



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 15, 2020 – 11:15 pm BST

PDB ID : 3FZN  
Title : Intermediate analogue in benzoylformate decarboxylase  
Authors : Bruning, M.; Berheide, M.; Meyer, D.; Golbik, R.; Bartunik, H.; Liese, A.;  
Tittmann, K.  
Deposited on : 2009-01-26  
Resolution : 1.62 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	<b>FAILED</b>
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	<b>FAILED</b>
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

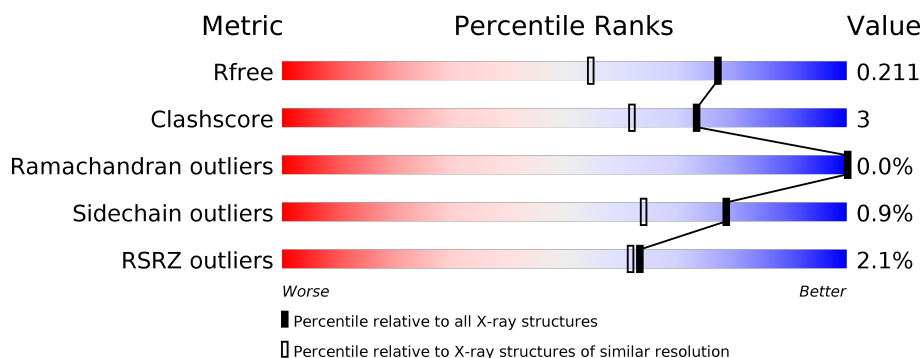
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.62 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	4693 (1.64-1.60)
Clashscore	141614	5002 (1.64-1.60)
Ramachandran outliers	138981	4888 (1.64-1.60)
Sidechain outliers	138945	4887 (1.64-1.60)
RSRZ outliers	127900	4609 (1.64-1.60)

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  $\geq 3$ , 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	534	<div> <div>2%</div> <div> <div></div> <div>93%</div> <div>6%</div> <div></div> </div> <div>• •</div> </div>
1	B	534	<div> <div>%</div> <div> <div></div> <div>94%</div> <div></div> </div> <div>• •</div> </div>
1	C	534	<div> <div>2%</div> <div> <div></div> <div>91%</div> <div>6%</div> <div></div> </div> <div>•</div> </div>
1	D	534	<div> <div>3%</div> <div> <div></div> <div>93%</div> <div>5%</div> <div></div> </div> <div>•</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	CL	A	709	-	-	X	-
4	CL	B	703	-	-	X	-
4	CL	C	711	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 18507 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 Benzoylformate decarboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	523	Total	C	N	O	S	0	18	0
			4013	2546	686	759	22			
1	B	523	Total	C	N	O	S	3	17	0
			4014	2545	689	758	22			
1	C	523	Total	C	N	O	S	0	21	0
			4033	2561	688	762	22			
1	D	523	Total	C	N	O	S	0	18	0
			4010	2544	686	758	22			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	529	HIS	-	EXPRESSION TAG	UNP P20906
A	530	HIS	-	EXPRESSION TAG	UNP P20906
A	531	HIS	-	EXPRESSION TAG	UNP P20906
A	532	HIS	-	EXPRESSION TAG	UNP P20906
A	533	HIS	-	EXPRESSION TAG	UNP P20906
A	534	HIS	-	EXPRESSION TAG	UNP P20906
B	529	HIS	-	EXPRESSION TAG	UNP P20906
B	530	HIS	-	EXPRESSION TAG	UNP P20906
B	531	HIS	-	EXPRESSION TAG	UNP P20906
B	532	HIS	-	EXPRESSION TAG	UNP P20906
B	533	HIS	-	EXPRESSION TAG	UNP P20906
B	534	HIS	-	EXPRESSION TAG	UNP P20906
C	529	HIS	-	EXPRESSION TAG	UNP P20906
C	530	HIS	-	EXPRESSION TAG	UNP P20906
C	531	HIS	-	EXPRESSION TAG	UNP P20906
C	532	HIS	-	EXPRESSION TAG	UNP P20906
C	533	HIS	-	EXPRESSION TAG	UNP P20906
C	534	HIS	-	EXPRESSION TAG	UNP P20906
D	529	HIS	-	EXPRESSION TAG	UNP P20906
D	530	HIS	-	EXPRESSION TAG	UNP P20906
D	531	HIS	-	EXPRESSION TAG	UNP P20906

