XXXVI Int. Congress of Phys. Sci.
TU-3 July 29, 2009

Tutorial-3 #2 Document

Structural modeling of proteins: Principle and application to an ion channel

Haruki Nakamura

PDBj, *Institute for Protein Research, Osaka University*

Daron M. Standley

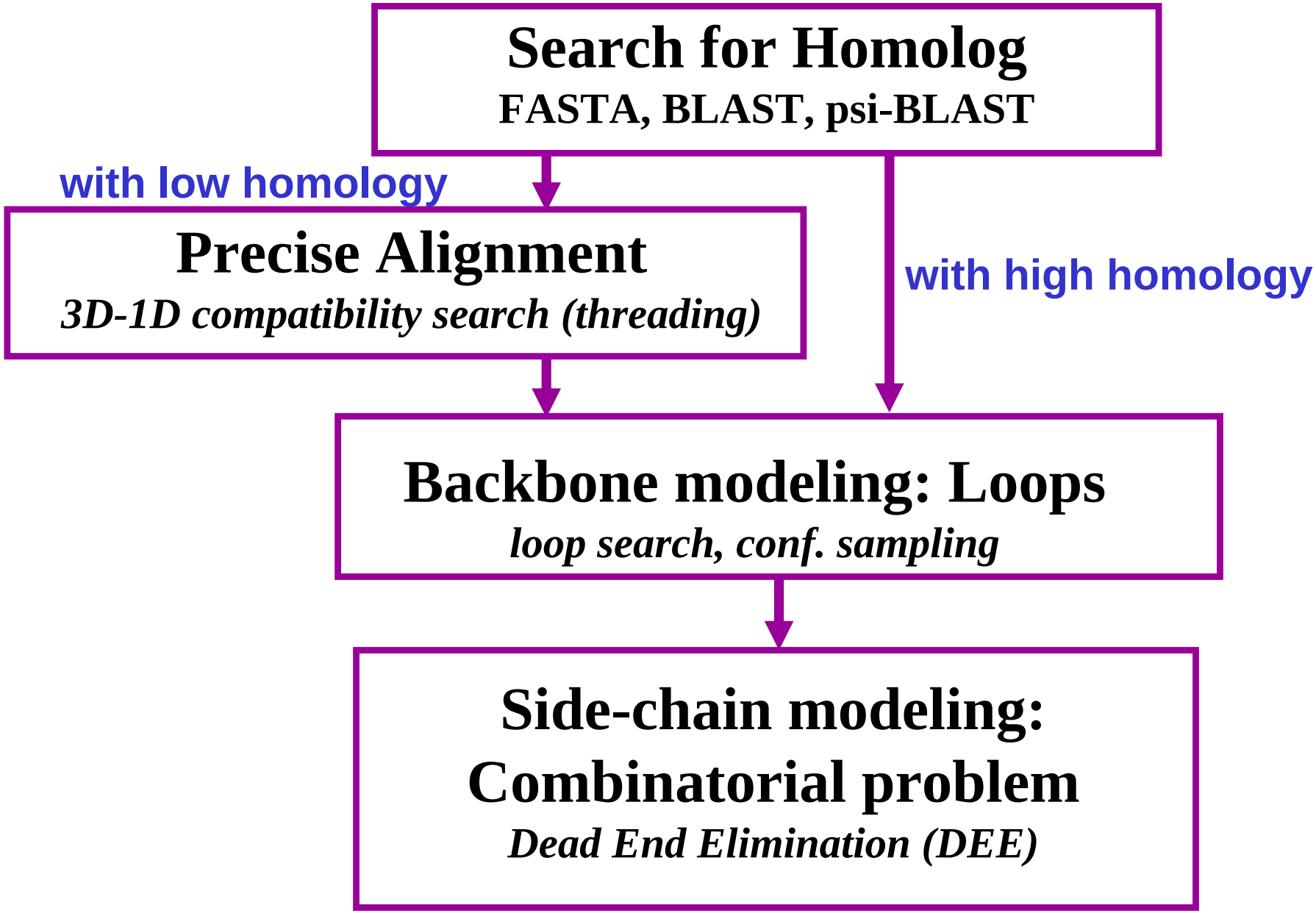
Immunology Frontier Research Center, Osaka University

Narutoshi Kamiya

The center for Advanced Medical Engineering and Informatics, Osaka University

http://www.protein.osaka-u.ac.jp/rcsfp/pi/
http://www.pdbj.org/

Goal of the Tutorial:
To construct the homology model of hERG channel



Step 1. Get amino-acid sequence of hERG channel from NCBI.

Access to NCBI (http://www.ncbi.nlm.nih.gov/), and get an amino acid sequence of Q12809 (region from S5 to S6: residues from 550 to 671).

Step 2. Make the alignment by threading.

Access to SFAS: Sequence to Function Annotation Server (*http://sysimm100.protein.osaka-u.ac.jp/sfas/*), and get the best alignment to the putative template structure.

The query may take several minutes to hours depending on the status of the Web site, please visit the example result page: (*http://sysimm100.protein.osaka-u.ac.jp/tmp/SFAS16483/hERG_top.html*)

Step 3. Make the tetramer structure and see the result.

Access to Quarternary Structure Service (*http://sysimm100.protein.osaka-u.ac.jp/pdb_quat/*), and get & see the tetramer model structure of hERG channel.

