

Protein Data Bank of molecular structures

Web: <https://www.rcsb.org/>

Select protein by code

The screenshot shows the RCSB PDB homepage. At the top, there is a search bar with the text "8osj" and a dropdown menu "3D Structures". Below the search bar, there is a "Help" link. The main content area features a "Welcome" section with text about the RCSB Protein Data Bank and a "December Molecule of the Month" section featuring a 3D molecular model.

Download

The screenshot shows the RCSB PDB Structure Summary page for protein 8OSJ. It displays a 3D ribbon diagram of the biological assembly. On the right, there is detailed information about the structure, including its classification as "GENE REGULATION", organism "Homo sapiens, synthetic construct, Mus musculus", and expression system "Escherichia coli, Escherichia coli 'BL21-CodonPlus(R) DE3'". The page also includes sections for "Experimental Data Snapshot" and "wwPDB V".

Data information:

```
HEADER      OXIDOREDUCTASE          10-JUN-99    1QQ2
TITLE       CRYSTAL STRUCTURE OF A MAMMALIAN 2-CYS PEROXIREDOXIN, HBP23.
COMPND     MOL_ID: 1;
COMPND     2 MOLECULE: THIOREDOXIN PEROXIDASE 2;
COMPND     3 CHAIN: A, B;
COMPND     4 SYNONYM: HEME-BINDING PROTEIN 23 KD, HBP23;
COMPND     5 ENGINEERED: YES;
COMPND     6 MUTATION: YES;
COMPND     7 OTHER_DETAILS: HBP23 IN OXIDATIVE FORM REVEALED A UNIQUE DIMER
COMPND     8 STRUCTURE IN WHICH THE ACTIVE SITE CYSTEIN FORMS A DISULFIDE BOND
COMPND     9 WITH OTHER CYSTEIN FROM ANOTHER SUBUNIT.
SOURCE     MOL_ID: 1;
SOURCE     2 ORGANISM_SCIENTIFIC: RATTUS NORVEGICUS;
SOURCE     3 ORGANISM_COMMON: NORWAY RAT;
SOURCE     4 ORGANISM_TAXID: 10116;
SOURCE     5 ORGAN: LIVER;
SOURCE     6 CELLULAR_LOCATION: CYTOPLASM;
SOURCE     7 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
```

Remarks

```
KEYWDS    THIOREDOXIN FOLD, OXIDOREDUCTASE
EXPDTA    X-RAY DIFFRACTION
AUTHOR    S.HIROTSU, Y.ABE, K.OKADA, N.NAGAHARA, H.HORI, T.NISHINO, T.HAKOSHIMA
REVDAT    3 03-NOV-21 1QQ2   1      REMARK SEQADV
REVDAT    2 24-FEB-09 1QQ2   1      VERSN
REVDAT    1 29-OCT-99 1QQ2   0      the vicinity of relevant SNPs will be identified. For further validation
JRNL      AUTH  S.HIROTSU, Y.ABE, K.OKADA, N.NAGAHARA, H.HORI, T.NISHINO, T.HAKOSHIMA from DisGeNet79
JRNL      AUTH 2 T.HAKOSHIMA disease relevance. The method will be compared with similar approaches
JRNL      TITL  CRYSTAL STRUCTURE OF A MULTIFUNCTIONAL 2-CYS PEROXIREDOXIN
JRNL      TITL 2 HEME-BINDING PROTEIN 23 KDA/PROLIFERATION-ASSOCIATED GENE
JRNL      TITL 3 PRODUCT.
JRNL      REF   PROC.NATL.ACAD.SCI.USA V. 96 12333 1999
JRNL      REFN  ISSN 0027-8424
JRNL      PMID 10535922
JRNL      DOI  10.1073/PNAS.96.22.12333
REMARK    2      mechanisms by which these variants affect or cause the pathology from a structural
REMARK    2 RESOLUTION. 2.60 ANGSTROMS.
REMARK    3      in the case of TFs by noticing the changes of the regulatory network, affecting
REMARK    3 REFINEMENT.  the native conformation of the TF, causing a conformational change
REMARK    3      the native conformation of the TF, causing a conformational change
REMARK    3      the native conformation of the TF, causing a conformational change
REMARK    3 DATA USED IN REFINEMENT.
```