*Continued on next page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	532	HIS	-	EXPRESSION TAG	UNP P20906
D	533	HIS	-	EXPRESSION TAG	UNP P20906
D	534	HIS	-	EXPRESSION TAG	UNP P20906

- [illegible]

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Mg 1 1	0	0
3	A	1	Total Mg 1 1	0	0
3	D	1	Total Mg 1 1	0	0



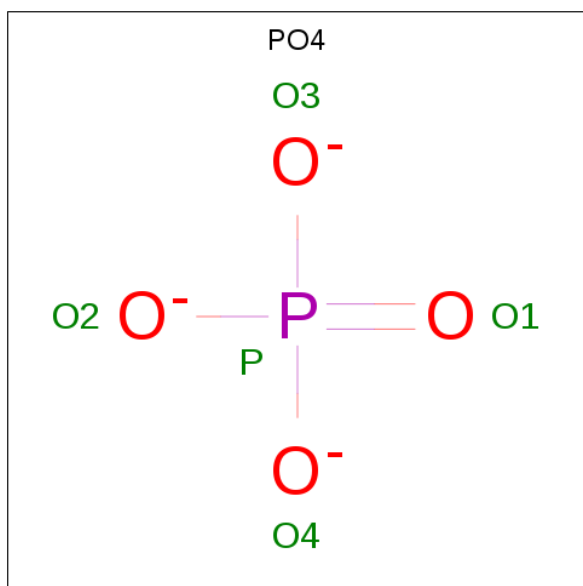
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	1	Total	Mg	0	0
			1	1		

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

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

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



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

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			7	4	3		
6	C	1	Total	C	O	0	1
			14	8	6		
6	D	1	Total	C	O	0	1
			14	8	6		

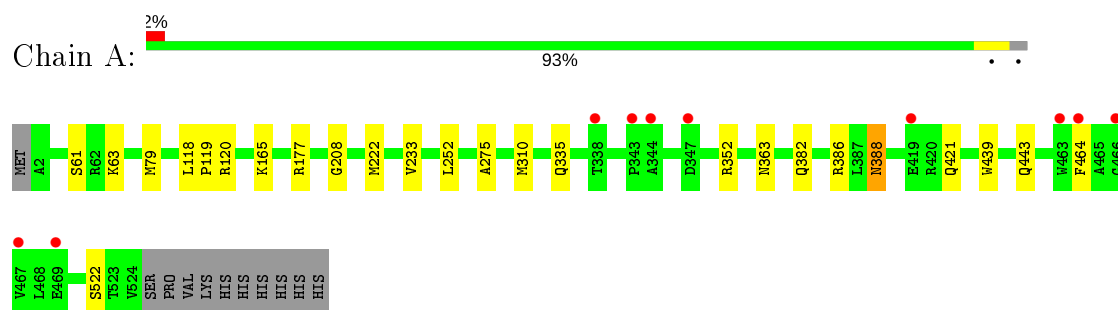
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	556	Total	O	0	0
			556	556		
7	B	560	Total	O	0	0
			560	560		
7	C	554	Total	O	0	0
			554	554		
7	D	546	Total	O	0	0
			546	546		

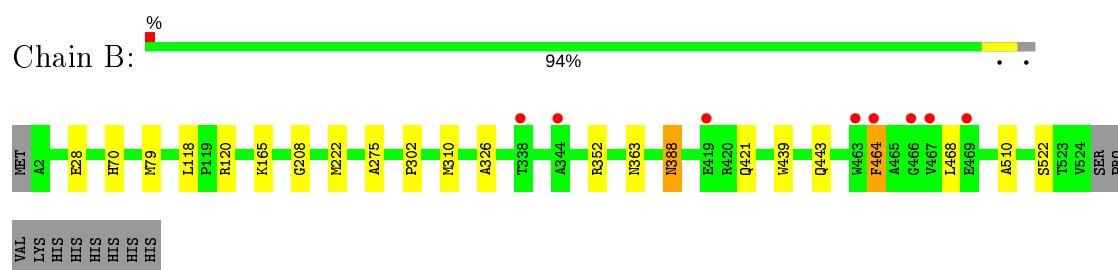
### 3 Residue-property plots [i](#)

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.

