



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:16 PM GMT

PDB ID : 3WKR  
Title : Crystal structure of the SepCysS-SepCysE complex from Methanocaldococcus jannaschii  
Authors : Nakazawa, Y.; Asano, N.; Nakamura, A.; Yao, M.  
Deposited on : 2013-10-30  
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

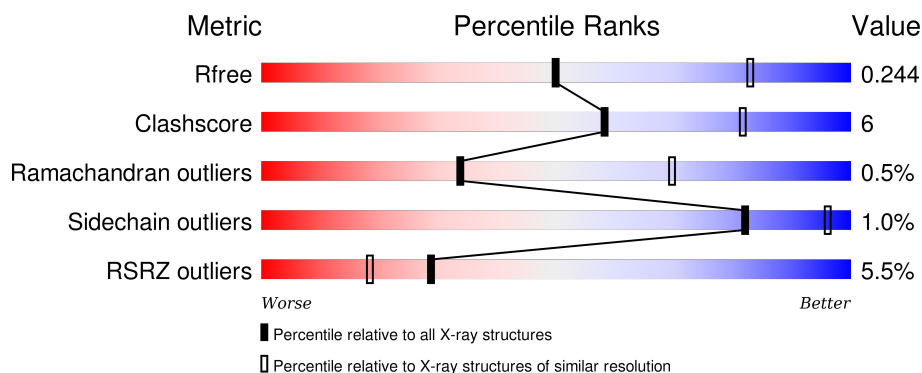
# 1 Overall quality at a glance

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	416	<div> <div>5%</div> <div>73%</div> <div>14%</div> <div>13%</div> </div>
1	B	416	<div> <div>3%</div> <div>72%</div> <div>15%</div> <div>13%</div> </div>
1	E	416	<div> <div>5%</div> <div>71%</div> <div>16%</div> <div>12%</div> </div>
1	F	416	<div> <div>6%</div> <div>71%</div> <div>16%</div> <div>13%</div> </div>
2	C	216	<div> <div>%</div> <div>27%</div> <div>69%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	216	<div><div><div>%</div><div><div></div><div></div><div></div></div><div>29%</div><div>69%</div></div></div>
2	G	216	<div><div><div>%</div><div><div></div><div></div><div></div></div><div>29%</div><div>68%</div></div></div>
2	H	216	<div><div><div>3%</div><div><div></div><div></div><div></div></div><div>29%</div><div>69%</div></div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 14030 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 O-phospho-L-seryl-tRNA:Cys-tRNA synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	362	Total	C	N	O	P	S	0	0	0
			2903	1865	485	539	1	13			
1	B	364	Total	C	N	O	P	S	0	0	0
			2916	1873	487	541	1	14			
1	E	365	Total	C	N	O	P	S	0	0	0
			2923	1877	488	543	1	14			
1	F	363	Total	C	N	O	P	S	0	0	0
			2910	1870	486	540	1	13			

There are 160 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP Q59072
A	-18	GLY	-	EXPRESSION TAG	UNP Q59072
A	-17	SER	-	EXPRESSION TAG	UNP Q59072
A	-16	SER	-	EXPRESSION TAG	UNP Q59072
A	-15	HIS	-	EXPRESSION TAG	UNP Q59072
A	-14	HIS	-	EXPRESSION TAG	UNP Q59072
A	-13	HIS	-	EXPRESSION TAG	UNP Q59072
A	-12	HIS	-	EXPRESSION TAG	UNP Q59072
A	-11	HIS	-	EXPRESSION TAG	UNP Q59072
A	-10	HIS	-	EXPRESSION TAG	UNP Q59072
A	-9	SER	-	EXPRESSION TAG	UNP Q59072
A	-8	SER	-	EXPRESSION TAG	UNP Q59072
A	-7	GLY	-	EXPRESSION TAG	UNP Q59072
A	-6	LEU	-	EXPRESSION TAG	UNP Q59072
A	-5	VAL	-	EXPRESSION TAG	UNP Q59072
A	-4	PRO	-	EXPRESSION TAG	UNP Q59072
A	-3	ARG	-	EXPRESSION TAG	UNP Q59072
A	-2	GLY	-	EXPRESSION TAG	UNP Q59072
A	-1	SER	-	EXPRESSION TAG	UNP Q59072
A	0	HIS	-	EXPRESSION TAG	UNP Q59072
A	1	MET	-	EXPRESSION TAG	UNP Q59072

