



wwPDB X-ray Structure Validation Summary Report

Mar 23, 2024 – 04:18 PM EDT

PDB ID : 3PEA
Title : Crystal structure of enoyl-CoA hydratase from Bacillus anthracis str. 'Ames Ancestor'
Authors : Filippova, E.V.; Wawrzak, Z.; Kudritska, M.; Edwards, A.; Savchenko, A.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CS-GID)
Deposited on : 2010-10-25
Resolution : 1.82 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 : 2.36.1
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) : 2.36.1

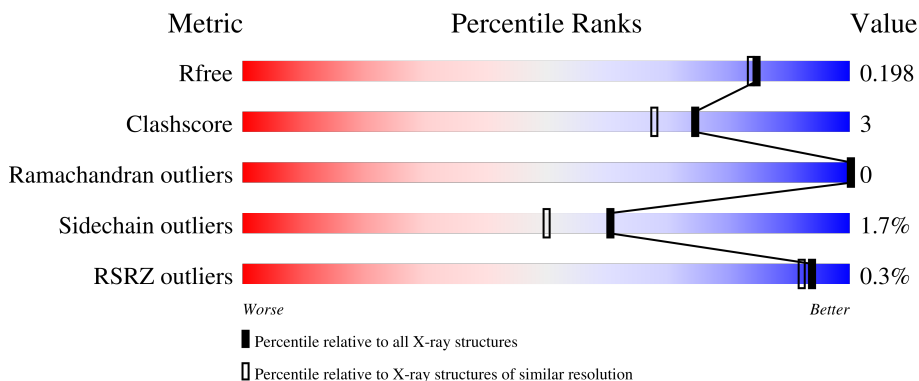
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.82 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	7484 (1.84-1.80)
Clashscore	141614	8401 (1.84-1.80)
Ramachandran outliers	138981	8290 (1.84-1.80)
Sidechain outliers	138945	8290 (1.84-1.80)
RSRZ outliers	127900	7371 (1.84-1.80)

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	261	95%
1	B	261	92%
1	C	261	93%
1	D	261	93%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	261	 92% 7% .
1	F	261	 91% 6% ..

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	FLC	A	259	-	X	X	-
2	FLC	B	259	-	-	X	-
2	FLC	D	259	-	X	X	-
2	FLC	E	259	-	-	X	-
4	ACT	E	260	-	-	X	-

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 12808 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 Enoyl-CoA hydratase/isomerase family protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	257	1983	1254	343	376	3	7	0	4	0
1	B	257	1958	1238	338	372	3	7	0	1	0
1	C	258	1962	1240	338	374	3	7	0	1	0
1	D	257	1958	1238	338	372	3	7	0	1	0
1	E	257	1958	1238	338	372	3	7	0	1	0
1	F	256	1941	1227	334	370	3	7	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

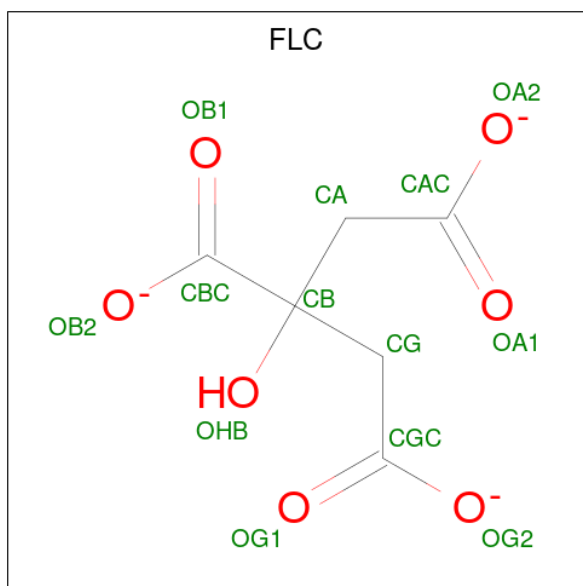
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q81L70
A	-1	ASN	-	expression tag	UNP Q81L70
A	0	ALA	-	expression tag	UNP Q81L70
B	-2	SER	-	expression tag	UNP Q81L70
B	-1	ASN	-	expression tag	UNP Q81L70
B	0	ALA	-	expression tag	UNP Q81L70
C	-2	SER	-	expression tag	UNP Q81L70
C	-1	ASN	-	expression tag	UNP Q81L70
C	0	ALA	-	expression tag	UNP Q81L70
D	-2	SER	-	expression tag	UNP Q81L70
D	-1	ASN	-	expression tag	UNP Q81L70
D	0	ALA	-	expression tag	UNP Q81L70
E	-2	SER	-	expression tag	UNP Q81L70
E	-1	ASN	-	expression tag	UNP Q81L70
E	0	ALA	-	expression tag	UNP Q81L70
F	-2	SER	-	expression tag	UNP Q81L70
F	-1	ASN	-	expression tag	UNP Q81L70

