WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Jan 20, 2024 – 03:47 pm GMT

PDB ID	:	7ASI
Title	:	Fixed-target serial femtosecond crystallography using in cellulo grown Neurospora crassa HEX-1 microcrystals. (Chips 1+2)
Authors	:	Lahey-Rudolph, J.M.; Schoenherr, R.; Barthelmess, M.; Fischer, P.; Seuring, C.; Wagner, A.; Meents, A.; Redecke, L.
Deposited on	:	2020-10-27
Resolution	:	1.70 Å(reported)

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

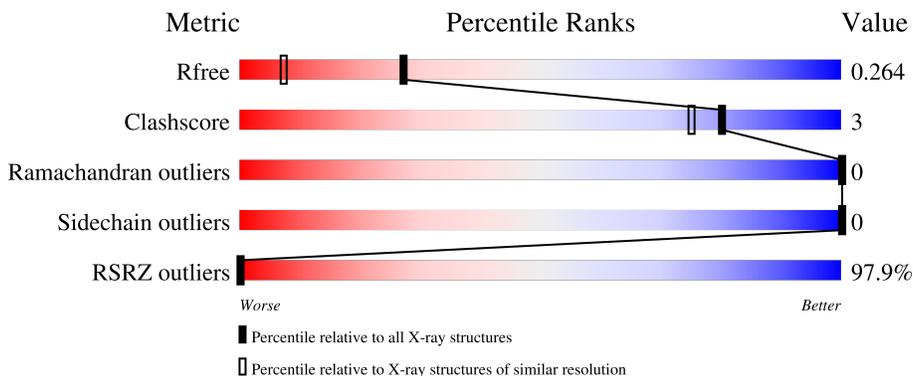|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.70 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	Percentile Ranks	Value
Rfree		0.264
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		97.9%

Worse   *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 <=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	177	78% / 74% 5% · 20%

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2303 atoms, of which 1093 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 eIF-5a domain-containing protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	141	2209	699	1093	201	209	7	14	3	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	177	ALA	-	cloning artifact	UNP A0A0B0EDT5

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	94	94	94	0	0

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: eIF-5a domain-containing protein

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	58.73Å 58.73Å 192.83Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.36 – 1.70 29.37 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.8 (29.36-1.70) 99.8 (29.37-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.15 (at 1.71Å)	Xtriage
Refinement program	PHENIX v1.16-3546, PHENIX v1.9-1692	Depositor
R, R_{free}	0.219 , 0.264 0.220 , 0.264	Depositor DCC
R_{free} test set	1787 reflections (7.95%)	wwPDB-VP
Wilson B-factor (Å^2)	18.3	Xtriage
Anisotropy	0.055	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 55.5	EDS
L-test for twinning[2]	$< |L| > = 0.34, < L^2 > = 0.18$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	2303	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.98% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.55	0/1134	0.72	2/1537 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	41	ARG	NE-CZ-NH2	-9.61	115.50	120.30
1	A	41	ARG	NE-CZ-NH1	7.08	123.84	120.30

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	1116	1093	1132	6	1
2	A	94	0	0	3	1
All	All	1210	1093	1132	6	2

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 3.

All (6) 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:163:GLU:OE1	2:A:201:HOH:O	2.14	0.66
1:A:124:ASP:OD2	2:A:202:HOH:O	2.16	0.62
1:A:51:ARG:NH2	1:A:159:ASP:OD2	2.45	0.50
1:A:33:VAL:HG13	1:A:102:VAL:HG22	2.00	0.43
1:A:102:VAL:HG13	2:A:289:HOH:O	2.18	0.43
1:A:145:PHE:O	1:A:149:ARG:HG2	2.18	0.43

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:41:ARG:HH22	1:A:44:ASP:OD1[11_655]	1.53	0.07
2:A:241:HOH:O	2:A:249:HOH:O[1_655]	2.19	0.01

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	142/177 (80%)	138 (97%)	4 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	128/148 (86%)	128 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	141/177 (79%)	4.50	138 (97%)	0	0	7, 20, 43, 54	1 (0%)

