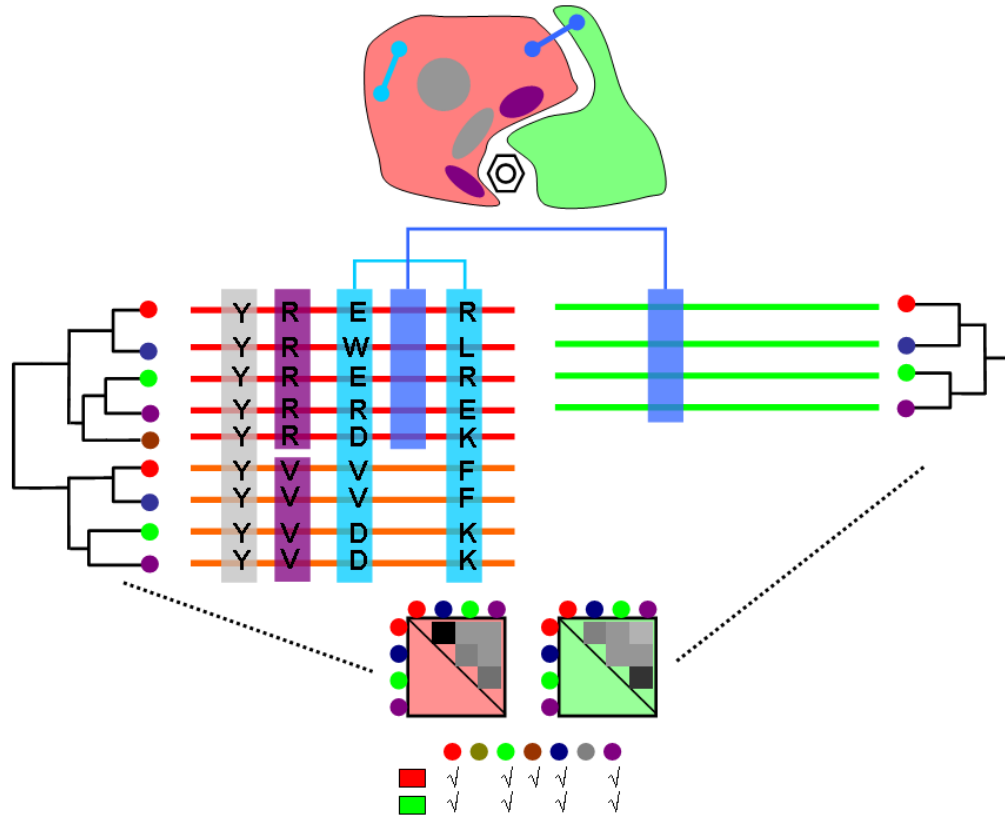
swiss model(<https://swissmodel.expasy.org/>)

I-tasser(<https://zhanggroup.org/I-TASSER/>)

The screenshot shows the 'Start a New Modelling Project' section of the SWISS-MODEL website. It includes a text area for 'Target Sequence(s)' with a green circular icon on the right. Below this is a green button 'Upload Target Sequence File...' and a grey button 'Validate'. There are also input fields for 'Project Title' (containing 'Untitled Project') and 'Email' (containing 'Optional'). At the bottom are two large blue buttons: 'Search For Templates' and 'Build Model'. A small disclaimer at the bottom states: 'By using the SWISS-MODEL server, you agree to comply with the following terms of use and to cite the corresponding articles.'

The screenshot shows the I-TASSER website. On the left is a vertical sidebar titled 'Online Services' with a list of tools: I-TASSER, I-TASSER-MTD, C-I-TASSER, CR-I-TASSER, QUARK, C-QUARK, LOMETS, MUSTER, CETHREADER, SEGMENT, DeepFold, DeepFoldRNA, FoldDesign, COFACTOR, COACH, MetaGO, TripletGO, IonCom, FG-MD, ModRefiner, and REMO. The main content area features the I-TASSER logo (an hourglass) and the text 'Protein Structure & Function Predictions'. Below this, it states: '(The server completed predictions for 691935 proteins submitted by 168453 users from 160 countries) (The template library was updated on 2022/06/14)'. A paragraph describes the I-TASSER method as a hierarchical approach for protein structure prediction. At the bottom, there is a section for 'I-TASSER On-line Server (View an example of I-TASSER output):' which includes a text box for pasting a sequence and a link to a sample input.

