

立体構造の可視化と立体構造予測

平成20年7月

川端 猛

1. 演習の準備

まず、次のコマンドを入力して、演習に必要なファイルをコピーして、演習用のディレクトリ **3DSTR** に移動してください。

```
cd  
cp -r /mandara/lecture/takawaba/3DSTR .  
cd 3DSTR
```

2. 立体構造データの取得

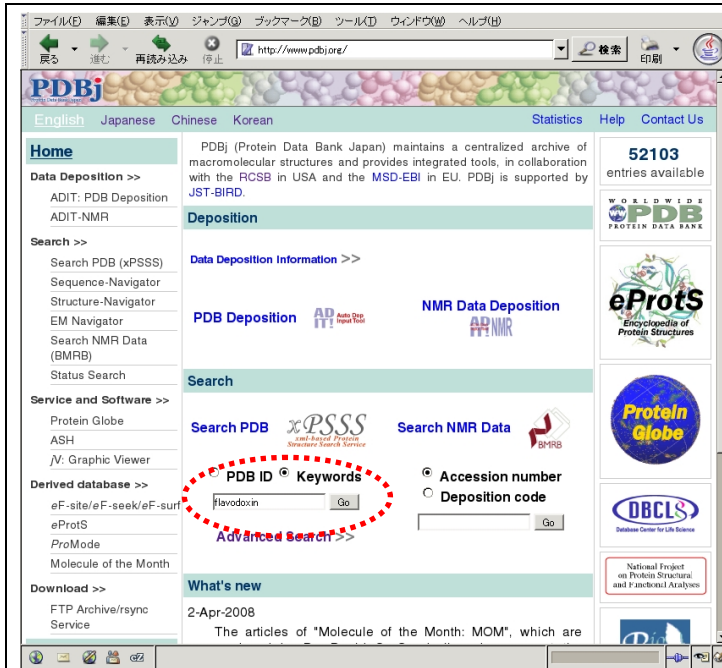
蛋白質の立体構造データの登録収集と配布業務は、現在 Worldwide Protein Data Bank (wwPDB <http://www.wwpdb.org>) という組織で行っており、配列データと同じように米国とヨーロッパと日本がそれぞれ拠点を持ち、協力して運営する形になっています。

データベース名	WWW アドレス	運営拠点
米国：RCSB-PDB	http://www.rcsb.org	Rutgers State University/San Diego Supercomputer Center
欧州：MSD-EBI	http://www.ebi.ac.uk/msd	英国 EBI
日本：PDBj	http://www.pdbj.org	大阪大学

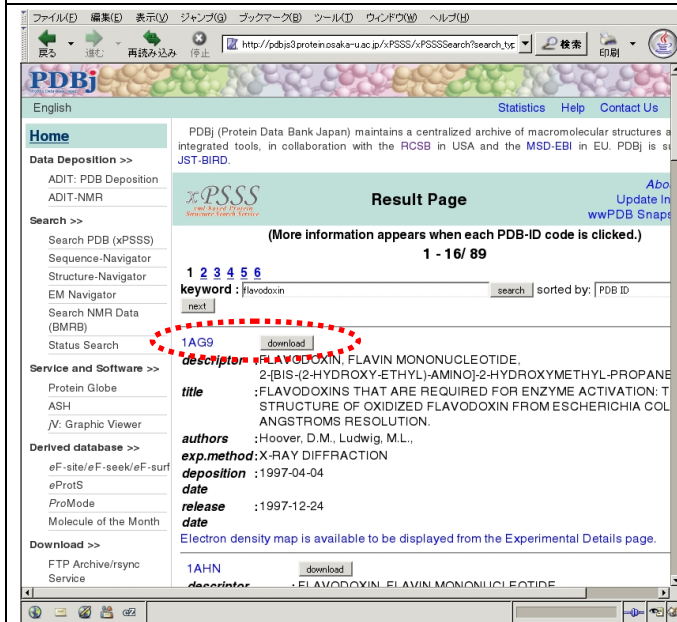
2.1 PDBj の WEB サーバを使用する場合

PDBj の WEB サイトにアクセスして、PDB 形式の立体構造データをダウンロードしてみましよう。

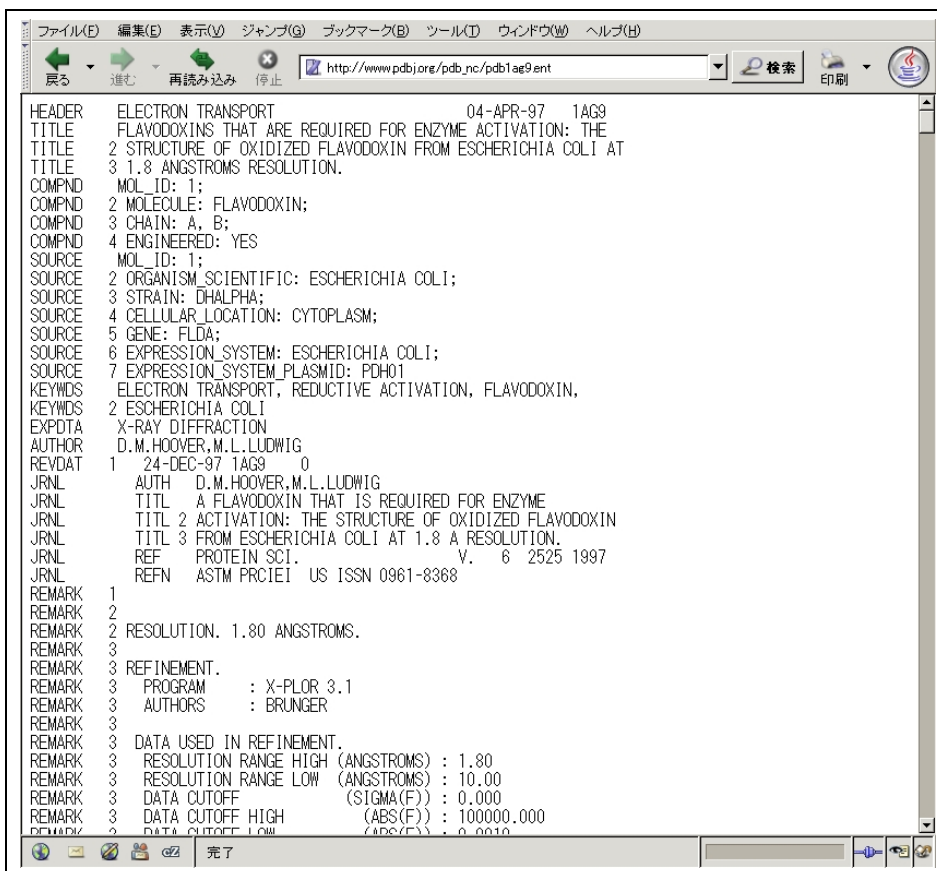
(1) PDBj の WEB ページ <http://www.pdbj.org> にアクセスする。



(2) フォームに適切なキーワード(例えば flavodoxin)を入力して、**Go** をクリックする。



(3) ヒットした PDB データの一つの **download** をクリックする。



ファイル(E) 編集(E) 表示(V) ジャンプ(G) ブックマーク(B) ツール(T) ウィンドウ(W) ヘルプ(H)

