



# Small molecule data and tools at RCSB

John Westbrook  
Rutgers University

September 24, 2008



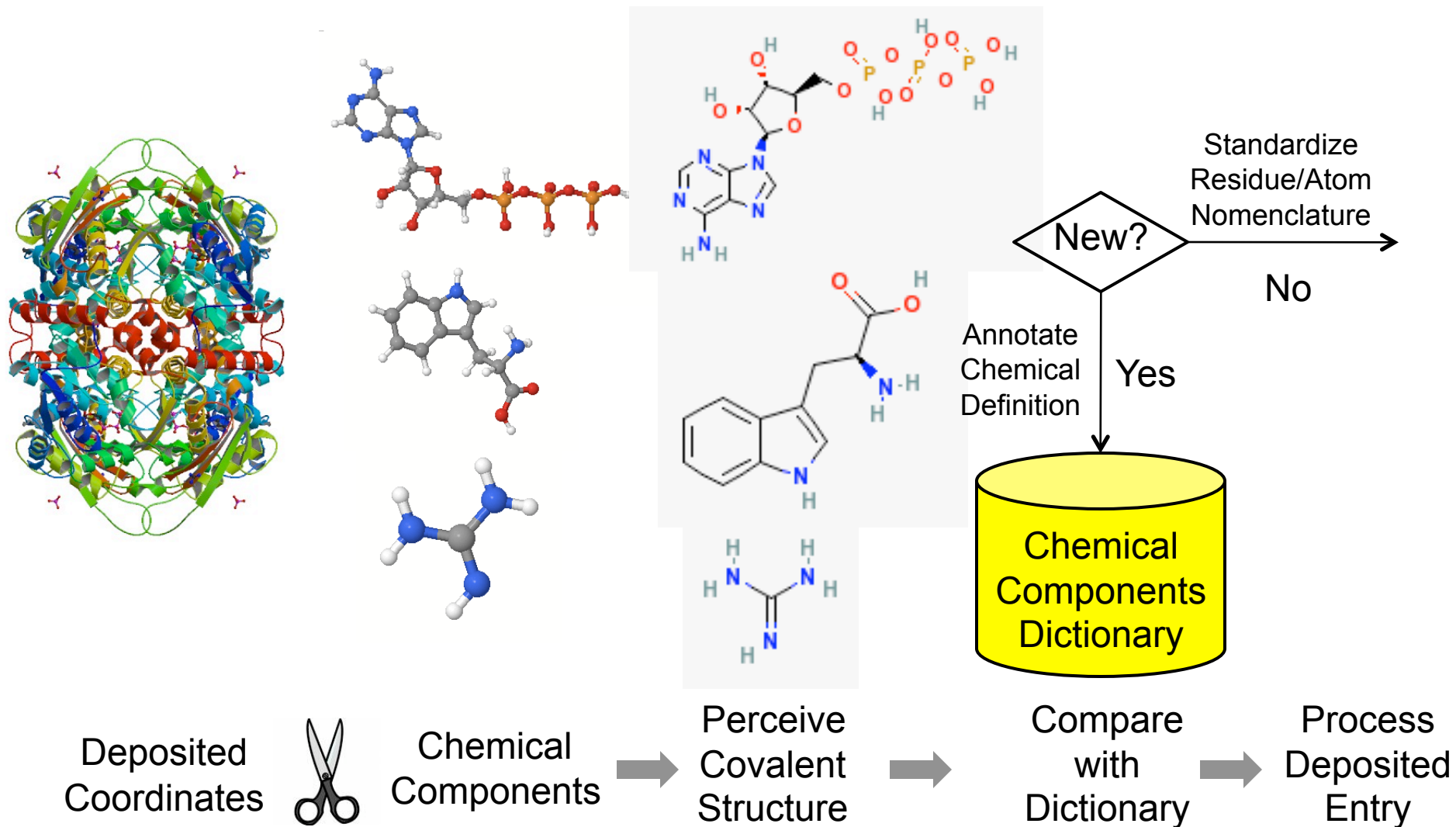
Computational Aspects of the Protein Target Selection, Protein  
Production Management and Structure Analysis Pipeline

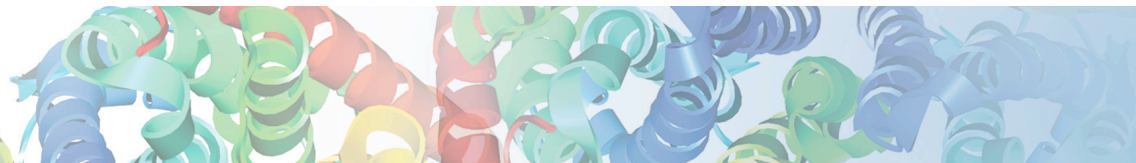


## Outline

- Chemical Components Dictionary
- Ligand Expo
- RCSB PDB website

# How does new chemistry enter PDB?





## Chemical Components Dictionary

- Library of all polymer and non-polymer chemical components in PDB (~8700).
- Organized by PDB chemical component 3-letter identifier.
- Used to resolve residue and ligand identity and to standardize residue, ligand and atom nomenclature in PDB data files.
- Update and maintenance of the dictionary is shared by the wwPDB deposition and annotation sites.

# Chemical annotation is a global effort

<http://www.wwpdb.org/>



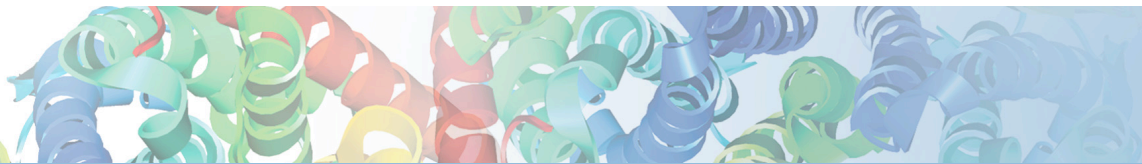
The **RCSB PDB** is managed by Rutgers, The State University of New Jersey and the University of California, San Diego.

It is supported by grants from the NSF, NIGMS, Office of Science-DOE, NLM, NCI, NCRR, NIBIB, NINDS, NIDDKD.

The **BMRB** (Madison, Wisconsin) is supported by NIH grant LM05799 from the NLM.

**PDBe/MSD** (Cambridge, UK) is supported by grants from the Wellcome Trust, the EU (TEMBLOR, NMRQUAL and IIMS), CCP4, the BBSRC, the MRC and EMBL.

**PDBj** (Osaka, Japan) is supported by grant-in-aid from the Institute for Bioinformatics Research and Development, Japan Science and Technology Agency (BIRD-JST), and the Ministry of Education, Culture, Sports, Science and Technology (MEXT).



# wwPDB Resources

<http://www.wwpdb.org/>



Home **wwPDB Agreement** Statistics FAQ News Contact Us

**Access the PDB FTP:**

RCSB PDB | PDBe | PDBj

Archive Download

**Deposit Data to the PDB:**

RCSB PDB | PDBe

PDBj | BMRB

**Search wwPDB Websites:**

RCSB PDB | PDBe

PDBj | BMRB

**PDB Archive Snapshots**

**Instructions to Journals**

**PDB Remediation**

Description

Chemical Component  
Dictionary

Software

**Documentation**

Format

Annotation

Remediation

**Workshops**

X-ray Validation

The Worldwide Protein Data Bank (wwPDB) consists of organizations that act as deposition, data processing and distribution centers for PDB data. The founding members are **RCSB PDB** (USA), **PDBe** (Europe) and **PDBj** (Japan)<sup>1</sup>. The **BMRB** (USA) group joined the wwPDB in 2006. The mission of the wwPDB is to maintain a single Protein Data Bank Archive of macromolecular structural data that is freely and publicly available to the global community.

This site provides information about services provided by the individual member organizations and about projects undertaken by the wwPDB.

Please note: <ftp://ftp.rcsb.org> is no longer updated. Please access the PDB archive using one of the FTP sites listed in the left menu.

**15-September-2008**

**Announcement: Comprehensive Format Guide Version 3.2**

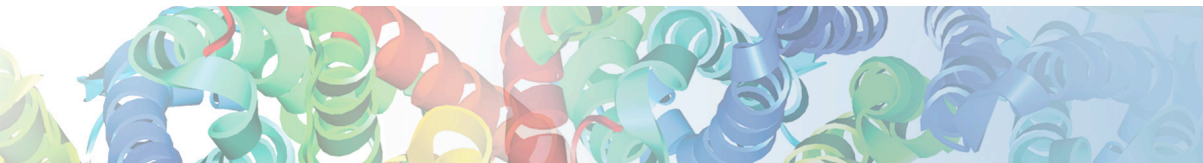
During the past year, the wwPDB annotators have collaborated on a project to clarify the details and procedures related to data processing and annotation. The result is a PDB Contents Guide Version 3.2 that more fully describes the PDB file format. This document is available as a **PDF** and in **HTML**, and is accompanied by a **document** highlighting these clarifications.

In the coming months, all files released by the wwPDB will follow the format as described in this document. Details will be made available on this website and at [www.wwpdb.org](http://www.wwpdb.org).

**14-August-2008**

**IUCr: wwPDB Exhibition Stand and Presentations**

The wwPDB partners will be exhibiting at the XXI Congress & General Assembly of the International Union of Crystallography (IUCr; August 23 - 31 in Osaka, Japan) at booth #14. Please stop by for website demonstrations and to meet with wwPDB members from around the globe.



## How is chemical information delivered by PDB?

### Dictionaries and Data Files 3 Formats

PDB (ca. 1972)

mmCIF (ca. 1997)

PDBML (ca. 2005)

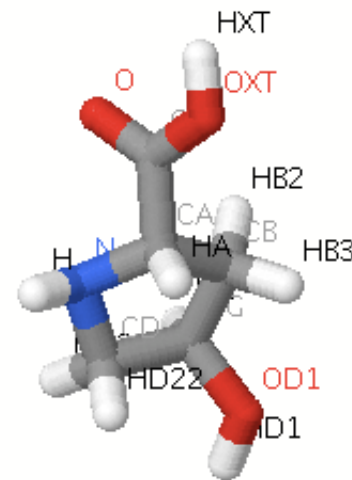
# PDB HET Dictionary Example

<ftp://ftp.wwpdb.org/pub/pdb/data/monomers/>

```

RESIDUE  HYP      18
CONNECT  N        3 CA   CD   H
CONNECT  CA       4 N    C    CB   HA
CONNECT  C        3 CA   O    OXT
CONNECT  O        1 C
CONNECT  CB       4 CA   CG   HB2  HB3
CONNECT  CG       4 CB   CD   OD1  HG
CONNECT  CD       4 N    CG   HD22 HD23
CONNECT  OD1      2 CG   HD1
CONNECT  OXT      2 C    HXT
CONNECT  H        1 N
CONNECT  HA       1 CA
CONNECT  HB2     1 CB
CONNECT  HB3     1 CB
CONNECT  HG      1 CG
CONNECT  HD22    1 CD
CONNECT  HD23    1 CD
CONNECT  HD1     1 OD1
CONNECT  HXT     1 OXT
END
  
```

Connectivity



```

HET      HYP      18
HETSYN   HYP HYDROXYPROLINE
HETNAM   HYP 4-HYDROXYPROLINE
FORMUL   HYP    C5 H9 N1 O3
  
