



**PDBj**  
Protein Data Bank Japan

# OneDep: Unified System for Deposition, Biocuration, and Validation of Macromolecular Structures and Ligand Library based on CSD data

**Atsushi NAKAGAWA**

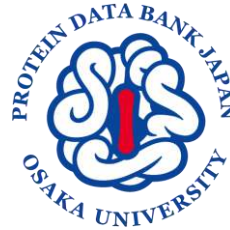
PDBj: Protein Data Bank Japan

Institute for Protein Research, Osaka University



[wwpdb.org](http://wwpdb.org)

第二十七屆生物物理研討會  
2023.5.18



**PDBj**  
Protein Data Bank Japan

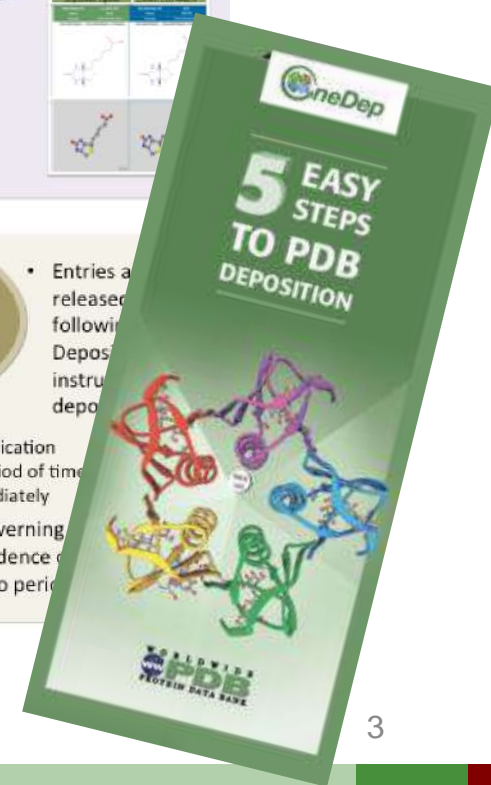
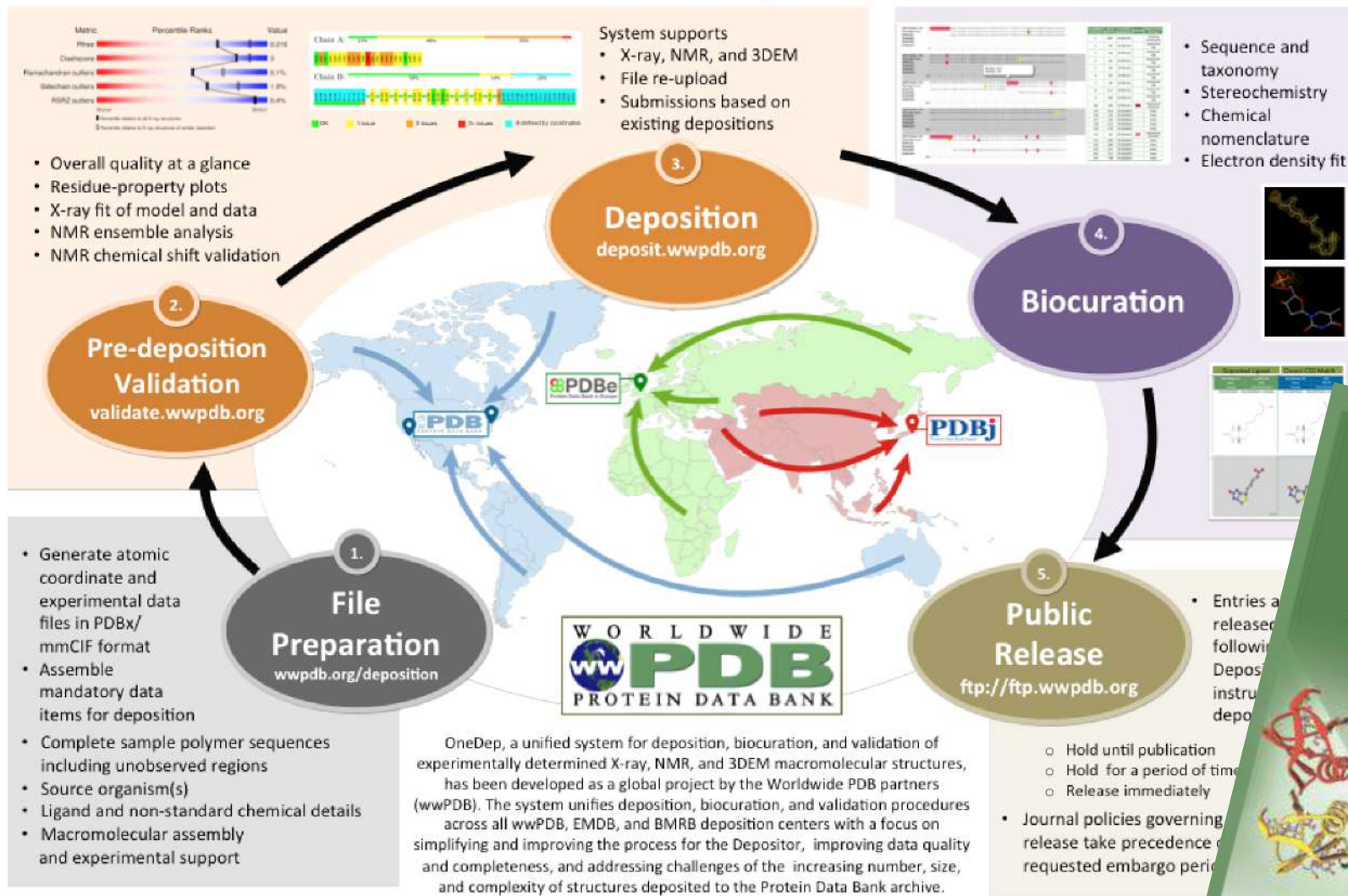
# OneDep: Unified System for Deposition, Biocuration, and Validation of Macromolecular Structures

WORLDWIDE  
**wwPDB**  
PROTEIN DATA BANK

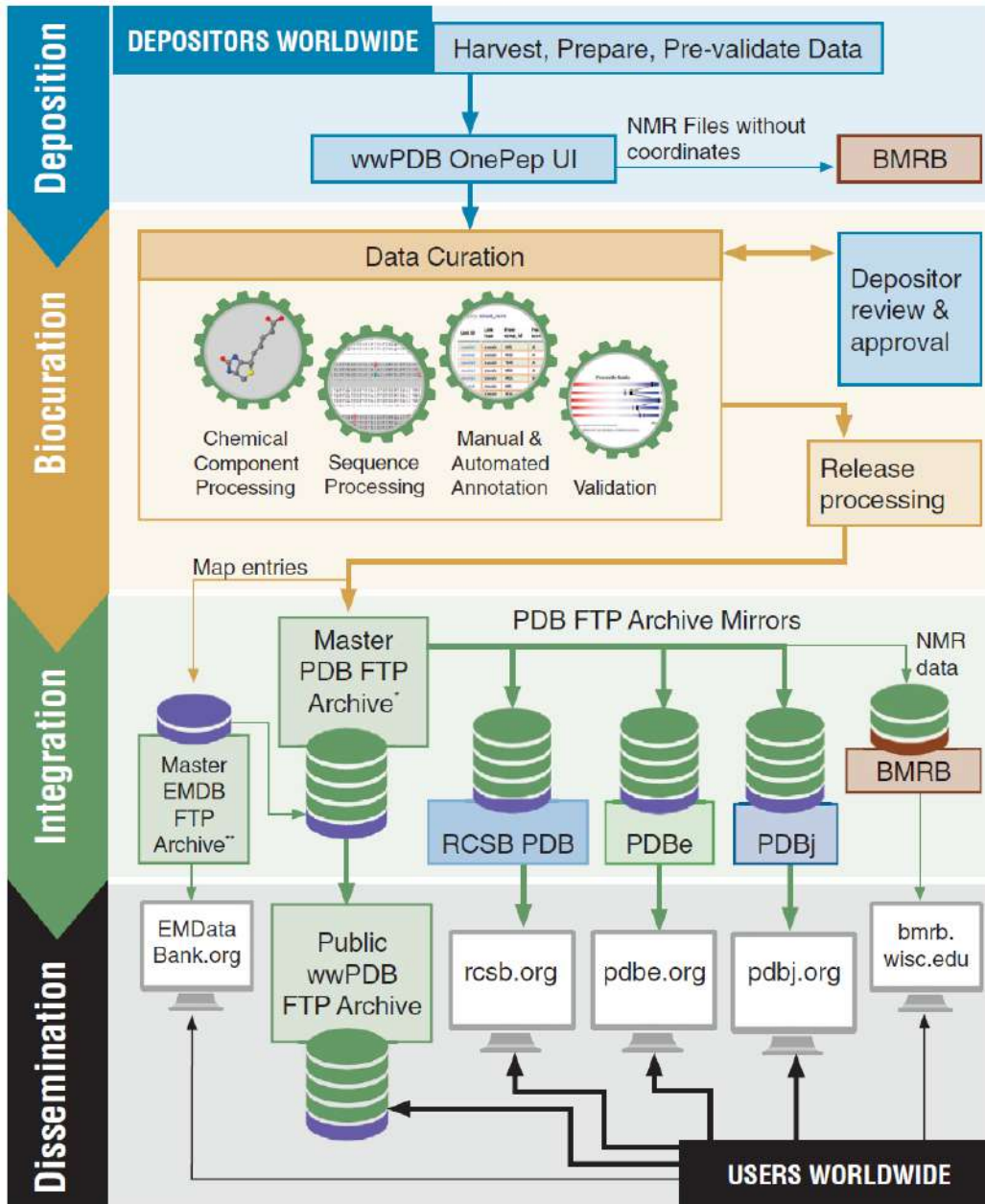
[wwpdb.org](http://wwpdb.org)

# OneDep: Unified Deposition Portal for the Protein Data Bank

wwPDB Partners - RCSB PDB, PDBe, PDBj, and BMRB



# Overview of OneDep system



\*Maintained by RCSB PDB \*\*Maintained by PDBe

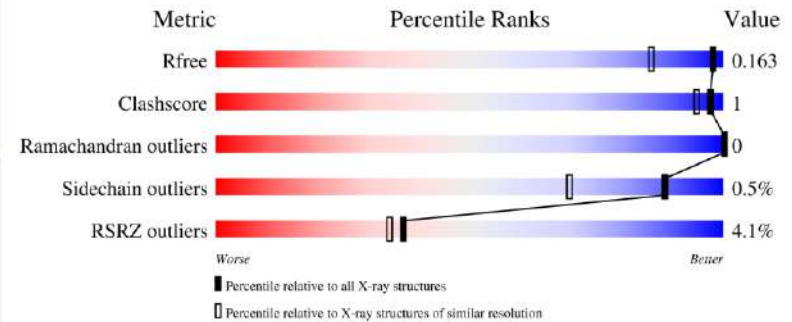
WORLDWIDE PDB PROTEIN DATA BANK  
Full wwPDB X-ray Structure Validation Report ⓘ

Jul 26, 2021 – 01:36 PM EDT

PDB ID : 7RGS  
Title : Crystal Structure of a Stable Heparanase Mutant  
Deposited on : 2021-07-14  
Resolution : 1.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report.  
This report is produced by the wwPDB biocuration pipeline after annotation of the structure.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

## wwPDB validation report (PDF)




## Percentile view of validation report

Young, JY et al., Structure, 25(3), 536-545, 2017


# wwPDB Common Deposition & Annotation

<https://wwpdb.org>





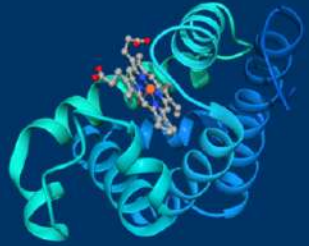
[VALIDATION](#) + [DEPOSITION](#) + [DICTIONARIES](#) + [DOCUMENTATION](#) + [TASK FORCES](#) + [DOWNLOADS](#) + [STATISTICS](#) + [ABOUT](#)






Since 1971, the Protein Data Bank archive (PDB) has served as the single repository of information about the 3D structures of proteins, nucleic acids, and complex assemblies.