戻る 進む 再読み込み 停止 http://www.pdbj.org/pdb_nc/pdb1ag9.ent 検索 印刷

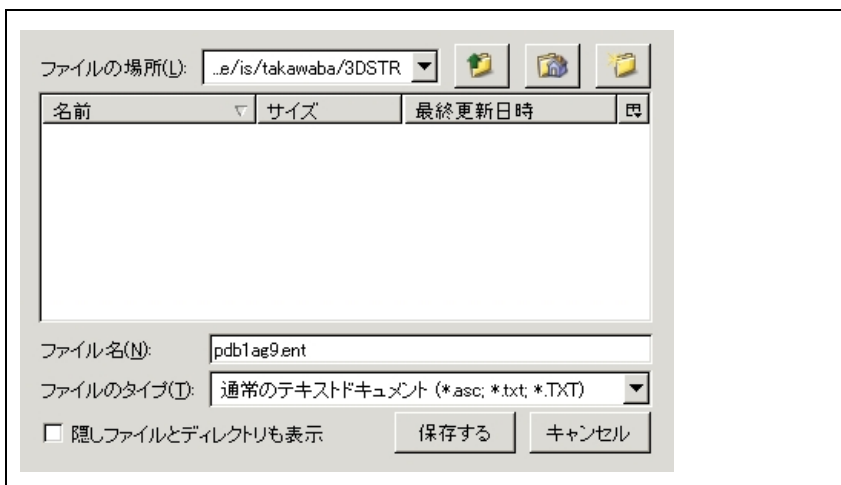
```

HEADER    ELECTRON TRANSPORT                                04-APR-97  1AG9
TITLE     FLAVODOXINS THAT ARE REQUIRED FOR ENZYME ACTIVATION: THE
TITLE     2 STRUCTURE OF OXIDIZED FLAVODOXIN FROM ESCHERICHIA COLI AT
TITLE     3 1.8 ANGSTROMS RESOLUTION.
COMPND    MOL_ID: 1;
COMPND    2 MOLECULE: FLAVODOXIN;
COMPND    3 CHAIN: A, B;
COMPND    4 ENGINEERED: YES
SOURCE    MOL_ID: 1;
SOURCE    2 ORGANISM_SCIENTIFIC: ESCHERICHIA COLI;
SOURCE    3 STRAIN: DHALPHA;
SOURCE    4 CELLULAR_LOCATION: CYTOPLASM;
SOURCE    5 GENE: FLDA;
SOURCE    6 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE    7 EXPRESSION_SYSTEM_PLASMID: PDH01
KEYWDS    ELECTRON TRANSPORT, REDUCTIVE ACTIVATION, FLAVODOXIN,
KEYWDS    2 ESCHERICHIA COLI
EXPDTA    X-RAY DIFFRACTION
AUTHOR    D.M.HOOVER,M.L.LUDWIG
REVDAT    1 24-DEC-97 1AG9 0
JRNL      AUTH  D.M.HOOVER,M.L.LUDWIG
JRNL      TITL  A FLAVODOXIN THAT IS REQUIRED FOR ENZYME
JRNL      TITL  2 ACTIVATION: THE STRUCTURE OF OXIDIZED FLAVODOXIN
JRNL      TITL  3 FROM ESCHERICHIA COLI AT 1.8 A RESOLUTION.
JRNL      REF   PROTEIN SCI.                                V.   6 2525 1997
JRNL      REFN  ASTM PRClEI  US ISSN 0961-8368
REMARK    1
REMARK    2
REMARK    3 RESOLUTION. 1.80 ANGSTROMS.
REMARK    3
REMARK    3 REFINEMENT.
REMARK    3 PROGRAM      : X-PLOR 3.1
REMARK    3 AUTHORS      : BRUNGER
REMARK    3
REMARK    3 DATA USED IN REFINEMENT.
REMARK    3 RESOLUTION RANGE HIGH (ANGSTROMS) : 1.80
REMARK    3 RESOLUTION RANGE LOW  (ANGSTROMS) : 10.00
REMARK    3 DATA CUTOFF          (SIGMA(F)) : 0.000
REMARK    3 DATA CUTOFF HIGH      (ABS(F)) : 100000.000
REMARK    3 DATA CUTOFF LOW       (ABS(F)) : 0.0010

```

完了

(4) PDB 形式のファイルがブラウザ内に表示される。メニューの **ファイル(F)** から **ページに名前を付けて保存(A)** を選ぶ。



ファイルの場所(L): ..e/is/takawaba/3DSTR

名前	サイズ	最終更新日時

ファイル名(N): pdb1ag9.ent

ファイルのタイプ(T): 通常のテキストドキュメント (*.asc;*.txt;*.TXT)

隠しファイルとディレクトリも表示

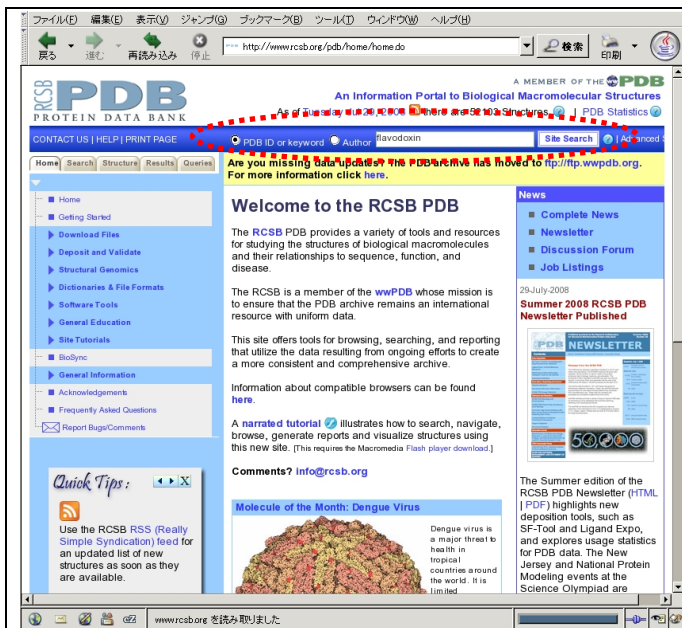
保存する キャンセル

(5) ファイル保存のウィンドウが表示されるので、適当なディレクトリ、ファイル名を指定して、**保存する** をクリックします。

2.2 RCSB-PDB の WEB サーバを使用する場合

RCSB-PDB の WEB サイトにアクセスして、PDB 形式の立体構造データをダウンロードしてみましょう。

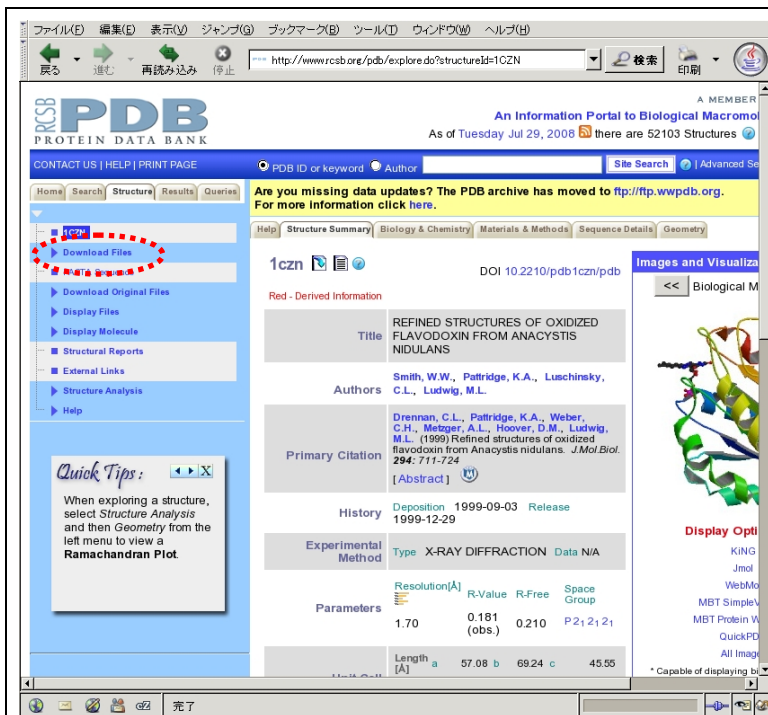
(2) RCSB-PDB の WEB ページ <http://www.rcsb.org> にアクセスする。



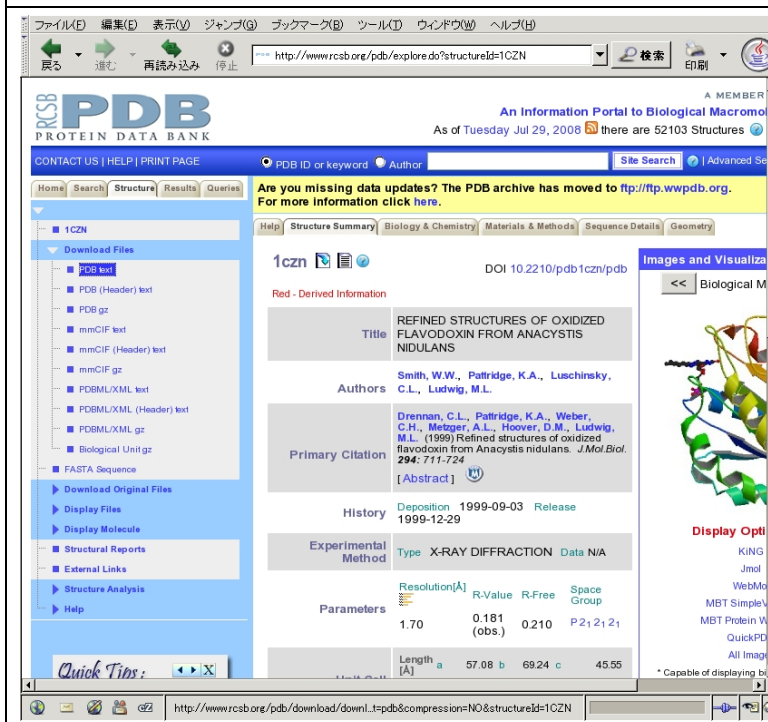
(2) フォームに適当なキーワード(例えば flavodoxin)を入力して、**Site Search** をクリックする。



