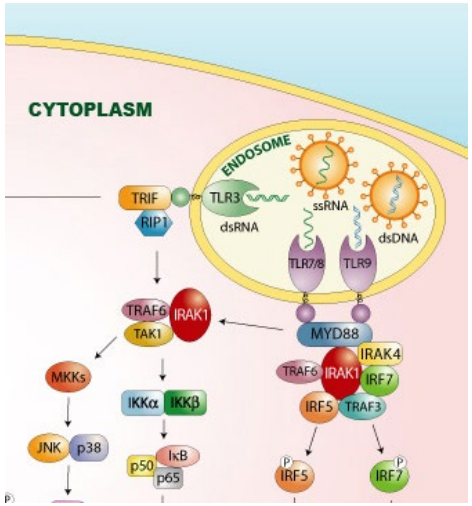
生物信息学

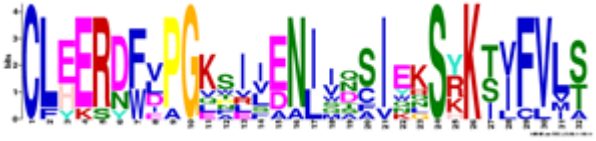
蛋白质三维结构预测 1

蛋白质结构



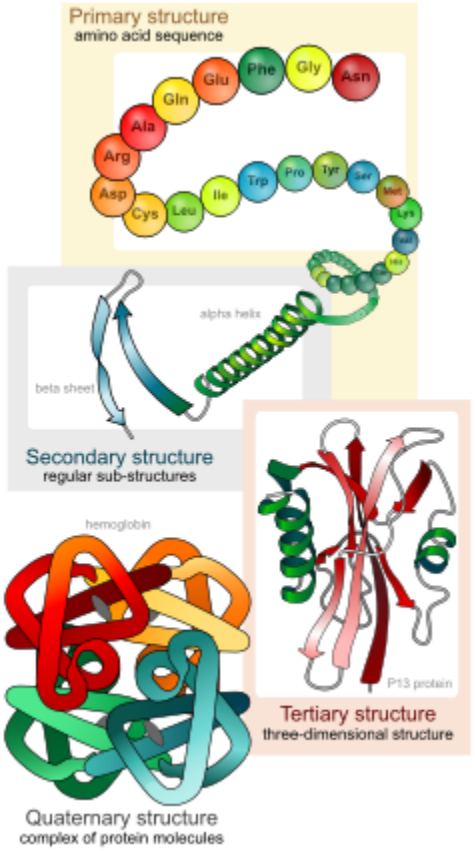
功能

结构



序列

蛋白质结构



蛋白质结构分四级：

一级结构

氨基酸序列

二级结构

周期性的局部结构构象

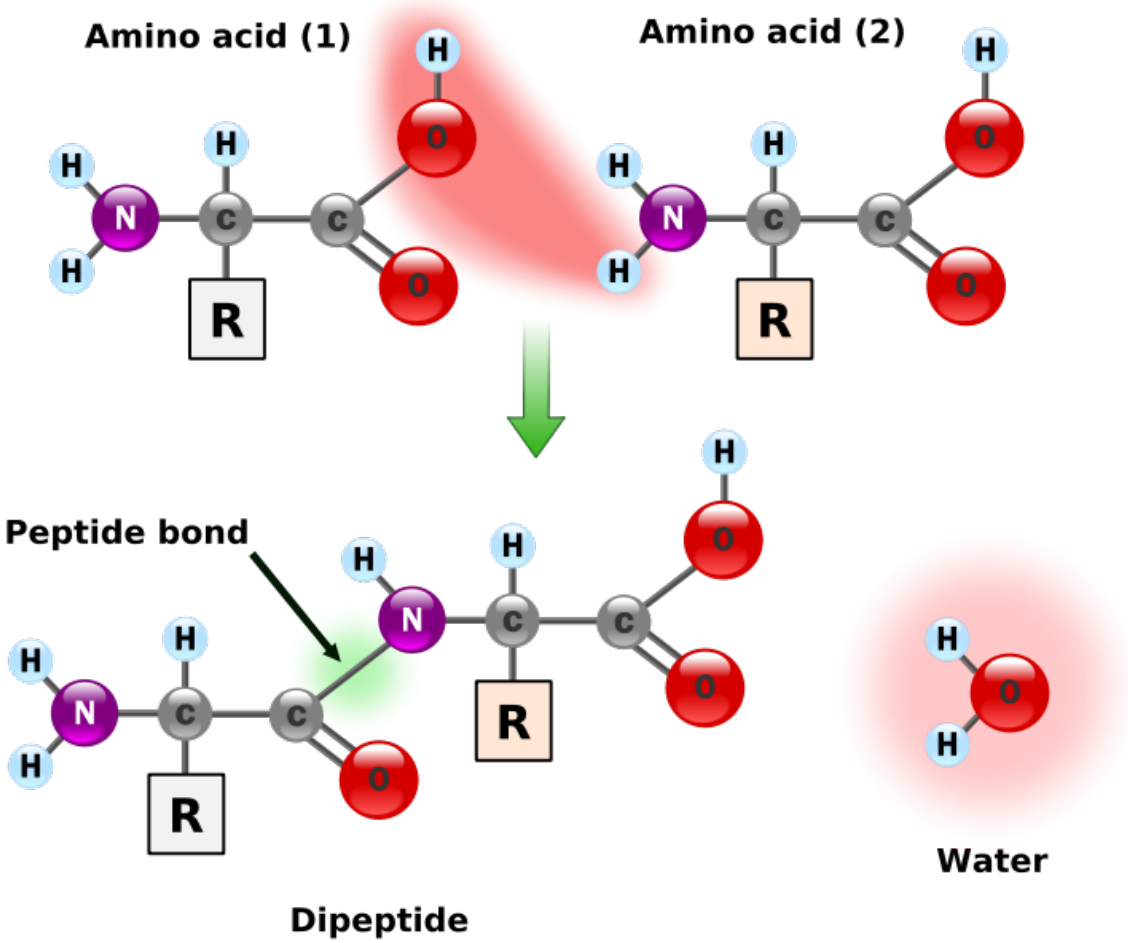
三级结构

三维空间结构

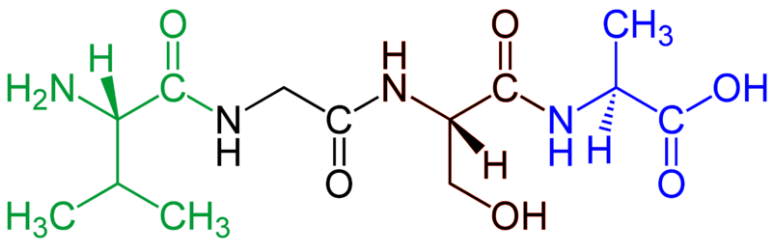
四级结构

几个蛋白质分子形成的复合体或含有多结构域单个蛋白质分子

蛋白质结构


Amino acid (1)
Amino acid (2)
Peptide bond
Dipeptide
Water
H
N
C
O
R

蛋白质结构 --- 一级结构

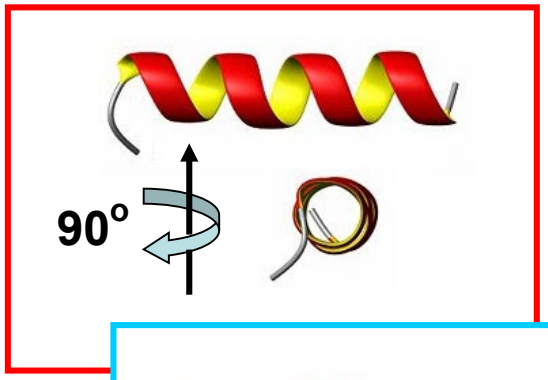


V G S A

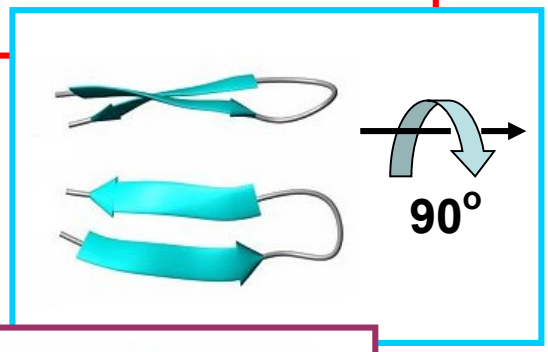
一级结构 – 蛋白质的氨基酸序列

```
>sp|P06968|DUT_ECOLI
MKKIDVKILDPRVGKEFPLPTYATSGSAGLDLRACLNDAVELAPGDTTLVPTGL
AIHIADPSLAAMMLPRSGLGHKHGIVLGNLVGLIDSDYQGQLMISVWNRGQDSF
TIQPGERIAQMIFVPVVQAEFNLVEDFDATDRGEGGFGHSGRQ
```

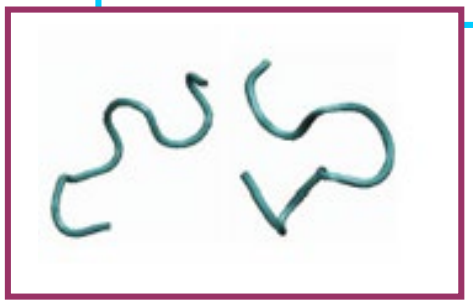
蛋白质结构 --- 二级结构



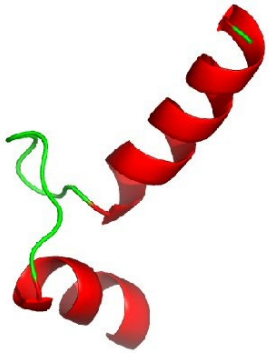
螺旋: 最常见的就是α螺旋。



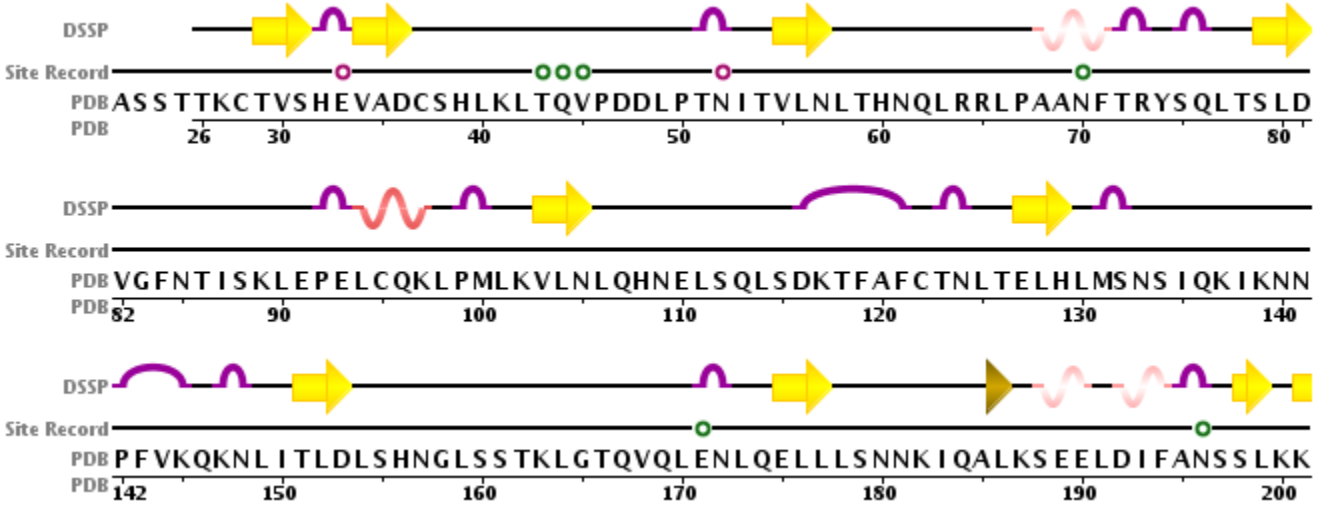
β折叠（β sheet）: β折叠由β折片（β strand）平行排列而成。



无规卷曲 (loop): 无规律松散结构.
转角（turn）: 如果肽链发生了急转弯 (角度大于90°)，这个转弯结构叫转角。



蛋白质结构 --- 二级结构




```
1    ASSTTKCTVS HEVADCSHLK LTQVPDDLPT NITVLNLTHN QLRRLPAANF
            EEE TTEEE  SS   SS  SSS T T SEEE  SS       GGGG

51   TRYSQLTSLD VGFNTISKLE PELCQKLPML KVLNLQHNEL SQLSDKTFAF
     TT TT SEEE   SS       TTHHHH TT   EEE  SS       TTTTT

101  CTNLTELHLM SNSIQKIKNN PFVKQKNLIT LDLSHNGLSS TKLGTQVQLE
      TT SEEE T TS      S  TTTT TT  E EE  SS         SSS  T

151  NLQELLLSNN KIQALKSEEL DIFANSSLKK LELSSNQIKE FSPGCFHAIG
     T  EEE  SS     B GGGT GGGTT EESE EE  S      B TTTTTTSS

201  RLFGLFLNNV QLGPSLTEKL CLELANTSIR NLSLSNSQLS TTSNTTFLGL
     EE EEE TT     HHHHHHH HHHTTTS    EEE TTS    EE TTTTGGG

251  KWTNLTMLDL SYNNLNVVGN DSFAWLPQLE YFFLEYNNIQ HLFSHSLHGL
     GGS   EEE  TTS   EE T TTTTT TT   EEE  S  B  EE TTTTTT
```