Step1: Get amino-acid sequence of hERG channel from NCBI (*http://www.ncbi.nlm.nih.gov/*)

NCBI Entrez Protein

My NCBI [Sign In] [Register]

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | OMIM | PMC | Journals | Books

Search Protein for Go Clear

Limits | Preview/Index | History | Clipboard | Details

Format: GenPept FASTA Graphics More Formats ▼ Download ▼ Save ▼ Links ▼

Try the Graphics report for a more informative view of the biological features.

Swiss-Prot: Q12809.1

RecName: Full=Potassium voltage-gated channel subfamily H member 2; AltName: Full=Voltage-gated potassium channel subunit Kv11.1; AltName: Full=Ether-a-go-go-related gene potassium channel 1; Short=H-ERG; Short=Erg1; Short=Ether-a-go-go-related protein 1; Shor...

Comment Features Sequence

```
LOCUS       Q12809                  1159 aa            linear   PRI 07-JUL-2009
DEFINITION  RecName: Full=Potassium voltage-gated channel subfamily H member 2;
            AltName: Full=Voltage-gated potassium channel subunit Kv11.1;
            AltName: Full=Ether-a-go-go-related gene potassium channel 1;
            Short=H-ERG; Short=Erg1; Short=Ether-a-go-go-related protein 1;
            Short=Eag-related protein 1; AltName: Full=eag homolog.
ACCESSION   Q12809
VERSION     Q12809.1  GI:7531135
DBSOURCE    UniProtKB: locus KCNH2_HUMAN, accession Q12809;
            class: standard.
            extra accessions:O75418,O75680,Q9BT72,Q9BUT7,Q9H3P0
            created: May 30, 2000.
            sequence updated: Nov 1, 1996.
            annotation updated: Jul 7, 2009.
            xrefs: U04270.1, AAA62473.1, AB009071.2, BAA37096.1, AF363636.1,
            AAL37559.1, AB044806.1, BAB19682.1, AJ512214.1, CAD54447.1,
            AJ010538.1, CAA09232.1, AJ010539.1, AJ010540.1, AJ010541.1,
            AJ010542.1, AJ010543.1, AJ010544.1, AJ010545.1, AJ010546.1,
            AJ010547.1, AJ010548.1, AJ010549.1, AJ010550.1, AJ010551.1,
            AF052728.1, AAC69709.1, BC001914.1, AAH01914.2, BC004311.2,
            AAH04311.2, I38465, NP_000229.1, NP_742053.1, NP_742054.1, 1BYW_A,
            1UJL_A
            xrefs (non-sequence databases): IPI:IPI00029662, IPI:IPI00172614,
            IPI:IPI00221190, IPI:IPI00221191, UniGene:Hs.647099, PDBsum:1BYW,
            PDBsum:1UJL, IntAct:Q12809, TCDB:1.A.1.20.1, PhosphoSite:Q12809,
            PRIDE:Q12809, Ensembl:ENSG00000055118, GeneID:3757, KEGG:hsa:3757,
            UCSC:uc003wib.1, UCSC:uc003wic.1, UCSC:uc003wie.1,
            GeneCards:GC07M150272, H-InvDB:HIX0007217, HGNC:6251,
            HPA:CAB006838, MIM:152427, MIM:609620, Orphanet:130, Orphanet:768,
            Orphanet:51083, PharmGKB:PA212, HOGENOM:Q12809, HOVERGEN:Q12809,
            OMA:Q12809, DrugBank:DB01118, DrugBank:DB00276, DrugBank:DB00637,
            DrugBank:DB01136, DrugBank:DB00604, DrugBank:DB00204,
            DrugBank:DB01218, DrugBank:DB00308, DrugBank:DB01100,
```

Change Region Shown

Customize View

Sequence Analysis Tools

- BLAST Sequence
- Conserved Domains

Articles about the KCNH2 gene

- Genetic Polymorphism of KCNH2 Confers Predispos [J Cardiovasc Electrophysiol. 2009]
- Breaking the gene barrier in schizophrenia. [Nat Med. 2009]
- Interactions of H562 in the S5 helix with T618 and S621 in the pore helix a [Biophys J. 2009]

» See all...

Identical Proteins for Q12809.1

- Sequence 2 from patent US 7541 [ACS12627]
- Sequence 5 from patent US 7537 [ACS08477]
- Sequence 2 from patent US 7510 [ACQ19114]

» See all...

RefSeq Protein Isoforms

See 3 reference sequence protein isoforms for the KCNH2 gene.

More about the KCNH2 gene

This gene encodes a voltage-activated potassium channel belonging to the eag family. It shares sequence similarity with the Drosophila ether...

Also Known As: ERG1, HERG, HERG1, Kv11...

Homologs of the KCNH2 gene

The KCNH2 gene is conserved in chimpanzee,

Step2: Get homologs in PDB and have alignments with 3D modes.

