



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:00 AM GMT

PDB ID : 3CW2  
Title : Crystal structure of the intact archaeal translation initiation factor 2 from *Sulfolobus solfataricus* .  
Authors : Stolboushkina, E.A.; Nikonov, S.V.; Nikulin, A.D; Blaesi, U; Manstein, D.J.; Fedorov, R.V.; Garber, M.B.; Nikonov, O.S.  
Deposited on : 2008-04-21  
Resolution : 2.80 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865



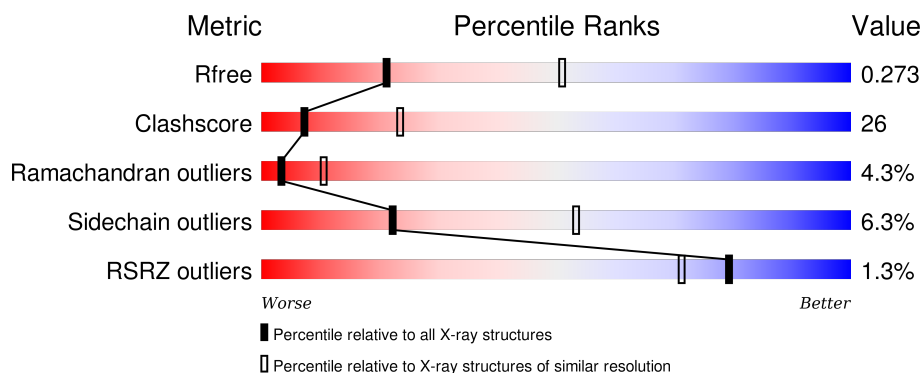
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
$R_{free}$	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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 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	415	<div> <div>53%</div> <div>42%</div> <div>5%</div> </div>
1	B	415	<div> <div>51%</div> <div>43%</div> <div>• •</div> </div>
1	E	415	<div> <div>51%</div> <div>45%</div> <div>•</div> </div>
1	F	415	<div> <div>44%</div> <div>50%</div> <div>5%</div> </div>
2	C	266	<div> <div>53%</div> <div>38%</div> <div>8%</div> <div>•</div> </div>

*Continued on next page...*



Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	266	<div><div></div><div>55%40%5%</div></div>
2	G	266	<div><div></div><div>54%40%6%</div></div>
2	H	266	<div><div></div><div>50%41%8%•</div></div>
3	K	139	<div><div>6%</div><div></div><div>42%37%9%•8%</div></div>
3	L	139	<div><div>7%8%</div><div></div><div>85%</div></div>
3	M	139	<div><div>7%</div><div></div><div>52%36%9%••</div></div>
3	N	139	<div><div>%</div><div></div><div>11%9%•78%</div></div>



## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 23823 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 Translation initiation factor 2 subunit gamma.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	414	Total	C	N	O	S	0	0	0
			3213	2058	548	595	12			
1	B	400	Total	C	N	O	S	0	0	0
			3096	1987	524	574	11			
1	E	414	Total	C	N	O	S	0	0	0
			3213	2058	548	595	12			
1	F	414	Total	C	N	O	S	0	0	0
			3213	2058	548	595	12			

- Molecule 2 is a protein called Translation initiation factor 2 subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	266	Total	C	N	O	S	0	0	0
			2134	1364	365	403	2			
2	D	266	Total	C	N	O	S	0	0	0
			2134	1364	365	403	2			
2	G	266	Total	C	N	O	S	0	0	0
			2134	1364	365	403	2			
2	H	266	Total	C	N	O	S	0	0	0
			2134	1364	365	403	2			

- Molecule 3 is a protein called Translation initiation factor 2 subunit beta.