Mutational strategy



	C	S	T	P	A	G	N	D	E	Q	H	R	K	M	I	L	V	F	Y	W	
C	9																				C
S	-1	4																			S
T	-1	1	5																		T
P	-3	-1	-1	7																	P
A	0	1	0	-1	4																A
G	-3	0	-2	-2	0	6															G
N	-3	1	0	-2	-2	0	6														N
D	-3	0	-1	-1	-2	-1	1	6													D
E	-4	0	-1	-1	-1	-2	0	2	5												E
Q	-3	0	-1	-1	-1	-2	0	0	2	5											Q
H	-3	-1	-2	-2	-2	-2	1	-1	0	0	8										H
R	-3	-1	-1	-2	-1	-2	0	-2	0	1	0	5									R
K	-3	0	-1	-1	-1	-2	0	-1	1	1	-1	2	5								K
M	-1	-1	-1	-2	-1	-3	-2	-3	-2	0	-2	-1	-1	5							M
I	-1	-2	-1	-3	-1	-4	-3	-3	-3	-3	-3	-3	-3	1	4						I
L	-1	-2	-1	-3	-1	-4	-3	-4	-3	-2	-3	-2	-2	2	2	4					L
V	-1	-2	0	-2	0	-3	-3	-3	-2	-2	-3	-3	-2	1	3	1	4				V
F	-2	-2	-2	-4	-2	-3	-3	-3	-3	-3	-1	-3	-3	0	0	0	-1	6			F
Y	-2	-2	-2	-3	-2	-3	-2	-3	-2	-1	2	-2	-2	-1	-1	-1	-1	3	7		Y
W	-2	-3	-2	-4	-3	-2	-4	-4	-3	-2	-3	-3	-1	-3	-2	-3	1	2	11		W

(a)

Target sequence

A L E D N

Functionally Identical or Related Enzymes (property A proteins)

H L E P L

Sequence-related Proteins (~A proteins)

L L P D L

Score

-1 0 6 -7 0

(b)

Physicochemical assessment analysis

<https://web.expasy.org/protparam/>

Expasy ProtParam Home | Contact

ProtParam tool

ProtParam (References / Documentation) is a tool which allows the computation of various physical and chemical parameters for a given protein stored in [Swiss-Prot](#) or [TrEMBL](#) or for a user entered protein sequence. The computed parameters include the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) ([Disclaimer](#)).

Please note that you may only fill out **one** of the following fields at a time.

Enter a [Swiss-Prot/TrEMBL](#) accession number (AC) (for example **P05130**) or a sequence identifier (ID) (for example **KPC1_DROME**):

Or you can paste your own amino acid sequence (in one-letter code) in the box below:

RESET Compute parameters

<https://protein-sol.manchester.ac.uk/>

 **Protein-Sol** Warwick and Curtis Groups  

Sequence Patches Heatmap Abpred pka phosIDP Software About

Sequence Prediction

The protein-sol software will take a single amino acid sequence and return the result of a set of solubility prediction calculations, compared to a solubility database.

Please enter a single sequence of single letter amino acid codes in the FASTA format.

For example

```
> P00547
MKVYAPASSANMSYGFVDVLAQVTPYDGLLGDVVTVEAETFSLNNLGRFADRLPSEPRENIYVOCWERFCQELGKQIPVAMTEKNMPIGSLGSSACSVVAALMAMNEHCGKPLNDTRLALMGELEG
RISGSIHYDINVAFCFLGGMLMIEENDIIISQVTPGFDENLWVLAYPGIKVSTAEARAILPAQYRRQDCIAGHLAGFTIACYSRQPELAALKMKDVIAEPYRERLLPGFRQARQAVAEIGAVASGISGSGP
TLFALCDKPEAQRVADWLGNLYLQNGEGFVHCRLDTAGARVLEN
```