(*http://sysimm100.protein.osaka-u.ac.jp/sfas/*)

PDBj

WPI Osaka University iFReC

Sequence to Function Annotation Server

Please enter your query

Name: hERG(550-67

Seqeunce:

```
>hERG | Q12809 | residues 550-671
LFLLMCTFALIAHWLACIWYAIGNMEQPHMDSRIGWLHNLGDQIGKPYNSSGLGGPSIKDKY
VTALYFTFSSLTSVGFGNVSPNTNSEKIFSICVMLIGSLMYASIFGNVSAIIQRLYSGTA
```

Or

Upload a FASTA-formatted sequence file: ファイルを選択 ファイルが選...ていません

Select Alignment methods

☑ Blast	◉ Whole PDB	○ Rep. Domains
☑ PsiBlast	◉ Whole PDB	○ Rep. Domains
☑ HHpred	◉ PDB + SCOP	

Send results to this email address

Input your e-mail address

Submit Clear Form

Send feedback About SFAS

hERG Results

Method	Template		E-value	Coverage	Alignment		Model
HHpred	2r9rB	jV	1.10e-19	75	Text	Jalview	Spanner
HHpred	1orqC	jV	9.70e-20	46	Text	Jalview	Spanner
HHpred	3behA	jV	2.20e-19	75	Text	Jalview	Spanner
HHpred	1xl4A	jV	8.20e-17	47	Text	Jalview	Spanner
HHpred	2a9hA	jV	3.70e-17	49	Text	Jalview	Spanner
HHpred	2ih3C	jV	4.10e-16	48	Text	Jalview	Spanner
HHpred	2q67A	jV	2.40e-16	67	Text	Jalview	Spanner
HHpred	2qksA	jV	5.60e-13	45	Text	Jalview	Spanner
Psiblast	1lnqG	jV	1.74e-08	40	Text	Jalview	Spanner
Psiblast	1lnqH	jV	1.74e-08	40	Text	Jalview	Spanner
Psiblast	1lnqB	jV	1.74e-08	40	Text	Jalview	Spanner
Psiblast	1lnqA	jV	1.74e-08	40	Text	Jalview	Spanner
Psiblast	1lnqF	jV	1.74e-08	40	Text	Jalview	Spanner
Psiblast	1lnqE	jV	1.74e-08	40	Text	Jalview	Spanner
Psiblast	1lnqC	jV	1.74e-08	40	Text	Jalview	Spanner
Psiblast	1lnqD	jV	1.74e-08	40	Text	Jalview	Spanner
Psiblast	1orqC	jV	1.67e-07	42	Text	Jalview	Spanner
Psiblast	2a0lB	jV	4.19e-07	41	Text	Jalview	Spanner
Blast	1ujlA	jV	8.03e-20	33	Text	Jalview	Spanner
Blast	2q6aB	jV	7.93e-04	29	Text	Jalview	Spanner
Blast	2q67B	jV	7.48e-04	29	Text	Jalview	Spanner
Blast	2q67A	jV	7.24e-04	29	Text	Jalview	Spanner
Blast	2ahzB	jV	8.62e-04	29	Text	Jalview	Spanner
Blast	3e89B	jV	1.03e-03	31	Text	Jalview	Spanner
Blast	3e8hB	jV	1.03e-03	31	Text	Jalview	Spanner
Blast	2q6aA	jV	8.13e-04	29	Text	Jalview	Spanner
Blast	2ahyA	jV	8.84e-04	29	Text	Jalview	Spanner
Blast	3e83B	jV	1.03e-03	31	Text	Jalview	Spanner

(*http://sysimm100.protein.osaka-u.ac.jp/tmp/SFAS16483/hERG_top.html*)

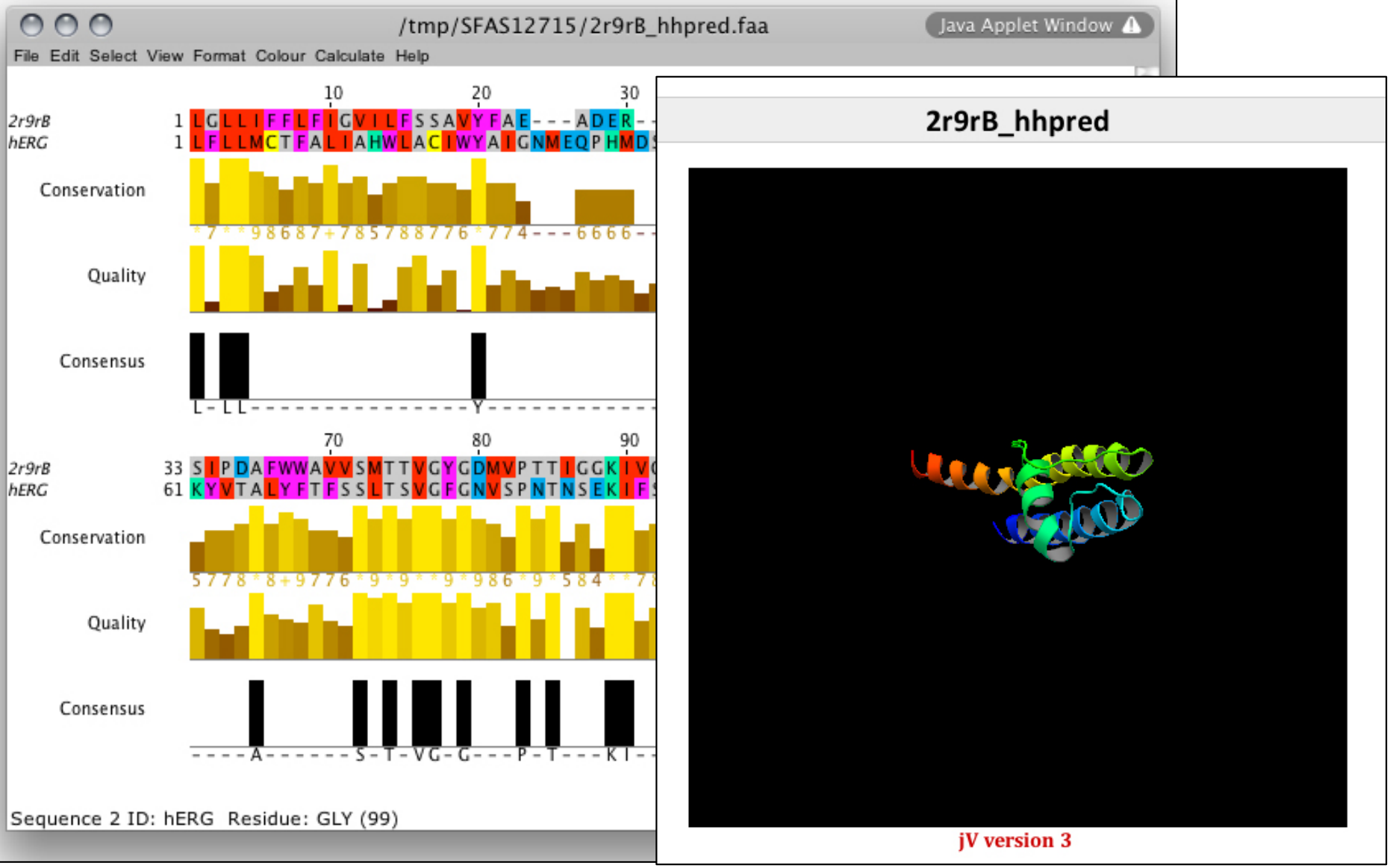
Step2: Result of SFAS: The best template is 2r9rB

