



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 07:25 pm GMT

PDB ID : 6B07  
Title : Crystal structure of CfFPPS2, a lepidopteran type-II farnesyl diphosphate synthase, complexed with [1-phosphono-2-(1-propylpyridin-2-yl)ethyl]phosphonic acid (inhibitor 1d)  
Authors : Picard, M.-E.; Cusson, M.; Shi, R.  
Deposited on : 2017-09-13  
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](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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

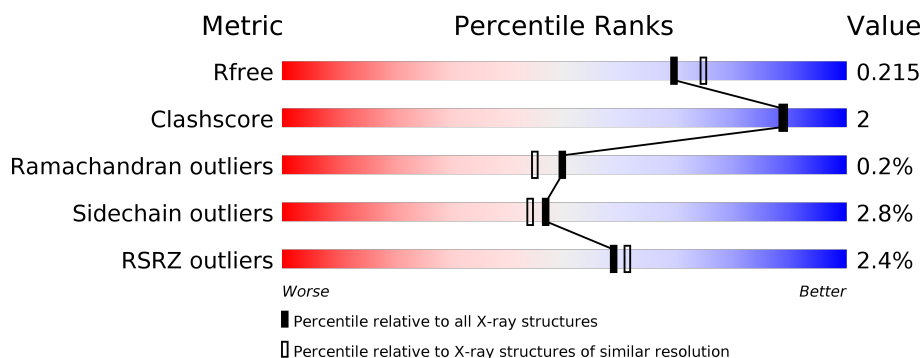
# 1 Overall quality at a glance

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))
$R_{free}$	111664	10189 (2.00-1.96)
Clashscore	122126	11405 (2.00-1.96)
Ramachandran outliers	120053	11281 (2.00-1.96)
Sidechain outliers	120020	11280 (2.00-1.96)
RSRZ outliers	108989	9953 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	341	<div> <div>3%</div> <div> <div></div> <div>91%</div> <div>8%</div> <div>.</div> </div> </div>
1	B	341	<div> <div>2%</div> <div> <div></div> <div>92%</div> <div>7%</div> <div>.</div> </div> </div>
1	C	341	<div> <div>2%</div> <div> <div></div> <div>88%</div> <div>8%</div> <div>...</div> </div> </div>

## 2 Entry composition [i](#)

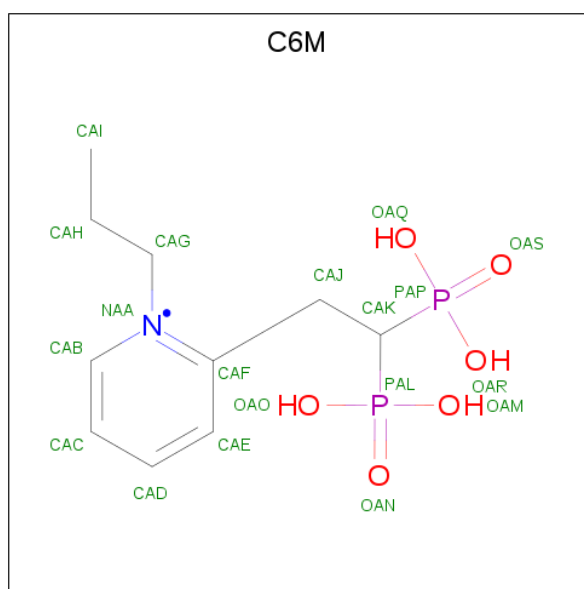
There are 5 unique types of molecules in this entry. The entry contains 9174 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Farnesyl diphosphate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	341	Total	C	N	O	S	0	2	0
			2756	1769	447	519	21			
1	B	338	Total	C	N	O	S	0	1	0
			2722	1749	439	513	21			
1	C	334	Total	C	N	O	S	0	2	0
			2691	1729	432	509	21			

- Molecule 2 is [1-phosphono-2-(1-propylpyridin-2-yl)ethyl]phosphonic acid (three-letter code: C6M) (formula:  $C_{10}H_{18}NO_6P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			19	10	1	6	2		
2	B	1	Total	C	N	O	P	0	0
			19	10	1	6	2		
2	C	1	Total	C	N	O	P	0	0
			19	10	1	6	2		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

