



Full wwPDB X-ray Structure Validation Report ⓘ

May 17, 2020 – 01:41 pm BST

PDB ID : 5F13
Title : Structure of Mn bound DUF89 from *Saccharomyces cerevisiae*
Authors : Nocek, B.; Skarina, T.; Joachimiak, A.; Savchenko, A.; Yakunin, A.
Deposited on : 2015-11-30
Resolution : 2.39 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 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 as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

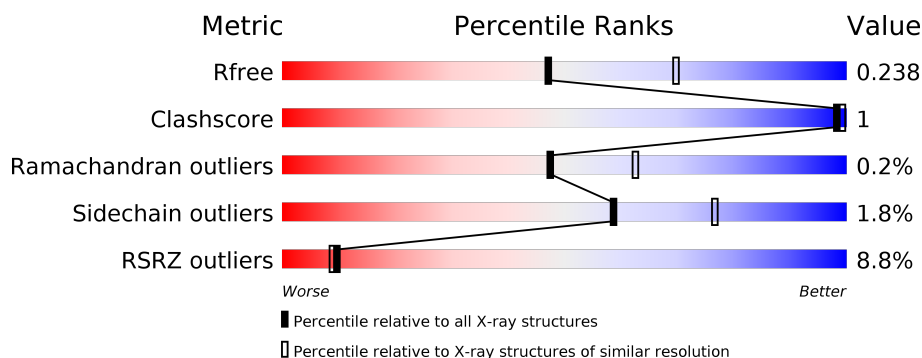
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.39 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 on 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	471	<div> <div>4%</div> <div> <div></div> <div>93%</div> <div>6%</div> </div> </div>
1	B	471	<div> <div>7%</div> <div> <div></div> <div>92%</div> <div>5%</div> </div> </div>
1	C	471	<div> <div>14%</div> <div> <div></div> <div>90%</div> <div>6%</div> </div> </div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 20964 atoms, of which 10167 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 Protein-glutamate O-methyltransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	445	Total	C	H	N	O	S	0	0	0
			6948	2277	3414	587	657	13			
1	B	449	Total	C	H	N	O	S	0	0	0
			6905	2279	3363	584	666	13			
1	C	444	Total	C	H	N	O	S	0	0	0
			6908	2272	3384	585	654	13			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	HIS	-	expression tag	UNP Q04371
B	0	HIS	-	expression tag	UNP Q04371
C	0	HIS	-	expression tag	UNP Q04371

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mn	0	0
			1	1		
2	A	1	Total	Mn	0	0
			1	1		
2	C	1	Total	Mn	0	0
			1	1		

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	P	0	0
			5	4	1		
3	A	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		
3	C	1	Total	O	P	0	0
			5	4	1		
3	C	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Cl	0	0
			1	1		
4	A	2	Total	Cl	0	0
			2	2		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	H	O	0	0
			10	2	6	2		

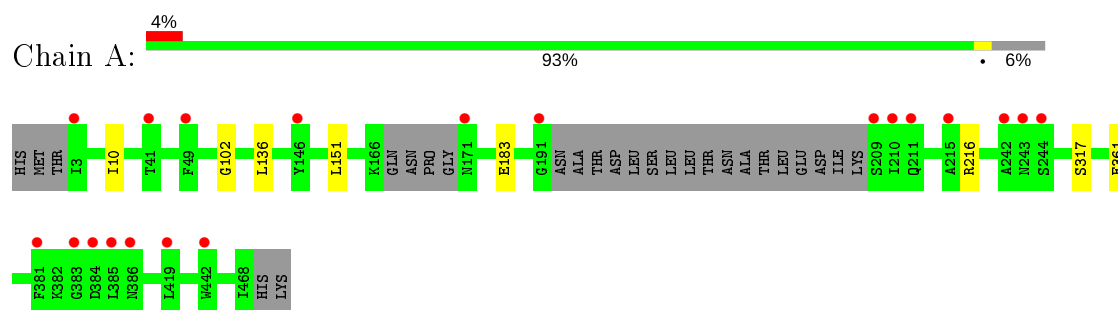
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	89	Total	O	0	1
			90	90		
6	B	45	Total	O	0	0
			45	45		
6	C	22	Total	O	0	0
			22	22		