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	ALA	-	expression tag	UNP Q81L70

- Molecule 2 is CITRATE ANION (three-letter code: FLC) (formula: $C_6H_5O_7$).



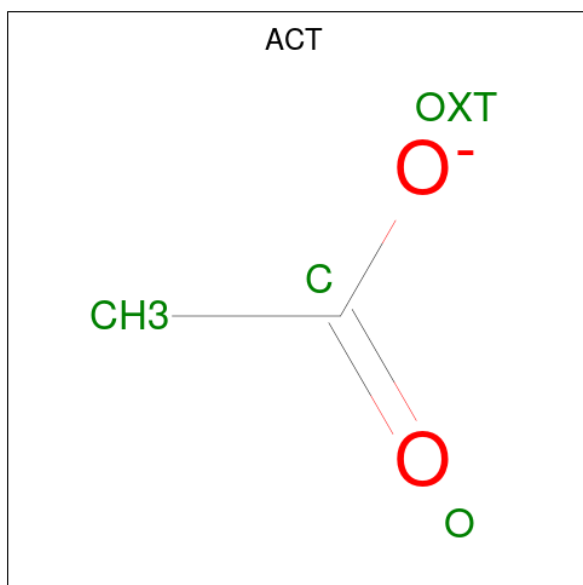
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	13	6	7	0	0
2	B	1	13	6	7	0	0
2	D	1	13	6	7	0	0
2	E	1	13	6	7	0	0

- Molecule 3 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			13	8	5		
3	C	1	Total	C	O	0	0
			13	8	5		
3	D	1	Total	C	O	0	0
			13	8	5		
3	F	1	Total	C	O	0	0
			13	8	5		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	139	Total	O	0	1
			140	140		
5	B	173	Total	O	0	1
			174	174		
5	C	151	Total	O	0	2
			153	153		
5	D	133	Total	O	0	1
			134	134		
5	E	183	Total	O	0	0
			183	183		
5	F	151	Total	O	0	1
			152	152		

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.

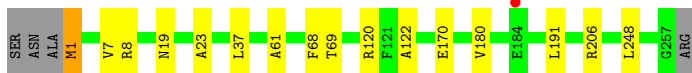
- Molecule 1: Enoyl-CoA hydratase/isomerase family protein

Chain A:  95%



- Molecule 1: Enoyl-CoA hydratase/isomerase family protein

Chain B:  92%



- Molecule 1: Enoyl-CoA hydratase/isomerase family protein

Chain C:  93%



- Molecule 1: Enoyl-CoA hydratase/isomerase family protein

Chain D:  93%




- Molecule 1: Enoyl-CoA hydratase/isomerase family protein

Chain E:  92%



- Molecule 1: Enoyl-CoA hydratase/isomerase family protein

Chain F:  91% 6% ..