Submit protein sequence

<http://bioinf.uab.es/aggrescan/>

 **AGGRESKAN**
The Hot Spot Finder [Help File](#) [References](#) [contact](#)

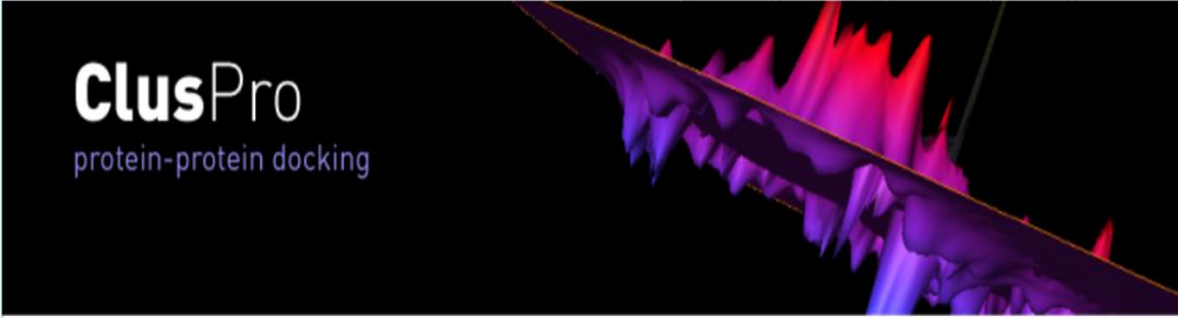
Prediction of "hot spots" of aggregation in polypeptides

Enter the peptide sequences in [FASTA format](#)

Molecular docking

Cluspro(<https://cluspro.bu.edu/login.php>)

[Login](#) [Sign Up](#) [Papers](#) [Help](#) [Contact](#)



ClusPro

protein-protein docking

Welcome to Cluspro 2.0

Recent news: [ClusPro server featured on the cover of February 2017 issue of Nature Protocols](#)

Use Without an Account

[Use the server without the benefits of your own account](#)

--or--

Login

Username:

Password:

Create Account

First Name:

Last Name:

Username:

Examples: JSmith, John.Smith

Affiliation:

Example: Boston University

Email:

Your email must be an educational or government address. Your password will be sent to this address (you can change this password after you've logged in).

Word Verification:

Type the characters seen in the picture



☐ I agree to use ClusPro only for noncommercial purposes.

Already Have an Account?
[Login](#)

Molecular docking Cluspro

Dock

Job Name:

Server:

Accepted PDB Input:
20 standard amino acids and RNA (as receptor only), ref: [RNA](#) Select Heparin Mode to use Heparin as Ligand.

Receptor

PDB ID:

[Upload PDB](#)

Chains:

Ligand

PDB ID:

[Upload PDB](#)

Chains:

Whitespace separate desired chains. Leave chains blank to use all chains.

Advanced Options



Dock

Advanced Options

Attraction and Repulsion

Enter attraction and repulsion of residues as whitespace separated "chain-residue" entries.
eg. a-23 a-25 a-26 a-27

Attraction:

Attraction:

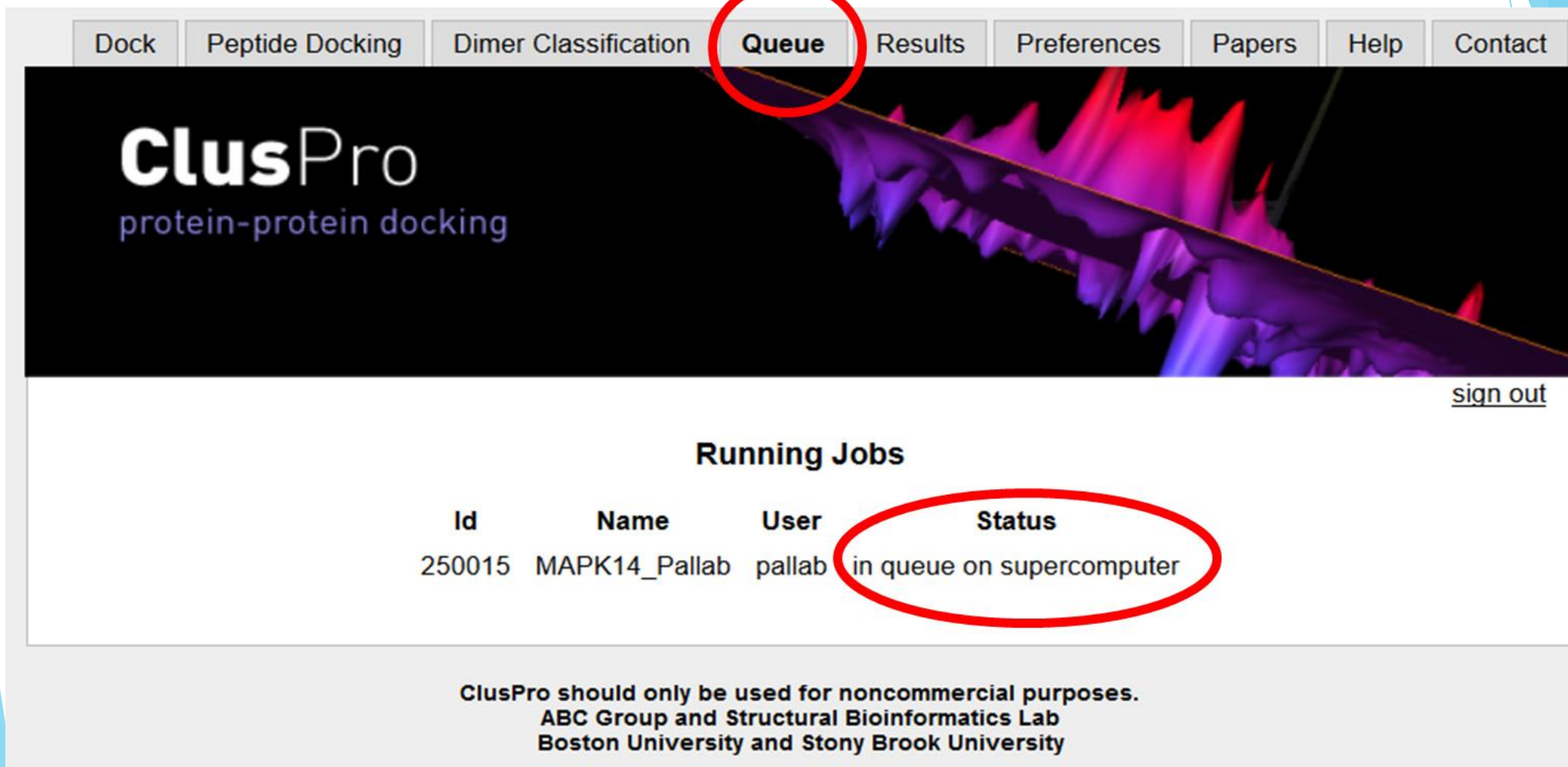
Repulsion:

Repulsion:

[Use PDB Masking File](#)

[Use PDB Masking File](#)

Molecular docking Cluspro



ClusPro
protein-protein docking

[Dock](#) [Peptide Docking](#) [Dimer Classification](#) **[Queue](#)** [Results](#) [Preferences](#) [Papers](#) [Help](#) [Contact](#)

[sign out](#)

Running Jobs

Id	Name	User	Status
250015	MAPK14_Pallab	pallab	in queue on supercomputer

ClusPro should only be used for noncommercial purposes.
ABC Group and Structural Bioinformatics Lab
Boston University and Stony Brook University

Molecular docking Cluspro

View Model Scores

[Download all Models for all Coefficients](#)

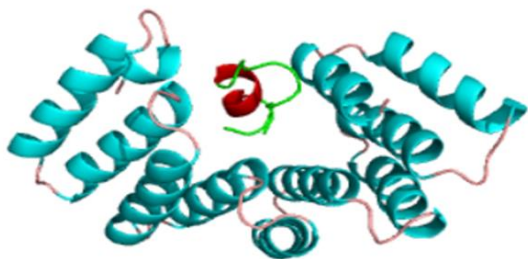
Balanced | [Electrostatic-favored](#) | [Hydrophobic-favored](#) | [VdW+Elec](#)

Display Models:

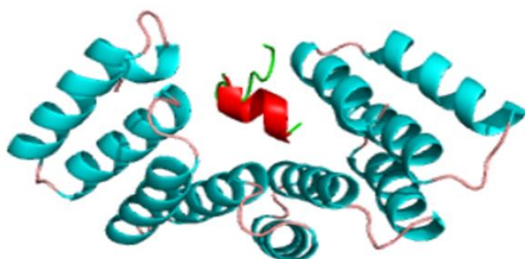
[Download Displayed Models](#)

If you use these models in a paper, please cite our [papers](#)

0



1



View Models

Balanced | [Electrostatic-favored](#) | [Hydrophobic-favored](#) | [VdW+Elec](#)

