



wwPDB X-ray Structure Validation Summary Report

Mar 2, 2024 – 10:25 PM EST

PDB ID : 5VOL
Title : Bacint_04212 ferulic acid esterase
Authors : Koropatkin, N.M.; Cann, I.; Mackie, R.I.
Deposited on : 2017-05-03
Resolution : 1.98 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
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

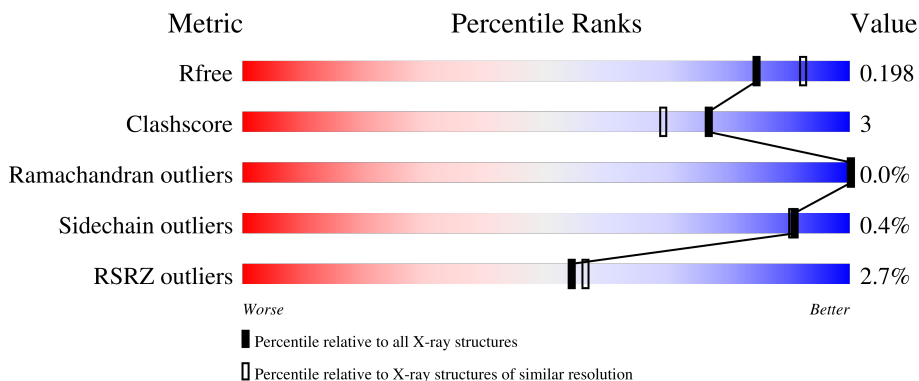
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.98 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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	275	 91% 5%
1	B	275	 88% 8%
1	C	275	 86% 11%
1	D	275	 88% 8%
1	E	275	 88% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	275	 <p>% 86% 8% 5%</p>
1	G	275	 <p>3% 87% 7% 6%</p>
1	H	275	 <p>10% 83% 10% 7%</p>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 18636 atoms, of which 144 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 Putative esterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	260	2113	1354	359	384	16	0	1	0
1	B	264	2142	1371	364	391	16	0	1	0
1	C	268	2171	1392	369	394	16	0	0	0
1	D	264	2138	1368	364	390	16	0	0	0
1	E	264	2146	1377	364	389	16	0	0	0
1	F	260	2114	1355	359	383	17	0	1	0
1	G	258	2097	1345	357	379	16	0	0	0
1	H	257	2088	1340	355	377	16	0	0	0

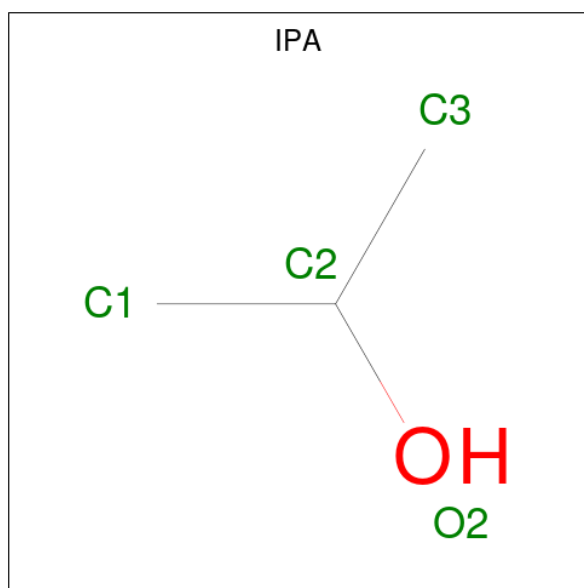
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	14	LYS	-	expression tag	UNP B3CET1
B	14	LYS	-	expression tag	UNP B3CET1
C	14	LYS	-	expression tag	UNP B3CET1
D	14	LYS	-	expression tag	UNP B3CET1
E	14	LYS	-	expression tag	UNP B3CET1
F	14	LYS	-	expression tag	UNP B3CET1
G	14	LYS	-	expression tag	UNP B3CET1
H	14	LYS	-	expression tag	UNP B3CET1