(3) ヒットしたエントリの一つをダブルクリックする。



(4) 画面左の
Download Files のメニュー
 を選ぶ。



(5) **PDB text** のボタンを
 クリックする。

このサイトは「1CZNPdb」を添付として処理するよう動いています。これはタイプ application/download であり、次の場所にあります:
<http://www.rcsb.org>

このファイルの処理方法を Web ブラウザ に指定してください。

次のアプリケーションで開く

ディスクに保存

このタイプのファイルを処理する際にも常にこのアクションを実行する

(5) ポップアップウィンドウが表示される。ディスクに保存を選択し、されるので、メニューの **OK** をクリックする。

ファイルの場所(L):

名前	サイズ	最終更新日時	開
1a0a.pdb	166698	2007年07月30日 09:26:29	▲
1aa9.pdb	236682	2007年07月30日 09:26:29	
1be2.pdb	233685	2007年07月30日 09:26:29	
1fdn.pdb	44550	2007年07月30日 09:26:29	
1flv.pdb	127089	2007年07月30日 09:26:27	
1flv_orig.pdb	118665	2007年07月30日 09:26:29	
1j77.pdb	300186	2007年07月30日 09:26:29	▼

ファイル名(N):

ファイルのタイプ(T):

隠しファイルとディレクトリも表示

(6) ファイル保存のウィンドウが表示されるので、適当なディレクトリ、ファイル名を指定して、**保存する** をクリックします。

データがダウンロードできたら、lessなどでファイルの内容を確認してください。

less [PDB ファイル名]

PDBフォーマットの例: 1flv (FLAVODOXIN)											
HEADER	ELECTRON TRANSPORT			02-JUL-92	1FLV			1FLV	2		
COMPND	FLAVODOXIN							1FLV	3		
SOURCE	(ANABAENA 7120)							1FLV	4		
AUTHOR	S.T.RAO,M.SUNDARALINGAM							1FLV	5		
REVDAT	1 31-OCT-93 1FLV 0							1FLV	6		
JENL	AUTH S.T.RAO,F.SHAFFIE,C.YU,K.A.SATYSHUR,B.J.STOCKMAN,							1FLV	7		
JENL	AUTH 2 J.L.MARKLEY,M.SUNDARALINGAM							1FLV	8		
JENL	TITL STRUCTURE OF THE OXIDIZED LONG CHAIN FLAVODOXIN							1FLV	9		
JENL	TITL 2 FROM ANABAENA 7120 AT 2 ANGSTROMS RESOLUTION							1FLV	10		
JENL	REF PROTEIN SCI.			V.	1	1413	1993	1FLV	11		
JENL	REFN US ISSN 0961-8368						795	1FLV	12		
REMARK	1							1FLV	13		
REMARK	2							1FLV	14		
REMARK	2	RESOLUTION. 2.0	ANGSTROMS.					1FLV	15		
REMARK	3							1FLV	16		
REMARK	3	REFINEMENT.						1FLV	17		
REMARK	3	PROGRAM	X-PLOR					1FLV	18		
REMARK	3	AUTHORS	BRUNGER					1FLV	19		
REMARK	3	R VALUE	0.185					1FLV	20		
REMARK	3	RMSD BOND DISTANCES	0.02	ANGSTROMS				1FLV	21		
REMARK	3	RMSD BOND ANGLES	3.	DEGREES				1FLV	22		
REMARK	4							1FLV	23		
REMARK	4	SHEETS S1 AND S2 REPRESENT ONE BIFURCATED SHEET AND DIFFER						1FLV	24		
REMARK	4	ONLY IN THE FIFTH STRAND.						1FLV	25		
REMARK	5							1FLV	26		
REMARK	5	NO ELECTRON DENSITY VISIBLE FOR N-TERMINAL RESIDUE.						1FLV	27		
REMARK	6	THE MODEL IS UNCERTAIN FOR RESIDUES 28 - 30 DUE TO POOR						1FLV	28		
REMARK	6	ELECTRON DENSITY.						1FLV	29		
SEQRES	1	169	SER LYS LYS ILE GLY LEU PHE TYR GLY THR GLN THR GLY					1FLV	30		
SEQRES	2	169	LYS THR GLU SER VAL ALA GLU ILE ILE ARG ASP GLU PHE					1FLV	31		
:											
SEQRES	13	169	LYS SER TRP VAL ALA GLN LEU LYS SER GLU PHE GLY LEU					1FLV	42		
HET	FMN	170	31	FLAVIN MONONUCLEOTIDE PROSTHETIC GROUP				1FLV	43		
FORMUL	2	FMN	C17 H21 N4 O9 P1					1FLV	44		
FORMUL	3	HQH	*104 (H2 O1)					1FLV	45		
HELIX	1	HL	GLY 13 PHE 26 1					1FLV	46		
:											
SHEET	1	S1	5 VAL 30 VAL 36 0					1FLV	50		
:											
CRYST1	48.000	32.000	51.600	90.00	92.00	90.00	P 21	2	1FLV	69	
ORIGX1	1.000000	0.000000	0.000000			0.000000			1FLV	70	
ORIGX2	0.000000	1.000000	0.000000			0.000000			1FLV	71	
ORIGX3	0.000000	0.000000	1.000000			0.000000			1FLV	72	
SCALE1	0.020833	0.000000	0.000728			0.000000			1FLV	73	
SCALE2	0.000000	0.031250	0.000000			0.000000			1FLV	74	
SCALE3	0.000000	0.000000	0.019392			0.000000			1FLV	75	
ATOM	1	N	LYS	2	-1.899	0.013	24.235	1.00	29.59	1FLV	76
ATOM	2	CA	LYS	2	-2.131	-0.941	25.290	1.00	29.27	1FLV	77
ATOM	3	C	LYS	2	-0.935	-1.842	25.037	1.00	27.53	1FLV	78
ATOM	4	O	LYS	2	0.169	-1.317	25.134	1.00	28.86	1FLV	79
ATOM	5	CB	LYS	2	-3.466	-1.690	25.099	1.00	30.28	1FLV	80
ATOM	6	CG	LYS	2	-4.224	-1.938	26.401	1.00	31.53	1FLV	81
ATOM	7	CD	LYS	2	-3.725	-3.084	27.272	1.00	32.62	1FLV	82
ATOM	8	CE	LYS	2	-4.673	-3.157	28.462	1.00	33.71	1FLV	83
ATOM	9	NZ	LYS	2	-4.785	-4.513	28.969	1.00	35.18	1FLV	84
ATOM	10	N	LYS	3	-1.052	-3.047	24.506	1.00	24.18	1FLV	85
ATOM	11	CA	LYS	3	0.093	-3.911	24.498	1.00	20.53	1FLV	86
ATOM	12	C	LYS	3	1.170	-3.726	23.433	1.00	17.82	1FLV	87
ATOM	13	O	LYS	3	2.111	-4.513	23.486	1.00	16.57	1FLV	88
ATOM	14	CB	LYS	3	-0.442	-5.340	24.505	1.00	20.54	1FLV	89
ATOM	15	CG	LYS	3	-1.494	-5.693	23.496	1.00	21.30	1FLV	90
ATOM	16	CD	LYS	3	-1.769	-7.176	23.543	1.00	22.20	1FLV	91
ATOM	17	CE	LYS	3	-2.562	-7.543	24.784	1.00	23.14	1FLV	92
ATOM	18	NZ	LYS	3	-2.902	-8.951	24.728	1.00	23.75	1FLV	93
ATOM	19	N	ILE	4	1.167	-2.763	22.496	1.00	14.53	1FLV	94
ATOM	20	CA	ILE	4	2.297	-2.675	21.552	1.00	11.64	1FLV	95
ATOM	21	C	ILE	4	2.957	-1.303	21.705	1.00	10.70	1FLV	96
ATOM	22	O	ILE	4	2.262	-0.265	21.724	1.00	10.53	1FLV	97
ATOM	23	CB	ILE	4	1.840	-2.891	20.046	1.00	11.34	1FLV	98
ATOM	24	CG1	ILE	4	1.250	-4.302	19.810	1.00	10.61	1FLV	99
ATOM	25	CG2	ILE	4	3.067	-2.727	19.135	1.00	10.50	1FLV	100
ATOM	26	CD1	ILE	4	0.658	-4.448	18.407	1.00	9.31	1FLV	101