The Worldwide PDB (wwPDB) organization manages the PDB archive and ensures that the PDB is freely and publicly available to the global community.

**Celebrating 50 Years of the PDB**



-  **Validate Structure**  
or View validation reports
-  **Deposit Structure**  
All Deposition Resources
-  **Download Archive**  
Instructions

### Vision and Mission

**Vision**

Sustain freely accessible, interoperating Core Archives of Structure data and metadata for biological macromolecules as an enduring public good to promote basic and applied research and education across the sciences.

**Mission**

- Manage the wwPDB Core Archives as a public good according to the FAIR Principles.
- Provide expert deposition, validation, biocuration, and remediation services at no charge to Data Depositors worldwide.
- Ensure universal open access to public domain structural biology data with no limitations on usage.
- Develop and promote community-endorsed data standards for archiving and exchange of global structural biology data.

### wwPDB Resources

**Data Dictionaries**

- **Macromolecular Dictionary (PDBx/mmCIF)**
- **Small Molecule Dictionary (CCD)**
- **Peptide-like antibiotic and inhibitor molecules (BIRD)**

**Biocuration**

- **Procedures and policies**
- **Improvements for consistency and accuracy**

**Community Input: Task Forces and Working Groups**

- **Validation Task Forces (X-ray, NMR, 3DEM)**
- **Small Angle Scattering Task Force**
- **PDBx/mmCIF Working Group**
- **Hybrid/Integrative Methods Task Force**
- **Ligand Validation Workshop**
- **ModelCIF Working Group**

**PDB Data Growth & Usage Statistics**

- **Depositions: by data center, by year, and by depositor location**
- **Downloads: by year for all entries**

**Workshops & Symposia**

- **2021: Celebrating 50 Years of the PDB**
- **Summaries and presentations from past meetings and events**

**Information for Journals**

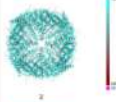
- **Policies, procedures, coordination with publishers, and preferred Instructions to Authors**

Cite wwPDB:  
*Nature Structural Biology* **10**, 980 (2003)  
doi: 10.1038/nsb1203-980

### News & Announcements

09/26/2022

➢ **Improved EM validation with Q-score**



wwPDB validation of EM structures for which there is both a model and an EM volume will include the Q-score metric.

[Read more](#)

09/05/2022

➢ **Future Planning: PDB entries with extended CCD or PDB IDs will be distributed in the PDBx/mmCIF format only**


pdb\_00099xyz

PDB users and related software developers should review code and begin to prepare for this change. Entries containing extended CCD or PDB IDs will not be distributed in legacy PDB file format.

[Read more](#)


08/22/2022

➢ **DNS name changes for PDB archive downloads from wwPDB**




Programmatic users (ftp, rsync or https) should update scripts as soon as possible.


### wwPDB Members

**Biological Magnetic Resonance Data Bank** 

Collects NMR data from any experiment and captures assigned chemical shifts, coupling constants, and peak lists for a variety of macromolecules; contains derived annotations such as hydrogen exchange rates, pKa values, and relaxation parameters.

**Protein Data Bank Japan** 

Supports browsing in multiple languages such as Japanese, Chinese, and Korean; SeSAW identifies functionality or evolutionarily conserved motifs by locating and annotating sequence and structural similarities, tools for bioinformaticians, and more.

**Protein Data Bank in Europe** 

# wwPDB Common Deposition & Annotation

<https://pd bj.org>



The screenshot shows the PDBj website interface. A red circle highlights the 'OneDep' (One-Deposition) service in the left sidebar. The main content area includes a search bar, a 'Find the service you need' section with a table of services, and a 'Latest news' section with a list of recent updates.

PDB	BMRB	EMDB
search	deposition	viewer
NMR	electron microscopy	secondary structure
similarity	function prediction	chemical component
binding site	surface structure	3D structure
RDF	SPARQL	gene
drug		disease

**Find the service you need**

Propose a keyword listed below or input keywords into the textbox at the right of the keyword list. The brief explanation of the matched services will be displayed.

- Click the 'Show all services' button to display the explanation for all services.
- Input some keywords into the 'Word Search Box' to narrow down the search results.

**OneDep**

幫助  
Deposition to PDB, EMDB or BMRB

下載  
下載PDB存檔 / 快照存檔

**標準格式**

PDBx/mmCIF Resources  
格式轉換  
PDBx/mmCIF editor

**Quick links**

幫助  
PDB格式不兼容數據  
Group Depositions  
Chemical Component entries  
Latest entries

**查詢**

幫助  
Search PDB (PDBj Mine)  
Search PDBj RDB  
Chemie search  
Search BMRB  
Sequence-Navigator  
DASH  
EM Navigator  
Omokage search  
Omokage搜索  
wwPDB/RDF  
NBDC RDF Portal  
Status search

**關於PDBj**

日本蛋白質結構數據庫 (PDBj: Protein Data Bank Japan), 是一個依托 [日本大阪大學蛋白質研究所](#), 與其聯合開發及共同研究的重點項目, 在運營維護全球統一生物大分子立體結構數據庫的同時, 還提供各種與結構解析相關的應用工具及數據庫。PDBj由 [JST-NBDC](#) (JPMJND2205) 和 [AMED-BINDS](#) 聯合提供支持, 其 [靈活的設計](#) 靈感來自於細胞色素c (PDBID: [1cyt](#)) 的立體結構, 它也是第一個在日本成功解析的蛋白質分子結構。

**PDBj 臨時暫停服務**

- 由於網絡維護, 一些 PDBj 服務 ([OneDep數據登記](#), [BSMA](#), [XRDa](#), [EMPIAR](#), [ef-site](#), [ef-seek](#), [ef-surf](#), [IV](#), [Omokage 搜索](#), [CRNPRED](#), [格式轉換](#), [HOMCOS](#)) 將無法使用
  - 5月3日星期三 (日本標準時間) 凌晨 0:00 到 6:00 之間, 最多 90 分鐘
  - 5月4日星期四晚上 10:00 到 5月5日星期五早上 6:00 (日本標準時間) 之間, 最多 50 分鐘
- 由於2023年5月3日 (星期三) 到5月5日 (星期五) 為日本國定假日, PDBj將暫停數據處理註冊服務, [wwPDB / PDBj站點的數據登記](#)可照常進行。

**最新消息**

- 2023-04-26 [We'll hold a luncheon seminar at the 27th Biophysical Society of Taiwan](#)
- 2023-04-26 [242 new PDB entries have been released on 2023-04-26.](#)
- 2023-04-04 [\[wwPDB\] Removal of Is-IR index file from the PDB archive](#)
- 2023-03-27 [\[wwPDB\] Access Depositions Using ORCID](#)
- 2023-03-24 [PDBj Newsletter Vol.23 is now available](#)
- 2023-03-11 [\[wwPDB\] Tribute to Dr. Olga Kennard](#)
- 2023-03-08 [\[wwPDB\] PDB entries with extended CCD or PDB IDs will be distributed in PDBx/mmCIF format only.](#)
- 2023-03-06 [\[wwPDB\] Deprecation of FTP File Download Protocol in the PDB Archive](#)
- 2023-02-15 [\[wwPDB\] Small Angle Scattering News](#)
- 2023-02-07 [\[wwPDB\] Prototype of PDB NextGen Archive now available](#)

[All news](#)

幫助 | 聯絡我們 | 使用條款 | 個人資料保護方針

PDBj 由 [JST](#) [NBDC](#) 支持, 並且是 [EBI](#) [portal](#) 的成員



deposit-pdbj.wwpdb.org

wwPDB Deposition

# wwPDB OneDep System

FAQ Tutorials

Existing deposition

Deposition ID

Password

Log in


Forgot Password

Sign in with ORCID

Validation server

Have you checked your data at the stand-alone validation server? [validate.wwpdb.org](http://validate.wwpdb.org)

wwPDB regions



wwPDB news and announcements

## ORCID Login

Contact authors can now use ORCIDiDs to authenticate OneDep access without the need for password sharing to view and access all their depositions in which the ORCID has been provided for the contact author. (See Tutorial [\[日本語\]](#) [\[English\]](#))

## Carbohydrate News

Carbohydrates will be renumbered and reassigned chain ids to provide consistent representation. For more details: Modernized uniform representation of carbohydrate molecules in the Protein Data Bank, (2021) *Glycobiology* 31: 1204-1218, doi: [10.1093/glycob/cwab039](https://doi.org/10.1093/glycob/cwab039).

[ [Help](#) | [ヘルプ](#) | [帮助](#) | [幫助](#) | [도움말](#) ]

Start a new deposition

### Welcome to the wwPDB OneDep system!

To continue with an existing deposition, please login on the left.  
Please note that un-submitted sessions will expire 3 months after last login. Un-submitted sessions and uploaded files will be removed once they expire.

**For requests such as entry release or citation updates, please login to the deposition system and send us a message through the communications section**

If you have any other feedback, please write to us at [deposit-help@mail.wwpdb.org](mailto:deposit-help@mail.wwpdb.org)  
At this time this deposition system does not work with Internet Explorer versions 8 or less.

**Warning: Please note that the current system does not support having multiple sessions open at the same time. To switch between existing sessions please log out using the "Log out" button in the bottom left corner of the opened deposition.**

---

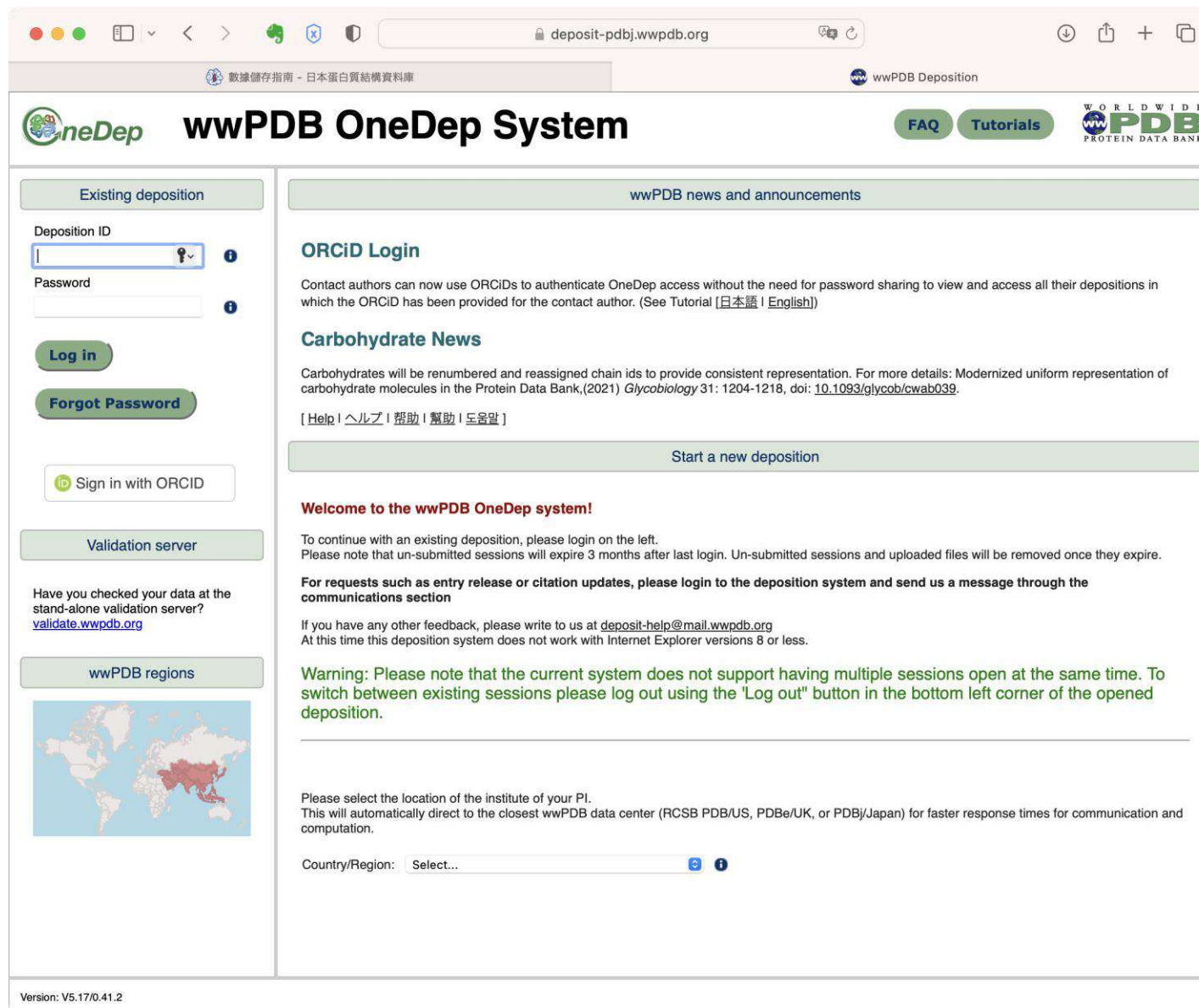
Please select the location of the institute of your PI.  
This will automatically direct to the closest wwPDB data center (RCSB PDB/US, PDBe/UK, or PDBj/Japan) for faster response times for communication and computation.

