

BCHS 6229
Protein Structure and Function

Lecture 7 (November, 2011)

**From Sequence to Function (II):
Sequences and Topology
Structural Biology Knowledgebase**

Protein diversity has accumulated over a looonngg time

3D structures more conserved in evolution than aa sequence!

Alignment of protein sequences reveals conserved closely and distantly related families

Alignment of protein structures reveals convergence of fold and function, or extreme divergence of sequence

mechanisms accounting for structural irregularities?

Mutation can create altered function of proteins; gene duplication; recombination

Circular permutation of genes (N and C-termini are close); inteins

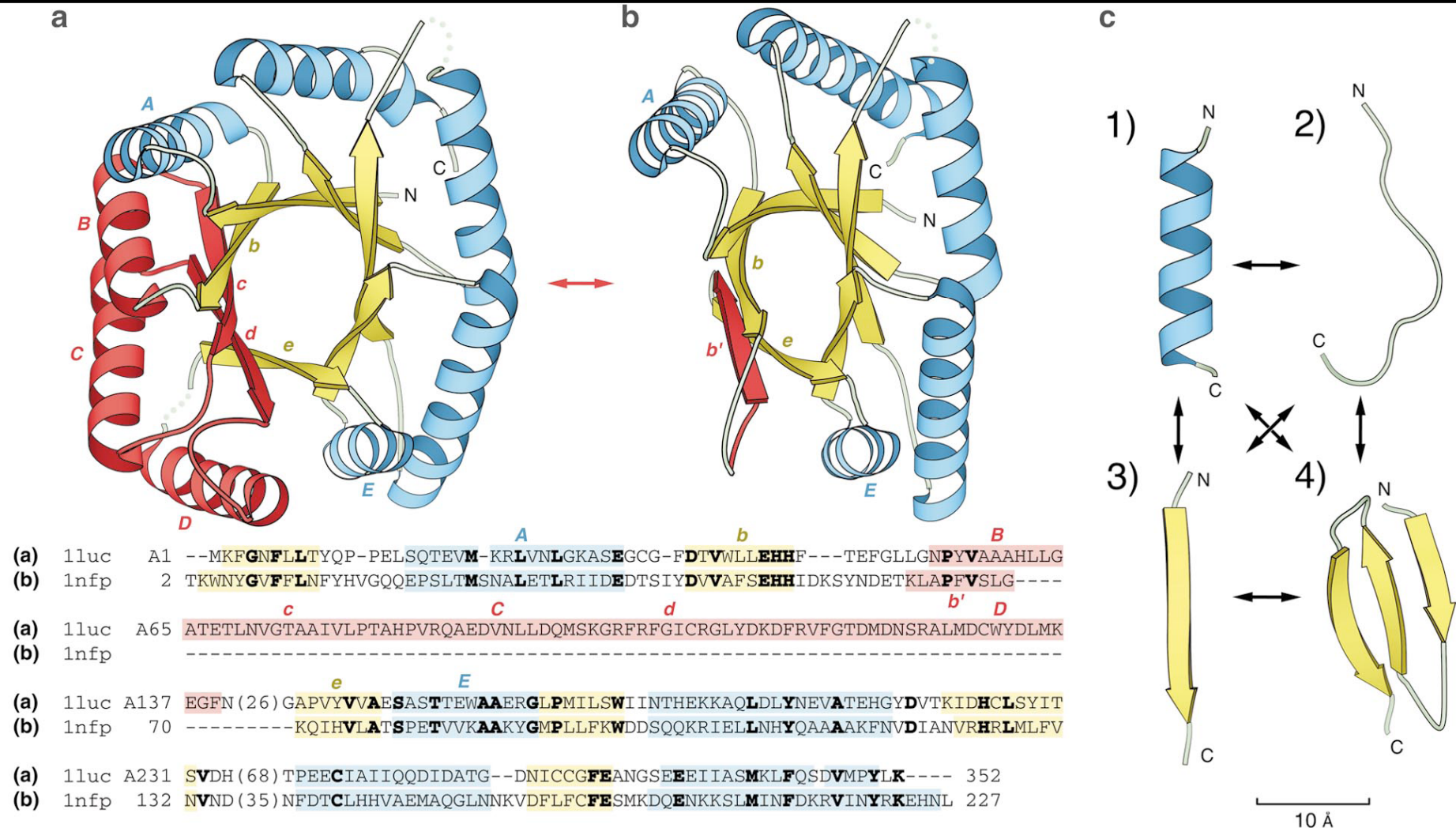
Deletions and insertions

Mutations that result in loss of function cannot accumulate unless the gene/protein is non-essential (or duplicated); may later provide a selective advantage

Fold Change in Evolution of Protein Structures

Nick V. Grishin

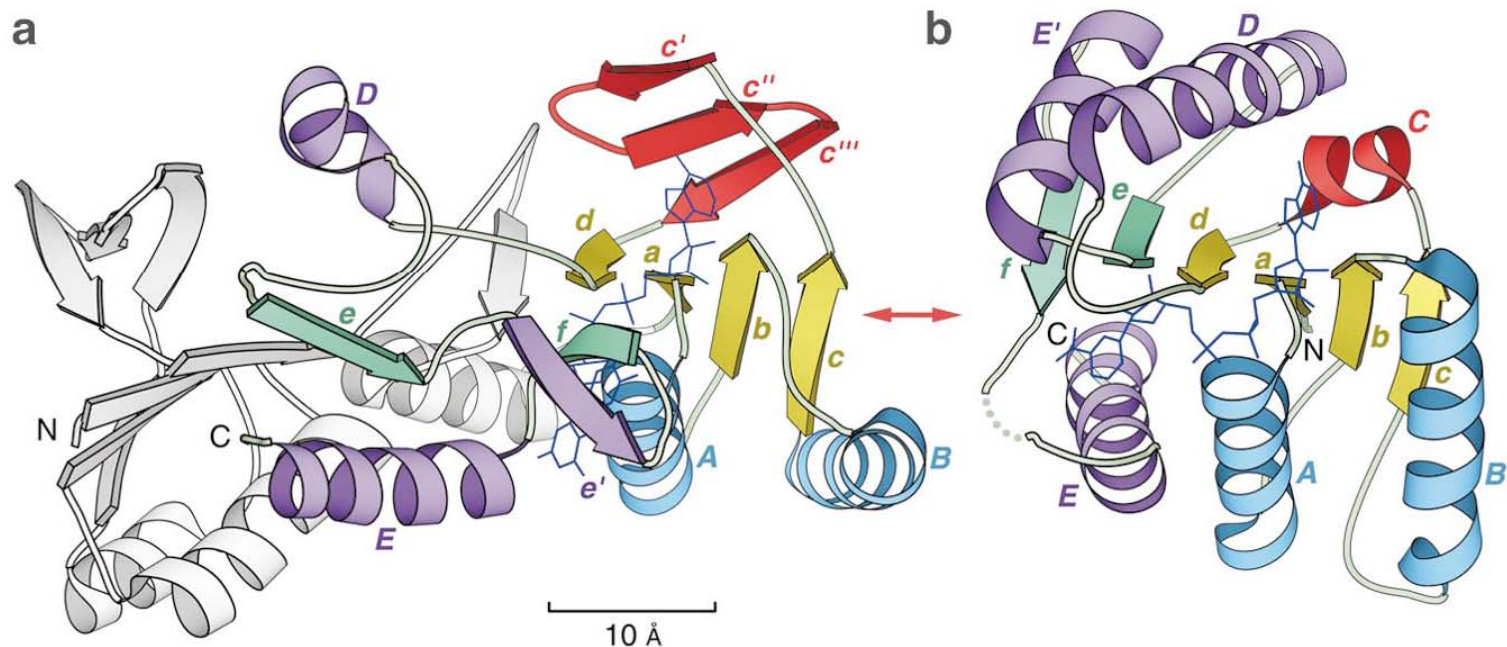
[Journal of Structural Biology 134, 167–185 (2001)]



1

1luc - Bacterial Luciferase

1nfp - Nonfluorescent Flavoprotein



			a	A	b	B	c
(a)	1npv	149	VNN VVVIGSG YIGIEAAEA FAKAG --KKVT VID ILDRPLGVYLDKEFTDVLTEEMEA-----NN IT				
(b)	1ldn	A20	GAR VVVIGAG GFVGASYV FALMNQG IADEIVL IDA -----NESKAIGDAMDFNHGKVFAPKPV DIW				
			c'	c''	c'''	d	D
(a)	1npv	208	IATGETVERYEGDGRVQKVVTDKNAY DADLVVAVG VRP-----NT AWL KG TLE LHPNGLIK				
(b)	1ldn	A80	HGDYD-----DCR DADLVVICAG ANQKPGETRLDLVDK NIA IFRSIV ES -----				
			C	e	e'	f	E
(a)	1npv	265	TDEY MRT SEPDV FAVGD --ATLIK YN -----PADTEVNIA----LATN ARKQGRFAVKNLEE 315				
(b)	1ldn	A124	--- V MASGFQGL FLV ATNPVDILT YATWKF SGL PHERVIGSG (85) Y YGI AMGLARVTRAILHN A265				
					E'		

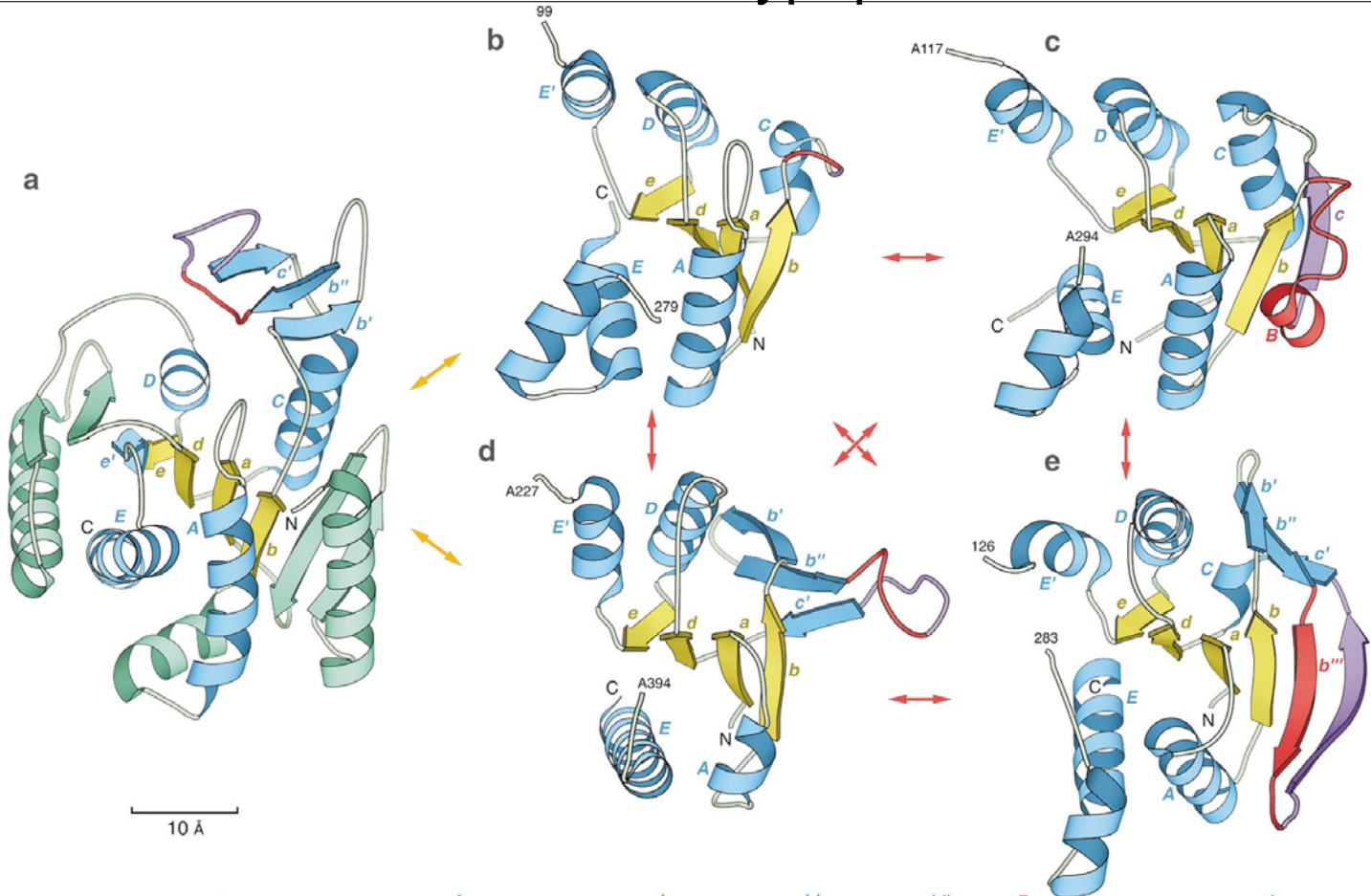
2

1npv - NADH Peroxidase

1ldn Lactate Dehydrogenase

Rossmann fold-like proteins

Rossmann fold-like domains of ATP-grasp proteins and zinc-carboxypeptidase.