Jalview

Start Jalview

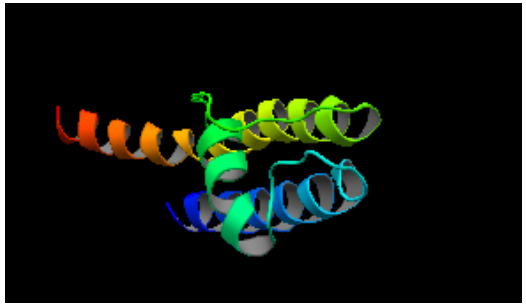
Return to SFAS Results

```
>2r9rB
LGLLIFFLFIGVILFSSAVYFAE---ADER-------------------------
DSQFPSIPDAFWWAVVSMTTVGYGDMVPTTIGGKIVGSLCAIAGVLTIALPVPVI
VSNFNYFYHRET
>hERG
LFLLMCTFALIAHWLACIWYAIGNMEQPHMDSRIGWLHNLGDQIGKPYNSSGLGG
PSIKDKYVTALYFTFSSLTSVGFGNVSPNTNSEKIFSICVMLIGSLMYASIFGNV
SAIIQRLYSGTA
```

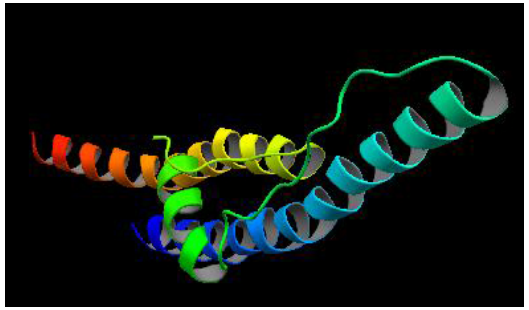


jV version 3

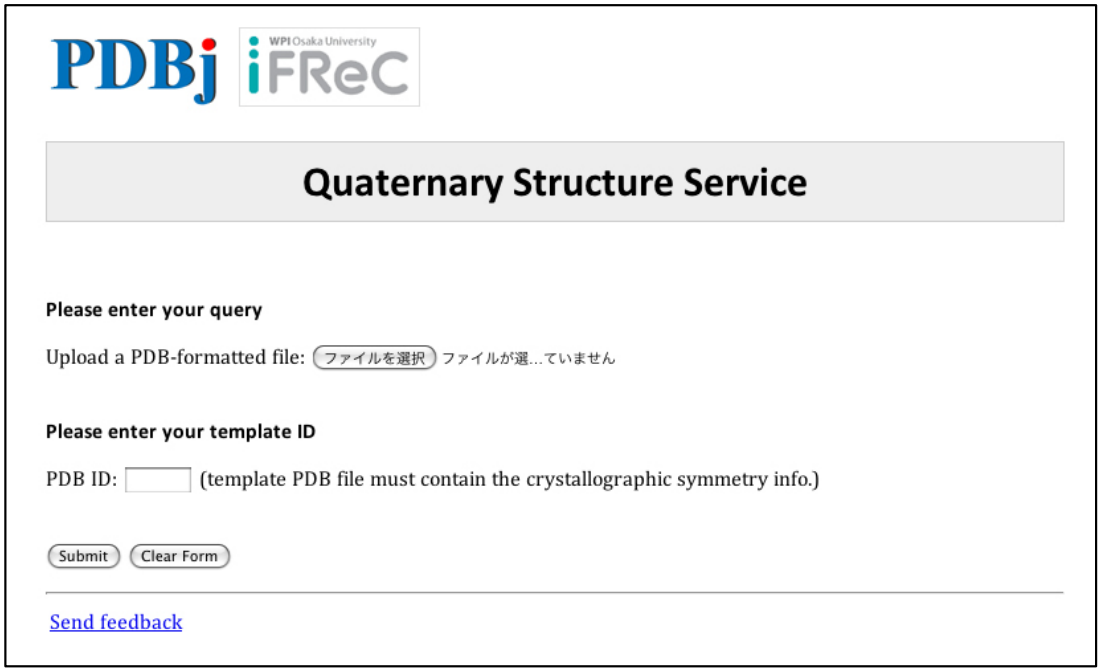
Step3: Make the tetramer structure



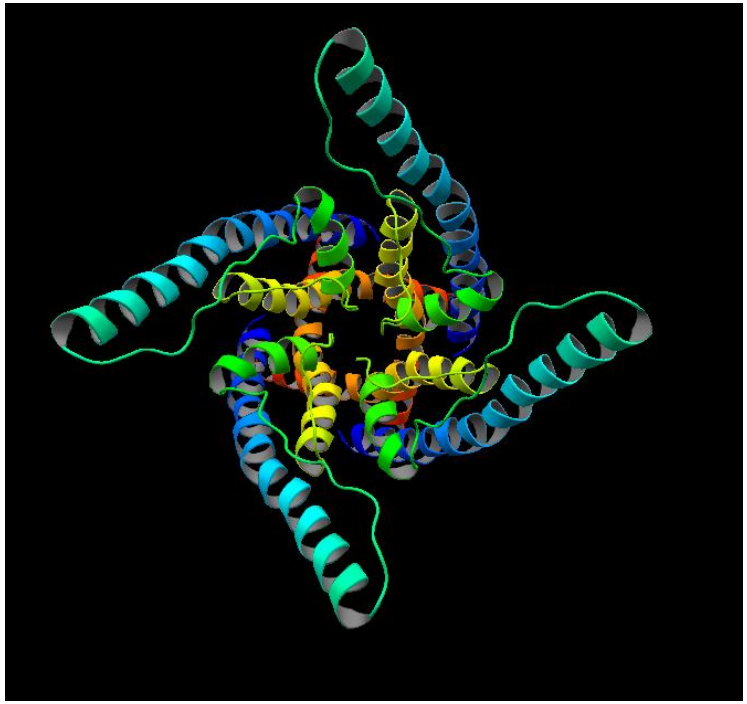
Template: 2r9rB
(550-670)



Model structure
(monomer)



(http://sysimm100.protein.osaka-u.ac.jp/pdb_quat/)



Model structure
(tetramer)

Other options

Step O1. Search PDB structures with similar amino acid sequences.

Access to Sequence Navigator (http://www.pdbj.org/), and get the PDBID information for the homologs.

Step O2. Collect the strutural information of the homologs.

Access to PDBj (*http://www.pdbj.org/*), and get the structural information for the above PDBIDs.

Step O3. Make the sequences-structures multiple alignment.

Access to MAFFTash (*http://www.pdbj.org/*), and input the PDBIDs and the sequence of the hERG channel, to get the multiple alignment.

Step O4. Make the homology model with the alignment.

Access to Spanner (*http://www.pdbj.org/spanner/*), and input the template structure with the above PDBID and the alignment information to get the homology model.

Step O1: Search for a Homolog protein in PDB

Use of Sequence Navigator at PDBj (http://www.pdbj.org/)

PDBj

English

Home

Data Deposition >>
- ADIT: PDB Deposition
- ADIT-NMR

Search >>
- Search PDB (xPSSS)
- Latest Released Search
- Sequence-Navigator
- Structure-Navigator
- SeSAW
- Ligand Binding Sites (GIRAF)
- EM Navigator
