



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 29, 2022 – 04:07 PM JST

PDB ID : 7XB5
Title : Structure of the ligand-binding domain of *S. cerevisiae* Upc2 in fusion with T4 lysozyme
Authors : Tan, L.; Im, Y.J.
Deposited on : 2022-03-20
Resolution : 3.44 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at 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
Xtriage (Phenix)	:	1.13
EDS	:	2.30
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.30

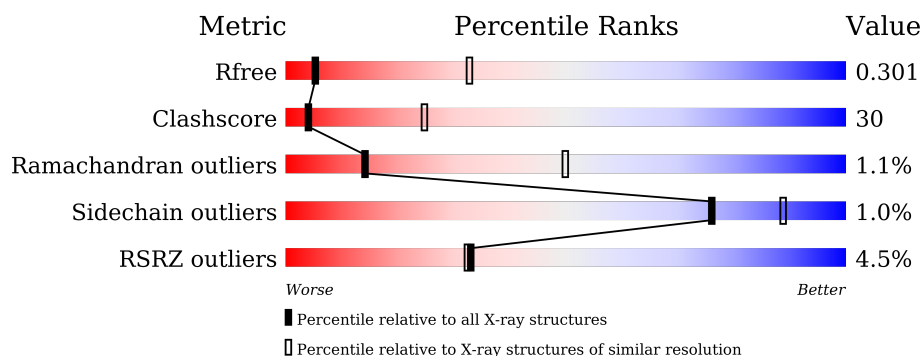
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.44 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1278 (3.50-3.38)
Clashscore	141614	1361 (3.50-3.38)
Ramachandran outliers	138981	1327 (3.50-3.38)
Sidechain outliers	138945	1328 (3.50-3.38)
RSRZ outliers	127900	1192 (3.50-3.38)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	465	<div> <div>3%</div> <div>37%</div> <div>37%</div> <div>•</div> <div>23%</div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2859 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called fusion protein of Sterol uptake control protein 2 and Endolysin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	358	Total	C	N	O	S	0	0	0
			2859	1840	485	521	13			

There are 35 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	581	MET	-	initiating methionine	UNP Q12151
A	582	GLY	-	expression tag	UNP Q12151
A	583	SER	-	expression tag	UNP Q12151
A	584	SER	-	expression tag	UNP Q12151
A	585	HIS	-	expression tag	UNP Q12151
A	586	HIS	-	expression tag	UNP Q12151
A	587	HIS	-	expression tag	UNP Q12151
A	588	HIS	-	expression tag	UNP Q12151
A	589	HIS	-	expression tag	UNP Q12151
A	590	HIS	-	expression tag	UNP Q12151
A	591	SER	-	expression tag	UNP Q12151
A	592	SER	-	expression tag	UNP Q12151
A	593	GLY	-	expression tag	UNP Q12151
A	594	LEU	-	expression tag	UNP Q12151
A	595	VAL	-	expression tag	UNP Q12151
A	596	PRO	-	expression tag	UNP Q12151
A	597	ARG	-	expression tag	UNP Q12151
A	598	GLY	-	expression tag	UNP Q12151
A	599	SER	-	expression tag	UNP Q12151
A	600	HIS	-	expression tag	UNP Q12151
A	601	MET	-	expression tag	UNP Q12151
A	715	VAL	-	linker	UNP Q12151
A	716	ASP	-	linker	UNP Q12151
A	727	GLY	ARG	engineered mutation	UNP P00720
A	769	THR	CYS	engineered mutation	UNP P00720
A	812	ALA	CYS	engineered mutation	UNP P00720
A	852	ARG	ILE	engineered mutation	UNP P00720

Continued on next page...

