



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 1, 2024 – 07:16 pm BST

PDB ID : 8PJT  
Title : Crystal structure of the computationally designed SAKe6DEref protein  
Authors : Wouters, S.M.L.  
Deposited on : 2023-06-23  
Resolution : 1.70 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](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 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  
Xtriage (Phenix) : 1.13  
EDS : 2.37.1  
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) : 2.37.1

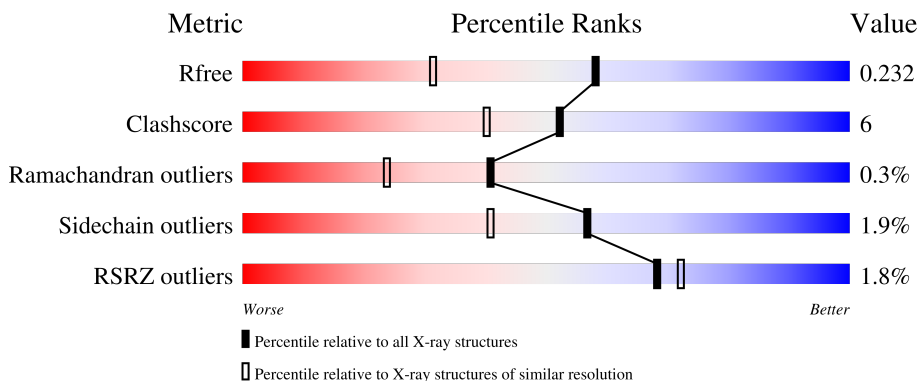
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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  $\geq 3$ , 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	310	 2% 84% 11% . .
1	B	310	 2% 81% 14% . 5%
1	C	310	 % 84% 11% 5%
1	D	310	 2% 78% 15% . 5%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9850 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 SAKe6DEref.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	299	2323	1432	431	454	6	0	6	0
1	B	295	2284	1409	429	440	6	0	1	0
1	C	296	2279	1407	424	442	6	0	0	0
1	D	293	2242	1387	414	435	6	0	0	0

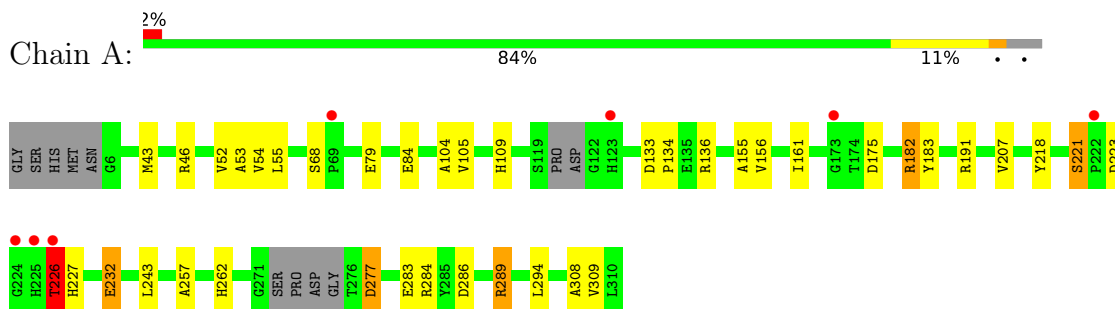
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	172	172	172	0	0
2	B	187	190	190	0	3
2	C	185	185	185	0	0
2	D	172	175	175	0	3

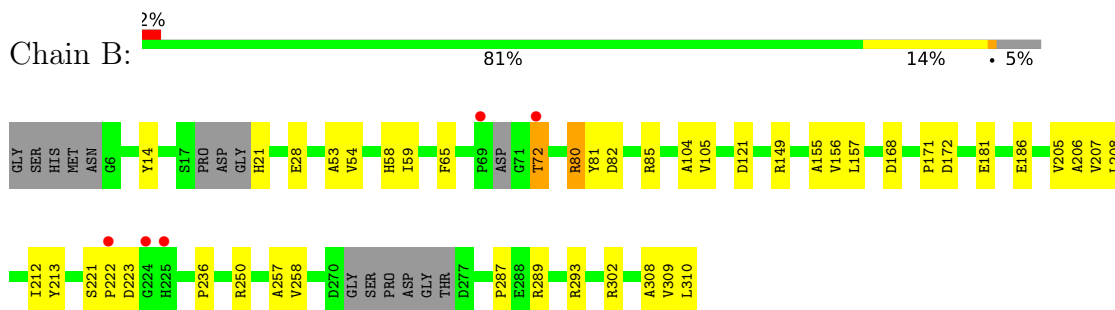
### 3 Residue-property plots [i](#)