Secondary Structure

Coordinates

ATOM	1	N	SER	A	3	51.504	49.004	26.114	1.00	83.57	N
ATOM	2	CA	SER	A	3	51.025	47.656	26.528	1.00	83.34	C
ATOM	3	C	SER	A	3	49.489	47.682	26.625	1.00	82.36	C
ATOM	4	O	SER	A	3	48.991	48.253	27.604	1.00	85.28	O
ATOM	5	CB	SER	A	3	51.580	47.260	27.903	1.00	81.65	C
ATOM	6	OG	SER	A	3	52.930	46.823	27.848	1.00	75.96	O
ATOM	7	N	GLY	A	4	48.755	47.076	25.658	1.00	79.10	N
ATOM	8	CA	GLY	A	4	47.278	47.050	25.676	1.00	69.66	C
ATOM	9	C	GLY	A	4	46.789	48.493	25.679	1.00	65.20	C
ATOM	10	O	GLY	A	4	46.548	49.097	24.614	1.00	69.57	O
ATOM	11	N	ASN	A	5	46.745	49.098	26.864	1.00	58.30	N

Format mmCIF

Correspondences PDB -> mmCIF

https://mmcif.wwpdb.org/docs/pdb_to_pdbx_correspondences.html

Databases of FOLDS: SCOP and CATH

Web: <https://scop.mrc-lmb.cam.ac.uk/>

About

Database description.

The SCOP database aims to provide a detailed and comprehensive description of the structural and evolutionary relationships between proteins whose three-dimensional structure is known and deposited in the Protein Data Bank. The main levels of the classification are:

Family groups closely related proteins with a clear evidence for their evolutionary origin. In most cases, their relationship is detectable with current sequence comparison methods, e.g. BLAST, PSI-BLAST, HMMER.

Superfamily brings together more distantly related protein domains. Their similarity is frequently limited to common structural features that along with a conserved architecture of active or binding sites, or similar modes of oligomerization suggest a probable evolutionary ancestry.

These relationships can sometimes span over structural regions of different size, therefore we provide domain boundaries for both, family and superfamily levels.

Fold groups superfamilies on the basis of the global structural features shared by the majority of their members. These features are the composition of the secondary structures in the domain core, their architecture and topology. Fold is an attribute of a superfamily but the constituent families of some superfamilies that have evolved distinct structural features can belong to a different fold.

IUPR (Intrinsically Unstructured Protein Region) organises superfamilies of proteins or protein regions that do not adopt globular folded structure. Some of these proteins exists in ensemble of different conformations or are unstructured in free state but adopt an ordered conformation upon binding to other macromolecules.

Classes bring together folds and IUPRs with different secondary structural content. These include all-alpha and all-beta proteins, containing predominantly alpha-helices and beta-strands, respectively, and 'mixed' alpha and beta classes (a/b) and (a+b) with respectively alternating and segregated alpha-helices and beta-strands, and the fifth class of small proteins with little or no secondary structures.

Protein type groups folds and IUPRs into four groups: soluble, membrane, fibrous and intrinsically disordered. Each of these types to a large extent correlates with characteristic sequence and structural features.

Each node in the SCOP classification can be uniquely identified by a seven digit identifier. The SCOP node identifiers will be stable across releases but may become obsolete if significant changes are introduced to the node.

The SCOP database is built as a classification of non-redundant protein domains. A representative is selected based on its sequence (UniProtKB) and structure (PDB) and used for manual SCOP classification. The SCOP domain boundaries are assigned to both, the PDB and UniProtKB entry. The manual classification of this selected representative is then automatically extended to related entries using SIFTS.

Usage

We hope that SCOP will have broad utility that will attract a wide range of users. Experimental structural biologists may wish to explore the structures similar to their proteins of research, while theoreticians will likely find it most useful to browse the wide range of protein structural and evolutionary relationships or use it for benchmarking of prediction methods. Molecular biologists may find the classification helpful because the categorization assists in locating protein domains of interest on the entire protein. We also hope that SCOP will find pedagogical use, for it organizes and annotates structures in an easily comprehensible manner and makes them accessible over the web.