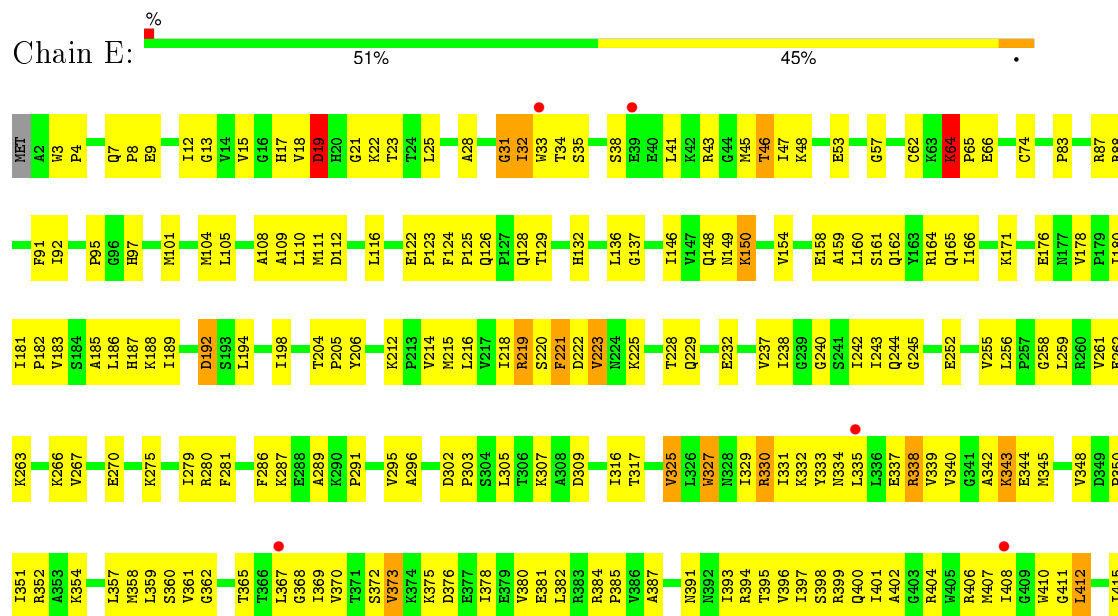
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	K	128	Total	C	N	O	S	0	0	0
			1032	655	174	194	9			
3	L	21	Total	C	N	O	S	0	0	0
			174	110	27	36	1			
3	M	138	Total	C	N	O	S	0	0	0
			1102	699	186	207	10			
3	N	30	Total	C	N	O	S	0	0	0
			244	152	41	50	1			