蛋白质结构 --- 三级结构

三级结构是指整条多肽链的三维空间结构，即，包括骨架和侧链在内的所有原子的空间排列。

X射线衍射法
X-ray Crystallography

>130000

核磁共振法
Nuclear Magnetic Resonance (NMR)

>12000

第一个蛋白质的三维空间结构是1958 由英国的Kendrew和Perutz博士用 X-射线晶体衍射法测定的。

Max Ferdinand Perutz (1914-2002) nobel prize 1962

John Cowdery Kendrew (1917-1997) nobel prize 1962

RCSB PDB PROTEIN DATA BANK

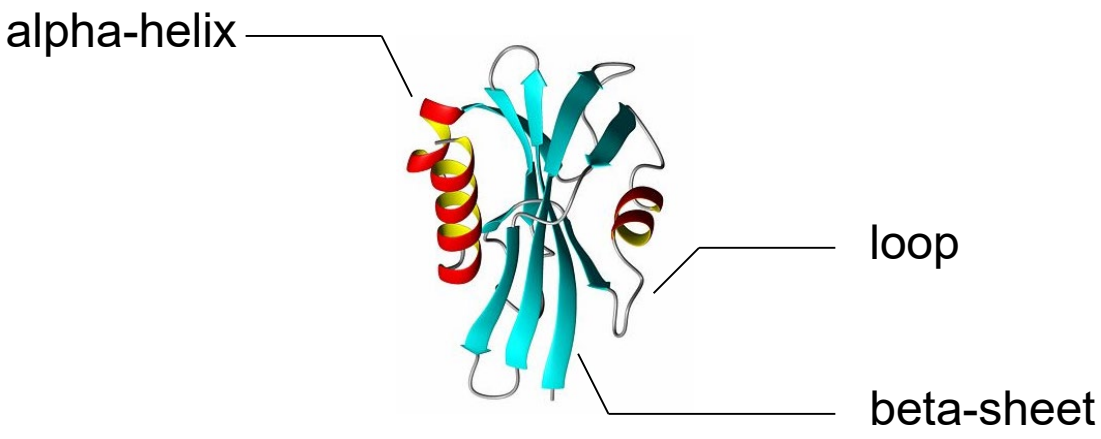
蛋白质结构 --- 二级结构的获取

如何获得蛋白质的二级结构

1. PDB/DSSP – 对于结构已知的蛋白质

DSSP (Definition of Secondary Structure of Proteins)，即，蛋白质二级结构词典。DSSP并不预测二级结构，而是根据二级结构的定义对已经测定结构的蛋白质的各个位置指出是哪种二级结构。

主页：**http://swift.cmbi.ru.nl/gv/dssp**



蛋白质结构 --- 二级结构获取

如何获得蛋白质的二级结构

1. PDB/DSSP – 对于结构已知的蛋白质

```
  66    92 A P  T >> S+     0   0   32      0, 0.0     3,-1.3     0, 0.0
  67    93 A E  H >> S+     0   0   79      1,-0.3     4,-3.1     2,-0.2
  68    94 A L  H <> S+     0   0   13     -3,-1.3     4,-1.4     1,-0.3
  69    95 A b  H <4 S+     0   0    4     -3,-1.3     6,-0.3    -4,-0.4
  70    96 A Q  H << S+     0   0   73     -3,-0.7    -2,-0.2    -4,-0.7
  71    97 A I  H  < S+     0   0   60     -4,-3.1    -1,-0.2     1,-0.3
  72    98 A L  S >< S+     0   0    6     -4,-1.4     3,-1.2    -5,-0.4
  73    99 A P  T 3  S+     0   0   67      0, 0.0    25,-0.2     0, 0.0
  74   100 A L  T 3    +     0   0   88    -23,-0.2     2,-0.6   -25,-0.2
  75   101 A L    <    +     0   0    7     -3,-1.2   -22,-0.2    -6,-0.3
  76   102 A K         +     0   0   83    -24,-3.2    24,-3.1    -2,-0.6
  77   103 A V  E      -de   53 101A  31    -25,-1.3   -23,-0.7    23,-0.2
  78   104 A L  E      -de   54 102A   6     23,-2.0    25,-1.3    -2,-0.6
  79   105 A N  E      +de   55 103A  17    -25,-1.9   -23,-1.9    -2,-0.4
  80   106 A L    >    +     0   0    0     23,-2.8     3,-0.9    -2,-0.4
  81   107 A Q  T 3    +     0   0   23      1,-0.2   -23,-2.2    24,-0.1
```