How to access SCOP data.

The SCOP data can be accessed either by browsing or by searching. More details on how-to are provided below.

Using SCOP search box.

To perform a search use the search box at the top right corner of the browser window or the main search box at the home page. The SCOP search engine allows queries with free text, PDB ID, UniProt ID, SCOP ID. If the search term contains two or more words, these have to be concatenated by % sign, e.g. binding%domain. The search is case-insensitive, but the order of the words is important.

If the search is with a free text, the search results are grouped and displayed for folds, superfamilies and families if any search hits are found. If the search is with PDB ID that is a representative, the search result lists the domains for this PDB entry. If the PDB entry is not a SCOP representative, then the search result shows the classification of the entry that represents the query PDB. In all instances, the search result provides a link to the page of the node that matches the search criteria. If the search matches description or name of family, superfamily or fold then a link to their pages is provided. If the search match is a domain then a link to the domain page is provided. From the domain page you can navigate through the classification using the clickable ancestry chart.

Browsing SCOP.

When browsing SCOP, there are two ways of entry into the classification: by structural class or by protein type. Each class and type page lists the corresponding folds and IUPRs. A small folder under the fold name allows to display its constituent nodes (superfamilies or families).

The fold page provides the SCOP annotation for a given fold and lists the nodes that share this fold e.g. superfamily and/or family. Links are provided first for the constituent superfamilies and then for families if any exist. Under the superfamily name a small folder allows to display its constituent families. Next to the superfamily name is displayed the unique SCOP identifier for this node followed by a small icon that allows to view the structure of a superfamily domain that is selected to represent this superfamily.

Example: 2QL2

 About Contact Download

Search SCOP

[Keyword and ID search](#) Sequence search

Enter free text, SCOP ID, PDB ID or UniProt ID Go

Search results for 2ql2

[Domains \[4\]](#) Represented [4]

- 8076575 [2QL2 B:102-159](#)
- 8076576 [2QL2 B:102-159](#)
- 8076579 [2QL2 A:543-601](#)
- 8076580 [2QL2 A:543-601](#)

 About Contact Download

FAMILY HLH, helix-loop-helix DNA-binding domain

DOMAIN

2QL2 B:102-159

SCOP ID: 8076575

Neurogenic differentiation factor 1

Species *Mus musculus*

Show ancestry □

[Representative sequence](#) Other domains for this sequence External links

UniProtQ60867 This domain

MTKSYSSESGL MGEQPQPGPP SWTDECLSSQ DEEHEADKKE DELEAMNAEE DSLRNNGEEE EEEDEDLEEEE EEEEEEEEDQK PKRRGPKKK MTKARLERFK **LRNRMKANARE RNRMHGGLNAA LDNLRKVVPC YSKTQKLISKI ETLRLAKNYI WALSEILRNSC KSPDLVSVFVQ TLCKGLSQPT TNLVAGCQLQ NRPTFLPEQN PDMPPHLPTA SASFVVPHYS YQSPGLPSPP YTGMDSHHVF HVKFPPhAYS AALEPFFESP LTDCTSPSFID GPLSPPLSIN GNFSFKHEPS AEFEKNYAFT MHYPAATLAG PQSHGSIFSS GAAAPRCEIP IDNIMSFDSH SHHERVMSAQ LNAIFHD**

DOMAIN

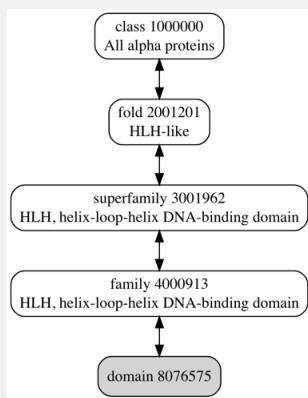
2QL2 B:102-159

SCOP ID: 8076575

Neurogenic differentiation factor 1

Species *Mus musculus*

Show ancestry 



Web: <https://www.cathdb.info/>

What is CATH?

The CATH database is a free, publicly available online resource that provides information on the evolutionary relationships of protein domains. It was created in the mid-1990s by [Professor Christine Orengo](#) and colleagues, and continues to be developed by the [Orengo group](#) at University College London.

