



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2018 – 12:49 pm GMT

PDB ID : 2B9U  
Title : Crystal structure of dTDP-4-dehydrorhamnose 3,5-epimerase from *sulfolobus tokodaii*  
Authors : Rajakannan, V.; Kondo, K.; Mizushima, T.; Suzuki, A.; Yamane, T.  
Deposited on : 2005-10-13  
Resolution : 2.07 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
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

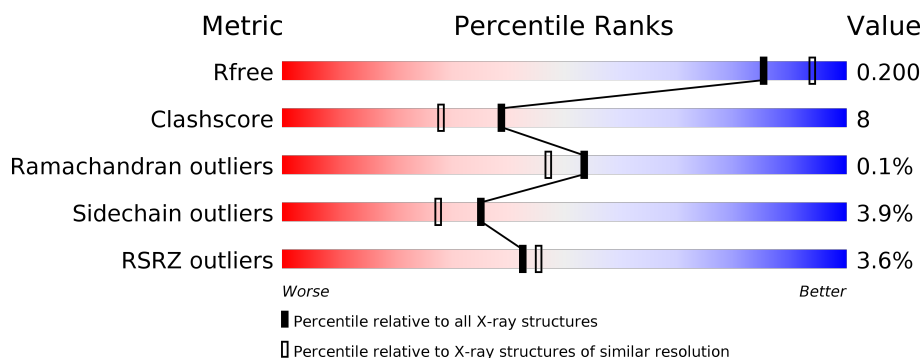
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.07 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	2250 (2.08-2.04)
Clashscore	122126	2358 (2.08-2.04)
Ramachandran outliers	120053	2339 (2.08-2.04)
Sidechain outliers	120020	2339 (2.08-2.04)
RSRZ outliers	108989	2211 (2.08-2.04)

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	176	<div> <div>3%</div> <div>85%</div> <div>15%</div> </div>
1	B	176	<div> <div>%</div> <div>82%</div> <div>17%</div> <div>.</div> </div>
1	C	176	<div> <div>%</div> <div>77%</div> <div>22%</div> <div>..</div> </div>
1	D	176	<div> <div>3%</div> <div>76%</div> <div>21%</div> <div>.</div> </div>
1	E	176	<div> <div>%</div> <div>83%</div> <div>16%</div> <div>.</div> </div>
1	F	176	<div> <div>3%</div> <div>84%</div> <div>14%</div> <div>.</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	176	<div><div></div><div>9%</div><div>85%</div><div>14%</div><div></div></div>
1	H	176	<div><div></div><div>%</div><div>81%</div><div>17%</div><div></div></div>
1	I	176	<div><div></div><div>4%</div><div>80%</div><div>20%</div><div></div></div>
1	J	176	<div><div></div><div>6%</div><div>79%</div><div>20%</div><div></div></div>
1	K	176	<div><div></div><div>6%</div><div>84%</div><div>14%</div><div></div></div>
1	L	176	<div><div></div><div>5%</div><div>84%</div><div>15%</div><div></div></div>

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 19111 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 hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			
1	C	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			
1	D	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			
1	B	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			
1	E	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			
1	F	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			
1	G	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			
1	H	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			
1	I	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			
1	J	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			
1	K	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			
1	L	176	Total	C	N	O	S	0	0	0
			1447	945	238	257	7			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	148	Total	O	0	0
			148	148		
2	C	168	Total	O	0	0
			168	168		

*Continued on next page...*

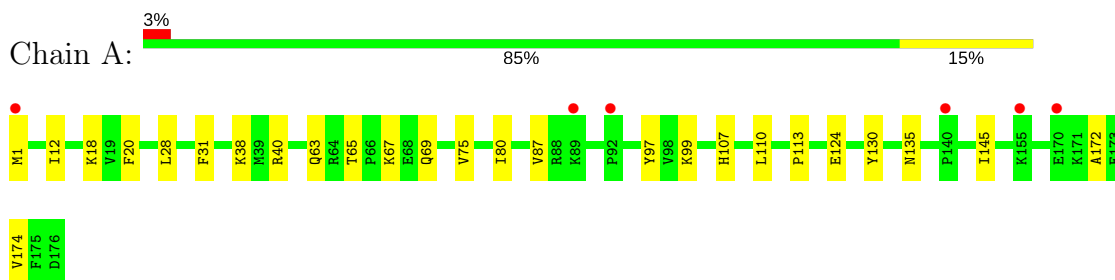
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	151	Total 151	O 151	0	0
2	B	152	Total 152	O 152	0	0
2	E	165	Total 165	O 165	0	0
2	F	140	Total 140	O 140	0	0
2	G	105	Total 105	O 105	0	0
2	H	170	Total 170	O 170	0	0
2	I	134	Total 134	O 134	0	0
2	J	124	Total 124	O 124	0	0
2	K	143	Total 143	O 143	0	0
2	L	147	Total 147	O 147	0	0