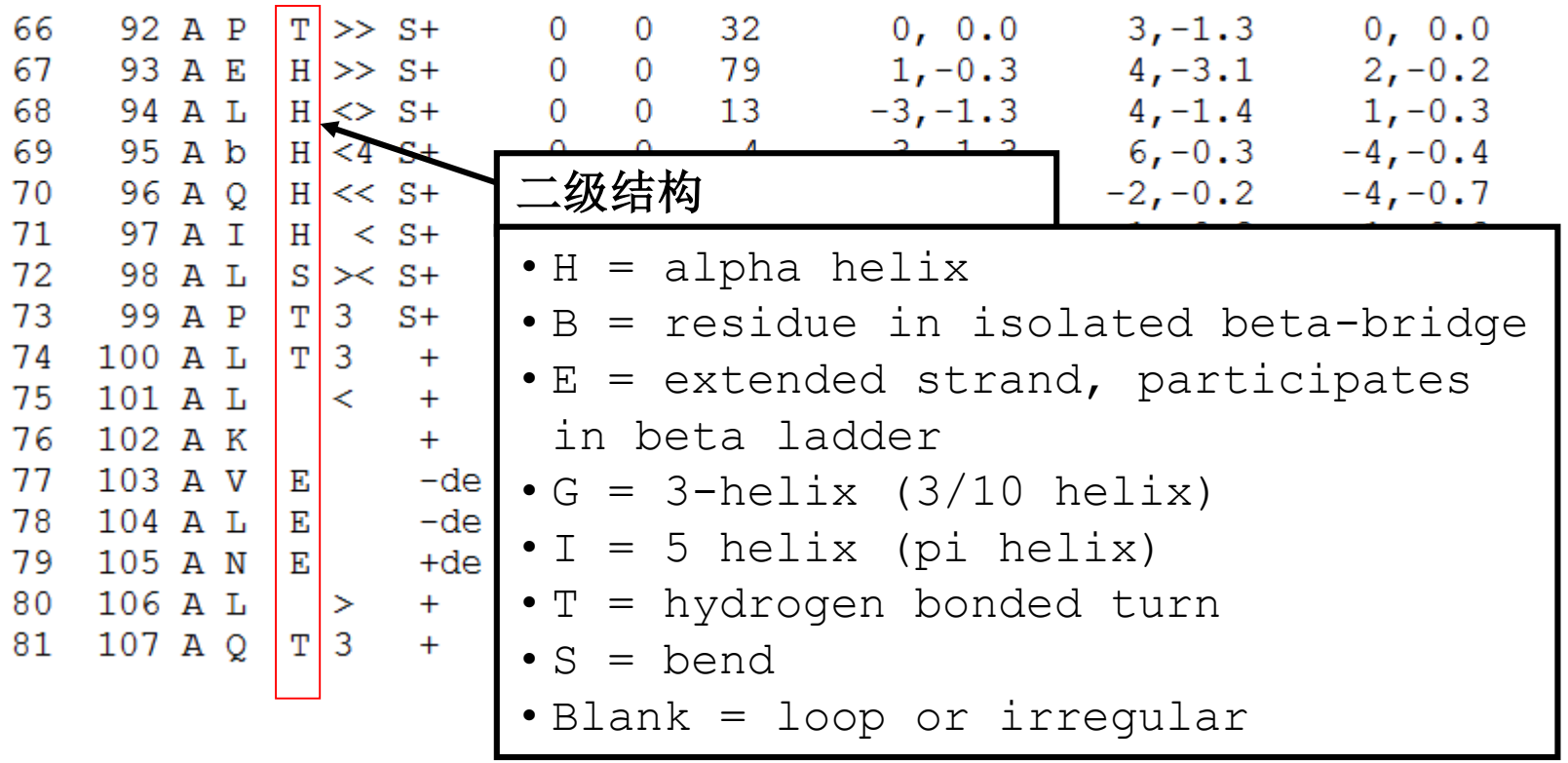
二级结构

一个**DSSP**文件的一部分

蛋白质结构 --- 二级结构获取

如何获得蛋白质的二级结构

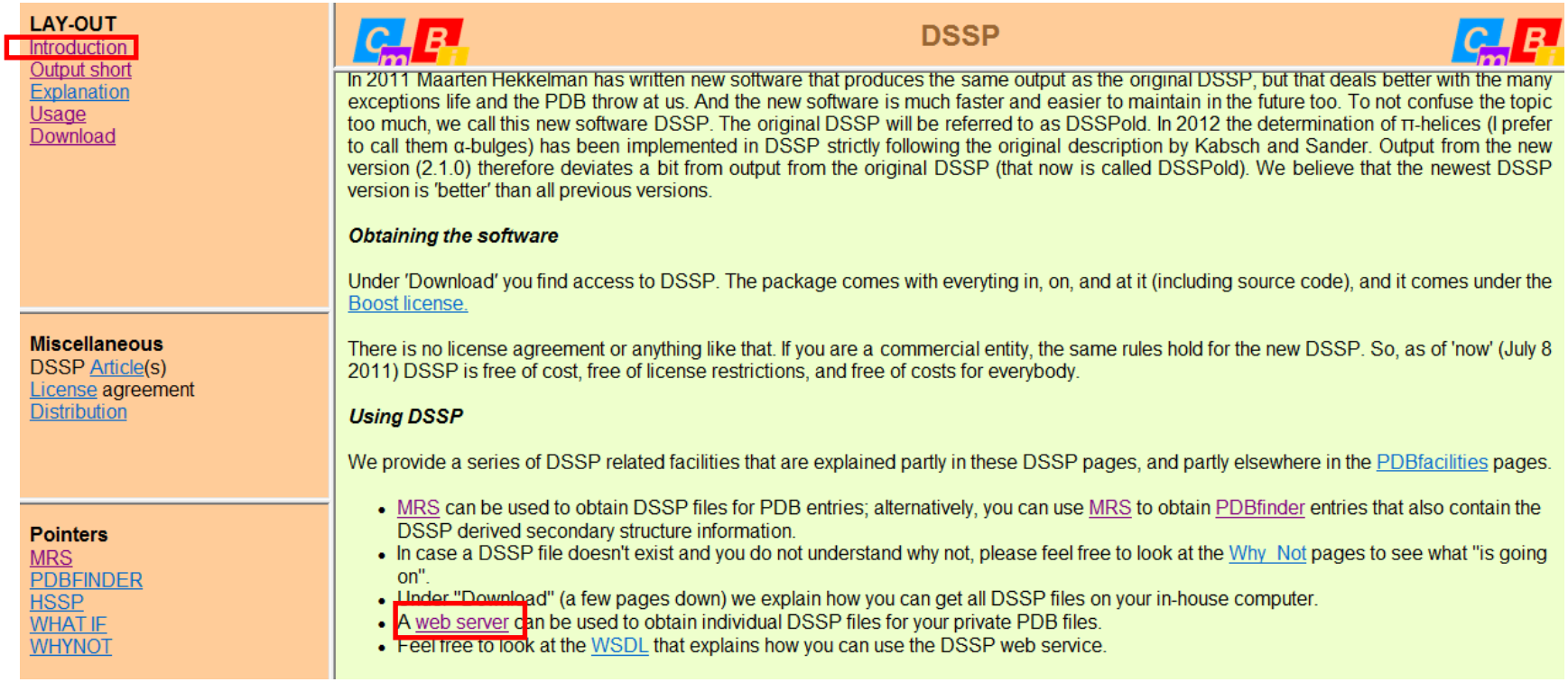
1. PDB/DSSP – 对于结构已知的蛋白质



蛋白质结构 --- 二级结构获取

如何获得蛋白质的二级结构

1. PDB/DSSP – 对于结构已知的蛋白质



蛋白质结构 --- 二级结构获取

如何获得蛋白质的二级结构

1. PDB/DSSP – 对于结构已知的蛋白质

蛋白质结构 --- 二级结构获取

如何获得蛋白质的二级结构

2. 预测软件 – 对于结构未知的蛋白质

已知一个蛋白质的氨基酸序列，预测其二级结构。

常用软件：PSIPRED, Jpred，APSSP2, NNPREDICT, PREDICTPROTEIN

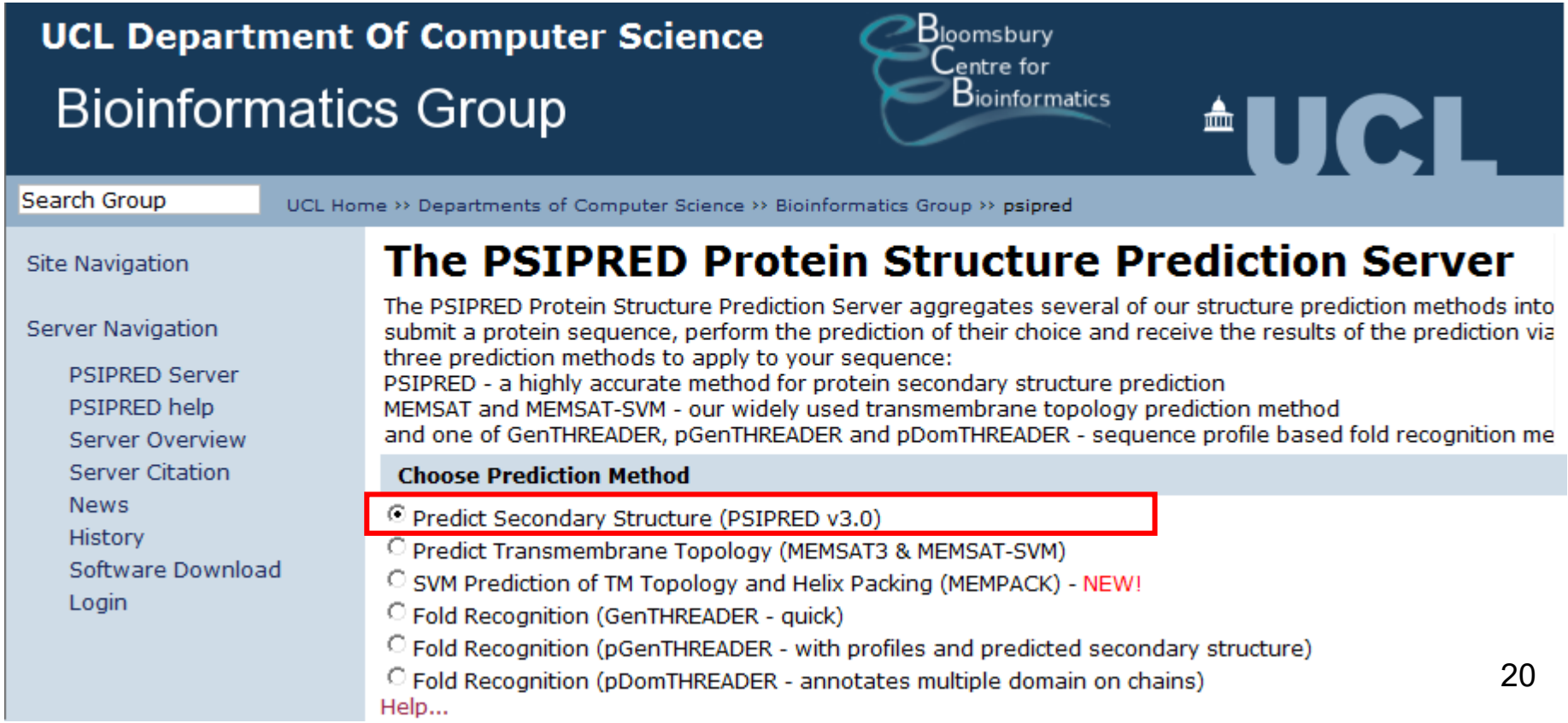
PSIPRED：**http://bioinf.cs.ucl.ac.uk/psipred**

Jpred：**http://www.compbio.dundee.ac.uk/www-jpred**

蛋白质结构 --- 二级结构获取

如何获得蛋白质的二级结构

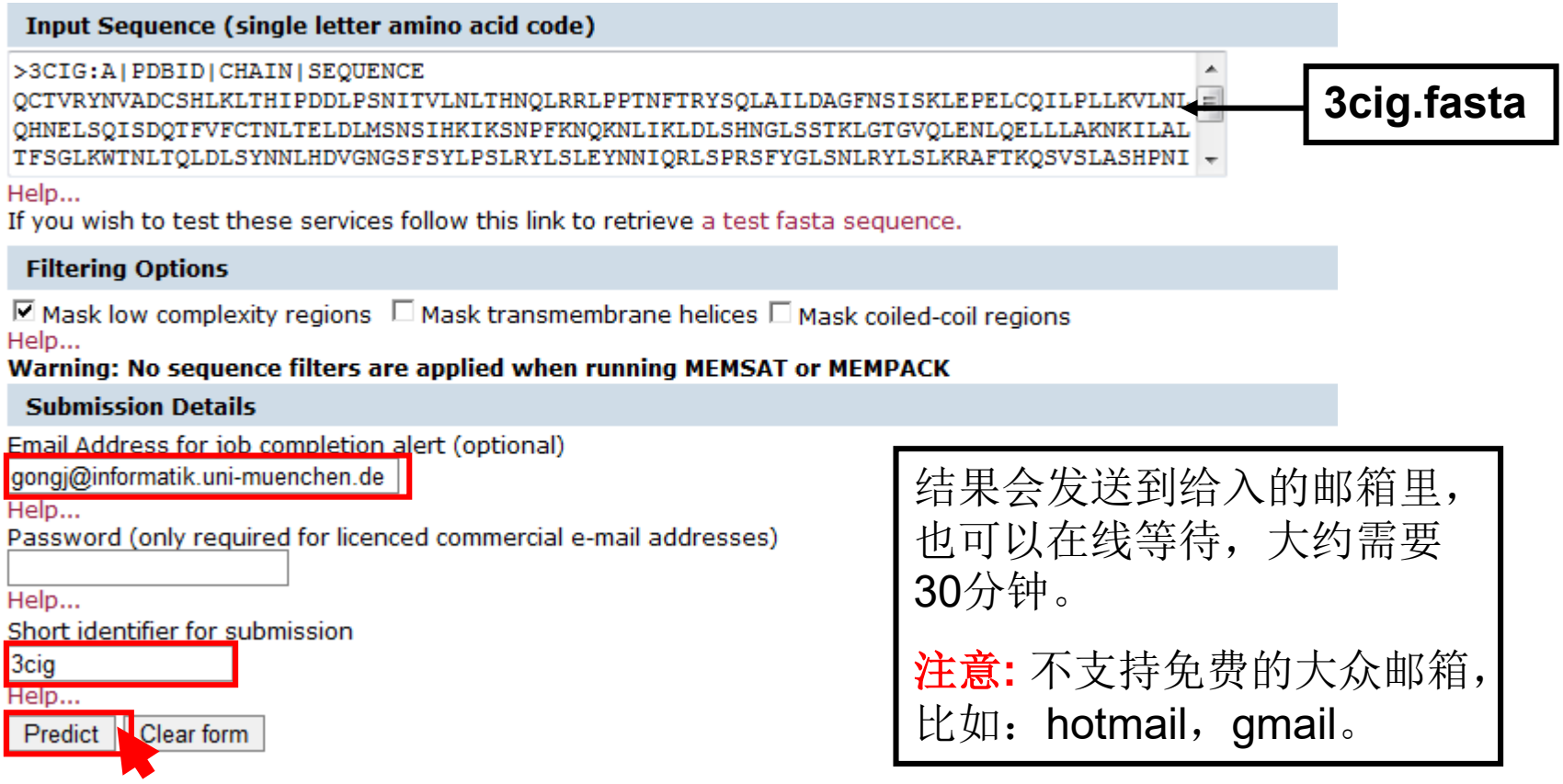
2. 预测软件 **–** 对于结构未知的蛋白质



蛋白质结构 --- 二级结构获取

如何获得蛋白质的二级结构

2. 预测软件 **–** 对于结构未知的蛋白质



结果会发送到给入的邮箱里，也可以在线等待，大约需要30分钟。

注意: 不支持免费的大众邮箱，比如：hotmail，gmail。

蛋白质结构 --- 二级结构获取

如何获得蛋白质的二级结构

2. 预测软件 **–** 对于结构未知的蛋白质

网页显示的预测结果

Results for psipred job 305290

Output for PSIPRED predictions

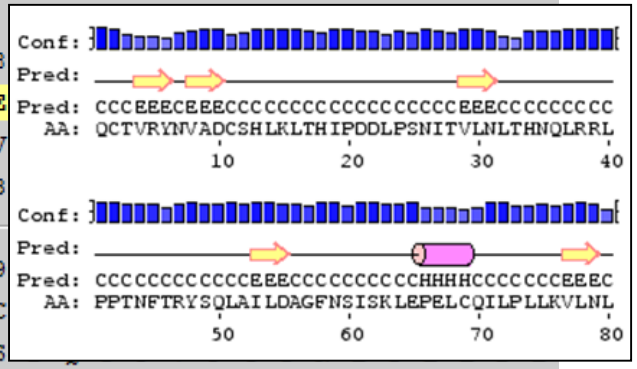
Psipred Residue Predictions

Conf: Confidence 0 =low 9 =high

Pred: Predicted secondary structure H =helix E =strand C =coil

AA: Target sequence

```
Conf: 9 8 3 2 2 0 4 7 9 8 4 6 8 9 9 9 8 6 8 9 8 8 5 8 6 9 7 5 8
Pred: C C C E E E C E E E C C C C C C C C C C C C C C C C C C E
  AA: Q C T V R Y N V A D C S H L K L T H I P D D L P S N I T V
                      1 0                 2 0                 3

Conf: 9 9 6 7 9 3 3 4 2 4 9 9 6 5 8 5 8 8 7 2 2 8 8 4 8 8 8 8 9
Pred: C C C C C H H H H C C C C C C C C E E E C C C C C C C C C C
  AA: I S K L E P E L C Q I L P L L K V L N L Q H N E L S Q I S
                      7 0                 8 0                 9 0                 1 0 0
```