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	GLU	-	EXPRESSION TAG	UNP Q59072
A	3	LEU	-	EXPRESSION TAG	UNP Q59072
A	4	GLU	-	EXPRESSION TAG	UNP Q59072
A	5	GLY	-	EXPRESSION TAG	UNP Q59072
A	6	PRO	-	EXPRESSION TAG	UNP Q59072
A	7	TYR	-	EXPRESSION TAG	UNP Q59072
A	8	SER	-	EXPRESSION TAG	UNP Q59072
A	9	LYS	-	EXPRESSION TAG	UNP Q59072
A	10	LYS	-	EXPRESSION TAG	UNP Q59072
A	11	PHE	-	EXPRESSION TAG	UNP Q59072
A	12	GLU	-	EXPRESSION TAG	UNP Q59072
A	13	VAL	-	EXPRESSION TAG	UNP Q59072
A	14	ILE	-	EXPRESSION TAG	UNP Q59072
A	15	THR	-	EXPRESSION TAG	UNP Q59072
A	16	LEU	-	EXPRESSION TAG	UNP Q59072
A	17	ASP	-	EXPRESSION TAG	UNP Q59072
A	18	ILE	-	EXPRESSION TAG	UNP Q59072
A	19	ASN	-	EXPRESSION TAG	UNP Q59072
A	20	LEU	-	EXPRESSION TAG	UNP Q59072
B	-19	MET	-	EXPRESSION TAG	UNP Q59072
B	-18	GLY	-	EXPRESSION TAG	UNP Q59072
B	-17	SER	-	EXPRESSION TAG	UNP Q59072
B	-16	SER	-	EXPRESSION TAG	UNP Q59072
B	-15	HIS	-	EXPRESSION TAG	UNP Q59072
B	-14	HIS	-	EXPRESSION TAG	UNP Q59072
B	-13	HIS	-	EXPRESSION TAG	UNP Q59072
B	-12	HIS	-	EXPRESSION TAG	UNP Q59072
B	-11	HIS	-	EXPRESSION TAG	UNP Q59072
B	-10	HIS	-	EXPRESSION TAG	UNP Q59072
B	-9	SER	-	EXPRESSION TAG	UNP Q59072
B	-8	SER	-	EXPRESSION TAG	UNP Q59072
B	-7	GLY	-	EXPRESSION TAG	UNP Q59072
B	-6	LEU	-	EXPRESSION TAG	UNP Q59072
B	-5	VAL	-	EXPRESSION TAG	UNP Q59072
B	-4	PRO	-	EXPRESSION TAG	UNP Q59072
B	-3	ARG	-	EXPRESSION TAG	UNP Q59072
B	-2	GLY	-	EXPRESSION TAG	UNP Q59072
B	-1	SER	-	EXPRESSION TAG	UNP Q59072
B	0	HIS	-	EXPRESSION TAG	UNP Q59072
B	1	MET	-	EXPRESSION TAG	UNP Q59072
B	2	GLU	-	EXPRESSION TAG	UNP Q59072
B	3	LEU	-	EXPRESSION TAG	UNP Q59072