```

Name & Formula





## PDB Data File Example

HET HYP E 12 8

HETSYN HYP HYDROXYPROLINE  
HETNAM HYP 4-HYDROXYPROLINE  
FORMUL HYP C5 H9 N1 O3



Name & Formula

ATOM	9302	N	HYP	E	12	-6.687	-2.308	28.039	1.00	31.26	N
ATOM	9303	CA	HYP	E	12	-7.953	-3.047	27.929	1.00	31.93	C
ATOM	9304	C	HYP	E	12	-8.133	-3.707	26.555	1.00	32.45	C
ATOM	9305	O	HYP	E	12	-7.784	-3.116	25.531	1.00	36.79	O
ATOM	9306	CB	HYP	E	12	-9.041	-2.009	28.163	1.00	32.46	C
ATOM	9307	CG	HYP	E	12	-8.383	-0.818	28.833	1.00	34.37	C
ATOM	9308	CD	HYP	E	12	-6.885	-0.945	28.586	1.00	33.39	C
ATOM	9309	OD1	HYP	E	12	-8.866	0.380	28.269	1.00	48.65	O

CONNECT 9302 9303 9308  
CONNECT 9303 9302 9304 9306  
CONNECT 9304 9303 9305  
CONNECT 9305 9304  
CONNECT 9306 9303 9307  
CONNECT 9307 9306 9308 9309  
CONNECT 9308 9302 9307  
CONNECT 9309 9307



Connectivity

END

REMARK 600 – Chemical details  
REMARK 610/615 – Missing & zero occupancy atoms  
REMARK 620 – Metal coordination  
REMARK 630 – Inhibitor description

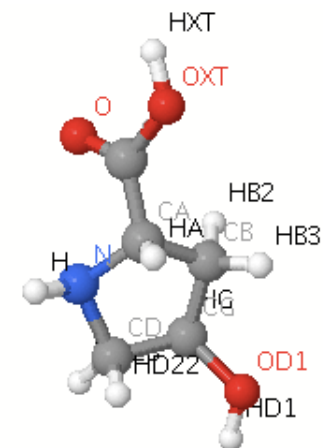
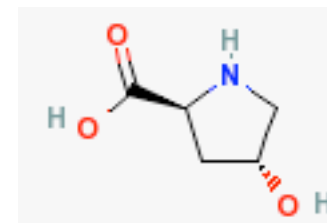
# mmCIF Component Definition Example

<ftp://ftp.wwpdb.org/pub/pdb/data/monomers/components.cif>

```

data_HYP
#
_chem_comp.id          HYP
_chem_comp.name        4-HYDROXYPROLINE
_chem_comp.type        "L-PEPTIDE LINKING"
_chem_comp.pdbx_type   ATOMP
_chem_comp.formula     "C5 H9 N O3"
_chem_comp.mon_nstd_parent_comp_id  PRO
_chem_comp.pdbx_synonyms  HYDROXYPROLINE
_chem_comp.pdbx_formal_charge  0
_chem_comp.pdbx_initial_date  1999-07-08
_chem_comp.pdbx_modified_date  2008-04-29
_chem_comp.pdbx_ambiguous_flag  N
_chem_comp.pdbx_release_status  ?
_chem_comp.pdbx_replaced_by  ?
_chem_comp.pdbx_replaces  ?
_chem_comp.formula_weight  131.130
_chem_comp.one_letter_code  ?
_chem_comp.three_letter_code  HYP
_chem_comp.pdbx_model_coordinates_details  ?
_chem_comp.pdbx_model_coordinates_missing_flag  N
_chem_comp.pdbx_ideal_coordinates_details  Corina
_chem_comp.pdbx_ideal_coordinates_missing_flag  N
_chem_comp.pdbx_model_coordinates_db_code  10B6
_chem_comp.pdbx_processing_site  RCSB
#

```



# mmCIF Definition Atom List Example

```

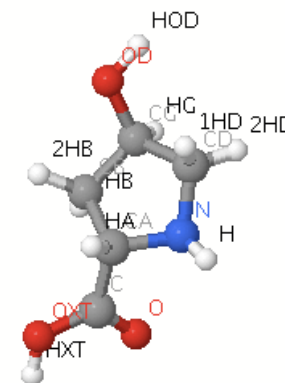
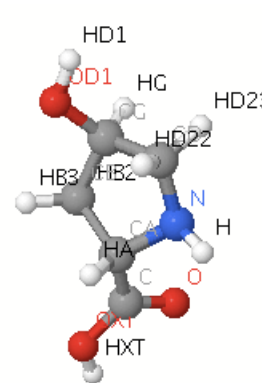
loop_
  _chem_comp_atom.comp_id
  _chem_comp_atom.atom_id
  _chem_comp_atom.alt_atom_id
  _chem_comp_atom.type_symbol
  _chem_comp_atom.charge
  _chem_comp_atom.pdbx_align
  _chem_comp_atom.pdbx_aromatic_flag
  _chem_comp_atom.pdbx_leaving_atom_flag
  _chem_comp_atom.pdbx_stereo_config
  _chem_comp_atom.model_Cartn_x
  _chem_comp_atom.model_Cartn_y
  _chem_comp_atom.model_Cartn_z
  _chem_comp_atom.pdbx_model_Cartn_x_ideal
  _chem_comp_atom.pdbx_model_Cartn_y_ideal
  _chem_comp_atom.pdbx_model_Cartn_z_ideal
  _chem_comp_atom.pdbx_ordinal
HYP N      N      N  0  1  N  N  N  -3.366  16.585  44.188  0.168  1.360  -0.282  1
HYP CA     CA     C  0  1  N  N  S  -2.955  15.768  43.044  -0.384  -0.003  -0.493  2
HYP C      C      C  0  1  N  N  N  -1.447  15.609  43.030  -1.811  -0.072  -0.013  3
HYP O      O      O  0  1  N  N  N  -0.722  16.484  43.503  -2.233  0.764  0.750  4
HYP CB     CB     C  0  1  N  N  N  -3.408  16.578  41.829  0.515  -0.924  0.359  5
HYP CG     CG     C  0  1  N  N  R  -4.437  17.482  42.330  1.847  -0.159  0.505  6
HYP CD     CD     C  0  1  N  N  N  -4.068  17.803  43.753  1.640  1.159  -0.271  7
HYP OD1    OD     O  0  1  N  N  N  -5.693  16.815  42.294  2.917  -0.911  -0.071  8
HYP OXT    OXT    O  0  1  N  Y  N  -0.976  14.502  42.469  -2.614  -1.063  -0.433  9
HYP H      H      H  0  1  N  Y  N  -3.980  16.047  44.765  -0.107  1.981  -1.028  10
HYP HA     HA     H  0  1  N  N  N  -3.385  14.756  43.068  -0.325  -0.278  -1.546  11
HYP HB2    1HB    H  0  1  N  N  N  -2.567  17.141  41.398  0.066  -1.092  1.337  12
HYP HB3    2HB    H  0  1  N  N  N  -3.790  15.930  41.026  0.678  -1.873  -0.153  13
HYP HG     HG     H  0  1  N  N  N  -4.508  18.399  41.726  2.052  0.048  1.555  14
HYP HD22   1HD    H  0  0  N  N  N  -4.956  18.005  44.370  2.018  1.065  -1.289  15
HYP HD23   2HD    H  0  0  N  N  N  -3.457  18.713  43.848  2.132  1.985  0.243  16
HYP HD1    HOD    H  0  1  N  N  N  -5.999  16.666  43.181  3.780  -0.479  -0.009  17
HYP HXT    HXT    H  0  1  N  N  N  -0.027  14.511  42.499  -3.520  -1.066  -0.098  18
#
  
```

← Atom names

← Stereochemistry & aromaticity

← Model coordinates

← Ideal coordinates



## mmCIF Definition Connectivity Table

```

loop_
  _chem_comp_bond.comp_id
  _chem_comp_bond.atom_id_1
  _chem_comp_bond.atom_id_2
  _chem_comp_bond.value_order
  _chem_comp_bond.pdbx_aromatic_flag
  _chem_comp_bond.pdbx_stereo_config
  _chem_comp_bond.pdbx_ordinal

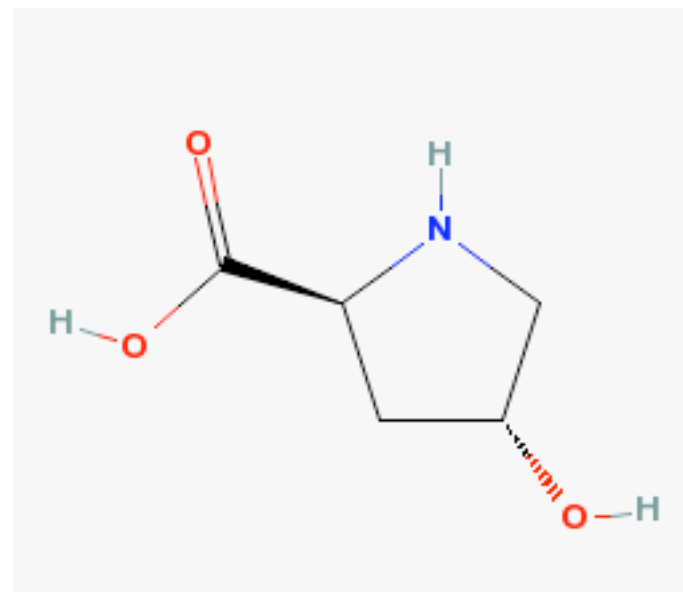
```

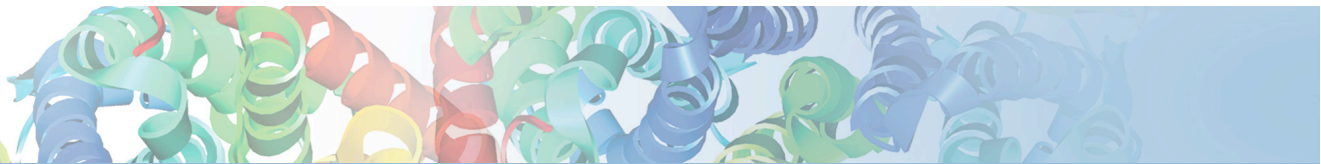
← Connected atoms  
 ← Bond type  
 ← Stereochemistry & aromaticity

```

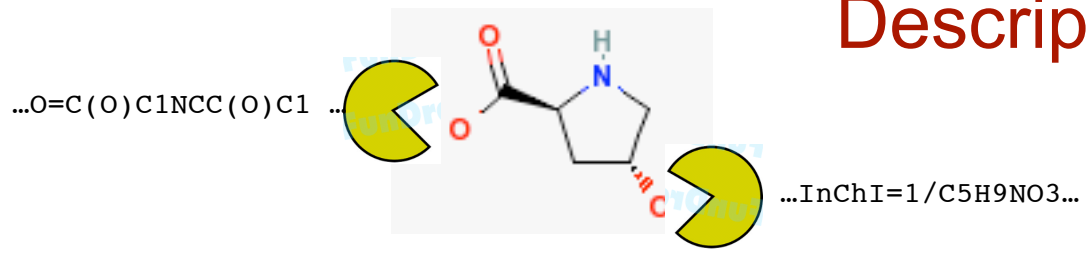
HYP N CA SING N N 1
HYP N CD SING N N 2
HYP N H SING N N 3
HYP CA C SING N N 4
HYP CA CB SING N N 5
HYP CA HA SING N N 6
HYP C O DOUB N N 7
HYP C OXT SING N N 8
HYP CB CG SING N N 9
HYP CB HB2 SING N N 10
HYP CB HB3 SING N N 11
HYP CG CD SING N N 12
HYP CG OD1 SING N N 13
HYP CG HG SING N N 14
HYP CD HD22 SING N N 15
HYP CD HD23 SING N N 16
HYP OD1 HD1 SING N N 17
HYP OXT HXT SING N N 18