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.

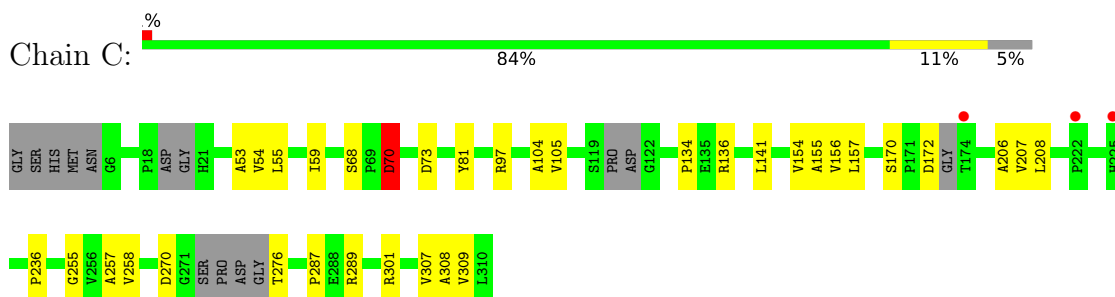
- Molecule 1: SAKe6DEref



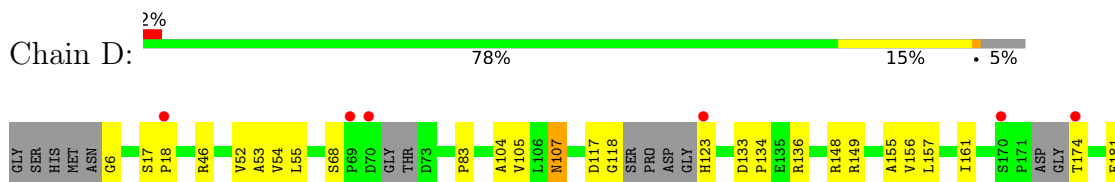
- Molecule 1: SAKe6DEref



- Molecule 1: SAKe6DEref



- Molecule 1: SAKe6DEref





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.53Å 58.53Å 78.71Å 90.00° 90.00° 82.15°	Depositor
Resolution (Å)	39.35 – 1.70 46.68 – 1.70	Depositor EDS
% Data completeness (in resolution range)	97.0 (39.35-1.70) 92.4 (46.68-1.70)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.28 (at 1.70Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.180 , 0.223 0.195 , 0.232	Depositor DCC
$R_{free}$ test set	5238 reflections (4.74%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.3	Xtriage
Anisotropy	0.244	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 37.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.458 for -h,-k,l 0.459 for -k,-h,-l 0.459 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	9850	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.37 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.1150e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

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	0.61	1/2384 (0.0%)	0.81	2/3249 (0.1%)
1	B	0.58	0/2345	0.76	0/3193
1	C	0.59	0/2338	0.77	0/3183
1	D	0.58	0/2299	0.76	1/3132 (0.0%)
All	All	0.59	1/9366 (0.0%)	0.77	3/12757 (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	A	0	4
1	B	0	2
1	C	0	1
1	D	0	2
All	All	0	9

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	232	GLU	CB-CG	-6.02	1.40	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	107	ASN	CB-CA-C	6.71	123.83	110.40
1	A	226[A]	THR	C-N-CA	-5.17	108.77	121.70
1	A	226[B]	THR	C-N-CA	-5.17	108.77	121.70

There are no chirality outliers.