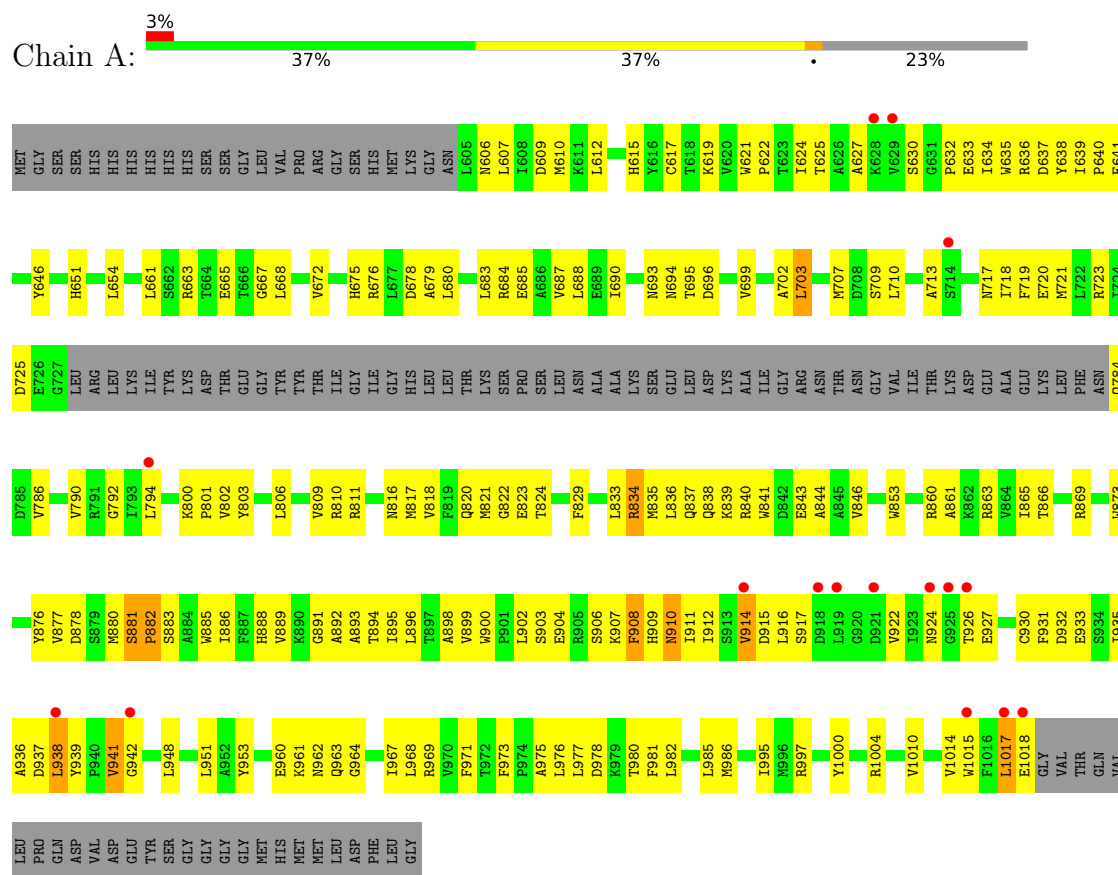
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	877	VAL	-	linker	UNP P00720
A	878	ASP	-	linker	UNP P00720
A	?	-	PRO	deletion	UNP Q12151
A	?	-	ASP	deletion	UNP Q12151
A	?	-	VAL	deletion	UNP Q12151
A	?	-	GLY	deletion	UNP Q12151
A	?	-	THR	deletion	UNP Q12151
A	925	GLY	ILE	engineered mutation	UNP Q12151

3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: fusion protein of Sterol uptake control protein 2 and Endolysin



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	65.74Å 125.08Å 155.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.98 – 3.44 32.98 – 3.44	Depositor EDS
% Data completeness (in resolution range)	96.0 (32.98-3.44) 96.1 (32.98-3.44)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.60 (at 3.47Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.258 , 0.301 0.258 , 0.301	Depositor DCC
R_{free} test set	398 reflections (4.68%)	wwPDB-VP
Wilson B-factor (Å ²)	105.9	Xtriage
Anisotropy	0.501	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 84.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	2859	wwPDB-VP
Average B, all atoms (Å ²)	120.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.29% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.58	1/2921 (0.0%)	0.88	11/3965 (0.3%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	617	CYS	CB-SG	-5.07	1.73	1.81

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	914	VAL	N-CA-C	-8.85	87.10	111.00
1	A	916	LEU	C-N-CA	7.80	141.20	121.70
1	A	908	PHE	N-CA-C	6.85	129.49	111.00
1	A	976	LEU	CB-CG-CD2	-6.76	99.50	111.00
1	A	985	LEU	CA-CB-CG	6.63	130.54	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2859	0	2872	172	0
All	All	2859	0	2872	172	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including

hydrogen atoms). The all-atom clashscore for this structure is 30.

The worst 5 of 172 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:882:PRO:HB2	1:A:1000:TYR:OH	1.20	1.28
1:A:930:CYS:HA	1:A:939:TYR:HE2	1.06	1.12
1:A:878:ASP:CG	1:A:997:ARG:HH12	1.63	1.01
1:A:930:CYS:HA	1:A:939:TYR:CE2	1.98	0.99
1:A:720:GLU:CD	1:A:881:SER:HB2	1.84	0.98

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	354/465 (76%)	299 (84%)	51 (14%)	4 (1%)	14	50

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	882	PRO
1	A	917	SER
1	A	931	PHE
1	A	667	GLY

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	307/395 (78%)	304 (99%)	3 (1%)	76 89

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	834	ARG
1	A	881	SER
1	A	910	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	615	HIS
1	A	717	ASN
1	A	847	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å ²)	Q<0.9
1	A	358/465 (76%)	0.11	16 (4%)	33 32	58, 113, 203, 341	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	942	GLY	5.1
1	A	925	GLY	4.5
1	A	1017	LEU	4.5
1	A	628	LYS	4.2
1	A	921	ASP	3.7

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.