Click to download these results in plain text format

蛋白质结构 --- 二级结构获取

如何获得蛋白质的二级结构

2. 预测软件 **–** 对于结构未知的蛋白质

发送至邮箱的预测结果

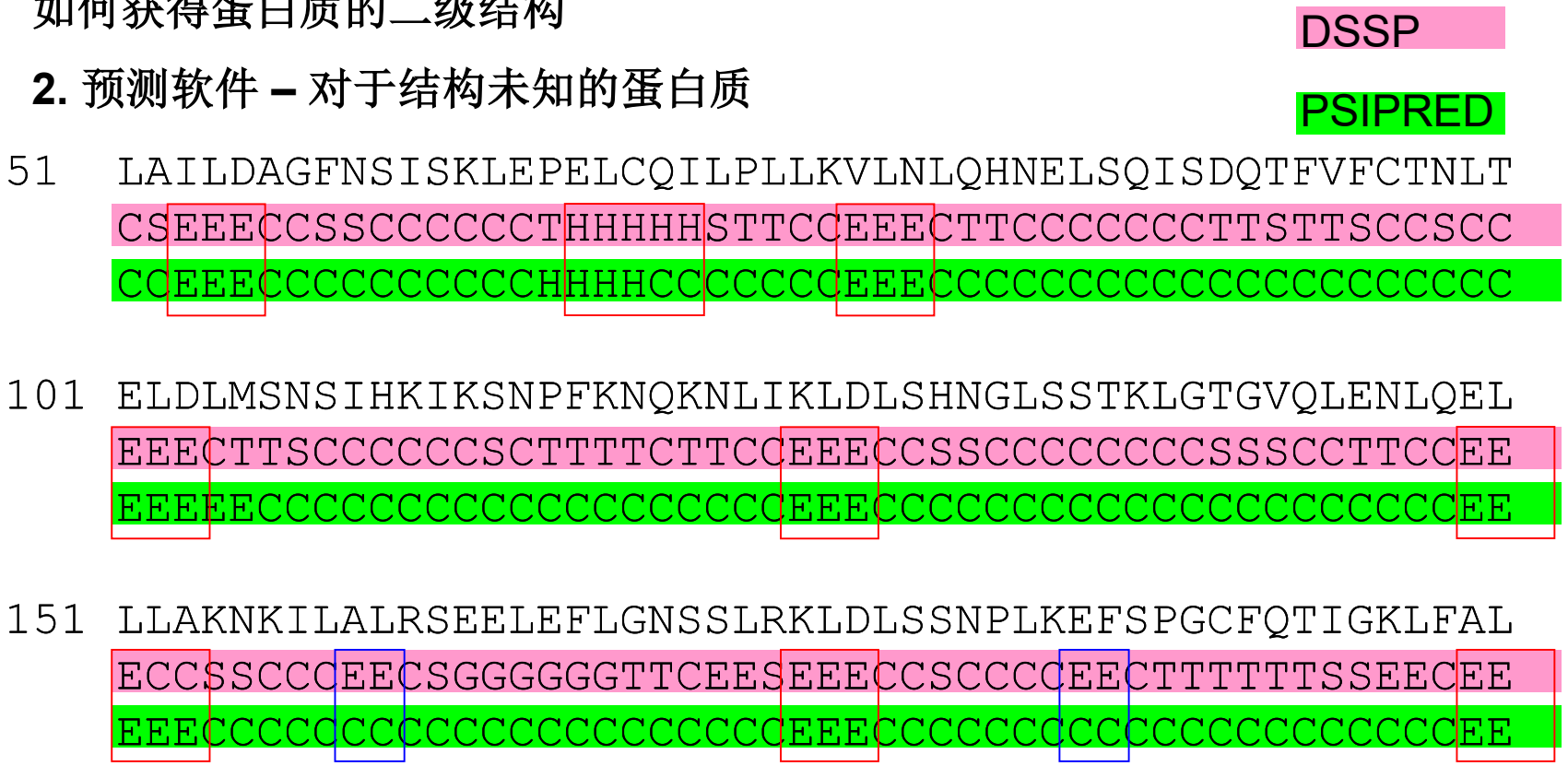
Subject: Psipred results for job ID:305293/3cig
From: psipred@cs.ucl.ac.uk
Date: Sat, October 1, 2011 5:49 am
To: gongj@informatik.uni-muenchen.de
Priority: Normal
Signature: Unsigned
Options: View Full Header | View Printable Version |

```
# PSIPRED HFORMAT (PSIPRED V3.0)

Conf: 983220479846899986898858697588822786998999888689889
Pred: CCCEEECEEECCCCCCCCCCCCCCCCCCCEEECCCCCCCCCCCCCCCCCCCC
  AA: QCTVRYNVADCSHLKLTHIPDDLPSNITVLNLTHNQLRRLPPTNFTRYSQLAILDAGFNS
              10        20        30        40        50        60

Conf: 996793342499658588722884888891014799988999840654884599888888
Pred: CCCCCHHHHCCCCCCCEEECCCCCCCCCCCCCCCCCCCCCCCEEEEECCCCCCCCCCCCCC
  AA: ISKLEPELCQILPLLKVLNLQHNELSQISDQTFVFCTNLTELDLMSNSIHKIKSNPFKNQ
              70        80        90       100       110       120

Conf: 777679677883886899979899998989810886986786821237999746999538
Pred: CCCCEEECCCCCCCCCCCCCCCCCCCCCCEEEEECCCCCCCCCCCCCCCCCCCCCEEECCC
  AA: KNLIKLDLSHNGLSSTKLGTGVQLENLQELLLAKNKILALRSEELEFLGNSSLRKLDLSS
             130       140       150       160       170       180

Conf: 98988899636789988699876887990225765201699985979922783897892 1
Pred: CCCCCCCCCCCCCCCCCCCEEECCCCCCCCHHHHHHHHHHCCCCCCCCCEEEEECCCCCCCCC
  AA: NPLKEFSPGCFQTIGKLFALLLNNAQLNPHLTEKLCWELSNTSIQNLSLANNQLLATSES
             190       200       210       220       230       240
```

