



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 10:55 pm BST

PDB ID : 5DBO  
Title : Crystal structure of the tetrameric eIF2B-beta2-delta2 complex from *C. thermophilum*  
Authors : Kuhle, B.; Ficner, R.  
Deposited on : 2015-08-21  
Resolution : 3.00 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 : 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

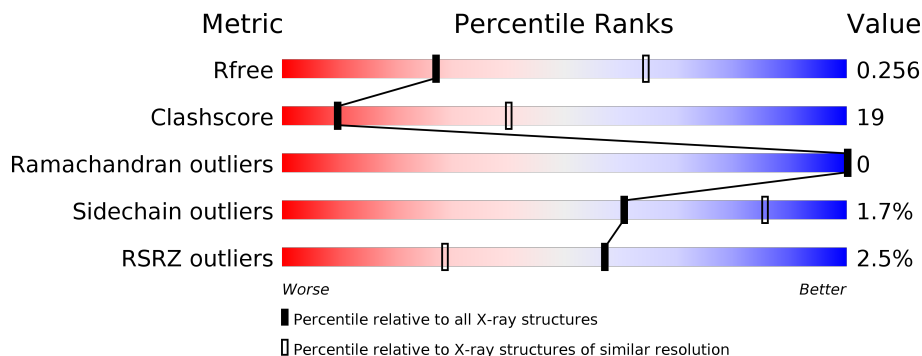
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	466	
1	C	466	
2	B	419	
2	D	419	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 9388 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 eif-2b-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	315	2451	1551	434	456	10	0	1	0
1	C	312	2414	1527	427	450	10	0	0	0

- Molecule 2 is a protein called Translation initiation factor eIF2b-like protein, Translation initiation factor eIF2b-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	303	2293	1453	402	432	6	0	0	0
2	D	296	2230	1410	392	422	6	0	0	0