```





## Descriptors and Identifiers



```
#
loop_
_pdbx_chem_comp_descriptor.comp_id
_pdbx_chem_comp_descriptor.type
_pdbx_chem_comp_descriptor.program
_pdbx_chem_comp_descriptor.program_version
_pdbx_chem_comp_descriptor.descriptor
```

### SMILES

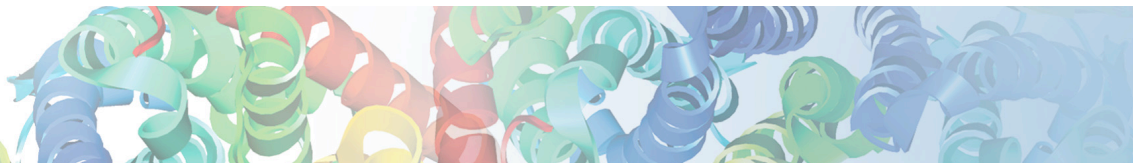
HYP SMILES	ACDLabs	10.04	O=C(O)C1NCC(O)C1
HYP SMILES_CANONICAL	CACTVS	3.341	O[C@H]1CN[C@@H](C1)C(O)=O
HYP SMILES	CACTVS	3.341	O[CH]1CN[CH](C1)C(O)=O
HYP SMILES_CANONICAL	"OpenEye OEToolkits"	1.5.0	C1[C@H](CN[C@@H]1C(=O)O)O
HYP SMILES	"OpenEye OEToolkits"	1.5.0	C1C(CNC1C(=O)O)O
HYP InChI	InChI	1.02b	"InChI=1/C5H9NO3/c7 ... (H,8,9)/t3-,4+/m1/s1/f/h8H"
HYP InChIKey	InChI	1.02b	PMMYEEVYMWASQN-FJFIVYRMDE

```
#
loop_
_pdbx_chem_comp_identifier.comp_id
_pdbx_chem_comp_identifier.type
_pdbx_chem_comp_identifier.program
_pdbx_chem_comp_identifier.program_version
_pdbx_chem_comp_identifier.identifier
```

### InChI

HYP "SYSTEMATIC NAME"	ACDLabs	10.04	(4R)-4-hydroxy-L-proline
HYP "SYSTEMATIC NAME"	"OpenEye OEToolkits"	1.5.0	"(2S,4R)-4-hydroxypyrrolidine-2-carboxylic acid"

### Systematic Names



## Descriptors

- **SMILES - Simplified molecular input line entry specification**
  - Weininger, D. (1988), SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules, J. Chem. Inf. Comput. Sci. 28, 31-36
  - <http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html>.
- **InChI - IUPAC International Chemical Identifier**
  - Stephen E. Stein, Stephen R. Heller, and Dmitrii Tchekhovskoi, An Open Standard for Chemical Structure Representation: The IUPAC Chemical Identifier, in Proceedings of the 2003 International Chemical Information Conference (Nimes), pp. 131-143.
  - <http://old.iupac.org/inchi/>
- **InChiKey – InChI digital signature**
  - <http://old.iupac.org/inchi/release102.html>



## Other Features

New in V3.2

loop\_

`_pdbx_chem_comp_feature.comp_id`

`_pdbx_chem_comp_feature.type`

`_pdbx_chem_comp_feature.value`

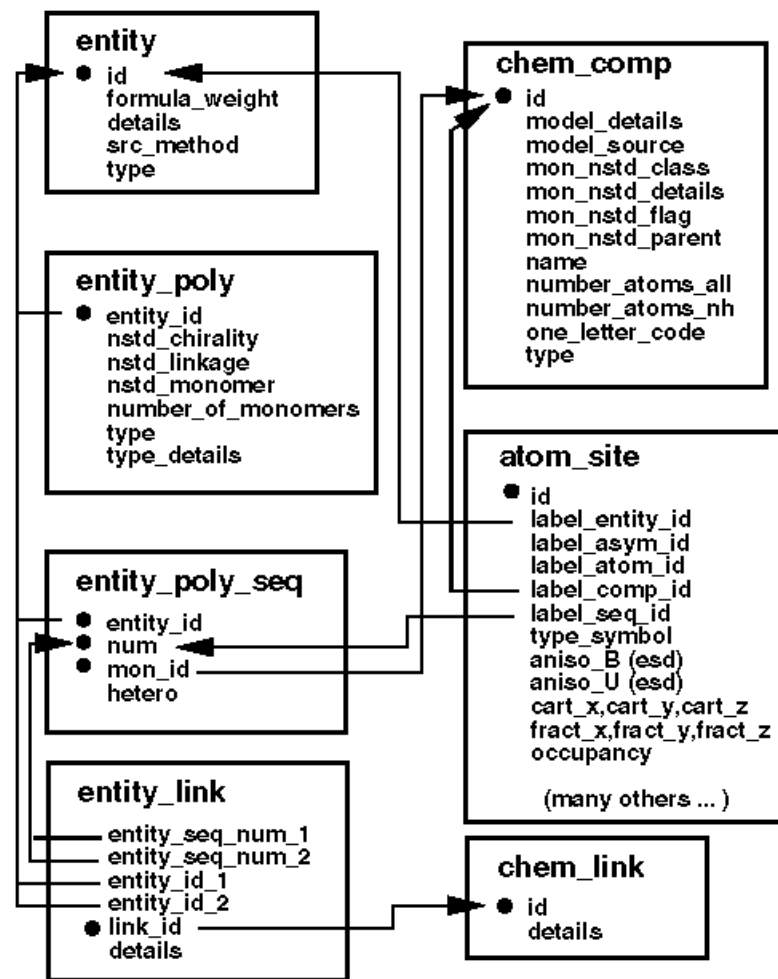
`_pdbx_chem_comp_feature.source`

00X	'ENZYME INHIBITED'	'ASPARTIC PROTEINASE'	PDB
00X	'FUNCTION'	'Transition-state analogue inhibitor'	PDB
00X	'STRUCTURE IMAGE URL'	' <a href="http://journals.iucr.org/00X.jpg">http://journals.iucr.org/00X.jpg</a> '	IUCR

## Key Categories in mmCIF Data Files

- entity – unique polymer and non-polymer molecules
- entity\_poly, entity\_poly\_seq – polymer features and sequence
- chem\_comp – chemical component names, identifiers and features
- atom\_site – Cartesian coordinates
- struct\_asym – list of molecular instances
- struct\_conn – intermolecular linkages

Dictionary details are not reproduced in each data file!







## PDBML Schemas

<http://pdbml.pdb.org/>

- Schemas provided for the PDB Exchange dictionary and component dictionaries
- Schemas and data dictionaries are updated synchronously

PDBML: the representation of archival macromolecular structure data in XML. John Wesbrook, Nobutoshi Ito, Haruki Nakamura, Kim Henrick and Helen M. Berman, *Bioinformatics*, 21(7), 988-992, 2005.



## PDBML Component Definition Example

```

<PDBx:datablock datablockName="HYP" xsi:schemaLocation=
  http://deposit.pdb.org/pdbML/pdbx.xsd pdbx.xsd>
<PDBx:chem_compCategory>
  <PDBx:chem_comp id="HYP">
    <PDBx:name>4-HYDROXYPROLINE</PDBx:name>
    <PDBx:type>L-peptide linking</PDBx:type>
    <PDBx:pdbx_type>ATOMP</PDBx:pdbx_type>
    <PDBx:formula>C5 H9 N O3</PDBx:formula>
    <PDBx:mon_nstd_parent_comp_id>PRO</PDBx:mon_nstd_parent_comp_id>
    <PDBx:pdbx_synonyms>HYDROXYPROLINE</PDBx:pdbx_synonyms>
    <PDBx:pdbx_formal_charge>0</PDBx:pdbx_formal_charge>
    <PDBx:pdbx_initial_date>1999-07-08</PDBx:pdbx_initial_date>
    <PDBx:pdbx_modified_date>2008-04-29</PDBx:pdbx_modified_date>
    <PDBx:pdbx_ambiguous_flag>N</PDBx:pdbx_ambiguous_flag>
    <PDBx:pdbx_release_status>REL</PDBx:pdbx_release_status>
    <PDBx:formula_weight>131.130</PDBx:formula_weight>
    <PDBx:three_letter_code>HYP</PDBx:three_letter_code>
    <PDBx:pdbx_model_coordinates_missing_flag>N</PDBx:pdbx_model_coordinates_missing_flag>
    <PDBx:pdbx_ideal_coordinates_details>Corina</PDBx:pdbx_ideal_coordinates_details>
    <PDBx:pdbx_ideal_coordinates_missing_flag>N</PDBx:pdbx_ideal_coordinates_missing_flag>
  </PDBx:chem_comp>
</PDBx:chem_compCategory>
  
```



## PDBML Data File Example

```

loop_
_entity_poly.entity_id
_entity_poly.type
_entity_poly.nstd_linkage
_entity_poly.nstd_monomer
_entity_poly.pdbx_seq_one_letter_code
_entity_poly.pdbx_seq_one_letter_code_can
1 polypeptide(L) no no
;DIVLTQSPASLSASVGETVTTITCRASGNIHNYLAWYQQKQGKSPQLLVYYTTTLADG
VPSRFSGSGSGTQYSLKINSLQPEDFGSYQCQHFWSPTPRTFGGGTKLEIK
;
;DIVLTQSPASLSASVGETVTTITCRASGNIHNYLAWYQQKQGKSPQLLVYYTTTLADG
VPSRFSGSGSGTQYSLKINSLQPEDFGSYQCQHFWSPTPRTFGGGTKLEIK
;

```

```

<PDBx:entity_polyCategory>
  <PDBx:entity_poly entity_id="1">
    <PDBx:type>polypeptide(L)</PDBx:type>
    <PDBx:nstd_linkage>no</PDBx:nstd_linkage>
    <PDBx:nstd_monomer>no</PDBx:nstd_monomer>
    <PDBx:pdbx_seq_one_letter_code>
      DIVLTQSPASLSASVGETVTTITCRASGNIHNYLAWYQQKQGKSPQLLVYYTTTLADG
      VPSRFSGSGSGTQYSLKINSLQPEDFGSYQCQHFWSPTPRTFGGGTKLEIK
    </PDBx:pdbx_seq_one_letter_code>
    <PDBx:pdbx_seq_one_letter_code_can>
      DIVLTQSPASLSASVGETVTTITCRASGNIHNYLAWYQQKQGKSPQLLVYYTTTLADG
      VPSRFSGSGSGTQYSLKINSLQPEDFGSYQCQHFWSPTPRTFGGGTKLEIK
    </PDBx:pdbx_seq_one_letter_code_can>
  </PDBx:entity_poly>
</PDBx:entity_polyCategory>

```

Standard container for each category

Keys mapped as attributes

# Dictionary and Schema Resources

<http://mmcif.pdb.org> & <http://pdbml.pdb.org>

## Dictionary Resources

The Protein Data Bank (PDB) uses macromolecular Crystallographic Information File (mmCIF) data dictionaries to describe the information content of PDB entries. The PDB Exchange data dictionary consolidates content from a variety of crystallographic dictionaries including: the IUCr Core, mmCIF, Image and symmetry dictionaries. The PDB Exchange Dictionary also includes extensions describing NMR, Cryo-EM, and protein production data. PDB data processing, data exchange, annotation, and database management operations all make heavy use of the dictionary format. The content of the PDB Exchange Dictionary, Software Tools, and mmCIF data files to the older PDB format and to PDBML/XML.