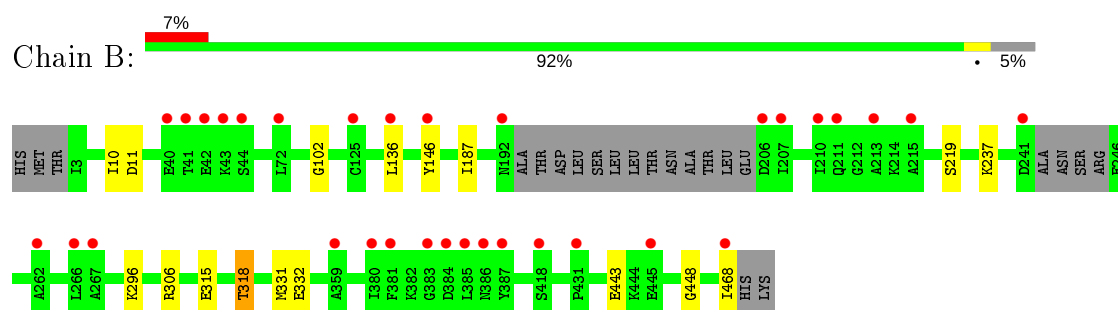
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

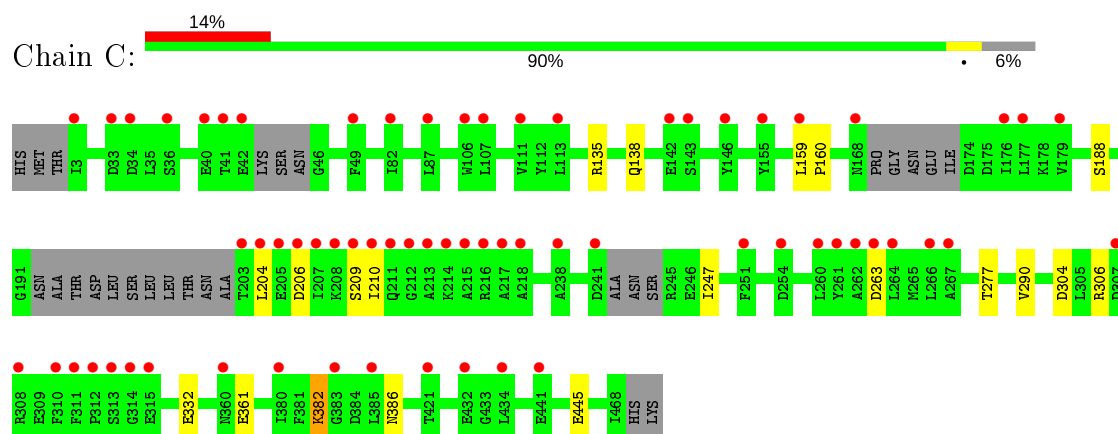
• Molecule 1: Protein-glutamate O-methyltransferase



• Molecule 1: Protein-glutamate O-methyltransferase



• Molecule 1: Protein-glutamate O-methyltransferase



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	223.03Å 95.55Å 78.10Å 90.00° 105.84° 90.00°	Depositor
Resolution (Å)	30.17 – 2.39 30.17 – 2.39	Depositor EDS
% Data completeness (in resolution range)	96.2 (30.17-2.39) 86.2 (30.17-2.39)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.07 (at 2.39Å)	Xtriage
Refinement program	PHENIX dev_1888	Depositor
R, R_{free}	0.199 , 0.237 0.201 , 0.238	Depositor DCC
R_{free} test set	2730 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	34.1	Xtriage
Anisotropy	0.089	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 37.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	20964	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.09% 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

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, MN, EDO, CL

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	0.24	0/3617	0.37	0/4899
1	B	0.22	0/3626	0.36	0/4923
1	C	0.22	0/3605	0.36	0/4885
All	All	0.23	0/10848	0.36	0/14707

There are no bond length outliers.

There are no bond angle outliers.

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	3534	3414	3412	2	0
1	B	3542	3363	3362	4	0
1	C	3524	3384	3381	6	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	10	0	0	0	0
3	B	10	0	0	0	0
3	C	10	0	0	0	0
4	A	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1	0	0	0	0
5	B	4	6	6	0	0
6	A	90	0	0	0	0
6	B	45	0	0	0	0
6	C	22	0	0	0	0
All	All	10797	10167	10161	12	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 1.