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Ca 1 1	0	0
2	B	1	Total Ca 1 1	0	0
2	C	1	Total Ca 1 1	0	0
2	D	1	Total Ca 1 1	0	0
2	E	1	Total Ca 1 1	0	0
2	F	1	Total Ca 1 1	0	0
2	G	1	Total Ca 1 1	0	0
2	H	1	Total Ca 1 1	0	0

- Molecule 3 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C₃H₈O).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C H O 12 3 8 1	0	0
3	A	1	Total C H O 12 3 8 1	0	0
3	A	1	Total C H O 12 3 8 1	0	0
3	B	1	Total C H O 12 3 8 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	H	O	0	0
			12	3	8	1		
3	B	1	Total	C	H	O	0	0
			12	3	8	1		
3	C	1	Total	C	H	O	0	0
			12	3	8	1		
3	C	1	Total	C	H	O	0	0
			12	3	8	1		
3	C	1	Total	C	H	O	0	0
			12	3	8	1		
3	D	1	Total	C	H	O	0	0
			12	3	8	1		
3	D	1	Total	C	H	O	0	0
			12	3	8	1		
3	D	1	Total	C	H	O	0	0
			12	3	8	1		
3	D	1	Total	C	H	O	0	0
			12	3	8	1		
3	E	1	Total	C	H	O	0	0
			12	3	8	1		
3	F	1	Total	C	H	O	0	0
			12	3	8	1		
3	F	1	Total	C	H	O	0	0
			12	3	8	1		
3	G	1	Total	C	H	O	0	0
			12	3	8	1		
3	H	1	Total	C	H	O	0	0
			12	3	8	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	222	Total	O	0	0
			222	222		
4	B	216	Total	O	0	0
			216	216		
4	C	189	Total	O	0	0
			189	189		
4	D	197	Total	O	0	0
			197	197		
4	E	184	Total	O	0	0
			184	184		

Continued on next page...

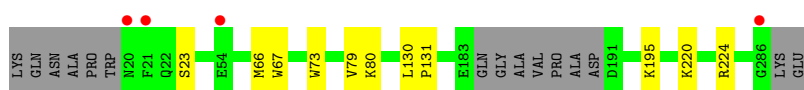
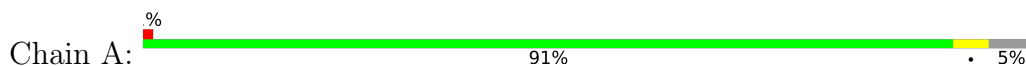
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	F	160	Total 160	O 160	0	0
4	G	151	Total 151	O 151	0	0
4	H	84	Total 84	O 84	0	0

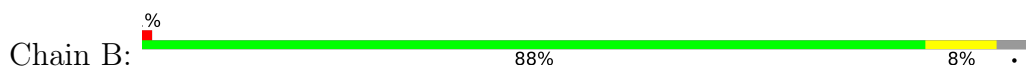
3 Residue-property plots [i](#)

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.

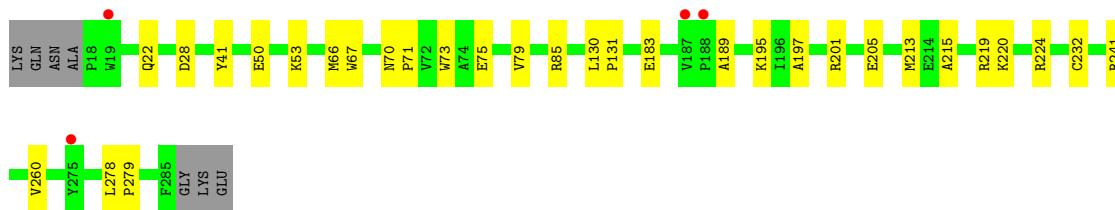
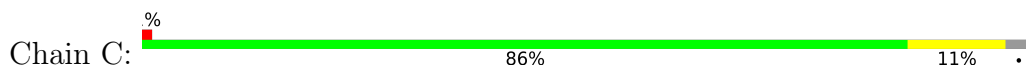
- Molecule 1: Putative esterase