[Download Model Scores for this Coefficient](#)

Coefficient Weights

See *Kozakov et. al.* in [Papers](#) for a description of these terms

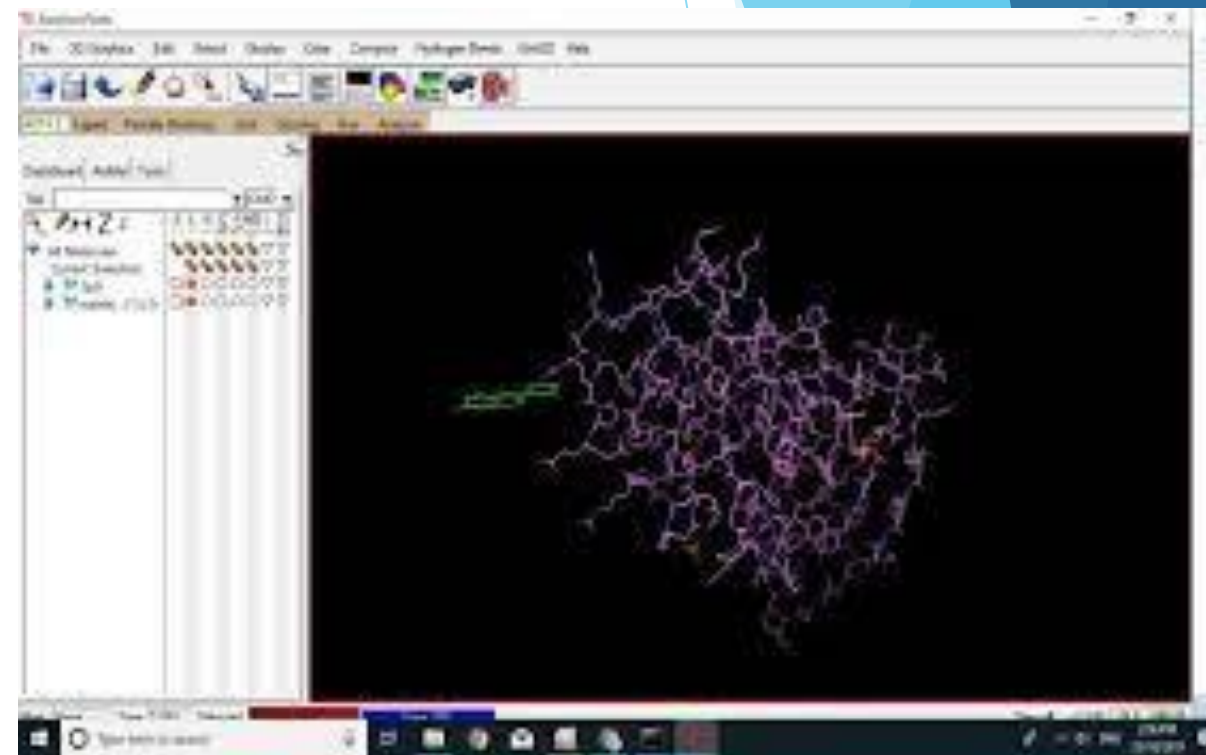
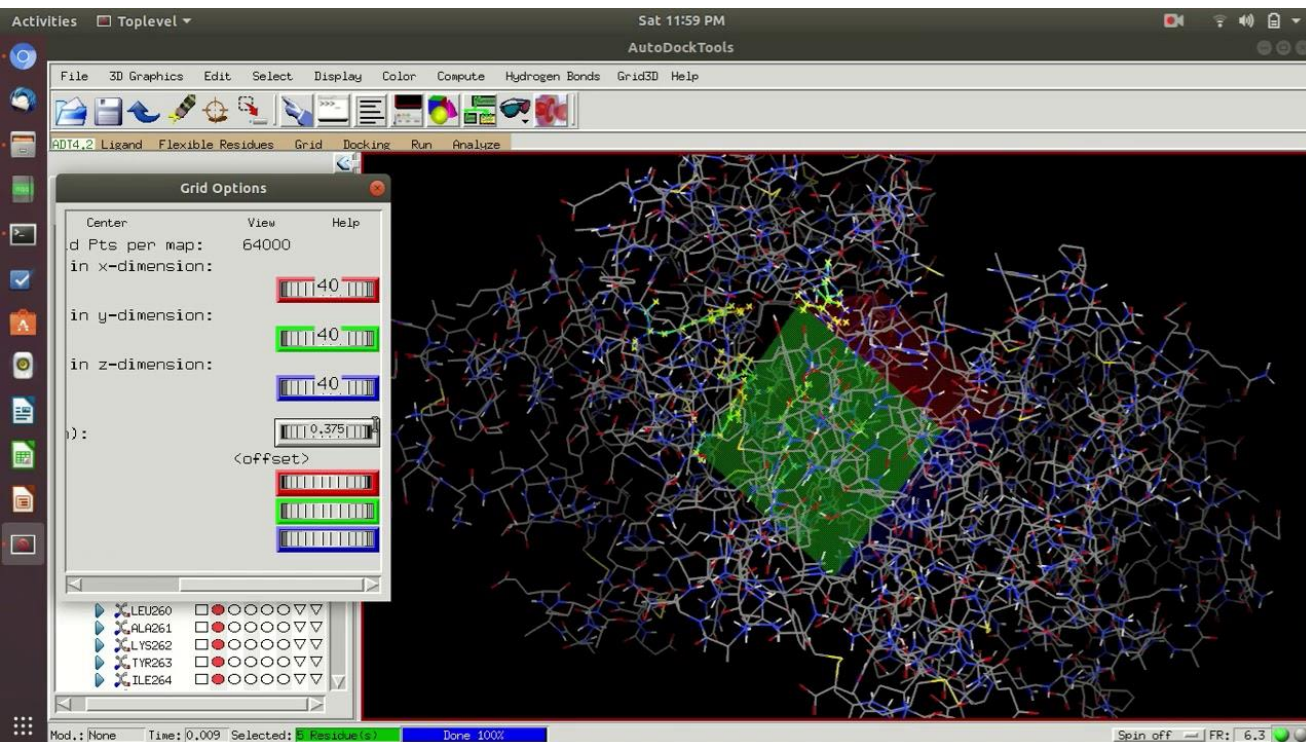
$$E = 0.40E_{rep} + -0.40E_{att} + 600E_{elec} + 1.00E_{DARS}$$