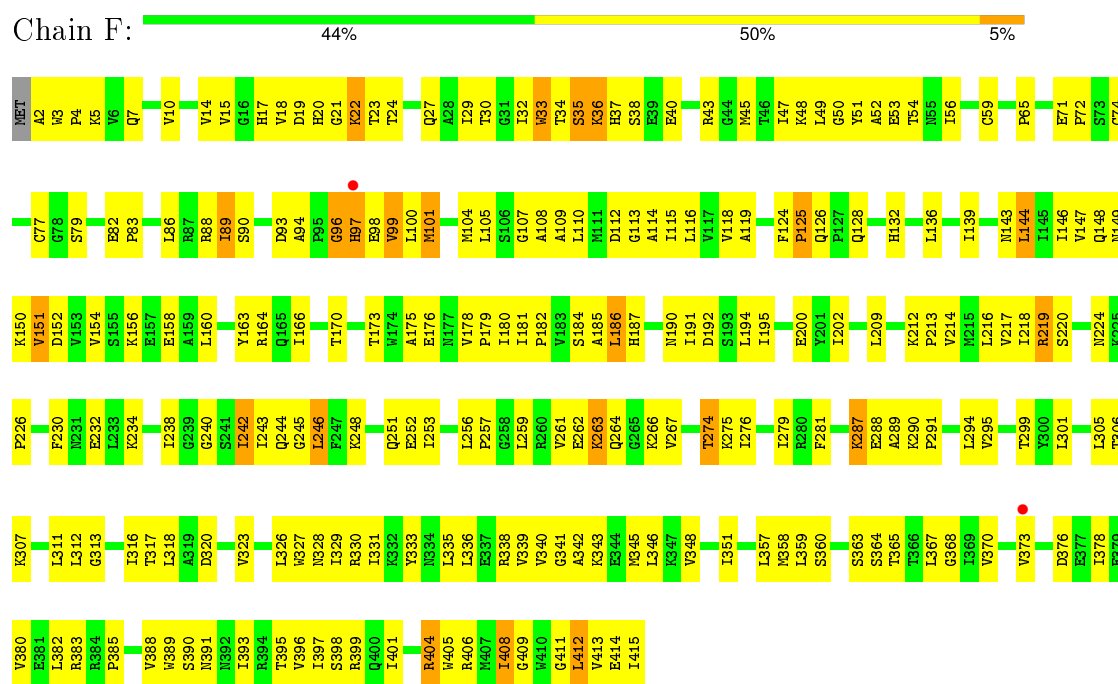




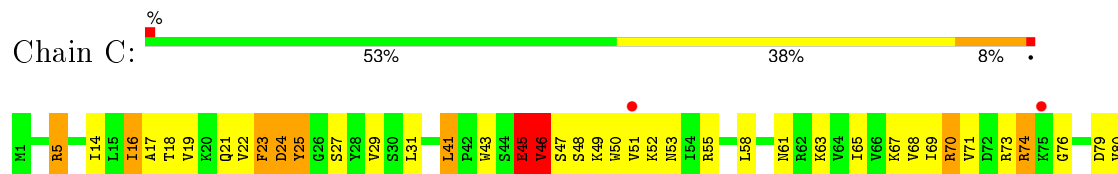
• Molecule 1: Translation initiation factor 2 subunit gamma



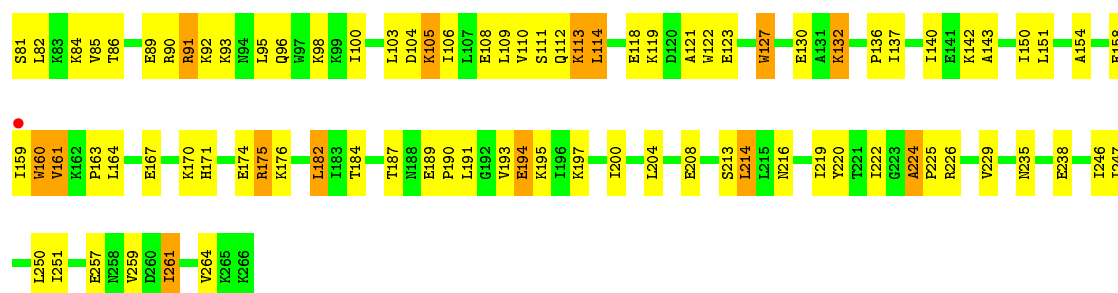
• Molecule 1: Translation initiation factor 2 subunit gamma



• Molecule 2: Translation initiation factor 2 subunit alpha

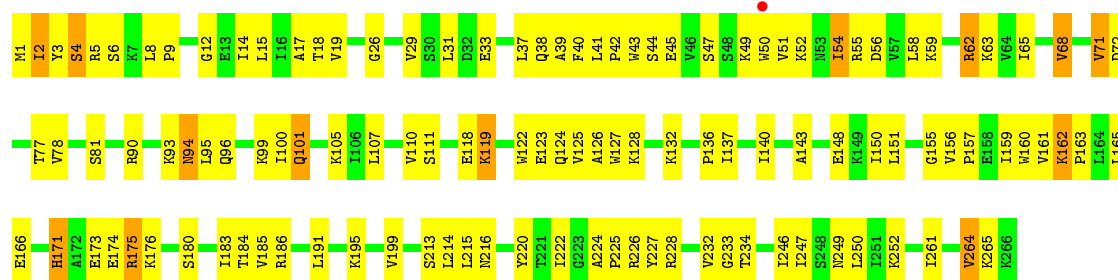






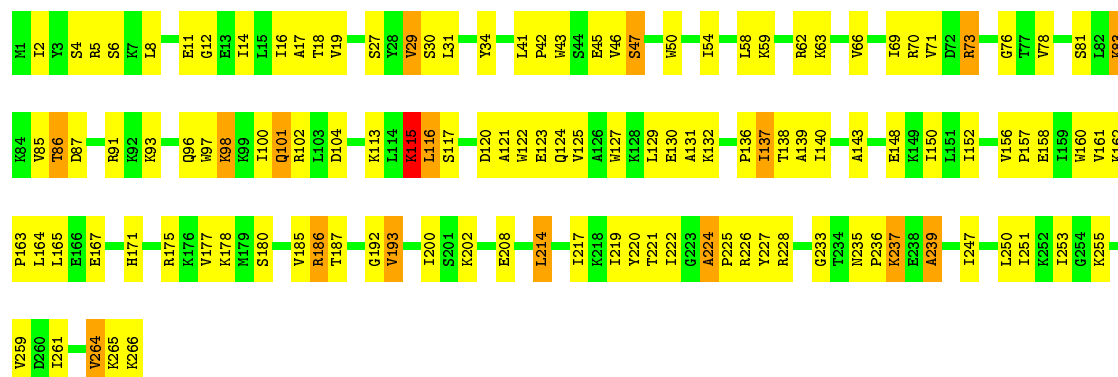
• Molecule 2: Translation initiation factor 2 subunit alpha