- Data files in mmCIF format can be downloaded
- Software tools are available for preparing and
- Software tools are available for converting mm
- A complete list of PDB software tools for mana

## Dictionary Content and Representation

- [Background and Introduction](#) about mmCIF
- [The Macromolecular Crystallographic Information File](#)
- [STAR/mmCIF: An Extensive Ontology for Macromolecular Crystallography](#)
- [mmCIF Software Developers Workshop 1997](#)
- [mmCIF Dictionary Templates](#)
- [mmCIF Examples](#)
- [References](#)

## Data Dictionaries

## PDBML Resources

### PDBML

The Protein Data Bank Markup Language (PDBML) provides a representation of PDB data in XML format. The description of this format is provided in XML schema of the PDB Exchange Data Dictionary. This schema is produced by direct translation of the mmCIF format PDB Exchange Data Dictionary. Other data dictionaries used by the PDB have been electronically translated into XML/XSD schemas and these are also presented in the list below.

- PDBML data files are provided in three forms:
  - fully marked-up files,
  - files without atom records
  - files with a more space efficient encoding of atom records
- Data files in PDBML format can be downloaded from the [RCSB PDB website](#) or by [ftp](#).
- Software tools for manipulating PDB data in XML format can be found [here](#).
- An [article](#) describing PDBML is available.  
PDBML: the representation of archival macromolecular structure data in XML.  
John Wesbrook, Nobutoshi Ito, Haruki Nakamura, Kim Henrick and Helen M. Berman, Bioinformatics, 21(7), 988-992, 2005.

### PDBML Schema

- [PDB Exchange Dictionary](#) | [current version](#) | [previous versions](#) | [alternative atom record markup](#) || XML Schema for Exchange Data dictionary developed as a collaboration between MSD-EBI, PDBj and RCSB and used by wwPDB members for data exchange.

## Other Supporting Software Tools

<http://sw-tools.rcsb.org>

- Validating Parsers for Files and Dictionaries
- Dictionary access and presentation tools
- File format translation tools
- Database Builder and Loader
- XML translation tool for data files and dictionaries
- Others: BioPerl, BioPython, mmLib (py), CCP4



## Ligand Expo

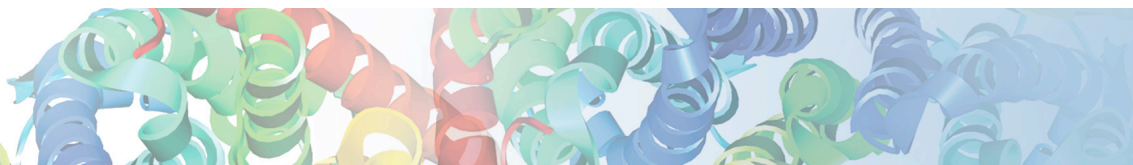
<http://ligand-expo.rcsb.org>

**Ligand Expo** presents a view of PDB data focusing chemical and structural information about small molecules within the structure entries of the Protein Data Bank. Tools are provided to browse and search the PDB dictionary of chemical components, to identify structure entries containing particular small molecules, and to download the 3D structures of the small molecule components within PDB entries. Tools are also provided to build new chemical definitions for PDB chemical dictionary.



## Chemical Dictionary Search

- Molecular Name
- Formula
- SMILES
- InChI/InChIKey
- PDB Component Identifier
- Chemical substructure



## Search Options

MOLECULAR NAME, FORMULA, AND DESCRIPTOR SEARCH OPTIONS ?

Search term  Search type

SKETCH INPUT AND/OR STRUCTURE SEARCH OPTIONS ?

File name   File f

-- OR --

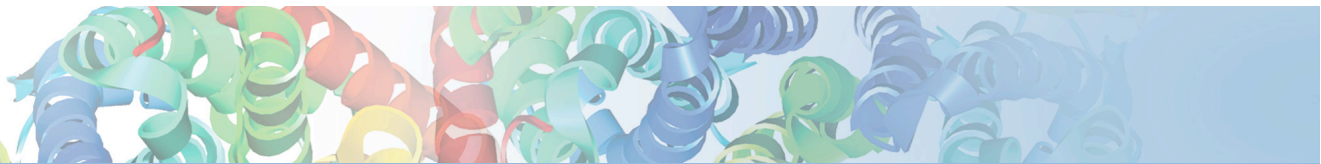
SMILES string

SEARCH FOR INSTANCES OF CHEMICAL COMPONENTS BY 3-LETTER ID CODE ?

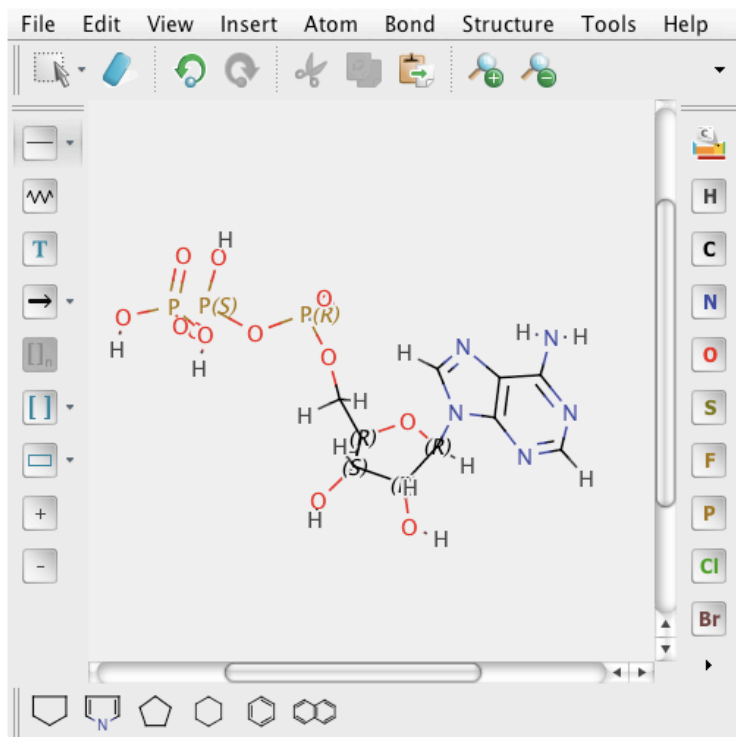
Component ID code  Display

- SMILES
- Component identifier (3-letter code)
- Similar to component (3-letter code)
- Formula (all atom exact)
- Formula (heavy atom exact)
- Formula (exact subset)
- Formula (subset)
- Formula (close)
- Molecular name (exact)
- Molecular name (exact sub-string)
- Molecular name (similar)
- SMILES
- Chemically similar to SMILES
- InChI
- InChIKey





### Chemical Component Sketch Tool



## Component Building & Substructure Search

#### SUBSTRUCTURE SEARCH OPTIONS

Search type

#### SAVE CHEMICAL COMPONENT

Export format   
Optional Component ID (3-letter-code)





## Browsing Options

Predefined queries based on SMILES patterns are provided in menus to permit convenient browsing of molecules containing:

- Amino acids
- Nucleotides
- Selected top-selling pharmaceuticals
- Common aromatic ring systems

## Example: Browsing Popular Drugs

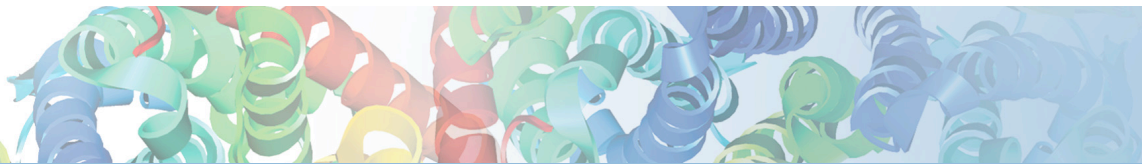
### Ligand Expo

Version 1 Updated: 2-May-2008

Ligand Expo (formerly Ligand Depot) provides chemical and structural information about small molecules within the structure entries of the Protein Data Bank. Tools are provided to search the PDB dictionary for chemical components, to identify structure entries containing particular small molecules, and to download the 3D structures of the small molecule components in the PDB entry. A sketch tool is also provided for building new chemical definitions from reported PDB chemical components.

#### Getting Started:

The [Ligand Expo Tutorial](#) contains a summary of query, report and download features.  
Select the [Search](#) menu to find a chemical component or build a new chemical component.  
Select the [Browse](#) menu to view tabulations of modified amino acids and nucleotides.  
Select the [Download](#) for links to chemical dictionaries and resource files.



Amino Acids Nucleotides **Popular Drugs** Ring Systems

Ligand Expo Browser

BROWSER SELECTION ?