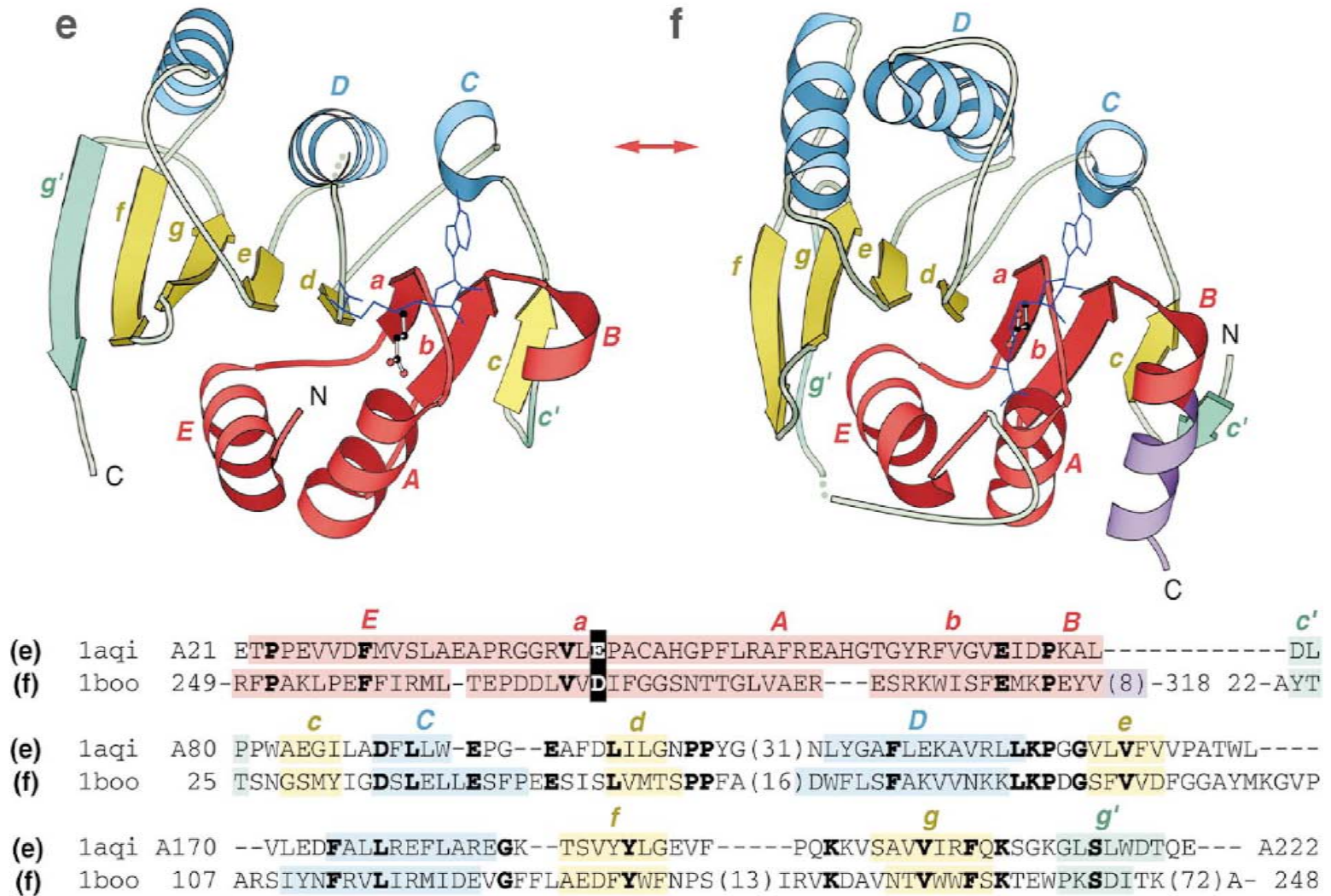


		a	A	b	b'	b''	B	c	c'
(a)	2ctc	A58	NRPAI WID LGIHSREW-ITQATGVVFAKKFTEDYQ(9)SMDIFLEIVT(14)WRKTRSVTSSSLCVGVDA ANRNWD AGE--GKAGASSSPCSETYHGKYANSEV						
(b)	2dln	A1	MTDKIAVLLGGT SA EREVSLNSGAAVLAGLREGG-----IDAYPVD PK -----						
(c)	1bnc	A1	MLDKIVIAN-----RGEIALRILRACKELG-----IKTVAVH SS -----ADRD LKHVL -----LADE TVCI GPAPSVKSYL						
(d)	1auv	A114	--RVLLVIDEPH-----TDWAKYFKGKKIHGEI--DIKVEQAEFS--DLNLVAHAN--GGFSVDMEVLRN-----GVKVVR-----						
(e)	1gsa	2	--IKLGIVMDPIANI-NIKKDSSFAMLEAQRRG-----YELHYMEMG--DLYLIN-----GEARAH TRTLNVKQNYE EWFSFVGEQDL-----						
		C	d	D	e	e'	E		
(a)	2ctc	A175	EVKSIVDFVKDHGNFKAPLS SI HSYSQ(47)QASGGSIDWSYNQGIKY SFT FELRDTGRYGFLL--PASQI PTA QETWL--GVL TIME HTLNN--A307						
(b)	2dln	A44	--EVD VTQLK SMGFQKV FIAL HGRG-----GED GT LQGMLELMGL FYT SGSGVMASALSMD KL (180)SHSLV PMA ARQAGMSFSQ LV RI LE LAD--A306						
(c)	1bnc	A62	NIPAIISAAEITGAVAIHPGYGFL-----SENANFAEQVERSG FIF GPKAETIRLMGDKV(176)-QVEHPVTEMITGV D --LIKEQLRIAAGQP--A320						
(d)	1auv	A177	-----SLK---PDFVLIRQHAFSMAR--NGDYRSLVIGLQYAGIP SI -NSLHSVYNFCDKP(167)-LIGDHQDE DKQL -----IVELV VNK MAQA--A417						
(e)	1gsa	76	-----PLAD--LDVILMRKDPPFD TE --FIYATYILERAEEK GT LIV-NKPQSLRDCNEKL(160)SPTCIREIEAEFPVS--ITGMLMDAIEAR LQ 314						

3

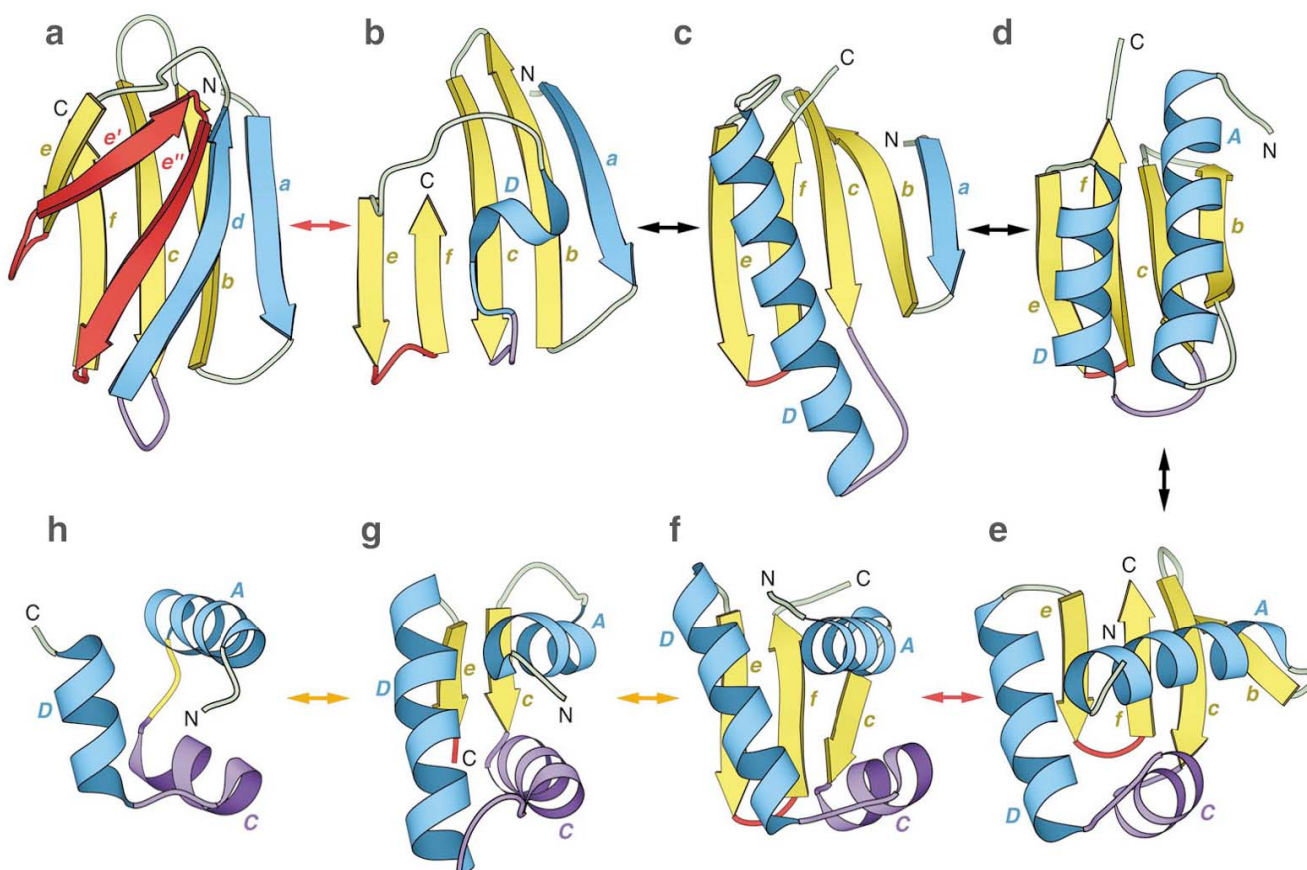
2ctc - Carboxypeptidase A 2dln - Dala-Dala Ligase 1bnc - Biotin Carboxylase 1auv - Synapsin 1gsa - Glutathione Synthase

Circular permutations



1aqi - Adenine-N6-Dna-Methyltransferase Taqi

1boo - PvuII Dna Methyltransferase



4

FIG. 4. A path from all- β to all- α proteins.

- (a) *Bacillus licheniformis* α -amylases, C terminal domain (1bpl);
- (b) *Pseudomonas stutzeri* G4-amylase C-terminal domain (2amg);
- (c) γ -subunit of glycogen phosphorylase kinase N-terminal domain (1phk);
- (d) sonic hedgehog N-terminal signaling domain (1vhh);
- (e) catabolite gene activator protein (CAP), C-terminal domain (1cgp);
- (f) biotin repressor N-terminal domain (1bia);
- (g) ribosomal protein L11 C-terminal domain (1fow);
- (h) HIN recombinase DNA-binding domain (1hcr).

Protein Design Principles

Need a stable hydrophobic core with constrained rotamer conformation for individual side chains (high stability can be achieved without a well-ordered core)

Core depends on three-dimensional arrangement of secondary structure elements and vice versa (breaking α helix, over/underpacking a core)

Natural deviations in core and surface can be associated with disease states resulted from misassociated to misfolded proteins

Amino acid diversity is ultimately required to reflect naturally occurring proteins

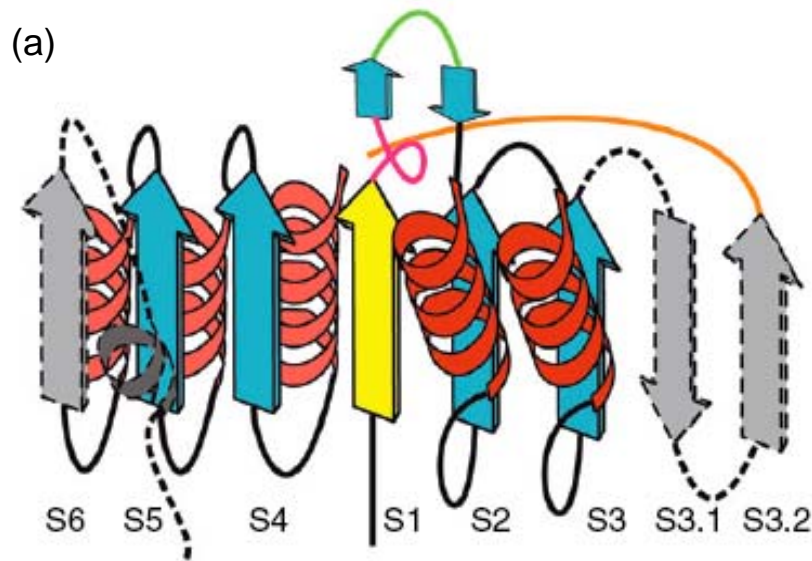
Solution experiments on designed proteins can test the computational methods

Topological variation in functionally diverse enzyme superfamily

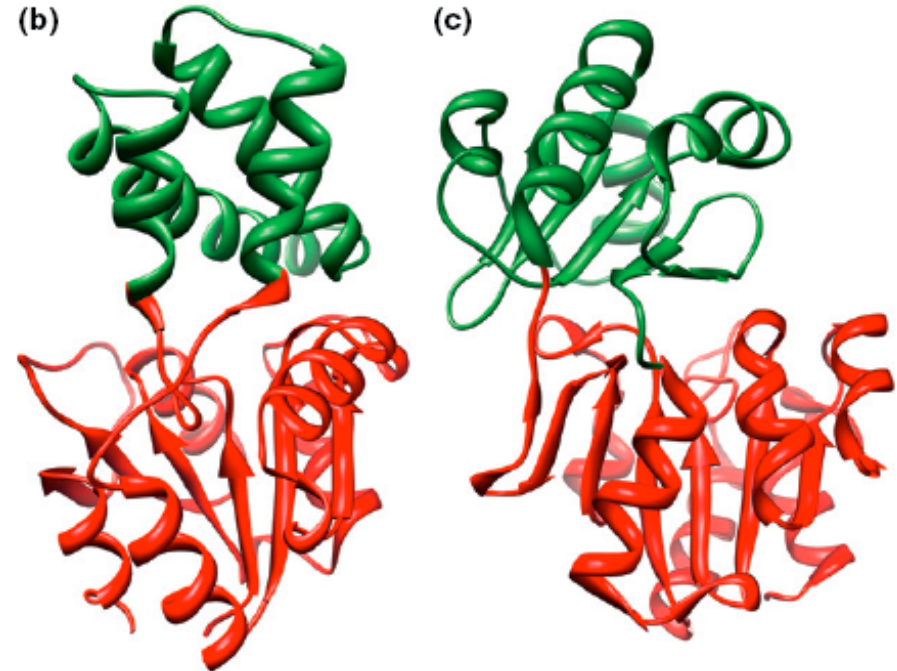
(Current Opinion in Structural Biology 2011, 21:391–397)