Table of Contents

- [What is CATH?](#)
- [How is CATH-Gene3D created?](#)
- [CATH Releases](#)
 - [CATH \(daily snapshot\)](#)
 - [CATH-Plus \(full release\)](#)
- [Open Source Software](#)
- [Contact us](#)

How is CATH-Gene3D created?

Experimentally-determined protein three-dimensional structures are obtained from the Protein Data Bank and split into their consecutive polypeptide chains, where applicable. Protein domains are identified within these chains using a mixture of automatic methods and manual curation. The domains are then classified within the CATH structural hierarchy: at the [Class](#) (C) level, domains are assigned according to their secondary structure content, i.e. all alpha, all beta, a mixture of alpha and beta, or little secondary structure; at the [Architecture](#) (A) level, information on the secondary structure arrangement in three-dimensional space is used for assignment; at the [Topology/fold](#) (T) level, information on how the secondary structure elements are connected and arranged is used; assignments are made to the [Homologous superfamily](#) (H) level if there is good evidence that the domains are related by evolution, i.e. they are homologous. To browse the classification hierarchy, see [CATH hierarchy](#).

Additional sequence data for domains with no experimentally determined structures are provided by our sister resource, [Gene3D](#), which are used to populate the homologous superfamilies. Protein sequences from UniProtKB and Ensembl are scanned against CATH HMMs to predict domain sequence boundaries and make homologous superfamily assignments.

Example: 2QL2

CATH

Home

Search

Browse

Download

About

Support

Search CATH by keywords or ID

Search CATH

2ql2

[Clear](#)

[Search](#)

Search CATH by text or ID

Type in general text or biological identifiers in the box and click "search" to perform a general text search on CATH data.

Examples: "protease", "1cuk"

[Search by Text or ID](#)

[Search by Sequence](#)

[Search by Structure](#)

Results

Currently displaying the top ranked hits from three separate search queries: CATH Superfamilies, CATH domains and PDB entries. Click "View all entries" to expand each section and show all the hits. Use the panel on the right to add additional filters to this query.

Current Search Filters

Remove search filters by clicking on the 'X'

2ql2^X

1 Matching CATH Superfamilies

[Info](#)

4.10.280.10

HLH, helix-loop-helix DNA-binding domain

Upstream stimulatory factor 1, cellular response to insulin stimulus, negative regulation of pro-B cell differentiation, negative regulation of transcription from RNA polymerase II promoter, transcription factor activity, sequence-specific DNA binding, RNA polymerase II transcription factor binding, erythrocyte differentiation, positive regulation of chromatin assembly or disassembly, positive regulation of mitotic cell cycle, positive regulation of protein complex assembly, positive regulation of NF-kappaB transcription factor activity, positive regulation of inflammatory response, *Saccharomyces cerevisiae* S288C

[View all entries](#)

4 Matching CATH Domains

[Info](#)

2ql2A00

PDB code 2ql2, chain A, domain 00

Superfamily: 4.10.280.10

gastrulation, positive regulation of muscle cell differentiation, response to lipopolysaccharide, sequence-specific DNA binding, regulation of transcription from RNA polymerase II promoter, lymphocyte differentiation, RNA polymerase II distal enhancer sequence-specific DNA binding, natural killer cell differentiation, RNA polymerase II core promoter proximal region sequence-specific DNA binding, protein heterodimerization activity, nucleus, protein homodimerization activity, protein stabilization, positive regulation

[View all entries](#)

DSSP

Execute:

```
dssp target.pdb target.dssp
```

- 1) Download a PDB structure, for example 2ql2.pdb
- 2) Execute dssp as

```
dssp 2ql2.pdb 2ql2.dssp
```
- 3) Open 2ql2.dssp (with gedit)