- Search NMR Data (BMRB)
- Status Search

Service and Software >>
- Protein Globe
- ASH
- MAFFTash
- JV: Graphic Viewer

Derived database >>
- eF-site/eF-seek/eF-surf
- eProtS
- ProMode
- Molecule of the Month

Download >>
- PDB Archive/Snapshot Archive

About Remediation Data

Links

PDBj (Protein Data Bank Japan) maintains a centralized archive of macromolecul
in collaboration with the RCSB in USA and the PDBe in EU. PDBj is supported by JS

Sequence Navigator

About Sequence Navigator

Sequence Navigator Soap Service

To Enter Navigator, Input a PDB ID [ ] **and Chain ID** [ ]

OR Input an AA Sequence

```
>Influenza A virus (A/Texas/04/2009(H1N1)) segment 6 neuraminidase(NA)
gene, complete cds. FJ981614
MNPNQKIITIGSVCMTIGMANLILQIGNIISIWISHSIQLGNQNQIETCNQSVITYENNTWV
NQTYVNISNTNFAAGQSVVSVKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVFVIREPFIS
CSPLECRTFFLTQGALLNDKHSNGTIKDRSPYRTLMSCPIGEVPSPYNSRFESVAWSASAC
```

Clustering Options
- (•) No Clustering
- ( ) Cluster by E-value $10^{\wedge}$ [ ]
- (+) (more clusters)
- (-) (fewer clusters)

Find All Homologs

Clear Form

Terms of Use | Newsletter | About Us

PDBj

English

Home

Data Deposition >>
- ADIT: PDB Deposition
- ADIT-NMR

Search >>
- Search PDB (xPSSS)
- Latest Released Search
- Sequence-Navigator
- Structure-Navigator
- SeSAW
- Ligand Binding Sites (GIRAF)
- EM Navigator
- Search NMR Data (BMRB)
- Status Search

Service and Software >>
- Protein Globe
- ASH
- MAFFTash
- JV: Graphic Viewer

Derived database >>
- eF-site/ eF-seek/ eF-surf
- eProtS
- ProMode
- Molecule of the Month

Download >>

PDBj (Protein Data Bank Japan) maintains a centralized archive of macromolecular structures
in collaboration with the RCSB in USA and the PDBe in EU. PDBj is supported by JST-BIRD.