Chain D: 55% 40% 5%



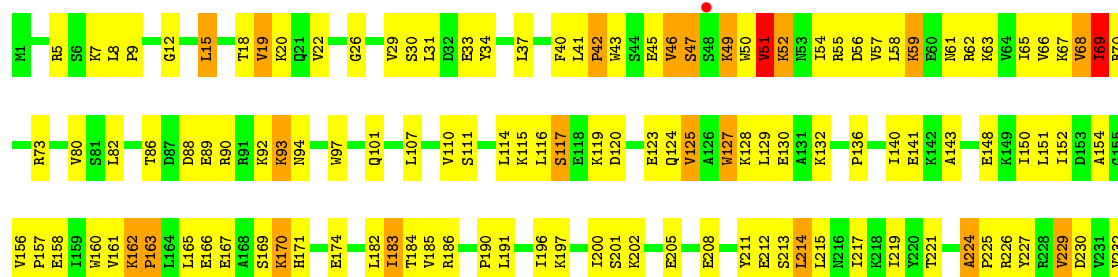
• Molecule 2: Translation initiation factor 2 subunit alpha

Chain G: 54% 40% 6%

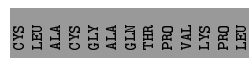


• Molecule 2: Translation initiation factor 2 subunit alpha

Chain H: 50% 41% 8%









## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.20 Å 162.92 Å 161.28 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.91 – 2.80 19.91 – 2.80	Depositor EDS
% Data completeness (in resolution range)	92.0 (19.91-2.80) 98.5 (19.91-2.80)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.00 (at 2.79 Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.225 , 0.276 0.229 , 0.273	Depositor DCC
$R_{free}$ test set	4940 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	72.0	Xtriage
Anisotropy	0.470	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 49.6	EDS
Estimated twinning fraction	0.000 for -h,-l,-k 0.000 for -h,l,k 0.449 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 98811 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	23823	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	106.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.72% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.23	0/3272	0.42	0/4430
1	B	0.22	0/3152	0.41	0/4269
1	E	0.23	0/3272	0.42	0/4430
1	F	0.24	0/3272	0.41	0/4430
2	C	0.22	0/2164	0.39	0/2914
2	D	0.22	0/2164	0.39	0/2914
2	G	0.23	0/2164	0.40	0/2914
2	H	0.21	0/2164	0.38	0/2914
3	K	0.23	0/1045	0.39	0/1399
3	L	0.22	0/176	0.38	0/233
3	M	0.23	0/1117	0.47	0/1497
3	N	0.23	0/247	0.38	0/328
All	All	0.23	0/24209	0.41	0/32672

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	1
1	F	0	1
2	C	0	3
2	D	0	1
3	K	0	10
3	M	0	3
All	All	0	19

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 19 planarity outliers are listed below:



Mol	Chain	Res	Type	Group
1	A	338	ARG	Peptide
2	C	158	GLU	Peptide
2	C	45	GLU	Peptide
2	C	46	VAL	Peptide
2	D	2	ILE	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	3213	0	3332	159	0
1	B	3096	0	3207	160	0
1	E	3213	0	3332	203	0
1	F	3213	0	3334	197	0
2	C	2134	0	2242	96	0
2	D	2134	0	2242	101	0
2	G	2134	0	2242	109	0
2	H	2134	0	2242	106	0
3	K	1032	0	1073	59	0
3	L	174	0	173	10	0
3	M	1102	0	1148	89	0
3	N	244	0	246	12	0
All	All	23823	0	24813	1261	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 26.