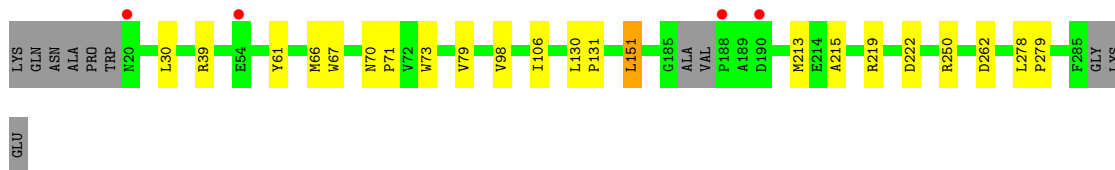
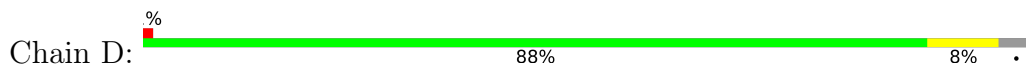
- Molecule 1: Putative esterase



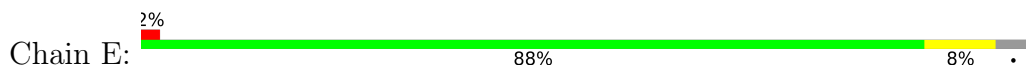
- Molecule 1: Putative esterase

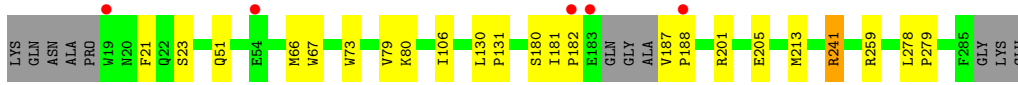


- Molecule 1: Putative esterase

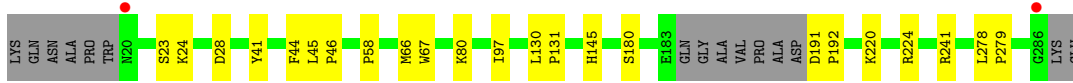
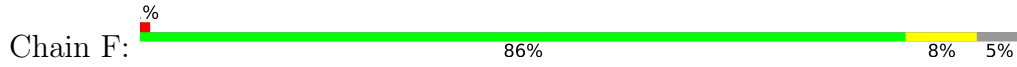


- Molecule 1: Putative esterase

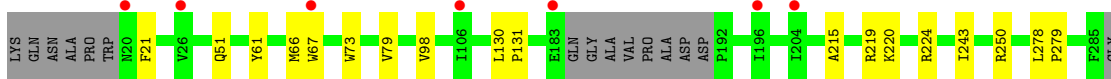
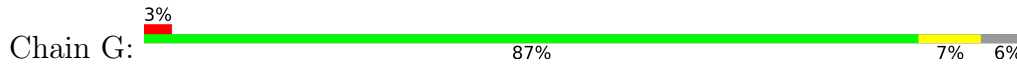




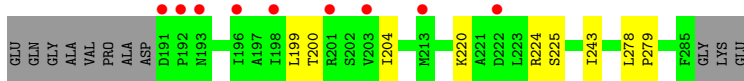
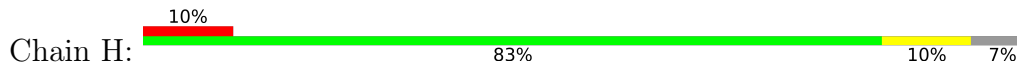
● Molecule 1: Putative esterase



● Molecule 1: Putative esterase



● Molecule 1: Putative esterase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	127.92Å 135.79Å 172.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	56.65 – 1.98 56.65 – 1.98	Depositor EDS
% Data completeness (in resolution range)	99.6 (56.65-1.98) 99.6 (56.65-1.98)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.44 (at 1.98Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.167 , 0.198 0.169 , 0.198	Depositor DCC
R_{free} test set	10450 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	26.5	Xtrriage
Anisotropy	0.559	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 49.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	18636	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.74% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CA, IPA