All (12) 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:183:GLU:OE2	1:A:216:ARG:NH1	2.30	0.64
1:C:188:SER:OG	1:C:263:ASP:OD1	2.26	0.53
1:B:443:GLU:OE2	1:B:448:GLY:N	2.38	0.53
1:C:306:ARG:NH1	1:C:332:GLU:OE2	2.38	0.51
1:C:382:LYS:O	1:C:386:ASN:ND2	2.42	0.51
1:A:10:ILE:HG23	1:A:102:GLY:O	2.13	0.49
1:B:315:GLU:O	1:B:318:THR:OG1	2.32	0.47
1:B:10:ILE:HG23	1:B:102:GLY:O	2.16	0.45
1:C:247:ILE:N	1:C:277:THR:OG1	2.42	0.43
1:B:306:ARG:NH1	1:B:332:GLU:OE2	2.51	0.42
1:C:159:LEU:HB3	1:C:160:PRO:HD3	2.02	0.40
1:C:138:GLN:NE2	1:C:304:ASP:OD2	2.52	0.40

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	439/471 (93%)	427 (97%)	12 (3%)	0	100	100
1	B	443/471 (94%)	429 (97%)	13 (3%)	1 (0%)	47	62
1	C	434/471 (92%)	420 (97%)	12 (3%)	2 (0%)	29	41
All	All	1316/1413 (93%)	1276 (97%)	37 (3%)	3 (0%)	47	62

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	11	ASP
1	C	209	SER
1	C	290	VAL

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	369/416 (89%)	365 (99%)	4 (1%)	73	87
1	B	366/416 (88%)	357 (98%)	9 (2%)	47	67
1	C	365/416 (88%)	358 (98%)	7 (2%)	57	75
All	All	1100/1248 (88%)	1080 (98%)	20 (2%)	59	76

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	136	LEU
1	A	151	LEU
1	A	317	SER
1	A	361	GLU
1	B	136	LEU
1	B	146	TYR
1	B	187	ILE
1	B	219	SER
1	B	237	LYS
1	B	296	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	318	THR
1	B	331	MET
1	B	468	ILE
1	C	135	ARG
1	C	204	LEU
1	C	206	ASP
1	C	210	ILE
1	C	361	GLU
1	C	382	LYS
1	C	445	GLU

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

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 6 are monoatomic - leaving 7 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	PO4	A	502	2	4,4,4	1.00	0	6,6,6	0.45	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	EDO	B	505	-	3,3,3	0.49	0	2,2,2	0.35	0
3	PO4	B	502	2	4,4,4	0.97	0	6,6,6	0.45	0
3	PO4	A	503	-	4,4,4	0.94	0	6,6,6	0.46	0
3	PO4	B	503	-	4,4,4	0.93	0	6,6,6	0.44	0
3	PO4	C	503	-	4,4,4	0.96	0	6,6,6	0.43	0
3	PO4	C	502	2	4,4,4	0.98	0	6,6,6	0.46	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	EDO	B	505	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	505	EDO	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

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	445/471 (94%)	0.36	20 (4%) 33 31	17, 31, 57, 114	0
1	B	449/471 (95%)	0.56	32 (7%) 16 14	27, 45, 73, 118	0
1	C	444/471 (94%)	0.89	66 (14%) 2 2	27, 52, 96, 129	0
All	All	1338/1413 (94%)	0.60	118 (8%) 10 9	17, 44, 86, 129	0

