



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 2, 2020 – 02:37 pm BST

PDB ID : 4FTB
Title : Crystal structure of the authentic Flock House virus particle
Authors : Speir, J.A.; Chen, Z.; Reddy, V.S.; Johnson, J.E.
Deposited on : 2012-06-27
Resolution : 2.70 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 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	:	2.11
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

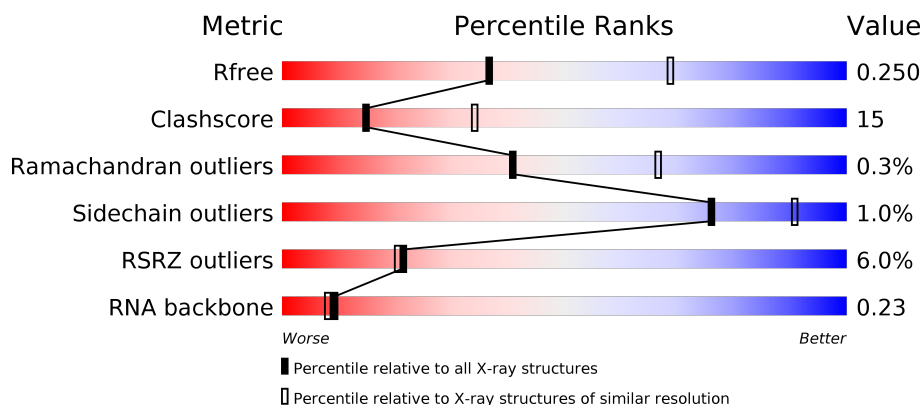
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)
RNA backbone	3102	1159 (3.00-2.40)

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 ≥ 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	363	<div> <div>2%</div> <div>67% 20% 12%</div> </div>
1	B	363	<div> <div>2%</div> <div>64% 21% 14%</div> </div>
1	C	363	<div> <div>2%</div> <div>69% 16% 15%</div> </div>
2	D	44	<div> <div>20%</div> <div>23% 25% 52%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	44	<p>25% 25% 18% 57%</p>
2	F	44	<p>11% 36% 9% 55%</p>
3	R	15	<p>87% 33% 47% 20%</p>

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 8321 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 Capsid protein beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	321	Total	C	N	O	S	0	0	0
			2407	1531	398	468	10			
1	B	311	Total	C	N	O	S	0	0	0
			2325	1481	381	452	11			
1	C	309	Total	C	N	O	S	0	0	0
			2312	1473	379	450	10			

- Molecule 2 is a protein called Capsid protein gamma.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	21	Total	C	N	O	S	0	0	0
			157	99	28	29	1			
2	E	19	Total	C	N	O	S	0	0	0
			141	89	25	26	1			
2	F	20	Total	C	N	O	S	0	0	0
			149	93	27	28	1			

- Molecule 3 is a RNA chain called Flock House virus genomic RNA.

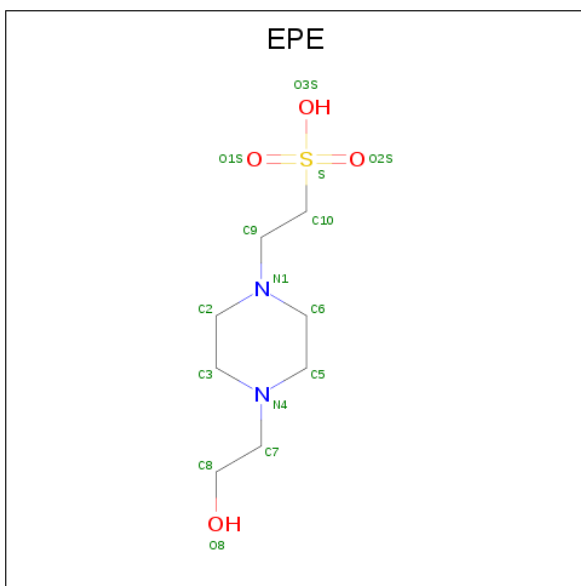
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	R	15	Total	C	N	O	P	0	0	0
			302	136	36	115	15			