ヘッダー、
著者、
雑誌名など
一般的な情報

配列情報

蛋白質の
原子座標

```

ATOM      20  CA  ILE      4      2.297 -2.675 21.552  1.00 11.64      1FLV  95
ATOM      21  C   ILE      4      2.957 -1.303 21.705  1.00 10.70      1FLV  96
ATOM      22  O   ILE      4      2.262 -0.265 21.724  1.00 10.53      1FLV  97
ATOM      23  CB  ILE      4      1.840 -2.891 20.046  1.00 11.34      1FLV  98
ATOM      24  CG1 ILE      4      1.250 -4.302 19.810  1.00 10.61      1FLV  99
ATOM      25  CG2 ILE      4      3.067 -2.727 19.135  1.00 10.50      1FLV 100
ATOM      26  CD1 ILE      4      0.658 -4.448 18.407  1.00  9.31      1FLV 101
ATOM      27  N   GLY      5      4.288 -1.250 21.879  1.00  8.56      1FLV 102
ATOM      28  CA  GLY      5      4.937  0.026 22.018  1.00  7.30      1FLV 103
ATOM      29  C   GLY      5      5.648  0.326 20.733  1.00  6.95      1FLV 104
ATOM      30  O   GLY      5      6.488 -0.456 20.298  1.00  6.56      1FLV 105
:
ATOM     1314  N   GLY     168      0.426 -12.492 17.537  1.00 18.11      1FLV1389
ATOM     1315  CA  GLY     168     -0.636 -13.326 18.086  1.00 19.41      1FLV1390
ATOM     1316  C   GLY     168     -1.677 -12.488 18.790  1.00 20.72      1FLV1391
ATOM     1317  O   GLY     168     -1.789 -12.542 20.018  1.00 21.03      1FLV1392
ATOM     1318  N   LEU     169     -2.392 -11.709 18.002  1.00 21.69      1FLV1393
ATOM     1319  CA  LEU     169     -3.443 -10.849 18.515  1.00 23.02      1FLV1394
ATOM     1320  C   LEU     169     -4.744 -11.482 18.031  1.00 23.47      1FLV1395
ATOM     1321  O   LEU     169     -4.910 -11.653 16.810  1.00 23.73      1FLV1396
ATOM     1322  CB  LEU     169     -3.323  -9.417 17.953  1.00 23.23      1FLV1397
ATOM     1323  CG  LEU     169     -1.973  -8.703 18.102  1.00 23.93      1FLV1398
ATOM     1324  CD1 LEU     169     -2.095  -7.313 17.470  1.00 23.75      1FLV1399
ATOM     1325  CD2 LEU     169     -1.557  -8.649 19.577  1.00 23.32      1FLV1400
ATOM     1326  OXT LEU     169     -5.556 -11.819 18.892  1.00 23.98      1FLV1401
TER      1327      LEU     169                                1FLV1402
HETATH  1328  N1  FMN     170      17.131  7.277  2.489  1.00  2.79      1FLV1403
HETATH  1329  C2  FMN     170      17.699  6.054  2.587  1.00  3.35      1FLV1404
HETATH  1330  O2  FMN     170      17.059  5.007  2.520  1.00  4.65      1FLV1405
HETATH  1331  N3  FMN     170      19.079  6.035  2.807  1.00  4.05      1FLV1406
:
HETATH  1357  OP2 FMN     170      12.954 12.006  8.738  1.00  7.50      1FLV1432
HETATH  1358  OP3 FMN     170      10.807 10.790  9.154  1.00  5.09      1FLV1433
HETATH  1359  O   HOH     201      14.870  -8.090 16.970  1.00 11.81      1FLV1434
HETATH  1360  O   HOH     202      10.010  7.800  2.920  1.00 18.92      1FLV1435
HETATH  1361  O   HOH     203      23.170  5.470  -1.310  1.00  2.68      1FLV1436
HETATH  1362  O   HOH     204      13.280 17.710 13.850  1.00  7.21      1FLV1437
HETATH  1363  O   HOH     205         5.820  9.150  -4.210  1.00 10.62      1FLV1438
HETATH  1364  O   HOH     206     -4.590  -4.270  8.280  1.00  7.65      1FLV1439
HETATH  1365  O   HOH     207      28.020  4.240  -3.140  1.00 11.58      1FLV1440
HETATH  1366  O   HOH     208         6.410  -8.820  1.880  1.00 13.18      1FLV1441
HETATH  1367  O   HOH     209      15.470  3.140  7.640  1.00  3.80      1FLV1442
:
HETATH  1461  O   HOH     303      21.300  3.430 29.340  1.00 38.14      1FLV1536
HETATH  1462  O   HOH     304      13.070 14.300  3.760  1.00 21.69      1FLV1537
CONNECT 1328 1329 1345                                1FLV1538
CONNECT 1329 1328 1330 1331                                1FLV1539
CONNECT 1330 1329                                1FLV1540
CONNECT 1331 1329 1332                                1FLV1541
CONNECT 1332 1331 1333 1334                                1FLV1542
CONNECT 1333 1332                                1FLV1543
CONNECT 1334 1332 1335 1345                                1FLV1544
:
CONNECT 1357 1355                                1FLV1567
CONNECT 1358 1355                                1FLV1568
MASTER   17   0   1   4  10   9   0   6 1461   1  31  13  1FLV1569
END                                             1FLV1570

```

蛋白質の
原子座標

FMNの
原子座標

水の
原子座標

共有結合の
接続情報

PDBフォーマットの例: 1a0a COMPLEX(TRANSCRIPTION FACTOR/DNA)

```

HEADER    COMPLEX (TRANSCRIPTION FACTOR/DNA)      27-NOV-97  1A0A
TITLE     CRYSTAL STRUCTURE OF PHO4 BHLH DOMAIN COMPLEXED WITH UASP2
TITLE     2 (17)
COMPND    MOL_ID: 1;
COMPND    2 MOLECULE: PHOSPHATE SYSTEM POSITIVE REGULATORY PROTEIN
COMPND    3 PHO4;
COMPND    4 CHAIN: A, B;
COMPND    5 FRAGMENT: DNA BINDING DOMAIN;
COMPND    6 SYNONYM: BHLH;
COMPND    7 ENGINEERED: YES;
COMPND    8 BIOLOGICAL_UNIT: DIMER;
COMPND    9 MOL_ID: 2;
COMPND    10 MOLECULE: DNA;
COMPND    11 CHAIN: C, D;
COMPND    12 FRAGMENT: UPSTREAM ACTIVATION SITE P2;
COMPND    13 SYNONYM: UASP2 (17);
COMPND    14 ENGINEERED: YES
SOURCE    MOL_ID: 1;
SOURCE    2 ORGANISM_SCIENTIFIC: SACCHAROMYCES CEREVISIAE;
SOURCE    3 ORGANISM_COMMON: BAKER'S YEAST;
SOURCE    4 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE    5 EXPRESSION_SYSTEM_STRAIN: BL21 (DE3);
SOURCE    6 MOL_ID: 2;
SOURCE    7 SYNTHETIC: YES
KEYWDS    TRANSCRIPTION FACTOR, BASIC HELIX LOOP HELIX,
KEYWDS    2 COMPLEX (TRANSCRIPTION FACTOR/DNA)
EXPDTA    X-RAY DIFFRACTION
AUTHOR    T. SHIMIZU, A. TOUMOTO, K. IHARA, M. SHIMIZU, Y. KYOGOKU, N. OGAWA,
AUTHOR    2 Y. OSHIMA, T. HAKOSHIMA
:
SEQRES   1 A  63  MET LYS ARG GLU SER HIS LYS HIS ALA GLU GLN ALA ARG
SEQRES   2 A  63  ARG ASN ARG LEU ALA VAL ALA LEU HIS GLU LEU ALA SER
SEQRES   3 A  63  LEU ILE PRO ALA GLU TRP LYS GLN GLN ASN VAL SER ALA
SEQRES   4 A  63  ALA PRO SER LYS ALA THR THR VAL GLU ALA ALA CYS ARG
SEQRES   5 A  63  TYR ILE ARG HIS LEU GLN GLN ASN GLY SER THR
SEQRES   1 B  63  MET LYS ARG GLU SER HIS LYS HIS ALA GLU GLN ALA ARG
SEQRES   2 B  63  ARG ASN ARG LEU ALA VAL ALA LEU HIS GLU LEU ALA SER
SEQRES   3 B  63  LEU ILE PRO ALA GLU TRP LYS GLN GLN ASN VAL SER ALA
SEQRES   4 B  63  ALA PRO SER LYS ALA THR THR VAL GLU ALA ALA CYS ARG
SEQRES   5 B  63  TYR ILE ARG HIS LEU GLN GLN ASN GLY SER THR
SEQRES   1 C  17  C  T  C  A  C  A  C  C  G  T  C  G  G  C  A
SEQRES   2 C  17  C  T  A  G
SEQRES   1 D  17  C  T  A  G  T  C  C  C  A  C  G  T  G
SEQRES   2 D  17  T  G  A  G
:
ATOM      1  N   MET  A  0      3.430 -2.059 57.593  1.00 14.05
ATOM      2  CA  MET  A  0      4.785 -2.490 57.148  1.00 20.64
ATOM      3  C   MET  A  0      4.821 -2.460 55.629  1.00 21.14
ATOM      4  O   MET  A  0      5.138 -3.464 54.967  1.00 18.32
ATOM      5  CB  MET  A  0      5.095 -3.902 57.652  1.00 20.07
ATOM      6  CG  MET  A  0      6.575 -4.178 57.727  1.00 25.32
ATOM      7  SD  MET  A  0      7.338 -2.826 58.659  1.00 32.78
ATOM      8  CE  MET  A  0      6.868 -3.310 60.397  1.00 31.75
ATOM      9  N   LYS  A  1      4.503 -1.273 55.106  1.00 28.78
ATOM     10  CA  LYS  A  1      4.430 -0.953 53.669  1.00 28.21
ATOM     11  C   LYS  A  1      5.793 -0.890 52.960  1.00 25.31
ATOM     12  O   LYS  A  1      5.940 -0.201 51.946  1.00 19.97
:
ATOM     497  CG2 THR  A  62     30.577 45.665 31.365  1.00 20.01
ATOM     498  OXT THR  A  62     33.796 47.459 32.723  1.00 28.20
TER      499   THR  A  62
ATOM     500  N   MET  B  0      22.256 -8.416 25.512  1.00 39.45
ATOM     501  CA  MET  B  0      21.339 -7.252 25.587  1.00 36.86
ATOM     502  C   MET  B  0      21.971 -6.163 26.439  1.00 34.79
ATOM     503  O   MET  B  0      21.986 -6.267 27.679  1.00 26.40
ATOM     504  CB  MET  B  0      20.001 -7.659 26.195  1.00 39.85
ATOM     505  CG  MET  B  0      19.037 -6.504 26.232  1.00 43.14
ATOM     506  SD  MET  B  0      17.509 -6.771 27.125  1.00 40.85
ATOM     507  CE  MET  B  0      16.294 -6.079 25.909  1.00 47.06
ATOM     508  N   LYS  B  1      22.372 -5.080 25.762  1.00 34.28
ATOM     509  CA  LYS  B  1      22.059 -3.931 26.369  1.00 35.28