The worst 5 of 1261 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)
2:D:5:ARG:HA	2:D:127:TRP:HE1	1.09	1.13
3:M:109:CYS:SG	3:M:127:CYS:HA	1.91	1.11
2:C:48:SER:HB2	2:C:50:TRP:HD1	1.18	1.08
3:M:111:SER:HB3	3:M:112:LEU:CG	1.84	1.07
1:E:330:ARG:HH11	1:E:330:ARG:HG3	1.16	1.06

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	412/415 (99%)	312 (76%)	82 (20%)	18 (4%)	3	10
1	B	394/415 (95%)	322 (82%)	67 (17%)	5 (1%)	15	44
1	E	412/415 (99%)	325 (79%)	73 (18%)	14 (3%)	5	16
1	F	412/415 (99%)	323 (78%)	77 (19%)	12 (3%)	6	19
2	C	264/266 (99%)	212 (80%)	40 (15%)	12 (4%)	3	10
2	D	264/266 (99%)	224 (85%)	34 (13%)	6 (2%)	8	26
2	G	264/266 (99%)	219 (83%)	34 (13%)	11 (4%)	3	11
2	H	264/266 (99%)	202 (76%)	42 (16%)	20 (8%)	1	2
3	K	126/139 (91%)	74 (59%)	41 (32%)	11 (9%)	1	2
3	L	19/139 (14%)	12 (63%)	6 (32%)	1 (5%)	2	7
3	M	136/139 (98%)	85 (62%)	36 (26%)	15 (11%)	0	1
3	N	28/139 (20%)	20 (71%)	5 (18%)	3 (11%)	0	1
All	All	2995/3280 (91%)	2330 (78%)	537 (18%)	128 (4%)	3	10

5 of 128 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	101	MET
1	A	307	LYS
1	B	102	ALA
2	C	24	ASP
2	D	54	ILE

### 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	356/357 (100%)	335 (94%)	21 (6%)	24	57
1	B	343/357 (96%)	330 (96%)	13 (4%)	40	74
1	E	356/357 (100%)	338 (95%)	18 (5%)	29	63
1	F	356/357 (100%)	328 (92%)	28 (8%)	15	40
2	C	237/239 (99%)	215 (91%)	22 (9%)	11	32
2	D	237/239 (99%)	223 (94%)	14 (6%)	24	57
2	G	237/239 (99%)	222 (94%)	15 (6%)	22	53
2	H	237/239 (99%)	219 (92%)	18 (8%)	16	42
3	K	117/126 (93%)	108 (92%)	9 (8%)	16	41
3	L	20/126 (16%)	20 (100%)	0	100	100
3	M	125/126 (99%)	119 (95%)	6 (5%)	31	66
3	N	28/126 (22%)	26 (93%)	2 (7%)	18	46
All	All	2649/2888 (92%)	2483 (94%)	166 (6%)	22	53

5 of 166 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	K	105	GLU
1	E	338	ARG
2	H	191	LEU
3	K	112	LEU
1	E	171	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 68 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	216	ASN
1	E	162	GLN
2	G	249	ASN
2	D	244	ASN
3	K	46	ASN

### 5.3.3 RNA ⓘ

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	414/415 (99%)	-0.27	4 (0%) 84 77	47, 105, 152, 162	0
1	B	400/415 (96%)	-0.29	4 (1%) 84 77	37, 109, 149, 173	0
1	E	414/415 (99%)	-0.28	5 (1%) 81 73	46, 106, 150, 165	0
1	F	414/415 (99%)	-0.31	2 (0%) 91 88	47, 107, 147, 163	0
2	C	266/266 (100%)	-0.38	3 (1%) 82 74	42, 95, 133, 151	0
2	D	266/266 (100%)	-0.37	1 (0%) 93 90	38, 98, 135, 169	0
2	G	266/266 (100%)	-0.44	0 100 100	41, 90, 120, 138	0
2	H	266/266 (100%)	-0.31	1 (0%) 93 90	36, 102, 135, 146	0
3	K	128/139 (92%)	0.32	8 (6%) 23 14	124, 148, 168, 175	0
3	L	21/139 (15%)	-0.24	0 100 100	89, 119, 133, 134	0
3	M	138/139 (99%)	0.23	10 (7%) 18 10	108, 147, 172, 181	0
3	N	30/139 (21%)	0.01	1 (3%) 50 38	100, 139, 154, 155	0
All	All	3023/3280 (92%)	-0.27	39 (1%) 79 71	36, 106, 153, 181	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	336	LEU	7.4
3	M	128	LEU	6.6
3	K	36	LEU	4.8
3	N	28	GLN	4.6
3	K	126	VAL	4.0

### 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](#)

There are no ligands in this entry.

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

There are no such residues in this entry.