There are 112 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	93	VAL	-	linker	UNP G0SEE6
B	94	ARG	-	linker	UNP G0SEE6
B	95	ARG	-	linker	UNP G0SEE6
B	96	VAL	-	linker	UNP G0SEE6
B	97	LEU	-	linker	UNP G0SEE6
B	98	GLY	-	linker	UNP G0SEE6
B	99	LEU	-	linker	UNP G0SEE6
B	100	ILE	-	linker	UNP G0SEE6
B	101	ARG	-	linker	UNP G0SEE6
B	102	ASP	-	linker	UNP G0SEE6
B	103	GLU	-	linker	UNP G0SEE6
B	104	ALA	-	linker	UNP G0SEE6
B	105	SER	-	linker	UNP G0SEE6
B	106	GLU	-	linker	UNP G0SEE6
B	107	ASN	-	linker	UNP G0SEE6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	108	ARG	-	linker	UNP G0SEE6
B	109	ASN	-	linker	UNP G0SEE6
B	110	ALA	-	linker	UNP G0SEE6
B	111	ASP	-	linker	UNP G0SEE6
B	112	ASP	-	linker	UNP G0SEE6
B	113	ILE	-	linker	UNP G0SEE6
B	114	ALA	-	linker	UNP G0SEE6
B	115	SER	-	linker	UNP G0SEE6
B	116	ASP	-	linker	UNP G0SEE6
B	117	ALA	-	linker	UNP G0SEE6
B	118	ALA	-	linker	UNP G0SEE6
B	119	SER	-	linker	UNP G0SEE6
B	120	ASP	-	linker	UNP G0SEE6
B	121	ILE	-	linker	UNP G0SEE6
B	122	GLN	-	linker	UNP G0SEE6
B	123	SER	-	linker	UNP G0SEE6
B	124	LEU	-	linker	UNP G0SEE6
B	125	ALA	-	linker	UNP G0SEE6
B	126	PRO	-	linker	UNP G0SEE6
B	127	SER	-	linker	UNP G0SEE6
B	128	GLN	-	linker	UNP G0SEE6
B	129	PRO	-	linker	UNP G0SEE6
B	130	PRO	-	linker	UNP G0SEE6
B	131	PRO	-	linker	UNP G0SEE6
B	132	GLN	-	linker	UNP G0SEE6
B	133	GLN	-	linker	UNP G0SEE6
B	134	ARG	-	linker	UNP G0SEE6
B	135	PRO	-	linker	UNP G0SEE6
B	136	PRO	-	linker	UNP G0SEE6
B	137	PRO	-	linker	UNP G0SEE6
B	138	PRO	-	linker	UNP G0SEE6
B	139	ALA	-	linker	UNP G0SEE6
B	140	ARG	-	linker	UNP G0SEE6
B	141	THR	-	linker	UNP G0SEE6
B	142	LEU	-	linker	UNP G0SEE6
B	143	THR	-	linker	UNP G0SEE6
B	144	SER	-	linker	UNP G0SEE6
B	145	GLY	-	linker	UNP G0SEE6
B	146	LEU	-	linker	UNP G0SEE6
B	147	GLN	-	linker	UNP G0SEE6
B	148	VAL	-	linker	UNP G0SEE6
D	93	VAL	-	linker	UNP G0SEE6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	94	ARG	-	linker	UNP G0SEE6
D	95	ARG	-	linker	UNP G0SEE6
D	96	VAL	-	linker	UNP G0SEE6
D	97	LEU	-	linker	UNP G0SEE6
D	98	GLY	-	linker	UNP G0SEE6
D	99	LEU	-	linker	UNP G0SEE6
D	100	ILE	-	linker	UNP G0SEE6
D	101	ARG	-	linker	UNP G0SEE6
D	102	ASP	-	linker	UNP G0SEE6
D	103	GLU	-	linker	UNP G0SEE6
D	104	ALA	-	linker	UNP G0SEE6
D	105	SER	-	linker	UNP G0SEE6
D	106	GLU	-	linker	UNP G0SEE6
D	107	ASN	-	linker	UNP G0SEE6
D	108	ARG	-	linker	UNP G0SEE6
D	109	ASN	-	linker	UNP G0SEE6
D	110	ALA	-	linker	UNP G0SEE6
D	111	ASP	-	linker	UNP G0SEE6
D	112	ASP	-	linker	UNP G0SEE6
D	113	ILE	-	linker	UNP G0SEE6
D	114	ALA	-	linker	UNP G0SEE6
D	115	SER	-	linker	UNP G0SEE6
D	116	ASP	-	linker	UNP G0SEE6
D	117	ALA	-	linker	UNP G0SEE6
D	118	ALA	-	linker	UNP G0SEE6
D	119	SER	-	linker	UNP G0SEE6
D	120	ASP	-	linker	UNP G0SEE6
D	121	ILE	-	linker	UNP G0SEE6
D	122	GLN	-	linker	UNP G0SEE6
D	123	SER	-	linker	UNP G0SEE6
D	124	LEU	-	linker	UNP G0SEE6
D	125	ALA	-	linker	UNP G0SEE6
D	126	PRO	-	linker	UNP G0SEE6
D	127	SER	-	linker	UNP G0SEE6
D	128	GLN	-	linker	UNP G0SEE6
D	129	PRO	-	linker	UNP G0SEE6
D	130	PRO	-	linker	UNP G0SEE6
D	131	PRO	-	linker	UNP G0SEE6
D	132	GLN	-	linker	UNP G0SEE6
D	133	GLN	-	linker	UNP G0SEE6
D	134	ARG	-	linker	UNP G0SEE6
D	135	PRO	-	linker	UNP G0SEE6