All (138) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	160	HIS	9.6
1	A	113	ASP	9.0
1	A	170	VAL	7.9
1	A	171	VAL	7.8
1	A	138	TRP	7.7
1	A	147	SER	7.2
1	A	114	GLY	7.2
1	A	89	PRO	6.9
1	A	53	CYS	6.6
1	A	33	VAL	6.6
1	A	91	PRO	6.6
1	A	161	GLY	6.6
1	A	90	ALA	6.5
1	A	156	VAL	6.4
1	A	146[A]	GLU	6.3
1	A	37	CYS	6.2
1	A	159	ASP	6.2
1	A	165	ALA	6.1
1	A	145	PHE	6.1
1	A	85	PHE	6.1
1	A	72	VAL	6.0
1	A	35[A]	ILE	6.0
1	A	106	TYR	5.9
1	A	112	GLN	5.8
1	A	108	VAL	5.8
1	A	152	VAL	5.7
1	A	70	LEU	5.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	143	LYS	5.6
1	A	55	VAL	5.6
1	A	103	PHE	5.6
1	A	118	ALA	5.6
1	A	137	LEU	5.5
1	A	47	ILE	5.5
1	A	144	ALA	5.5
1	A	95	VAL	5.5
1	A	158[A]	SER	5.5
1	A	157	VAL	5.4
1	A	93	VAL	5.3
1	A	117	VAL	5.3
1	A	102	VAL	5.3
1	A	154	VAL	5.3
1	A	56	ILE	5.2
1	A	48	LEU	5.2
1	A	120	THR	5.1
1	A	52	PRO	5.1
1	A	139	ASN	5.1
1	A	142	GLN	5.1
1	A	79	LEU	5.0
1	A	149	ARG	5.0
1	A	130	PRO	5.0
1	A	132	ILE	5.0
1	A	99	LEU	5.0
1	A	69	TYR	5.0
1	A	75	PHE	4.9
1	A	115	SER	4.9
1	A	162	ARG	4.9
1	A	46	LEU	4.9
1	A	141	LEU	4.9
1	A	45	ILE	4.8
1	A	36	PRO	4.8
1	A	136	SER	4.8
1	A	101	PRO	4.7
1	A	64	THR	4.7
1	A	166	VAL	4.7
1	A	58	ILE	4.7
1	A	129	LEU	4.6
1	A	116	ILE	4.6
1	A	148	GLY	4.6
1	A	88	ASN	4.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	31	GLN	4.6
1	A	151	SER	4.6
1	A	74	LEU	4.5
1	A	122	THR	4.5
1	A	131	VAL	4.4
1	A	42	LEU	4.3
1	A	155	LEU	4.3
1	A	125	VAL	4.2
1	A	135	SER	4.2
1	A	84	SER	4.1
1	A	76	THR	4.1
1	A	168	MET	4.1
1	A	94	VAL	4.1
1	A	62	ALA	4.1
1	A	40	ILE	4.1
1	A	164	MET	4.0
1	A	65	GLY	4.0
1	A	109	LEU	4.0
1	A	71	GLY	4.0
1	A	92	SER	3.9
1	A	63	ALA	3.9
1	A	128	ASN	3.9
1	A	87	SER	3.9
1	A	98	MET	3.9
1	A	111	MET	3.9
1	A	86	VAL	3.9
1	A	59	SER	3.9
1	A	60	THR	3.7
1	A	32	THR	3.7
1	A	97	THR	3.7
1	A	140	ARG	3.7
1	A	78	GLN	3.7
1	A	107	ARG	3.6
1	A	34	THR	3.5
1	A	67	HIS	3.5
1	A	100	GLY	3.4
1	A	54	GLN	3.4
1	A	61	SER	3.4
1	A	51	ARG	3.3
1	A	57	ARG	3.3
1	A	83	SER	3.3
1	A	119	MET	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	150	GLY	3.2
1	A	66	GLN	3.2
1	A	44	ASP	3.2
1	A	39	HIS	3.0
1	A	73	ASP	3.0
1	A	167	ASP	2.9
1	A	121	GLU	2.9
1	A	41	ARG	2.9
1	A	133	ASP	2.9
1	A	68	ARG	2.8
1	A	169	LYS	2.8
1	A	80	HIS	2.8
1	A	124	ASP	2.8
1	A	110	ASP	2.7
1	A	38	HIS	2.7
1	A	134	GLN	2.7
1	A	123	GLY	2.7
1	A	96	GLN	2.7
1	A	153	ARG	2.5
1	A	126	LYS	2.5
1	A	127	GLN	2.4
1	A	43	GLY	2.4
1	A	81	GLU	2.3
1	A	104	LYS	2.3
1	A	49	GLN	2.3
1	A	105	GLN	2.2
1	A	82	GLU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.