



Full wwPDB X-ray Structure Validation Report ⓘ

May 25, 2020 – 05:41 am BST

PDB ID : 5K20
Title : Caspase-7 S239E Phosphomimetic
Authors : Eron, S.J.; Hardy, J.A.
Deposited on : 2016-05-18
Resolution : 2.20 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

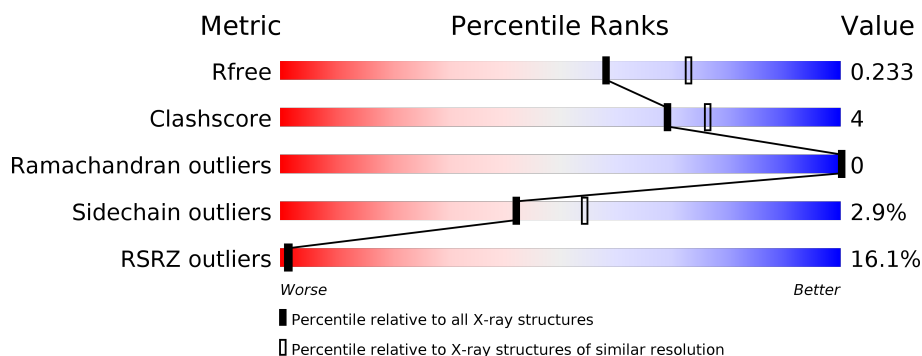
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	198	<div> <div>8%</div> <div>64%</div> <div>6%</div> <div>29%</div> </div>
1	C	198	<div> <div>4%</div> <div>63%</div> <div>8%</div> <div>29%</div> </div>
2	B	113	<div> <div>24%</div> <div>73%</div> <div>6%</div> <div>19%</div> </div>
2	D	113	<div> <div>22%</div> <div>71%</div> <div>10%</div> <div>18%</div> </div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3831 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 Caspase-7 large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	140	Total	C	N	O	S	3	0	0
			1099	691	188	209	11			
1	C	140	Total	C	N	O	S	0	0	0
			1099	691	188	209	11			

- Molecule 2 is a protein called Caspase-7 small subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	92	Total	C	N	O	S	4	0	0
			760	489	128	139	4			
2	D	93	Total	C	N	O	S	8	0	0
			772	498	129	141	4			

There are 18 discrepancies between the modelled and reference sequences:

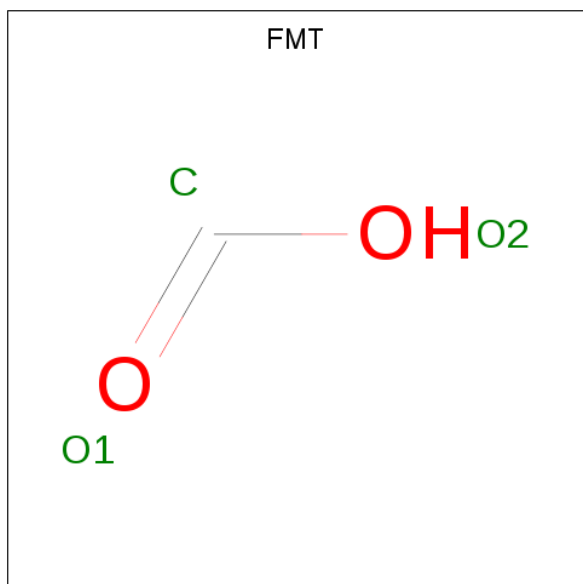
Chain	Residue	Modelled	Actual	Comment	Reference
B	239	GLU	SER	engineered mutation	UNP P55210
B	304	LEU	-	expression tag	UNP P55210
B	305	GLU	-	expression tag	UNP P55210
B	306	HIS	-	expression tag	UNP P55210
B	307	HIS	-	expression tag	UNP P55210
B	308	HIS	-	expression tag	UNP P55210
B	309	HIS	-	expression tag	UNP P55210
B	310	HIS	-	expression tag	UNP P55210
B	311	HIS	-	expression tag	UNP P55210
D	539	GLU	SER	engineered mutation	UNP P55210
D	604	LEU	-	expression tag	UNP P55210
D	605	GLU	-	expression tag	UNP P55210
D	606	HIS	-	expression tag	UNP P55210
D	607	HIS	-	expression tag	UNP P55210
D	608	HIS	-	expression tag	UNP P55210
D	609	HIS	-	expression tag	UNP P55210

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	610	HIS	-	expression tag	UNP P55210
D	611	HIS	-	expression tag	UNP P55210

- Molecule 3 is FORMIC ACID (three-letter code: FMT) (formula: CH₂O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			3	1	2		
3	C	1	Total	C	O	0	0
			3	1	2		
3	C	1	Total	C	O	0	0
			3	1	2		