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	68.95Å 75.12Å 89.49Å 89.04° 90.01° 75.47°	Depositor
Resolution (Å)	29.02 – 1.82 29.01 – 1.82	Depositor EDS
% Data completeness (in resolution range)	96.8 (29.02-1.82) 92.4 (29.01-1.82)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.40 (at 1.82Å)	Xtrriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.162 , 0.194 0.167 , 0.198	Depositor DCC
R_{free} test set	7600 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	26.6	Xtrriage
Anisotropy	0.071	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 26.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.329 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	12808	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 5.47% 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: ACT, PG4, FLC

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	0.78	0/2011	0.73	0/2703
1	B	0.82	0/1986	0.78	2/2670 (0.1%)
1	C	0.79	0/1988	0.74	1/2671 (0.0%)
1	D	0.77	0/1986	0.74	0/2670
1	E	0.86	1/1986 (0.1%)	0.76	1/2670 (0.0%)
1	F	0.82	0/1967	0.76	0/2642
All	All	0.81	1/11924 (0.0%)	0.75	4/16026 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	233	GLU	CG-CD	5.27	1.59	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	120	ARG	NE-CZ-NH2	-5.72	117.44	120.30
1	B	206	ARG	NE-CZ-NH2	5.55	123.07	120.30
1	C	206	ARG	NE-CZ-NH1	5.50	123.05	120.30
1	E	206	ARG	NE-CZ-NH1	-5.18	117.71	120.30

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	1983	0	1995	13	0
1	B	1958	0	1968	9	0
1	C	1962	0	1974	12	0
1	D	1958	0	1968	10	0
1	E	1958	0	1968	11	0
1	F	1941	0	1952	12	0
2	A	13	0	5	9	0
2	B	13	0	5	7	0
2	D	13	0	5	8	0
2	E	13	0	5	7	0
3	A	13	0	18	1	0
3	C	13	0	18	0	0
3	D	13	0	18	3	0
3	F	13	0	18	1	0
4	B	4	0	3	1	0
4	E	4	0	3	2	0
5	A	140	0	0	0	0
5	B	174	0	0	2	0
5	C	153	0	0	2	0
5	D	134	0	0	1	0
5	E	183	0	0	4	0
5	F	152	0	0	2	0
All	All	12808	0	11923	76	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 76 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:74:ALA:HB2	1:A:236:THR:HG21	1.41	1.02
1:A:153:LYS:NZ	1:C:216:LYS:HZ2	1.73	0.86
1:A:153:LYS:NZ	1:C:216:LYS:NZ	2.24	0.85
1:F:74:ALA:HB2	1:F:236:THR:HG21	1.58	0.82
1:D:74:ALA:HB2	1:D:236:THR:HG21	1.61	0.82

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	259/261 (99%)	250 (96%)	9 (4%)	0	100	100
1	B	256/261 (98%)	251 (98%)	5 (2%)	0	100	100
1	C	257/261 (98%)	251 (98%)	6 (2%)	0	100	100
1	D	256/261 (98%)	248 (97%)	8 (3%)	0	100	100
1	E	256/261 (98%)	251 (98%)	5 (2%)	0	100	100
1	F	252/261 (97%)	246 (98%)	6 (2%)	0	100	100
All	All	1536/1566 (98%)	1497 (98%)	39 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	208/200 (104%)	206 (99%)	2 (1%)	76	70
1	B	205/200 (102%)	202 (98%)	3 (2%)	65	55
1	C	205/200 (102%)	202 (98%)	3 (2%)	65	55
1	D	205/200 (102%)	202 (98%)	3 (2%)	65	55
1	E	205/200 (102%)	199 (97%)	6 (3%)	42	28
1	F	203/200 (102%)	199 (98%)	4 (2%)	55	43
All	All	1231/1200 (103%)	1210 (98%)	21 (2%)	60	50

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	135	LEU
1	F	73	GLU
1	F	191	LEU
1	F	135	LEU
1	E	191	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	19	ASN
1	F	76	GLN
1	D	40	GLN
1	F	145	GLN
1	E	85	GLN

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 [i](#)

10 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
2	FLC	B	259	-	12,12,12	2.72	2 (16%)	17,17,17	2.60	7 (41%)
3	PG4	D	260	-	12,12,12	0.67	0	11,11,11	0.69	0
3	PG4	F	259	-	12,12,12	0.77	0	11,11,11	0.77	0
3	PG4	A	260	-	12,12,12	1.16	2 (16%)	11,11,11	1.18	2 (18%)
3	PG4	C	259	-	12,12,12	0.52	0	11,11,11	0.64	0
2	FLC	E	259	-	12,12,12	2.65	3 (25%)	17,17,17	3.30	7 (41%)
4	ACT	B	260	-	3,3,3	0.82	0	3,3,3	1.27	0
4	ACT	E	260	-	3,3,3	0.90	0	3,3,3	1.00	0
2	FLC	A	259	-	12,12,12	3.42	5 (41%)	17,17,17	4.10	11 (64%)
2	FLC	D	259	-	12,12,12	3.87	3 (25%)	17,17,17	4.12	9 (52%)