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	4	GLU	-	EXPRESSION TAG	UNP Q59072
B	5	GLY	-	EXPRESSION TAG	UNP Q59072
B	6	PRO	-	EXPRESSION TAG	UNP Q59072
B	7	TYR	-	EXPRESSION TAG	UNP Q59072
B	8	SER	-	EXPRESSION TAG	UNP Q59072
B	9	LYS	-	EXPRESSION TAG	UNP Q59072
B	10	LYS	-	EXPRESSION TAG	UNP Q59072
B	11	PHE	-	EXPRESSION TAG	UNP Q59072
B	12	GLU	-	EXPRESSION TAG	UNP Q59072
B	13	VAL	-	EXPRESSION TAG	UNP Q59072
B	14	ILE	-	EXPRESSION TAG	UNP Q59072
B	15	THR	-	EXPRESSION TAG	UNP Q59072
B	16	LEU	-	EXPRESSION TAG	UNP Q59072
B	17	ASP	-	EXPRESSION TAG	UNP Q59072
B	18	ILE	-	EXPRESSION TAG	UNP Q59072
B	19	ASN	-	EXPRESSION TAG	UNP Q59072
B	20	LEU	-	EXPRESSION TAG	UNP Q59072
E	-19	MET	-	EXPRESSION TAG	UNP Q59072
E	-18	GLY	-	EXPRESSION TAG	UNP Q59072
E	-17	SER	-	EXPRESSION TAG	UNP Q59072
E	-16	SER	-	EXPRESSION TAG	UNP Q59072
E	-15	HIS	-	EXPRESSION TAG	UNP Q59072
E	-14	HIS	-	EXPRESSION TAG	UNP Q59072
E	-13	HIS	-	EXPRESSION TAG	UNP Q59072
E	-12	HIS	-	EXPRESSION TAG	UNP Q59072
E	-11	HIS	-	EXPRESSION TAG	UNP Q59072
E	-10	HIS	-	EXPRESSION TAG	UNP Q59072
E	-9	SER	-	EXPRESSION TAG	UNP Q59072
E	-8	SER	-	EXPRESSION TAG	UNP Q59072
E	-7	GLY	-	EXPRESSION TAG	UNP Q59072
E	-6	LEU	-	EXPRESSION TAG	UNP Q59072
E	-5	VAL	-	EXPRESSION TAG	UNP Q59072
E	-4	PRO	-	EXPRESSION TAG	UNP Q59072
E	-3	ARG	-	EXPRESSION TAG	UNP Q59072
E	-2	GLY	-	EXPRESSION TAG	UNP Q59072
E	-1	SER	-	EXPRESSION TAG	UNP Q59072
E	0	HIS	-	EXPRESSION TAG	UNP Q59072
E	1	MET	-	EXPRESSION TAG	UNP Q59072
E	2	GLU	-	EXPRESSION TAG	UNP Q59072
E	3	LEU	-	EXPRESSION TAG	UNP Q59072
E	4	GLU	-	EXPRESSION TAG	UNP Q59072
E	5	GLY	-	EXPRESSION TAG	UNP Q59072