Cluster Scores

We strongly encourage you to read the [FAQ related to these scores](#) before using them.

Cluster	Members	Representative	Weighted Score
0	433	Center	-847.2
		Lowest Energy	-956.6
1	322	Center	-853.3
		Lowest Energy	-959.0
2	140	Center	-842.2
		Lowest Energy	-915.7
3	53	Center	-844.7

Molecular docking Autodock (<https://autodock.scripps.edu/>)



LigPLOT

<https://www.ebi.ac.uk/thorntonsrv/software/LIGPLOT>

EMBL-EBI

Services Research Training About us

LIGPLOT v.4.5.3 - Program for automatically plotting protein-ligand interactions

Groups > Thornton > Software > LIGPLOT

Written by Andrew Wallace and Roman Laskowski

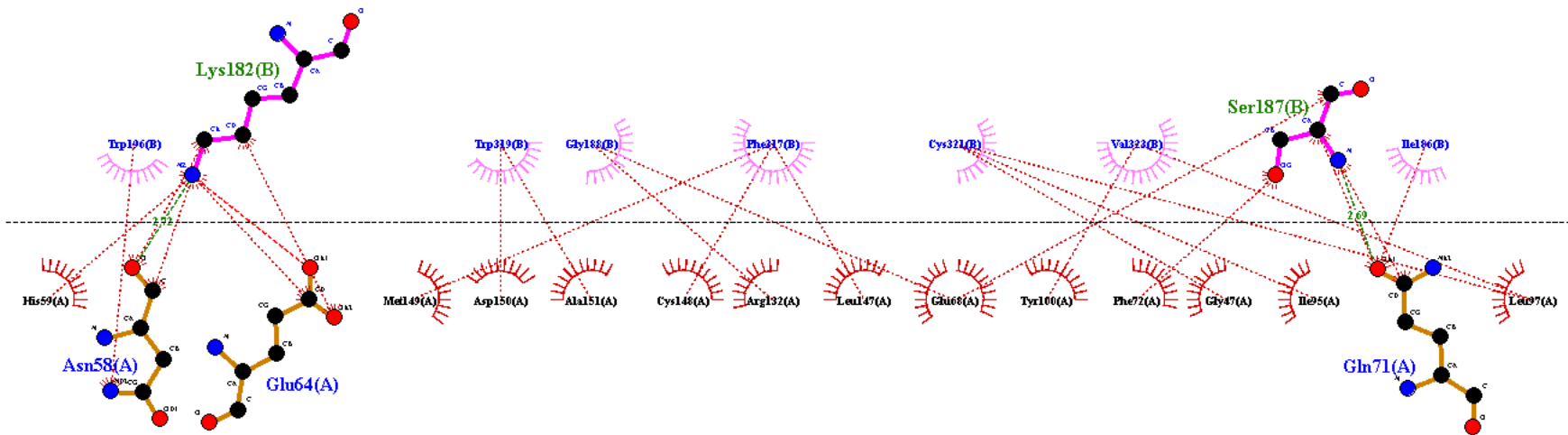
Automatically generates schematic diagrams of protein-ligand interactions for a given PDB file. (Click on the example on the left).