Results of H-Bonds to determine strands and helices:

```
REFERENCE W. KABSCH AND C.SANDER, BIOPOLYMERS 22 (1983) 2577-2637
HEADER TRANSCRIPTION/DNA 12-JUL-07 2QL2
COMPND MOL_ID: 1; MOLECULE: TRANSCRIPTION FACTOR E2-ALPHA; CHAIN: A, C; FRAGM
SOURCE MOL_ID: 1; ORGANISM_SCIENTIFIC: MUS MUSCULUS; ORGANISM_COMMON: MOUSE;
AUTHOR R.B.ROSE,A.LONGO,G.P.GUANGA
230 6 0 0 TOTAL NUMBER OF RESIDUES, NUMBER OF CHAINS, NUMBER OF SS-BRIDGES(TOTAL,INTRACHAIN,INTERCHAIN)
13829.6 ACCESSIBLE SURFACE OF PROTEIN (ANGSTROM**2)
182 79.1 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(J) , SAME NUMBER PER 100 RESIDUES
0 0.0 TOTAL NUMBER OF HYDROGEN BONDS IN PARALLEL BRIDGES, SAME NUMBER PER 100 RESIDUES
0 0.0 TOTAL NUMBER OF HYDROGEN BONDS IN ANTIPARALLEL BRIDGES, SAME NUMBER PER 100 RESIDUES
0 0.0 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(I-5), SAME NUMBER PER 100 RESIDUES
0 0.0 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(I-4), SAME NUMBER PER 100 RESIDUES
0 0.0 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(I-3), SAME NUMBER PER 100 RESIDUES
0 0.0 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(I-2), SAME NUMBER PER 100 RESIDUES
0 0.0 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(I-1), SAME NUMBER PER 100 RESIDUES
0 0.0 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(I+0), SAME NUMBER PER 100 RESIDUES
0 0.0 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(I+1), SAME NUMBER PER 100 RESIDUES
1 0.4 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(I+2), SAME NUMBER PER 100 RESIDUES
13 5.7 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(I+3), SAME NUMBER PER 100 RESIDUES
168 73.0 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(I+4), SAME NUMBER PER 100 RESIDUES
0 0.0 TOTAL NUMBER OF HYDROGEN BONDS OF TYPE O(I)-->H-N(I+5), SAME NUMBER PER 100 RESIDUES
```

Structural features determine secondary structure and accessible surface area:

#	RESIDUE	AA	STRUCTURE	BP1	BP2	ACC	N-H-->O	O-->H-N	N-H-->O	O-->H-N	TCO	KAPPA	ALPHA	PHI	PSI	X-CA	Y-CA	Z-CA
1	543	A	R	>>	0	0	113	0, 0.0	3,-1.1	0, 0.0	4,-1.0	0.000	360.0	360.0	360.0-141.7	8.3	56.9	-12.8
2	544	A	R	H 3> +	0	0	198	1,-0.3	4,-2.4	2,-0.2	3,-0.3	0.822	360.0	63.4	-56.6 -33.4	7.1	55.6	-16.1
3	545	A	M	H 3> S+	0	0	150	1,-0.2	4,-1.9	2,-0.2	-1,-0.3	0.861	98.7	53.7	-60.8 -36.3	10.6	56.1	-17.5
4	546	A	A	H <> S+	0	0	38	-3,-1.1	4,-1.7	2,-0.2	-1,-0.2	0.816	108.6	49.9	-69.6 -30.4	12.0	53.5	-15.1
5	547	A	N	H X S+	0	0	83	-4,-1.0	4,-3.0	-3,-0.3	5,-0.3	0.961	107.3	51.2	-72.5 -52.1	9.5	51.0	-16.3
6	548	A	N	H X S+	0	0	86	-4,-2.4	4,-2.6	1,-0.3	-2,-0.2	0.909	110.4	52.9	-49.9 -41.8	10.2	51.5	-19.9
7	549	A	A	H X S+	0	0	34	-4,-1.9	4,-1.5	2,-0.2	-1,-0.3	0.899	109.3	47.1	-60.5 -45.2	13.8	51.0	-19.0
8	550	A	R	H X S+	0	0	153	-4,-1.7	4,-2.5	-3,-0.2	3,-0.2	0.946	112.2	50.2	-62.8 -47.8	13.1	47.7	-17.2
9	551	A	E	H X S+	0	0	92	-4,-3.0	4,-2.6	1,-0.2	5,-0.2	0.898	107.6	53.0	-59.1 -42.5	11.0	46.4	-20.1
10	552	A	R	H X S+	0	0	157	-4,-2.6	4,-1.3	-5,-0.3	-1,-0.2	0.860	112.1	46.4	-62.4 -34.1	13.8	47.3	-22.7

α -helix is H,
 β -strand is E,
other definitions are for different types of turns.