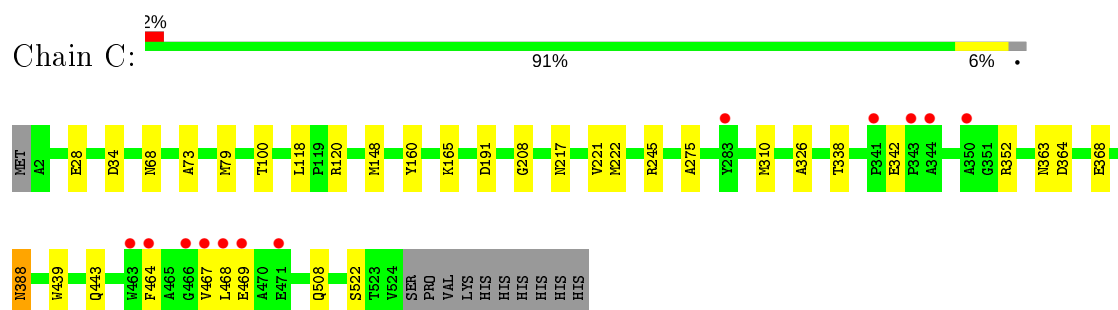
- Molecule 1: Benzoylformate decarboxylase



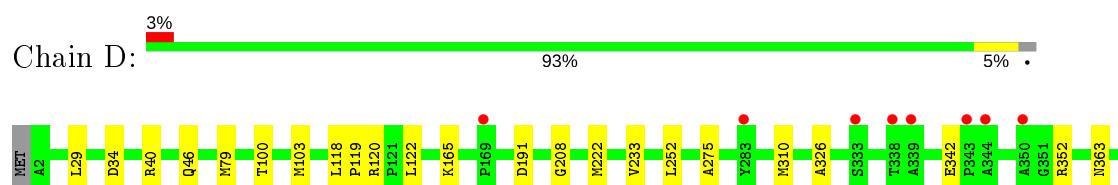
- Molecule 1: Benzoylformate decarboxylase



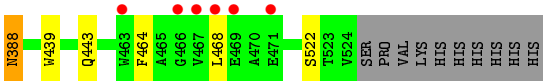
- Molecule 1: Benzoylformate decarboxylase



- Molecule 1: Benzoylformate decarboxylase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.51Å 93.34Å 94.53Å 63.46° 72.82° 73.21°	Depositor
Resolution (Å)	19.96 – 1.62 19.93 – 1.62	Depositor EDS
% Data completeness (in resolution range)	93.0 (19.96-1.62) 92.9 (19.93-1.62)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.25 (at 1.62Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.170 , 0.203 0.178 , 0.211	Depositor DCC
$R_{free}$ test set	12259 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.0	Xtriage
Anisotropy	0.190	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 71.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.339 for -h,-l,-k	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	18507	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.75% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: D7K, MG, PEG, PO4, 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	0.75	0/4167	0.68	1/5694 (0.0%)
1	B	0.73	0/4165	0.69	0/5689
1	C	0.70	0/4194	0.69	0/5729
1	D	0.71	0/4162	0.67	0/5685
All	All	0.72	0/16688	0.68	1/22797 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	1	0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	177	ARG	NE-CZ-NH2	-5.19	117.70	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	C	469	GLU	CA

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	4013	0	3974	26	0
1	B	4014	0	3984	26	0
1	C	4033	0	4004	40	2
1	D	4010	0	3977	30	0
2	A	39	0	24	1	0
2	B	39	0	24	1	0
2	C	39	0	24	0	0
2	D	39	0	24	3	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	4	0	0	5	0
4	B	4	0	0	3	0
4	C	2	0	0	4	0
4	D	1	0	0	0	0
5	A	5	0	0	0	0
5	C	10	0	0	0	0
6	C	21	0	30	2	0
6	D	14	0	20	1	0
7	A	556	0	0	4	2
7	B	560	0	0	4	0
7	C	554	0	0	12	0
7	D	546	0	0	5	0
All	All	18507	0	16085	108	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.

The worst 5 of 108 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:D:191[B]:ASP:OD2	7:D:3088:HOH:O	1.55	1.20
1:A:79[A]:MET:HE1	1:B:79[A]:MET:HE1	1.18	1.16
1:C:222[A]:MET:CE	7:C:3176:HOH:O	1.92	1.15

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:222[A]:MET:HE1	7:C:3176:HOH:O	1.50	1.09
1:A:222[A]:MET:HE1	7:A:3115:HOH:O	0.92	1.08

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:C:368[A]:GLU:OE2	7:A:2784:HOH:O[1_465]	1.71	0.49
1:C:338[A]:THR:OG1	7:A:2494:HOH:O[1_565]	2.19	0.01