The interactions shown are those mediated by **hydrogen bonds** and by **hydrophobic contacts**. Hydrogen bonds are indicated by dashed lines between the atoms involved, while hydrophobic contacts are represented by an arc with spokes radiating towards the ligand atoms they contact. The contacted atoms are shown with spokes radiating back.

Availability

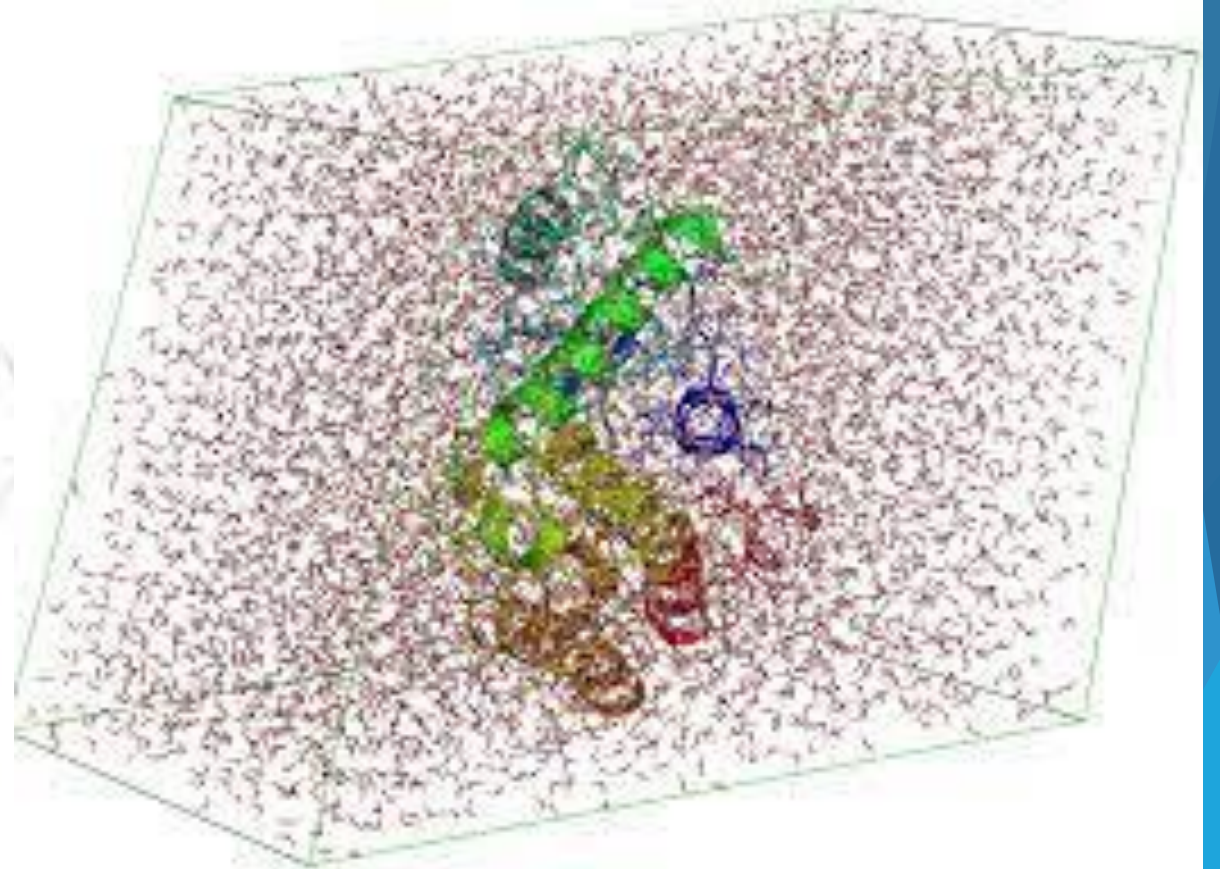
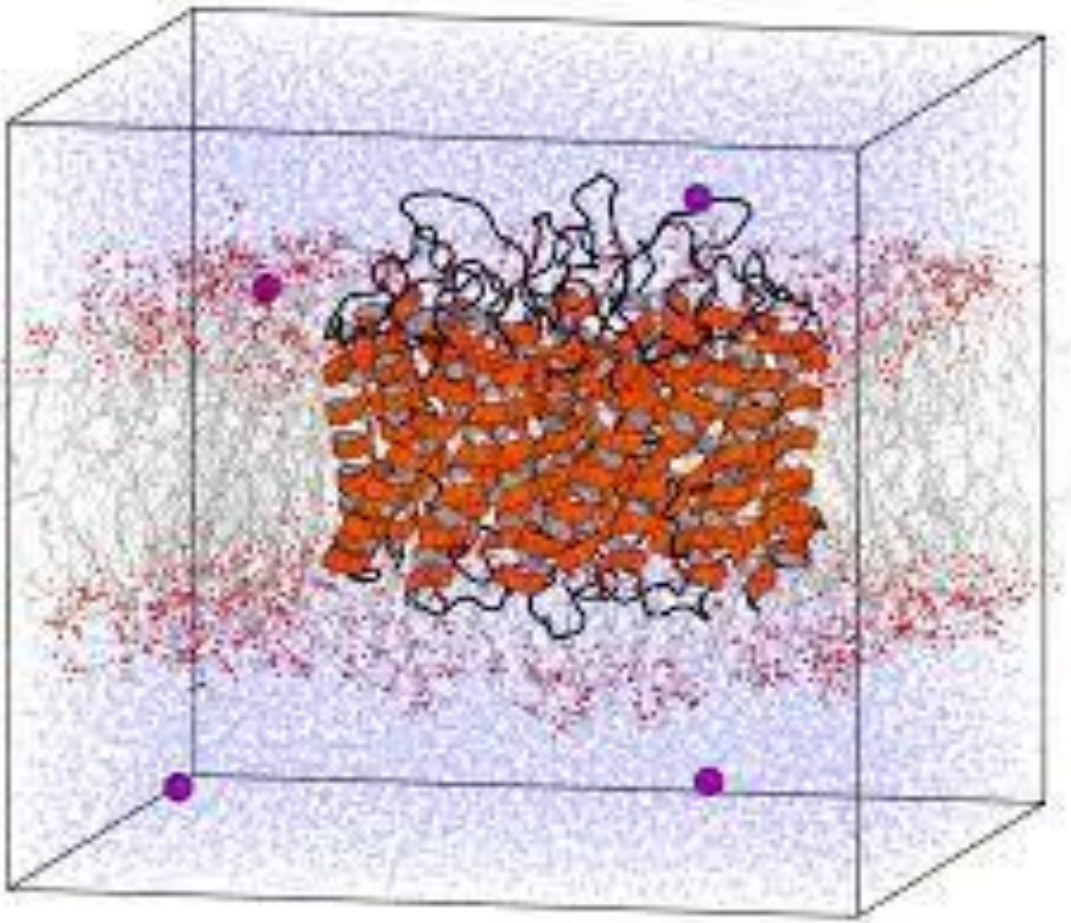
LIGPLOT has now been superseded by a new, GUI-based version called [LigPlot+](#), which has a number of novel features, including automatic superposition of related plots. Please see for more information.

Click to enlarge



Molecular dynamics simulation

Gromacs(<https://www.gromacs.org/>)



Workflow of computational protein engineering

- 1 • In silico analysis and design
- 2 • Gene library construct
- 3 • Protein display/expression
- 4 • High throughput screening/selection
- 5 • Validation and characterization
- 6 • Selection the most potential protein