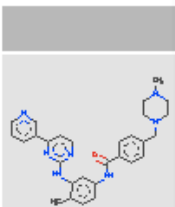
Now browsing: Molecules like Gleevec (Imatinib)

Comparison type: Molecules like Actonel (Risedronate)

Molecules like

Comparison type

Result count:



- Molecules like Gleevec (Imatinib)
- Molecules like Actonel (Risedronate)
- Molecules like Advil (Ibuprofen)
- Molecules like Biaxin (Clarithromycin)
- Molecules like Boniva (Ibandronic acid)
- Molecules like Cipro (Ciprofloxacin)
- Molecules like Combivir (Lamivudine)
- Molecules like Cialis (Tadalafil)
- Molecules like Diflucan (Fluconazole)
- Molecules like Diovan (Valsartan)
- Molecules like Evista (Raloxifene)
- Molecules like Gemzar (Gemcitabine)
- Molecules like Gleevec (Imatinib)
- Molecules like Kaletra (Lopinavir)
- Molecules like Levitra (Vardenafil)
- Molecules like Lipitor (Atorvastatin)
- Molecules like Mevalotin (Pravastatin)
- Molecules like Namenda (Memantine)
- Molecules like Nasonex (Mometasone furoate)
- Molecules like Pravachol (Pravastatin)

Name: 4-(4-methyl-piperazin-1-ylmethyl)-n-[4-methyl-3-(4-pyridin-3-yl-pyrimidin-2-ylamino)-phenyl]-benzamide

Synonyms: sti-571;imatinib  
sti-571;imatinib

4-[(4-methylpiperazin-1-yl)methyl]-n-[4-methyl-3-[(4-pyridin-3-ylpyrimidin-2-yl)amino]phenyl]benzamide

SMILES: Cc1ccc(cc1Nc2nccc(n2)c3cccnc3)NC(=O)c4ccc(cc4)CN5CCN(CC5)C  
Formula: C29 H31 N7 O

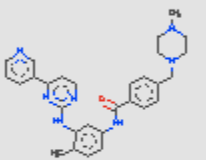
Browse

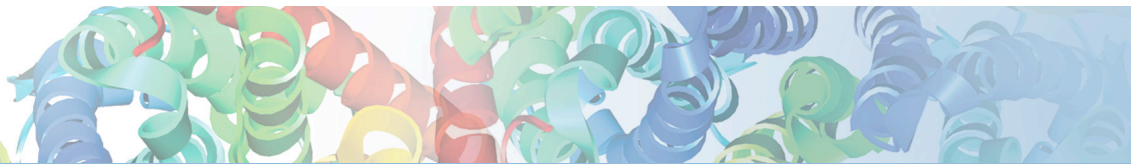
## Search report summary

### Molecules like Gleevec (Imatinib)

**Comparison type:** Containing SMILES pattern

**Result count:** 1

	ID	View Options	Description
	STI	<a href="#">Chemical details</a> <a href="#">Coordinates files</a>	<p><b>Name:</b> 4-(4-methyl-piperazin-1-ylmethyl)-n-[4-methyl-3-(4-pyridin-3-yl-pyrimidin-2-ylamino)-phenyl]-benzamide</p> <p><b>Synonyms:</b> sti-571;imatinib sti-571;imatinib 4-[(4-methylpiperazin-1-yl)methyl]-n-{4-methyl-3-[(4-pyridin-3-ylpyrimidin-2-yl)amino]phenyl}benzamide 4-[(4-methylpiperazin-1-yl)methyl]-n-[4-methyl-3-[(4-pyridin-3-ylpyrimidin-2-yl)amino]phenyl]benzamide</p> <p><b>SMILES:</b> <chem>Cc1ccc(cc1Nc2nccc(n2)c3ccnc3)NC(=O)c4ccc(cc4)CN5CCN(CC5)C</chem></p> <p><b>Formula:</b> C29 H31 N7 O</p>

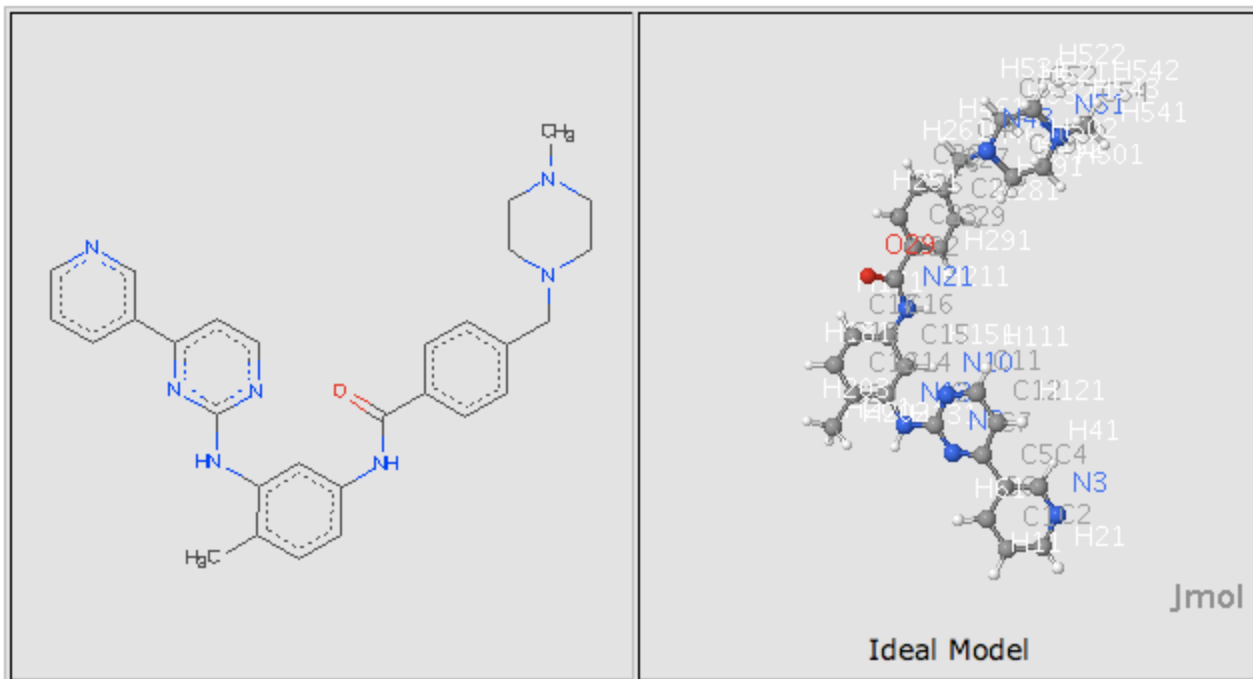


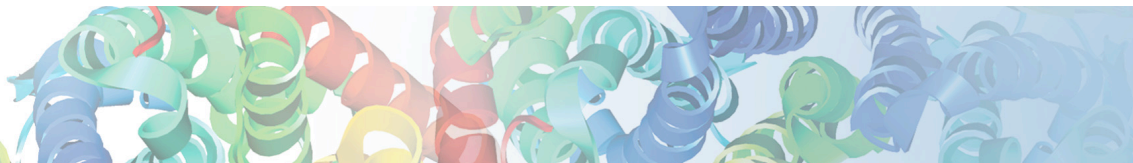
# Chemical Details

Ligand Expo

- Chemical Details**
- Geometry
- Atom Nomenclature
- Downloads
- Related Resources

## PDB Chemical Component STI





## Chemical Details

### Chemical Description

<b>Name</b>	4-(4-METHYL-PIPERAZIN-1-YLMETHYL)-N-[4-METHYL-3-(4-PYRIDIN-3-YL-PYRIMIDIN-2-YLAMINO)-PHENYL]-BENZAMIDE
<b>Synonyms</b>	STI-571;IMATINIB
<b>Formula</b>	C29 H31 N7 O
<b>Formal charge</b>	0
<b>Molecular weight</b>	493.603 g/mol
<b>Component type</b>	NON-POLYMER

### Chemical features

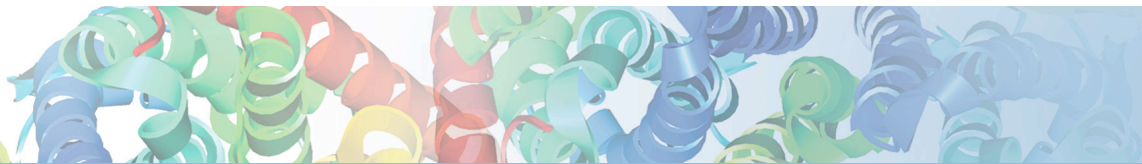
<b>Atom count</b>	68
<b>Chiral atom count</b>	0
<b>Bond count</b>	72
<b>Aromatic bond count</b>	24

### Chemical Identifiers

<b>Systematic name (ACDLabs)</b>	4-[[4-methylpiperazin-1-yl)methyl]-N-{4-methyl-3-[(4-pyridin-3-yl)pyrimidin-2-yl)amino]phenyl}benzamide
<b>Systematic name (OpenEye OEToolkits)</b>	4-[[4-methylpiperazin-1-yl)methyl]-N-[4-methyl-3-[(4-pyridin-3-yl)pyrimidin-2-yl)amino]phenyl]benzamide

### Chemical Descriptors

<b>Stereo SMILES (CACTVS)</b>	<chem>CN1CCN(CC1)Cc2ccc(cc2)C(=O)Nc3ccc(C)c(Nc4nccc(n4)c5cccnc5)c3</chem>
<b>SMILES (CACTVS)</b>	<chem>CN1CCN(CC1)Cc2ccc(cc2)C(=O)Nc3ccc(C)c(Nc4nccc(n4)c5cccnc5)c3</chem>
<b>Stereo SMILES (OpenEye)</b>	<chem>Cc1ccc(cc1Nc2nccc(n2)c3cccnc3)NC(=O)c4ccc(cc4)CN5CCN(CC5)C</chem>
<b>InChI descriptor</b>	InChI=1/C29H31N7O/c1-21-5-10-25(18-27(21)34-29-31-13-11-26(33-29)24-4-3-12-30-19-24)32-28(37)23-8-6-22(7-9-23)20-36-16-14-35(2)15-17-36/h3-13,18-19H,14-17,20H2,1-2H3,(H,32,37)(H,31,33,34)/f/h32,34H
<b>InChIKey descriptor</b>	KTUFNOKKBVMGRW-RPGFBOUCX



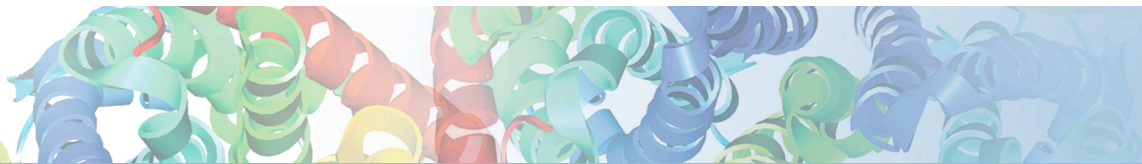
# Downloads

## PDB Chemical Component STI

### Download Coordinate Files and Chemical Diagrams

- Component definition in mmCIF format
- SDF/MOL format (experimental model coordinates)
- SDF/MOL format (ideal coordinates)
- PDB format (experimental model coordinates)
- PDB format (ideal coordinates)
- Chemical diagram without atom labels
- Chemical diagram with format V3 atom labels (heavy atoms) ←
- Chemical diagram with format V3 atom labels (all atoms)
- Chemical diagram with format V2 atom labels (heavy atoms)
- Chemical diagram with format V2 atom labels (all atoms)





# Model Geometry

**Ligand Expo**

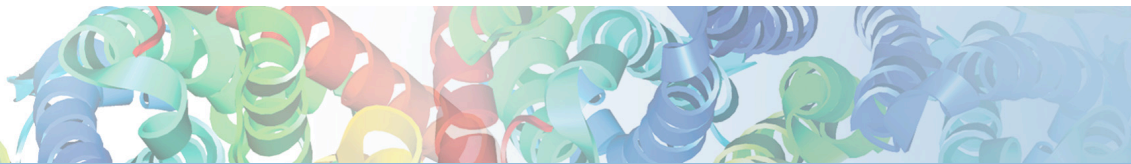
[Chemical Details](#) **[Geometry](#)** [Atom Nomenclature](#) [Downloads](#) [Related Resources](#)

## PDB Chemical Component STI

The following tables include covalent bond distance, bond angle and torsion angle values for the experimental model and ideal geometries stored in the definition for this chemical component.

### Bond Distances for STI

Atom I	Atom J	Distance (Experimental Model)	Distance (Ideal Model)
C1	C6	1.446	1.384
C1	C2	1.436	1.383
C1	H11	1.104	1.080
C6	C5	1.458	1.398
C6	H61	1.104	1.080
C5	C4	1.452	1.392
C5	C7	1.501	1.483
C4	N3	1.391	1.318
C4	H41	1.103	1.080




## Related Resources

### Related Data Resources for STI

The following table contains links to resources and identifiers which have additional information about this chemical component. The first column of the table contains resource names. Each resource name is linked to the top-level of the resource website. The identifier used by the resource is given in the second column. Clicking on a highlighted identifier will launch a query at the related resource and display the results in a new window.

