



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 12, 2024 – 08:44 PM EDT

PDB ID : 8TTC  
Title : Structure of retromer VPS29-VPS35 (483-796) complexed with Fam21A repeat 20 (1289-1302)  
Authors : Chen, K.-E.; Guo, Q.; Collins, B.M.  
Deposited on : 2023-08-13  
Resolution : 3.01 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](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 as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.37.1  
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) : 2.37.1

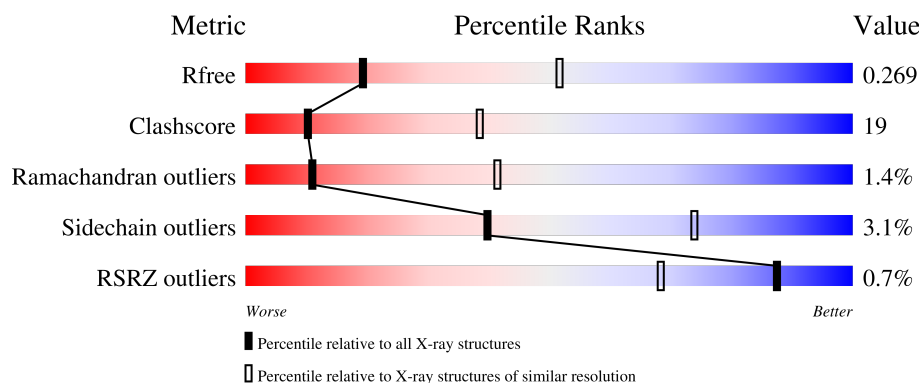
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.01 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	2399 (3.04-3.00)
Clashscore	141614	2734 (3.04-3.00)
Ramachandran outliers	138981	2640 (3.04-3.00)
Sidechain outliers	138945	2643 (3.04-3.00)
RSRZ outliers	127900	2287 (3.04-3.00)

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  $\geq 3$ , 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	182	<div> <div>7%</div> <div>62%</div> <div>36%</div> <div>•</div> </div>
1	C	182	<div> <div>55%</div> <div>43%</div> <div>••</div> </div>
2	B	316	<div> <div>57%</div> <div>32%</div> <div>•</div> <div>6%</div> </div>
2	D	316	<div> <div>59%</div> <div>34%</div> <div>•</div> <div>5%</div> </div>
3	E	14	<div> <div>7%</div> <div>21%</div> <div>71%</div> </div>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 7807 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 Vacuolar protein sorting-associated protein 29.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	182	Total	C	N	O	S	4	0	0
			1445	934	242	263	6			
1	C	181	Total	C	N	O	S	0	0	0
			1437	929	241	262	5			

- Molecule 2 is a protein called Vacuolar protein sorting-associated protein 35.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	296	Total	C	N	O	S	6	1	0
			2402	1522	422	447	11			
2	D	299	Total	C	N	O	S	31	1	0
			2428	1536	429	452	11			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	481	GLY	-	expression tag	UNP Q9EQH3
B	482	SER	-	expression tag	UNP Q9EQH3
D	481	GLY	-	expression tag	UNP Q9EQH3
D	482	SER	-	expression tag	UNP Q9EQH3

- Molecule 3 is a protein called SER-ILE-PHE-ASP-ASP-ASP-MET-ASP-ASP-ILE-PHE-SER-SER-GLY.

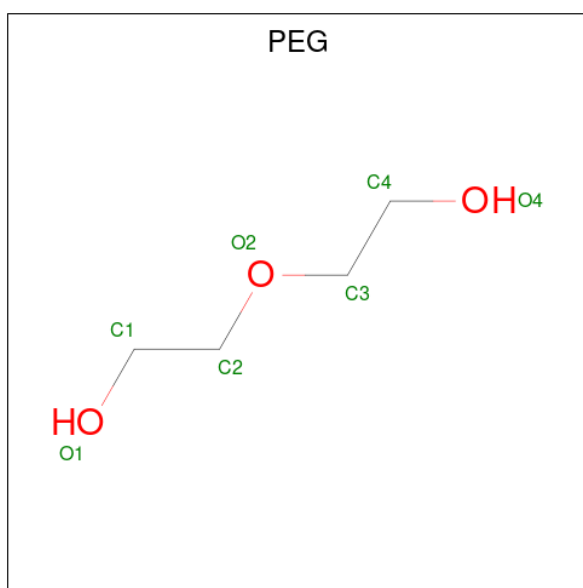
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	E	4	Total	C	N	O	0	0	0
			35	23	4	8			