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	6	PRO	-	EXPRESSION TAG	UNP Q59072
E	7	TYR	-	EXPRESSION TAG	UNP Q59072
E	8	SER	-	EXPRESSION TAG	UNP Q59072
E	9	LYS	-	EXPRESSION TAG	UNP Q59072
E	10	LYS	-	EXPRESSION TAG	UNP Q59072
E	11	PHE	-	EXPRESSION TAG	UNP Q59072
E	12	GLU	-	EXPRESSION TAG	UNP Q59072
E	13	VAL	-	EXPRESSION TAG	UNP Q59072
E	14	ILE	-	EXPRESSION TAG	UNP Q59072
E	15	THR	-	EXPRESSION TAG	UNP Q59072
E	16	LEU	-	EXPRESSION TAG	UNP Q59072
E	17	ASP	-	EXPRESSION TAG	UNP Q59072
E	18	ILE	-	EXPRESSION TAG	UNP Q59072
E	19	ASN	-	EXPRESSION TAG	UNP Q59072
E	20	LEU	-	EXPRESSION TAG	UNP Q59072
F	-19	MET	-	EXPRESSION TAG	UNP Q59072
F	-18	GLY	-	EXPRESSION TAG	UNP Q59072
F	-17	SER	-	EXPRESSION TAG	UNP Q59072
F	-16	SER	-	EXPRESSION TAG	UNP Q59072
F	-15	HIS	-	EXPRESSION TAG	UNP Q59072
F	-14	HIS	-	EXPRESSION TAG	UNP Q59072
F	-13	HIS	-	EXPRESSION TAG	UNP Q59072
F	-12	HIS	-	EXPRESSION TAG	UNP Q59072
F	-11	HIS	-	EXPRESSION TAG	UNP Q59072
F	-10	HIS	-	EXPRESSION TAG	UNP Q59072
F	-9	SER	-	EXPRESSION TAG	UNP Q59072
F	-8	SER	-	EXPRESSION TAG	UNP Q59072
F	-7	GLY	-	EXPRESSION TAG	UNP Q59072
F	-6	LEU	-	EXPRESSION TAG	UNP Q59072
F	-5	VAL	-	EXPRESSION TAG	UNP Q59072
F	-4	PRO	-	EXPRESSION TAG	UNP Q59072
F	-3	ARG	-	EXPRESSION TAG	UNP Q59072
F	-2	GLY	-	EXPRESSION TAG	UNP Q59072
F	-1	SER	-	EXPRESSION TAG	UNP Q59072
F	0	HIS	-	EXPRESSION TAG	UNP Q59072
F	1	MET	-	EXPRESSION TAG	UNP Q59072
F	2	GLU	-	EXPRESSION TAG	UNP Q59072
F	3	LEU	-	EXPRESSION TAG	UNP Q59072
F	4	GLU	-	EXPRESSION TAG	UNP Q59072
F	5	GLY	-	EXPRESSION TAG	UNP Q59072
F	6	PRO	-	EXPRESSION TAG	UNP Q59072
F	7	TYR	-	EXPRESSION TAG	UNP Q59072

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	8	SER	-	EXPRESSION TAG	UNP Q59072
F	9	LYS	-	EXPRESSION TAG	UNP Q59072
F	10	LYS	-	EXPRESSION TAG	UNP Q59072
F	11	PHE	-	EXPRESSION TAG	UNP Q59072
F	12	GLU	-	EXPRESSION TAG	UNP Q59072
F	13	VAL	-	EXPRESSION TAG	UNP Q59072
F	14	ILE	-	EXPRESSION TAG	UNP Q59072
F	15	THR	-	EXPRESSION TAG	UNP Q59072
F	16	LEU	-	EXPRESSION TAG	UNP Q59072
F	17	ASP	-	EXPRESSION TAG	UNP Q59072
F	18	ILE	-	EXPRESSION TAG	UNP Q59072
F	19	ASN	-	EXPRESSION TAG	UNP Q59072
F	20	LEU	-	EXPRESSION TAG	UNP Q59072

- Molecule 2 is a protein called Uncharacterized protein MJ1481.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	68	Total	C	N	O	S	0	0	0
			562	368	89	103	2			
2	D	68	Total	C	N	O	S	0	0	0
			565	371	89	103	2			
2	G	69	Total	C	N	O	S	0	0	0
			573	377	90	104	2			
2	H	68	Total	C	N	O	S	0	0	0
			565	371	89	103	2			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	MET	-	EXPRESSION TAG	UNP Q58876
C	-1	ASN	-	EXPRESSION TAG	UNP Q58876
C	0	HIS	-	EXPRESSION TAG	UNP Q58876
D	-2	MET	-	EXPRESSION TAG	UNP Q58876
D	-1	ASN	-	EXPRESSION TAG	UNP Q58876
D	0	HIS	-	EXPRESSION TAG	UNP Q58876
G	-2	MET	-	EXPRESSION TAG	UNP Q58876
G	-1	ASN	-	EXPRESSION TAG	UNP Q58876
G	0	HIS	-	EXPRESSION TAG	UNP Q58876
H	-2	MET	-	EXPRESSION TAG	UNP Q58876
H	-1	ASN	-	EXPRESSION TAG	UNP Q58876
H	0	HIS	-	EXPRESSION TAG	UNP Q58876