Country/Region:

Please select your country, here.

Version: V5.17/0.41.2

# After choosing your country/region



The screenshot shows the wwPDB OneDep System interface. The browser address bar is `deposit-pdbj.wwpdb.org`. The page title is "wwPDB OneDep System". The navigation bar includes "FAQ" and "Tutorials" buttons, and the "WORLDWIDE PDBj PROTEIN DATA BANK" logo. The main content is divided into two columns. The left column has three sections: "Existing deposition" with input fields for "Deposition ID" and "Password", "Log in", and "Forgot Password" buttons, and a "Sign in with ORCID" button; "Validation server" with a link to `validate.wwpdb.org`; and "wwPDB regions" with a world map. The right column has a "wwPDB news and announcements" section with "ORCID Login" and "Carbohydrate News" articles, and a "Start a new deposition" button. Below this is a "Welcome to the wwPDB OneDep system!" section with instructions and a "Warning" about multiple sessions. At the bottom, there is a "Country/Region" dropdown menu.

Existing deposition

Deposition ID

Password

Log in

Forgot Password

Sign in with ORCID

Validation server

Have you checked your data at the stand-alone validation server?  
[validate.wwpdb.org](http://validate.wwpdb.org)

wwPDB regions

wwPDB news and announcements

### ORCID Login

Contact authors can now use ORCIDiDs to authenticate OneDep access without the need for password sharing to view and access all their depositions in which the ORCID has been provided for the contact author. (See Tutorial [\[日本語\]](#) [\[English\]](#))

### Carbohydrate News

Carbohydrates will be renumbered and reassigned chain ids to provide consistent representation. For more details: Modernized uniform representation of carbohydrate molecules in the Protein Data Bank, (2021) *Glycobiology* 31: 1204-1218, doi: [10.1093/glycob/cwab039](https://doi.org/10.1093/glycob/cwab039).

[ [Help](#) | [ヘルプ](#) | [帮助](#) | [도움말](#) ]

Start a new deposition

### Welcome to the wwPDB OneDep system!

To continue with an existing deposition, please login on the left.  
Please note that un-submitted sessions will expire 3 months after last login. Un-submitted sessions and uploaded files will be removed once they expire.

**For requests such as entry release or citation updates, please login to the deposition system and send us a message through the communications section**

If you have any other feedback, please write to us at [deposit-help@mail.wwpdb.org](mailto:deposit-help@mail.wwpdb.org)  
At this time this deposition system does not work with Internet Explorer versions 8 or less.

**Warning: Please note that the current system does not support having multiple sessions open at the same time. To switch between existing sessions please log out using the "Log out" button in the bottom left corner of the opened deposition.**

Please select the location of the institute of your PI.  
This will automatically direct to the closest wwPDB data center (RCSB PDB/US, PDBe/UK, or PDBj/Japan) for faster response times for communication and computation.

Country/Region:

Version: V5.17/0.41.2




# Recent updates of OneDep system

- Validation report needed for peer review
- Improvements to visualization of ligand validation and electron density maps in the wwPDB validation report
- Mandatory PDBx/mmCIF format files submission for MX depositions
- Improve your previously-released coordinates and keep your original PDB ID
- Carbohydrate remediation
- PDBx/mmCIF data files to include PI information
- Extended PDB IDs and PDB DOIs now available in PDBx/mmCIF files
- Future Planning: PDB entries with extended CCD or PDB IDs will be distributed in the PDBx/mmCIF format only

# Official wwPDB Validation Report is requested for peer review



  
 Full wwPDB X-ray Structure Validation Report ⓘ

Jul 26, 2021 01:36 PM EDT


PDB ID : 7RG8  
 Title : Crystal Structure of a Stable Heparanase Mutant  
 Deposited on : 2021-07-14  
 Resolution : 1.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report.  
 This report is produced by the wwPDB biocuration pipeline after annotation of the structure.  
 We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
 A user guide is available at  
<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>  
 with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see references ⓘ) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as411be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.22
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrow)
Ideal geometry (proteins)	:	Engl & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.22

  
 Preliminary Full wwPDB X-ray Structure Validation Report ⓘ

Oct 14, 2022 - 02:15 PM JST

Deposition ID : D\_1300032869

**This wwPDB validation report is NOT for manuscript review**

This is a Preliminary Full wwPDB X-ray Structure Validation Report.  
 This report is produced by the wwPDB Deposition System during initial deposition but before annotation of the structure.  
 We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
 A user guide is available at  
<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>  
 with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at  
<http://www.wwpdb.org/validation/2017/FAQs/#types>.

---

The following versions of software and data (see references ⓘ) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as411be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrow)
Ideal geometry (proteins)	:	Engl & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

# Validation metrics in wwPDB validation reports

## X-ray/EM/NMR

- Geometric & conformational
  - bond, angle, planarity
  - protein backbone conformation
  - protein side-chain conformation
- Atomic & molecular interaction
  - all-atom contacts
  - under packing
  - hydrogen bond quality
- Non-protein
  - nucleic acids (RNA pucker, suite)
  - carbohydrates (N-glycan core)
  - ligands (CSD)
  - ions & other solvent
- Incomplete model (e.g. CA\_ONLY)

## X-ray

- Structure factor & electron density
  - Wilson plot, outliers, tNCS
  - wrong symmetry
  - twinning
  - agreement ( $R_{\text{free}}$ , RSR, RSCC)

## NMR

- Chemical shifts
  - completeness
  - outliers
  - estimated reference error
  - random coil index
- Structure ensembles
  - representative model (medoid)
  - domain detection

### Caveat:

LLDF (Local Ligand Density Fit) has been replaced by a combination of RSR (Real-space R factor) > 0.4 and RSCC (Real-space correlations coefficient) < 0.8 since this March.

# wwPDB validation report PDF

## • Standard geometry

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.47	0/1107	0.71	0/1491

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## • Too close contacts

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1091	0	1106	7	0
2	A	22	0	27	2	0
3	A	100	0	0	2	0
All	All	1213	0	1133	9	0

## • Protein backbones

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	135/137 (98%)	132 (98%)	3 (2%)	0	100 100

## • Protein sidechains

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	123/123 (100%)	120 (98%)	3 (2%)	52 38

## • Ligand geometry

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	REA	A	200	-	19,22,22	1.05	1 (5%)	26,30,30	1.02	2 (7%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	200	REA	C1-C6	2.25	1.56	1.53

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	200	REA	C11-C10-C9	-2.40	123.89	127.31
2	A	200	REA	C18-C5-C6	2.08	126.83	124.51

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	200	REA	2	0

# Validation software utilized for generation of wwPDB validation report (2022)

Component Software Packages Included in the 2022 Version of the Validation Pipeline

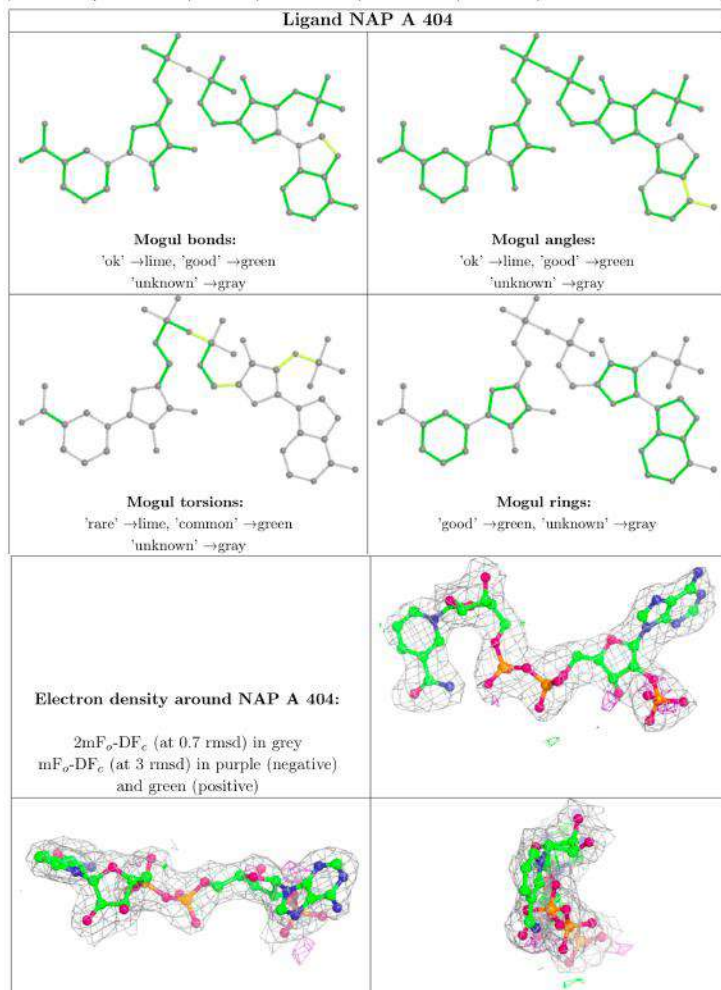
Software Package	Version
MolProbity	4.02b-467
Mogul	1.8.5 (274361), CSD as541be (2020)
Xtrriage (Phenix)	1.13
EDS	2.31.2
buster-report	1.1.7 (2018)
Percentile statistics	20191225.v01
Refmac	5.8.0158
CCP4	7.0.044 (Gargrove)
Ideal geometry (proteins)	Engh & Huber (2001)

# Ligand Validation

- Adapted software from Global Phasing Ltd. under a formal agreement
- Provides geometrical quality in 2D depiction
- Provides electron density fit for X-ray
- Now mandatory at deposition: identification of Ligand/s Of Interest (LOI, author's research focus)