1. The haloalkanoic acid dehalogenase (HAD) SF: cap domain variations enable divergent evolution of many different reaction and substrate specificities.
2. The vicinal oxygen chelate (VOC) SF: mixing and matching subdomains for functional versatility.
3. The thioredoxin (Trx)-fold like SFs: varied inserts and domain additions extend the redox repertoire of the canonical Trx-fold

Topologies of HAD and cap domains



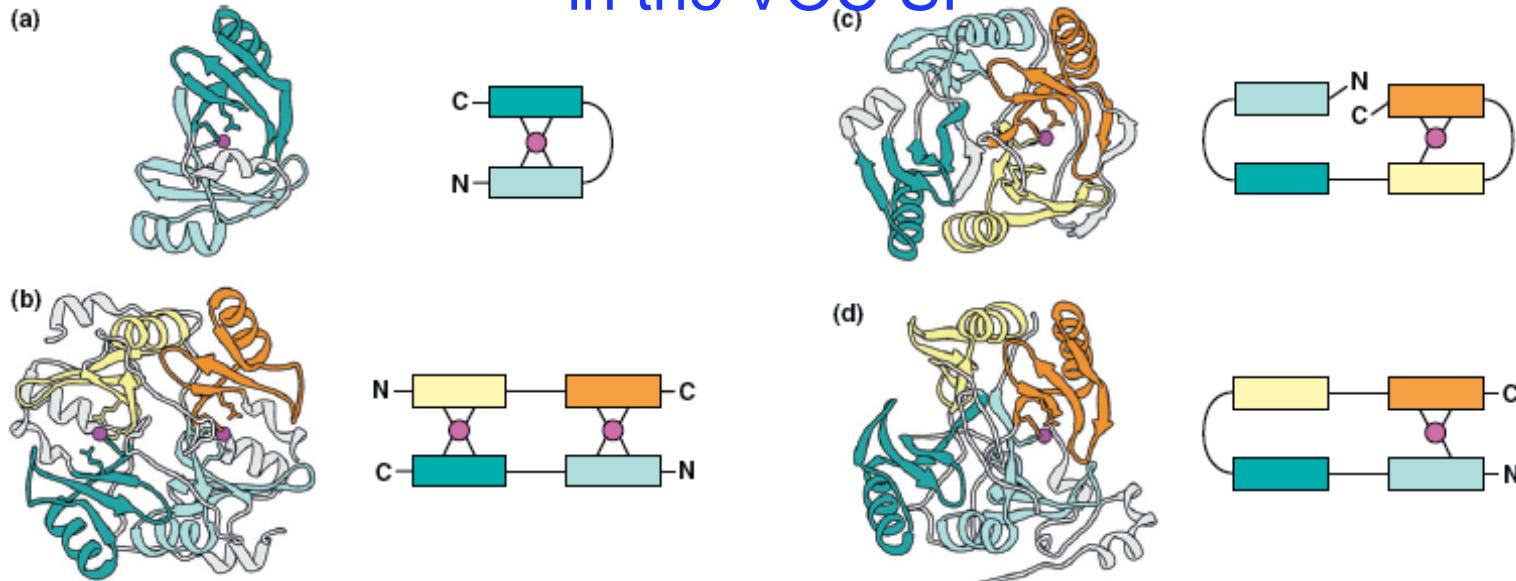
(a) Schematic diagram of the classic HAD domain: yellow, strand containing the catalytic Asp residue; blue, core strands conserved in all HAD SF members; gray, structural elements that may not have occurred in the ancestral structure; green line shows the insertion point for C1 caps; orange line shows the insertion point for C2 caps. Broken lines indicate secondary structure elements not present in all members containing the HAD domain.



(b) Crystal structure of L-2-haloacid dehalogenase with 2-chloro-N-butyrate (ligand not shown) (PDB 1ZRM) showing the HAD domain in red and the C1 cap domain in green.

(c) Crystal structure of the HAD subclass IIB sugar phosphatases (PDB 1YMQ) showing the HAD domain in red and the C2 cap domain in green.

Examples of alternate arrangements of paired $\beta\alpha\beta\beta\beta$ modules in the VOC SF



(a) Glyoxalase I from *Clostridium* (PDB 3HDP), in which two modules from a single chain pair to form a metal site.

(b) Human glyoxalase I (PDB 1QIN), in which two chains each containing two modules pair in head-to-tail fashion to form two metal sites.

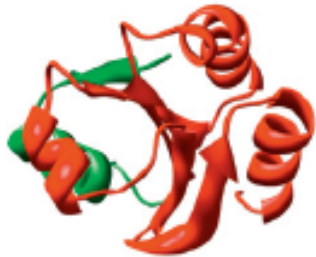
(c) 2,3 Dihydroxybiphenyl 1,2-dioxygenase from *Burkholderia* (PDB 1KMY): the four modules in a single chain pair in the order 1–2 and 3–4, and only the latter pair forms a metal site.

(d) Protein of unknown function from *Bacillus* (PDB 1ZSW): the four modules in a single chain pair in the order 1–4 and 2–3, and only the former pair forms a metal site.

metal ions are magenta and different modules are shown in different colors to highlight the repeating $\beta\alpha\beta\beta\beta$ unit. Metal-ligating residues may occur in the first and/or last beta-strands of a module. Parts of the structures not within $\beta\alpha\beta\beta\beta$ modules are shown in light gray.

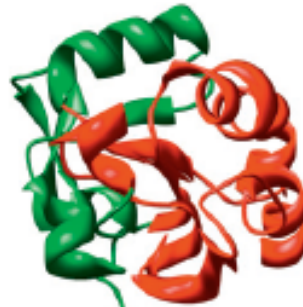
Examples of topological variations for Trx-fold SFs

(a)



Human Trx
(PDB 1AUC)

(b)



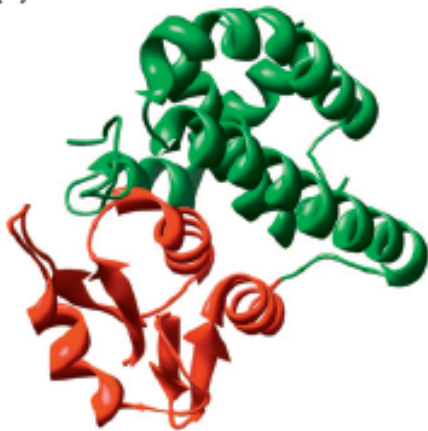
Human Prx V
(PDB 1HD2)

(c)



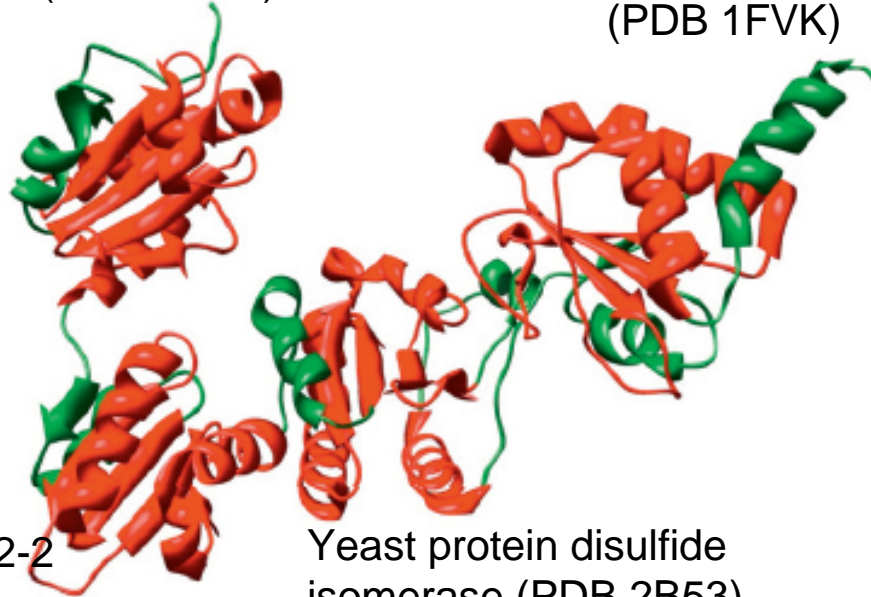
E. coli disulfide binding protein A
(PDB 1FVK)

(d)



Human glutathione transferase M2-2
(PDB 1XW5).

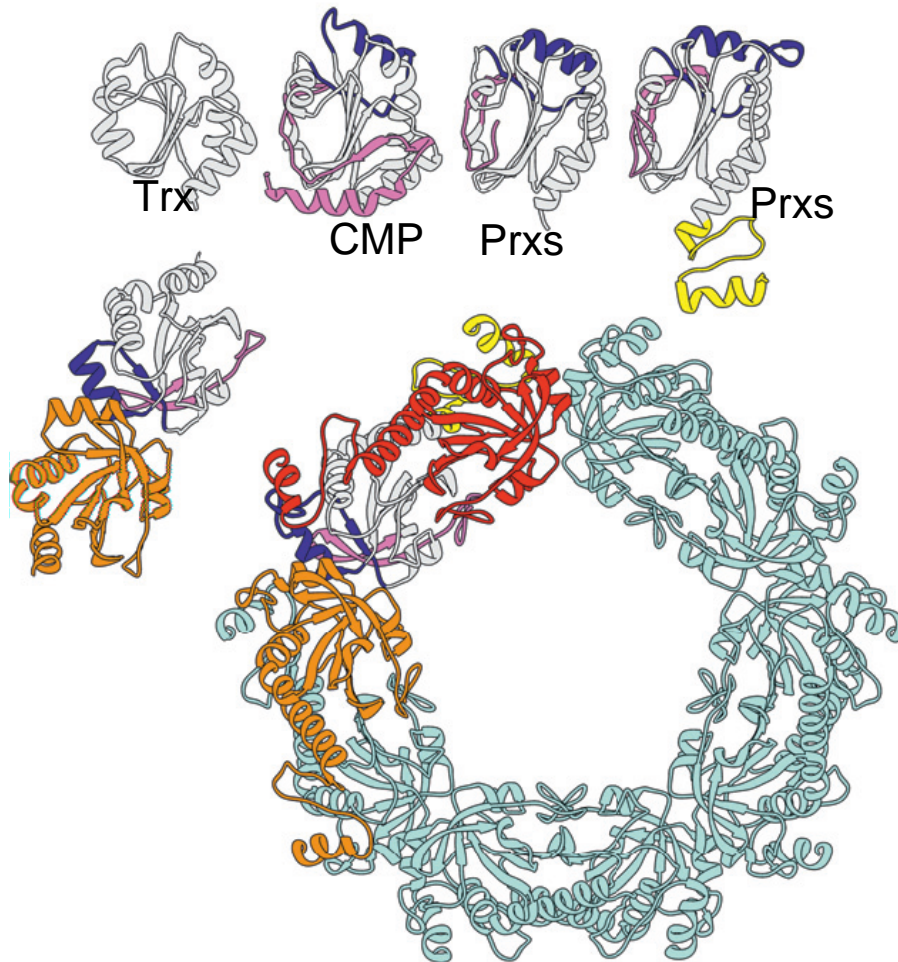
(e)



Yeast protein disulfide
isomerase (PDB 2B53).

Red: Trx-fold, green: inserts or domain pairing additions to the Trx-fold.

Structures of Trx, CMP, and Prxs are consistent with their connections in sequence space, in which CMPs link the other two SFs



The canonical Trx-fold is shown in light gray, an N-terminal extension in darkpink, an insert that includes a helix in blue, and a C-terminal extension in yellow. The bottom row shows the same two Prxs as in the top row, but in homo multimeric assemblies. Bottom left, Prx dimer (PDB 1XXU chains A and B) with one subunit colored as in the top row, the other subunit in orange. This dimer uses one type of interface (marked I) that primarily involves the insert helix (blue). Right, Prx decamer (PDB 1QMV) with one monomer colored as in the top row, a second monomer in orange interacting with the first through the type I interface, a third monomer in red interacting with the first through the type II interface, and the remaining seven monomers in aqua. The type II interface is primarily a concatenation of the central beta-sheet of the Trx-fold, but in this decamer, the C-terminal extension (yellow) is also involved.

PSI SBKB

PSI | nature
StructuralBiology Knowledgebase

The Structural Biology Knowledgebase by Protein Structure Initiative (PSI) and Nature Publishing Group (NPG)

Adopted from Materials prepared by Jennifer Williams, Ph.D.

(Updated: Q1 2011)

www.openhelix.com

PSI SBKB

1. Introduction and Credits
2. Structural Biology Update
3. Sequence or Structure Search
4. Text Searches
5. Additional Features
6. Summary
7. Exercises

PSI SBKB: <http://www.sbkb.org/>

1. PSI SBKB Homepage

“This ‘one-stop shop’ provides users with the available genetic, structural, functional and experimental information about a particular protein of interest.”