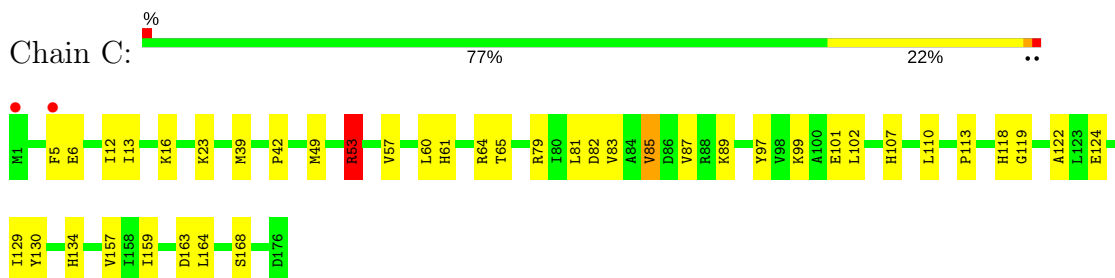
### 3 Residue-property plots [i](#)

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.

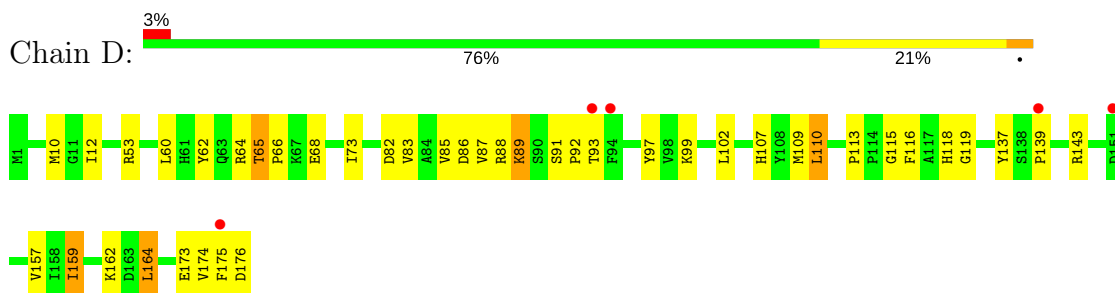
- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase



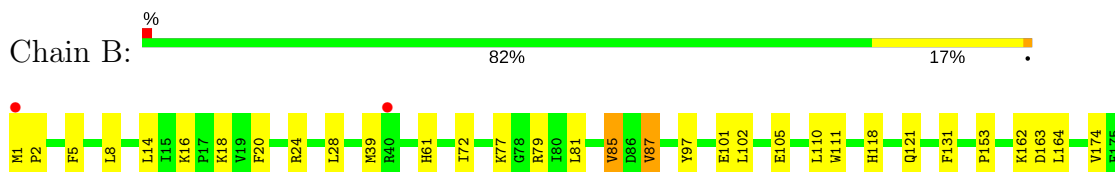
- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase



- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase




- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase



D176


- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase

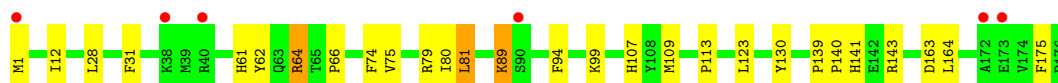
Chain E:  83% 16%




D176

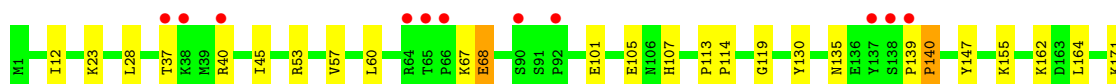
- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase

Chain F:  84% 14% 3%




- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase

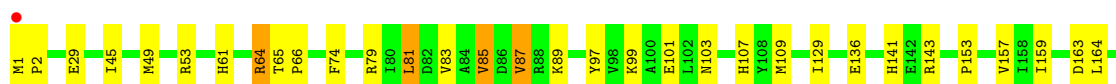
Chain G:  85% 14% 9%



A172  
E173  
V174  
F175  
D176

- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase

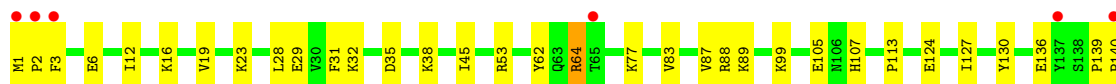
Chain H:  81% 17% 2%

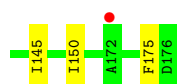


V174  
F175  
D176

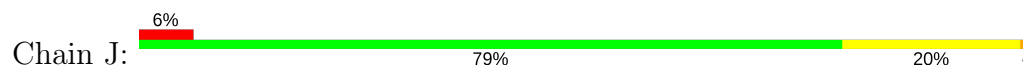
- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase

Chain I:  80% 20% 4%

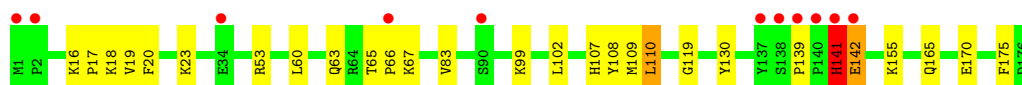
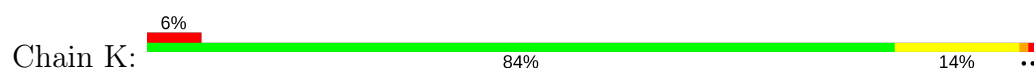




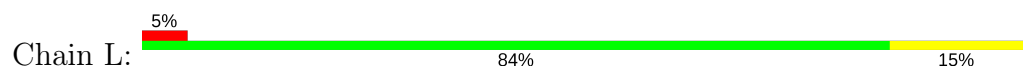
- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase



- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase



- Molecule 1: hypothetical dTDP-4-dehydrorhamnose 3,5-epimerase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.79Å 229.50Å 74.58Å 90.00° 96.95° 90.00°	Depositor
Resolution (Å)	20.00 – 2.07 19.86 – 2.07	Depositor EDS
% Data completeness (in resolution range)	99.5 (20.00-2.07) 99.5 (19.86-2.07)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.09 (at 2.07Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.203 , 0.257 0.208 , 0.200	Depositor DCC
$R_{free}$ test set	14888 reflections (9.89%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.9	Xtriage
Anisotropy	0.032	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 47.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.019 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	19111	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.08% 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

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	0.54	0/1489	0.66	0/2010
1	B	0.61	0/1489	0.67	0/2010
1	C	0.62	0/1489	0.73	2/2010 (0.1%)
1	D	0.60	0/1489	0.68	0/2010
1	E	0.59	0/1489	0.66	0/2010
1	F	0.57	0/1489	0.65	0/2010
1	G	0.54	0/1489	0.61	0/2010
1	H	0.61	0/1489	0.70	0/2010
1	I	0.59	0/1489	0.68	0/2010
1	J	0.56	0/1489	0.66	0/2010
1	K	0.55	0/1489	0.70	1/2010 (0.0%)
1	L	0.56	0/1489	0.67	0/2010
All	All	0.58	0/17868	0.67	3/24120 (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	K	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	53	ARG	NE-CZ-NH2	-7.66	116.47	120.30
1	C	53	ARG	NE-CZ-NH1	5.57	123.09	120.30
1	K	141	HIS	N-CA-C	5.51	125.88	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	K	141	HIS	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	1447	0	1444	12	0
1	B	1447	0	1444	18	0
1	C	1447	0	1444	25	0
1	D	1447	0	1444	43	0
1	E	1447	0	1444	31	0
1	F	1447	0	1444	22	0
1	G	1447	0	1444	17	0
1	H	1447	0	1444	26	0
1	I	1447	0	1444	29	0
1	J	1447	0	1444	29	0
1	K	1447	0	1444	18	0
1	L	1447	0	1444	26	0
2	A	148	0	0	2	2
2	B	152	0	0	2	0
2	C	168	0	0	5	0
2	D	151	0	0	8	0
2	E	165	0	0	8	0
2	F	140	0	0	3	0
2	G	105	0	0	7	1
2	H	170	0	0	1	0
2	I	134	0	0	3	0
2	J	124	0	0	3	0
2	K	143	0	0	0	0
2	L	147	0	0	7	1
All	All	19111	0	17328	280	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 8.