Using Chimera

Web: <http://www.cgl.ucsf.edu/chimera/>

Description: In this seminar we will learn how to use Chimera for interactive visualization and analysis of molecular structures through a set of tutorials (based on the official tutorials).

Required:

Download files 1zik.pdb, 1d86.pdb and 3w7f.pdb from PDB: <http://www.rcsb.org>

Tutorial 1: Molecular representations Menu Version

Easy access to features without knowledge of commands and their syntax.

1. Start Chimera from the system

2. Open the 1zik.pdb

-> File...Open from the menu and file browse to locate the file

3. Apply a preset (predefined combinations of display settings):

-> Presets...Interactive 2 (all atoms). Color codes: red -> oxygens, blue -> nitrogens, tan (default for model) -> carbons

4. Move the structure using the mouse:

- Left mouse button: controls rotation
- Middle mouse button: controls XY translation
- Right mouse button: performs scaling (zooming)

Hovering the mouse cursor over an atom or bond (without clicking any buttons) will show identifying information in a pop-up “balloon.” The balloon will disappear when the cursor is moved away.

5. Use selections and actions menus:

The selection menu specifies atoms, bonds, residues, etc. for subsequent operations with the actions menu. When nothing is selected, the actions menu applies to everything.

-> Select the water (red dots): Select... Structure... solvent or Select... Residue... HOH

-> Hide the selected atoms: Actions... Atoms/Bonds... hide

Even though the water atoms are hidden, they are still selected.

-> Clear the selection: Select... Clear Selection

Include only alpha carbons (CA atoms)

-> Thicken the lines: Actions... Atoms/Bonds... wire width... 3

-> Display only the chain trace: Actions... Atoms/Bonds... backbone only... chain trace

Picking from the screen is done by clicking on the atom or bond of interest with the left mouse button while pressing the Ctrl key. To add to an existing selection, also press Shift.

-> Pick two atoms, one from each peptide chain: Ctrl-click the first, Shift-Ctrl-click the second

Label the atoms

-> Residue type, number and chain ID: Actions... Label...residue... name + specifier

Clear the selection

-> Ctrl-clicking in empty space, as if picking “nothing.”

Turn off the labels:

-> Actions... Label... residue... off

Color the two chains:

- > Select... Chain... A
- > Actions... Color... cyan
- > Select... Chain... B
- > Actions... Color... yellow

Select both chains and display its full backbone:

- > Pick any atom or bond of chain A (Shift-Ctrl-click) and press the up arrow key twice
- > Actions... Atoms/Bonds... backbone only... full

Display all atoms and color them by element:

- > Actions... Atoms/Bonds... show
- > Actions... Color... by element

Close the protein

Each file of coordinates opened in Chimera becomes a model with an associated model ID number (starting from 0). Models are listed in the Model Panel

-> Click Tools... General Controls... Model Panel.

A checkbox in the Active column of the Model Panel shows that the model is activated for motion; unchecking the box makes it impossible to move the model.

-> Click on 1zik.pdb and click close in the list of functions

6. Open the 1d86.pdb:

-> File...Open from the menu and file browse to locate the file

7. Apply the preset 2 (show the DNA as wire and netropsin as spheres)

-> Presets...Interactive 2 (all atoms).

8. Color carbons white and then undisplay the water:

- > Select... Chemistry... element... C
- > Actions... Color... white
- > Select... Structure... solvent
- > Actions... Atoms/Bonds... hide

9. Color the different nucleotides with different colours

- > Select... Residue... DA
- > Actions... Color... blue
- > Select... Residue... DC
- > Actions... Color... cyan
- > Select... Residue... DG
- > Actions... Color... yellow
- > Select... Residue... DT
- > Actions... Color... magenta

10. Display the structure in different styles:

- > Actions... Atoms/Bonds... sphere
- > Select... Chain... A
- > Actions... Atoms/Bonds... ball & stick
- > Select... Clear Selection
- > Actions... Atoms/Bonds... stick

11. Show ribbons (automatically hides backbone atoms):

- > Actions... Ribbon... show
- > Actions... Ribbon... edged
- > Actions... Ribbon... rounded

Nucleotide residues can be shown as “lollipops,” boxes, and a ladder (nucleotide objects).