```

ヘッダー、
著者、
雑誌名など
全般的な情報

配列情報

蛋白質A鎖の
原子座標

蛋白質B鎖の
原子座標

ATOM	509	CA	LYS B	1	23.059	-3.931	26.369	1.00	35.28	C	蛋白質B鎖の 原子座標
:											
ATOM	994	CB	THR B	62	9.653	51.796	41.377	1.00	20.97	C	
ATOM	995	CG1	THR B	62	10.175	51.057	42.493	1.00	12.90	O	
ATOM	996	CG2	THR B	62	8.457	51.016	40.835	1.00	16.72	C	
ATOM	997	OXT	THR B	62	12.527	52.988	41.599	1.00	19.13	O	
TER	998		THR B	62							
ATOM	999	O5*	C C	1	6.718	3.528	61.794	1.00	25.61	O	
ATOM	1000	C5*	C C	1	5.329	3.838	61.554	1.00	34.46	C	
ATOM	1001	C4*	C C	1	5.200	5.341	61.510	1.00	35.40	C	
ATOM	1002	O4*	C C	1	6.381	5.841	62.177	1.00	39.11	O	
ATOM	1003	C3*	C C	1	5.243	5.922	60.100	1.00	38.67	C	
ATOM	1004	O3*	C C	1	4.517	7.153	59.981	1.00	38.84	O	
ATOM	1005	C2*	C C	1	6.717	6.161	59.855	1.00	37.34	C	
ATOM	1006	Cl*	C C	1	7.308	6.358	61.244	1.00	32.10	C	
ATOM	1007	N1	C C	1	8.552	5.588	61.354	1.00	26.96	N	
ATOM	1008	C2	C C	1	9.601	6.112	62.080	1.00	23.30	C	
ATOM	1009	O2	C C	1	9.429	7.176	62.680	1.00	26.10	O	
ATOM	1010	N3	C C	1	10.776	5.449	62.117	1.00	24.36	N	
ATOM	1011	C4	C C	1	10.905	4.288	61.471	1.00	28.17	C	
ATOM	1012	N4	C C	1	12.086	3.666	61.518	1.00	30.86	N	
ATOM	1013	C5	C C	1	9.830	3.710	60.748	1.00	25.20	C	
ATOM	1014	O6	C C	1	8.683	4.385	60.717	1.00	24.85	O	
ATOM	1015	P	T C	2	4.283	7.786	58.519	1.00	36.24	P	
ATOM	1016	O1P	T C	2	2.972	7.313	57.965	1.00	33.14	O	
ATOM	1017	O2P	T C	2	5.528	7.537	57.753	1.00	34.01	O	
ATOM	1018	O5*	T C	2	4.150	9.352	58.776	1.00	40.74	O	
ATOM	1019	C5*	T C	2	5.144	10.089	59.519	1.00	37.48	C	
:											
ATOM	1341	C2	G C	17	20.174	-4.315	10.268	1.00	35.57	C	DNA D鎖の 原子座標
ATOM	1342	N2	G C	17	19.925	-5.605	10.599	1.00	35.65	N	
ATOM	1343	N3	G C	17	21.430	-3.878	10.217	1.00	32.09	N	
ATOM	1344	C4	G C	17	21.493	-2.578	9.883	1.00	21.03	C	
TER	1345		G C	17							
ATOM	1346	O5*	C D	1	11.703	-8.680	14.224	1.00	28.24	O	
ATOM	1347	C5*	C D	1	12.057	-8.948	12.859	1.00	26.74	C	
ATOM	1348	C4*	C D	1	13.503	-9.377	12.762	1.00	31.72	C	
ATOM	1349	O4*	C D	1	14.140	-8.600	11.718	1.00	39.60	O	
ATOM	1350	C3*	C D	1	14.298	-9.090	14.036	1.00	34.74	C	
ATOM	1351	O3*	C D	1	15.288	-10.094	14.266	1.00	32.94	O	
ATOM	1352	C2*	C D	1	14.959	-7.756	13.749	1.00	41.21	C	
ATOM	1353	Cl*	C D	1	15.223	-7.847	12.256	1.00	45.23	C	
ATOM	1354	N1	C D	1	15.249	-6.533	11.586	1.00	49.84	N	
ATOM	1355	C2	C D	1	16.489	-6.008	11.149	1.00	50.89	C	
ATOM	1356	O2	C D	1	17.527	-6.667	11.337	1.00	47.15	O	
ATOM	1357	N3	C D	1	16.513	-4.797	10.536	1.00	50.87	N	
ATOM	1358	O4	C D	1	15.377	-4.120	10.349	1.00	49.62	O	
ATOM	1359	N4	C D	1	15.450	-2.936	9.739	1.00	49.15	N	
ATOM	1360	C5	C D	1	14.112	-4.630	10.780	1.00	53.28	C	
ATOM	1361	O6	C D	1	14.094	-5.827	11.389	1.00	49.92	O	
ATOM	1362	P	T D	2	16.352	-9.882	15.446	1.00	21.84	P	
ATOM	1363	O1P	T D	2	17.225	-11.068	15.516	1.00	29.29	O	
ATOM	1364	O2P	T D	2	15.635	-9.440	16.656	1.00	27.06	O	
:											
ATOM	1688	N2	G D	17	11.323	9.011	62.477	1.00	54.12	N	水の 原子座標
ATOM	1689	N3	G D	17	13.569	9.498	62.545	1.00	50.24	N	
ATOM	1690	C4	G D	17	14.787	8.977	62.297	1.00	45.74	C	
TER	1691		G D	17							
HETATH	1692	O	HOH	1	13.098	-7.247	26.613	1.00	40.32	O	
HETATH	1693	O	HOH	2	16.197	-9.171	25.226	1.00	22.22	O	
HETATH	1694	O	HOH	3	12.437	-7.744	49.623	1.00	38.95	O	
HETATH	1695	O	HOH	4	33.849	16.525	38.170	1.00	6.14	O	
HETATH	1696	O	HOH	5	6.669	21.212	29.263	1.00	35.12	O	
HETATH	1697	O	HOH	6	18.135	0.105	59.664	1.00	13.39	O	
HETATH	1698	O	HOH	7	10.540	-5.597	58.865	1.00	19.04	O	
HETATH	1699	O	HOH	8	20.325	16.039	22.103	1.00	2.07	O	
HETATH	1700	O	HOH	9	8.348	29.524	49.561	1.00	48.19	O	
:											
HETATH	1770	O	HOH	79	12.012	24.584	29.542	1.00	2.00	O	
HETATH	1771	O	HOH	80	8.102	32.342	26.634	1.00	2.07	O	
MASTER	218	O	O	7	0	0	0	6.1767	4	0	14
END											