Sequence Navigator

Query Sequence:
[MNPNQKIITIGSVCMTIGMANLILQIGNIISIWISHSIQLGNQNQIETCNQSVITYENNTW
Clustering Option: [No Clustering]

Re-Clustering Options
- (•) No Clustering
- ( ) Cluster by E-value $10^{\wedge}$ [ ]
- (+) (more clusters)
- (-) (fewer clusters)

Find All Homologs

Results (1-33) / 33

Seq. Identity: 91% Seq. Positives: 97% E-value: 0.0 Score: 776 Compound: NEURAMINIDASE

New Search [2HTYC]

```
Query  83 385 VKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVFVIREPFISCSPLECRTFFLTQGALLNDKHSNG
2HTYC   1 385 VKLAGNSSLCPINGWAVYSKDNSIRIGSKGDVFVIREPFISCSHLECRTFFLTQGALLNDKHSNG
```

2HTYC Exact Matches: 2HU0F 2HU4C 2HTYG 2HU4G 2HU0B 2HU4A 2HTYA 2HU

Seq. Identity: 91% Seq. Positives: 97% E-value: 0.0 Score: 774 Compound: NEURAMINIDASE

New Search [3CL2B]

```
Query  83 385 VKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVFVIREPFISCSPLECRTFFLTQGALLNDKHSNG
3CL2B   1 385 VKLAGNSSLCPINGWAVYSKDNSIRIGSKGDVFVIREPFISCSHLECRTFFLTQGALLNDKHSNG
```

3CL2B Exact Matches: 3CL2G 3CL2F 3CL2C 3CL2H 3CL2A 3CL2D 3CL2E

Seq. Identity: 91% Seq. Positives: 97% E-value: 0.0 Score: 774 Compound: NEURAMINIDASE

New Search [3CL0A]

```
Query  83 385 VKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVFVIREPFISCSPLECRTFFLTQGALLNDKHSNG
3CL0A   1 385 VKLAGNSSLCPINGWAVYSKDNSIRIGSKGDVFVIREPFISCSHLECRTFFLTQGALLNDKHSNG
```

3CL0A Exact Matches: 3CKZA

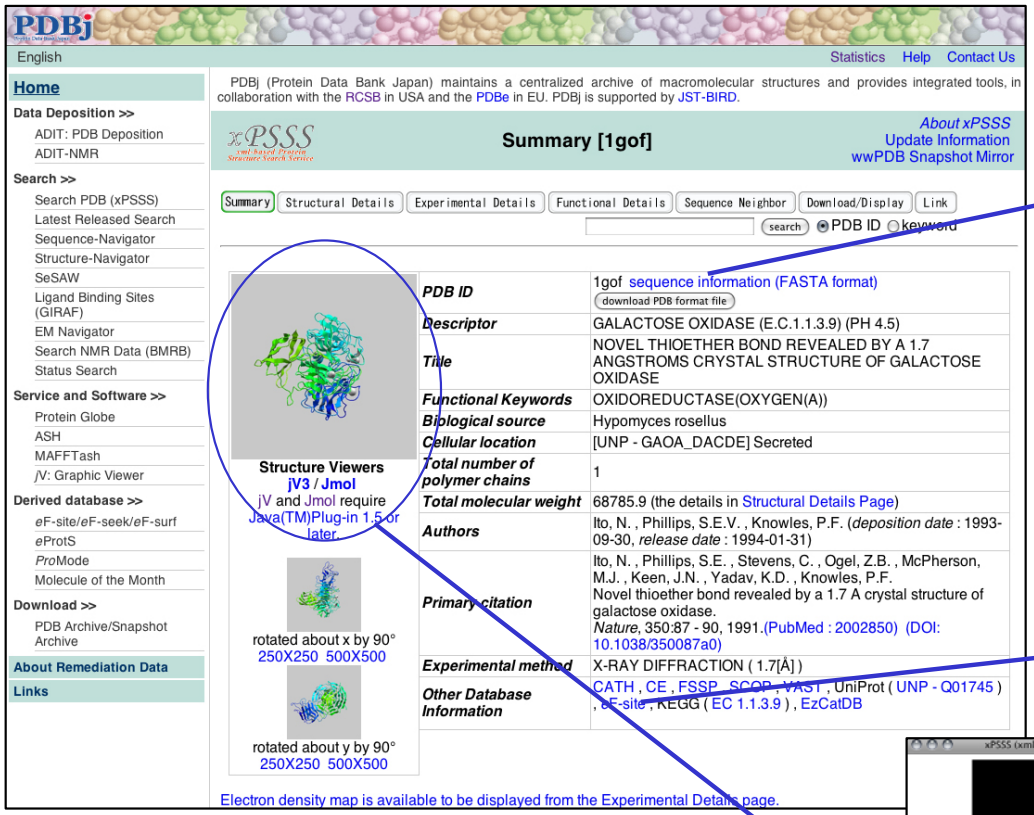
Seq. Identity: 88% Seq. Positives: 95% E-value: 0.0 Score: 763 Compound: NEURAMINIDASE