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	0.41	0/2174	0.53	0/2941
1	B	0.40	0/2204	0.52	0/2982
1	C	0.41	0/2234	0.50	0/3027
1	D	0.40	0/2197	0.52	1/2972 (0.0%)
1	E	0.37	0/2207	0.49	0/2989
1	F	0.35	0/2175	0.49	0/2941
1	G	0.32	0/2155	0.45	0/2914
1	H	0.31	0/2146	0.45	0/2903
All	All	0.37	0/17492	0.50	1/23669 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	151	LEU	CA-CB-CG	6.05	129.22	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	2113	0	2030	6	0
1	B	2142	0	2055	14	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2171	0	2080	16	0
1	D	2138	0	2048	12	0
1	E	2146	0	2055	15	0
1	F	2114	0	2032	13	0
1	G	2097	0	2017	10	0
1	H	2088	0	2008	17	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	12	24	24	0	0
3	B	12	24	24	0	0
3	C	12	24	24	0	0
3	D	16	32	32	1	0
3	E	4	8	8	0	0
3	F	8	16	16	0	0
3	G	4	8	8	0	0
3	H	4	8	8	0	0
4	A	222	0	0	1	0
4	B	216	0	0	1	0
4	C	189	0	0	1	0
4	D	197	0	0	3	0
4	E	184	0	0	1	0
4	F	160	0	0	0	0
4	G	151	0	0	1	0
4	H	84	0	0	0	0
All	All	18492	144	16469	102	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 3.

The worst 5 of 102 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:B:106:ILE:HD12	1:B:196:ILE:HG21	1.59	0.84
1:D:39:ARG:HD3	4:D:410:HOH:O	1.84	0.77
1:H:130:LEU:HB3	1:H:131:PRO:HD3	1.69	0.74

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:106:ILE:HD12	1:B:196:ILE:CG2	2.19	0.72
1:G:130:LEU:HB3	1:G:131:PRO:HD3	1.76	0.66

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	257/275 (94%)	251 (98%)	6 (2%)	0	100	100
1	B	261/275 (95%)	255 (98%)	6 (2%)	0	100	100
1	C	266/275 (97%)	260 (98%)	6 (2%)	0	100	100
1	D	260/275 (94%)	255 (98%)	5 (2%)	0	100	100
1	E	260/275 (94%)	253 (97%)	7 (3%)	0	100	100
1	F	257/275 (94%)	252 (98%)	5 (2%)	0	100	100
1	G	254/275 (92%)	249 (98%)	5 (2%)	0	100	100
1	H	253/275 (92%)	246 (97%)	6 (2%)	1 (0%)	34	22
All	All	2068/2200 (94%)	2021 (98%)	46 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	117	PRO

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	224/234 (96%)	224 (100%)	0	100	100
1	B	227/234 (97%)	227 (100%)	0	100	100
1	C	229/234 (98%)	226 (99%)	3 (1%)	69	64
1	D	226/234 (97%)	223 (99%)	3 (1%)	69	64
1	E	227/234 (97%)	226 (100%)	1 (0%)	91	90
1	F	224/234 (96%)	224 (100%)	0	100	100
1	G	222/234 (95%)	222 (100%)	0	100	100
1	H	221/234 (94%)	221 (100%)	0	100	100
All	All	1800/1872 (96%)	1793 (100%)	7 (0%)	91	90

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	151	LEU
1	D	213	MET
1	E	241	ARG
1	D	250	ARG
1	C	213	MET

Sometimes 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 monosaccharides in this entry.