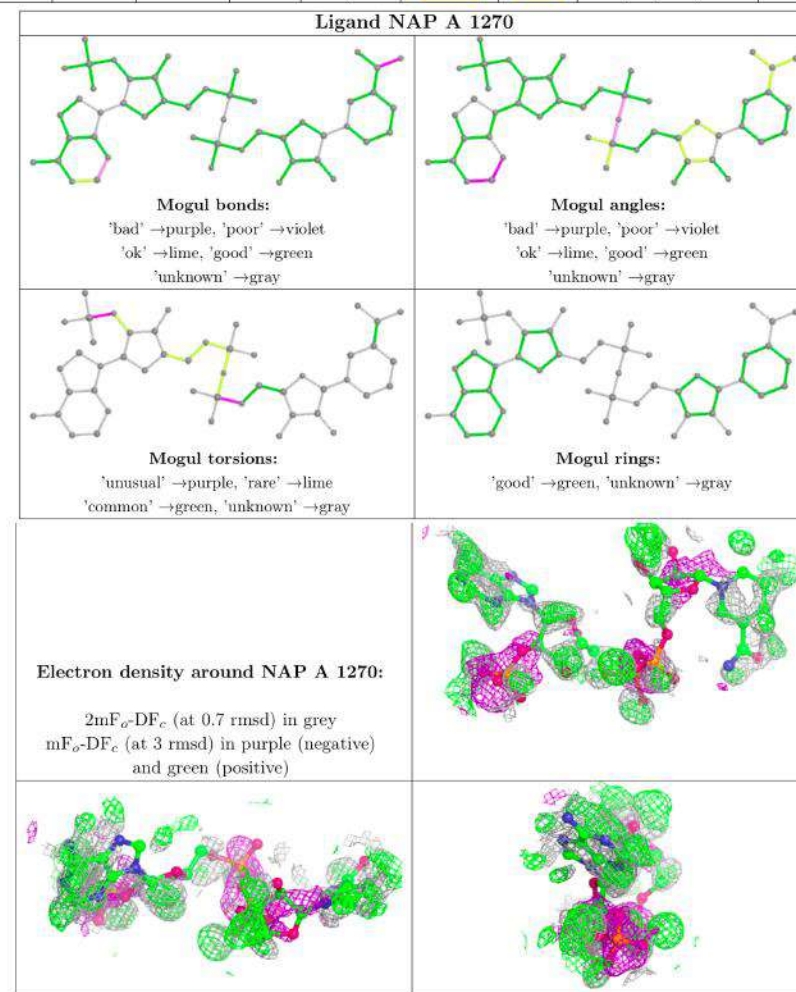
# Ligand Validation- Examples on NADP Ligands

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAP	A	404	48/48	0.96	0.14	31,43,66,70	0



PDB entry 5zix (Better data quality)

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAP	A	1270	48/48	-0.06	0.67	87,96,100,100	0



PDB entry 1zk4 (Worse data quality)

# Mandatory PDBx/mmCIF format files submission for MX depositions

## PDB File formats from wwPDB

- (Legacy) PDB format
  - *NOT SUPPORTED !*
- **PDBx/mmCIF**
  - The canonical format of the wwPDB.
  - Ver. 5.361 released. (2022-08-31)



# Coordinate Versioning

- To improve data quality in the PDB archive
- Depositors can now make corrections to existing structures in the PDB Core Archive by updating the atomic coordinates while preserving the original PDB identifier.
- Deployed on July 27<sup>th</sup> 2019 for all structures that were originally deposited via OneDep (Phase one)
- First coordinate replacement (PDB entry 5T26) versioned at FTP on August 7<sup>th</sup> 2019 following reviewers' comments based on the wwPDB validation reports
  - With PDB prefix and extension of 4 characters (e.g., from “1ABC” to “PDB\_00001ABC”)
  - Example: PDB\_00001ABC\_XYZ\_V2-2.cif.gz

# Deposition Notes

- Model coordinates
  - TER records
  - Coordinates for Chimera protein
  - ALA models
- Sample sequence
  - Complete polymer sequence
  - Non-poly residues
  - UNK
  - Source organism
- Assembly details
- Deposition efficiency
- Communication, contact authors & ORCID
- Release

# Model coordinates

**TER cards must follow each protein chain,  
but can only follow a protein (or nucleotide) chain.**

Chain A protein

TER ←

Chain A ligand

Chain B protein

TER ←

Chain B ligand

Chain C protein

TER ←

Chain C ligand

Chain D ligand

# ALA model

- When an amino acid residue is disordered due to low density

Side chain atoms cannot be assigned, and the residue is often modeled as:

- ALA model
- GLY model
- SER model

Etc.

**However, the residue name in the coordinates should not be changed to MATCH with the sequence even without the atoms in the side chain.**

# Sample Sequence

Fill in a complete polymer sequence used for experiment

- Please include

HIS- or other Expression tags, Linker, Residues missing from the coordinates due to disorder

Coordinates: . . . . . LVVVTNNLR . . . RIPGIRIED . . . ITLMELILEH . . . . .

Sequence: **GSHMALVVVTNNLREFERIPGIRIEDGSGITLMELILEHHHHH**

- Please don't include

Residues cleaved from the macromolecules prior to or during the experiment

# DO NOT include ligands

Sample sequence is a list of the consecutive chemical components covalently linked in a linear fashion to form a polymer.

Please don't include:

- Metal ions, Chemical components or groups covalently linked to side-chains (in peptides)
- Floating Metal ions, Chemical components or groups

incorrect sequence: NLREFERIPG(NAG)IRIEDYTYITLMELILEHHH(NAG)(ZN)

(Should be) NLREFERIPGIRIEDYTYITLMELILEHHH

## Use "UNK" or "N" ONLY in TWO special cases

UNK : unknown amino acid

N or DN : unknown nucleotide

1. You don't know the sequence
2. You know the sequence but don't know how the coordinates align with the sequence

(Example)

You know the sequence but are not sure if the first residue seen in the density is really the first residue of the sequence.

(Note)

Please use "UNK" or "N" in the coordinates only when you use "UNK" or "N" in sequence.

# Provide details of source information

For an exact sequence database reference match,  
please provide detail information (e.g. strain)  
in addition to source organism name.

Example)

*Escherichia coli* K-12

*Saccharomyces cerevisiae* S288c



# Deposition efficiency

- Deposit Multiple related entries
- Deposit an entry to supersede an existing entry
- Re-submit withdrawn entries



OneDep provides an option to copy information from a deposited entry

**Navigation**

- ✓ Instructions
- ✓ Communication
- ! **File upload**

**Log out**

**wwPDB Deposition: D\_800091 -- Requested ID: PDB**

**File upload**

**X-Ray upload information**

Mandatory submission:

- Model coordinate file.
- Structure factor file used for the refinement. This can either be in mmCIF or mtz formats and should at least include h, k, l, F, SigmaF (and/or ...)

The processed files will be made publicly available upon release of the entry.

**Based on a previous wwPDB deposition**

Do you want to import information from a previous wwPDB deposition?  Yes  No

Previous deposition ID:

Previous deposition password:

What data items would you like to transfer from a previous deposition?

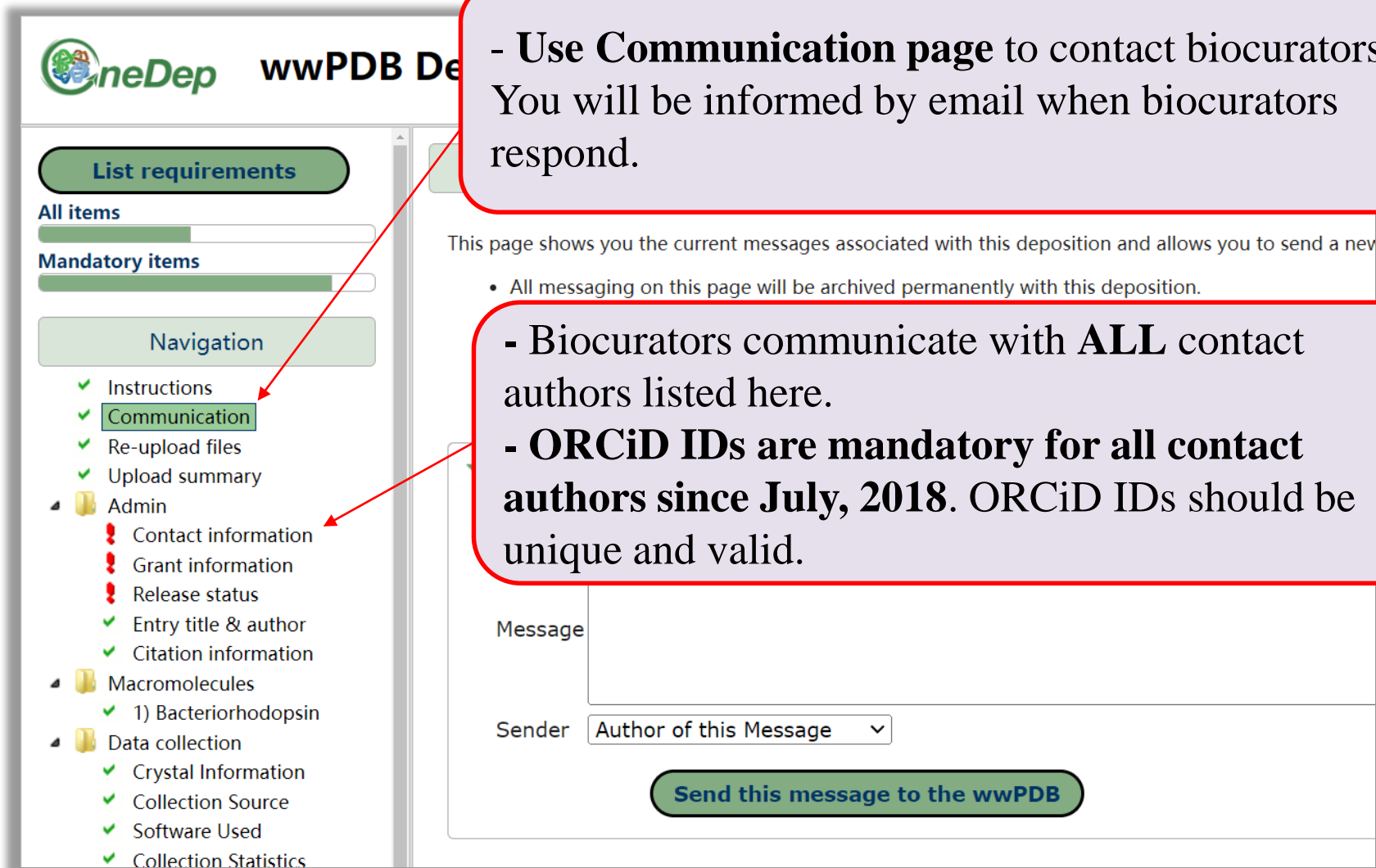
- Contact information
- Entry title & author
- Citation information
- Grant information

**Check**

**Input PDB ID and password of an existing entry**

**Check the information needed**

# Communication, Contact authors & ORCiD



**- Use Communication page to contact biocurators. You will be informed by email when biocurators respond.**

This page shows you the current messages associated with this deposition and allows you to send a new message.

- All messaging on this page will be archived permanently with this deposition.