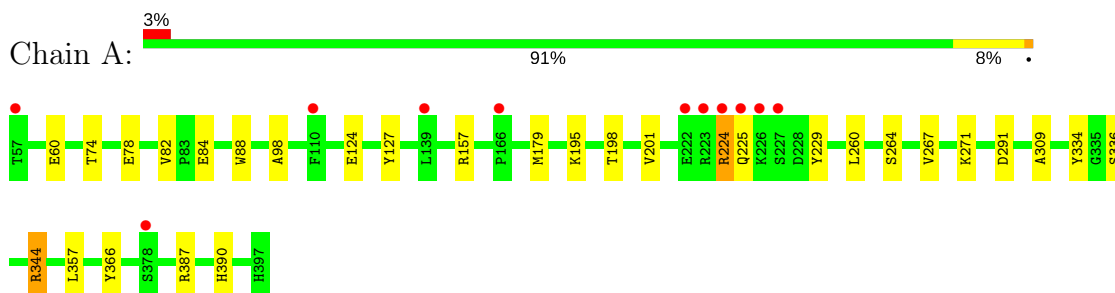
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	310	Total 310	O 310	0	0
5	B	304	Total 304	O 304	0	0
5	C	297	Total 297	O 297	0	0

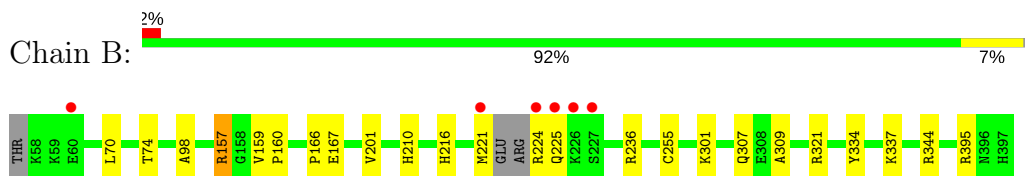
### 3 Residue-property plots [i](#)

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.

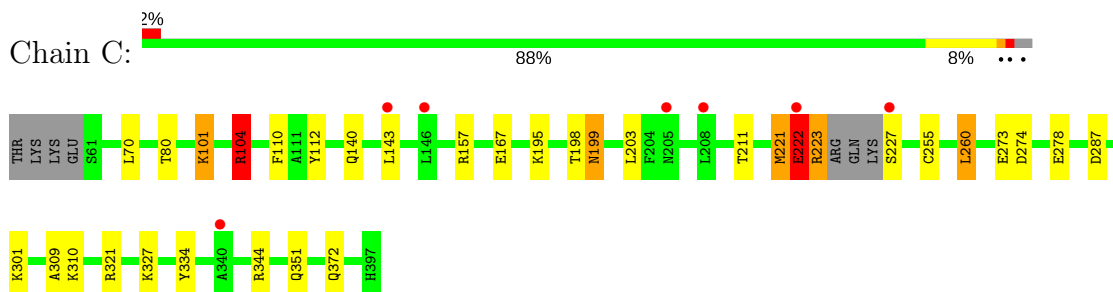
- Molecule 1: Farnesyl diphosphate synthase



- Molecule 1: Farnesyl diphosphate synthase



- Molecule 1: Farnesyl diphosphate synthase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	186.91Å 122.58Å 69.06Å 90.00° 106.51° 90.00°	Depositor
Resolution (Å)	101.18 – 1.98 33.22 – 1.98	Depositor EDS
% Data completeness (in resolution range)	95.9 (101.18-1.98) 95.9 (33.22-1.98)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.05 (at 1.98Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.171 , 0.209 0.180 , 0.215	Depositor DCC
$R_{free}$ test set	4966 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.8	Xtriage
Anisotropy	0.089	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 50.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9174	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.92% 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: MG, C6M, EDO