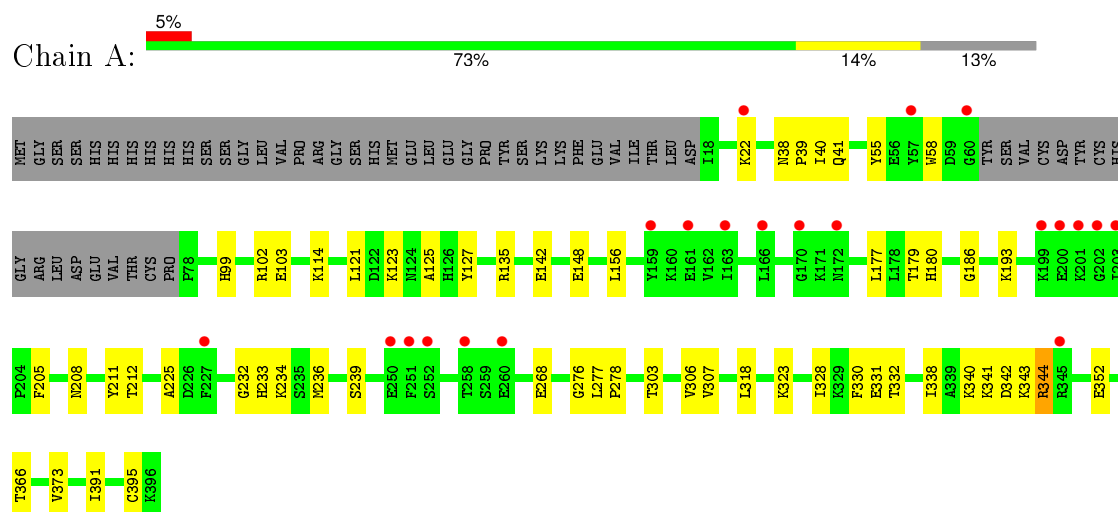
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	24	Total 24	O 24	0	0
3	B	26	Total 26	O 26	0	0
3	C	4	Total 4	O 4	0	0
3	D	4	Total 4	O 4	0	0
3	E	22	Total 22	O 22	0	0
3	F	20	Total 20	O 20	0	0
3	G	6	Total 6	O 6	0	0
3	H	7	Total 7	O 7	0	0

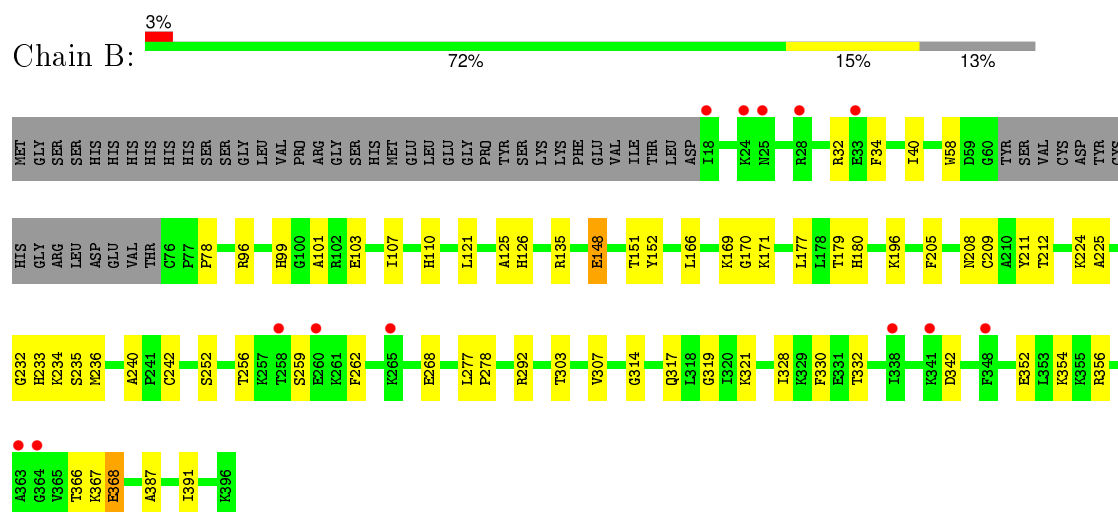
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