New Search [3CYEA]

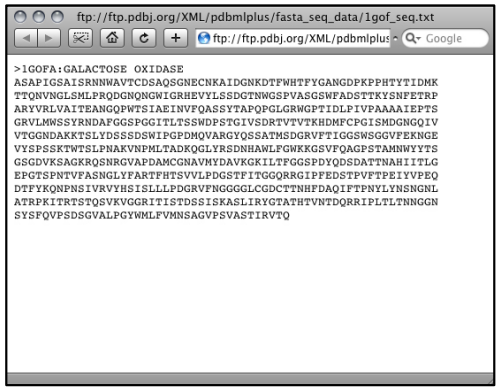
```
Query  83 385 VKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVFVIREPFISCSPLECRTFFLTQGALLNDKHSNG
3CYEA   1 385 VILTGNSSLCPISGWAIYSKDNGIRIGSKGDVFVIREPFISCSHLECRTFFLTQGALLNDKHSNG
```

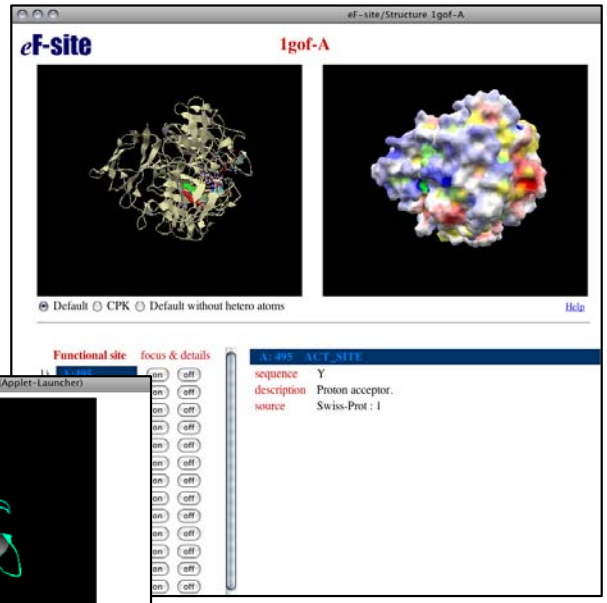
Step O2: Get Entry Data from PDBj

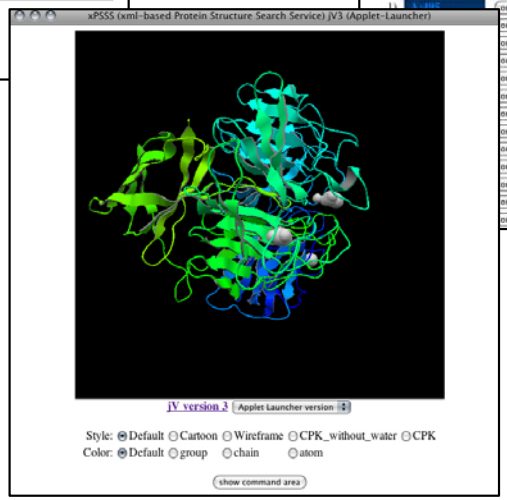
(*http://www.pdbj.org/*)