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	1.04	3/2821 (0.1%)	0.94	5/3816 (0.1%)
1	B	0.98	0/2782	0.91	4/3762 (0.1%)
1	C	1.06	3/2754 (0.1%)	0.99	12/3728 (0.3%)
All	All	1.03	6/8357 (0.1%)	0.95	21/11306 (0.2%)

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	0	1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	273	GLU	CD-OE1	-7.42	1.17	1.25
1	A	124	GLU	CD-OE2	-5.69	1.19	1.25
1	C	80	THR	CB-CG2	-5.68	1.33	1.52
1	A	127	TYR	CE1-CZ	5.59	1.45	1.38
1	C	112	TYR	CB-CG	-5.22	1.43	1.51

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	223	ARG	NE-CZ-NH1	8.57	124.58	120.30
1	A	344	ARG	NE-CZ-NH1	8.08	124.34	120.30
1	C	260	LEU	CB-CG-CD2	6.89	122.71	111.00
1	B	236	ARG	NE-CZ-NH1	6.43	123.52	120.30
1	B	321	ARG	NE-CZ-NH2	-6.20	117.20	120.30



There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	222	GLU	Peptide

## 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	2756	0	2754	12	0
1	B	2722	0	2720	13	0
1	C	2691	0	2679	13	0
2	A	19	0	0	0	0
2	B	19	0	0	0	0
2	C	19	0	0	0	0
3	A	3	0	0	0	0
3	B	3	0	0	0	0
3	C	3	0	0	0	0
4	A	8	0	12	0	0
4	B	8	0	12	0	0
4	C	12	0	18	0	0
5	A	310	0	0	2	1
5	B	304	0	0	5	2
5	C	297	0	0	6	1
All	All	9174	0	8195	35	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 2.

The worst 5 of 35 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:255[B]:CYS:SG	5:B:666:HOH:O	2.18	0.95
1:C:104:ARG:NH1	1:C:140:GLN:OE1	2.15	0.78
1:B:157:ARG:NH2	5:B:502:HOH:O	2.17	0.77
1:A:198:THR:HG21	5:A:762:HOH:O	1.86	0.76
1:B:98:ALA:O	1:B:157:ARG:NH1	2.20	0.75

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)
5:B:767:HOH:O	5:C:730:HOH:O[1_554]	1.99	0.21
5:A:628:HOH:O	5:B:746:HOH:O[4_545]	2.18	0.02

## 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	341/341 (100%)	331 (97%)	9 (3%)	1 (0%)	43	37
1	B	335/341 (98%)	326 (97%)	9 (3%)	0	100	100
1	C	332/341 (97%)	322 (97%)	9 (3%)	1 (0%)	43	37
All	All	1008/1023 (98%)	979 (97%)	27 (3%)	2 (0%)	49	45

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	222	GLU
1	A	74	THR

### 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	306/304 (101%)	297 (97%)	9 (3%)	45	42

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	302/304 (99%)	296 (98%)	6 (2%)	58	59
1	C	299/304 (98%)	289 (97%)	10 (3%)	41	36
All	All	907/912 (100%)	882 (97%)	25 (3%)	47	44

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	221	MET
1	B	307	GLN
1	C	227	SER
1	B	225	GLN
1	B	337	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	225	GLN
1	C	372	GLN
1	C	199	ASN
1	B	199	ASN
1	C	188	HIS