- Molecule 4 is CITRIC ACID (three-letter code: CIT) (formula: C<sub>6</sub>H<sub>8</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			13	6	7		
4	B	1	Total	C	O	0	0
			13	6	7		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 7 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	D	1	Total	C	O	0	0
			4	2	2		

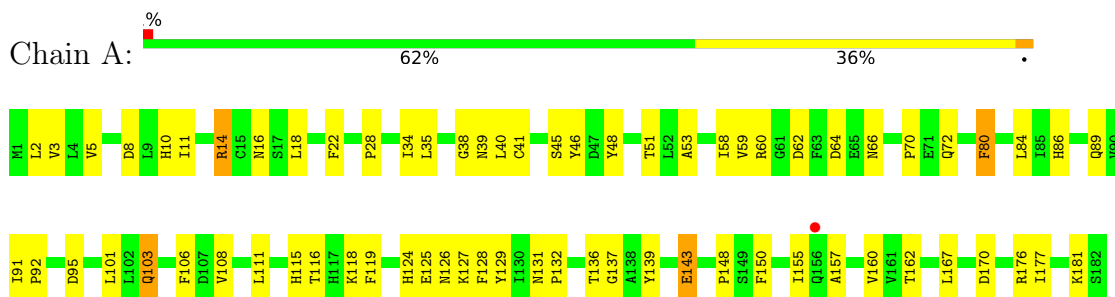
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	2	Total 2	O 2	0	0
8	B	4	Total 4	O 4	0	0
8	C	1	Total 1	O 1	0	0
8	D	10	Total 10	O 10	0	0

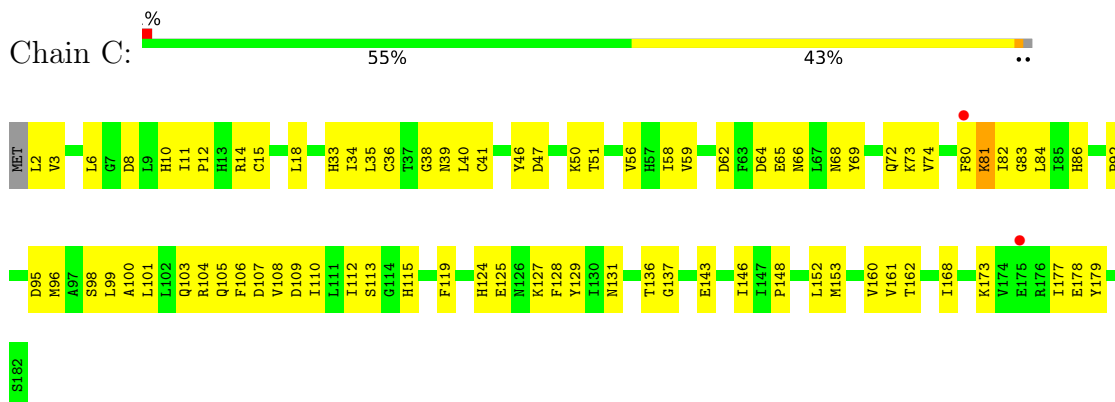
### 3 Residue-property plots

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: Vacuolar protein sorting-associated protein 29