The screenshot shows the PSI SBKB Knowledgebase homepage. The header includes the PSI | nature Structural Biology Knowledgebase logo and navigation icons. A left sidebar contains a navigation menu with links: home, structural biology update, targets, protein structures, sequences, and function, homology models, methods, publications, about PSI, and NPG resources. The main content area features a welcome message, a search bar with options to search by sequence, text, or structure (PDB id), and a featured molecule section for Isoxanthopterin deaminase. Below this is a research advances section with articles on COX inhibition, ABA receptor diversity, and solving homodimeric structures with NMR. The right sidebar includes e-alerts, RSS feeds, a propose targets section, a functional sleuth section, and latest PSI statistics. Annotations with arrows point to the navigation menu (labeled 'navigation'), the search bar (labeled 'search'), and the structural biology update section (labeled 'SB update'). A 'sign up' button is also highlighted with an arrow.

PSI | nature
Structural Biology Knowledgebase

home
structural biology update
targets
protein structures, sequences, and function
homology models
methods
publications
about PSI
NPG resources

Welcome to the Structural Biology Knowledgebase

navigation

search

sign up

SB update

search

structural biology update November 2010
Research advances, news and events in structural biology

featured molecule

Isoxanthopterin deaminase
Researchers at NYSGXRC have discovered the function of a new enzyme from the Global Ocean Sampling Project.

research advances

COX inhibition: Naproxen by proxy
Structural basis of COX inhibition suggests that adverse effects of NSAIDs will not be easily avoided.

ABA receptor diversity
There are 14 members in the ABA receptor family in *Arabidopsis thaliana*. How these receptors differ in their response to a synthetic, seed-specific ABA agonist is now explored in two independent papers, using a combination of crystallography, NMR and biochemical approaches.

Solving homodimeric structures with NMR
Complexes are very difficult to solve in solution but are important biologically. Combining multiple approaches is the way forward.

Scaling up mutational scanning

e-alerts
Receive news of monthly updates by e-mail
sign up

RSS (monthly updates)
RSS (new molecules)

propose targets
Evaluate your sequences or nominate them to the PSI
evaluate targets

functional sleuth
Functional Sleuth presents PSI structures that lack full functional annotation
begin exploring

latest PSI statistics
New structures last month: 94
Total structures to date: 5097
Total distinct structures: 4371
more
see latest structures
get test datasets

PSI | nature
StructuralBiologyKnowledgebase

home

structural biology update

targets

protein structures, sequences, and function

homology models

methods

publications

about this site

about PSI

NPG resources

Welcome to the Structural Biology Knowledgebase

Keep informed about advances in structural genomics. Discover new dimensions of protein structure. Stay up to date with the latest technologies.

search

by sequence ☒ by text ☐ by structure (PDB id) ☐

example query

structural biology Research

featured molecule

Isoxanthopterin
Researchers are studying a new enzyme function.

research advances

COX inhibition
Structural basis of the effects of NSAIDs on COX.

ABA receptor
There are 14 members of the ABA receptor family in *Arabidopsis thaliana* that respond to a range of ligands. Two of these have been explored in two crystallographic studies.

Solving homodimers
Homodimers are very difficult to solve in solution but are important biologically. Combining multiple approaches is the way forward.

Scaling up mutational scanning

home

structural biology update

targets

protein structures, sequences, and function

homology models

methods

publications

about this site

about PSI

NPG resources

get test datasets

Next section

Access to pages organized by scientific focus

Next slides

home

structural biology update

targets

protein structures,
sequences, and function

homology models

methods

publications

about this site

getting started

contact us

site map

terms of use

how to cite us

about

NPG

about this site

The Structural Biology Knowledgebase resource produced in a collaboration between the Protein Structure Initiative (PSI) and Nature Publishing Group (NPG) provides an easy way of keeping abreast of developments generally in the fields of structural biology.

The SBKB serves as a continually updated resource from the PSI. NPG provides synopses of important research advances in a categorized library of research articles in structural biology. You can [register to receive our newsletter](#) and [subscribe to our RSS feeds](#).

The Protein Structure Initiative

The PSI is a federal, university, and industry effort aimed at dramatically reducing the costs and lessening the time it takes to determine a three-dimensional protein structure. The long-range goal of the PSI is to make the three-dimensional atomic-level structures of most proteins easily obtainable from knowledge of their corresponding DNA sequences. The PSI strives to gain biological insights from new structures and to help the broad biomedical research community make use of PSI research findings. For more information, please visit our [About PSI](#) page.

Nature Publishing Group

NPG is the scientific publishing arm of Macmillan Publishers Ltd, combining the excellence of Nature, Nature Research Journals, Nature Review Journals and NPG Academic Journals. In recent years, NPG's presence in the scientific and medical communities has been further enhanced by the launch of many new online resources that provide users with easy access to research results, news, events, and job lists, together with features that facilitate communications and social interactions between scientists.

The Editors at Nature Publishing Group appreciate the central importance of structural biology research to molecular and cell biology as well as therapeutic development. The collaboration with PSI to publish the SBKB builds on a strong publication record in structural biology and further cements NPG's links with the research community.

"... is a federal, university, and industry effort aimed at dramatically reducing the costs and lessening the time it takes to determine a three-dimensional protein structure."

<http://www.nigms.nih.gov/Initiatives/PSI/>

Click
"getting
started"

"NPG is the scientific publishing arm of Macmillan Publishers Ltd, ... The Editors at Nature Publishing Group appreciate the central importance of structural biology research to molecular and cell biology as well as therapeutic development."

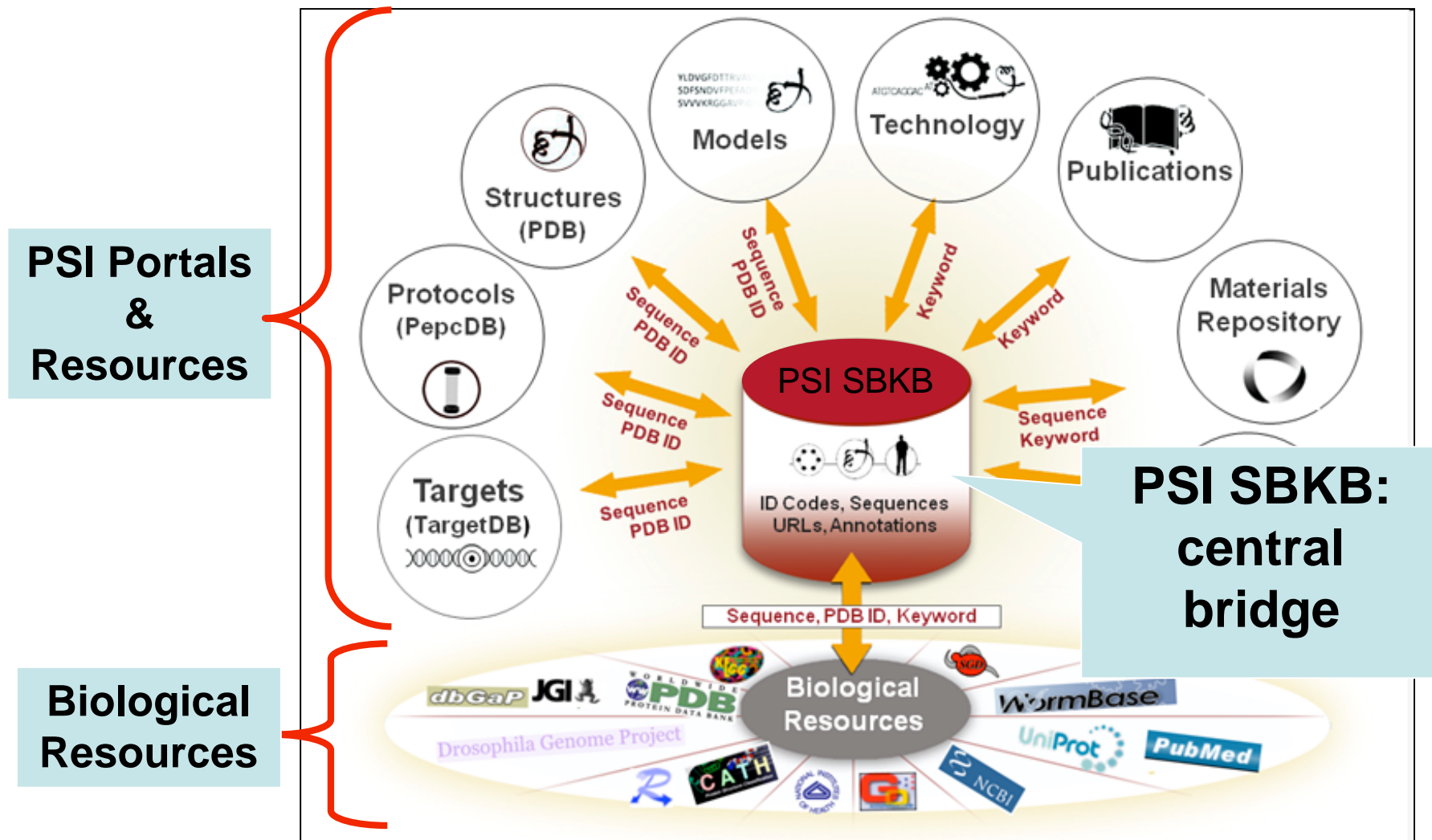


Image - http://www.sbkb.org/about/getting_started.html

Resource list - <http://www.sbkb.org/KB/seqstrucfunc hub.html>

2. Structural Biology Update

PSI | nature
StructuralBiologyKnowledgebase

home
structural biology update
targets
protein structures, sequences, and function
homology models
methods
publications
about this site
getting started
contact us
site map
terms of use
how to cite us
about PSI
NPG resources

site map

Home

Instructions

Structural B

- research
- research
- featured
- research
- news
- events c
- events a

Targets

- PSI Targ
- TargetDB
- Peptide
- Commun

Protein St

- Description of Query Functions and Annotation Module

Published online 14 November 2008

Nucleic Acids Research, 2009, Vol. 37, Database issue D365–D368
doi:10.1093/nar/gkn790

The protein structure initiative structural genomics knowledgebase

Helen M. Berman^{1,*}, John D. Westbrook¹, Margaret J. Gabanyi¹, Wendy Tao¹, Raship Shah¹, Andrei Kouranov¹, Torsten Schwede², Konstantin Arnold², Florian Kiefer², Lorenza Bordoli², Jürgen Kopp^{2,3}, Michael Podvinec², Paul D. Adams⁴, Lester G. Carter⁴, Wladek Minor⁵, Rajesh Nair⁶ and Joshua La Baer⁷

¹Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, ²Swiss Institute of Bioinformatics & Biozentrum, University of Basel, CH-4056 Basel, Switzerland, ³Biochemie-Zentrum, Heidelberg University, D-69120 Heidelberg, Germany, ⁴Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, ⁵Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, VA 22908-0736, ⁶Department of Biochemistry & Molecular Biophysics, Columbia University, New York, NY 10027 and ⁷Harvard Institute of Proteomics & Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115

Received September 16, 2008; Accepted October 8, 2008

ABSTRACT

The Protein Structure Initiative Structural Genomics Knowledgebase (PSI SGBK, <http://kb.psi-structuralgenomics.org>) has been created to turn the prod-

scale. Since 2001, these efforts have resulted in more than 6772 structure depositions to the PDB (1), 3251 of which are from the National Institutes of Health-sponsored Protein Structure Initiative Centers. In order to determine these structures is a high-throughput

http://nar.oxfordjournals.org/cgi/reprint/37/suppl_1/D365

More info

events

events

Featured Molecules

Isoxanthopterin Deaminase

Introduction

SBKB [doi:10.3942/psi_sbkb]

How do you discover the function of a new protein?
[Structural genomics](#)

Structure search

they have developed a toolbox full of methods to help solve it. The recent discovery the new enzyme isoxanthopterin deaminase shows some of the tools that are currently available for this search.

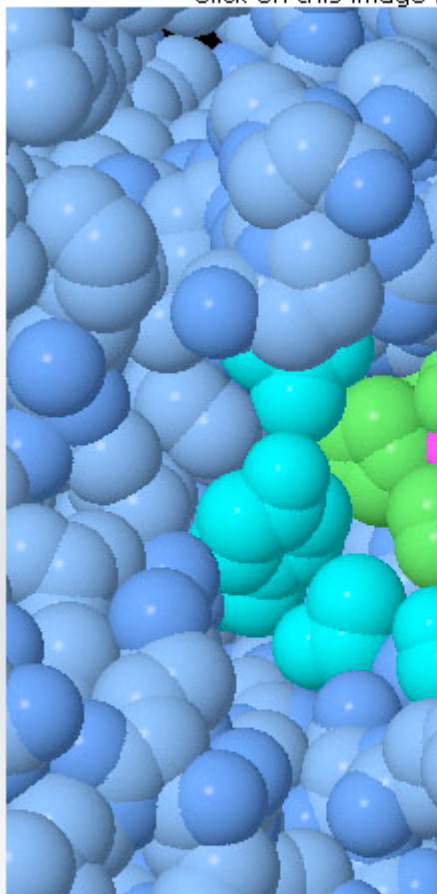
From Gene...

The process of discovery began with genomics, by sifting through the sequences of ocean microorganisms at [NYSGXRC](#) have been identified enzymes, which includes several adenosine deaminase, as well as perform similar, and occasionally through the database of DN Project, which picks up organisms determines the sequence of likely members of the superfamily.