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Ca	0	0
			1	1		
4	A	3	Total	Ca	0	0
			3	3		

- Molecule 5 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID

(three-letter code: EPE) (formula: C₈H₁₈N₂O₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
5	B	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
5	C	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	Cl	0	0
			1	1		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	162	Total	O	0	0
			162	162		
7	D	5	Total	O	0	0
			5	5		
7	B	144	Total	O	0	0
			144	144		
7	E	4	Total	O	0	0
			4	4		
7	C	157	Total	O	0	0
			157	157		

Continued on next page...

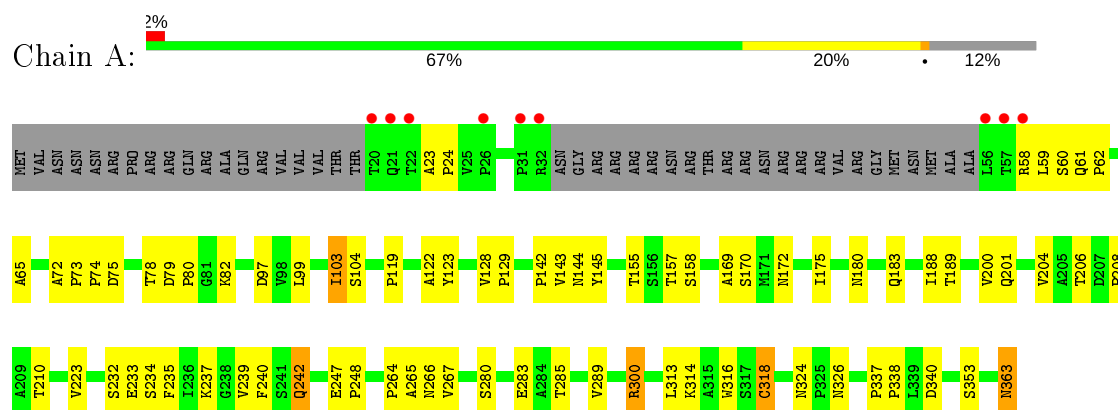
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	F	5	Total	O	0	0
			5	5		
7	R	1	Total	O	0	0
			1	1		

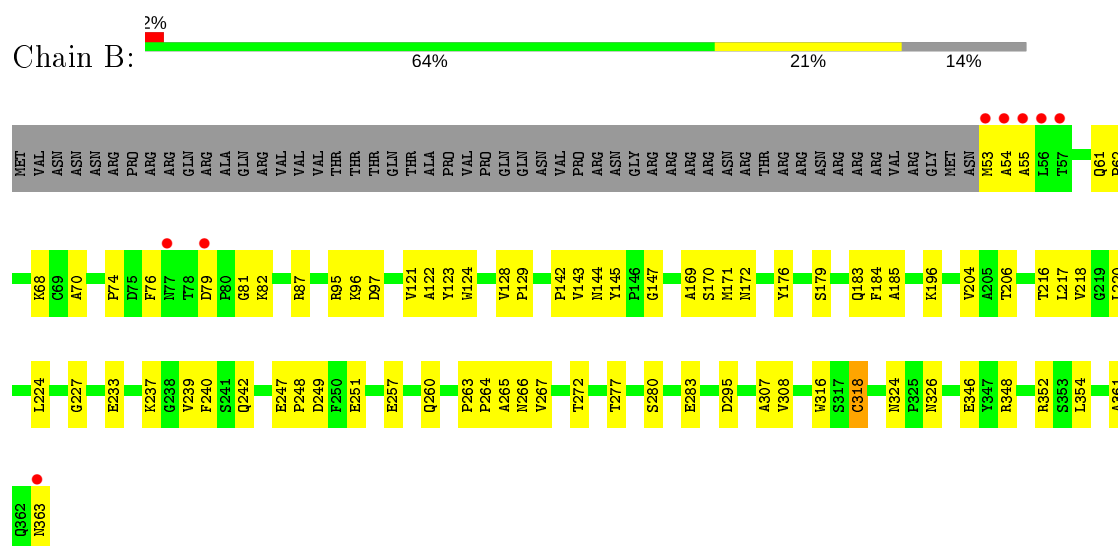
3 Residue-property plots [i](#)

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.