**- Biocurators communicate with ALL contact authors listed here.**

**- ORCiD IDs are mandatory for all contact authors since July, 2018. ORCiD IDs should be unique and valid.**

Message

Sender

**Send this message to the wwPDB**

# Release

- Do not ask for replacement of coordinates just before the release
- When you refer to PDB in your paper, please refer as **“the coordinates are deposited to the wwPDB”**.
  - Your entries are processed by PDBj, RCSB or PDBe
- Release instructions (REL, HPUB, HOLD) cannot be changed by depositors
  - Ask biocurators after submission

# DOI Landing Page Layout



[VALIDATION](#) ▾ [DEPOSITION](#) ▾ [DICTIONARIES](#) ▾ [DOCUMENTATION](#) ▾ [TASK FORCES](#) ▾ [DOWNLOADS](#) ▾ [STATISTICS](#) ▾ [ABOUT](#) ▾



PDB Entry - 5ZIL

(Status - Released)

## Summary information:

**Title:** Crystal structure of bacteriorhodopsin at 1.29 Å resolution

**DOI:** [10.2210/pdb5zil/pdb](https://doi.org/10.2210/pdb5zil/pdb)

**Primary publication DOI:** [10.1038/s41598-018-31370-0](https://doi.org/10.1038/s41598-018-31370-0)

**Entry authors:** Hasegawa, N., Jonotsuka, H., Miki, K., Takeda, K.

**Initial deposition on:** 16 March 2018

**Initial release on:** 10 October 2018

**Latest revision on:** 5 May 2021

## Downloads:

[Structure coordinates \(PDBx/mmCIF\)](#)

[Structure coordinates \(PDBML\)](#)

[Structure coordinates \(PDB\)](#)

[X-ray diffraction data \(PDBx/mmCIF\)](#)

[Validation report \(PDF\)](#)

[Validation report \(XML\)](#)

## Links to more resources for 5ZIL at:



[News & Announcements](#)

[Download Archive](#)

[RCSB PDB ftp](#) | [PDBe ftp](#) | [PDBj ftp](#)  
[Instructions](#)

[Cite wwPDB:](#)

*Nature Structural Biology* **10**, 980 (2003)  
[doi: 10.1038/nsb1203-980](https://doi.org/10.1038/nsb1203-980)



**PDBj**  
Protein Data Bank Japan

# Ligand Library based on Cambridge Structural Database (CSD) data



[wwpdb.org](http://wwpdb.org)

# Refinement of structures

It is usually difficult to obtain a sufficient number of observations for refinement of atom parameters due to low resolution in X-ray crystallography

Observations/parameter ratio (isotropic B-factor: 4 parameters/atom)

5Å ... 0.14

3.5Å ... 0.4

2.5Å ... 1

2Å ... 2.6

1.6Å ... 4

1.4Å ... 20

1.2Å ... 35



Require restraints

Geometry restraints ... `refmac5`, `buster`

Energy restraints ... `phenix.refine`

# Refinement of structures

(eg. refmac5, BUSTER)

- Restrained parameter

- Stereochemical restrains are combined with X-ray term

- $f_{\text{total}} = f_{\text{geom}} + w \cdot f_{\text{xray}}$

- (eg.  $\sum_{\text{bond}} \frac{1}{\sigma_b^2} (b_m - b_i)$ )

(eg. phenix.refine)

- Energy minimization (with simulated annealing)

- Observations (structure factors) are treated as a part of energy term

- $T_{xyz} = wXC_{\text{scale}} \cdot T_{\text{exp}} + wC \cdot T_{\text{xyz\_restraints}}$

Ideal (target) geometries obtained by CSD data (Engh & Huber, 1991)

# Restraints Library (dictionary)

## refmac5

- \$CLIBD/standard\_geometry.cif
  - Amino acids
    - Engh and Huber, 1999
      - » Engh, R.A. & Huber, R. (2001). International Tables for Crystallography, Vol F, pp 382-392.
    - 他... (水素原子)
  - Nucleic acids
    - Clowney *et al.*, 1996
      - » Clowney *et al.* (1996), J. Am. Chem. Soc., 118, 509-518.
      - » Parkinson *et al.* (1996) Acta Cryst. D 52, 57-64.
    - NDB (Nucleic Acid Database)
      - » <http://ndbserver.rutgers.edu/>
- \$CLIBD/monomers/...
  - Ligands

## phenix.refine

- Conformation Dependent Library (CDL)
- \$CLIBD/monomers/...
  - Ligands



# Preparation of Ligand Library

AceDRG 

F. Long *et al.* ACEDRG: A stereo-chemical description generator for ligands.  
*Acta Crystallogr. Sect. D Biol. Crystallogr.*, **73**, 112–122 (2017).

eLBOW 

N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams. Electronic ligand builder and optimization workbench (eLBOW): a tool for ligand coordinate and restraint generation.  
*Acta Crystallogr. D Biol. Crystallogr.* **65**, 1074–1080 (2009)

Grade 

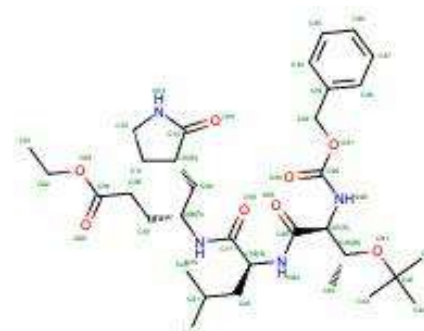
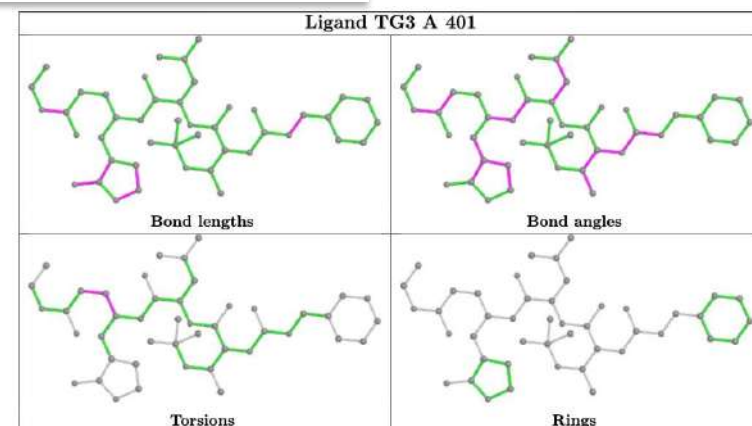
Smart, O. S., Womack, T. O., Sharff, A., Flensburg, C., Keller, P., Paciorek, W., Vonrhein, C. and Bricogne, G. (2011).  
Grade, version 1.2.20. Cambridge, United Kingdom, Global Phasing Ltd.

Grade2 

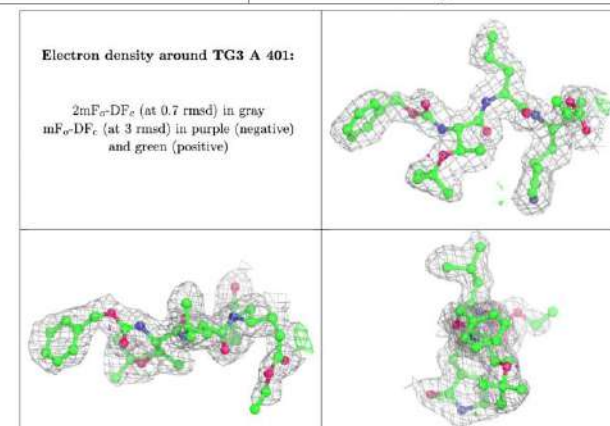
Smart, O.S., Sharff A., Holstein, J., Womack, T.O., Flensburg, C., Keller, P., Paciorek, W., Vonrhein, C. and Bricogne G. (2021).  
Grade2 version 1.0.0. Cambridge, United Kingdom: Global Phasing Ltd.

# wwPDB Validation Improvement: Ligand

- Partnership with Global Phasing Ltd.
- Ligand of Interest (LOI) highlighted
- 2D views of geometrical quality
- 3D views of electron density fits for X-ray



Ligand ID TG3



PDB ID 7JT7

Feng, Z. *et al.*, Structure, 2021 doi: [10.1016/j.str.2021.02.004](https://doi.org/10.1016/j.str.2021.02.004)

# wwPDB validation report uses buster-report (by Global Phasing Inc.)

06/11/2019

## Improvements to visualization of ligand validation and electron density maps in the wwPDB validation report

Our recent update to the wwPDB validation reports provides much clearer validation information for ligands.

We now include 2-dimensional diagrams of ligands, highlighting geometric validation criteria and, for structures determined by crystallography, 3-dimensional views of electron density.

We also provide calculated electron density map coefficients which were used to generate the analysis in the validation reports.

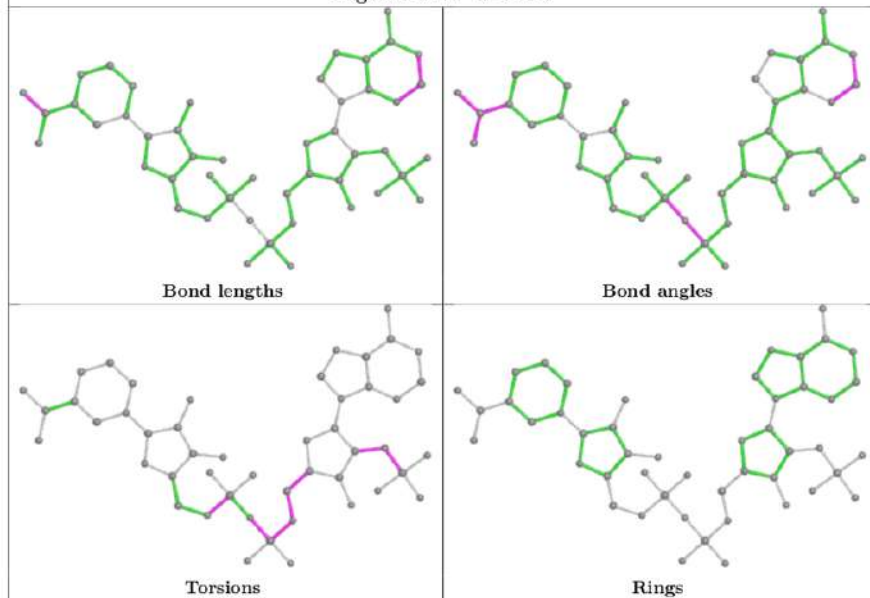
### Ligand Validation

We have collaborated with **Global Phasing Ltd** to integrate the ligand visualization from **buster-report** into the wwPDB validation report, as recommended by the **wwPDB/CCDC/D3R Ligand Validation Workshop**. The ligand visualization will be available for ligands that have been designated as "Ligand of Interest" by the depositor and ligands with a molecular weight greater than 250 Daltons that have outliers.

The following ligand instance of NAP was chosen intentionally as a representative of sub-optimal quality in both the ligand model and its agreement with the X-ray data.

Geometric analysis provided by CCDC Mogul will be highlighted on a 2D diagram of the ligand, as shown below.

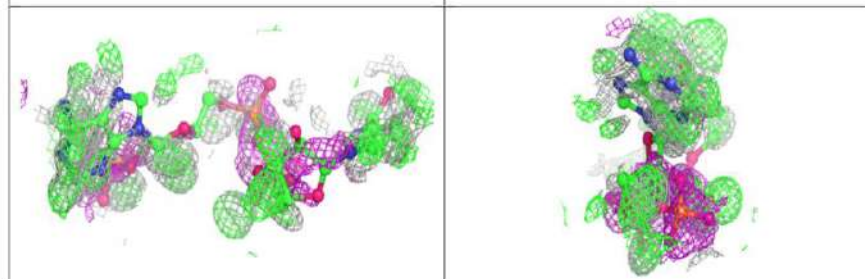
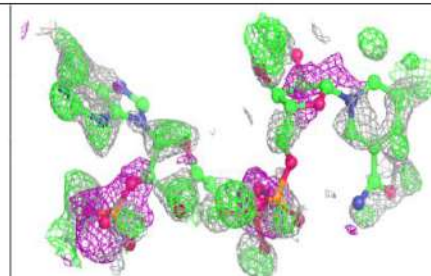
Ligand NAP A 1270



In addition to geometric validation for ligands, for X-ray diffraction PDB entries the wwPDB validation report also presents images displaying the ligand and the surrounding electron density map.