- > Actions... Atoms/Bonds... nucleotide objects... settings...
 - > Show side (sugar/base) as to ladder
 - > Click Ladder Options, set Rung radius to 0.3 Å
 - > Click Apply
 - > Show side (sugar/base) as to tube/slab
 - > Show base orientation to false
 - > Click Slab Style tab, set slab style to skinny
 - > Click Slab Options tab, set Slab object to ellipsoid
 - > Click Apply
- Apply Nucleotide setting to a residue selection:
- > Use the sequence tool: Favorites... Sequence
 - > Show the sequence of chain A
 - > Click over a few nucleotides and quit sequence window
 - > Show base orientation to true
 - > Click Slab Options tab, set Slab object to box
 - > Click Apply

12. Show surface:

- > Actions... Atoms/Bonds... nucleotide objects... off
 - > Actions... Ribbon... hide
 - > Actions... Atoms/Bonds... ball & stick
- There are categories within structures such as main and ligand (default main)
- > Actions... Surface... show
 - > Actions... Surface... hide
 - > Select... Structure... ligand
 - > Actions... Surface... show
 - > Actions... Surface... mesh
- Surface colour can be specified separately from the colours of the underlying atoms
- > Actions... Color... all options... to open the colour dialog
 - > Change the Coloring applies to surfaces
 - > Click red
 - > Click close (which will automatically reset the coloring target back to all of the above)
- Remove surface representation:
- > Select... Clear Selection
 - > Actions... Surface... solid
 - > Actions... Surface... hide
- Show the surface of a selected region (adenine and thymine nucleotides in chain B)
- > Change the selection mode: Select... Selection Mode...
- append
- > Select... Residue... DA
 - > Select... Residue... DT
- > Change the selection mode: Select... Selection Mode...
- intersect
- > Select... Chain... B
 - > Actions... Surface... show
 - > Actions... Surface... transparency... 50%
- To prepare for any subsequent operations, restore the selection mode and clear the selection:
- > Select... Selection Mode... replace
 - > Select... Clear Selection (or Ctrl-click in empty space)
- Close the 1d86.pdb (as done with 1zik.pdb)

Tutorial 2: Molecular representations Command Version

Commands are more concise and powerful than menus (we can code macros).

1. Use the Favorites menu to show the Command Line.

2. Open the 1zik.pdb:

> open

Browse to locate the file

3. Apply the interactive preset #2

-> preset apply int 2. Color codes: red -> oxygens,
blue -> nitrogens, tan (default for model) -> carbons

4. Use actions and selections:

> color hot pink :lys

Chimera commands may include arguments and/or atom specifications (the color name hot pink is an argument and :lys is an atom specification for all residues named lys). A blank specification is interpreted as all applicable items.

> color hot pink

Many commands have “~” versions that perform the opposite function. The following will change a structure back to its default color

> ~color

The command help can be used to show the manual page for any command.

> help color

Include only alpha carbons (CA atoms)

> Thicken the lines: linewidth 3

> Display only the chain trace: show @ca

The word selected, sel, or picked can be used in commands to specify the current selection.

-> Pick two atoms, one from each peptide chain: Ctrl-click the first, Shift-Ctrl-click the second

Label the atoms

> label sel

> ~label sel

> rlabel sel

Clear the selection and turn off labels

-> Ctrl-clicking in empty space, as if picking “nothing.”