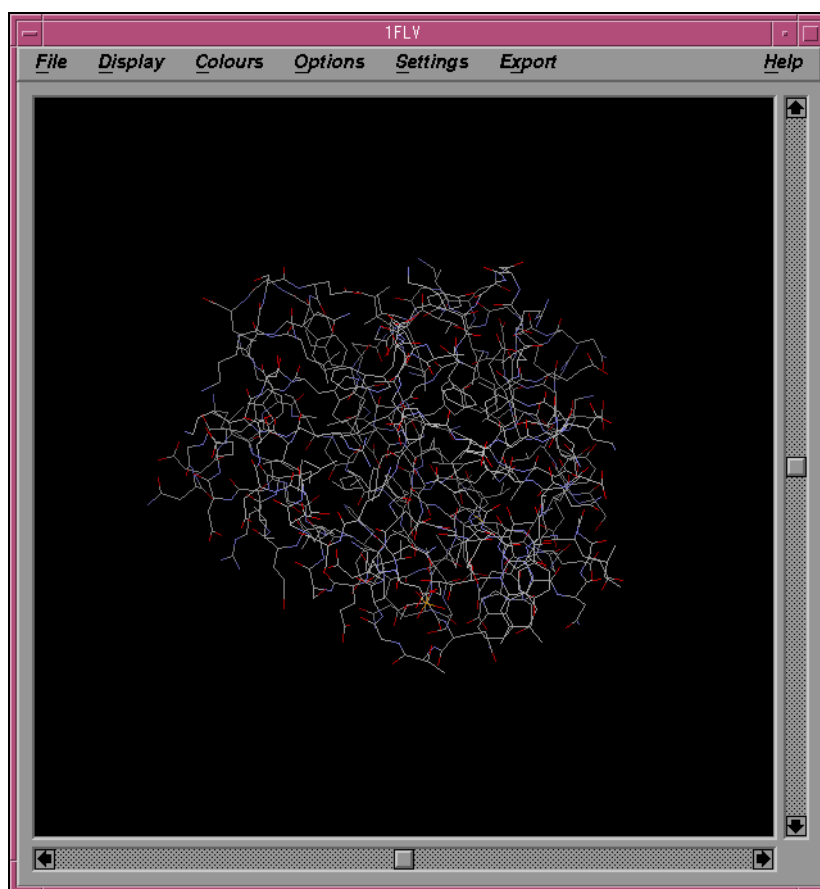
3. RasMol を使ったPDBデータの可視化

PDB 形式の立体構造データを表示するソフトは多数存在しますが、最もよく使われているソフトウェアはRasMol です。グラフィックスは比較的簡素ですが、描画は高速で、幅広いプラットフォームで同様に動作するのが特長です。http://www.openrasmol.org にアクセスすれば、Windows 版、Mac 版、Linux 版の実行ファイル、ソースコードを入手することが可能です。

起動するには以下のコマンドを打ちます。

```
rasmol [PDB ファイル名]
```

次のようなウィンドウが起動するはずですが。

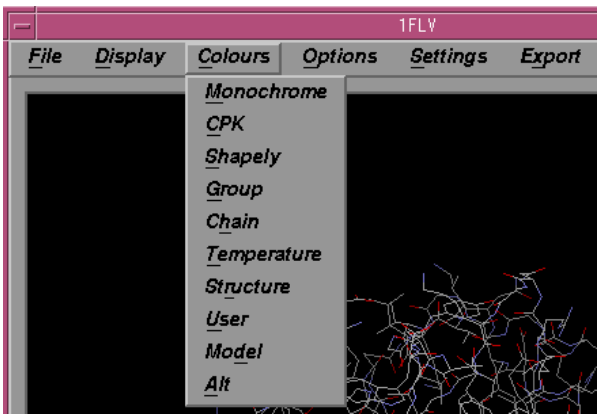


3.1 メニューによる RasMol の簡易操作

まず、マウスによる回転・並進操作になれてください。マウスを分子が描画している画面でボタンを押しながら動かすことで、分子を回転・並進させることができます。基本は以下の3つです。

左ボタン	回転 (X 軸、Y 軸まわりの回転)
中ボタン	移動 (X、Y 方向への並進)
<code>Shift</code> キー + 左ボタン	ズーム (拡大、縮小)

また、画面上部のメニューを操作することで、いろいろな描画法、色付け法を試すことができます。

 <p>The screenshot shows the 'File' menu in the RasMol application. The menu items are: Open..., Save As..., Close, Exit, and a file named 1.pdb1flv.ent. The application window title is '1FLV'.</p>	<p>[File]メニューから、[Open], [Save As...], [Close], [Exit]を選ぶことができます。</p>
 <p>The screenshot shows the 'Display' menu in the RasMol application. The menu items are: Wireframe, Backbone, Sticks, Spacefill, Ball & Stick, Ribbons, Strands, and Cartoons. The application window title is '1FLV'.</p>	<p>[Display]メニューから、[Wireframe], [Backbone], [Sticks], [Spacefill], [Ball & Stick], [Ribbons], [Strands], [Cartoons]の表示様式を選ぶことができます。</p>
 <p>The screenshot shows the 'Colours' menu in the RasMol application. The menu items are: Monochrome, CPK, Shapely, Group, Chain, Temperature, Structure, User, Model, and Alt. The application window title is '1FLV'.</p>	<p>[Colours]メニューから、[Monochrome], [CPK], [Shapely], [Group], [Chain], [Temperature], [Structure], [User], [Model], [Alt]の色付けを選ぶことができます。</p>

また、画面上で原子をマウスでクリックすると、RasMol を立ち上げたターミナルにクリックされた原子名が以下のように表示されます。

<pre>RasMol> (画面上でマウスをクリックする) Atom: OD1 129 Group: ASN 14 Chain:A RasMol> (画面上でマウスをクリックする) Atom:CG 679 Group:HIS 21 Chain:B</pre>

3.2 コマンドラインによる RasMol の操作

RasMol は、ターミナルから、コマンドを入力することで、より細やかな表示の指定が可能になります。RasMol を立ち上げると、分子グラフィックスのウィンドウが表示されると同時に、立ち上げたターミナルに、以下のようなコマンドプロンプトが現れるはずです。

```
RasMol>
```

ここにコマンドを入力すると、様々な表示法を指定することができます。RasMol のコマンドはキャラクターベースのインターフェース (CUI) という点で、UNIX コマンドや gnuplot のコマンドと似ているところもあります。基本的に、[選択] をしてから、選択された原子に対して [アクション] をするという作法です。

>> RasMol のコマンドラインの使い方 <<

- [1. 選択] : select コマンド等を使って、対象とする原子を指定する。
- [2. アクション] : 選択された原子について、なんらかの表示をするコマンドを実行する。

よって、常に、今、何が選択されているか、意識して使う必要があります。最初に立ち上げた状態では、全ての原子が選択された状態になっています。コマンドの詳細については、付録の RasMol コマンドのリファレンスのページを参考にしてください。以下に、いくつか、選択、アクションの組み合わせの例を示します。

(1) 全てを空間充填モデルで表示する

```
RasMol> select all  
RasMol> spacefill true
```

こうすると、たいていの場合、大量の水の酸素の球が表示されてしまいます。水の空間充填モデルの表示を止めるには以下のようにします。

```
RasMol> select water  
RasMol> spacefill false
```

そうすると、水の空間充填モデルが解除されます。この spacefill のところを変えるといろいろ他の表示法が指定できます。重要なものは以下の4つです。

spacefill : 空間充填モデル
wireframe : ワ이어フレームモデル
backbone : バックボーンモデル
catroon : カートゥーンモデル

また、分子の選択についても、select water 以外に、以下のような選択ができます。

select water : 水を選択
select protein : タンパク質を選択
select dna : DNA を選択
select ligand : 結合している低分子を選択

(2) 選択した種類のアミノ酸だけ、色を変える

Cysteine だけを選択して、色を黄色に変えたい場合、以下のようなコマンドを打ちます。

```
RasMol> select cys  
RasMol> color yellow
```

色としては、**color yellow** の他に **color red**、**color blue**、**color green** など使えます。また、**color chain** とすると鎖ごとの色分け、**color structure** とすると2次構造ごとの色分けが可能です。

(3) ある残基番号のアミノ酸だけ、色を変える

12番目のアミノ酸を、赤にしたい場合は、以下のようなコマンドを打ちます。

```
RasMol> select 12  
RasMol> color red
```

12番目から16番目のアミノ酸を、緑にしたい場合、以下のようなコマンドを打ちます。

```
RasMol> select 12-16  
RasMol> color green
```

(4) ある鎖だけ選択して色を変える

A鎖の色をオレンジに変えたい場合、以下のコマンドを打ちます。

```
RasMol> select *:A  
RasMol> color orange
```