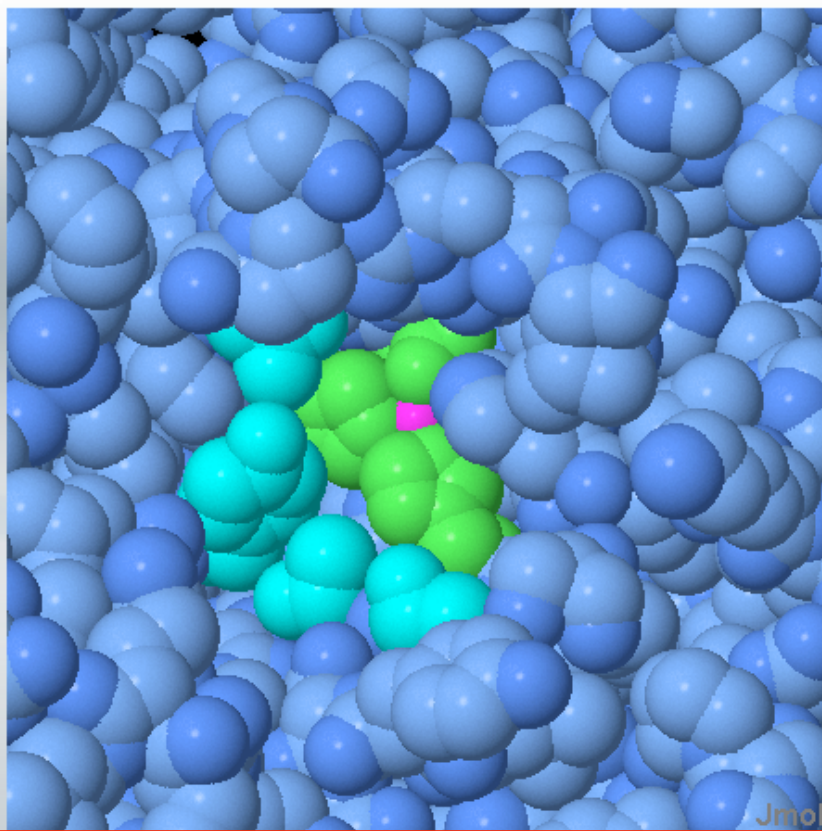
...To Protein...

The first step was to make a

Click on this image for



Isoxanthopterin Deaminase (PDB entry 2paj)



- ☒ hide catalytic water ☐ show catalytic water
- ☒ spacefilling representation ☐ backbone representation
- ☒ close-up on active site ☐ show entire protein

The active site of isoxanthopterin deaminase includes a zinc ion (magenta) coordinated by three histidines and an aspartate (green). Based on comparisons to similar enzymes, several amino acids (shown in turquoise) are predicted to be important for recognition of the substrate. A water molecule is used in the deamination reaction--use the buttons below to turn it on and off.

[Go Back](#)

Interactive Jmol View

- [Featured PSI Structures by David S. Goodsell](#)

3. PSI SBKB Search Form

The image shows a screenshot of the PSI Structural Biology Knowledgebase website. The header features the logo "PSI | nature StructuralBiology Knowledgebase" and three circular icons representing molecular structures, a protein, and a person. A left sidebar contains navigation links: home, structural biology update, targets, protein structures, sequences, and function, homology models, methods, publications, about this site, about PSI, and NPG resources. The main content area has a welcome message and a search section. The search section includes radio buttons for "by sequence" (selected), "by text", and "by structure (PDB id)". A text input field contains the sequence "ENILLNEDMHIQITDFGTAKVLSPEKQARANSFVGTAQYV". Below the input field is a link labeled "example query" circled in red. To the right of the input field are "help" and "search" buttons. A red box highlights the search input area. Annotations include a blue arrow labeled "search" pointing to the search section, a blue arrow labeled "help" pointing to the help button, and a blue arrow labeled "click" pointing to the search button. The right sidebar contains sections for e-alerts, RSS feeds, propose targets, and functional sleuth.

PSI | nature
StructuralBiology Knowledgebase

home
structural biology update
targets
protein structures, sequences, and function
homology models
methods
publications
about this site
about PSI
NPG resources

Welcome to the
Structural Biology Knowledgebase

Keep informed about advances in structural biology and structural genomics. Discover how protein sequences, three-dimensional structures and models relate to biological function. Stay up to date with the latest protocols, materials and technologies.

[SBKB tutorial](#) [What's new?](#)

search

Explore proteins and this website

by sequence ☒
by text ☐
by structure (PDB id) ☐

example query

ENILLNEDMHIQITDFGTAKVLSPEKQARANSFVGTAQYV

help

search

evaluate targets

propose targets

Evaluate your sequences or nominate them to the PSI

functional sleuth

Functional Sleuth presents PSI

Sequence Search Results

The screenshot shows the PSI | nature Structural Biology Knowledgebase interface. The header includes the logo and navigation icons. A left sidebar contains links like 'home', 'structural biology update', 'targets', 'protein structure', 'homology', 'methods', 'publications', 'about this site', 'about PSI', and 'NPG resources'. The main content area is titled 'Search Results' and features a tabbed interface with 'Summary', 'Structures', 'Targets', 'Protocols', and 'Materials'. The 'Summary' tab is selected and highlighted with a red border. A blue arrow labeled 'click' points to the 'Structures' tab. Below the tabs, a message says 'Select a tab from the top or follow the links below for detailed results'. The search details show 'Search type: Sequence Search' and 'Your query: NGELLKYIRKIGSFDETCTRFYTAIEIVSALEYLHGKGIIHRDLKPENILL NEDMHIQITDFGTAKVLSPEKQARANSFVGTAQYVSPEL'. The 'Results:' section is highlighted with a red border and contains a list of similar items with counts: 'Similar protein structures from the Protein Data Bank: 371', 'Similar theoretical models from the Protein Model Portal: 33', 'Similar protein targets in TargetDB: 500', 'Similar experimental protocols in PepcDB: 475', and 'Similar materials available from PSI Materials Repository: 4'. A blue arrow labeled 'click' points to the 'Results:' section. Three blue arrows with labels point to specific elements: 'Search type' points to the 'Search type:' label, 'Search input' points to the 'Your query:' text, and 'Result summary' points to the 'Results:' section.

PSI | nature
Structural Biology Knowledgebase

home
structural biology update
targets
protein structure
homology
methods
publications
about this site
about PSI
NPG resources

Search Results

Summary Structures Targets Protocols Materials

Select a tab from the top or follow the links below for detailed results

Search type: Sequence Search

Your query: NGELLKYIRKIGSFDETCTRFYTAIEIVSALEYLHGKGIIHRDLKPENILL
NEDMHIQITDFGTAKVLSPEKQARANSFVGTAQYVSPEL

Results:

- Similar protein structures from the Protein Data Bank: **371**
- Similar theoretical models from the Protein Model Portal: **33**
- Similar protein targets in TargetDB: **500**
- Similar experimental protocols in PepcDB: **475**
- Similar materials available from PSI Materials Repository: **4**

Structures Report

Annotation Quick Reference - 2pe2 (A)						
NMPDR	PROSITE	DisProt	PDB	PROFUNC	BioCyc	OMIM
Ensembl	Pfam	PRINTS	SMR	ProKnow	HPA	Orphanet
UniGene	InterPro	Phos-PTMs	PDBSUM	IntEnz	Reactome	PharmGKB
Nucleotide	SMART	Gene3D	SCOP	EC	KEGG	DrugBank
GeneID	ProDom	UniProt	CATH	GO		
DNA Sequence		Protein Sequence		Protein Structure	Functions	Pathways
					Medicine	

Go to available annotations (green) by clicking on their name.

Search Results

Summary

Structures

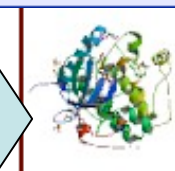
Mod

Gene Sequence
Protein Sequence
Protein Structure
Functions
Localization
Pathways
Medicine
References

2pe2 Chain ids | A

navigation

click



Launch Viewer »

(Similarity: I = 100%E = 9.5E-47) [View matching sequence alignment](#)

[PDB](#) [Download](#)

Title

CRYSTAL STRUCTURE OF H-DEPENDENT PROTEIN KINase-1, 2-dihydro-indol-3-yl)-(2-piperidin-1-yl)

Authors

Whitley, M., Adler, M.

Experimental method

X-RAY DIFFRACTION

Release date

2007-06-19

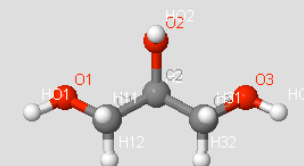
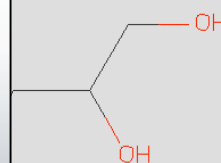
RCSB PDB PROTEIN DATA BANK

[RCSB PDB](#) | [Contact Us](#)

Home Search Browse Download Ligand Expo Help

Chemical Details Geometry Atom Nomenclature Downloads Related Resources

PDB Chemical Component GOL



Ideal Model

Jmol

Description

GLYCEROL

C3H8O3

Report

Sequence alignment for Id: 2pe2

id = 2pe2_A I = 100% E = 9.5E-47

Query: 1 NGELLKYIRKIGSFDETCTRFYTAIEIVSALEYLHGKGIIHRDLKPENILLNEDMHIQITD
NGELLKYIRKIGSFDETCTRFYTAIEIVSALEYLHGKGIIHRDLKPENILLNEDMHIQITD
Sbjct: 91 NGELLKYIRKIGSFDETCTRFYTAIEIVSALEYLHGKGIIHRDLKPENILLNEDMHIQITD

Query: FGTAKVLSPEKQARANSFVGTAQYVSP 90
FGTAKVLSPEKQARANSFVGTAQYVSP
Sbjct: FGTAKVLSPEKQARANSFVGTAQYVSP 180

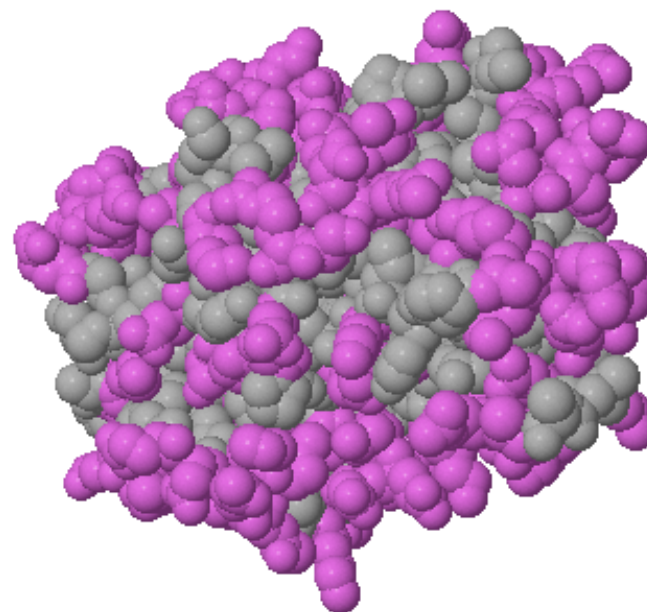
Structure Viewer

Image of
Structure

Image
Controls

Secondary Structure [Cartoon](#)
[N→C Rainbow](#)
[Composition](#) [Hydrophobic/Polar](#)
[Charge](#)
[Contacts..](#) [Vines..](#) [All Models](#)
[Hide..](#) [Find..](#)
☐ Ligands+..
☐ Water..
☐ Slab..
☐ Background
☒ Spin
Zoom
[More Views..](#) 2PE2: **PQS**
PDB **OCA** **Gaps?**
[Center Atom..](#) [Troubleshooting](#)
[Reset](#) [Close](#)

2PE2



Jmol

Image
Notes

Hydrophobic/Polar: Amino acids are colored either

- **Hydrophobic** or
- **Polar**
(charged or uncharged)

Nucleic acid is shown as [Cartoons](#) with each chain given a different color. Ligand and solvent are initially hidden.

Clicking the **Ligands+** or **Water**

[Can't see the molecule?](#)

[FirstGlance in Jmol](#)

After resizing this window, use the browser's Reload/Refresh to force Jmol to resize accordingly. Please wait until you see the molecule before reloading/refreshing. ([Why?](#))

Many visualization options for exploring your protein!

Structure Annotation Reports, Gene-Level View

The screenshot displays a web interface for a Gene Sequence report. At the top, there are two buttons: "Back to Summary" and "Back to search". The main title is "Gene Sequence", and the subtitle is "Gene Sequence information about entry 2pe2". On the right side, there is a vertical menu with options: "Gene-level view", "Protein-level view", "Structural view", "Functions", "Localization", "Pathways", "Medicine", and "References". The main content area is divided into two sections: "Primary gene sequence sources" and "Additional gene sequence information". The "Primary gene sequence sources" section lists: Gene name (PDPK1, PDK1), Organism (Homo Sapiens), Taxonomy ID (9606), GenBank nucleotide ID(s) (AF017995 | Y15056 | BC012103), and GenBank gene ID (5170). The "Additional gene sequence information" section lists: Ensembl: (ENSG00000140992), HOVERGEN: (vertebrate homologs), KEGG entry, KEGG orthology, and H-InvDB:human (HIX0012730 | HIX0027031). At the bottom, there is a section titled "Model organism databases" which lists: MGI: Mouse (MGI:1338068) and RGD: Rat (620307). Annotations on the left side of the image point to various parts of the interface: "Navigation area" points to the top buttons, "Primary sources" points to the "Primary gene sequence sources" section, "Additional info" points to the "Additional gene sequence information" section, "Model organism database" points to the "Model organism databases" section, and "Other informational views" points to the right-hand menu.