*Continued on next page...*

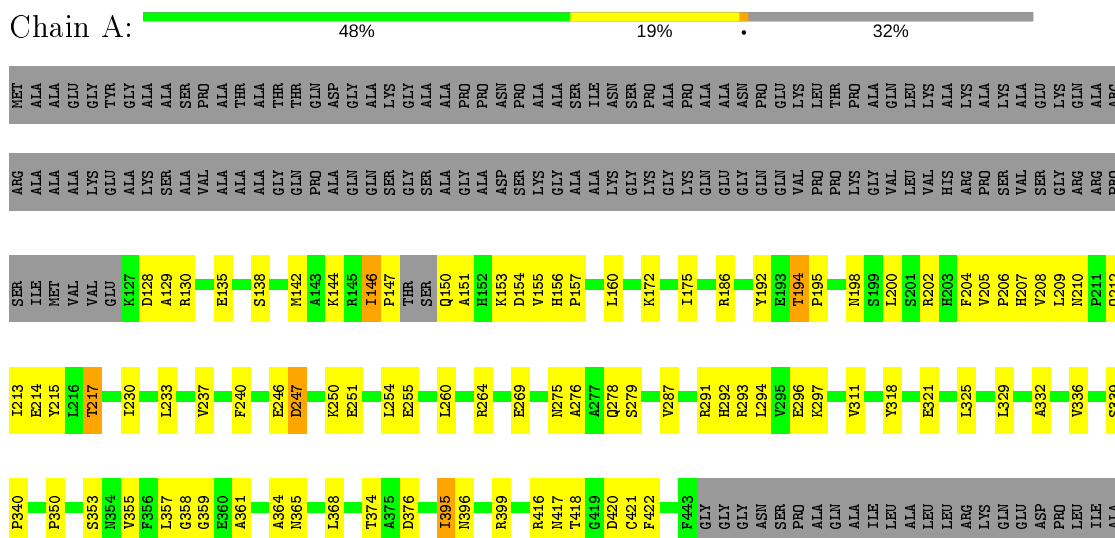
*Continued from previous page...*

<b>Chain</b>	<b>Residue</b>	<b>Modelled</b>	<b>Actual</b>	<b>Comment</b>	<b>Reference</b>
D	136	PRO	-	linker	UNP G0SEE6
D	137	PRO	-	linker	UNP G0SEE6
D	138	PRO	-	linker	UNP G0SEE6
D	139	ALA	-	linker	UNP G0SEE6
D	140	ARG	-	linker	UNP G0SEE6
D	141	THR	-	linker	UNP G0SEE6
D	142	LEU	-	linker	UNP G0SEE6
D	143	THR	-	linker	UNP G0SEE6
D	144	SER	-	linker	UNP G0SEE6
D	145	GLY	-	linker	UNP G0SEE6
D	146	LEU	-	linker	UNP G0SEE6
D	147	GLN	-	linker	UNP G0SEE6
D	148	VAL	-	linker	UNP G0SEE6

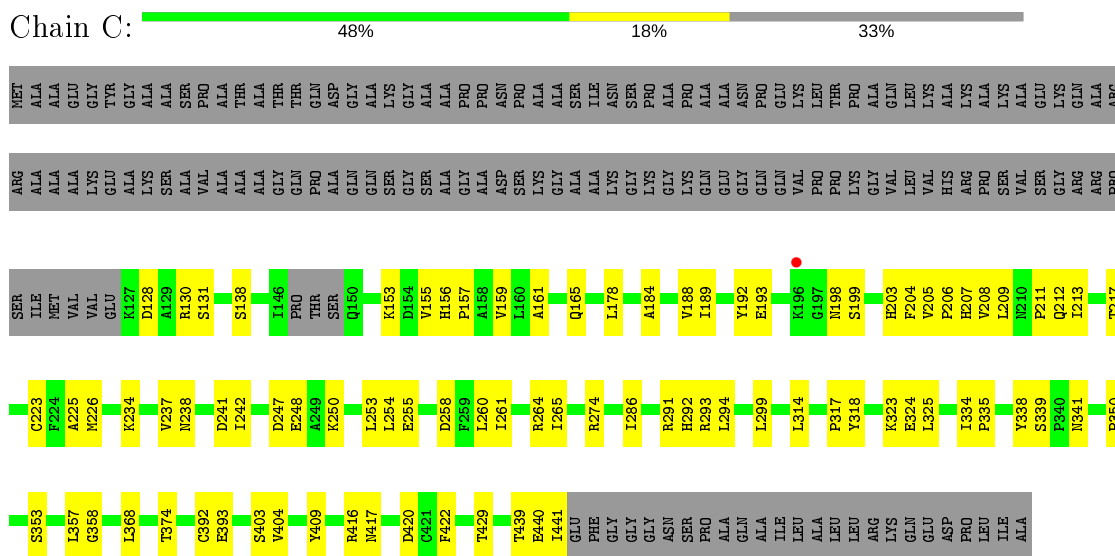
### 3 Residue-property plots

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.