All (9) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	221	SER	Mainchain
1	A	226[A]	THR	Mainchain
1	A	226[B]	THR	Mainchain
1	B	250	ARG	Sidechain
1	B	80	ARG	Sidechain
1	C	301	ARG	Sidechain
1	D	148	ARG	Sidechain
1	D	293	ARG	Sidechain

## 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	2323	0	2127	28	0
1	B	2284	0	2101	23	0
1	C	2279	0	2099	19	0
1	D	2242	0	2059	27	0
2	A	172	0	0	1	0
2	B	190	0	0	2	0
2	C	185	0	0	1	0
2	D	175	0	0	3	0
All	All	9850	0	8386	97	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 6.

All (97) 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:6:GLY:HA2	2:D:515:HOH:O	1.46	1.14
1:D:156:VAL:HG22	1:D:161:ILE:HD13	1.66	0.76
1:A:156:VAL:HG22	1:A:161:ILE:HD13	1.68	0.75
1:A:232:GLU:HG3	1:A:243:LEU:HD12	1.73	0.70
1:D:117:ASP:O	1:D:123:HIS:HA	1.93	0.68
1:D:118:GLY:C	2:D:493:HOH:O	2.31	0.68
1:D:283:GLU:HG2	1:D:294:LEU:HD13	1.74	0.68
1:C:55:LEU:HD21	1:C:134:PRO:HG3	1.77	0.66