### 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 19 ligands modelled in this entry, 9 are monoatomic - leaving 10 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	$\# Z  > 2$
2	C6M	A	401	3	18,19,19	3.67	11 (61%)	24,28,28	3.13	13 (54%)
4	EDO	A	405	-	3,3,3	0.69	0	2,2,2	0.80	0
4	EDO	A	406	-	3,3,3	0.94	0	2,2,2	0.59	0
2	C6M	B	401	3	18,19,19	4.00	13 (72%)	24,28,28	2.96	12 (50%)
4	EDO	B	405	-	3,3,3	0.87	0	2,2,2	0.24	0
4	EDO	B	406	-	3,3,3	0.60	0	2,2,2	0.33	0
2	C6M	C	401	3	18,19,19	4.25	14 (77%)	24,28,28	3.49	15 (62%)
4	EDO	C	405	-	3,3,3	0.56	0	2,2,2	0.51	0
4	EDO	C	406	-	3,3,3	0.16	0	2,2,2	0.71	0
4	EDO	C	407	-	3,3,3	0.72	0	2,2,2	1.13	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
2	C6M	A	401	3	-	0/19/30/30	0/1/1/1
4	EDO	A	405	-	-	0/1/1/1	0/0/0/0
4	EDO	A	406	-	-	0/1/1/1	0/0/0/0
2	C6M	B	401	3	-	0/19/30/30	0/1/1/1
4	EDO	B	405	-	-	0/1/1/1	0/0/0/0
4	EDO	B	406	-	-	0/1/1/1	0/0/0/0
2	C6M	C	401	3	-	0/19/30/30	0/1/1/1
4	EDO	C	405	-	-	0/1/1/1	0/0/0/0
4	EDO	C	406	-	-	0/1/1/1	0/0/0/0
4	EDO	C	407	-	-	0/1/1/1	0/0/0/0

The worst 5 of 38 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	401	C6M	CAF-NAA	-7.01	1.36	1.48
2	B	401	C6M	CAF-NAA	-6.88	1.36	1.48
2	B	401	C6M	PAP-OAQ	-6.85	1.43	1.54

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	C6M	CAF-NAA	-6.81	1.36	1.48
2	A	401	C6M	CAD-CAE	-5.98	1.37	1.53

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	401	C6M	CAH-CAG-NAA	-5.46	103.22	114.87
2	B	401	C6M	CAH-CAG-NAA	-4.28	105.74	114.87
2	A	401	C6M	OAR-PAP-OAS	-4.13	102.94	113.43
2	A	401	C6M	OAO-PAL-OAN	-3.54	104.44	113.43
2	C	401	C6M	OAS-PAP-CAK	-3.36	102.43	112.13

There are no chirality outliers.

There are no torsion outliers.

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 [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	341/341 (100%)	0.09	11 (3%) 47 50	22, 31, 57, 91	0
1	B	338/341 (99%)	-0.16	6 (1%) 68 70	18, 28, 54, 95	0
1	C	334/341 (97%)	-0.12	7 (2%) 63 65	18, 29, 51, 86	0
All	All	1013/1023 (99%)	-0.06	24 (2%) 59 61	18, 29, 53, 95	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	227	SER	4.7
1	B	227	SER	4.4
1	A	225	GLN	4.1
1	A	226	LYS	3.9
1	B	226	LYS	3.9

### 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
4	EDO	B	405	4/4	0.77	0.25	44,56,56,64	0
4	EDO	A	405	4/4	0.81	0.14	45,46,46,48	0
4	EDO	C	407	4/4	0.85	0.13	35,36,37,42	0
4	EDO	A	406	4/4	0.91	0.12	28,28,29,31	0
4	EDO	C	406	4/4	0.92	0.13	38,48,49,53	0
4	EDO	B	406	4/4	0.94	0.10	44,48,49,51	0
3	MG	A	403	1/1	0.96	0.05	26,26,26,26	0
4	EDO	C	405	4/4	0.97	0.12	27,31,33,34	0
3	MG	C	403	1/1	0.98	0.05	20,20,20,20	0
3	MG	A	402	1/1	0.98	0.04	24,24,24,24	0
2	C6M	C	401	19/19	0.98	0.12	21,25,37,42	0
2	C6M	A	401	19/19	0.98	0.10	21,23,36,38	0
3	MG	B	404	1/1	0.99	0.06	20,20,20,20	0
3	MG	B	402	1/1	0.99	0.06	20,20,20,20	0
3	MG	B	403	1/1	0.99	0.05	21,21,21,21	0
3	MG	C	404	1/1	0.99	0.10	27,27,27,27	0
2	C6M	B	401	19/19	0.99	0.09	18,20,27,28	0
3	MG	C	402	1/1	0.99	0.05	24,24,24,24	0
3	MG	A	404	1/1	0.99	0.07	26,26,26,26	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.