- Molecule 1: Vacuolar protein sorting-associated protein 29

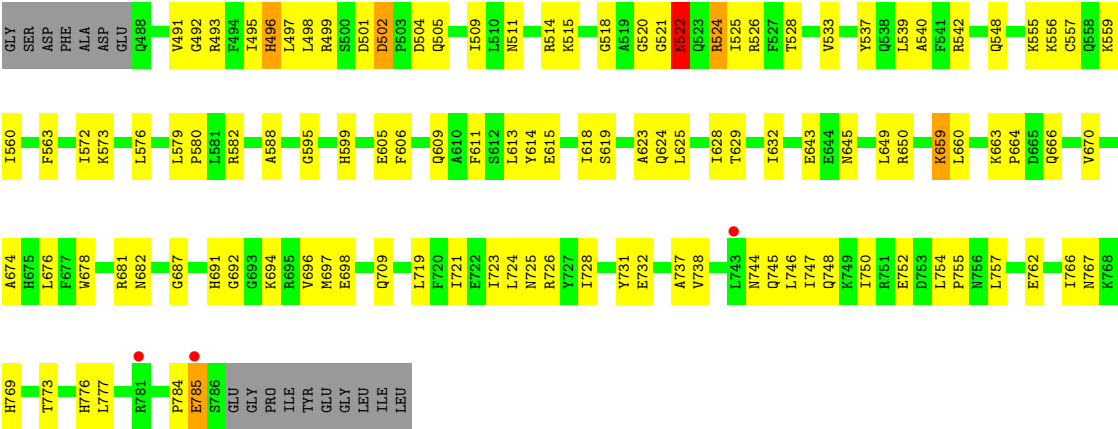


- Molecule 2: Vacuolar protein sorting-associated protein 35

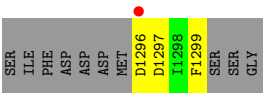


SER  
GLY  
GLY  
PRO  
ILE  
TYR  
GLU  
GLY  
ILE  
LEU

● Molecule 2: Vacuolar protein sorting-associated protein 35



● Molecule 3: SER-ILE-PHE-ASP-ASP-ASP-MET-ASP-ASP-ILE-PHE-SER-SER-GLY





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.76Å 88.76Å 328.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.58 – 3.01 46.58 – 3.01	Depositor EDS
% Data completeness (in resolution range)	99.5 (46.58-3.01) 99.5 (46.58-3.01)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.24 (at 3.01Å)	Xtriage
Refinement program	PHENIX (1.19.2_4158: ???)	Depositor
R, $R_{free}$	0.238 , 0.271 0.247 , 0.269	Depositor DCC
$R_{free}$ test set	1343 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	90.6	Xtriage
Anisotropy	0.531	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 67.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	7807	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	101.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: PEG, ACT, GOL, CIT

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	0.28	0/1478	0.60	0/2004
1	C	0.32	0/1470	0.61	0/1994
2	B	0.34	0/2452	0.68	2/3301 (0.1%)
2	D	0.32	0/2478	0.64	0/3335
3	E	0.24	0/35	0.35	0/46
All	All	0.32	0/7913	0.64	2/10680 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	690	LEU	CA-CB-CG	5.25	127.36	115.30
2	B	618	ILE	CG1-CB-CG2	-5.10	100.18	111.40

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	1445	0	1457	50	0
1	C	1437	0	1445	65	0
2	B	2402	0	2389	105	0
2	D	2428	0	2415	80	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	35	0	27	3	0
4	A	13	0	5	1	0
4	B	13	0	5	2	0
5	B	7	0	10	0	0
6	B	6	0	8	0	0
7	D	4	0	3	0	0
8	A	2	0	0	0	0
8	B	4	0	0	1	0
8	C	1	0	0	0	0
8	D	10	0	0	0	0
All	All	7807	0	7764	290	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 19.

The worst 5 of 290 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)
2:B:714:SER:O	2:B:717:VAL:HG23	1.31	1.24
2:B:713:PRO:O	2:B:717:VAL:HG22	1.35	1.21
1:C:8:ASP:H	1:C:38:GLY:HA3	1.17	1.07
1:A:8:ASP:H	1:A:38:GLY:HA3	1.26	1.01
1:C:110:ILE:HG22	1:C:128:PHE:HB3	1.45	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	180/182 (99%)	171 (95%)	7 (4%)	2 (1%)	14	48
1	C	179/182 (98%)	169 (94%)	10 (6%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	295/316 (93%)	274 (93%)	16 (5%)	5 (2%)	9	37
2	D	298/316 (94%)	274 (92%)	18 (6%)	6 (2%)	7	32
3	E	2/14 (14%)	2 (100%)	0	0	100	100
All	All	954/1010 (94%)	890 (93%)	51 (5%)	13 (1%)	11	41