> ~rlabel

Turn off the labels:

> ~rlabel

Color the two chains:

> color cyan ::a

> color yellow ::b

Residues and atoms can also be specified, along with or independent of chain

> color orange :5-9.a,12.a,8.b

> color magenta :14-18

> disp :leu.b

> color green :leu.b@cb

Display the water

> disp solvent or disp :hoh

Display the full backbone of chain A:

> disp ::a@n,ca,c,o

Display all atoms and color them by element:

> show ::a

> disp

> color byelement

Close the protein
> close 0

5. Open the 1d86.pdb:

> open
Browse to locate the file

6. Apply the interactive preset #2

-> preset apply int 2

7. Color carbons white and then undisplay the water:

> color white C
> ~disp solvent

8. Color the different nucleotides with different colours

> color blue :da
> color cyan :dc
> color yellow :dg
> color magenta :dt

9. Display the structure in different styles:

> represent sphere
> repr bs ::a
> rep stick

10. Show ribbons (automatically hides backbone atoms):

> ribbon
> ribrep edged
> ribr rounded

Nucleotide residues can be shown as “lollipops,” boxes, and a ladder (nucleotide objects).
> nuc side ladder radius 0.3

> nuc side tube/slab shape ellipsoid orient false style skinny

Apply Nucleotide setting to a residue selection:

> nuc side tube/slab shape box orient true style skinny :8-10.a

11. Show surface:

> ~nuc
> ~ribbon
> rep bs

There are categories within structures such as main and ligand (default main)

> surface
> ~surf
> surf ligand or surf :nt
> surfreps mesh

Surface colour can be specified separately from the colours of the underlying atoms

> color red,s ligand
> surfreps solid

Remove surface representation:

> ~surf

Show the surface of a selected region (adenine and thymine nucleotides in chain B)

> surf :da,b,dt,b
> transp 50,s

Close the 1d86.pdb
> close 0

Tutorial 3: Structure analysis

1 Open the 3w7f.pdb
> open 3w7f

2. Apply preset 1 and delete chain b:

-> Presets...Interactive 1 (ribbons)
> delete :.b

3. Display labels:

> rlabel ligand
> del solvent
> rlab @/display
> setattr m residueLabelPos 2 (near CA instead of residue centroids)

4. Find distances between atoms

-> Ctrl-click to pick the sidechain oxygen of Ser 21
-> Shift-Ctrl-doubleclick on the nearest phosphate oxygen
-> Click Show Distance in the resulting context menu
-> Measure the distance between the sidechain oxygen of Tyr 248 and the same phosphate oxygen.

5. Find H-bonds

-> Select the FPS residues: Select... Residue... FPS
-> Start FindHBond: Tools... Structure Analysis... FindHBond

In that dialog, turn on the options:

-> Only find H-bonds with at least one end selected
-> Write information to reply log
-> Set Line width to 3
-> Click ok

Remove HBonds
>~hbond

6. Find Clashes and Contacts

- Clashes - unfavourable interactions where atoms are too close together; close contacts
- Contacts - all kinds of direct interactions: polar and nonpolar, favourable and unfavourable (including clashes)

-> Start FindHBond: Tools... Structure Analysis... Clashes/Contacts
-> With the FPS residues still selected, click Designate. Now, 48 atoms should be designated for checking against all other atoms.
-> Set the Clash/Contact Parameters to the default contact criteria
-> Set the Treatment of Clash/Contact Atoms to: Select, If endpoint atom hidden, show endpoint residue, Write information to reply log and turn off any other treatment options
-> Click Apply.
Close 3w7f.pdb
> close 0

7. Open 3eeb.pdb
> open 3eeb

8. Surface properties

Hydrophobicity:

```
> delete ::a  
> del solvent  
> surface  
> ~disp ions  
> setattr m stickScale 2  
> color yellow ligand & C  
> rangecolor kdHydrophobicity min dodger blue 0 white max orange red  
> rangecolor kdHydrophobicity min medium purple 0 white max tan  
> preset apply pub 1
```

Coulombic Electrostatic Potential:

```
> preset apply pub 1  
-> Start Tools... Surface/Binding Analysis... Coulombic Surface Coloring  
-> Click ok  
> set silhouetteWidth 3  
> light mode ambient  
> light mode two-point
```

PRACTICE WITH EXAMPLES

- 1) Download file of fold examples
- 2) Identify the secondary structures with Chimera
- 3) Change color of residues by hydrophobicity
- 4) Identify the hydrophobic core
- 5) Identify the number of domains and the names of the folds of the protein named "problem"
- 6) Identify the folds of the PDB structure 8osj and compare it with 2ql2