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:133:ASP:OD1	1:A:136:ARG:HG3	1.97	0.64
1:A:55:LEU:HD21	1:A:134:PRO:HG3	1.79	0.64
1:A:283:GLU:HG2	1:A:294:LEU:HD13	1.79	0.64
1:D:133:ASP:OD2	1:D:136:ARG:HG3	1.99	0.63
1:B:14:TYR:OH	1:B:21:HIS:ND1	2.30	0.63
1:B:65:PHE:HZ	1:B:72:THR:HG22	1.64	0.62
1:D:262:HIS:HB3	1:D:284:ARG:HG3	1.83	0.60
1:C:257:ALA:HB1	1:C:309:VAL:HG23	1.85	0.58
1:B:82:ASP:OD2	1:B:85:ARG:HG3	2.04	0.57
1:D:55:LEU:HD21	1:D:134:PRO:HG3	1.86	0.57
1:C:136:ARG:NE	1:C:136:ARG:HA	2.20	0.57
1:C:136:ARG:HA	1:C:136:ARG:HE	1.70	0.56
1:A:286:ASP:OD2	1:A:289:ARG:HD2	2.06	0.56
1:D:83:PRO:HG3	1:D:310:LEU:HD21	1.87	0.55
1:D:208:LEU:HD21	1:D:287:PRO:HG3	1.87	0.55
1:A:218:TYR:HE1	1:A:226[A]:THR:HG1	1.54	0.54
1:A:161:ILE:HB	1:A:183:TYR:HB3	1.89	0.53
1:B:58:HIS:HB3	1:B:80:ARG:HG3	1.91	0.53
1:D:269:PHE:HZ	1:D:276:THR:HG23	1.73	0.53
1:A:43:MET:HB3	1:A:79:GLU:OE1	2.10	0.52
1:D:52:VAL:HG12	1:D:308:ALA:HB2	1.92	0.52
1:D:157:LEU:HD22	1:D:212:ILE:HG13	1.93	0.51
1:D:17:SER:HB2	1:D:18:PRO:HD2	1.92	0.51
1:D:54:VAL:HG23	1:D:308:ALA:HB1	1.93	0.51
1:C:73:ASP:OD2	1:C:73:ASP:N	2.43	0.50
1:A:182:ARG:HH12	1:A:191:ARG:NH2	2.10	0.49
1:C:104:ALA:HB1	1:C:156:VAL:HG23	1.94	0.49
1:C:155:ALA:HB1	1:C:207:VAL:HG23	1.95	0.49
1:A:221:SER:O	1:A:223[B]:ASP:N	2.39	0.48
1:B:155:ALA:HB1	1:B:207:VAL:HG23	1.96	0.48
1:D:155:ALA:HB1	1:D:207:VAL:HG23	1.95	0.48
1:D:104:ALA:HB1	1:D:156:VAL:HG23	1.95	0.48
1:A:156:VAL:HG22	1:A:161:ILE:CD1	2.40	0.48
1:A:104:ALA:HB1	1:A:156:VAL:HG23	1.96	0.47
1:C:104:ALA:HB2	1:C:154:VAL:HG12	1.96	0.47
1:D:53:ALA:HB1	1:D:105:VAL:HG23	1.97	0.47
1:A:262:HIS:CG	1:A:284:ARG:HD2	2.50	0.47
1:D:46:ARG:HD2	1:D:68:SER:HB3	1.96	0.46
1:A:155:ALA:HB1	1:A:207:VAL:HG23	1.97	0.46
1:A:54:VAL:HG23	1:A:308:ALA:HB1	1.98	0.45
1:A:109:HIS:HE1	2:A:557:HOH:O	1.99	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:157:LEU:HD21	1:D:236:PRO:HG3	1.98	0.45
1:A:43:MET:HB3	1:A:79:GLU:CD	2.37	0.45
1:B:104:ALA:HB1	1:B:156:VAL:HG23	1.97	0.45
1:C:54:VAL:HG23	1:C:308:ALA:HB1	1.99	0.45
1:B:257:ALA:HB1	1:B:309:VAL:HG23	1.99	0.45
1:D:257:ALA:HB1	1:D:309:VAL:HG23	1.98	0.45
1:B:53:ALA:HB1	1:B:105:VAL:HG23	1.98	0.44
1:C:206:ALA:HB1	1:C:258:VAL:HG23	1.99	0.44
1:D:269:PHE:CZ	1:D:276:THR:HG23	2.53	0.44
1:B:289:ARG:NH2	2:B:405:HOH:O	2.50	0.44
1:D:206:ALA:HB1	1:D:258:VAL:HG23	2.00	0.44
1:C:270:ASP:O	1:C:276:THR:HA	2.19	0.43
1:A:257:ALA:HB1	1:A:309:VAL:HG23	2.00	0.43
1:A:52:VAL:HG12	1:A:308:ALA:HB2	1.99	0.43
1:D:107:ASN:O	2:D:401:HOH:O	2.21	0.43
1:C:157:LEU:HD21	1:C:236:PRO:HG3	2.00	0.43
1:C:255:GLY:HA3	1:C:307:VAL:HG13	1.99	0.43
1:D:149:ARG:NH2	1:D:181:GLU:HG3	2.33	0.43
1:D:278:HIS:O	1:D:301:ARG:HD3	2.19	0.43
1:B:206:ALA:HB1	1:B:258:VAL:HG23	1.99	0.43
1:A:105:VAL:O	1:A:156:VAL:HG21	2.19	0.43
1:B:157:LEU:HD21	1:B:236:PRO:HG3	2.00	0.42
1:C:68:SER:O	1:C:70:ASP:N	2.53	0.42
1:C:59:ILE:HB	1:C:81:TYR:HB3	2.02	0.42
1:B:28:GLU:HG3	1:B:302:ARG:HH21	1.84	0.42
1:B:54:VAL:HG23	1:B:308:ALA:HB1	2.01	0.42
1:B:149:ARG:NH2	1:B:181:GLU:HG3	2.35	0.42
1:B:208:LEU:HD21	1:B:287:PRO:HG3	2.00	0.42
1:B:171:PRO:O	1:B:172:ASP:HB2	2.20	0.42
1:B:221:SER:O	1:B:222:PRO:C	2.58	0.42
1:B:59:ILE:HG13	1:B:310:LEU:HD22	2.03	0.41
1:A:53:ALA:HB1	1:A:105:VAL:HG23	2.02	0.41
1:D:221:SER:O	1:D:222:PRO:C	2.59	0.41
1:A:226[A]:THR:OG1	1:A:227:HIS:CD2	2.73	0.41
1:B:59:ILE:HB	1:B:81:TYR:HB3	2.01	0.41
1:B:80:ARG:HG2	1:B:80:ARG:HH11	1.85	0.41
1:C:97:ARG:HG2	1:C:97:ARG:HH11	1.86	0.41
1:A:262:HIS:HB3	1:A:284:ARG:HG2	2.02	0.41
1:C:208:LEU:HD21	1:C:287:PRO:HG3	2.02	0.41
1:A:277:ASP:OD1	1:A:277:ASP:N	2.54	0.41
1:B:186:GLU:OE2	1:B:186:GLU:HA	2.21	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:53:ALA:HB1	1:C:105:VAL:HG23	2.03	0.41
1:A:283:GLU:HG2	1:A:294:LEU:CD1	2.50	0.40
1:B:168:ASP:OD2	2:B:401:HOH:O	2.22	0.40
1:B:205:VAL:HA	1:B:213:TYR:O	2.21	0.40
1:A:46:ARG:HD2	1:A:68:SER:HB3	2.02	0.40
1:C:289:ARG:NH2	2:C:410:HOH:O	2.54	0.40