Related Resource	Identifier
<a href="#">PubChem</a>	<a href="#">5291</a>
<a href="#">DrugBank</a>	<a href="#">DB03261</a>
<a href="#">ChEBI</a>	<a href="#">45783</a>
<a href="#">BindingDB</a>	<a href="#">13530</a>
<a href="#">NCI Chemical Lookup Service</a>	<a href="#">SMILES query</a>
<a href="#">eMolecules</a>	<a href="#">SMILES query</a>
<a href="#">CAS Registry Number</a>	<a href="#">152459-95-5</a>

# Related Resources



PubChem

Imat

inhibit

Target

Search

Se

Se

ChE

Adv

Bro

Dow

Doc

Dev

Con

Ch

Na

Sh

Nu

Dat

Au

Ins

Pu

Oth

PC


Un

Pu

Do

All

Tar



**ChemSpider**<sup>TM</sup>

Building a Structure Centric Community for Chemists

GO

Search...

Home

Search

Services

Resources

About

Login

Results 1

Simple Search

Structure Search

LASSO Search

Advanced LASSO

Chemical Elements

Properties Search

Predicted Properties

Data Source Search

Literature Search

NCBI Entrez Search

PubChem Search

Advanced Search

Searches History

**1 hit(s) found in 0.09 seconds**

**Search term: Cc1ccc(cc1Nc2nccc(n2)c3cccnc3)NC(=O)c4ccc(cc4)CN5CCN(CC5)C**

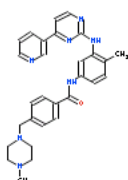
**Found by conversion query string to chemical structure (full match)**

Please [login](#) to be able to add spectra, identifiers, links and publications.

Add: Comments

**INHERENT PROPERTIES, IDENTIFIERS AND REFERENCES**

2D



load save zoom

**ChemSpider ID:** 5101

**Empirical Formula:** C<sub>29</sub>H<sub>31</sub>N<sub>7</sub>O

**Molecular Weight:** 493.6027

**Nominal Mass:** 493 Da

**Average Mass:** 493.6027 Da

**Monoisotopic Mass:** 493.259009 Da

**Quick Links:** [Permalink](#) [Similar](#) [Isomers](#)

*Place Your Ad Here*

or

*Claim this Molecule*

Click for Details

**Systematic Name:** 4-[[4-(4-methylpiperazin-1-yl)methyl]-N-[4-methyl-3-[[4-(3-pyridyl)pyrimidin-2-yl]amino]phenyl]benzamide

**SMILES:** O=C(Nc3ccc(c(Nc2nc(c1cccnc1)ccn2)c3)C)c4ccc(cc4)CN5CCN(CC5)C

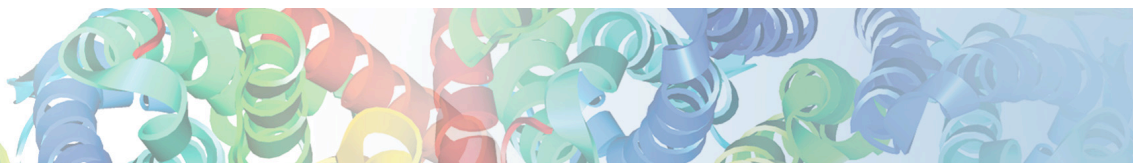
**InChI:** InChI=1/C29H31N7O/c1-21-5-10-25(18-27(21)34-29-31-13-11-26(33-29)24-4-3-12-30-19-24)32-28(37)23-8-6-22(7-9-23)20-36-16-14-35(2)15-17-36/h3-13,18-19H,14-17,20H2,1-2H3,(H,32,37)(H,31,33,34)

**InChIKey:** KTUFNOKKBVMGRW-UHFFFAOYAJ

**WIKIPEDIA ARTICLE(S)**

**LICENSE**

**Imatinib** is a drug used to treat certain types of cancer. It is currently marketed by Novartis as **Gleevec (USA)** or **Glivec (Europe/Australia)** as its mesylate salt, **imatinib mesilate (INN)**. It was originally coded during development as CGP57148B or STI-571 (these terms are used in early preclinical publications). It is used in treating chronic myelogenous leukemia (CML), gastrointestinal stromal tumors (GISTs) and a number of other malignancies. It is the first member of a new class of agents that act by inhibiting particular tyrosine kinase enzymes, instead of



## Experimental Examples

### Search Result Summary

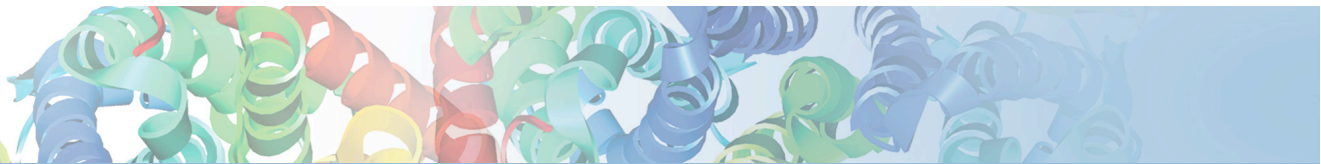
Target ID: STI

Count in released entries: 12

### Instances of STI in Released Entries

PDB ID	CC ID	Method	Title/Authors	Resolution	Downloads & Views
1xbb	STI	X-RAY	Crystal structure of the syk tyrosine kinase domain with Gleevec Nienaber, V.L., Atwell, S., Adams, J.M., Badger, J., Buchanan, M.D., Feil, I.K., Froning, K.J., Gao, X., Hendle, J., Keegan, K., Leon, B.C., Muller-Deickmann, H.J., Noland, B.W., Post, K., Rajashankar, K.R., Ramos, A., Russell, M., Burley, S.K., Buchanan, S.G.	1.57	<a href="#">Component definition</a> <a href="#">mmCIF</a> <a href="#">SDF</a> <a href="#">PDB</a> <a href="#">Launch viewer</a>
1t46	STI	X-RAY	STRUCTURAL BASIS FOR THE AUTOINHIBITION AND STI-571 INHIBITION OF C-KIT TYROSINE KINASE Mol, C.D., Dougan, D.R., Schneider, T.R., Skene, R.J., Kraus, M.L., Scheibe, D.N., Snell, G.P., Zou, H., Sang, B.C., Wilson, K.P.	1.6	<a href="#">Component definition</a> <a href="#">mmCIF</a> <a href="#">SDF</a> <a href="#">PDB</a> <a href="#">Launch viewer</a>
1opj	STI	X-RAY	Structural basis for the auto-inhibition of c-Abl tyrosine kinase Nagar, B., Hantschel, O., Young, M.A., Scheffzek, K., Veach, D., Bornmann, W., Clarkson, B., Superti-Furga, G., Kuriyan, J.	1.75	<a href="#">Component definition</a> <a href="#">mmCIF</a> <a href="#">SDF</a> <a href="#">PDB</a> <a href="#">Launch viewer</a>
1opj	STI	X-RAY	Structural basis for the auto-inhibition of c-Abl tyrosine kinase Nagar, B., Hantschel, O., Young, M.A., Scheffzek, K., Veach, D., Bornmann, W., Clarkson, B., Superti-Furga, G., Kuriyan, J.	1.75	<a href="#">Component definition</a> <a href="#">mmCIF</a> <a href="#">SDF</a> <a href="#">PDB</a> <a href="#">Launch viewer</a>
2oiq	STI	X-RAY	Crystal Structure of chicken c-Src kinase domain in complex with the cancer drug imatinib.	2.07	<a href="#">Component definition</a> <a href="#">mmCIF</a>





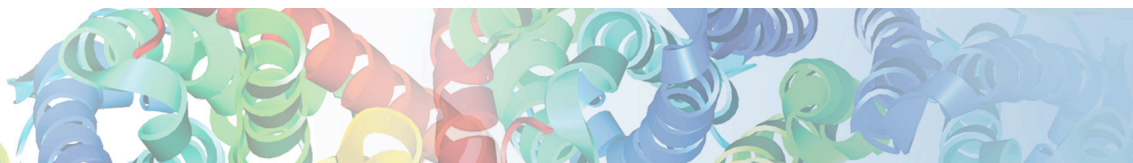
**Search Result Summary**

AstexViewer(TM)@MSD-EBI

**Molecular View**

	Resolution	Downloads & Views
vec , M.D., , B.C., , K.R.,	1.57	Component definition mmCIF SDF PDB <a href="#">Launch viewer</a>
1 M.L.,	1.6	Component definition mmCIF SDF PDB





## Experimental Examples

### Search Result Summary

Target ID: STI

Count in released entries: 12

### Instances of STI in Released Entries

PDB ID	CC ID	Method	Title/Authors	Resolution	Downloads & Views
<a href="#">1xbb</a>	STI	X-RAY	Crystal structure of the syk tyrosine kinase domain with Gleevec Nienaber, V.L., Atwell, S., Adams, J.M., Badger, J., Buchanan, M.D., Feil, I.K., Froning, K.J., Gao, X., Hendle, J., Keegan, K., Leon, B.C., Muller-Deickmann, H.J., Noland, B.W., Post, K., Rajashankar, K.R., Ramos, A., Russell, M., Burley, S.K., Buchanan, S.G.	1.57	<a href="#">Component definition</a> <a href="#">mmCIF</a> <a href="#">SDF</a> <a href="#">PDB</a> <a href="#">Launch viewer</a>
<a href="#">1t46</a>	STI	X-RAY	STRUCTURAL BASIS FOR THE AUTOINHIBITION AND STI-571 INHIBITION OF C-KIT TYROSINE KINASE Mol, C.D., Dougan, D.R., Schneider, T.R., Skene, R.J., Kraus, M.L., Scheibe, D.N., Snell, G.P., Zou, H., Sang, B.C., Wilson, K.P.	1.6	<a href="#">Component definition</a> <a href="#">mmCIF</a> <a href="#">SDF</a> <a href="#">PDB</a> <a href="#">Launch viewer</a>
<a href="#">1opj</a>	STI	X-RAY	Structural basis for the auto-inhibition of c-Abl tyrosine kinase Nagar, B., Hantschel, O., Young, M.A., Scheffzek, K., Veach, D., Bornmann, W., Clarkson, B., Superti-Furga, G., Kuriyan, J.	1.75	<a href="#">Component definition</a> <a href="#">mmCIF</a> <a href="#">SDF</a> <a href="#">PDB</a> <a href="#">Launch viewer</a>
<a href="#">1opj</a>	STI	X-RAY	Structural basis for the auto-inhibition of c-Abl tyrosine kinase Nagar, B., Hantschel, O., Young, M.A., Scheffzek, K., Veach, D., Bornmann, W., Clarkson, B., Superti-Furga, G., Kuriyan, J.	1.75	<a href="#">Component definition</a> <a href="#">mmCIF</a> <a href="#">SDF</a> <a href="#">PDB</a> <a href="#">Launch viewer</a>
<a href="#">2oiq</a>	STI	X-RAY	Crystal Structure of chicken c-Src kinase domain in complex with the cancer drug imatinib.	2.07	<a href="#">Component definition</a> <a href="#">mmCIF</a>

# RCSB PDB Structure Explorer Page

Are you missing data updates? The PDB archive has moved to <ftp://ftp.wwpdb.org>.  
For more information click [here](#).