### Electron density around NAP A 1270:

$2mF_o - DF_c$  (at 0.7 rmsd) in gray  
 $mF_o - DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



### Electron Density Map Coefficient Files

We are now providing depositors with electron density map coefficient files ( $2mF_o - DF_c$  and  $mF_o - DF_c$ ) from the wwPDB validation pipeline alongside the wwPDB validation report. The electron density map coefficients generated for wwPDB validation reports will be made available to end users in the PDB archive as new entries are released and for existing entries when validation reports for the PDB archive are recalculated.

We hope that these changes to the wwPDB validation pipeline will help depositors to interpret the validation information provided for PDB entries more easily. If you have any queries, please contact the wwPDB at [deposit-help@mail.wwpdb.org](mailto:deposit-help@mail.wwpdb.org).

# BUSTER

- **BUSTER** is a refinement package developed by **Global Phasing Ltd.** (Director: Gérard Bricogne)
- s a maximum-likelihood macromolecular refinement package based on **TNT** (Dale Trounrud and Lyn TenEyck) (requires TNT license)
- **BUSTER** assembles the structural model, scales observed and calculated structure-factor amplitudes and computes the model likelihood, whilst TNT handles the stereochemistry and NCS restraints/constraints and shifts the atomic coordinates, B factors and occupancies.
- Requires **CCP4** program suite

# BUSTER

- **grade** uses the RM1 semiempirical method as implemented in the **fdynamo** library written by Martin Field and co-workers.
- **visualise-geometry-coot** can visualize the results of a **BUSTER** refinement
- **grade** (and **grade2**) typically runs the CCP4 libcheck program to produce an initial dictionary for a given input SMILES string or structure.
- This initial dictionary is then improved: grade runs the **CCDC mogul** program (in batch mode) to obtain ideal bond lengths/sigma, bond angles/sigmas and planar torsions from similar structures in the **CSD**.
- Using **OpenBabel** (<http://openbabel.org>) to convert a PDB file containing 3D coordinates to a MOL2 file fills in the atom types and bond orders.
- **reduce** from **MolproBily** is used to add all hydrogens explicitly and to perform HQN flip

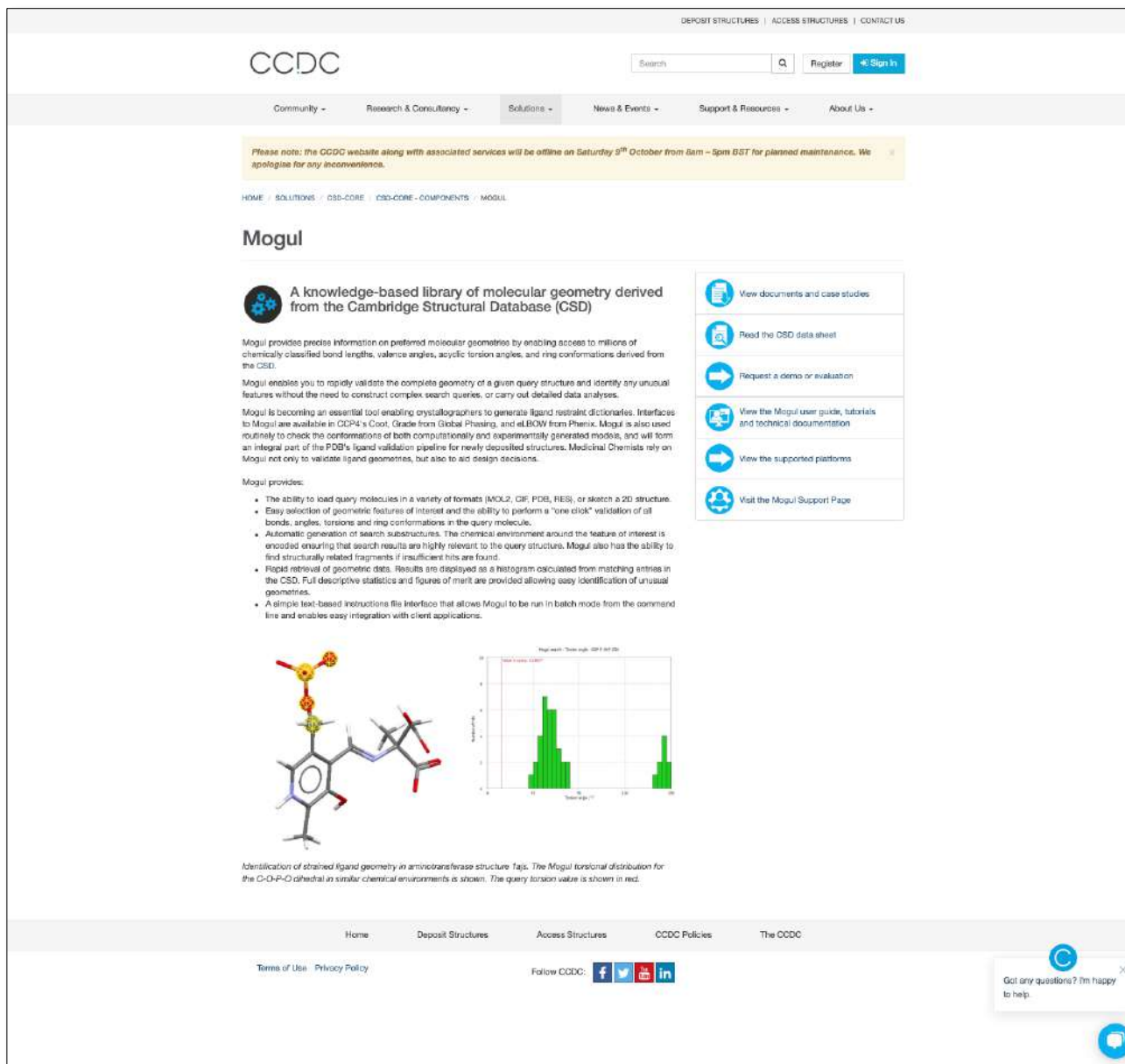
# BUSTER Input data

- PDB ... checked by pdbchk (in BUSTER)
  - each atom should have a chain identifier (e.g A, B and C for protein chains and W for water)
  - a correct CRYST1 card is required (see PDB format guide), especially the space group symbol.
  - although not enforced by the PDB standard, it seems sensible to use letters (A, B, C etc) in column 17 of ATOM/HETATM records to denote alternate conformations and numbers (1, 2, 3 etc) in column 27 of ATOM/HETATM records to denote insertion code.
  - standard SSBOND and LINK records are supported and often required (to describe correctly the molecular connectivity).

# BUSTER Input data

- MTZ
  - normal MTZ file with F/SIGF columns (any column name is possible, but the column types have to be F/Q - which they nearly always are anyway, unless something went really wrong)
  - the cell parameters for the refinement are taken from the MTZ file header (please note that it does not yet handle different cell entries for different datasets). The assumption is that the MTZ file usually contains only a single dataset.
  - if the MTZ file contains a set of columns with Hendrickson-Lattmann coefficients (usually named HLA, HLB, HLC and HLD) these can be used as additional, external phase information (unless the MTZ file is actually the output of a previous BUSTER run - which would not be a good idea). The user needs to set the parameter `autoBUSTER_hls` to the four column names, e.g. with `'refine autoBUSTER_hls="HLA HLB HLC HLD" ... '`.

# Mogul in CSD



DEPOSIT STRUCTURES | ACCESS STRUCTURES | CONTACT US

CCDC


Search [ ] Register Sign In

Community - Research & Consultancy - Solutions - News & Events - Support & Resources - About Us -

Please note: the CCDC website along with associated services will be offline on Saturday 9<sup>th</sup> October from 8am - 5pm BST for planned maintenance. We apologise for any inconvenience.

HOME / SOLUTIONS / CSD-CORE / CSD-CORE - COMPONENTS / MOGUL

## Mogul

 A knowledge-based library of molecular geometry derived from the Cambridge Structural Database (CSD)







Mogul provides precise information on preferred molecular geometries by enabling access to millions of chemically classified bond lengths, valence angles, torsion angles, and ring conformations derived from the CSD.


Mogul enables you to rapidly validate the complete geometry of a given query structure and identify any unusual features without the need to construct complex search queries, or carry out detailed data analyses.

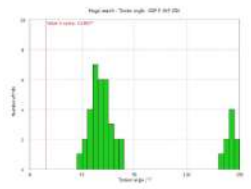
Mogul is becoming an essential tool enabling crystallographers to generate ligand restraint dictionaries. Interfaces to Mogul are available in CCP4's Coot, Grade from Global Phasing, and eLBOW from Phenix. Mogul is also used routinely to check the conformations of both computationally and experimentally generated models, and will form an integral part of the PDB's ligand validation pipeline for newly deposited structures. Medicinal Chemists rely on Mogul not only to validate ligand geometries, but also to aid design decisions.

Mogul provides:

- The ability to load query molecules in a variety of formats (MOL2, CIF, PDB, HES), or sketch a 2D structure.
- Easy selection of geometric features of interest and the ability to perform a "one click" validation of all bonds, angles, torsions and ring conformations in the query molecule.
- Automatic generation of search substructures. The chemical environment around the feature of interest is encoded ensuring that search results are highly relevant to the query structure. Mogul also has the ability to find structurally related fragments if insufficient hits are found.
- Rapid retrieval of geometric data. Results are displayed as a histogram calculated from matching entries in the CSD. Full descriptive statistics and figures of merit are provided allowing easy identification of unusual geometries.
- A simple text-based instructions file interface that allows Mogul to be run in batch mode from the command line and enables easy integration with client applications.

-  View documents and case studies
-  Read the CSD data sheet
-  Request a demo or evaluation
-  View the Mogul user guide, tutorials and technical documentation
-  View the supported platforms
-  Visit the Mogul Support Page









Identification of strained ligand geometry in an aminotransferase structure (ajs). The Mogul torsional distribution for the C-D-P-C dihedral in similar chemical environments is shown. The query torsion value is shown in red.

Home Deposit Structures Access Structures CCDC Policies The CCDC

Terms of Use Privacy Policy

Follow CCDC:    

Got any questions? I'm happy to help.



# Grade Web Server

GΦL Grade Web Server

<http://grade.globalphasing.org/cgi-bin/grade/server.cgi>

## Introduction

Grade generates geometrical restraints for novel small molecules. It does so by querying the Cambridge Structural Database (CSD) as a main source of restraint information, and then invoking quantum chemical procedures to obtain whatever information the CSD is unable to supply. Grade takes a SMILES string or a mol2 file containing 3D coordinates of all atoms as input. For compounds that are already in the [PDB chemical components dictionary](#), Grade can produce restraints given the PDB 3-letter code. For novel large and complex systems the preparation and upload of a mol2 file is recommended. Please see [preferred input formats help page](#) for further information. Please note that there are limitations regarding which [element types can be handled by Grade](#).

## News:

- Updated to latest Grade release (16 Jul 2021) and CSD Data Update May21. 16 July 2021
- Upgrade to Grade (29 Nov 2019, including multiple improvements and bug fixes), added electron-cloud option and updated CSD database. 11 December 2019'  
Please note that the option to generate dictionaries without the use of the CSD has been removed.
- Maintenance: Latest versions of Grade (27 Nov 2018) installed and updated to CSD 2019. 14 December 2018
- Maintenance: updated to CSD 2018. 3 March 2018
- Maintenance: Latest versions of Grade (20 Sep 2017) installed and CSD database updated. When giving a PDB compound identifier for which no ideal coordinates are available, we now fall back to the experimental coordinates to avoid a failure for those compounds. 22 September 2017
- Maintenance: Latest versions of Grade (7th Dec 2016) and Mogul (CSD 2017) installed. 17 February 2017
- Maintenance: Latest versions of Grade and Mogul installed: [details](#). 29 March 2015
- [Bug fixes](#). 20 February 2014
- Major upgrade to Grade Web Server with [new features and improvements](#). 20 December 2013

For further details please see [Grade Web Server Help Page](#) on the BUSTER wiki.