Navigation area

Primary sources

Additional info

Model organism database

Other informational views

Back to Summary

Back to search

Gene Sequence

Gene-level view

Protein-level view

Structural view

Functions

Localization

Pathways

Medicine

References

Gene Sequence information about entry 2pe2

Chain(s) : A

Primary gene sequence sources:

Gene name PDPK1, PDK1

Organism Homo Sapiens

Taxonomy ID 9606

GenBank nucleotide ID(s) AF017995 | Y15056 | BC012103

GenBank gene ID 5170

Additional gene sequence information:

Ensembl: ENSG00000140992

Genome browser

HOVERGEN: vertebrate homologs

KEGG entry

KEGG orthology

H-InvDB:human HIX0012730 | HIX0027031

Model organism databases:

MGI: Mouse MGI:1338068

RGD: Rat 620307

Models Report

PSI | nature Structural Biology Knowledgebase

home
structural biology update
targets
protein structures, sequences, and function
homology models
methods
publications
about this site
about PSI
NPG resources

Search Results

Summary Structures

Total 33 Model(s) found

Listed below are the computational models for your query sequence. Click on the "view" link to view the details.

33 models found

view

Providers

MODBASE:	15
NYSGXRC:	8
SWISSMODEL:	10

Model quality influenced by:

- Template flexibility & variation
- Methodology limitations
- Evolutionary distances

PMP | Query Result:

Summary **PMP Documentation**

Models found: 33

Models:

Index	Model	Rel.	Provid	Type	Templates	%Seq id	from	to	Target Proteins
1	[Show]								5530
2	[Show]								TRL2
3	[Show]								UGN6
4	[Show]								TRL2
5	[Show]								Z2A0
6	[Show]								K3L3
7	[Show]								Z2A0
8	[Show]								5530
9	[Show]								K3L3
10	[Show]								10Z4
11	[Show]								UGN6
12	[Show]								Z2A0
13	[Show]								5530
14	[Show]								TRL2
15	[Show]								UEW8
16	[Show]								UGN6
17	[Show]								10Z4
18	[Show]								UPJ8
19	[Show]								UPJ8
20	[Show]								K3L3
21	[Show]								UPJ8
22	[Show]		MODBASE	SC	1257	99%	83	340	Q8R2L4
23	[Show]		MODBASE	SC	3c5		4	394	O15530
24	[Show]		MODBASE	SC					
31	[Show]		MODBASE	SC					

Mouse over signal

The sequence identity between target protein and template structure is commonly seen as a first indicator for the expected accuracy of a model, as confirmed by various studies.

This model is based on target-template sequence alignment of 99% sequence identity (C). Typically, target-template sequence alignments of this degree of similarity are mainly correct. Structural variation in templates, and incorrect reconstruction of loops (insertions and deletions) are the main sources of model inaccuracies. Validation of the model quality and analysis of the variability among template structures is advised.

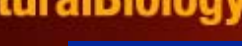
Experimental Structures:

PDB	Title	%Seq id	from	to
3iop	PDK-1 in complex with the inhibitor	100	48-359	
3ion	PDK1 in complex with Compound 8h	100	48-359	
2r7b	Crystal Structure of the Phosphoinositide-dependent Kinase-1 (PDK-1) Catalytic Domain bound to a dibenzonaphthyridine inhibitor	100	48-359	

Quality estimation and explanation for model

Fig.1: For individual model colors please refer to Fig. 2. The N termini of the models are marked with N. Please select individual superpositions onto the model with the highest sequence coverage, by clicking the links below in the Superposition Results section. **Reset**

Target Report



ID: HR3403B Latest update: 2006-09-06 TargetDB Status: active

- home
- structural biology
- targets
- protein structure sequences, and f
- homology models
- methods
- publications
- about this
- about PSI
- NPG resources

Annotation Quick Reference - HR3403B

NMPDR	PROSITE	Gene3D	PDB	IntE
Ensembl	Pfam	ProDom	SCOP	PRC
UniGene	InterPro	DisProt	CATH	Pro
Nucleotide	SMART	PRINTS	PDBSUM	EC
GeneID	UniProt	Phos-PTMs	SMR	GO
DNA Sequence	Protein Sequence	Protein Structure	Fun	

Go to available annotations (green) by clicking

TargetDB : HR3403B

Gene Sequence

Protein Sequence

Protein Structure

Functions

Localization

Pathways

Medicine

References

TargetDB

Target Status

Source organism

Pfam

Pfam

Pfam

Pfam

Pfam

Pfam

MEGA

Lab: NESG

Name: Ribosomal protein S6 kinase alpha-2 (EC 2.7.11.1) (S6K-alpha 2) (90kDa ribosomal protein S6 kinase 2) (p90-RSK 2) (Ribosomal S6 kinase 3)(RSK-3) (pp90RSK3) (MAP kinase-activated protein kinase 1c)(MAPKAPK1C)

Status: Selected

URL: <http://spine.nesg.org/target.cgi?id=HR3403B>

Database Reference: MEGA: [3.90.1200.10](#) PFAM: [PF00433](#) PFAM: [PF06293](#) PFAM: [PF07714](#) PFAM: [PF00069](#) MEGA: [3.30.200.20](#) MEGA: [1.10.510.10](#)

Source Organism: Homo sapiens

Sequence:

```
GIIYRDLKPENILDEEGHIKITD FGLSKEAIDHDKRAYSF CGTIEYMAPEVVMNRGHTQS
ADNWSFGVLMFEMLTGSLPFQ GKDRKETMALILKAKLGMPQFLSGEAQSLRLALFKRNPC
NRLGAGIDGVVEIKRHPFFVTIDWNTLYRKEIKPPFKPAVGRPEDTFHFDPEFTARTPTD
SPGVPPSANAHLFRGFSFVASSLIQEPSQDLHKVPVHPIVQQLHGNNIHFTDGYEIKE
DIGVGSYSVCKRCVHKATDTEYAVKIIDKSKRDPSEEIEILLRYGQHPNIITLKDVYDDG
KFVYLVMELMRGGELLDRILRQRYFSEREASDVLCTITTKTMDYLHSQGVVHRDLKPSNIL
YRDESGSPESIRVCD FGF AKQLRAGNGLLMTPCYTANFVAPEVLKRQGYDAACDIWSLGI
YTMLAGTFPFANGPDDTPEEILARIGSGKYALSGGNWDSISDAAKDVVSKMLHVDPHQ
RLTAMQVLKHPVWVMREYLSPNQLSRQDVHLVKGAMAATYFALNRTQPAPRLEPVLSNLL
AQRGMKRLTSTRL
```

Domain Annotation: [Pfam](#)

Other Sources: [PSI SBKB](#) | [Superfamily](#) | [TIGR Families](#) | [ProDom](#) | [iProClass](#) | [Prosit](#)

Protein Properties:

Number of Residues	Mol. Weight	Avg. Hydropathy Score	Charge	pI Value
555	62669	-0.352	9.5	7.7

[New Search](#)

Protocols Report

PSI | nature
StructuralBiology Knowledgebase

- home
- structural biology update
- targets
- protein structures, sequences, and function
- homology models
- methods
- publications
- about this site
- about PSI
- NPG resources

Search Results

Summary	Structures	Models	Ta
---------	------------	--------	----

Protocol(s) for 475 Target(s) found

Listed below are the experimental protocols for each determination pipeline.
Browse these protocols to learn what worked and View the entire PepcDB record by clicking on the Pe

PepcDB : HR3403B | (Similarity: I = 59% E = 2.0E-
[PepcDb Summary](#) [selection](#)

PepcDB : HR3377B | (Similarity: I = 55% E = 7.9E-
[PepcDb Summary](#) [selection](#)

PepcDB : HR3527B | (Similarity: I = 52% E = 8.5E-
[PepcDb Summary](#) [selection](#)

PepcDB : HR3107B | (Similarity: I = 51% E = 7.1E-
[PepcDb Summary](#) [selection](#) [cloning](#) [expression](#)

PepcDB : HR3107A | (Similarity: I = 51% E = 7.1E-19) [View alignment](#)

PepcDB | Target Query Results
There is **1** target that match your request.

ID: [HR3107B](#) **Site:** NESG
Protein Name: Serine/threonine-protein kinase PRKY (EC 2.7.11.1)
Target Sequence
Sequence Type: predicted dna

TATCGTTTACAAGATTGTGATGCTCTGGTTACGATGGGTACTGGTACATTTGGGCGTGTAC
ATTGGTCAAAGAGAAGACCGGAAACACTTCTTTGCCCTTAAAGTGATGTCATTCGGG
ACCTTATCCGCGAAGCAGGAGCAAGATTCGACACAGCAATCGTATCTAAGAAAG
CGGTTGACTGGTGGGCTTGGGAATTCTTATTTTGAATGCTGCTGCTGCCITCCCCCGT
TTTTCGATGACAATCCATTTGGTATCTATCAGAAGATACTGGCGGGCAAACCTCTATTTCC
CTCGTCATTAGATTTTCACGTGAAAACAGGTCGCATGATG

Sequence Type: protein

YRLQDCDALVTMGTTGFRVHLVKEKTAKHFFALKVMSIPDVIRRKQEQHVHNEKSVLKEV
SHPFLIRLFWTWEERFLYMLMEYVPGGELFSYLRNRGHFSSTTGLFYSAEIIICAEYHLH
SKEIVYRDLKPENILLDRDGHKILDFGFAKKLVDRWTLCGTPEYLAPEVIQSKGHGRA
VDWUWALGILIFEMLSGFPPFFDDNPFGIYQKILAGKLYFPRHLD FHVKTGRMM

Other Sources:
[PSI SBKB](#) | [Pfam](#) | [Superfamily](#) |
[TIGR Families](#) | [ProDom](#) | [iProClass](#) | [Prosit](#)

Site Target Info
<http://spine.nesg.org/target.cgi?id=HR3107B>
Source Organism: Homo sapiens
Database Reference: MEGA: [3.30.200.20](#) MEGA: [1.10.510.10](#)
PFAM: [PF07714](#) PFAM: [PF00069](#) MEGA: [3.90.1200.10](#)

Trial: Number of Trials: 2 ([view all trials](#))
Latest update: 2008-12-02
Current Status: expressed
Status History: selected, cloned, expressed ([view details](#))
Sequence of the Experimental Trial:
MGHHHHHSHYRLQDCDALVTMGTTGFRVHLVKEKTAKHFFALKVMSIPDVIRRKQEQHV
HNEKSVLKEVSHPFLLIRLFWTWEERFLYMLMEYVPGGELFSYLRNRGHFSSTTGLFYSA
EIIICAEYHLH SKEIVYRDLKPENILLDRDGHKILDFGFAKKLVDRWTLCGTPEYLAPE
VIQSKGHGRAVDWUWALGILIFEMLSGFPPFFDDNPFGIYQKILAGKLYFPRHLD FHVKTG
RMM


Protocol Reference:
[selection](#) [cloning](#) [expression](#)

[New Search](#)

PepcDB = Protein expression purification and crystallization DataBase

Materials Report

PSI | nature
StructuralBiologyKnowledgebase



home

structural biology update

targets

protein structures,
sequences, and function

homology models

methods

publications

about this site

about PSI


NPG resources



Search Results

SummaryStructuresModelsTargetsProtocolsMaterials

Total 4 Material(s)

Listed below are the materials that match the search criteria. Clicking the TargetDB link will take you to the TargetDB database.

TargetDB : GO
Target DB
Order Clone 
Clone Details
Vector Details

TargetDB : GO
Target DB
Order Clone 
Order Clone 
Clone Details
Clone Details
Vector Details

ASU BIODESIGN INSTITUTE
ARIZONA STATE UNIVERSITY

Home About PSKBiology-MR Search Deposit FAQ News Contact

LoginNew User

Search Results

Search Results




Matches found1Search terms not found0

List of search terms found : 1

No of Result Per Page: 50Display ResultPage: 1DisplayDownloadExplanation of Terms

Search Term	Clone ID	Original Clone ID	Clone Type	Species Specific ID	Gene Symbol	Keywords	Gene Name	Reference Sequence	Mutation/Discrepancy	Insert Format	Vector	Selection Markers	
1 ATCD000334209	ATCD000334209	5467.1.11_GO.23919	cDNA	930330	PDK1	full-length cds	PDK1 (3'-PHOSPHOINOSITIDE-DEPENDENT PROTEIN KINASE 1); 3-phosphoinositide-dependent protein kinase/ kinase/ phosphoinositide binding / protein binding / protein kinase	NM_120533	No/No	CLOSED	pVP13	bacterial : ampicillin bacterial : chloramphenicol	Add To Cart

No of Result Per Page: 50Display ResultPage: 1Display

ScCD00085124 
ScCD00332121 
pDONR221 

Structure (PDB ID) Search

The screenshot shows the PSI | nature Structural Biology Knowledgebase interface. On the left is a navigation menu with links like 'home', 'structural biology update', 'targets', 'protein structures, sequences, and function', 'homology models', 'methods', 'publications', 'about this site', 'about PSI', and 'NPG resources'. The main content area is titled 'Search Results' and features a tabbed interface with 'Summary', 'Structures', 'Models', 'Targets', 'Protocols', and 'Materials'. The 'Summary' tab is active, displaying search details: 'Search type: PDB ID', 'Your query: 3IHR', and a list of results. The results are highlighted with a red box and include: 'Similar protein structures from the Protein Data Bank: [1](#)', 'Similar theoretical models from the Protein Model Portal: [17](#)', 'Similar protein targets in TargetDB: [1](#)', 'Similar experimental protocols in PepcDB: [1](#)', and 'Similar materials available from PSI Materials Repository: [1](#)'. Below the search results is a search bar with radio buttons for 'by sequence', 'by text', and 'by structure (PDB id)'. The 'by structure (PDB id)' option is selected, and the query '3IHR' is entered in the search box. To the right of the search bar is a 'propose targets' section with an 'evaluate targets' button.