Summary for each PDBID







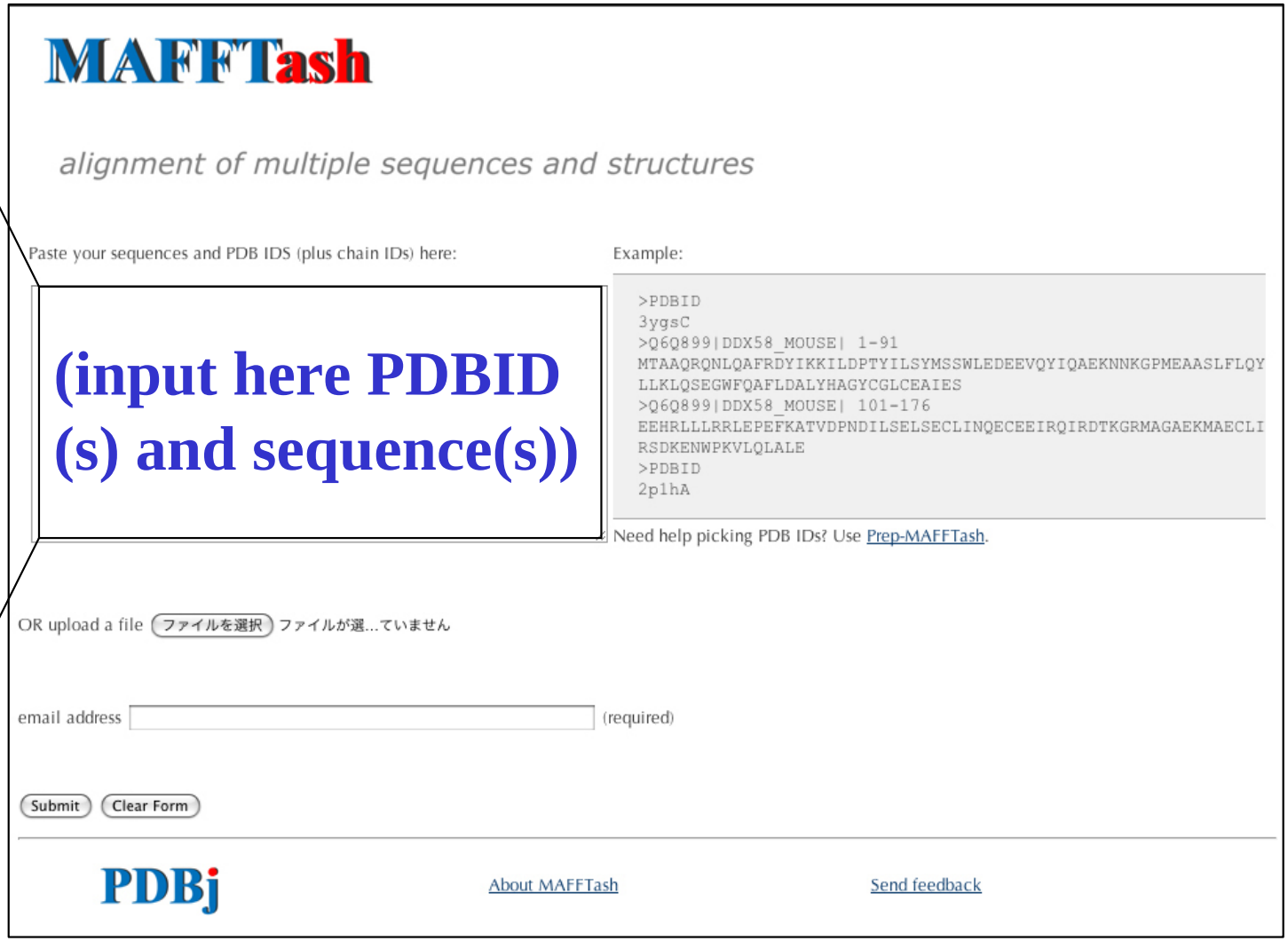
Graphic viewer: *jV*

http://www.pdbj.org/jV/

Step O3: Alignment of Sequences and Structures(i)

(*http://sysimm100.protein.osaka-u.ac.jp/MAFFTash.3/*)

```
>PDBID
2r9rB
>PDBID
1orqC
>Q12809|hERG;550-671
LFLLMCTFALIAHWLACIWY
AIGNMEQPHMDSRIGWLHNL
GDQIGKPYNSSGLGGPSIKD
KYVTALYFTFSSLTSVGFGN
VSPNTNSEKIFSICVMLIGS
LMYASIFGNVSAIIQRLYSG
TA
```



Step O3: Alignment of Sequences and Structures(ii)

(*http://sysimm100.protein.osaka-u.ac.jp/MAFFTash.3/*)

Use "Prep-MAFFTash" for automatic search of sequences and structures.

MAFFTash

alignment of multiple sequences and structures

Paste your sequences below. Missing PDB IDs will be automatically added.

```
>hERG | Q12809 | residues 550-671
LFLLMCTFALIAHWLACIWYAIGNMEQPHMDSRIGWLHNLGDQIGKPYNSSGLGGPSIKDKY
VTALYFTFSSLTSVGFGNVSPNTNSEKIFSICVMLIGSLMYASIFGNVSAIIQRLYSGTA
```

Example:

```
>PDBID
1fanA
>BPTI
RPDFCLEPPYTGPCKARIIRYFYNAKAGLCQTFVYGGCRAKRGNFKSAEDCMRTCG
>Alpha Dendrotoxin
EPRRKLCILHRNPGRCYDKIPAFYYNQKKKQCERFDWSGCGGNSNRFKTIEECRRTCIG
```

Note: You can add PDB IDs (plus chain IDs) too, if you want.

OR upload a file [ファイルを選択] ファイルが選...ていません

email address harukin@protein.osaka-u.ac.jp (required)

Add structures

☐ Use Blast to add structures (from PDB)

Max seq ID between added structures 90

Min seq ID from original input 20

☐ Add ASH structural neighbors

Max seq ID between added structures 90

Min seq ID from original input 0

Add sequences

☑ Use Blast to add sequences (from UniRef100)

Max seq ID between added sequences 90

Min seq ID from original input 20

[Submit] [Clear Form]

Step O4: Homology modeling with the alignment

(*http://www.pdbj.org/spanner/*)

Spanner

Spanner is a structural homology modeling program—that is, it threads a specific amino-acid sequence onto a specific PDB structure, patching up the gaps as best it can.

To create a model, you must provide a template structure, as well as an alignment of the sequence you wish to model onto the template sequence. Spanner will replace matching residues, fill any gaps caused by inserted or deleted residues, and thermodynamically optimize the resulting structure.

The resulting PDB, as well as a log file, will be emailed to you when the modeling task finishes. If an error prevented a homology model from being generated (for example, when the alignment you provided does not match the template structure), the log file will explain which part of the modeling sequence failed.

Template PDB structure (PDB format):
ファイルを選択 ファイルが選…ていません

Sequence alignment (FASTA format; first sequence is the template, second sequence is the query):
ファイルを選択 ファイルが選…ていません

Model: (not necessary if PDB file contains only one model)

Chain: (not necessary if PDB file contains only one chain)

Email address for results:

START!