## Run Grade

Name  email:

Use of the Grade Web Server is subject to you agreeing to the [conditions of use](#)

Tick box to confirm that you have read and agree to the conditions of use.

## Acknowledgements

Grade is developed at [Global Phasing Ltd.](#)

Big thanks to:

- [CCDC](#) for permission to use [Mogul](#) on the Grade Web Server
- [Garib Murshudov](#) and [CCP4](#) for [libcheck](#) currently used to interpret SMILES strings
- [Martin Field](#) for the [fdynamo](#) implementation of the [RMI](#) semiempirical QM method
- [PyMOL](#) and [Open Babel](#)
- Members of the Global Phasing Consortium and European Commission project [SILVER](#) (FP7-HEALTH-F3-2010-260644) for financial support
- BUSTER and grade users for suggestions, discussion, feedback and bug reports. Particular thanks to all the Grade web Server pre-release testers for their help!

## Current usage statistics (updated every 10 min):

YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TOTAL
2012		7	70	79	105	81	84	43	33	139	73	51	765
2013	91	54	70	53	63	38	69	31	39	61	85	58	712
2014	49	154	48	63	74	83	133	146	91	147	59	92	1139
2015	79	121	73	142	144	122	132	120	95	111	78	83	1300
2016	163	165	166	129	229	118	232	170	139	157	107	134	1849
2017	122	169	168	126	166	281	292	169	159	213	165	128	2078
2018	169	205	248	214	227	323	492	179	185	231	203	208	2884
2019	159	155	349	226	272	150	253	104	162	225	163	182	2400
2020	165	235	215	193	130	148	221	131	195	202	281	200	2316
2021	198	261	219	203	171	210	201	123	151	44			1781

For help and further details see [Grade Web Server Help Page](#) on the BUSTER wiki.

Please email [buster-develop@globalphasing.com](mailto:buster-develop@globalphasing.com) with any problems, queries or suggestions.

# Grade Web Server

**GΦL** Grade Web Server <http://grade.globalphasing.org>

---

**Introduction**

Grade generates geometric... obtain whatever informati... [dictionary](#). Grade can prod... information. Please note th...

**News:**

- Updated to latest Gr...
- Upgrade to Grade (...
- Please note that the...
- Maintenance: Latest...
- Maintenance: updat...
- Maintenance: Latest...
- coordinates to avoid...
- Maintenance: Latest...
- Maintenance: Latest...
- [Bug fixes](#). 20 Febru...
- Major upgrade to G...

For further details please s...

**Run Grade**

Name  email:

Use of the Grade Web Server is subject to you agreeing to the [conditions of use](#)

Tick box to confirm that you have read and agree to the conditions of use.

**Run Grade**

Name

Use of the Grade Web Ser...

Tick box to confirm that you have read and agree to the conditions of use.

---

**Acknowledgements**

Grade is developed at [Global Phasing Ltd.](#)

Big thanks to:

- [CCDC](#) for permission to use [Mogul](#) on the Grade Web Server
- [Garib Murshudov](#) and [CCP4](#) for [libcheck](#) currently used to interpret SMILES strings
- [Martin Field](#) for the [fdynamo](#) implementation of the [RMI](#) semiempirical QM method
- [PyMOL](#) and [Open Babel](#)
- Members of the Global Phasing Consortium and European Commission project [SILVER](#) (FP7-HEALTH-F3-2010-260644) for financial support
- [BUSTER](#) and grade users for suggestions, discussion, feedback and bug reports. Particular thanks to all the Grade web Server pre-release testers for their help!

---

**Current usage statistics (updated every 10 min):**

YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TOTAL
2012		7	70	79	105	81	84	43	33	139	73	51	765
2013	91	54	70	53	63	38	69	31	39	61	85	58	712
2014	49	154	48	63	74	83	133	146	91	147	59	92	1139
2015	79	121	73	142	144	122	132	120	95	111	78	83	1300
2016	163	165	166	129	229	118	232	170	139	157	107	134	1849
2017	122	169	168	126	166	281	292	169	159	213	165	128	2078
2018	169	205	248	214	227	323	492	179	185	231	203	208	2684
2019	159	155	349	226	272	150	253	104	162	225	163	182	2400
2020	165	235	215	193	130	148	221	131	195	202	281	200	2316
2021	198	261	219	203	171	210	201	123	151	44			1781

For help and further details see [Grade Web Server Help Page](#) on the [BUSTER](#) wiki.

Please email [buster-develop@globalphasing.com](mailto:buster-develop@globalphasing.com) with any problems, queries or suggestions.

version Release v1.107 Jul 16 2021

# Grade Web Server

SMILES input

## GΦL Grade Web Server <http://grade.globalphasing.org>

Atsushi Nakagawa ([atsushi@protein.osaka-u.ac.jp](mailto:atsushi@protein.osaka-u.ac.jp)) thank you for agreeing to the [conditions of use](#)

Please select an input type:

- SMILES string (for example from the Structure Editor at the [CACTVS Online SMILES Translator](#))
- Run grade on mol2 file (must include all [hydrogen atoms](#). This option is useful for getting correct atom names [further details](#))
- Produce dictionary for an existing PDB chemical component (for example "ATP") [further details](#)

(enter SMILES string)

### Options

Provide electron-cloud bond distances for hydrogens. (instead of using the default nuclear position)

Fuse planes together rather than having 4 atom planes. (Generally this is a bad idea!)

Output [SHELXL](#) type restraint file  .

three-letter code to be used for the ligand. (Defaults to "XXX")

name for the ligand (for example "PF31337" or "BENZONEOMYCIN")

Net overall charge. (Defaults - SMILES: work charge out from the SMILES string, mol2 file: use zero charge)

# Grade Web Server

MOL2 input

## GΦL Grade Web Server <http://grade.globalphasing.org>

Atsushi Nakagawa ([atsushi@protein.osaka-u.ac.jp](mailto:atsushi@protein.osaka-u.ac.jp)) thank you for agreeing to the [conditions of use](#)

Please select an input type:

- SMILES string (for example from the Structure Editor at the [CACTVS Online SMILES Translator](#))
- Run grade on mol2 file (must include all [hydrogen atoms](#). This option is useful for getting correct atom names [further details](#))
- Produce dictionary for an existing PDB chemical component (for example "ATP") [further details](#)

mol2  file upload

### Options

- Provide electron-cloud bond distances for hydrogens. (instead of using the default nuclear position)
- Fuse planes together rather than having 4 atom planes. (Generally this is a bad idea!)
- Output [SHELXL](#) type restraint file  .
- three-letter code to be used for the ligand. (Defaults to "XXX")
- name for the ligand (for example "PF31337" or "BENZONEOMYCIN")
- Net overall charge. (Defaults - SMILES: work charge out from the SMILES string, mol2 file: use zero charge)
- Molecule described in mol2 file really does not have any hydrogen atoms. (This sets the grade `-really_noH` option)

# Grade Web Server

PDB input

## GΦL Grade Web Server <http://grade.globalphasing.org>

Atsushi Nakagawa ([atsushi@protein.osaka-u.ac.jp](mailto:atsushi@protein.osaka-u.ac.jp)) thank you for agreeing to the [conditions of use](#)

Please select an input type:

- SMILES string (for example from the Structure Editor at the [CACTVS Online SMILES Translator](#))
- Run grade on mol2 file (must include all [hydrogen atoms](#). This option is useful for getting correct atom names [further details](#))
- Produce dictionary for an existing PDB chemical component (for example "ATP") [further details](#)

(enter 3 letter code - see [RCSB ligand expo](#))

### Options

- Provide electron-cloud bond distances for hydrogens. (instead of using the default nuclear position)
- Fuse planes together rather than having 4 atom planes. (Generally this is a bad idea!)
- Output [SHELXL](#) type restraint file   .

# Grade2

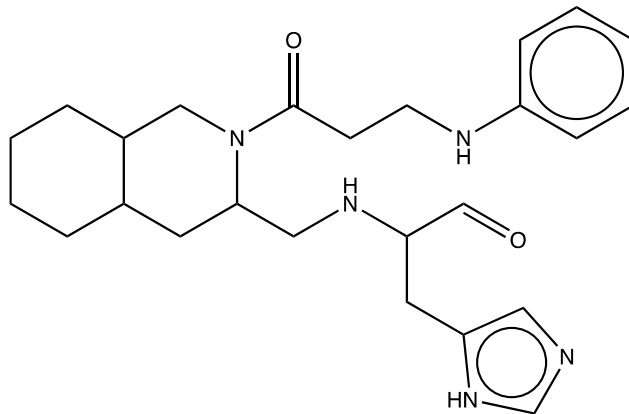
- The initial release of the new restraint generation tool, written by Oliver Smart.
- Grade2 is a clean-slate reimplementation using ideas from Grade, using the CSD Python API and the RDKit toolkit.
- Grade2 already has several advantages compared to Grade:
  - Grade2 makes better use of CSD information for rings and planes, using the CSD Python API to extend Mogul analysis.
  - Grade2 can handle a wider range of elements.
  - Grade2 is more robust. In a recent test of generating restraints-dictionaries for 2893 chemical components recently added to wwPDB, Grade2 produced dictionaries for all but 6 (0.2% failure) whereas Grade failed in 336 cases (12%).
  - Grade2 protonates or deprotonates groups commonly charged at neutral pH such as carboxylic acids, phosphates and alkylamines.

# Grade2 Input

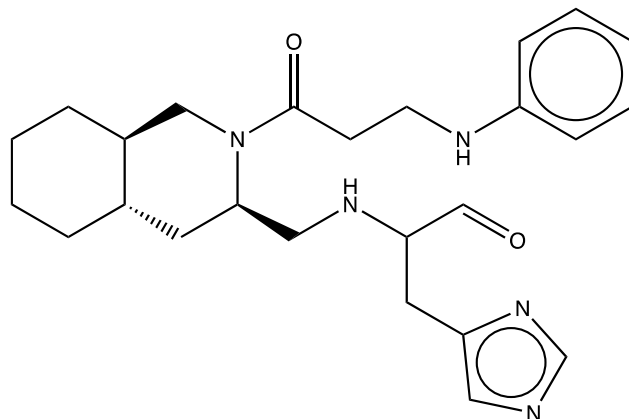
- PDB 3-letter code  
grade2 -P <pdb id>
- SMILES  
grade2 'Oc1ccccc1' --resname LIG
- mol/sdf  
grade2 --in Conformer3D\_CID\_135398744.sdf
- Tripos mol2
- Restraint dictionary CIF
- wwPDB CCD CIF
  - wwPDB Chemical Component Dictionary

# SLH phenyl-beta-alanyl (R,S)-N-decalin type inhibitor

(2S)-3-(1H-imidazol-5-yl)-2-({[(3R,4aS,8aR)-2-(N-phenyl-beta-alanyl)decahydroisoquinolin-3-yl]methyl}amino)propanal