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	495	ILE
2	B	499	ARG
2	B	500	SER
2	D	495	ILE
2	D	496	HIS

### 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	160/160 (100%)	157 (98%)	3 (2%)	57	83
1	C	159/160 (99%)	157 (99%)	2 (1%)	69	88
2	B	259/274 (94%)	243 (94%)	16 (6%)	18	50
2	D	262/274 (96%)	255 (97%)	7 (3%)	44	76
3	E	4/13 (31%)	4 (100%)	0	100	100
All	All	844/881 (96%)	816 (97%)	28 (3%)	40	72

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	681	ARG
2	D	726	ARG
2	B	726	ARG
2	D	582[A]	ARG
2	B	718	GLN

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	10	HIS
2	B	769	HIS
2	D	776	HIS

### 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 ⓘ

5 ligands are modelled in this entry.

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	$\# Z  > 2$
4	CIT	A	201	-	12,12,12	1.23	1 (8%)	17,17,17	1.32	2 (11%)
5	PEG	B	801	-	6,6,6	0.49	0	5,5,5	0.27	0
6	GOL	B	802	-	5,5,5	0.91	0	5,5,5	1.03	0
7	ACT	D	801	-	3,3,3	1.30	0	3,3,3	1.50	0
4	CIT	B	803	-	12,12,12	1.39	1 (8%)	17,17,17	1.53	3 (17%)

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
4	CIT	B	803	-	-	12/16/16/16	-
5	PEG	B	801	-	-	1/4/4/4	-
6	GOL	B	802	-	-	0/4/4/4	-
4	CIT	A	201	-	-	9/16/16/16	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	803	CIT	C3-C6	-3.35	1.49	1.53
4	A	201	CIT	C3-C6	-2.57	1.50	1.53

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	803	CIT	O6-C6-C3	3.66	119.40	113.05
4	A	201	CIT	O6-C6-C3	3.23	118.66	113.05
4	B	803	CIT	O5-C6-C3	-3.18	117.75	122.25
4	A	201	CIT	O5-C6-C3	-2.51	118.70	122.25
4	B	803	CIT	C3-C4-C5	-2.09	108.75	113.81

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	201	CIT	C1-C2-C3-O7
4	A	201	CIT	C1-C2-C3-C6
4	A	201	CIT	C2-C3-C6-O5
4	A	201	CIT	C2-C3-C6-O6
4	A	201	CIT	O7-C3-C6-O5

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	201	CIT	1	0
4	B	803	CIT	2	0

## 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	182/182 (100%)	-0.17	1 (0%) 91 75	65, 96, 134, 196	1 (0%)
1	C	181/182 (99%)	-0.12	2 (1%) 80 55	68, 109, 155, 193	0
2	B	296/316 (93%)	-0.18	0 100 100	61, 95, 149, 192	2 (0%)
2	D	299/316 (94%)	-0.15	3 (1%) 82 58	62, 91, 142, 188	8 (2%)
3	E	4/14 (28%)	1.34	1 (25%) 0 0	141, 149, 162, 183	0
All	All	962/1010 (95%)	-0.15	7 (0%) 87 68	61, 96, 149, 196	11 (1%)

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	781	ARG	3.2
3	E	1296	ASP	2.9
1	C	175	GLU	2.5
2	D	743	LEU	2.3
1	C	80	PHE	2.1

### 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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
5	PEG	B	801	7/7	0.44	0.24	126,141,149,152	0
6	GOL	B	802	6/6	0.68	0.14	120,129,136,139	0
7	ACT	D	801	4/4	0.82	0.19	94,102,108,109	0
4	CIT	A	201	13/13	0.93	0.36	20,20,20,20	0
4	CIT	B	803	13/13	0.94	0.42	20,20,20,20	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.