## 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	539/534 (101%)	532 (99%)	7 (1%)	0	100	100
1	B	539/534 (101%)	533 (99%)	6 (1%)	0	100	100
1	C	542/534 (102%)	537 (99%)	4 (1%)	1 (0%)	47	26
1	D	538/534 (101%)	530 (98%)	8 (2%)	0	100	100
All	All	2158/2136 (101%)	2132 (99%)	25 (1%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	469	GLU

### 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	427/420 (102%)	424 (99%)	3 (1%)	84	72
1	B	427/420 (102%)	424 (99%)	3 (1%)	84	72
1	C	430/420 (102%)	426 (99%)	4 (1%)	78	64
1	D	427/420 (102%)	422 (99%)	5 (1%)	71	52
All	All	1711/1680 (102%)	1696 (99%)	15 (1%)	78	64

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	217	ASN
1	C	310	MET
1	D	342	GLU
1	B	464	PHE
1	D	310	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	391	ASN
1	C	68	ASN
1	D	363	ASN
1	B	388	ASN
1	D	382	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

Mogul failed to run properly - this section is therefore empty.

## 5.5 Carbohydrates ⓘ

Mogul failed to run properly - this section is therefore empty.

## 5.6 Ligand geometry [i](#)

Mogul failed to run properly - this section is therefore empty.

## 5.7 Other polymers [i](#)

Mogul failed to run properly - this section is therefore empty.

## 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	523/534 (97%)	-0.17	10 (1%) 66 65	11, 16, 27, 49	3 (0%)
1	B	523/534 (97%)	-0.19	8 (1%) 73 72	11, 16, 27, 49	5 (0%)
1	C	523/534 (97%)	-0.14	12 (2%) 60 58	11, 16, 30, 72	2 (0%)
1	D	523/534 (97%)	-0.12	14 (2%) 54 51	11, 16, 30, 61	4 (0%)
All	All	2092/2136 (97%)	-0.15	44 (2%) 63 62	11, 16, 29, 72	14 (0%)

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	463[A]	TRP	5.6
1	C	344	ALA	5.2
1	C	463[A]	TRP	5.2
1	C	469	GLU	4.8
1	C	466	GLY	4.7

### 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 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, 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( $\text{\AA}^2$ )	Q<0.9
6	PEG	C	716[A]	7/7	0.84	0.21	14,22,39,41	7
6	PEG	C	716[B]	7/7	0.84	0.21	19,35,55,64	7
5	PO4	C	713	5/5	0.86	0.35	104,106,111,112	0
6	PEG	D	718[B]	7/7	0.89	0.15	18,30,34,34	7
5	PO4	C	714	5/5	0.89	0.15	15,19,27,31	5
6	PEG	D	718[A]	7/7	0.89	0.15	26,32,40,40	7
6	PEG	C	715	7/7	0.92	0.12	34,39,44,53	0
5	PO4	A	712	5/5	0.93	0.30	36,47,56,68	0
2	D7K	C	604	39/39	0.95	0.09	11,15,26,36	0
2	D7K	B	602	39/39	0.96	0.08	10,14,26,27	0
2	D7K	D	603	39/39	0.96	0.09	11,15,26,28	0
2	D7K	A	601	39/39	0.96	0.08	9,14,26,29	0
4	CL	B	708	1/1	0.97	0.10	23,23,23,23	1
4	CL	B	701	1/1	0.99	0.03	23,23,23,23	0
3	MG	B	606	1/1	0.99	0.05	13,13,13,13	0
3	MG	D	608	1/1	0.99	0.04	14,14,14,14	0
4	CL	D	706	1/1	0.99	0.02	21,21,21,21	0
4	CL	A	704	1/1	0.99	0.05	22,22,22,22	0
4	CL	A	710	1/1	0.99	0.04	21,21,21,21	0
3	MG	C	607	1/1	0.99	0.06	13,13,13,13	0
4	CL	B	702	1/1	0.99	0.03	23,23,23,23	0
4	CL	B	703	1/1	0.99	0.06	22,22,22,22	0
4	CL	A	705	1/1	0.99	0.03	23,23,23,23	0
4	CL	C	707	1/1	0.99	0.04	21,21,21,21	0
3	MG	A	605	1/1	1.00	0.05	13,13,13,13	0
4	CL	C	711	1/1	1.00	0.05	12,12,12,12	0
4	CL	A	709	1/1	1.00	0.04	12,12,12,12	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.