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	299/310 (96%)	282 (94%)	17 (6%)	0	100	100
1	B	288/310 (93%)	277 (96%)	10 (4%)	1 (0%)	41	24
1	C	286/310 (92%)	273 (96%)	12 (4%)	1 (0%)	41	24
1	D	283/310 (91%)	271 (96%)	11 (4%)	1 (0%)	34	18
All	All	1156/1240 (93%)	1103 (95%)	50 (4%)	3 (0%)	41	24

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	70	ASP
1	B	223	ASP
1	D	222	PRO

### 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	237/246 (96%)	229 (97%)	8 (3%)	37	18
1	B	235/246 (96%)	231 (98%)	4 (2%)	60	46
1	C	235/246 (96%)	231 (98%)	4 (2%)	60	46
1	D	229/246 (93%)	225 (98%)	4 (2%)	60	46
All	All	936/984 (95%)	916 (98%)	20 (2%)	57	36

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	84	GLU
1	A	175[A]	ASP
1	A	175[B]	ASP
1	A	182	ARG
1	A	226[A]	THR
1	A	226[B]	THR
1	A	277	ASP
1	A	289	ARG
1	B	72	THR
1	B	121	ASP
1	B	212	ILE
1	B	293	ARG
1	C	70	ASP
1	C	141	LEU
1	C	170	SER
1	C	172	ASP
1	D	174	THR
1	D	191	ARG
1	D	222	PRO
1	D	293	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	21	HIS
1	A	109	HIS
1	A	242	HIS
1	B	160	HIS
1	C	160	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	211	HIS
1	C	248	ASN

### 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](#)

There are no ligands in this entry.

### 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, 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	299/310 (96%)	-0.25	7 (2%) 60 65	19, 27, 62, 87	0
1	B	295/310 (95%)	-0.28	5 (1%) 70 74	19, 27, 58, 90	0
1	C	296/310 (95%)	-0.28	3 (1%) 82 85	18, 28, 67, 88	0
1	D	293/310 (94%)	-0.29	6 (2%) 65 69	19, 27, 57, 98	0
All	All	1183/1240 (95%)	-0.28	21 (1%) 68 72	18, 27, 62, 98	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	222	PRO	4.1
1	A	224[A]	GLY	4.0
1	A	69	PRO	3.8
1	A	225[A]	HIS	3.6
1	B	225	HIS	3.6
1	B	222	PRO	3.2
1	A	222[A]	PRO	3.1
1	D	69	PRO	3.0
1	D	123	HIS	3.0
1	D	18	PRO	2.7
1	A	226[A]	THR	2.6
1	D	70	ASP	2.6
1	A	123	HIS	2.5
1	C	225	HIS	2.3
1	B	224	GLY	2.3
1	D	170	SER	2.2
1	B	72	THR	2.2
1	C	174	THR	2.1
1	A	173	GLY	2.1
1	B	69	PRO	2.0
1	D	174	THR	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 monosaccharides in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.