5.6 Ligand geometry

Of 26 ligands modelled in this entry, 8 are monoatomic - leaving 18 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	IPA	C	304	-	3,3,3	0.60	0	3,3,3	0.22	0
3	IPA	C	302	-	3,3,3	0.61	0	3,3,3	0.29	0
3	IPA	D	304	-	3,3,3	0.57	0	3,3,3	0.21	0
3	IPA	B	303	-	3,3,3	0.57	0	3,3,3	0.32	0
3	IPA	D	305	-	3,3,3	0.62	0	3,3,3	0.21	0
3	IPA	F	302	-	3,3,3	0.54	0	3,3,3	0.18	0
3	IPA	F	303	-	3,3,3	0.64	0	3,3,3	0.21	0
3	IPA	B	304	-	3,3,3	0.58	0	3,3,3	0.24	0
3	IPA	G	302	-	3,3,3	0.55	0	3,3,3	0.24	0
3	IPA	A	302	-	3,3,3	0.58	0	3,3,3	0.20	0
3	IPA	E	302	-	3,3,3	0.58	0	3,3,3	0.27	0
3	IPA	B	302	-	3,3,3	0.57	0	3,3,3	0.29	0
3	IPA	D	303	-	3,3,3	0.55	0	3,3,3	0.19	0
3	IPA	A	303	-	3,3,3	0.60	0	3,3,3	0.29	0
3	IPA	D	302	-	3,3,3	0.61	0	3,3,3	0.28	0
3	IPA	H	302	-	3,3,3	0.62	0	3,3,3	0.22	0
3	IPA	A	304	-	3,3,3	0.52	0	3,3,3	0.25	0
3	IPA	C	303	-	3,3,3	0.59	0	3,3,3	0.22	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	303	IPA	1	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, 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	260/275 (94%)	-0.07	4 (1%) 73 75	17, 24, 40, 68	0
1	B	264/275 (96%)	-0.13	3 (1%) 80 82	17, 24, 42, 66	0
1	C	268/275 (97%)	-0.02	4 (1%) 73 75	18, 26, 49, 69	0
1	D	264/275 (96%)	-0.20	4 (1%) 73 75	19, 25, 46, 70	0
1	E	264/275 (96%)	-0.07	5 (1%) 66 68	20, 28, 51, 76	0
1	F	260/275 (94%)	0.05	2 (0%) 86 87	20, 31, 49, 71	0
1	G	258/275 (93%)	0.05	7 (2%) 54 56	24, 36, 52, 69	0
1	H	257/275 (93%)	0.48	27 (10%) 6 7	26, 44, 65, 78	0
All	All	2095/2200 (95%)	0.01	56 (2%) 54 56	17, 29, 53, 78	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	19	TRP	6.2
1	E	183	GLU	4.1
1	H	191	ASP	3.9
1	A	286	GLY	3.9
1	H	108	LEU	3.8

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, 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	IPA	G	302	4/4	0.62	0.20	53,64,70,74	0
3	IPA	F	303	4/4	0.69	0.20	50,60,62,67	0
3	IPA	F	302	4/4	0.73	0.23	51,61,69,71	0
3	IPA	E	302	4/4	0.75	0.18	44,53,66,66	0
3	IPA	C	303	4/4	0.77	0.30	46,55,65,66	0
3	IPA	C	304	4/4	0.77	0.23	41,50,58,58	0
3	IPA	A	304	4/4	0.80	0.19	38,46,52,53	0
3	IPA	A	303	4/4	0.82	0.13	45,54,61,65	0
3	IPA	C	302	4/4	0.83	0.14	41,52,63,66	0
3	IPA	D	302	4/4	0.84	0.23	43,53,63,72	0
3	IPA	B	303	4/4	0.85	0.14	45,54,59,61	0
3	IPA	H	302	4/4	0.85	0.24	47,57,58,59	0
3	IPA	B	304	4/4	0.86	0.15	45,54,61,70	0
3	IPA	D	303	4/4	0.87	0.12	39,49,58,62	0
3	IPA	D	304	4/4	0.88	0.13	45,59,59,63	0
3	IPA	B	302	4/4	0.90	0.36	47,56,57,59	0
3	IPA	A	302	4/4	0.91	0.23	41,53,54,57	0
3	IPA	D	305	4/4	0.92	0.25	43,58,59,59	0
2	CA	G	301	1/1	0.98	0.04	39,39,39,39	0
2	CA	B	301	1/1	0.99	0.06	29,29,29,29	0
2	CA	H	301	1/1	0.99	0.04	37,37,37,37	0
2	CA	C	301	1/1	0.99	0.06	30,30,30,30	0
2	CA	D	301	1/1	0.99	0.07	29,29,29,29	0
2	CA	E	301	1/1	0.99	0.05	34,34,34,34	0
2	CA	A	301	1/1	1.00	0.06	28,28,28,28	0
2	CA	F	301	1/1	1.00	0.06	31,31,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.