All (118) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	210	ILE	7.1
1	C	215	ALA	6.0
1	C	176	ILE	5.3
1	C	314	GLY	5.2
1	B	41	THR	5.1
1	C	204	LEU	5.0
1	C	203	THR	4.9
1	C	312	PRO	4.8
1	A	242	ALA	4.6
1	B	210	ILE	4.5
1	C	214	LYS	4.5
1	A	210	ILE	4.4
1	C	211	GLN	4.4
1	C	311	PHE	4.3
1	C	41	THR	4.3
1	C	146	TYR	4.3
1	C	206	ASP	4.2
1	C	238	ALA	4.1
1	B	192	ASN	4.0
1	C	209	SER	4.0
1	C	212	GLY	4.0
1	C	310	PHE	3.8
1	B	146	TYR	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	44	SER	3.6
1	C	213	ALA	3.5
1	C	360	ASN	3.5
1	B	383	GLY	3.5
1	C	308	ARG	3.5
1	B	207	ILE	3.4
1	C	383	GLY	3.3
1	C	315	GLU	3.3
1	C	36	SER	3.3
1	B	359	ALA	3.3
1	C	263	ASP	3.2
1	B	387	TYR	3.2
1	B	385	LEU	3.2
1	C	307	ASP	3.2
1	C	207	ILE	3.2
1	B	266	LEU	3.2
1	C	3	ILE	3.1
1	A	244	SER	3.1
1	C	111	VAL	3.1
1	B	267	ALA	3.1
1	B	386	ASN	3.1
1	C	267	ALA	3.1
1	B	136	LEU	3.1
1	C	177	LEU	3.0
1	C	143	SER	3.0
1	A	41	THR	3.0
1	A	442	TRP	2.9
1	C	179	VAL	2.9
1	C	262	ALA	2.8
1	C	49	PHE	2.8
1	B	43	LYS	2.8
1	A	385	LEU	2.7
1	C	168	ASN	2.7
1	C	313	SER	2.7
1	C	87	LEU	2.7
1	C	142	GLU	2.7
1	C	385	LEU	2.7
1	B	445	GLU	2.6
1	A	383	GLY	2.6
1	B	241	ASP	2.6
1	C	432	GLU	2.6
1	B	381	PHE	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	434	LEU	2.5
1	C	216	ARG	2.5
1	C	218	ALA	2.5
1	C	42	GLU	2.5
1	B	125	CYS	2.5
1	C	155	TYR	2.4
1	C	241	ASP	2.4
1	B	468	ILE	2.4
1	A	215	ALA	2.4
1	A	384	ASP	2.4
1	A	146	TYR	2.4
1	C	113	LEU	2.4
1	B	262	ALA	2.4
1	C	421	THR	2.4
1	A	3	ILE	2.4
1	A	171	ASN	2.4
1	C	251	PHE	2.3
1	B	418	SER	2.3
1	C	107	LEU	2.3
1	C	208	LYS	2.3
1	C	159	LEU	2.3
1	A	419	LEU	2.2
1	C	266	LEU	2.2
1	A	386	ASN	2.2
1	C	254	ASP	2.2
1	C	217	ALA	2.2
1	B	206	ASP	2.2
1	C	34	ASP	2.2
1	A	191	GLY	2.2
1	C	441	GLU	2.2
1	C	205	GLU	2.2
1	B	384	ASP	2.2
1	A	243	ASN	2.1
1	A	381	PHE	2.1
1	A	209	SER	2.1
1	A	49	PHE	2.1
1	B	72	LEU	2.1
1	B	215	ALA	2.1
1	C	264	LEU	2.1
1	B	380	ILE	2.1
1	A	211	GLN	2.1
1	B	431	PRO	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	261	TYR	2.1
1	B	42	GLU	2.0
1	B	213	ALA	2.0
1	C	40	GLU	2.0
1	C	260	LEU	2.0
1	C	380	ILE	2.0
1	C	106	TRP	2.0
1	B	40	GLU	2.0
1	C	82	ILE	2.0
1	B	211	GLN	2.0
1	C	33	ASP	2.0

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 carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	PO4	B	503	5/5	0.91	0.27	49,55,55,57	0
5	EDO	B	505	4/4	0.94	0.13	20,24,24,24	0
4	CL	A	504	1/1	0.95	0.04	55,55,55,55	0
3	PO4	C	502	5/5	0.95	0.41	64,68,70,72	0
3	PO4	B	502	5/5	0.96	0.26	46,48,49,52	0
4	CL	B	504	1/1	0.96	0.11	30,30,30,30	1
3	PO4	A	502	5/5	0.97	0.34	40,42,42,42	0
4	CL	A	505	1/1	0.97	0.12	36,36,36,36	0
2	MN	B	501	1/1	0.98	0.12	38,38,38,38	0
3	PO4	C	503	5/5	0.98	0.41	71,72,74,75	0
3	PO4	A	503	5/5	0.98	0.26	33,37,39,41	0
2	MN	C	501	1/1	0.98	0.14	48,48,48,48	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	MN	A	501	1/1	0.99	0.17	33,33,33,33	0

6.5 Other polymers [i](#)

There are no such residues in this entry.