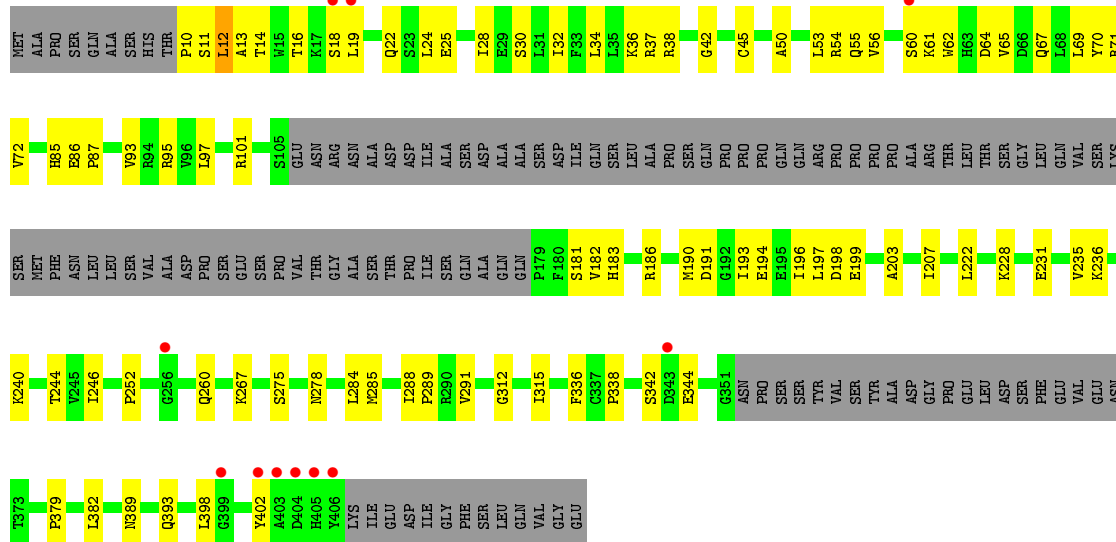
- Molecule 1: Translation initiation factor eif-2b-like protein



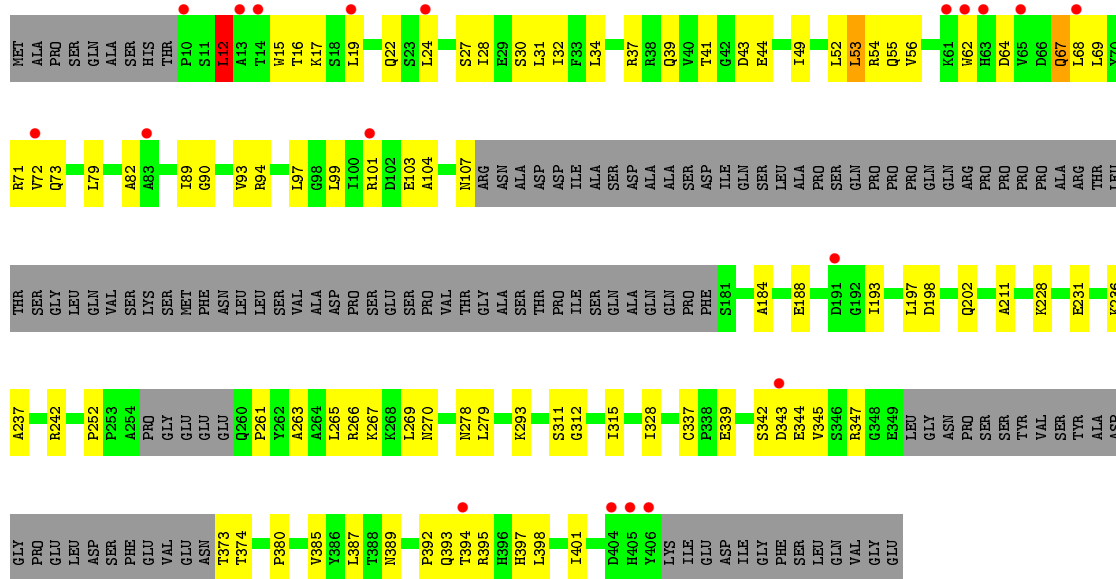
- Molecule 1: Translation initiation factor eif-2b-like protein