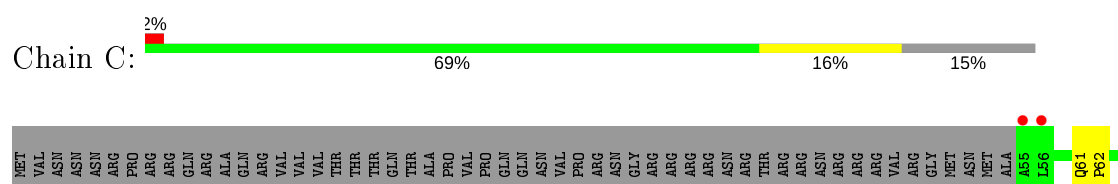
- Molecule 1: Capsid protein beta

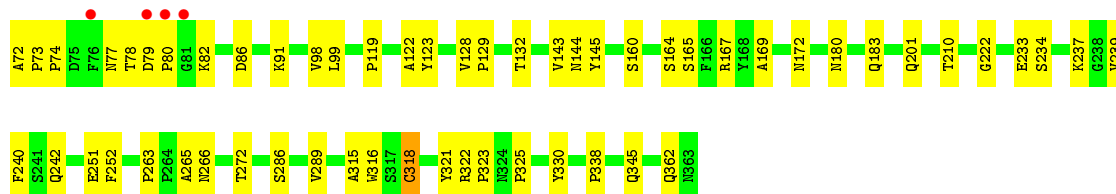


- Molecule 1: Capsid protein beta

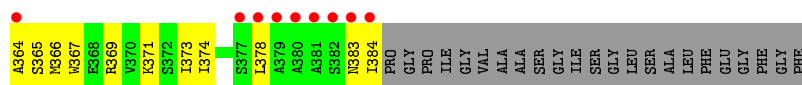


- Molecule 1: Capsid protein beta

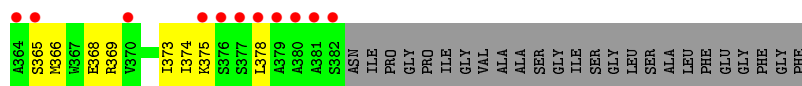




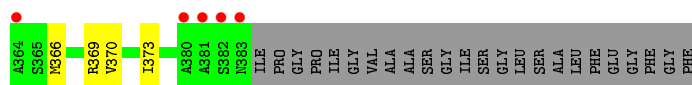
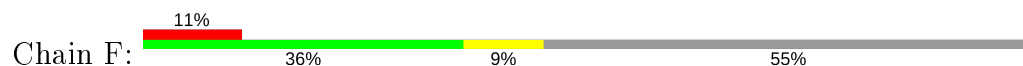
- Molecule 2: Capsid protein gamma



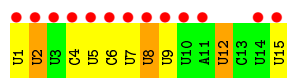
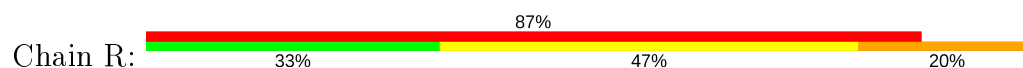
- Molecule 2: Capsid protein gamma