(5) ある分子の近傍だけ選択して、強調して表現する

HEM という分子の周辺7Åにある原子を選択して、太いワイアフレームで表示したい場合、以下のコマンドを打ちます。

```
RasMol> select within(7.0, HEM)  
RasMol> wireframe 100
```

(6) 画像ファイルの保存

画像ファイルを保存する場合は、次のコマンドを入力してください。様々なタイプの画像形式に対応しているはずですが、本演習では、Sun raster 形式の画像ファイルを使ってください。出力画像ファイル名は、必ず末尾が **.ras** となるように指定してください。

```
RasMol> write ras <画像ファイル名(末尾は必ず.ras)>
```

確認のために、保存された画像ファイルを確認するコマンドは、演習環境では **display** というプログラムが実装されています。UNIX のターミナルから以下のコマンドを入力してください。

```
display <画像ファイル名>
```

4. MATRAS サーバを用いた立体構造予測：簡易ホモロジーモデリング

アミノ酸配列だけから立体構造を推定することを立体構造予測といいます。立体構造予測には様々なアプローチがありますが、本演習では、最も単純で有効性の高いホモロジーモデリング法を行います。一般にホモロジーモデリング法は、以下のような2段階からなります。

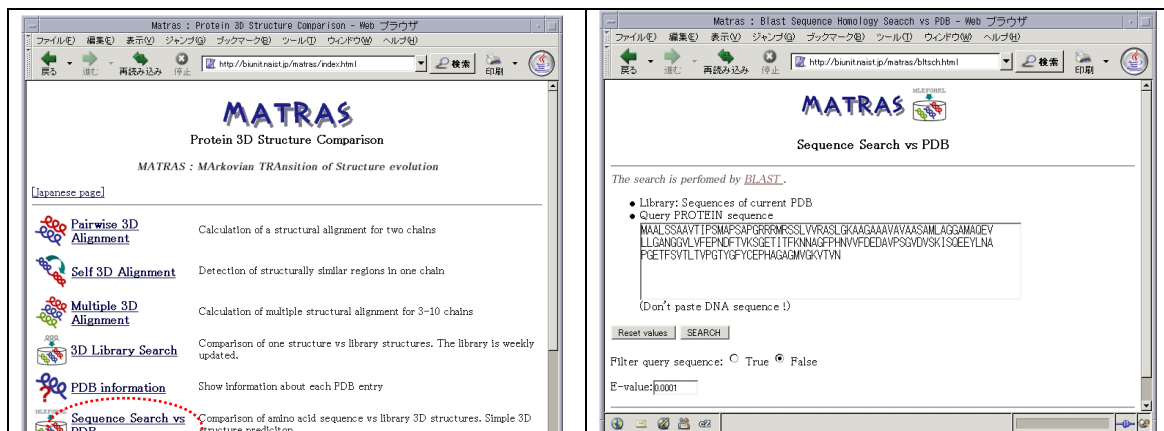
[ステップ1]：テンプレート構造の発見 (Fold recognition)

予測したい配列に最も近いと思われる立体構造（テンプレート構造）を既知の構造データベースの中から見つける。

[ステップ2]：全原子の構造の構築 (Modeling)

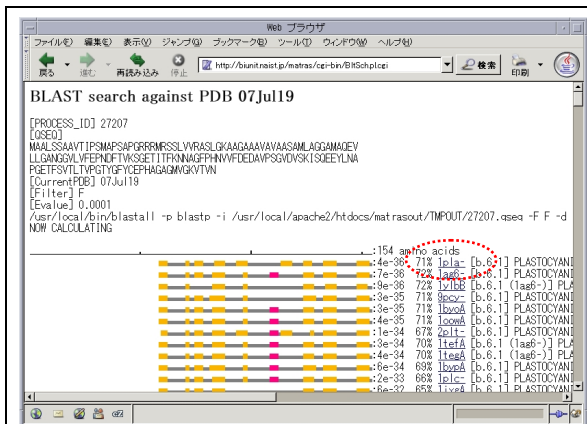
発見したテンプレート構造を元にして主鎖構造をできるだけ維持するようにして、側鎖構造の構築、削除・挿入されたアミノ酸の主鎖構造の構築を行う。

本演習では、[ステップ1]の部分を MATRAS サーバを用いて、blastp のプログラムで行い、[ステップ2]はアミノ酸名と残基番号を置き換えるだけの簡易法をやはり MATRAS サーバの機能を用いて行うことにします。

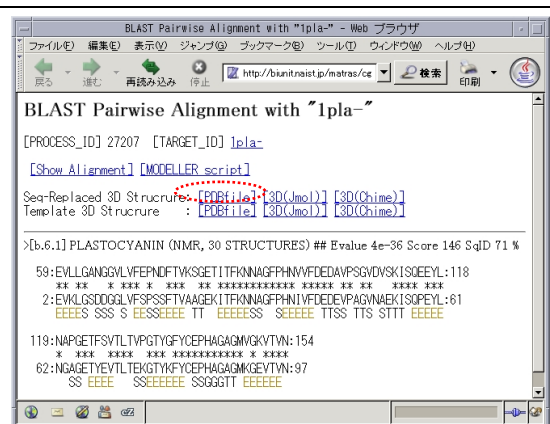


(1) MATRAS サーバ
<http://biunit.naist.jp/matras/>
にアクセスし、[Sequence Search vs PDB]をクリックします。

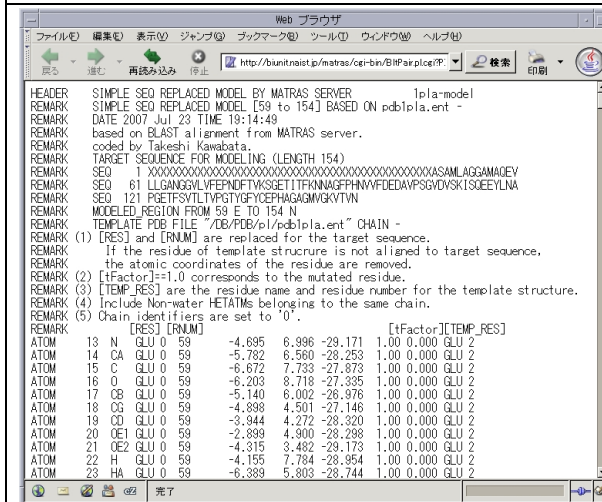
(2) 予測したいアミノ酸配列をペーストし、[Search]をクリックします。すると、最新の PDB のアミノ酸配列データベースに対して、blastp が実行されます。



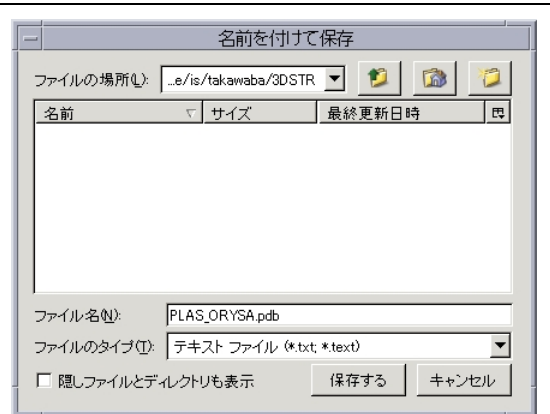
(3) blastp の結果が、2次構造のバーの形式で表示されます。次に E-value が最も低い PDB コード名をターゲット構造として選択して、クリックします。



(4) クエリ配列と、指定した PDB のエンタリとのアライメントが表示されます。次に、Seq-Replaced 3D Structure の [PDB file] をクリックします。



(5) クエリ配列のアミノ酸名・残基番号に書き換えたターゲット構造が PDB 形式で表示されます。次に、ブラウザの [ファイル] メニューから、[名前を付けて保存] を選択します。



(6) ファイル保存のためのウィンドウが表示されます。末尾が、.pdb であるようなファイル名を [ファイル名] に指定し、[ファイルのタイプ] を "テキスト ファイル" に設定して、保存してください。

保存された予測構造は、通常の PDB ファイルと同じように、RasMol で可視化することができます。

rasmol [予測構造の PDB ファイル (ファイル末尾が .pdb)]

付録：RasMol 2.7.2 簡易コマンド・リファレンス


本ページは、RasMol Quick Reference を参考に、部分的に改変・日本語訳したものです
(<http://biochemistry.utoronto.ca/steipe/bioinformatics/tutorials/refcard.pdf>)
RasMol のソフトウェアの入手、より詳細なコマンドの説明は、RasMol の Web サイト：
<http://www.openrasmol.org> をご覧ください。


マウスボタンによる操作

左ボタン	回転 (X 軸、Y 軸まわりの回転)
中ボタン	移動 (X、Y 方向への並進)
Shift キー + 左ボタン	ズーム (拡大、縮小)
Shift キー + 右ボタン	Z 軸まわりの回転
Control キー + 左ボタン	切断表示 (Slab) の切断面を Z 軸に対して調節。