蛋白质结构 --- 二级结构获取

如何获得蛋白质的二级结构

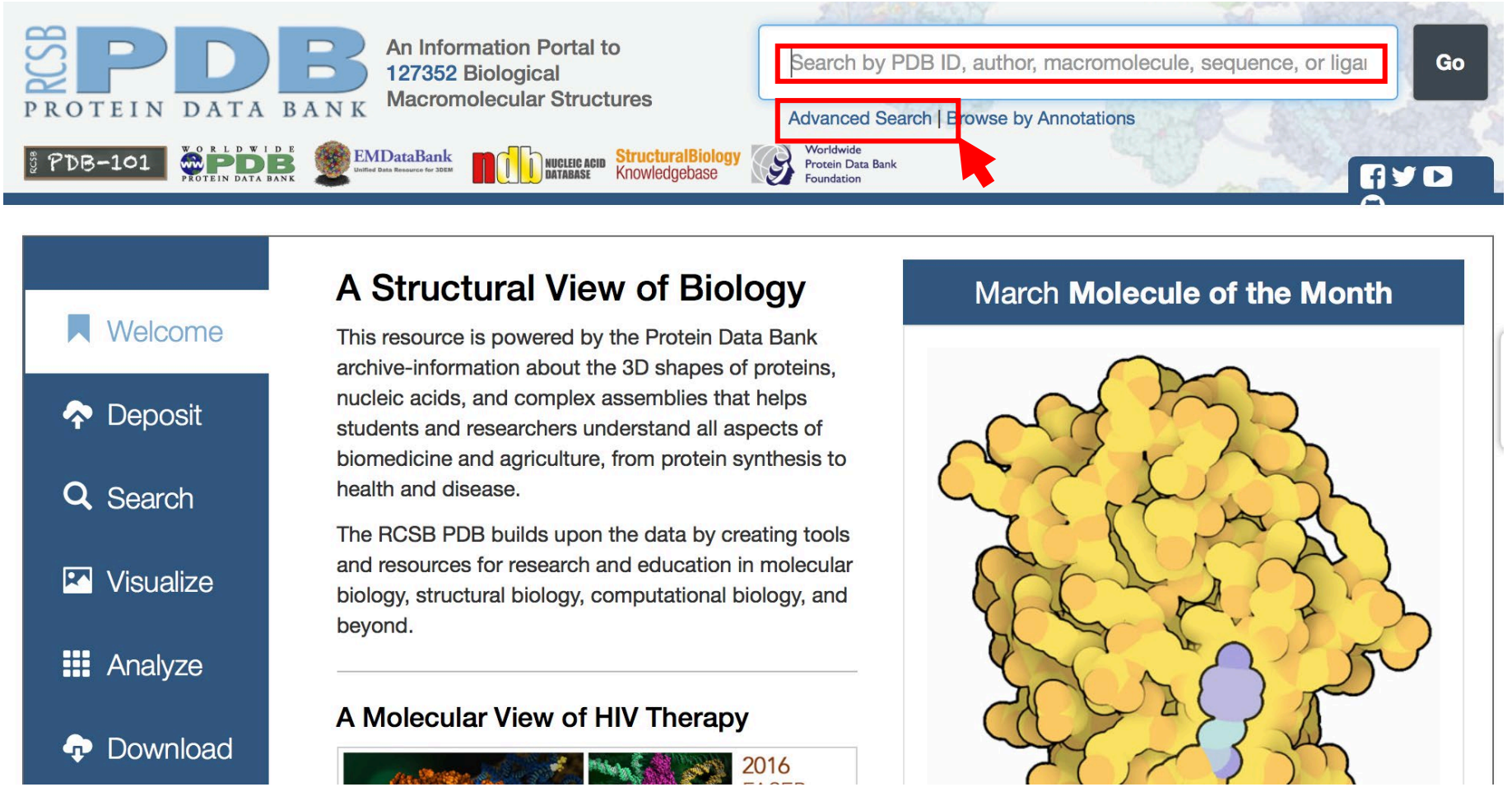
2. 预测软件 **–** 对于结构未知的蛋白质



准确率超过90%

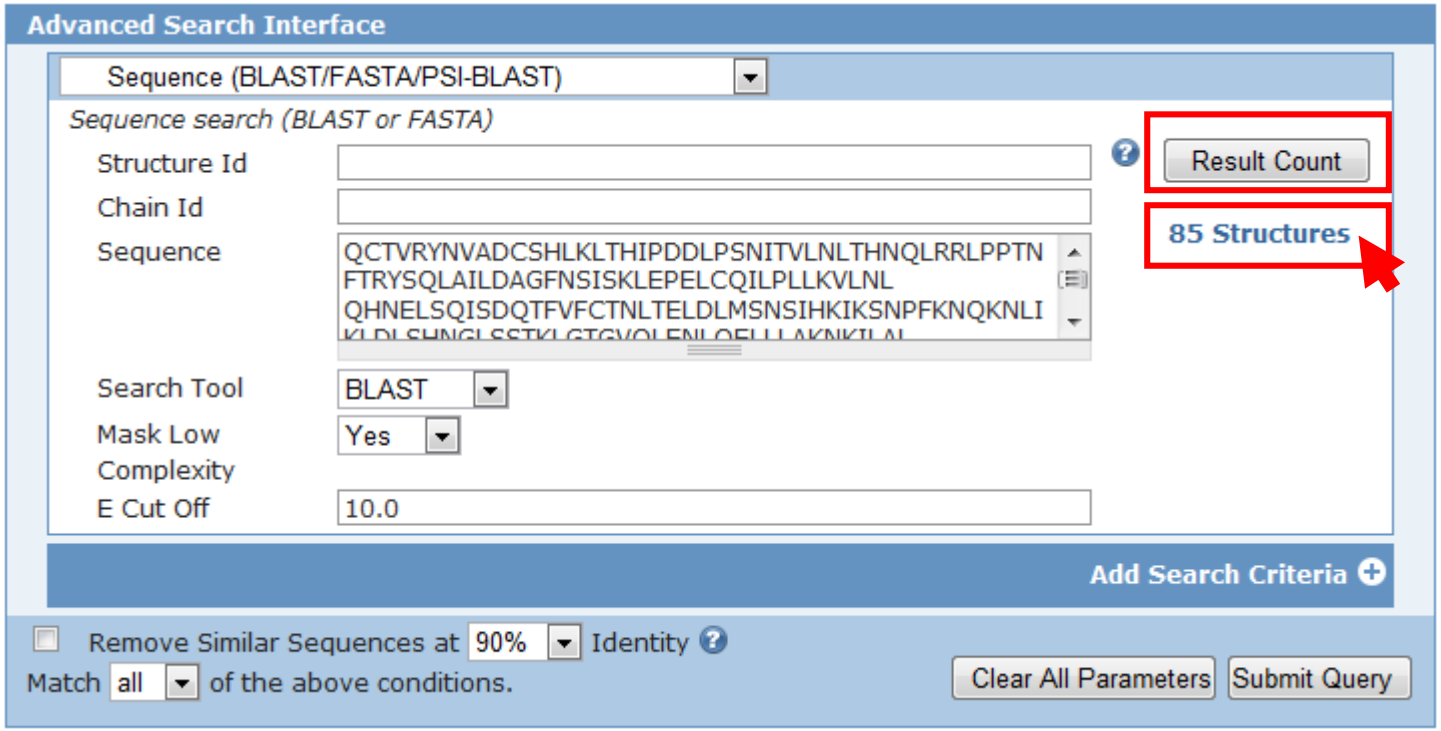
蛋白质结构 --- 三级结构

蛋白质结构数据库**PDB** **https://www.rcsb.org**



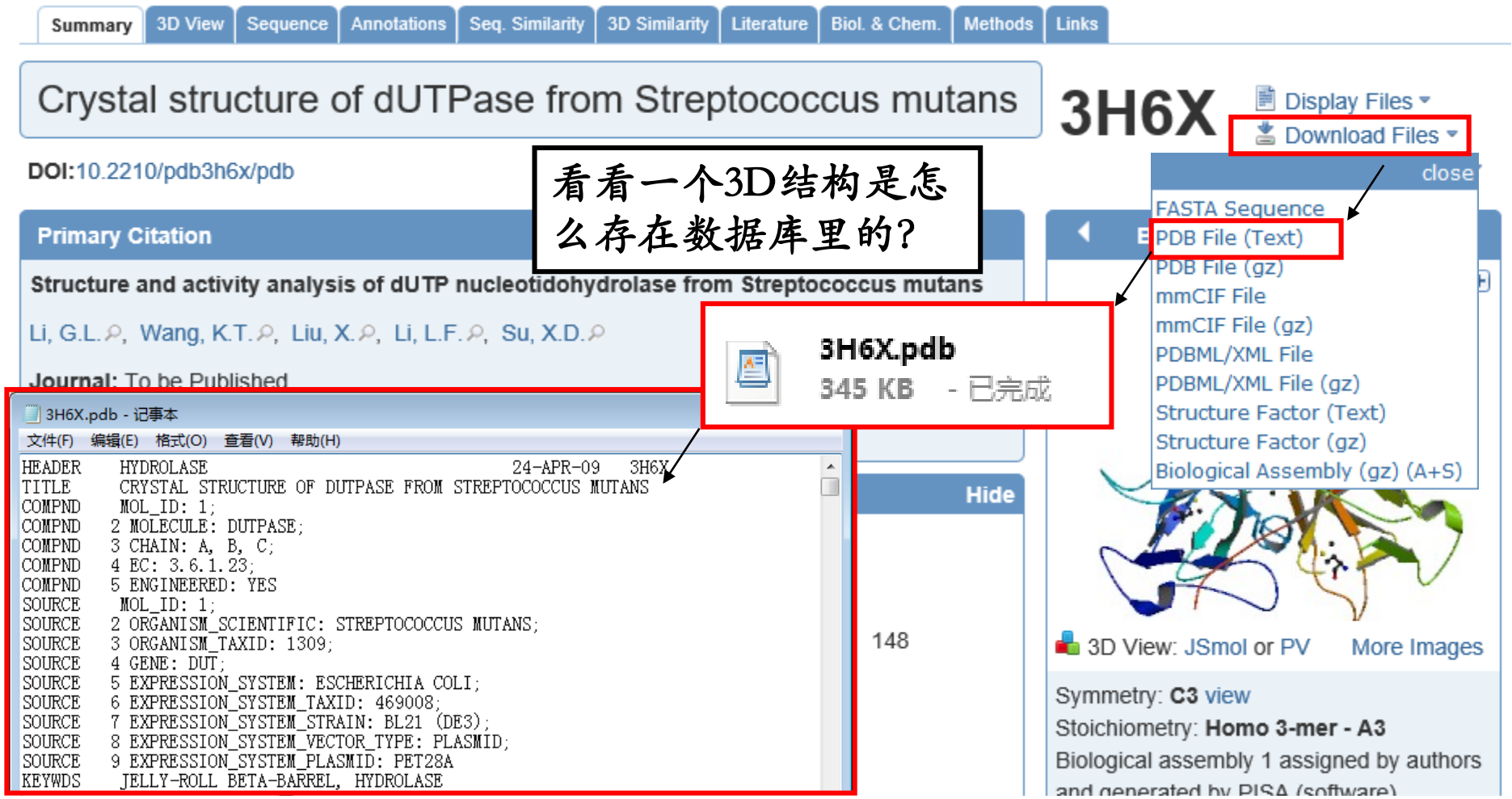
蛋白质结构 --- 三级结构

蛋白质结构数据库**PDB** **https://www.rcsb.org**



pdb.fasta

浏览dUTPase蛋白质的三级结构



浏览dUTPase蛋白质的三级结构

```
HEADER    HYDROLASE                                   24-APR-09   3H6X
TITLE     CRYSTAL STRUCTURE OF DUTPASE FROM STREPTOCOCCUS MUTANS
COMPND    MOL_ID: 1;
......
COMPND   5 ENGINEERED: YES
SOURCE    MOL_ID: 1;
......
SOURCE   9 EXPRESSION_SYSTEM_PLASMID: PE
KEYWDS    JELLY-ROLL BETA-BARREL, HYDROL
EXPDTA    X-RAY DIFFRACTION
AUTHOR    G.L.LI,K.T.WANG,X.LIU,L.F.LI,X
REVDAT   1   05-MAY-10 3H6X    0
JRNL        AUTH   G.L.LI,K.T.WANG,X.LIU
......
JRNL        REFN
REMARK   2
REMARK   2 RESOLUTION.    1.70 ANGSTROMS.
REMARK   3
REMARK   3 REFINEMENT.
```

基本信息部分

这部分对蛋白质进行了基本描述，包括以下索引词：

HEADER, OBSLTE, TITLE, SPLIT, CAVEAT, COMPND, SOURCE, KEYWDS, EXPDTA, AUTHOR, REVDAT, SPRSDE, JRNL, REMARK.

浏览dUTPase蛋白质的三级结构

```
HEADER    HYDROLASE                                24-APR-09   3H6X
TITLE     CRYSTAL STRUCTURE OF DUTPASE FROM STREPTOCOCCUS MUTANS
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: DUTPASE;
COMPND   3 CHAIN: A, B, C;
COMPND   4 EC: 3.6.1.23;
COMPND   5 ENGINEERED: YES
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: STREPTOCOCCUS MUTANS;
SOURCE   3 ORGANISM_TAXID: 1309;
SOURCE   4 GENE: DUT;
SOURCE   5 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE   6 EXPRESSION_SYSTEM_TAXID: 469008;
SOURCE   7 EXPRESSION_SYSTEM_STRAIN: BL21 (DE3);
SOURCE   8 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE   9 EXPRESSION_SYSTEM_PLASMID: PET28A
KEYWDS    JELLY-ROLL BETA-BARREL, HYDROLASE
```

HEADER ：分子类别，日期，数据库编号

浏览dUTPase蛋白质的三级结构

```
HEADER    HYDROLASE                                24-APR-09   3H6X
TITLE     CRYSTAL STRUCTURE OF DUTPASE FROM STREPTOCOCCUS MUTANS
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: DUTPASE;
COMPND   3 CHAIN: A, B, C;
COMPND   4 EC: 3.6.1.23;
COMPND   5 ENGINEERED: YES
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: STREPTOCOCCUS MUTANS;
SOURCE   3 ORGANISM_TAXID: 1309;
SOURCE   4 GENE: DUT;
SOURCE   5 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE   6 EXPRESSION_SYSTEM_TAXID: 469008;
SOURCE   7 EXPRESSION_SYSTEM_STRAIN: BL21 (DE3);
SOURCE   8 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE   9 EXPRESSION_SYSTEM_PLASMID: PET28A
KEYWDS    JELLY-ROLL BETA-BARREL, HYDROLASE
```

TITLE：题目，一般就是相关文献的题目。

浏览dUTPase蛋白质的三级结构

```
HEADER    HYDROLASE                                24-APR-09   3H6X
TITLE     CRYSTAL STRUCTURE OF DUTPASE FROM STREPTOCOCCUS MUTANS
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: DUTPASE;
COMPND   3 CHAIN: A, B, C;
COMPND   4 EC: 3.6.1.23;
COMPND   5 ENGINEERED: YES
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: STREPTOCOCCUS MUTANS;
SOURCE   3 ORGANISM_TAXID: 1309;
SOURCE   4 GENE: DUT;
SOURCE   5 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE   6 EXPRESSION_SYSTEM_TAXID: 469008;
SOURCE   7 EXPRESSION_SYSTEM_STRAIN: BL21 (DE3);
SOURCE   8 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE   9 EXPRESSION_SYSTEM_PLASMID: PET28A
KEYWDS    JELLY-ROLL BETA-BARREL, HYDROLASE
```

COMPND：对各个分子的描述。

浏览dUTPase蛋白质的三级结构

```
HEADER    HYDROLASE                                24-APR-09   3H6X
TITLE     CRYSTAL STRUCTURE OF DUTPASE FROM STREPTOCOCCUS MUTANS
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: DUTPASE;
COMPND   3 CHAIN: A, B, C;
COMPND   4 EC: 3.6.1.23;
COMPND   5 ENGINEERED: YES
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: STREPTOCOCCUS MUTANS;
SOURCE   3 ORGANISM_TAXID: 1309;
SOURCE   4 GENE: DUT;
SOURCE   5 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE   6 EXPRESSION_SYSTEM_TAXID: 469008;
SOURCE   7 EXPRESSION_SYSTEM_STRAIN: BL21 (DE3);
SOURCE   8 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE   9 EXPRESSION_SYSTEM_PLASMID: PET28A
KEYWDS    JELLY-ROLL BETA-BARREL, HYDROLASE
```