- Molecule 2: Capsid protein gamma



- Molecule 3: Flock House virus genomic RNA



4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	327.66Å 327.66Å 774.49Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.00 – 2.70 46.91 – 2.70	Depositor EDS
% Data completeness (in resolution range)	82.2 (40.00-2.70) 82.3 (46.91-2.70)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.87 (at 2.69Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.240 , (Not available) 0.245 , 0.250	Depositor DCC
R_{free} test set	34853 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	52.5	Xtriage
Anisotropy	0.233	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 42.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.55$, $\langle L^2 \rangle = 0.39$	Xtriage
Estimated twinning fraction	0.000 for -h,1/3*h-1/3*k-1/3*l,-4/3*h-8/3*k +1/3*l 0.000 for -1/3*h+1/3*k+1/3*l,-k,8/3*h+4/ 3*k+1/3*l 0.000 for -2/3*h-1/3*k-1/3*l,-1/3*h-2/3*k+ 1/3*l,-4/3*h+4/3*k+1/3*l 0.000 for 1/3*h+2/3*k-1/3*l,-k,-8/3*h-4/3* k-1/3*l 0.000 for -1/3*h-2/3*k+1/3*l,-2/3*h-1/3*k- 1/3*l,4/3*h-4/3*k-1/3*l 0.000 for -h,2/3*h+1/3*k+1/3*l,4/3*h+8/3 *k-1/3*l 0.000 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	8321	wwPDB-VP
Average B, all atoms (Å ²)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.32% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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: CA, EPE, 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	0.48	0/2471	0.70	0/3384
1	B	0.48	0/2387	0.72	0/3267
1	C	0.46	0/2374	0.69	0/3250
2	D	0.40	0/158	0.50	0/211
2	E	0.35	0/142	0.48	0/189
2	F	0.38	0/150	0.52	0/200
3	R	0.33	0/332	0.74	0/511
All	All	0.46	0/8014	0.69	0/11012

There are no bond length outliers.

There are no bond angle outliers.

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	2407	0	2341	77	0
1	B	2325	0	2260	75	0
1	C	2312	0	2246	65	0
2	D	157	0	167	8	0
2	E	141	0	150	12	0
2	F	149	0	156	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	R	302	0	155	18	0
4	A	3	0	0	0	0
4	B	1	0	0	0	0
5	A	15	0	17	1	0
5	B	15	0	17	2	0
5	C	15	0	17	0	0
6	B	1	0	0	0	0
7	A	162	0	0	3	0
7	B	144	0	0	1	0
7	C	157	0	0	4	0
7	D	5	0	0	0	0
7	E	4	0	0	0	0
7	F	5	0	0	0	0
7	R	1	0	0	0	0
All	All	8321	0	7526	233	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 15.

The worst 5 of 233 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)
3:R:1:U:H2'	3:R:2:U:C5'	1.42	1.45
3:R:1:U:C2'	3:R:2:U:H5''	1.47	1.42
1:A:180:ASN:OD1	1:A:183:GLN:HG2	1.42	1.16
1:C:128:VAL:HG13	1:C:129:PRO:HD2	1.38	1.04
3:R:7:U:H2'	3:R:8:U:H4'	1.39	1.04

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	317/363 (87%)	308 (97%)	9 (3%)	0	100	100
1	B	309/363 (85%)	294 (95%)	13 (4%)	2 (1%)	25	50
1	C	307/363 (85%)	299 (97%)	8 (3%)	0	100	100
2	D	19/44 (43%)	17 (90%)	1 (5%)	1 (5%)	2	3
2	E	17/44 (39%)	17 (100%)	0	0	100	100
2	F	18/44 (41%)	18 (100%)	0	0	100	100
All	All	987/1221 (81%)	953 (97%)	31 (3%)	3 (0%)	41	66

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	81	GLY
2	D	365	SER
1	B	361	ALA

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	268/305 (88%)	262 (98%)	6 (2%)	52	79
1	B	257/305 (84%)	256 (100%)	1 (0%)	91	97
1	C	256/305 (84%)	255 (100%)	1 (0%)	91	97
2	D	17/31 (55%)	17 (100%)	0	100	100
2	E	15/31 (48%)	15 (100%)	0	100	100
2	F	16/31 (52%)	16 (100%)	0	100	100
All	All	829/1008 (82%)	821 (99%)	8 (1%)	76	91

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	300	ARG
1	C	318	CYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	363	ASN
1	A	242	GLN
1	A	318	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	383	ASN
1	B	144	ASN
1	C	266	ASN
1	A	363	ASN
1	C	345	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	14/15 (93%)	6 (42%)	0

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	R	2	U
3	R	4	C
3	R	5	U
3	R	8	U
3	R	12	U

There are no RNA pucker outliers to report.