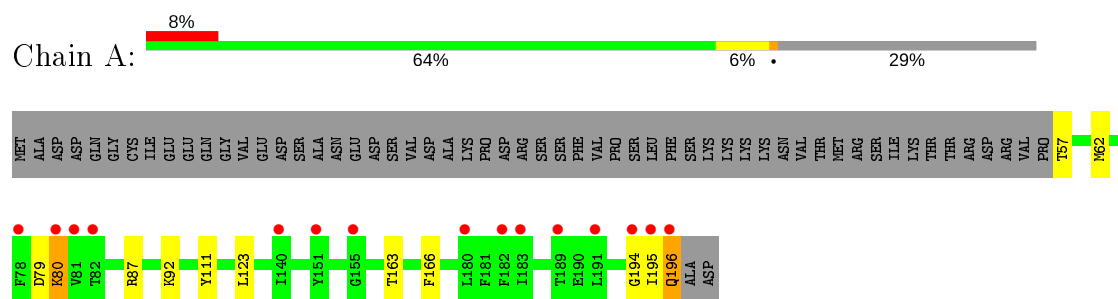
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	11	Total	O	0	0
			11	11		
4	B	13	Total	O	0	0
			13	13		
4	C	53	Total	O	0	0
			53	53		
4	D	15	Total	O	0	0
			15	15		

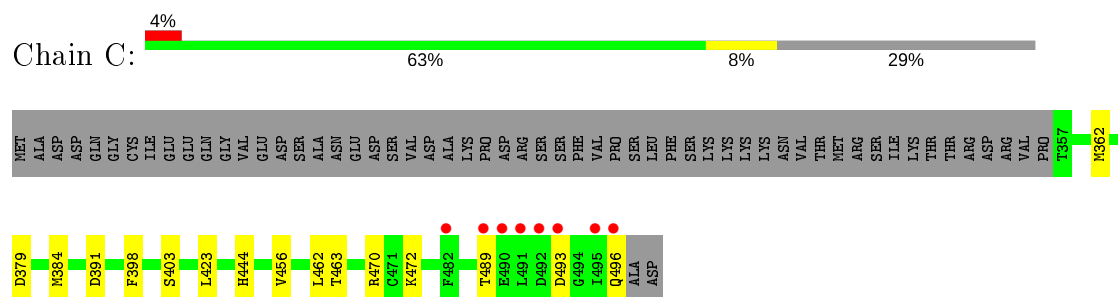
3 Residue-property plots [i](#)

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.

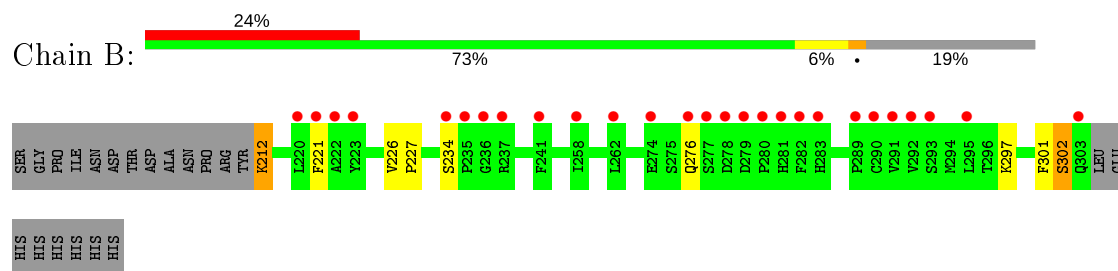
- Molecule 1: Caspase-7 large subunit



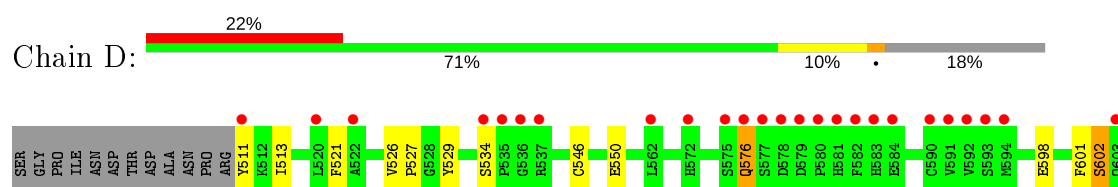
- Molecule 1: Caspase-7 large subunit



- Molecule 2: Caspase-7 small subunit



- Molecule 2: Caspase-7 small subunit



LEU
GLU
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	88.76 Å 88.76 Å 185.14 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	76.98 – 2.20 33.36 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.5 (76.98-2.20) 99.6 (33.36-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.74 (at 2.20 Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.191 , 0.228 0.200 , 0.233	Depositor DCC
R_{free} test set	4298 reflections (9.88%)	wwPDB-VP
Wilson B-factor (Å ²)	54.1	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 55.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.019 for -h,-k,l	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3831	wwPDB-VP
Average B, all atoms (Å ²)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% 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: FMT