- 1T46
- Download Files
  - FASTA Sequence
  - Download Original Files
- Display Files
  - Custom Structure Summary
  - PDB File
  - PDB File (Header)**
  - mmCIF File
  - mmCIF File (Header)
  - PDBML/XML File
  - PDBML/XML (Header)
- Display Molecule
  - Structural Reports
  - External Links
- Structure Analysis
- Help

Help Structure Summary Biology & Chemistry Materials & Methods Sequence Details Geometry


1t46  

DOI 10.2210/pdb1t46/pdb


Red - Derived Information

**Title** STRUCTURAL BASIS FOR THE AUTOINHIBITION AND STI-571 INHIBITION OF C-KIT TYROSINE KINASE

**Authors** Mol, C.D., Dougan, D.R., Schneider, T.R., Skene, R.J., Kraus, M.L., Scheibe, D.N., Snell, G.P., Zou, H., Sang, B.C., Wilson, K.P.


**Primary Citation** Mol, C.D., Dougan, D.R., Schneider, T.R., Skene, R.J., Kraus, M.L., Scheibe, D.N., Snell, G.P., Zou, H., Sang, B.C., Wilson, K.P. (2004) Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. *J.Biol.Chem.* **279**: 31655-31663  
[Abstract] 

**History** Deposition 2004-04-28 Release 2004-06-15


**Experimental Method** Type X-RAY DIFFRACTION Data 

Parameters	Resolution[Å]	R-Value	R-Free	Space Group
	1.60	0.190 (obs.)	0.213	P 3 <sub>2</sub> 2 1

Unit Cell	Length [Å]	a	b	c	Angles [°]
		70.09	70.09	127.88	
		90.00	90.00	120.00	

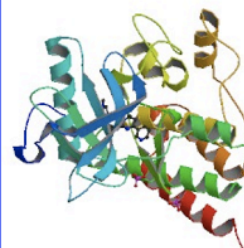
**Molecular Description** Polymer: 1 Molecule: Homo sapiens v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog Fragment: tyrosine kinase  
**Asymmetric Unit** Chains: A EC no.: 2.7.1.112   
Structure Weight: 36072.75

**Classification** **Transferase Activator**

**Source** Polymer: 1 Scientific Name: **Homo sapiens**  Common Name: **Human** Expression system: **Spodoptera frugiperda**

## Images and Visualization

<< Biological Molecule >>



### Display Options



- KING
- Jmol
- WebMol
- MBT Protein Workshop\*
- QuickPDB
- All Images

\* Capable of displaying biological molecules.









### Quick Tips:

To view **sequence details** of this structure click on the **Sequence Details** tab above the summary page.

# Structure Explorer Page for PDB 1T46

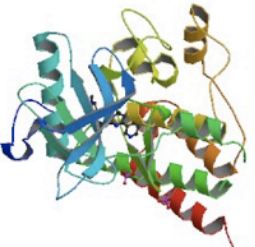
1t46   DOI 10.2210/pdb1t46/pdb


Red - Derived Information

<b>Title</b>	STRUCTURAL BASIS FOR THE AUTOINHIBITION AND STI-571 INHIBITION OF C-KIT TYROSINE KINASE						
<b>Authors</b>	Mol, C.D., Dougan, D.R., Schneider, T.R., Skene, R.J., Kraus, M.L., Scheibe, D.N., Snell, G.P., Zou, H., Sang, B.C., Wilson, K.P.						
<b>Primary Citation</b>	Mol, C.D., Dougan, D.R., Schneider, T.R., Skene, R.J., Kraus, M.L., Scheibe, D.N., Snell, G.P., Zou, H., Sang, B.C., Wilson, K.P. (2004) Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. <i>J.Biol.Chem.</i> 279: 31655-31663 [Abstract] 						
<b>History</b>	Deposition	2004-04-28	Release	2004-06-15			
<b>Experimental Method</b>	Type	X-RAY DIFFRACTION <a href="#">Data</a>					
<b>Parameters</b>	Resolution[Å]	R-Value	R-Free	Space Group			
	1.60	0.190 (obs.)	0.213	P 3 <sub>2</sub> 2 1			
<b>Unit Cell</b>	Length [Å]	a	b	c			
	Angles [°]	alpha	beta	gamma			
		70.09	70.09	127.88			
		90.00	90.00	120.00			
<b>Molecular Description Asymmetric Unit</b>	Polymer: 1 Molecule: Homo sapiens v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog Fragment: tyrosine kinase Chains: A EC no.: 2.7.1.112 						
	Structure Weight:	36072.75					
<b>Classification</b>	Transferase Activator						
<b>Source</b>	Polymer: 1	Scientific Name: <b>Homo sapiens</b> 	Common Name: <b>Human</b>	Expression system: <b>Spodoptera frugiperda</b>			
<b>Related PDB Entries</b>	Id	Details					
	1T45	Autoinhibited c-Kit					
	1PKG	Active c-Kit					
<b>Ligand Chemical Component</b>	Identifier	Name	Formula	Drug Similarity	Hapten Similarity	Ligand Structure	Ligand Interaction 
	PO4	PHOSPHATE ION	O <sub>4</sub> P			[ View ]	[ View ]
	STI	4-(4-METHYL-PIPERAZIN-1-YLMETHYL)-N-[4-METHYL-3-(4-PYRIDIN-3-YL-PYRIMIDIN-2-YLAMINO)-PHENYL]- BENZAMIDE	C <sub>29</sub> H <sub>31</sub> N <sub>7</sub> O			[ View ]	[ View ]

**Images and Visualization**

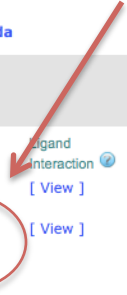
<< Biological Molecule >>



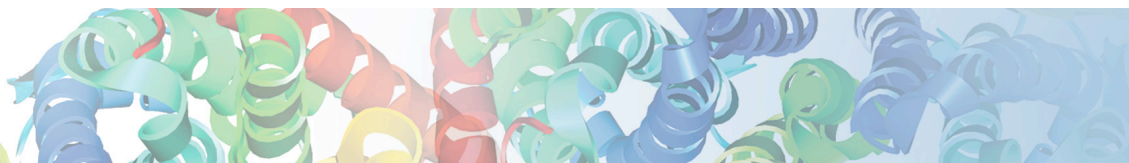
**Display Options** 

- KING
- Jmol
- WebMol
- MBT SimpleViewer\*
- MBT Protein Workshop
- QuickPDB
- All Images

\* Capable of displaying biological molecules.







## RCSB Chemical Details



### Chemical Component Summary [?](#)

1T46










<i>Name</i>	4-(4-METHYL-PIPERAZIN-1-YLMETHYL)-N-[4-METHYL-3-(4-PYRIDIN-3-YL-PYRIMIDIN-2-YLAMINO)-PHENYL]-BENZAMIDE 4-[(4-methylpiperazin-1-yl)methyl]-N-[4-methyl-3-[(4-pyridin-3-yl)pyrimidin-2-yl]amino]phenyl]benzamide
<i>ID</i>	STI
<i>Synonyms</i>	STI-571;IMATINIB
<i>Formula</i>	C <sub>29</sub> H <sub>31</sub> N <sub>7</sub> O
<i>Molecular Weight</i>	493.603 g/mol
<i>Formal Charge</i>	0
<i>Type</i>	NON-POLYMER
<i>Atom Count</i>	68
<i>Chiral Atom Count</i>	0
<i>Chiral Atoms</i>	
<i>Bond Count</i>	72
<i>Aromatic Bond Count</i>	24
<b>Descriptors</b>	
<i>SMILES (OpenEye OEToolkits)</i>	<chem>Cc1ccc(cc1Nc2nccc(n2)c3cccnc3)NC(=O)c4ccc(cc4)CN5CCN(CC5)C</chem>
<i>SMILES_CANONICAL (OpenEye OEToolkits)</i>	<chem>Cc1ccc(cc1Nc2nccc(n2)c3cccnc3)NC(=O)c4ccc(cc4)CN5CCN(CC5)C</chem>
<i>SMILES (CACTVS)</i>	<chem>CN1CCN(CC1)Cc2ccc(cc2)C(=O)Nc3ccc(C)c(Nc4nccc(n4)c5cccnc5)c3</chem>
<i>SMILES_CANONICAL (CACTVS)</i>	<chem>CN1CCN(CC1)Cc2ccc(cc2)C(=O)Nc3ccc(C)c(Nc4nccc(n4)c5cccnc5)c3</chem>
<i>InChI (InChI)</i>	InChI=1/C29H31N7O /c1-21-5-10-25(18-27(21)34-29-31-13-11-26(33-29)24-4-3-12-30-19-24)32-28(37)23-8-6-22(7-9-23)20-36-16-14-35(2)15-17-36 /h3-13,18-19H,14-17,20H2,1-2H3,(H,32,37)(H,31,33,34)
<i>SMILES (ACDLabs)</i>	<chem>O=C(Nc3ccc(c(Nc2nc(c1cccnc1)ccn2)c3)C)c4ccc(cc4)CN5CCN(CC5)C</chem>

# Structure Explorer Page for PDB 1T46

Help Structure Summary **Biology & Chemistry** Materials & Methods Sequence Details Geometry

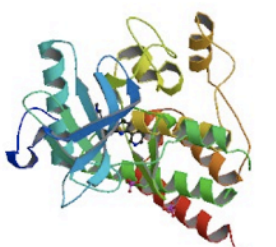
**1t46**   DOI 10.2210/pdb1t46/pdb


Red - Derived Information

<b>Title</b>	STRUCTURAL BASIS FOR THE AUTOINHIBITION AND STI-571 INHIBITION OF C-KIT TYROSINE KINASE						
<b>Authors</b>	Mol, C.D., Dougan, D.R., Schneider, T.R., Skene, R.J., Kraus, M.L., Scheibe, D.N., Snell, G.P., Zou, H., Sang, B.C., Wilson, K.P.						
<b>Primary Citation</b>	Mol, C.D., Dougan, D.R., Schneider, T.R., Skene, R.J., Kraus, M.L., Scheibe, D.N., Snell, G.P., Zou, H., Sang, B.C., Wilson, K.P. (2004) Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. <i>J.Biol.Chem.</i> 279: 31655-31663 [Abstract] 						
<b>History</b>	Deposition	2004-04-28	Release	2004-06-15			
<b>Experimental Method</b>	Type	X-RAY DIFFRACTION					Data 
<b>Parameters</b>	Resolution[Å]	R-Value	R-Free	Space Group			
	1.60	0.190 (obs.)	0.213	P 3 <sub>2</sub> 2 1			
<b>Unit Cell</b>	Length [Å]	a	70.09	b	70.09	c	127.88
	Angles [°]	alpha	90.00	beta	90.00	gamma	120.00
<b>Molecular Description Asymmetric Unit</b>	Polymer: 1 Molecule: Homo sapiens v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog Fragment: tyrosine kinase Chains: A EC no.: 2.7.1.112  Structure Weight: 36072.75						
<b>Classification</b>	Transferase Activator						
<b>Source</b>	Polymer: 1	Scientific Name:	Homo sapiens 		Common Name:	Human Expression system: Spodoptera frugiperda	
<b>Related PDB Entries</b>	Id	Details					
	1T45	Autoinhibited c-Kit					
	1PKG	Active c-Kit					
<b>Ligand Chemical Component</b>	Identifier	Name	Formula	Drug Similarity	Hapten Similarity	Ligand Structure	Ligand Interaction 
	PO4	PHOSPHATE ION	O <sub>4</sub> P			[ View ]	[ View ]
	STI	4-(4-METHYL-PIPERAZIN-1-YLMETHYL)-N-[4-METHYL-3-(4-PYRIDIN-3-YL-PYRIMIDIN-2-YLAMINO)-PHENYL]- BENZAMIDE	C <sub>29</sub> H <sub>31</sub> N <sub>7</sub> O			[ View ]	[ View ]

**Images and Visualization**


<< Biological Molecule >>



**Display Options** 

- KING
- Jmol
- WebMol
- MBT SimpleViewer\*
- MBT Protein Workshop
- QuickPDB
- All Images

\* Capable of displaying biological molecules.