※X軸：画面の右方向 Y軸：画面の下方向 Z軸：画面に垂直で手前から奥への方向。

コマンドラインの編集

Control+**B** あるいは  : カーソルの後退

Control+**F** あるいは  : カーソルの前進

Control+**D** : カーソルにある文字の消去

Delete あるいは **Backspace** : カーソルの前の文字の消去

Control+**P** あるいは  : 一つ前のコマンドの表示

Control+**N** あるいは  : 一つ先のコマンドの表示

一般的なコマンド

load [フォーマット] <ファイル名> : ファイルの読み込み

フォーマットのデフォルトは pdb 形式

zap : 読み込んだ分子の削除

exit : RasMol の終了

help [トピック] : オンラインヘルプの表示

選択

select <選択表現> : 分子の一部だけを<選択表現>によって選択する
restrict <選択表現> : 選択した部分だけを表示する。

and もしくは **&&** : 論理積 (AND)
or もしくは **||** : 論理和 (OR)
not もしくは **!** : 論理否定 (NOT)

>> 選択の例 <<

select * : 全ての原子を選択
select Cys : Cysteine の原子を選択
select hoh : 水の原子を選択
select protein : 蛋白質の原子を選択
select *:A : A 鎖の原子を選択
select *.CA : CA 原子を選択

>> 定義済み(predefined)の集合表現を使用する例 <<

select helix : ヘリックスの原子を選択
select hydrophobic : 疎水性アミノ酸の原子を選択
select backbone : 主鎖の原子
select sidechain : 側鎖の原子

>> 残基番号の範囲の選択の例 <<

select 5,10,16 : 5 番目と 3 番目と 16 番目の残基(group)を選択
select 5-16 : 5 番目から 16 番目の残基(group)を選択

>> 論理演算の例 <<

select hydrophobic and sidechain : 疎水性アミノ酸の側鎖の原子
select !HOH && !protein : 水でなく、蛋白質でもない原子

>> 比較演算子の例 <<

select atomno=1203 : 原子番号が 1203 番の原子を選択
select temperature>=800 : 温度因子が 900 以上の原子

>> 距離表現 <<

select within(7.0,HEM) : HEM から 7Å以下の原子を選択

定義済み(predefined)集合表現

at	A か T の塩基	acidic	Asp か Glu	acyclic	環を持たない
aliphatic	脂肪族の AGILV	alpha	C α 原子	amino	アミノ酸
aromatic	芳香族 HFWY	backbone	主鎖の原子	basic	Arg,His,Lys
bonded	他原子と結合	buried	ACILMFVV	cg	C か G の塩基
charged	電荷 RDEHK	cyclic	環を持つ HFPWV	helix	α ヘリックス構造
hetero	HETATM	hydrogen	水素	hydrophobic	AGILMFVWYV
ions	イオン金属	large	REQHILKMFVY	ligand	水でない hetero
medium	NDCPT	neutral	Charged 以外	nucleic	核酸
polar	RNDCEQHKST	protein	蛋白質	purine	A か G の塩基
pyrimidine	C か T の塩基	sheet	β シート構造	sidechain	側鎖の原子
small	Ala,Gly, Ser	solvent	water か ion	surface	RNDEAGHKPSTY
turn	ターン構造	water	水		

表示法（レンダリング）のコマンド

background <色> : 背景の色を指定
set ambient <値> : Depth-cue の背景光の強さを指定
set shadows <真偽> : 影の表示をする・しない
set specular <真偽> : highlight(原子の反射光)の表示をする・しない
set specpower <値> : highlight(原子の反射光)の強さを指定

表示のコマンド

※<値>の大きさは全て 1/250 Å 単位 (例: **spacefill 500** → 半径 2 Å の球で表示)

wireframe <真偽> : ワイアフレーム表示をする・しない
wireframe <値> : <値>の太さのシリンダでワイアフレーム表示をする

spacefill <真偽> : 空間充填球の表示をする・しない
spacefill <値> : <値>の半径で空間充填球の表示をする

backbone <真偽> : バックボーン表示 (Calpha 原子を通るシリンダ) をする・しない
backbone <値> : <値>の半径のシリンダでバックボーン表示をする。

ribbons <真偽> : リボン表示をする・しない
ribbons <値> : <値>の幅のリボン表示をする。

cartoon <真偽> : カートゥーン(厚みのあるリボン)表示をする・しない
cartoon <値> : <値>の幅のカートゥーン表示をする。

label <真偽> : 原子にラベル原子名等の文字列を表示する
label <文字列> : <文字列>を原子にラベル表示する
set fontsize <値> : 文字列のフォントの大きさを<値>に設定する

ssbonds <真偽> : SS 結合を表示する
ssbonds <値> : SS 結合を<値>の太さで表示する
set ssbonds backbone : SS 結合を主鎖の Calpha 原子間に設定する
set ssbonds sidechain : SS 結合を側鎖の硫黄原子間に設定する

hbonds <真偽> : 水素結合を表示する・しない
hbonds <値> : 水素結合を<値>の太さで表示する

monitor <原子番号>, <原子番号> : 2つの原子の間に点線を引き、
距離のラベルを表示
monitor <真偽> : monitor 表示をする・しない
set monitor <真偽> : monitor のラベルを表示する・しない

dots <真偽> : ドットサーフェースを表示する
dots <値> : <値>の大きさの点密度でドットサーフェースを表示

色のコマンド

color <オブジェクト> <色>: 色を指定

<オブジェクト>: 以下から選択できる。省略した場合は、全てが対象となる。

atoms bonds backbone ribbons labels hbonds
ssbonds dots axes

<色>:

定義済みの色

blue black cyan deeppink firebrick forestgreen
grey green greenblue maganta orange pink purple
red redorange skyblue violet white yellow

原子のカラーリング法

cpk	炭素:灰、酸素:赤、窒素:青	amino	アミノ酸ごとの色
shapely	アミノ酸ごとの色	group	残基番号順に青→赤
chain	鎖ごとに別の色を付ける	structure	ヘリックス magenta シート yellow
temperature	温度因子に従って青→赤	charge	温度因子に従って赤→青

回転・拡大操作などのコマンド

rotate <軸> <回転角(°)>: <軸>を回転軸として<回転角>° 回転

translate <軸> <値(Å)>: <軸>の方向に<値>Åだけ並進移動

<軸>: **x** あるいは **y** あるいは **z**

>>回転・並進の例<<

rotate x 30 : x 軸に対して30度回転

translate x 2.5 : x 軸に対して2.5Å並進移動

zoom <値(%)>: <値> (%) に拡大。 **zoom 100** がそのまま。

slab <値(0~100)>: 切断表示の調節。画面に垂直な z 軸に垂直な面で切断表示する。
切断面の位置を<値>で指定。分子の背を 0、分子の手前を 100 とする。

slab <真偽>: 切断表示をする・しない

reset: 回転、並進、拡大、切断をリセットし、初期表示に戻す。

出力コマンド

画像ファイルを出力

write <フォーマット> <ファイル名> : ファイルに画像を出力
<フォーマット>

gif : gif 画像フォーマット
ps : PostScript 画像フォーマット
epsf : Encapsulated Postscript 画像フォーマット
monops : Monochrome Postscript 画像フォーマット
bmp : Microsoft BMP 画像フォーマット
pict : Apple PICT 画像フォーマット
ppm : Portable Pixmap 画像フォーマット
ras : Sun raster 形式の画像フォーマット

画像以外のファイルを出力

write vectps <ファイル名> : Postscript 形式の線画フォーマットで出力。
※ワイアフレーム、空間充填、バックボーン表示のみ対応。

write molscript <ファイル名> : Molscript のスクリプトファイルを出力

write vrml <ファイル名> : VRML 形式のファイルを出力。
※ ワイアフレーム、空間充填の表示のみ対応。

write phipsi <ファイル名> : 主鎖の角度 ϕ 、 ψ の値をテキスト形式で出力

write ramachan <ファイル名> : ラマチャンドラン・マップ (ϕ 、 ψ の 2 次元分布図)
をテキスト形式で出力

write script <ファイル名> : 現在の表示を再現するためのコマンド群を、出力する。
※一度、RasMol を終了したあと、ターミナルのコマンドラインから、スクリプトファイルを指定して **rasmol -script** <ファイル名> と実行すると、全く同じ表示が再現される。

その他のコマンド

structure : 2 次構造の情報を表示

connect <真偽> : 共有結合の接続を再計算する・しない

show information : 構造の様々な情報を表示

show sequence : 配列を表示

show symmetry : 結晶の空間群の情報を表示

show selected : 選択されたグループ(残基)の情報を表示

show selected group : 選択されたグループ(残基)の情報を表示

show selected atom : 選択された原子の情報を表示

show selected atom : 選択された鎖の情報を表示