PSI | nature Structural Biology Knowledgebase

Search Results

Summary Structures Models Targets Protocols Materials

Select a tab from the top or follow the links below for detailed results

Search type: PDB ID

Your query: 3IHR

Results:

- Similar protein structures from the Protein Data Bank: [1](#)
- Similar theoretical models from the Protein Model Portal: [17](#)
- Similar protein targets in TargetDB: [1](#)
- Similar experimental protocols in PepcDB: [1](#)
- Similar materials available from PSI Materials Repository: [1](#)

search Explore proteins and this website

by sequence ☐ 3IHR

by text ☐

by structure (PDB id) ☒

example query

help search

propose targets

Evaluate your sequences or nominate them to the PSI

evaluate targets

PDB ID = very specific, only single result!

4. Text Searches

The screenshot displays the PSI search interface. At the top, a dark red header bar contains the word "search" in white and the text "Explore proteins and this website" in orange. Below this, a light gray search box contains the text "green fluorescent protein". To the left of the search box, there are three radio buttons: "by sequence" (unselected), "by text" (selected), and "by structure (PDB id)" (unselected). A light blue arrow points from the text "by text" to the "by text" radio button. Below the search box, there is a red button labeled "search" and a gray button labeled "help".

Below the search interface, the "Text Search Results" page is shown. It features a tabbed interface with tabs labeled "Summary", "Gateway", "PSI sites", "Structures", "Annotations", "Methods", and "Tech". The "Summary" tab is selected. Below the tabs, there is a section titled "Text Search Results" with the following information:

- Search type: Text Search
- Your query: GREEN FLUORESCENT PROTEIN
- Results:
 - 7 Articles from the SBKB **Gateway**
 - 4 Hits from **PSI Center Websites**
 - 239 **Structures** and associated articles
 - 235 Biological **Annotations** from PSI and other databases
 - 8 **Method** articles in the Publication Portal
 - 2 Reports from the PSI **Technology** portal

At the bottom of the results page, there is a note: "Models Portal, TargetDB (targets), PepcDB (protocols), and the Material Repository, temporarily available only by sequence or PDB ID search".

Text Search Results

Summary

Gateway

PSI sites

Structures

Annotations

Methods

Tech

found in 239 publications with protein structures.

1emc [↗]

Title : The structural basis for spectral variations in green fluorescent protein.

Authors : Palm GJ;Zdanov A;Gaitanaris GA;Stauber R;Pavlakakis GN;Wlodawer A

Journal : Nat Struct Biol

Volume : 4 | Issue : 5 | Pub

PubMed : **9145105** [↗]

Abstracts : 1: nat struct bio

green fluorescent...

Read More [↗]

same PubMed article

1eme [↗]

Title : The structural basis for spectral variations in green fluorescent protein.

Authors : Palm GJ;Zdanov A;Gaitanaris GA;Stauber R;Pavlakakis GN;Wlodawer A

Journal : Nat Struct Biol

Volume : 4 | Issue : 5 | Pub

PubMed : **9145105** [↗]

Abstracts : 1: nat struct bio

green fluorescent...

Read More [↗]

same PubMed article

1emf [↗]

Title : The structural basis for spectral variations in green fluorescent protein.

CATH results

Text Search Results

[Summary](#)
[Gateway](#)
[PSI sites](#)
[Structures](#)
[Annotations](#)
[Methods](#)
[Tech](#)

Total 235 Annotations found for query:


1B9C Chain: A
 CATH : **Class: Mainly Beta |Architecture: Beta Barrel |Topology: Green Fluorescent Protein |Homology: Green Fluorescent Protein (1B9Ca00) ***

1B9C Chain: B
 CATH : **Class:**
|Homology:

1B9C Chain: C
 CATH : **Class:**
|Homology:

1B9C Chain: D
 CATH : **Class:**
|Homology:

1BFP Chain: A
 CATH : **Class:**
|Homology:


 [Quick Search](#)

[Home](#)
[Search](#)
[Documentation](#)
[Tools](#)
[Download](#)
[About](#)

CATH v3.2.0

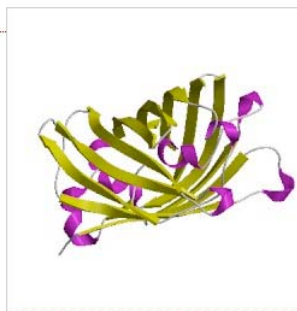
[Home](#) > [Search](#) > [Domain](#) > 1b9cA00

[CathDB: V3_2_0 \(change\)](#)

CATH Domain: 1b9cA00 [XML](#)

PDB 1b9c, Chain A, Domain 0


CATH Code	Level Description	Links
2	Mainly Beta	
2.40	Beta Barrel	
2.40.155	Green Fluorescent Protein	
2.40.155.10	Green Fluorescent Protein	[Gene3D]
2.40.155.10.1		
2.40.155.10.1.1		
2.40.155.10.1.1.1		
2.40.155.10.1.1.1.26		
2.40.155.10.1.1.1.26.1		[Gene3D]



1b9cA00

[Structure](#)
[Sequence](#)
[History](#)

Segment boundaries for domain 1b9cA00 [\[top\]](#)



Domain ID	Start Res	Stop Res	Name	Length
1b9cA00	4	230		224

Tech Report Results

Text Search Results

Summary	Gateway	PSI sites	Structure
<p>found in 2 PSI Technology reports.</p> <p>Title : Factorial evolved auto-induction medium Center : Center for Eukaryotic Structural Genomics Summary : Factorial design methods were used to optimize the auto-induction of two model proteins, Photinus luciferase and enhanced green fluorescent protein, including variations for production of both unlabeled and selenomethionine-labeled samples. Read More...</p> <p>Title : Small-scale semi-automated purification of eukaryotic proteins Center : Center for Eukaryotic Structural Genomics Summary : A simple approach that allows cost-effective automated purification of proteins in levels sufficient for functional characterization or structural studies of human stem cell proteins, an engineered version of green fluorescent protein was included. Read More...</p>			

PSI Center
Technology
Summary
Description
Figures
Publication
Contact

PSI | SBKB Technology Portal

Factorial evolved auto-induction medium

Center
Center for Eukaryotic Structural Genomics

Technology
Protein Expression

Summary
The auto-induction method of protein expression in E. coli is based on diauxic growth resulting from dynamic function of lac operon regulatory elements (lacO and LacI) in mixtures of glucose, glycerol, and lactose. Our results show that successful execution of auto-induction is strongly dependent on the plasmid promoter and repressor construction, on the oxygenation state of the culture, and on the composition of the auto-induction medium.

Description
The auto-induction method of protein expression in E. coli is based on diauxic growth resulting from dynamic function of lac operon regulatory elements (lacO and LacI) in mixtures of glucose, glycerol, and lactose. Our results show that successful execution of auto-induction is strongly dependent on the plasmid promoter and repressor construction, on the oxygenation state of the culture, and on the composition of the auto-induction medium. Thus expression hosts expressing high levels of LacI during aerobic growth exhibit reduced ability to effectively complete the auto-induction process. Manipulation of the promoter to decrease the expression of LacI altered the preference for lactose consumption in a manner that led to increased protein expression and partially relieved the sensitivity of the auto-induction process to the oxygenation state of the culture. Factorial design methods were used to optimize the chemically defined growth medium used for expression of two model proteins, Photinus luciferase and enhanced green fluorescent protein, including variations for production of both unlabeled and selenomethionine-labeled samples (Blommel et al., 2007). The optimization included studies of the expression from T7 and T7-lacI promoter plasmids and from T5 phage promoter plasmids expressing two levels of LacI. Upon the basis of the analysis of over 500 independent expression results, combinations of optimized expression media and expression plasmids that gave protein yields of greater than 1000 µg/mL of expression culture were identified. These approaches are incorporated into CESC Technology Dissemination Reports 010, 020, 021, and 022.

Figure

Publication
Blommel PG, Becker KJ, Duvnjak P, Fox BG. "Enhanced bacterial protein expression during auto-induction obtained by alteration of lac repressor dosage and medium composition." Biotechnol. Prog. 23, 585-98 (2007). [Pubmed:17506520](#) | [Search SBKB Publications portal](#)

Contact
Brian Fox (bgfox@biochem.wisc.edu)

Availability
Email for details.

[Related articles](#)

5. Additional Features

The screenshot shows the PSI | nature Structural Biology Knowledgebase website. The header includes the logo and navigation icons. A left sidebar contains a menu with items like 'home', 'structural biology update', 'targets', 'protein structures, sequences, and function', 'homology models', 'methods', 'publications', 'about this site', 'about PSI', and 'NPG resources'. The main content area features a welcome message, a search bar with options to search by sequence, text, or structure (PDB id), and a section for 'structural biology update' dated November 2010. This section includes articles on 'Isoxanthopterin deaminase', 'COX inhibition: Naproxen by proxy', 'ABA receptor diversity', 'Solving homodimeric structures with NMR', and 'Scaling up mutational scanning'. On the right, there are sections for 'e-alerts', 'propose targets', 'functional sleuth', and 'latest PSI statistics'. Callouts are present: a red box around the sidebar menu is labeled 'sources of additional information'; a light blue box points to the 'propose targets' section with the text 'evaluate or propose your amino acid sequences'; another light blue box points to the 'functional sleuth' section with the text 'help characterize novel protein sequences'; and a third light blue box points to the 'latest PSI statistics' section with the text 'see stats on PSI progress'.

PSI | nature
Structural Biology Knowledgebase

home
structural biology update
targets
protein structures, sequences, and function
homology models
methods
publications
about this site
about PSI
NPG resources

Welcome to the
Structural Biology Knowledgebase

Keep informed about advances in structural biology and structural genomics. Discover how protein sequences, three-dimensional structures and models relate to biological function. Stay up to date with the latest protocols, materials and technologies.

[SBKB tutorial](#) [What's new?](#)

search Explore proteins and this website

by sequence ☒
by text ☐
by structure (PDB id) ☐
[example query](#) [help](#) [search](#)

structural biology update November 2010
Research advances, news and events in structural biology

and molecule

[Isoxanthopterin deaminase](#)
Researchers at NYSGXRC have discovered the function of a new enzyme from the Global Ocean Sampling Project. [more](#)

research advances

[COX inhibition: Naproxen by proxy](#)
Structural basis of COX inhibition suggests that adverse effects of NSAIDs will not be easily avoided.

[ABA receptor diversity](#)
There are 14 members in the ABA receptor family in *Arabidopsis thaliana*. How these receptors differ in their response to a synthetic, seed-specific ABA agonist is now explored in two independent papers, using a combination of crystallography, NMR and biochemical approaches.

[Solving homodimeric structures with NMR](#)
Complexes are very difficult to solve in solution but are important biologically. Combining multiple approaches is the way forward.

[Scaling up mutational scanning](#)

e-alerts
Receive news of monthly updates by e-mail
[sign up](#)

RSS (monthly updates)
RSS (new molecules)

propose targets
Evaluate your sequences or nominate them to the PSI
[evaluate targets](#)

functional sleuth
Functional Sleuth presents PSI structures that lack full functional annotation
[begin exploring](#)

latest PSI statistics
New structures last month: 94
Total structures to date: 5097
Total distinct structures: 4371
[more](#)
[see latest structures](#)
[get test datasets](#)

evaluate or
propose your
amino acid
sequences

help characterize
novel protein
sequences

see stats on PSI
progress

propose targets

Evaluate your sequences or nominate them to the PSI



evaluate targets

functional sleuth

Functional Sleuth presents PSI structures that lack full

PSI Sequence Comparison and Analysis tool

Target Home

Your session id is: 1258479651511

Print Summary Report

Report submitted.