# Ligand Explorer View

Full Sequences Contact Map

Various Molecules:

Chain A: `GNNYVYIDPTQLPYDHKWEFPRNRLSFGKTLGAGAFGKVVVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSLKVLSYLGNHMNIIVNLLGACTIGGPTLVIIEYCYGDDLNFLLRRKRDSFLALDLEDLLSFSYQVAKGMAFLASKNCRDLAARNILLTHGRITKICFGLARDIKNDSNYVVKGNARLPVKWMAPEIFNCVYTFESDVWSYGIPLWELFSLGSSPYPGMPVDSKFYKMIKEGFRMLSPEHAPAEMYDIMKTCWDADPLKRPTFKQIVQLIEKQISESTN`

**Step 1: Choose a ligand to analyze...**

- PO4 (4)
- PO4 (5)
- STI (3)

**Step 2: Choose interactions and thresholds...**

Inter-ligand 0.0 - 5.0

**...protein-ligand interactions...**

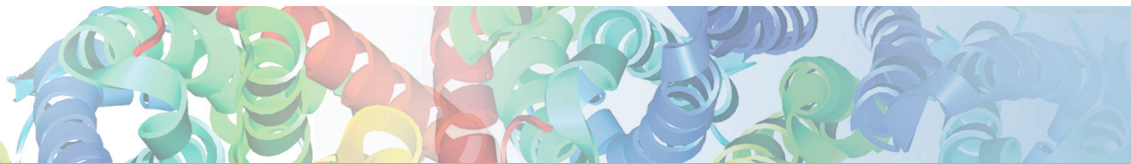
- Hydrophilic 2.7 - 3.3
- Hydrophobic 1.9 - 3.9
- Bridged H-Bond 0.0 - 5.0
- Hydrophobic 1.9 - 3.9

**Step 3: Miscellaneous...**

- Label interactions by distance

**Step 4: Finish...**

Apply Reset



## Dictionary Downloads

Ligand Expo

### Ligand Expo Downloads

#### Chemical component dictionaries

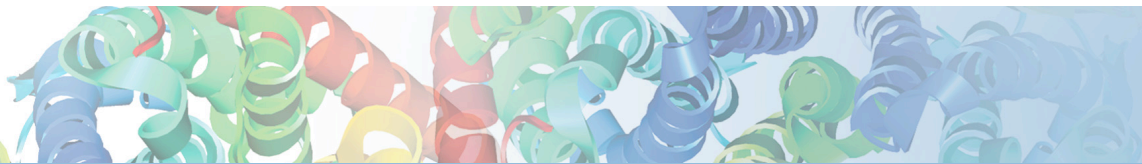
- [mmCIF format dictionary](#) ([text](#) | [gzip](#))
- [mmCIF format variants dictionary](#) ([text](#) | [gzip](#))
- [SDF/MOL format file of dictionary ideal geometries](#) ([gzip](#))
- [PDBML format dictionary](#) ([tar.gz](#))
- [PDBML format variants dictionary](#) ([tar.gz](#))

#### SMILES/InChI data files:

- [SMILES \(OpenEye with stereo\)](#) ([tab delimited text](#))
- [SMILES \(OpenEye\)](#) ([tab delimited text](#))
- [SMILES \(CACTVS with stereo\)](#) ([tab delimited text](#))
- [SMILES \(CACTVS\)](#) ([tab delimited text](#))
- [InChI](#) ([tab delimited text](#))
- [InChIKey](#) ([tab delimited text](#))

#### Chemical component coordinate data files<sup>1</sup>:

- Experimental coordinates for non-polymer molecules and non-standard amino acids and nucleotides are bundled in single data files  
([SDF/MOL \(gz\)](#) (>45MB) | [mmCIF \(gz\)](#) (>98MB) )



# RCSB PDB Ligand Search

CONTACT US | HELP | PRINT PAGE

PDB ID or keyword  Author

Site Search

Advanced Search

Are you missing data updates? The PDB archive has moved to <ftp://ftp.wwpdb.org>.  
For more information click [here](#).

Home Search

Search Database

- Advanced Search
- Latest Release
- Sequence
- Ligand Structure
- Ligand Name
- Ligand ID
- Models
- Unreleased Entries
- Structural Genomics
- Targets
- How to Search

Browse Database

Histograms

Quick Tips:

Custom reports can be generated in Microsoft Excel format. Click [here](#) for more information.

Match all of the following conditions:

Advanced Search Tutorial (Requires Flash)

- Ligand Name
- SCOP Classification (opens popup)
- CATH Classification (opens popup)
- Sequence Features
  - Sequence (Blast/Fasta)
  - Motif
  - Chain Length
  - Number Of Chains
  - Number Of Entities
  - in a Genome Location (opens popup)
- Modified Residues
- Ligand
  - Ligand Name
  - Ligand ID
  - Smiles
  - Has Ligand(s)
- Biology & Chemistry
  - Taxonomy (opens popup)
  - Enzyme Classification (opens popup)
  - Biological Process (GO) (opens popup)
  - Cell Component (GO) (opens popup)
  - Molecular Function (GO) (opens popup)

Contains:

Evaluate Subquery

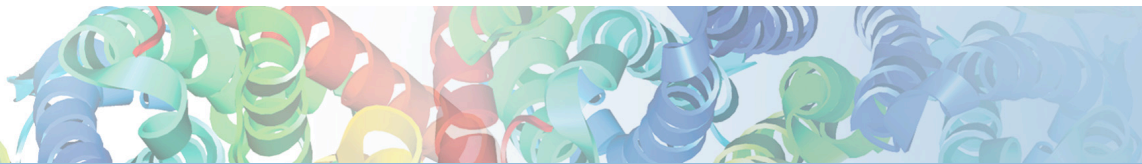
Clear All

Evaluate Query

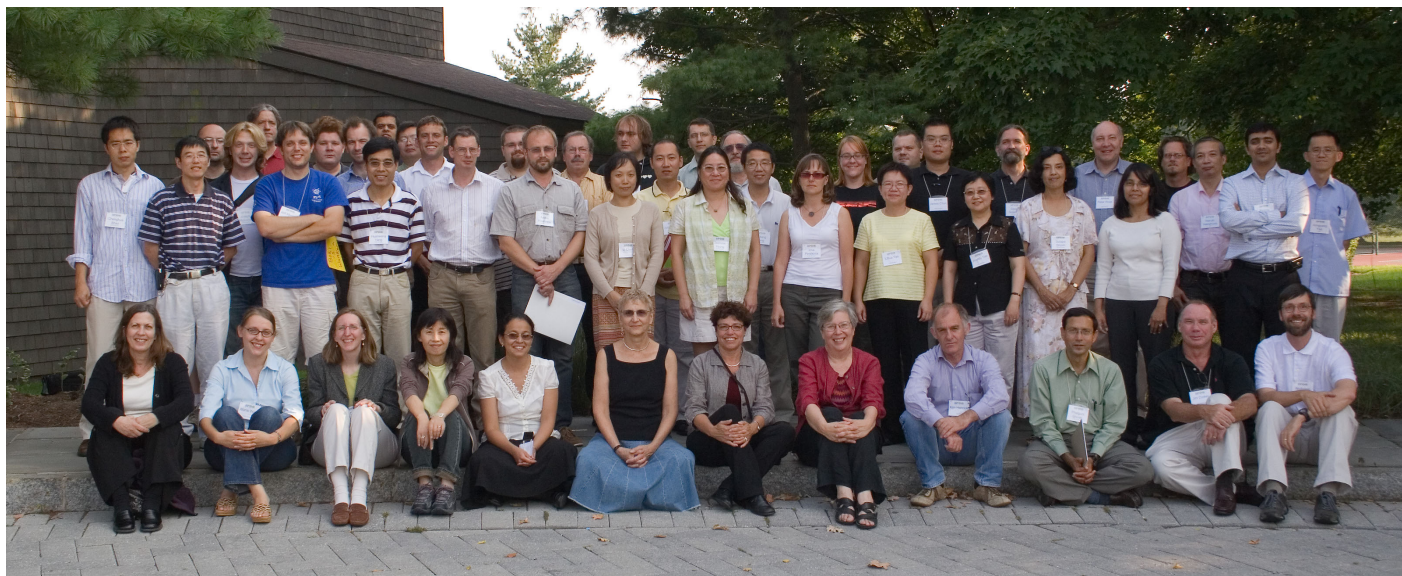


## Access

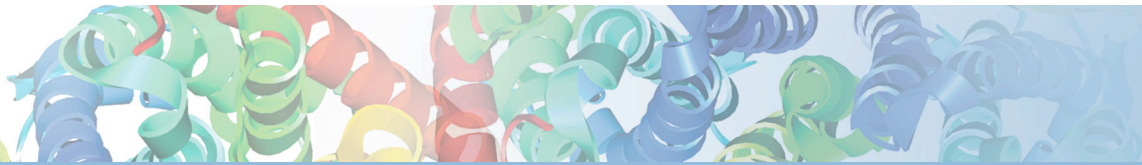
- RCSB Protein Data Bank Site
  - <http://www.rcsb.org/pdb/>
- wwPDB Site
  - <http://www.wwpdb.org/>
- Dictionary Resource Site
  - <http://mmcif.pdb.org/> & [http://pdbml.pdb.org /](http://pdbml.pdb.org/)
- Ligand Expo
  - <http://ligand-expo.rcsb.org>
- RCSB Software Download Site
  - [http://sw-tools.pdb.org /](http://sw-tools.pdb.org/)
  - CVS server [rcsb-cvs.rcsb.org](http://rcsb-cvs.rcsb.org) - anonymous access



W O R L D W I D E  
 P D B  
P R O T E I N D A T A B A N K



**wwPDB retreat September 2007**



## Acknowledgements

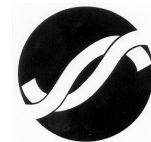
Operated by two  
members of the RCSB:



The RCSB PDB is  
a member of the



Supported by:



**NIGMS**