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	FLC	B	259	-	-	2/16/16/16	-
3	PG4	D	260	-	-	4/10/10/10	-
3	PG4	F	259	-	-	8/10/10/10	-
3	PG4	A	260	-	-	6/10/10/10	-
3	PG4	C	259	-	-	2/10/10/10	-
2	FLC	E	259	-	-	4/16/16/16	-
2	FLC	A	259	-	-	7/16/16/16	-
2	FLC	D	259	-	-	7/16/16/16	-

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	259	FLC	CB-CBC	-11.86	1.41	1.53
2	A	259	FLC	CB-CBC	-9.88	1.43	1.53
2	B	259	FLC	CB-CBC	-7.92	1.45	1.53
2	E	259	FLC	CB-CBC	-7.83	1.45	1.53
2	A	259	FLC	OHB-CB	4.74	1.52	1.43

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	259	FLC	OHB-CB-CBC	-12.01	92.00	108.86

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	259	FLC	OHB-CB-CBC	-11.64	92.52	108.86
2	E	259	FLC	OHB-CB-CBC	-9.58	95.41	108.86
2	D	259	FLC	OB1-CBC-CB	-7.89	111.08	122.25
2	B	259	FLC	OHB-CB-CBC	-6.84	99.26	108.86

There are no chirality outliers.

5 of 40 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	259	FLC	CAC-CA-CB-CG
3	F	259	PG4	O2-C3-C4-O3
3	A	260	PG4	C6-C5-O3-C4
2	A	259	FLC	CA-CB-CG-CGC
2	B	259	FLC	CAC-CA-CB-CBC

There are no ring outliers.

9 monomers are involved in 25 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	259	FLC	7	0
3	D	260	PG4	3	0
3	F	259	PG4	1	0
3	A	260	PG4	1	0
2	E	259	FLC	7	0
4	B	260	ACT	1	0
4	E	260	ACT	2	0
2	A	259	FLC	9	0
2	D	259	FLC	8	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	250/261 (95%)	-0.45	0 100 100	19, 30, 51, 58	0
1	B	250/261 (95%)	-0.47	1 (0%) 92 91	18, 26, 49, 65	0
1	C	251/261 (96%)	-0.40	1 (0%) 92 91	18, 30, 52, 64	0
1	D	250/261 (95%)	-0.47	1 (0%) 92 91	19, 31, 52, 58	0
1	E	250/261 (95%)	-0.52	1 (0%) 92 91	18, 26, 48, 60	0
1	F	249/261 (95%)	-0.47	0 100 100	18, 29, 48, 60	0
All	All	1500/1566 (95%)	-0.46	4 (0%) 94 92	18, 29, 51, 65	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	71	VAL	3.9
1	B	184	GLU	3.0
1	C	71	VAL	2.3
1	E	57	ARG	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, 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	PG4	A	260	13/13	0.80	0.13	35,41,47,48	0
2	FLC	B	259	13/13	0.83	0.12	36,41,44,45	0
3	PG4	F	259	13/13	0.86	0.11	27,38,45,50	0
3	PG4	D	260	13/13	0.90	0.10	29,33,47,50	0
2	FLC	E	259	13/13	0.90	0.10	29,33,35,36	0
2	FLC	A	259	13/13	0.92	0.08	23,31,34,36	0
3	PG4	C	259	13/13	0.92	0.10	34,37,51,51	0
2	FLC	D	259	13/13	0.93	0.08	22,26,29,29	0
4	ACT	B	260	4/4	0.95	0.10	38,42,42,43	0
4	ACT	E	260	4/4	0.99	0.09	31,35,35,37	0

6.5 Other polymers [i](#)

There are no such residues in this entry.