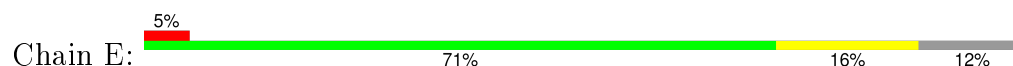
- Molecule 1: O-phospho-L-seryl-tRNA:Cys-tRNA synthase



- Molecule 1: O-phospho-L-seryl-tRNA:Cys-tRNA synthase



- Molecule 1: O-phospho-L-seryl-tRNA:Cys-tRNA synthase





GLY	GLY	ALA	LEU	TYR	ILE	VAL	SER	ASN	LYS	LYS	LEU	PHE	LYS	LYS	LEU	LYS	ASN	LYS	ASN	PRO	LYS	ASN	LEU	LYS	VAL	VAL	CYS	THR	GLU	GLY	MET	LEU	ASP	ILE	GLU	ASP	MET	ARG	ALA	ILE	GLY	VAL	PRO	PRO	GLU	GLY	LYS	LYS	VAL	GLU	ILE	ALA	ALA	ARG	LYS	ASN
VAL	GLU	ARG	PHE	ILE	GLU	LYS	TYR	LYS	PRO	GLU	GLU	LYS	ILE	PHE	VAL	VAL	GLU	ASP	ASP	LYS	ASP	GLU	LEU	LEU	TYR	LEU	ARG	ALA	LYS	ASN	LEU	TYR	ASN	ALA	GLU	GLU	ILE	ILE	LEU	ASP																

● Molecule 2: Uncharacterized protein MJ1481



ASN																										ASP	GLN	GLY	ALA	LEU	TYR	ILE	VAL	VAL	SER	LYS	ASN	LYS	ASP	LEU	ILE	ARG	LYS	GLY	ILE	SER	THR	ILE	SER	GLN	LEU	LYS	LYS	ALA	LYS	ILE	ARG	VAL	GLU	GLY	LYS	ASP	ASP	ILE	GLU	LYS	I34	E48	F49	Y54	I55	Y64	L72	K78	L83	H90	F102	GLU	GLY	LYS	ILE																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
ASN																										ASP	VAL	GLU	ARG	PHE	ILE	GLU	LYS	TYR	LYS	PRO	GLY	LYS	ILE	PHE	VAL	VAL	VAL	GLU	ASP	LYS	ASP	LYS	ASP	LYS	ASP	LYS	LYS	GLU	LEU	LYS	ASP	LYS	ASP	ASN	LYS	LEU	LEU	TYR	LYS	ASP	ASP	GLU	ALA	ALA	LYS	ARG	VAL	GLU	ILE	GLY	LYS	VAL	VAL	LYS	LYS	ALA	ILE	ILE	ARG	ALA	LEU	VAL	GLY	VAL	PRO	VAL	VAL	GLY	LEU	LYS	LYS	VAL	VAL	GLU	ILE	GLY	ILE	ALA	ARG	GLU	LYS	ILE	LYS	LYS	VAL	VAL	VAL	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY	GLY

● Molecule 2: Uncharacterized protein MJ1481



ALA	ARG	LYS	ASN	VAL	GLU	ARG	PHE	ILE	GLU	LYS	TYR	LYS	GLU	GLY	VAL	VAL	VAL	PHE	VAL	VAL	VAL	VAL	GLU	ASP	LYS	ASP	ASP	GLU	LEU	TYR	LEU	ARG	ALA	LYS	LEU	ASP	ALA	ALA	ASP	ASP	ILE	ILE	GLU	ILE	LEU	ASP						
MET	ASN	HIS	MET	ARG	VAL	GLU	TYR	SER	LYS	ASN	LEU	ASP	ILE	ARG	LYS	GLY	ILE	SER	THR	ILE	ILE	SER	GLN	LEU	LYS	LYS	ALA	LYS	ILE	ARG	VAL	GLU	LYS	LYS	ASP	LYS	ILE	S35	Y36	K37	I44	E48	A52	L63	H90	T94	F98	G99	F100	E101	F102	GLU