- Molecule 2: Translation initiation factor eIF2b-like protein, Translation initiation factor eIF2b-like protein



- Molecule 2: Translation initiation factor eIF2b-like protein, Translation initiation factor eIF2b-like protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	109.43Å 109.43Å 219.06Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.95 – 3.00 48.95 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.95-3.00) 99.7 (48.95-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.49 (at 3.01Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.208 , 0.255 0.214 , 0.256	Depositor DCC
$R_{free}$ test set	1554 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	76.9	Xtrriage
Anisotropy	0.403	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 77.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.037 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9388	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	95.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.68% 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 [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	0.36	0/2496	0.58	1/3380 (0.0%)
1	C	0.32	0/2457	0.56	0/3327
2	B	0.30	0/2332	0.58	0/3167
2	D	0.34	0/2266	0.61	1/3079 (0.0%)
All	All	0.33	0/9551	0.58	2/12953 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	12	LEU	CA-CB-CG	5.31	127.50	115.30
1	A	146	ILE	C-N-CD	5.05	139.01	128.40

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	2451	0	2463	99	0
1	C	2414	0	2431	71	0
2	B	2293	0	2306	85	0
2	D	2230	0	2228	115	0
All	All	9388	0	9428	361	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 19.

The worst 5 of 361 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:73:GLN:NE2	2:D:97:LEU:HD11	1.22	1.46
2:D:73:GLN:CD	2:D:97:LEU:HD11	1.40	1.38
2:D:22:GLN:HE21	2:D:27:SER:CB	1.39	1.33
2:B:13:ALA:O	2:B:16:THR:OG1	1.53	1.23
2:D:73:GLN:OE1	2:D:97:LEU:HD21	1.33	1.23

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	312/466 (67%)	307 (98%)	5 (2%)	0	100	100
1	C	308/466 (66%)	306 (99%)	2 (1%)	0	100	100
2	B	297/419 (71%)	293 (99%)	4 (1%)	0	100	100
2	D	288/419 (69%)	284 (99%)	4 (1%)	0	100	100
All	All	1205/1770 (68%)	1190 (99%)	15 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	261/362 (72%)	256 (98%)	5 (2%)	57	84
1	C	257/362 (71%)	253 (98%)	4 (2%)	62	86
2	B	239/346 (69%)	237 (99%)	2 (1%)	81	93
2	D	231/346 (67%)	225 (97%)	6 (3%)	46	78
All	All	988/1416 (70%)	971 (98%)	17 (2%)	60	85

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	193	GLU
1	C	247	ASP
2	D	67	GLN
2	B	393	GLN
2	D	71	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	22	GLN
2	D	270	ASN
2	D	39	GLN
1	A	396	ASN
2	D	67	GLN

### 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	315/466 (67%)	-0.14	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	40, 75, 123, 189	0
1	C	312/466 (66%)	0.00	1 (0%) <span style="border: 1px solid blue; padding: 2px;">94</span> <span style="border: 1px solid blue; padding: 2px;">84</span>	43, 82, 135, 190	0
2	B	303/419 (72%)	0.21	11 (3%) <span style="border: 1px solid red; padding: 2px;">42</span> <span style="border: 1px solid red; padding: 2px;">17</span>	51, 88, 172, 256	0
2	D	296/419 (70%)	0.44	19 (6%) <span style="border: 1px solid red; padding: 2px;">19</span> <span style="border: 1px solid red; padding: 2px;">6</span>	52, 105, 206, 299	0
All	All	1226/1770 (69%)	0.12	31 (2%) <span style="border: 1px solid blue; padding: 2px;">57</span> <span style="border: 1px solid red; padding: 2px;">29</span>	40, 85, 174, 299	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	405	HIS	8.4
2	B	404	ASP	6.2
2	B	403	ALA	5.2
2	B	402	TYR	4.9
2	D	405	HIS	4.6

### 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

There are no such residues in this entry.