SMILES: 3C1C(CCCC1)CC(CNC(C=O)Cc2ncnc2)N3C(=O)CCNc4ccccc4



SMILES: O=C[C@H](Cc1[nH]cnc1)NC[C@H]2C[C@@H]3CCCC[C@H]3CN2C(=O)CCNc4ccccc4

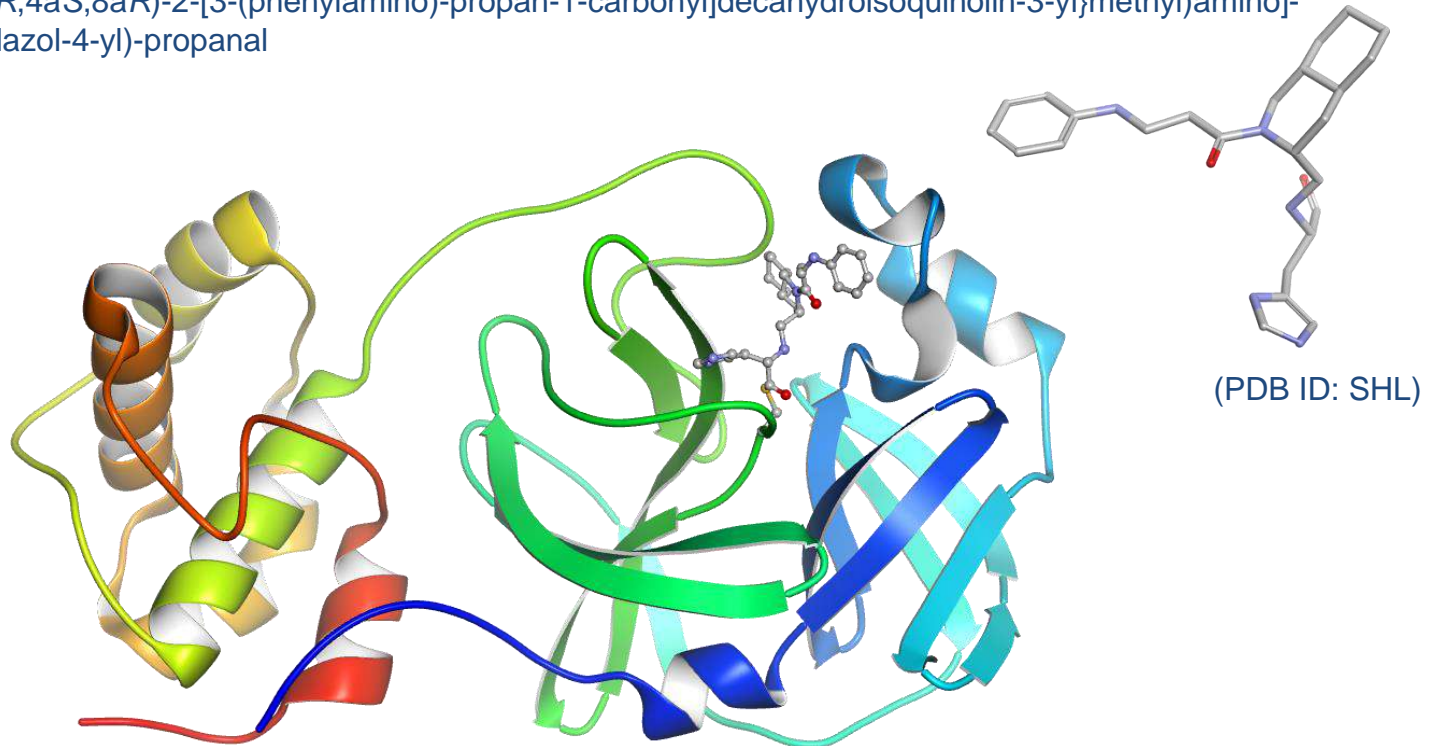


# Example 1:

## SARS-3CL protease complex with a phenyl-beta-alanyl (*R*)-N-decalin type inhibitor

(PDB ID: 5c5n)

(*S*)-2-[[[(3*R*,4*aS*,8*aR*)-2-[3-(phenylamino)-propan-1-carbonyl]decahydroisoquinolin-3-yl)methyl]amino]-3-(1*H*-imidazol-4-yl)-propanal



Teruya, K. *et al.* Structural basis for the development of SARS 3CL protease inhibitors from a peptide mimic to an aza-decaline scaffold. *Biopolymers* **106**, 391-403, doi:10.1002/bip.22773 (2016).

# Ligand Library by grade2

```

# GEN: Generated by grade2 version 1.0.0
# GEN: from PDB_chemical_components_definition_file ftp://ftp.ebi.ac.uk/pub/databases/msd/pdbechem_v2/S/SLH/SLH.cif
# GEN: using RDKit 2021.03.4
# GEN: ((DatabaseID PDB-SLH))
#
# BUSTER-KEYWORD TRUSTCOORD
# BUSTER-KEYWORD TRUSTTORS
#
data_comp_list
#
_chem_comp.id                SLH
_chem_comp.three_letter_code SLH
_chem_comp.name              "(2S)-3-(1H-imidazol-5-yl)-2-({[(3R,4aS,8aR)-2-(N-phenyl-beta-
alanyl)decahydroisoquinolin-3-yl]methyl}amino)propanal (CHARGED)"
_chem_comp.group             .
_chem_comp.number_atoms_all  68
_chem_comp.number_atoms_nh   32
_chem_comp.desc_level        .

#
data_comp_SLH
#
loop_
_chem_comp_atom.comp_id
_chem_comp_atom.atom_id
_chem_comp_atom.type_symbol
_chem_comp_atom.type_energy
_chem_comp_atom.partial_charge
_chem_comp_atom.charge
_chem_comp_atom.x
_chem_comp_atom.y
_chem_comp_atom.z
SLH  C1  C  CH2  0.024  0 -0.654 -1.761 -0.422
SLH  C2  C  CH1 -0.021  0 -1.592 -2.611  0.430
SLH  C3  C  CH2 -0.048  0 -1.998 -3.896 -0.292
SLH  N4  N  NH1 -0.385  0  4.527 -1.137 -0.603
SLH  C5  C  CH2 -0.053  0 -4.230 -3.881  0.870
SLH  C6  C  CH2 -0.050  0 -3.831 -2.604  1.604
SLH  C9  C  CH1  0.081  0 -1.375  0.300  0.791

```

# Example 1: 5c5n SARS-3CL

complexed with a phenyl-beta-alanyl (*R*)-N-decalin type inhibitor  
(1.69Å)

	CCP4	grade2
R	0.279517	0.279434
Rfree	0.304315	0.304152
Rxpct	0.254654	0.254535
Rxpctfree	0.282053	0.272112
BOND (Å)	0.0082	0.0082
ANGL (deg.)	1.03	1.02
TORS	15.63	15.46
SINTOR	15.63	15.46
PLAN (Å)	0.0180	0.0150
BCOR (Å <sup>2</sup> )	0.879	0.879

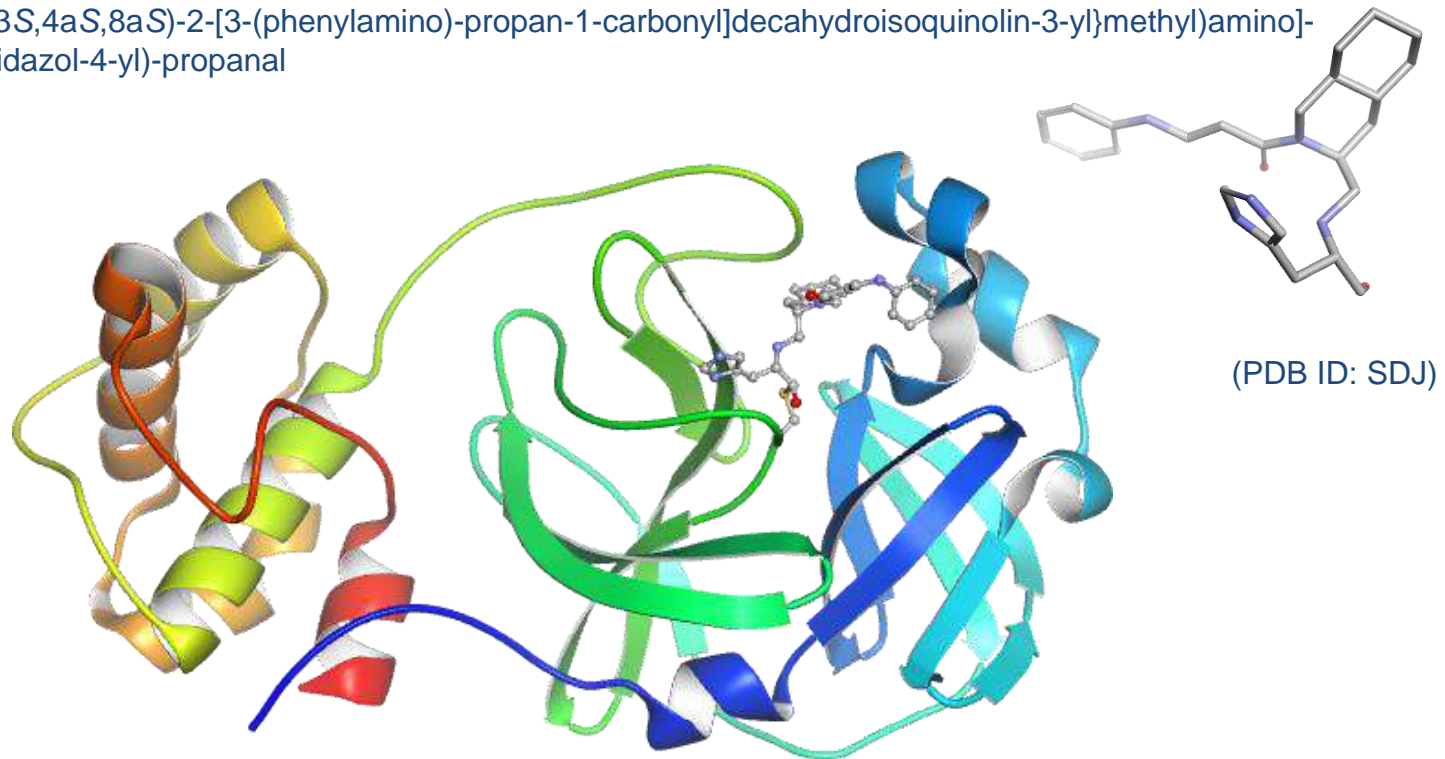
Bricogne G., Blanc E., Brandl M., Flensburg C., Keller P., Paciorek W.,  
Roversi P, Sharff A., Smart O.S., Vornrhein C., Womack T.O. (2017).  
BUSTER version 2.10.4. Cambridge, United Kingdom: Global Phasing Ltd.

# Example 2:

## SARS-3CL protease complex with a phenyl-beta-alanyl (S)-N-decalin type inhibitor

(PDB ID: 5c5o)

(S)-2-[[[(3S,4aS,8aS)-2-[3-(phenylamino)-propan-1-carbonyl]decahydroisoquinolin-3-yl)methyl]amino]-3-(1H-imidazol-4-yl)-propanal



Teruya, K. *et al.* Structural basis for the development of SARS 3CL protease inhibitors from a peptide mimic to an aza-decaline scaffold. *Biopolymers* **106**, 391-403, doi:10.1002/bip.22773 (2016).

# Example 2: 5c5o SARS-3CL

complexed with a phenyl-beta-alanyl (S)-N-decalin type inhibitor  
(1.50Å)

	CCP4	grade2
R	0.232363	0.232599
Rfree	0.256229	0.255772
Rxpct	0.223368	0.223642
Rxpctfree	0.244443	0.244695
BOND (Å)	0.082	0.082
ANGL (deg.)	0.98	0.97
TORS	15.07	14.97
SINTOR	15.07	14.97
PLAN (Å)	0.0201	0.0168
BCOR (Å <sup>2</sup> )	0.964	0.964

Bricogne G., Blanc E., Brandl M., Flensburg C., Keller P., Paciorek W.,  
Roversi P, Sharff A., Smart O.S., Vornrhein C., Womack T.O. (2017).  
BUSTER version 2.10.4. Cambridge, United Kingdom: Global Phasing Ltd.

## Please use CSD data to prepare Ligand Library

- ACEDRG ... CCP4
  - COD: Crystallography Open Database
  - <http://www.crystallography.net/cod/>
- eLBOW ... phenix
  - Mogul
- Grade, Grade2 ... Global Phasing Inc.
  - Mogul
- Pyrogen ... Coot (CCP4)
  - Mogul

# Acknowledgements

- Thanks to all PDBj members and wwPDB consortium