The worst 5 of 280 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:E:157:VAL:CG1	1:E:159:ILE:HD11	1.72	1.20
1:E:49:MET:HE2	1:F:74:PHE:HZ	1.16	1.10
1:E:53:ARG:CD	2:E:208:HOH:O	1.99	1.08
1:E:53:ARG:HD2	2:E:208:HOH:O	1.53	1.03
1:I:3:PHE:CZ	1:I:31:PHE:HA	1.95	1.01

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)
2:A:197:HOH:O	2:L:286:HOH:O[1_655]	1.98	0.22
2:A:318:HOH:O	2:G:258:HOH:O[1_655]	2.00	0.20

## 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	174/176 (99%)	171 (98%)	3 (2%)	0	100	100
1	B	174/176 (99%)	166 (95%)	8 (5%)	0	100	100
1	C	174/176 (99%)	169 (97%)	5 (3%)	0	100	100
1	D	174/176 (99%)	165 (95%)	9 (5%)	0	100	100
1	E	174/176 (99%)	166 (95%)	8 (5%)	0	100	100
1	F	174/176 (99%)	169 (97%)	5 (3%)	0	100	100
1	G	174/176 (99%)	166 (95%)	7 (4%)	1 (1%)	27	16
1	H	174/176 (99%)	168 (97%)	6 (3%)	0	100	100
1	I	174/176 (99%)	169 (97%)	5 (3%)	0	100	100
1	J	174/176 (99%)	167 (96%)	7 (4%)	0	100	100
1	K	174/176 (99%)	166 (95%)	7 (4%)	1 (1%)	27	16
1	L	174/176 (99%)	166 (95%)	7 (4%)	1 (1%)	27	16

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	2088/2112 (99%)	2008 (96%)	77 (4%)	3 (0%)	53 47

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	K	142	GLU
1	L	140	PRO
1	G	140	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	161/161 (100%)	155 (96%)	6 (4%)	37 30
1	B	161/161 (100%)	154 (96%)	7 (4%)	32 24
1	C	161/161 (100%)	151 (94%)	10 (6%)	20 11
1	D	161/161 (100%)	156 (97%)	5 (3%)	43 37
1	E	161/161 (100%)	153 (95%)	8 (5%)	27 18
1	F	161/161 (100%)	154 (96%)	7 (4%)	32 24
1	G	161/161 (100%)	154 (96%)	7 (4%)	32 24
1	H	161/161 (100%)	155 (96%)	6 (4%)	37 30
1	I	161/161 (100%)	157 (98%)	4 (2%)	50 45
1	J	161/161 (100%)	156 (97%)	5 (3%)	43 37
1	K	161/161 (100%)	155 (96%)	6 (4%)	37 30
1	L	161/161 (100%)	156 (97%)	5 (3%)	43 37
All	All	1932/1932 (100%)	1856 (96%)	76 (4%)	35 27

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	164	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	164	LEU
1	K	175	PHE
1	E	171	LYS
1	F	64	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	43	ASN
1	H	61	HIS
1	L	7	ASN
1	G	69	GLN
1	G	107	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](#)

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 [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	176/176 (100%)	0.05	6 (3%) 45 48	18, 29, 49, 59	0
1	B	176/176 (100%)	-0.07	2 (1%) 80 82	16, 26, 40, 46	0
1	C	176/176 (100%)	-0.02	2 (1%) 80 82	15, 26, 41, 60	0
1	D	176/176 (100%)	0.16	5 (2%) 53 57	20, 31, 48, 61	0
1	E	176/176 (100%)	-0.07	1 (0%) 89 90	18, 27, 37, 48	0
1	F	176/176 (100%)	0.08	6 (3%) 45 48	18, 29, 44, 54	0
1	G	176/176 (100%)	0.30	15 (8%) 11 11	16, 33, 57, 71	0
1	H	176/176 (100%)	-0.09	1 (0%) 89 90	17, 28, 39, 52	0
1	I	176/176 (100%)	0.20	7 (3%) 38 40	19, 30, 48, 64	0
1	J	176/176 (100%)	0.39	11 (6%) 20 20	21, 33, 54, 64	0
1	K	176/176 (100%)	0.16	11 (6%) 20 20	22, 31, 51, 63	0
1	L	176/176 (100%)	0.09	8 (4%) 33 34	22, 29, 48, 64	0
All	All	2112/2112 (100%)	0.10	75 (3%) 42 45	15, 29, 48, 71	0

The worst 5 of 75 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	K	1	MET	9.0
1	K	141	HIS	7.4
1	L	1	MET	7.2
1	C	1	MET	6.9
1	K	140	PRO	6.7

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.