Title

cellcycle1

Status

In Preparation

Submit Date

2010-10-26

Summary Report

Refresh All

Target description: yeastcellcycleprotein1 - Finished

135 similar structures found ([view results](#))

PDB ID	Similarity(I)	E-Value	PSI Center	Target ID	Release Date
3CMM	100%	4.4E-28			2008-08-05
2JWZ	100%	4.4E-28			2008-01-08
2JT4	100%	4.4E-28			2007-08-25
2G3Q	100%	4.4E-28			
1WR1	100%	4.4E-28			
1Q0W	100%	4.4E-28			
10TR	100%	4.4E-28			
3JSV	95%	1.0E-26			2009-10-27
3I3T	95%	1.0E-26	SGC		2009-07-21
3HM3	95%	1.0E-26			2009-08-25
3H7S	95%	1.0E-26			2009-09-22
3H7P	95%	1.0E-26			2009-09-22

Similarity may not always be this high

IFAGKQLEDGRTLSDYN

click

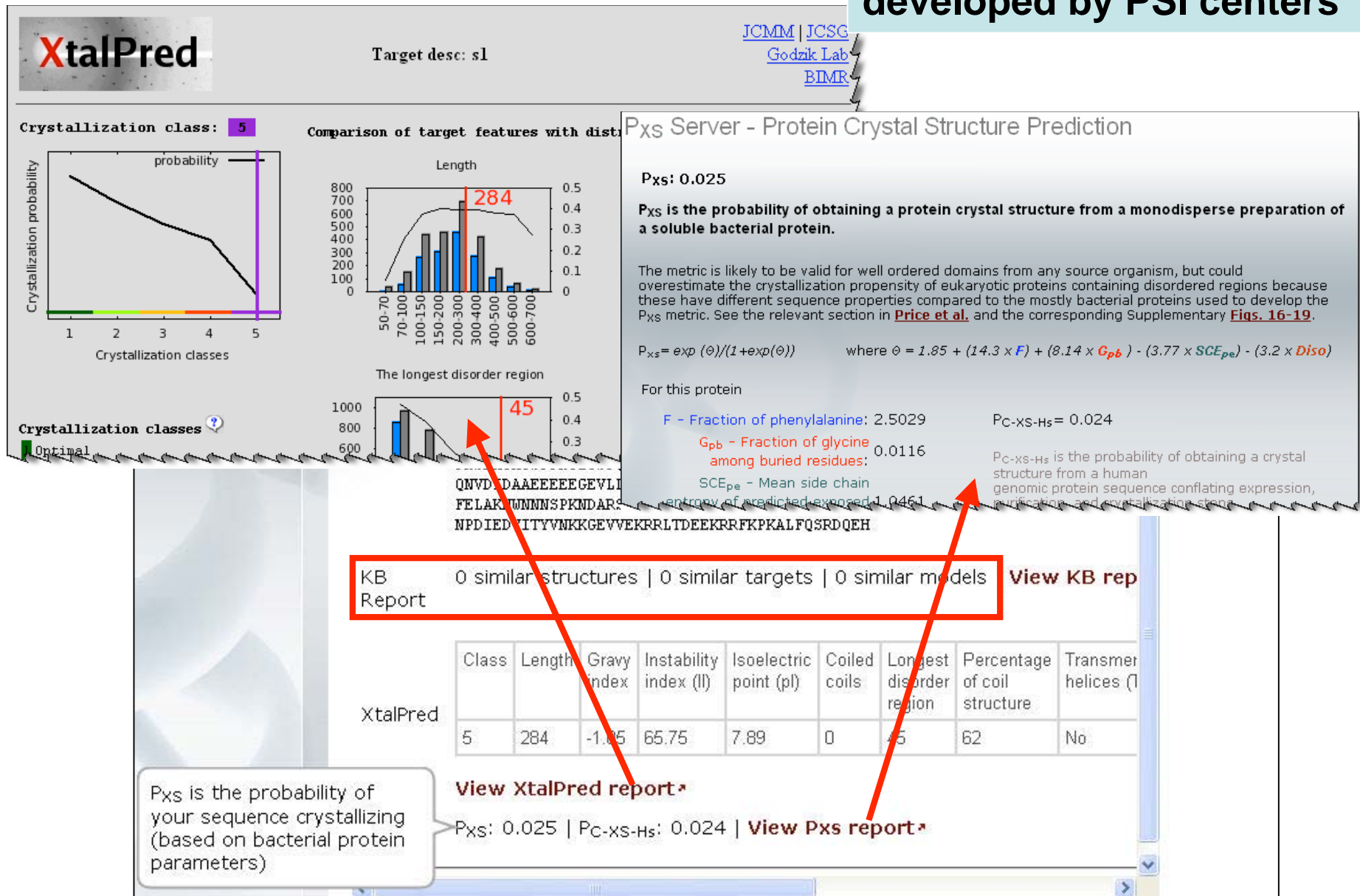
[View KB report](#)

d	Longest disorder region	Percentage of coil structure	Transmembrane helices (TM)	Signal peptide (SP)
	3	38	No	No

[report](#)

Proposed Targets, Example 2

XtalPred & Pxs: Two Sequence Analysis Tools developed by PSI centers



propose targets

Evaluate your sequences or nominate them to the PSI



evaluate targets

functional sleuth

Functional Sleuth presents PSI structures that lack full functional annotation



begin exploring

latest PSI statistics

New structures last month: 42

Total structures to date: 4378

Total distinct structures: 3714



more



see latest structures



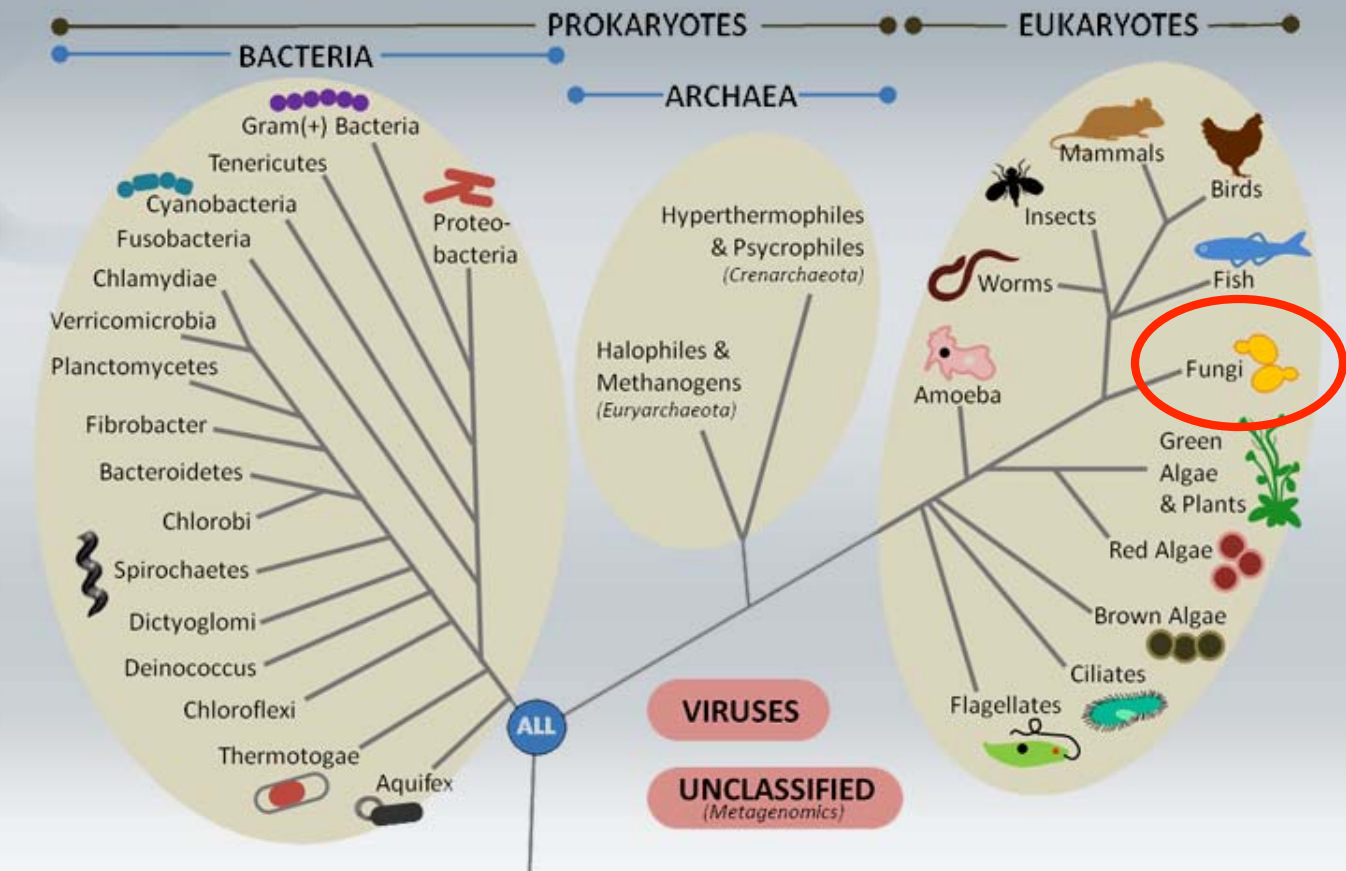
get test datasets

Functional Sleuth

Want to be a Functional Sleuth? Functional Sleuth enables further research for proteins in the Protein Data Bank archive whose functions are unknown or minimally characterized.

☒ View by Taxonomy Tree ☐ View by Center

These "structures of unknown function" (SUFs) are organized by source organism. Select a domain/superkingdom or group to begin; or jump to popular phyla or classes using our tree of life. Making a selection will launch our interactive tree browser. You can also [download a text file](#) of all SUF PDB IDs, updated weekly. Only organisms with structures are displayed.



Functional Sleuth



















Functional Sleuth

Want to be a Functional Sleuth? Functional Sleuth enables further exploration of the Protein Data Bank archive whose functions are unknown or minimally characterized.

☒ View by Taxonomy Tree ☐ View by Center

[Show Help](#) [Back to the Main Tree](#) [Go To Root](#)

Saccharomycetes

1f89	1jzt	1lxj	1n6z	1njr
				
1nkg	1nyn	1qvv	1qvw	1qvz
				
		1ty8	1wpi	1xb4
				
1xe7	1xe8	1y8m	1ycd	1yn8
				

Right click to view structures

PDB ID search

Click to progress through Taxonomy Tree

Mouse over for information

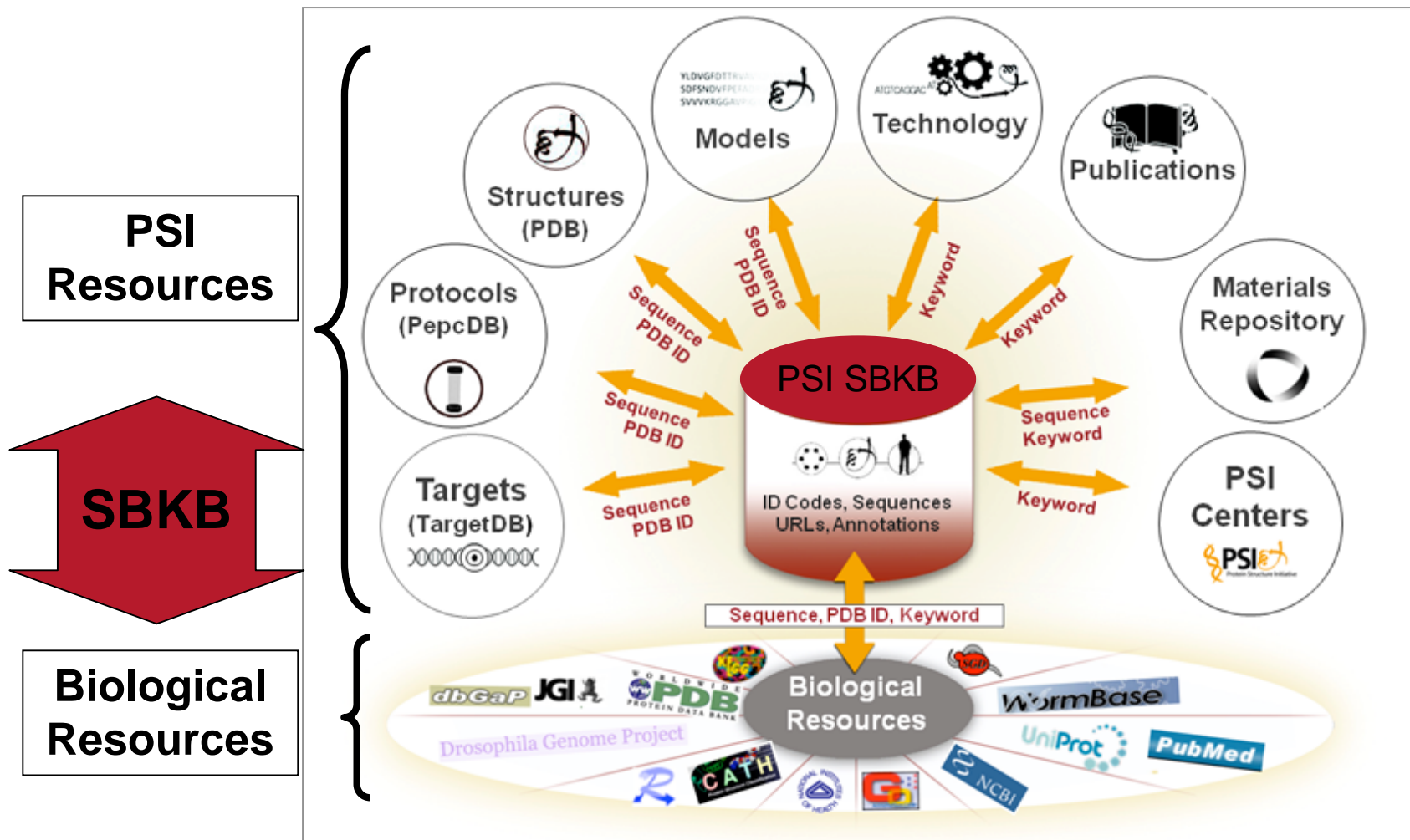
Ascomycota
Basidiomycota

Fungi

Microsporidia
Dikarya

Dikarya
Total Structures: 40
Right Click to see all the structures.

6. Summary



**SBKB: a Bridge between Biological and
PSI Structural Information**

7. Exercises

1. You recently read an interesting paper that mentioned the PDB structure [xxx](#). Search the PSI SBKB to learn more about this structure.

2. Imagine that you've joined your first rotation lab and have been offered the project of characterizing the protein product from domestic pig given below. You decide to begin your analysis of the sequence at the PSI SBKB to determine anything you can about this protein sequence. Are there any structures with similarity to your sequence? Are there any targets in other species with greater than 90% similarity?

```
>gi|259420073|emb|CBF63208.1| unnamed protein product [Sus scrofa domestica]  
MDPETCPCPTGGSCCTCAGSCKCEGCKCTSCKKSCCSCCPAECEKCAKDCVCKGGEGAEAEKSCCQ
```

3. You are working on a project trying to crystallize a membrane protein, but you are having trouble with your protocols. What helpful information can you find from the SBKB Research Library as well as the rest of the PSI SBKB?