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 carbohydrates in this entry.

5.6 Ligand geometry

Of 8 ligands modelled in this entry, 5 are monoatomic - leaving 3 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
5	EPE	B	403	-	15,15,15	1.49	3 (20%)	18,20,20	1.80	4 (22%)
5	EPE	C	401	-	15,15,15	1.73	4 (26%)	18,20,20	3.35	4 (22%)
5	EPE	A	404	-	15,15,15	1.64	2 (13%)	18,20,20	2.67	3 (16%)

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
5	EPE	B	403	-	-	2/9/19/19	0/1/1/1
5	EPE	C	401	-	-	2/9/19/19	0/1/1/1
5	EPE	A	404	-	-	0/9/19/19	0/1/1/1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	404	EPE	C10-S	2.91	1.81	1.77
5	C	401	EPE	O1S-S	2.65	1.52	1.45
5	C	401	EPE	C10-S	2.62	1.81	1.77
5	C	401	EPE	O2S-S	2.59	1.52	1.45
5	B	403	EPE	O1S-S	2.45	1.52	1.45

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	401	EPE	O3S-S-C10	13.31	127.29	105.77
5	A	404	EPE	O3S-S-C10	10.00	121.94	105.77
5	B	403	EPE	O3S-S-O2S	5.69	125.17	111.27
5	A	404	EPE	O2S-S-C10	-3.86	102.27	106.92

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	403	EPE	O2S-S-C10	-2.53	103.86	106.92

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	403	EPE	C9-C10-S-O3S
5	B	403	EPE	C9-C10-S-O1S
5	C	401	EPE	N4-C7-C8-O8
5	C	401	EPE	C9-C10-S-O1S

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	403	EPE	2	0
5	A	404	EPE	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	321/363 (88%)	-0.00	9 (2%) 53 54	40, 50, 95, 153	0
1	B	311/363 (85%)	-0.16	8 (2%) 56 57	39, 50, 88, 156	0
1	C	309/363 (85%)	-0.15	6 (1%) 66 69	39, 50, 70, 118	0
2	D	21/44 (47%)	2.90	9 (42%) 0 0	81, 97, 160, 164	0
2	E	19/44 (43%)	3.13	11 (57%) 0 0	99, 114, 154, 155	0
2	F	20/44 (45%)	1.04	5 (25%) 0 0	60, 77, 130, 135	0
3	R	15/15 (100%)	5.42	13 (86%) 0 0	150, 198, 248, 249	0
All	All	1016/1236 (82%)	0.12	61 (6%) 21 20	39, 51, 118, 249	0

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	54	ALA	12.5
1	B	53	MET	11.8
1	A	56	LEU	10.9
2	D	384	ILE	10.0
3	R	4	C	9.1

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 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, 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
6	CL	B	402	1/1	0.56	0.17	81,81,81,81	0
4	CA	A	403	1/1	0.64	0.36	115,115,115,115	0
4	CA	A	401	1/1	0.84	0.12	56,56,56,56	0
5	EPE	A	404	15/15	0.87	0.40	64,81,93,93	0
4	CA	A	402	1/1	0.87	0.12	51,51,51,51	0
5	EPE	C	401	15/15	0.88	0.31	60,83,92,93	0
4	CA	B	401	1/1	0.90	0.19	50,50,50,50	0
5	EPE	B	403	15/15	0.90	0.28	60,78,89,89	0

6.5 Other polymers [i](#)

There are no such residues in this entry.