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	1.03	1/1116 (0.1%)	0.96	3/1496 (0.2%)
1	C	0.91	1/1116 (0.1%)	0.97	2/1496 (0.1%)
2	B	0.92	1/782 (0.1%)	0.87	1/1058 (0.1%)
2	D	0.93	2/795 (0.3%)	0.96	3/1076 (0.3%)
All	All	0.95	5/3809 (0.1%)	0.95	9/5126 (0.2%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	196	GLN	CG-CD	21.33	2.00	1.51
2	D	602	SER	C-N	5.63	1.47	1.34
2	B	276	GLN	CB-CG	-5.56	1.37	1.52
1	C	403	SER	CB-OG	-5.34	1.35	1.42
2	D	598	GLU	CD-OE2	-5.08	1.20	1.25

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	196	GLN	CB-CG-CD	-9.63	86.55	111.60
1	A	87	ARG	NE-CZ-NH1	6.33	123.46	120.30
1	A	62	MET	CG-SD-CE	-6.19	90.29	100.20
1	C	362	MET	CG-SD-CE	-5.92	90.72	100.20
2	D	576	GLN	CB-CG-CD	5.76	126.59	111.60
2	B	302	SER	CB-CA-C	5.64	120.82	110.10
2	D	576	GLN	CA-CB-CG	-5.52	101.27	113.40
1	C	379	ASP	CB-CG-OD2	5.36	123.12	118.30
2	D	602	SER	O-C-N	5.14	130.92	122.70

There are no chirality outliers.

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	1099	0	1086	12	0
1	C	1099	0	1086	8	0
2	B	760	0	732	7	0
2	D	772	0	741	7	0
3	A	3	0	1	0	0
3	C	6	0	2	1	0
4	A	11	0	0	1	0
4	B	13	0	0	0	0
4	C	53	0	0	1	1
4	D	15	0	0	0	0
All	All	3831	0	3648	27	1

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 4.

All (27) 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:57:THR:O	2:B:297:LYS:HE3	1.88	0.73
2:B:301:PHE:O	2:B:302:SER:HB3	1.90	0.72
1:A:80:LYS:HE3	1:A:80:LYS:H	1.55	0.72
1:A:163:THR:HG21	2:B:221:PHE:HE1	1.60	0.67
2:D:601:PHE:O	2:D:602:SER:HB3	1.95	0.66
1:A:111:TYR:OH	4:A:301:HOH:O	2.04	0.61
2:B:212:LYS:HE2	1:C:493:ASP:O	2.07	0.55
1:C:423:LEU:HD12	1:C:462:LEU:HD22	1.89	0.55
1:C:463:THR:HG21	2:D:521:PHE:HE2	1.72	0.55
1:C:391:ASP:HB2	3:C:501:FMT:C	2.38	0.53
1:A:80:LYS:CE	1:A:80:LYS:H	2.21	0.52
1:A:123:LEU:HD12	1:A:166:PHE:CE1	2.47	0.50
2:D:526:VAL:HB	2:D:527:PRO:HD2	1.92	0.50
2:D:546:CYS:O	2:D:550:GLU:HG3	2.14	0.47
1:C:470:ARG:NH1	4:C:603:HOH:O	2.42	0.47
2:B:226:VAL:HB	2:B:227:PRO:HD2	1.97	0.47
2:D:526:VAL:HB	2:D:527:PRO:CD	2.47	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:79:ASP:HB3	1:A:80:LYS:HZ2	1.83	0.43
1:A:195:ILE:HD11	1:C:472:LYS:HE2	2.00	0.43
1:A:163:THR:HG21	2:B:221:PHE:CE1	2.48	0.43
1:A:79:ASP:HB3	1:A:80:LYS:NZ	2.33	0.43
2:D:526:VAL:HG23	2:D:529:TYR:CD1	2.53	0.43
1:C:384:MET:HB3	1:C:444:HIS:CD2	2.54	0.43
1:A:123:LEU:HD12	1:A:166:PHE:HE1	1.86	0.41
1:C:423:LEU:HA	1:C:423:LEU:HD23	1.95	0.41
1:A:194:GLY:HA2	2:D:513:ILE:O	2.21	0.41
2:B:226:VAL:HB	2:B:227:PRO:CD	2.52	0.40

All (1) 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)
4:C:651:HOH:O	4:C:651:HOH:O[5_557]	2.10	0.10