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.88Å 106.88Å 495.25Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.30 – 2.80 47.30 – 2.80	Depositor EDS
% Data completeness (in resolution range)	94.9 (47.30-2.80) 94.9 (47.30-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.32 (at 2.81Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, $R_{free}$	0.224 , 0.242 0.226 , 0.244	Depositor DCC
$R_{free}$ test set	4277 reflections (5.78%)	DCC
Wilson B-factor (Å <sup>2</sup> )	71.2	Xtriage
Anisotropy	0.006	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 70.8	EDS
Estimated twinning fraction	0.498 for h,-h-k,-l 0.467 for h,-h-k,-l	Xtriage
Reported twinning fraction	0.498 for h,-h-k,-l	Depositor
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.37$ , $\langle L^2 \rangle = 0.19$	Xtriage
Outliers	0 of 74017 reflections	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	14030	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	101.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.17% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: LLP

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	0.20	0/2933	0.35	0/3939
1	B	0.21	0/2947	0.36	0/3960
1	E	0.20	0/2954	0.36	0/3970
1	F	0.21	0/2941	0.35	0/3951
2	C	0.21	0/573	0.31	0/765
2	D	0.21	0/577	0.30	0/770
2	G	0.21	0/585	0.30	0/781
2	H	0.21	0/577	0.30	0/770
All	All	0.21	0/14087	0.35	0/18906

There are no bond length outliers.

There are no bond angle outliers.

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	2903	0	2953	36	0
1	B	2916	0	2964	44	0
1	E	2923	0	2971	45	0
1	F	2910	0	2960	44	0
2	C	562	0	571	7	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	565	0	569	5	0
2	G	573	0	580	6	0
2	H	565	0	569	4	0
3	A	24	0	0	0	0
3	B	26	0	0	0	0
3	C	4	0	0	0	0
3	D	4	0	0	0	0
3	E	22	0	0	0	0
3	F	20	0	0	0	0
3	G	6	0	0	0	0
3	H	7	0	0	0	0
All	All	14030	0	14137	164	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.

The worst 5 of 164 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:B:101:ALA:HB3	1:B:234:LLP:H5'2	1.73	0.71
1:A:268:GLU:HA	1:B:135:ARG:HD2	1.75	0.69
1:B:232:GLY:HA2	1:B:236:MET:HB2	1.75	0.68
1:E:268:GLU:HA	1:F:135:ARG:HD2	1.77	0.66
1:E:55:TYR:O	1:F:32:ARG:NH1	2.26	0.66

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	357/416 (86%)	333 (93%)	23 (6%)	1 (0%)	46 79

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	359/416 (86%)	342 (95%)	15 (4%)	2 (1%)	30	65
1	E	360/416 (86%)	335 (93%)	22 (6%)	3 (1%)	24	58
1	F	358/416 (86%)	340 (95%)	16 (4%)	2 (1%)	30	65
2	C	66/216 (31%)	65 (98%)	1 (2%)	0	100	100
2	D	66/216 (31%)	65 (98%)	1 (2%)	0	100	100
2	G	67/216 (31%)	66 (98%)	1 (2%)	0	100	100
2	H	66/216 (31%)	65 (98%)	1 (2%)	0	100	100
All	All	1699/2528 (67%)	1611 (95%)	80 (5%)	8 (0%)	34	69

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	148	GLU
1	E	115	GLU
1	E	148	GLU
1	F	314	GLY
1	A	148	GLU