SOURCE：结构中包括的每一个分子的来源（生物学/化学）。

浏览dUTPase蛋白质的三级结构

```
HEADER    HYDROLASE                                24-APR-09   3H6X
TITLE     CRYSTAL STRUCTURE OF DUTPASE FROM STREPTOCOCCUS MUTANS
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: DUTPASE;
COMPND   3 CHAIN: A, B, C;
COMPND   4 EC: 3.6.1.23;
COMPND   5 ENGINEERED: YES
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: STREPTOCOCCUS MUTANS;
SOURCE   3 ORGANISM_TAXID: 1309;
SOURCE   4 GENE: DUT;
SOURCE   5 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE   6 EXPRESSION_SYSTEM_TAXID: 469008;
SOURCE   7 EXPRESSION_SYSTEM_STRAIN: BL21 (DE3);
SOURCE   8 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE   9 EXPRESSION_SYSTEM_PLASMID: PET28A
KEYWDS    JELLY-ROLL BETA-BARREL, HYDROLASE
```

KEYWDS：一系列关键词，可用于数据库搜索。

浏览dUTPase蛋白质的三级结构

```
SOURCE    8 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE    9 EXPRESSION_SYSTEM_PLASMID: PET28A
KEYWDS     JELLY-ROLL BETA-BARREL, HYDROLASE
EXPDTA     X-RAY DIFFRACTION
AUTHOR     G.L.LI,K.T.WANG,X.LIU,L.F.LI,X.D.SU
REVDAT    1    05-MAY-10 3H6X     0
JRNL         AUTH    G.
JRNL         TITL    ST                                                  HYDROLASE
JRNL         TITL 2 FR
JRNL         REF     TO
JRNL         REFN
REMARK    2
REMARK    2 RESOLUTION
REMARK    3
REMARK    3 REFINEMENT
REMARK    3   PROGRAM
REMARK    3   AUTHORS
```

EXPDTA：测定结构所用的实验方法：

X-RAY DIFFRACTION X-射线衍射
FIBER DIFFRACTION 光纤点衍射
NEUTRON DIFFRACTION 中子粉末衍射
ELECTRON CRYSTALLOGRAPHY 电子晶体学
ELECTRON MICROSCOPY 电子显微镜
SOLID-STATE NMR 固态核磁共振
SOLUTION NMR 液态核磁共振
SOLUTION SCATTERING 液态散射

浏览dUTPase蛋白质的三级结构

```
SOURCE    8 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE    9 EXPRESSION_SYSTEM_PLASMID: PET28A
KEYWDS     JELLY-ROLL BETA-BARREL, HYDROLASE
EXPDTA     X-RAY DIFFRACTION
AUTHOR     G.L.LI,K.T.WANG,X.LIU,L.F.LI,X.D.SU
REVDAT    1    05-MAY-10 3H6X     0
JRNL         AUTH    G.L.LI,K.T.WANG,X.LIU,L.F.LI,X.D.SU
JRNL         TITL    STRUCTURE AND ACTIVITY ANALYSIS OF DUTP NUCLEOTIDOHYDROLASE
JRNL         TITL 2 FROM STREPTOCOCCUS MUTANS
JRNL         REF     TO BE PUBLISHED
JRNL         REFN
REMARK    2
REMARK    2 RESOLUTION.     1.70 ANGSTROMS.
REMARK    3
REMARK    3 REFINEMENT.
REMARK    3   PROGRAM      : PHENIX (PHENIX.REFINE)
REMARK    3   AUTHORS      : PAUL ADAMS,PAVEL AFONINE,VICENT CHEN,IAN
```

AUTHOR：结构测定者

浏览dUTPase蛋白质的三级结构

```
SOURCE    8 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE    9 EXPRESSION_SYSTEM_PLASMID: PET28A
KEYWDS     JELLY-ROLL BETA-BARREL, HYDROLASE
EXPDTA     X-RAY DIFFRACTION
AUTHOR     G.L.LI,K.T.WANG,X.LIU,L.F.LI,X.D.SU
REVDAT    1    05-MAY-10 3H6X     0
JRNL         AUTH    G.L.LI,K.T.WANG,X.LIU,L.F.LI,X.D.SU
JRNL         TITL    STRUCTURE AND ACTIVITY ANALYSIS OF DUTP NUCLEOTIDOHYDROLASE
JRNL         TITL 2 FROM STREPTOCOCCUS MUTANS
JRNL         REF     TO BE PUBLISHED
JRNL         REFN
REMARK    2
REMARK    2 RESOLUTION.     1.70 ANGSTROMS.
REMARK    3
REMARK    3 REFINEMENT.
REMARK    3   PROGRAM      : PHENIX (PHENIX.REFINE)
REMARK    3   AUTHORS      : PAUL ADAMS,PAVEL AFONINE,VICENT CHEN,IAN
```

REVDAT：历史上曾经对该数据库记录进行过的修改。

浏览dUTPase蛋白质的三级结构

```
SOURCE    8 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE    9 EXPRESSION_SYSTEM_PLASMID: PET28A
KEYWDS     JELLY-ROLL BETA-BARREL, HYDROLASE
EXPDTA     X-RAY DIFFRACTION
AUTHOR     G. L. LI, K. T. WANG, X. LIU, L. F. LI, X. D. SU
REVDAT    1    05-MAY-10 3H6X     0
JRNL         AUTH    G. L. LI, K. T. WANG, X. LIU, L. F. LI, X. D. SU
JRNL         TITL    STRUCTURE AND ACTIVITY ANALYSIS OF DUTP NUCLEOTIDOHYDROLASE
JRNL         TITL 2 FROM STREPTOCOCCUS MUTANS
JRNL         REF     TO BE PUBLISHED
JRNL         REFN
REMARK    2
REMARK    2 RESOLUTION.     1.70 ANGSTROMS.
REMARK    3
REMARK    3 REFINEMENT.
REMARK    3   PROGRAM      : PHENIX (PHENIX.REFINE)
REMARK    3   AUTHORS      : PAUL ADAMS, PAVEL AFONINE, VICENT CHEN, IAN
```

JRNL：发表这个结构的文献信息。

浏览dUTPase蛋白质的三级结构

REMARK： 注释，关于这个蛋白质的各种注释信息，这些信息无法归入其他部分，所以就都放这里了。

```
SOURCE    8 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE    9 EXPRESSION_SYSTEM_PLASMID: PET28A
KEYWDS     JELLY-ROLL BETA-BARREL, HYDROLASE
EXPDTA     X-RAY DIFFRACTION
AUTHOR     G.L.LI,K.T.WANG,X.LIU,L.F.LI,X.D.SU
REVDAT    1    05-MAY-10 3H6X     0
JRNL         AUTH    G.L.LI,K.T.WANG,X.LIU,L.F.LI,X.D.SU
JRNL         TITL    STRUCTURE AND ACTIVITY ANALYSIS OF DUTP NUCLEOTIDOHYDROLASE
JRNL         TITL 2 FROM STREPTOCOCCUS MUTANS
JRNL         REF     TO BE PUBLISHED
JRNL         REFN
REMARK    2
REMARK    2 RESOLUTION.     1.70 ANGSTROMS.
REMARK    3
REMARK    3 REFINEMENT.
REMARK    3   PROGRAM      : PHENIX (PHENIX.REFINE)
REMARK    3   AUTHORS      : PAUL ADAMS,PAVEL AFONINE,VICENT CHEN,IAN
```

浏览dUTPase蛋白质的三级结构

```
REMARK 525    HOH A 616        DISTANCE =  5.29 ANGSTROMS
REMARK 525    HOH A 677        DISTANCE =  5.20 ANGSTROMS
DBREF  3H6X A    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
DBREF  3H6X B    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
DBREF  3H6X C    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
SEQRES   1 A  148  MSE MSE LYS THR ARG GLY PHE GLU LEU ILE THR ASP TYR
SEQRES   2 A  148  THR ASP GLU ASN LEU LEU PRO LYS ARG GLU THR ALA HIS
SEQRES   3 A  148  ALA ALA GLY TYR ASP LEU LYS
SEQRES   4 A  148  ILE SER ALA GLY ALA ILE VAL
......
SEQRES   9 C  148  ASN MSE THR ASP GLN THR VAL
SEQRES  10 C  148  ARG VAL VAL GLN GLY VAL PHE
SEQRES  11 C  148  ASP GLY ASP LYS ALA THR GLY
SEQRES  12 C  148  GLY SER THR GLY GLY
MODRES 3H6X MSE A   56  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  103  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  106  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  125  MET  SELENOMETHIONINE
......
```

一级结构信息部分

这部分包括大分子（蛋白质）各个残基的信息，包括以下索引词：

DBREF, SEQADV, SEQRES, MODRES

浏览dUTPase蛋白质的三级结构

```
REMARK 525    HOH A 616        DISTANCE =  5.29 ANGSTROMS
REMARK 525    HOH A 677        DISTANCE =  5.20 ANGSTROMS
DBREF  3H6X A    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
DBREF  3H6X B    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
DBREF  3H6X C    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
SEQRES   1 A  148  MSE MSE LYS THR ARG GLY PHE GLU LEU ILE THR ASP TYR
SEQRES   2 A  148  THR ASP GLU ASN LEU LEU PRO LYS ARG GLU THR ALA HIS
SEQRES   3 A  148  ALA ALA GLY TYR ASP LEU LYS                      LU
SEQRES   4 A  148  ILE SER ALA GLY ALA ILE VAL                      AL
......
SEQRES   9 C  148  ASN MSE THR ASP GLN THR VAL                      LU
SEQRES  10 C  148  ARG VAL VAL GLN GLY VAL PHE MSE PRO PHE LEU LEU ILE
SEQRES  11 C  148  ASP GLY ASP LYS ALA THR GLY THR ARG THR GLY GLY PHE
SEQRES  12 C  148  GLY SER THR GLY GLY
MODRES 3H6X MSE A   56  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  103  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  106  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  125  MET  SELENOMETHIONINE
......
```

DBREF：其他蛋白质序列数据库的有关记录。

浏览dUTPase蛋白质的三级结构

```
REMARK 525    HOH A 616        DISTANCE =  5.29 ANGSTROMS
REMARK 525    HOH A 677        DISTANCE =  5.20 ANGSTROMS
DBREF  3H6X A    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
DBREF  3H6X B    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
DBREF  3H6X C    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
SEQRES   1 A  148  MSE MSE LYS THR ARG GLY PHE GLU LEU ILE THR ASP TYR
SEQRES   2 A  148  THR ASP GLU ASN LEU LEU PRO LYS ARG GLU THR ALA HIS
SEQRES   3 A  148  ALA ALA GLY TYR ASP LEU LYS VAL ALA GLU ARG THR GLU
SEQRES   4 A  148  ILE SER ALA GLY ALA ILE VAL LEU VAL PRO THR GLY VAL
......
SEQRES   9 C  148  ASN MSE THR ASP GLN THR VAL VAL LEU GLU ALA GLY GLU
SEQRES  10 C  148  ARG VAL VAL GLN GLY VAL PHE MSE PRO PHE LEU LEU ILE
SEQRES  11 C  148  ASP GLY ASP LYS ALA THR GLY THR ARG THR GLY GLY PHE
SEQRES  12 C  148  GLY SER THR GLY GLY
MODRES 3H6X MSE A   56  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  103  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  106  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  125  MET  SELENOMETHIONINE
......
```

SEQRES：氨基酸序列。

浏览dUTPase蛋白质的三级结构

```
REMARK 525    HOH A 616        DISTANCE =  5.29 ANGSTROMS
REMARK 525    HOH A 677        DISTANCE =  5.20 ANGSTROMS
DBREF  3H6X A    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
DBREF  3H6X B    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
DBREF  3H6X C    1   148  UNP    Q8DVY3   Q8DVY3_STRMU     1    148
SEQRES   1 A  148  MSE MSE LYS THR ARG GLY PHE GLU LEU ILE THR ASP TYR
SEQRES   2 A  148  THR ASP GLU ASN LEU LEU PRO LYS ARG GLU THR ALA HIS
SEQRES   3 A  148  ALA ALA GLY TYR ASP LEU LYS VAL ALA GLU ARG THR GLU
SEQRES   4 A  148  ILE SER ALA GLY ALA ILE VAL LEU VAL PRO THR GLY VAL
......
SEQRES   9 C  148  ASN MSE THR ASP GLN THR VAL VAL LEU GLU ALA GLY GLU
SEQRES  10 C  148  ARG VAL VAL GL
SEQRES  11 C  148  ASP GLY ASP LYS ALA THR GLY THR ARG THR GLY GLY PHE
SEQRES  12 C  148  GLY SER THR GLY GLY
MODRES 3H6X MSE A   56  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  103  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  106  MET  SELENOMETHIONINE
MODRES 3H6X MSE A  125  MET  SELENOMETHIONINE
......
```