5.3 Torsion angles

5.3.1 Protein backbone

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	138/198 (70%)	136 (99%)	2 (1%)	0	100	100
1	C	138/198 (70%)	136 (99%)	2 (1%)	0	100	100
2	B	90/113 (80%)	86 (96%)	4 (4%)	0	100	100
2	D	91/113 (80%)	85 (93%)	6 (7%)	0	100	100
All	All	457/622 (74%)	443 (97%)	14 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	120/172 (70%)	117 (98%)	3 (2%)	47	60
1	C	120/172 (70%)	116 (97%)	4 (3%)	38	49
2	B	84/103 (82%)	82 (98%)	2 (2%)	49	62
2	D	85/103 (82%)	82 (96%)	3 (4%)	36	46
All	All	409/550 (74%)	397 (97%)	12 (3%)	42	54

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	80	LYS
1	A	92	LYS
1	A	196	GLN
2	B	212	LYS
2	B	234	SER
1	C	398	PHE
1	C	456	VAL
1	C	489	THR
1	C	496	GLN
2	D	511	TYR
2	D	534	SER
2	D	576	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

3 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	$\# Z > 2$
3	FMT	A	201	-	0,2,2	0.00	-	0,1,1	0.00	-
3	FMT	C	502	-	0,2,2	0.00	-	0,1,1	0.00	-
3	FMT	C	501	-	0,2,2	0.00	-	0,1,1	0.00	-

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	C	501	FMT	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

6.1 Protein, DNA and RNA chains

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	140/198 (70%)	0.27	15 (10%) 6 5	48, 66, 92, 112	1 (0%)
1	C	140/198 (70%)	-0.01	8 (5%) 23 22	44, 53, 94, 123	0
2	B	92/113 (81%)	1.31	27 (29%) 0 0	45, 61, 121, 135	1 (1%)
2	D	93/113 (82%)	1.05	25 (26%) 0 0	44, 64, 123, 141	2 (2%)
All	All	465/622 (74%)	0.55	75 (16%) 1 1	44, 60, 112, 141	4 (0%)

All (75) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	278	ASP	6.5
2	B	281	HIS	6.2
2	B	280	PRO	6.0
2	D	580	PRO	5.9
2	D	582	PHE	5.8
1	C	495	ILE	5.6
2	B	277	SER	5.5
2	B	235	PRO	5.2
2	B	282	PHE	5.1
2	D	577	SER	4.7
2	B	291	VAL	4.7
2	D	581	HIS	4.6
2	B	237	ARG	4.6
1	A	82	THR	4.5
2	B	222	ALA	4.5
2	D	534	SER	4.5
2	B	220	LEU	4.5
2	D	536	GLY	4.3
2	B	303	GLN	4.3
2	D	579	ASP	4.3
2	B	292	VAL	4.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	489	THR	4.0
2	B	279	ASP	4.0
1	A	81	VAL	3.7
2	D	535	PRO	3.7
1	A	80	LYS	3.6
1	A	182	PHE	3.5
2	B	290	CYS	3.5
2	D	578	ASP	3.5
2	D	592	VAL	3.4
2	D	511	TYR	3.4
2	D	591	VAL	3.4
2	D	572	HIS	3.4
2	D	603	GLN	3.4
1	A	183	ILE	3.4
2	B	283	HIS	3.3
1	C	490	GLU	3.3
2	B	295	LEU	3.1
2	D	583	HIS	3.1
2	D	576	GLN	3.0
2	B	221	PHE	2.9
2	D	537	ARG	2.9
1	A	180	LEU	2.9
2	B	234	SER	2.9
2	D	562	LEU	2.8
2	B	293	SER	2.8
1	C	496	GLN	2.7
2	B	262	LEU	2.7
2	D	520	LEU	2.7
1	C	492	ASP	2.6
2	B	274	GLU	2.6
1	A	195	ILE	2.6
2	D	590	CYS	2.5
2	D	584	GLU	2.5
1	A	194	GLY	2.5
2	B	289	PRO	2.5
2	B	276	GLN	2.5
2	D	575	SER	2.4
1	A	191	LEU	2.4
1	C	482	PHE	2.4
1	A	155	GLY	2.4
1	A	140	ILE	2.3
1	C	491	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	196	GLN	2.3
2	B	236	GLY	2.2
2	B	223	TYR	2.2
1	A	189	THR	2.2
2	D	593	SER	2.2
2	D	522	ALA	2.1
2	D	594	MET	2.1
1	A	78	PHE	2.1
2	B	241	PHE	2.1
2	B	258	ILE	2.0
1	A	151	TYR	2.0
1	C	493	ASP	2.0

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
3	FMT	C	502	3/3	0.78	0.22	72,72,75,96	0
3	FMT	A	201	3/3	0.86	0.18	65,65,68,71	0
3	FMT	C	501	3/3	0.97	0.14	51,51,57,60	0

6.5 Other polymers [i](#)

There are no such residues in this entry.