### 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	309/358 (86%)	306 (99%)	3 (1%)	82	96
1	B	311/358 (87%)	307 (99%)	4 (1%)	76	94
1	E	312/358 (87%)	308 (99%)	4 (1%)	76	94
1	F	310/358 (87%)	306 (99%)	4 (1%)	76	94
2	C	60/195 (31%)	60 (100%)	0	100	100
2	D	60/195 (31%)	60 (100%)	0	100	100
2	G	61/195 (31%)	61 (100%)	0	100	100
2	H	60/195 (31%)	60 (100%)	0	100	100
All	All	1483/2212 (67%)	1468 (99%)	15 (1%)	82	96



5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	368	GLU
1	E	58	TRP
1	F	96	ARG
1	B	211	TYR
1	F	58	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	F	208	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

4 non-standard protein/DNA/RNA residues 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$
1	LLP	A	234	1	23,24,25	1.72	3 (13%)	28,32,34	1.79	5 (17%)
1	LLP	B	234	1	23,24,25	1.71	2 (8%)	28,32,34	1.88	5 (17%)
1	LLP	E	234	1	23,24,25	1.72	4 (17%)	28,32,34	1.93	4 (14%)
1	LLP	F	234	1	23,24,25	1.74	4 (17%)	28,32,34	1.77	6 (21%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	LLP	A	234	1	-	0/15/17/19	0/1/1/1
1	LLP	B	234	1	-	0/15/17/19	0/1/1/1
1	LLP	E	234	1	-	0/15/17/19	0/1/1/1
1	LLP	F	234	1	-	0/15/17/19	0/1/1/1

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	234	LLP	O3-C3	-5.81	1.23	1.37
1	F	234	LLP	O3-C3	-5.80	1.23	1.37
1	B	234	LLP	O3-C3	-5.78	1.23	1.37
1	A	234	LLP	O3-C3	-5.75	1.23	1.37
1	E	234	LLP	C3-C2	-2.01	1.39	1.40

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	234	LLP	CE-NZ-C4'	-3.01	110.28	118.97
1	F	234	LLP	CE-NZ-C4'	-2.99	110.35	118.97
1	E	234	LLP	CE-NZ-C4'	-2.95	110.45	118.97
1	B	234	LLP	CE-NZ-C4'	-2.93	110.51	118.97
1	E	234	LLP	C4-C4'-NZ	-2.81	109.45	125.06

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	234	LLP	3	0
1	B	234	LLP	5	0
1	E	234	LLP	1	0
1	F	234	LLP	4	0

## 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	361/416 (86%)	0.37	21 (5%) 26 16	36, 102, 154, 181	0
1	B	363/416 (87%)	0.15	13 (3%) 46 34	39, 74, 168, 203	0
1	E	364/416 (87%)	0.49	22 (6%) 25 15	42, 102, 163, 201	0
1	F	362/416 (87%)	0.51	23 (6%) 23 14	34, 114, 182, 213	0
2	C	68/216 (31%)	0.28	2 (2%) 55 43	43, 79, 146, 173	0
2	D	68/216 (31%)	0.25	3 (4%) 38 26	39, 71, 169, 182	0
2	G	69/216 (31%)	0.31	3 (4%) 39 27	45, 94, 154, 196	0
2	H	68/216 (31%)	0.52	7 (10%) 9 4	35, 116, 178, 193	0
All	All	1723/2528 (68%)	0.37	94 (5%) 29 18	34, 97, 171, 213	0

The worst 5 of 94 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	102	PHE	6.1
1	E	23	TYR	4.9
1	F	60	GLY	4.6
2	G	102	PHE	4.4
1	A	203	ILE	4.3

### 6.2 Non-standard residues in protein, DNA, RNA chains [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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	LLP	A	234	24/25	0.95	0.17	-	29,65,85,92	0
1	LLP	F	234	24/25	0.93	0.21	-	63,92,100,105	0
1	LLP	E	234	24/25	0.92	0.23	-	79,128,147,154	0
1	LLP	B	234	24/25	0.94	0.19	-	59,80,91,95	0

### 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.