MODRES：对标准残基上的修饰。

蛋氨酸 ⟶ 硒代蛋氨酸

浏览dUTPase蛋白质的三级结构

```
HET    MSE  B  56       8
HET    MSE  B 103       8
HET    MSE  B 106       8
HET    MSE  B 125       8
HET    MSE  C  56       8
HET    MSE  C 103       8
HET    MSE  C 106       8
HET    MSE  C 125       8
HETNAM     MSE SELENOMETHIONINE
FORMUL   1  MSE    12(C5 H11 N 02 SE)
FORMUL   4  HOH   *688(H2 O)
```

非标准残基部分，包括以下索引词：
HET：非标准残基及位置
HETNAM：非标准残基的化学名称
FORMUL：非标准残基的化学式

浏览dUTPase蛋白质的三级结构

二级结构部分，
包括以下索引词：
HELIX：螺旋
SHEET：折片
TURN：转角

```
HELIX    1   1 ASP A   15  LEU A   19  5                                    5
HELIX    2   2 SER A   68  GLY A   75  1                                    8
HELIX    3   3 ASP B   15  LEU B   19  5                                    5
HELIX    4   4 SER B   68  GLY B   75  1                                    8
HELIX    5   5 ASP C   15  LEU C   19  5                                    5
HELIX    6   6 SER C   68  GLY C   75  1                                    8
SHEET    1   A 2 GLY A   6  LEU A   9  0
SHEET    2   A 2 VAL A  52  TYR A  55 -1  O  LYS A  53   N  GLU A   8
SHEET    1   B 4 TYR A  30  LYS A  33  0
SHEET    2   B 4 ARG A 118  PRO A 126 -1  O  GLY A 122   N  TYR A  30
SHEET    3   B 4 GLU A  60  ASP A  66 -1  N  PHE A  65   O  GLN A 121
SHEET    4   B 4 VAL A  82  ILE A  85 -1  O  ILE A  85   N  LEU A  62
SHEET    1   C 2 THR A  38  ILE A  40  0
SHEET    2   C 2 VAL A 111  LEU A 113 -1  O  LEU A 113   N  THR A  38
SHEET    1   D 3 ILE A  45  PRO A  49  0
```

浏览dUTPase蛋白质的三级结构

LINK：残基间化学键，比如氢键，二硫键等。

```
LINK         C   MSE A 106                 N   THR A 107     1555   1555  1.32
LINK         C   PHE A 124                 N   MSE A 125     1555   1555  1.33
LINK         C   MSE A 125                 N   PRO A 126     1555   1555  1.34
LINK         C   TYR B  55                  N   MSE B  56     1555   1555  1.33
LINK         C   MSE B  56                  N   GLN B  57     1555   1555  1.33
LINK         C   GLN B 102                 N   MSE B 103     1555   1555  1.33
LINK         C   MSE B 103                 N   LYS B 104     1555   1555  1.33
LINK         C   ASN B 105                 N   MSE B 106     1555   1555  1.34
LINK         C   MSE B 106                 N   THR B 107     1555   1555  1.33
LINK         C   PHE B 124                 N   MSE B 125     1555   1555  1.34
LINK         C   MSE B 125                 N   PRO B 126     1555   1555  1.34
LINK         C   TYR C  55                  N   MSE C  56     1555   1555  1.33
LINK         C   MSE C  56                  N   GLN C  57     1555   1555  1.33
LINK         C   GLN C 102                 N   MSE C 103     1555   1555  1.33
LINK         C   MSE C 103                 N   LYS C 104     1555   1555  1.33
```

ATOM：原子坐标

浏览dUTPase蛋白质的三级结构

	原子号	原子名	残基名	分子链	残基号	3D坐标 X轴	3D坐标 Y轴	3D坐标 Z轴	占有率	温度因子	元素符号
ATOM	134	N	PRO	A	20	6.147	-10.140	21.368	1.00	10.34	N
ATOM	135	CA	PRO	A	20	7.611	-10.161	21.225	1.00	10.95	C
ATOM	136	C	PRO	A	20	8.320	-11.125	22.163	1.00	11.53	C
ATOM	137	O	PRO	A	20	7.895	-11.335	23.310	1.00	11.87	O
ATOM	138	CB	PRO	A	20	8.022	-8.719	21.567	1.00	9.74	C
ATOM	139	CG	PRO	A	20	6.790	-7.890	21.271	1.00	10.92	C
ATOM	140	CD	PRO	A	20	5.638	-8.787	21.670	1.00	7.65	C
ATOM	141	N	LYS	A	21	9.412	-11.701	21.670	1.00	11.67	N
ATOM	142	CA	LYS	A	21	10.297	-12.480	22.521	1.00	12.81	C
ATOM	143	C	LYS	A	21	11.732	-12.080	22.223	1.00	11.21	C
ATOM	144	O	LYS	A	21	12.020	-11.451	21.190	1.00	12.10	O
ATOM	145	CB	LYS	A	21	10.098	-13.978	22.289	1.00	16.83	C
ATOM	146	CG	LYS	A	21	10.533	-14.470	20.914	1.00	24.79	C
ATOM	147	CD	LYS	A	21	10.280	-15.970	20.773	1.00	30.83	C
ATOM	148	CE	LYS	A	21	10.625	-16.476	19.371	1.00	31.96	C
ATOM	149	NZ	LYS	A	21	10.319	-17.928	19.240	1.00	44.03	N

浏览dUTPase蛋白质的三级结构

```
CONECT  390  400
CONECT  400  390  401
CONECT  401  400  402  404
CONECT  402  401  403  408
CONECT  403  402
CONECT  404  401  405
CONECT  405  404  406
CONECT  406  405  407
CONECT  407  406
CONECT  408  402
CONECT  765  772
CONECT  772  765  773
CONECT  773  772  774  776
CONECT  774  773  775  780
CONECT  775  774
```

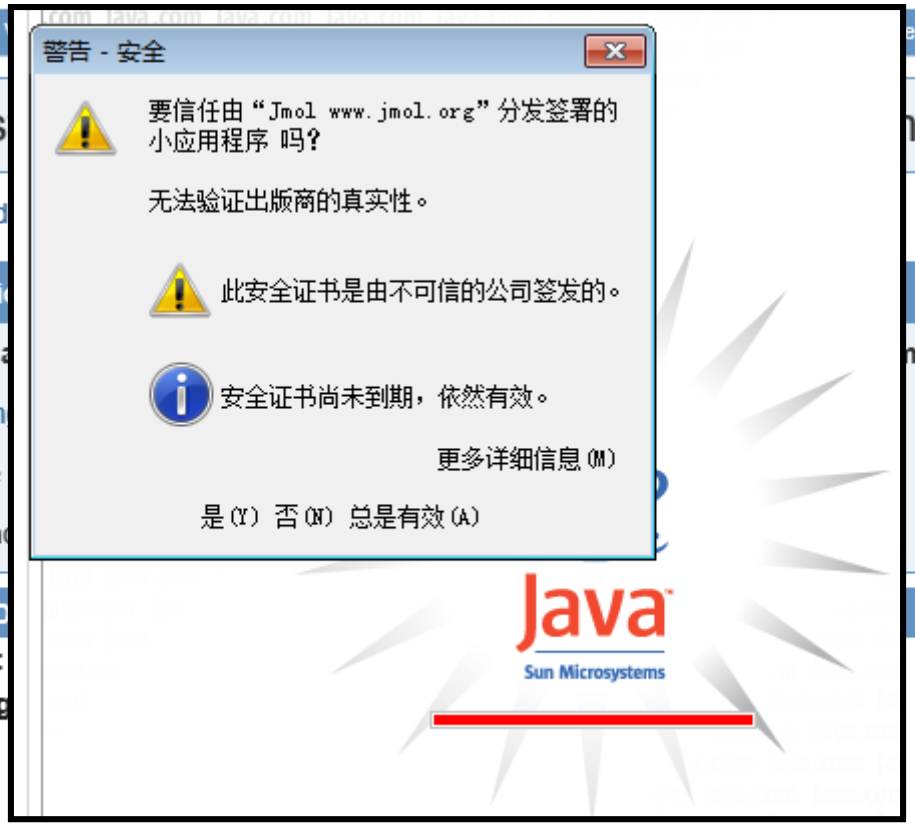
CONECT：原子间化学键连接

浏览dUTPase蛋白质的三级结构

```
CONECT 2940 2949
CONECT 2949 2940 2950
CONECT 2950 2949 2951 2953
CONECT 2951 2950 2952 2957
CONECT 2952 2951
CONECT 2953 2950 2954
CONECT 2954 2953 2955
CONECT 2955 2954 2956
CONECT 2956 2955
CONECT 2957 2951
MASTER      385    0   12    6   33    0    0    6 3696    3  120   36
END
```

档案管理部分
MASTER：版权拥有者
END：结束符

浏览dUTPase蛋白质的三级结构



Summary | 3D | Methods | Links

Crystal s ... mutans

DOI:10.2210/pd

3H6X

Display Files
Download Files
Download Citation

Primary Citati ... mutans

Structure and a

Li, G.L., Wan

Journal: To be

PubMed ID is n

Molecular D ... Hide

Classification:	
Structure Weig	
Molecule:	
Polymer:	
Chains:	A, B, C
EC#:	3.6.1.23
Organism:	Streptococcus mutans
Gene Name:	SMU_325

Biological Assembly

3D View: JSmol or PV More Images

Symmetry: **C3** view
Stoichiometry: **Homo 3-mer - A3**
Biological assembly 1 assigned by authors and generated by PISA (software)

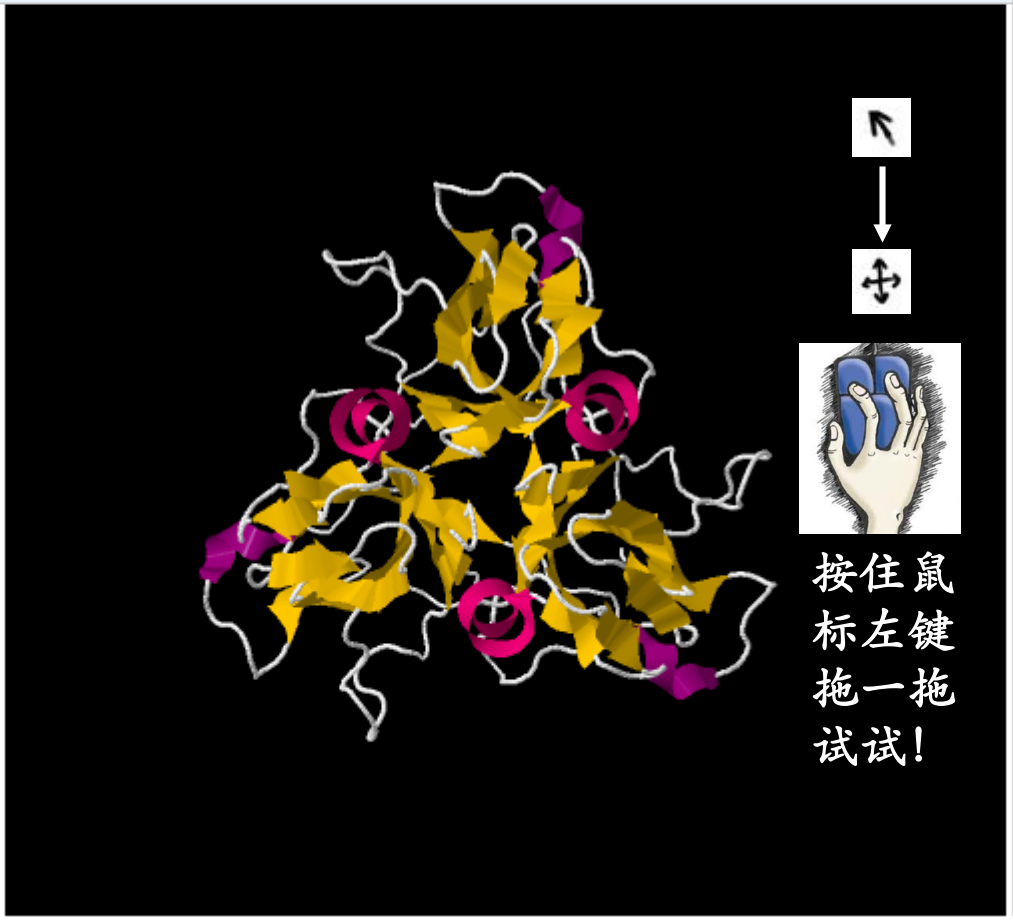
Summary | 3D View | Sequence | Annotations | Seq. Similarity | 3D Similarity | Literature | Biol. & Chem. | Methods | Geometry | Links

Crystal structure of dUTPase from Streptococcus mutans

3H6X

Display Files
Download Files
Share this Page

NOTE: Use your mouse to drag, rotate, and zoom in and out of the structure.



Structure Details

Structure	Biological Assembly 1
Symmetry Type	Global Symmetry
Symmetry	C3
Stoichiometry	A3

Select Orientation

Front C3 axis

Select Display Mode

Secondary Structure | Symmetry | Subunit | Custom View

Custom View options

Style	Cartoon
Color	Secondary Structure
Surface	None
Jmol mode	HTML5 (beta)

- [ ] H-Bonds
- [ ] SS Bonds
- [ ] Rotation
- [x] Black Background
- [ ] Polyhedron
- [ ] Axes

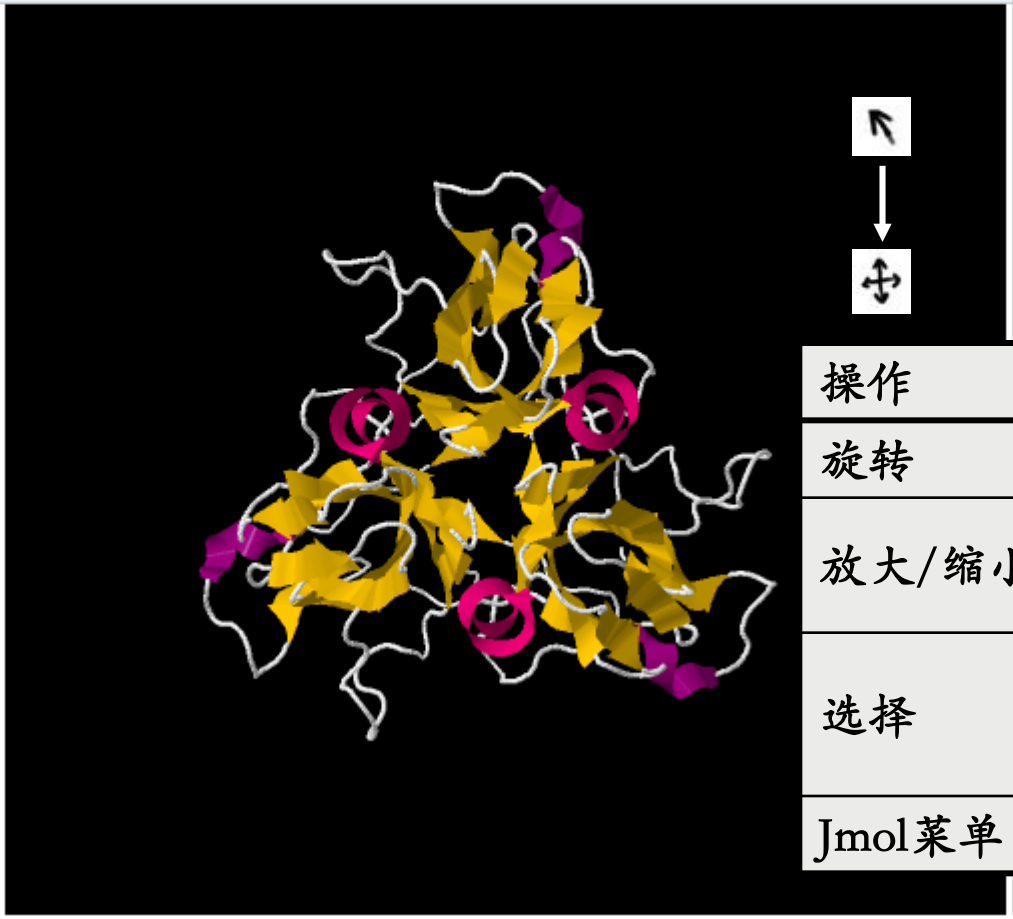
Summary | 3D View | Sequence | Annotations | Seq. Similarity | 3D Similarity | Literature | Biol. & Chem. | Methods | Geometry | Links

Crystal structure of dUTPase from Streptococcus mutans

3H6X

Display Files
Download Files
Share this Page

NOTE: Use your mouse to drag, rotate, and zoom in and out of the structure.



操作	
旋转	按住鼠标左键拖拽
放大/缩小	按住Shift + 鼠标左键拖拽 或者利用鼠标中键
选择	点击鼠标左键
Jmol菜单	点击鼠标右键

Structure Details

Structure: Biological Assembly 1
Symmetry Type: Global Symmetry
Symmetry: C3
Stoichiometry: A3

Select Orientation

Front C3 axis

Select Display Mode

Secondary Structure | Symmetry

Polyhedron | Axes

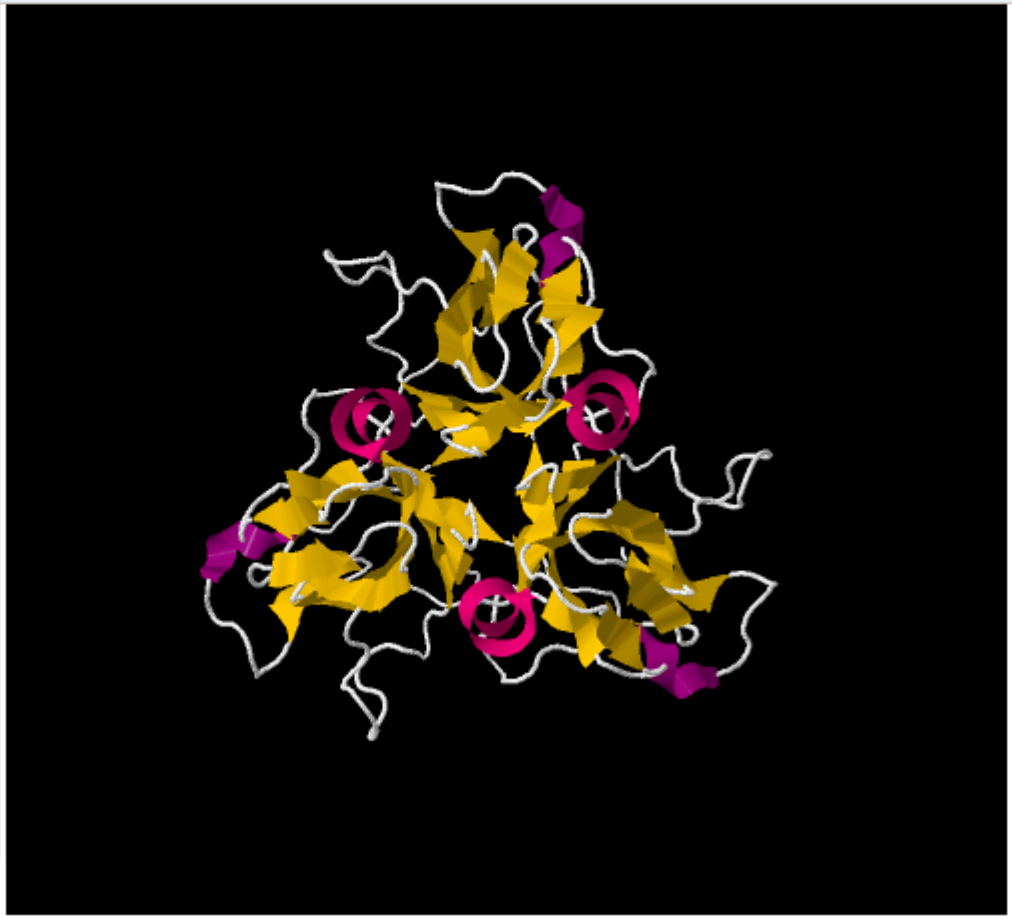
Summary | 3D View | Sequence | Annotations | Seq. Similarity | 3D Similarity | Literature | Biol. & Chem. | Methods | Geometry | Links

Crystal structure of dUTPase from Streptococcus mutans

3H6X

Display Files
Download Files
Share this Page

NOTE: Use your mouse to drag, rotate, and zoom in and out of the structure.



Structure Details

Structure	Biological Assembly 1
Symmetry Type	Global Symmetry
Symmetry	C3
Stoichiometry	A3

Select Orientation

◂◂ *Front C3 axis* ▸▸

Select Display Mode

Secondary Structure | Symmetry | Subunit | Custom View

Custom View options

Style	Cartoon
Color	Secondary Structure
Surface	None
Jmol mode	HTML5 (beta)

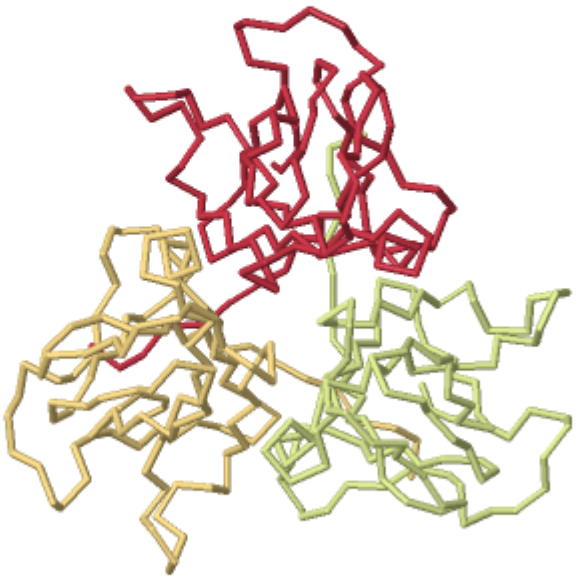
- ☐ H-Bonds
- ☐ SS Bonds
- ☐ Rotation
- ☑ Black Background
- ☐ Polyhedron
- ☐ Axes

Crystal structure of dUTPase from Streptococcus mutans

3H6X

Display Files
Download Files
Share this Page

NOTE: Use your mouse to drag, rotate, and zoom in and out of the structure.



Structure Details

Structure	Biological Assembly 1
Symmetry Type	Global Symmetry
Symmetry	C3
Stoichiometry	A3

Select Orientation

◂◂ *Front C3 axis* ▸▸

Select Display Mode

Secondary Structure | Symmetry | Subunit | Custom View

Custom View options

Style	Backbone
Color	Subunit
Surface	None
Jmol mode	HTML5 (beta)

☐ H-Bonds ☐ SS Bonds
☐ Rotation ☐ Black Background
☐ Polyhedron ☐ Axes