



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 01:27 am BST

PDB ID : 3TKN
Title : Structure of the Nup82-Nup159-Nup98 heterotrimer
Authors : Stuwe, T.T.; Hoelz, A.
Deposited on : 2011-08-28
Resolution : 3.40 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

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

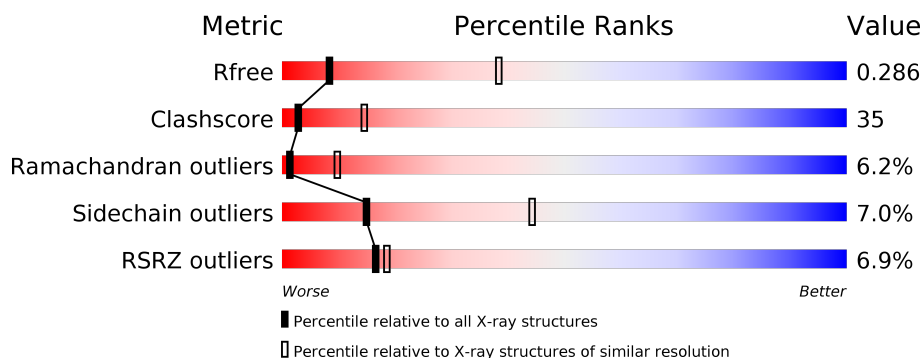
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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 ≥ 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	452	<div> <div>7%</div> <div> <div>44%</div> <div>48%</div> <div>8%</div> </div> </div>
1	D	452	<div> <div>%</div> <div> <div>44%</div> <div>47%</div> <div>8%</div> </div> </div>
1	G	452	<div> <div>6%</div> <div> <div>38%</div> <div>53%</div> <div>8%</div> </div> </div>
2	B	39	<div> <div>46%</div> <div>18%</div> <div>8%</div> <div>28%</div> </div>
2	E	39	<div> <div>36%</div> <div>28%</div> <div>8%</div> <div>28%</div> </div>
2	H	39	<div> <div>5%</div> <div> <div>38%</div> <div>26%</div> <div>8%</div> <div>28%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	152	<div><div></div><div>13%</div><div>38%</div><div>51%</div><div>8%</div><div></div><div></div></div>
3	F	152	<div><div></div><div>6%</div><div>40%</div><div>48%</div><div>8%</div><div></div><div></div></div>
3	I	152	<div><div></div><div>24%</div><div>39%</div><div>49%</div><div>8%</div><div></div><div></div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15021 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 Nucleoporin NUP82.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	451	Total	C	N	O	S	0	0	0
			3628	2323	587	707	11			
1	D	447	Total	C	N	O	S	0	0	0
			3601	2309	582	699	11			
1	G	447	Total	C	N	O	S	0	0	0
			3601	2309	582	699	11			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	396	SER	CYS	CONFLICT	UNP P40368
D	396	SER	CYS	CONFLICT	UNP P40368
G	396	SER	CYS	CONFLICT	UNP P40368

- Molecule 2 is a protein called Nucleoporin NUP159.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	28	Total	C	N	O	S	0	0	0
			221	144	37	37	3			
2	E	28	Total	C	N	O	S	0	0	0
			221	144	37	37	3			
2	H	28	Total	C	N	O	S	0	0	0
			221	144	37	37	3			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1422	GLY	-	EXPRESSION TAG	UNP P40477
B	1423	PRO	-	EXPRESSION TAG	UNP P40477
B	1424	HIS	-	EXPRESSION TAG	UNP P40477
E	1422	GLY	-	EXPRESSION TAG	UNP P40477
E	1423	PRO	-	EXPRESSION TAG	UNP P40477

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	1424	HIS	-	EXPRESSION TAG	UNP P40477
H	1422	GLY	-	EXPRESSION TAG	UNP P40477
H	1423	PRO	-	EXPRESSION TAG	UNP P40477
H	1424	HIS	-	EXPRESSION TAG	UNP P40477

- Molecule 3 is a protein called Nucleoporin 98.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	147	Total	C	N	O	S	0	0	0
			1176	745	202	226	3			
3	F	147	Total	C	N	O	S	0	0	0
			1176	745	202	226	3			
3	I	147	Total	C	N	O	S	0	0	0
			1176	745	202	226	3			

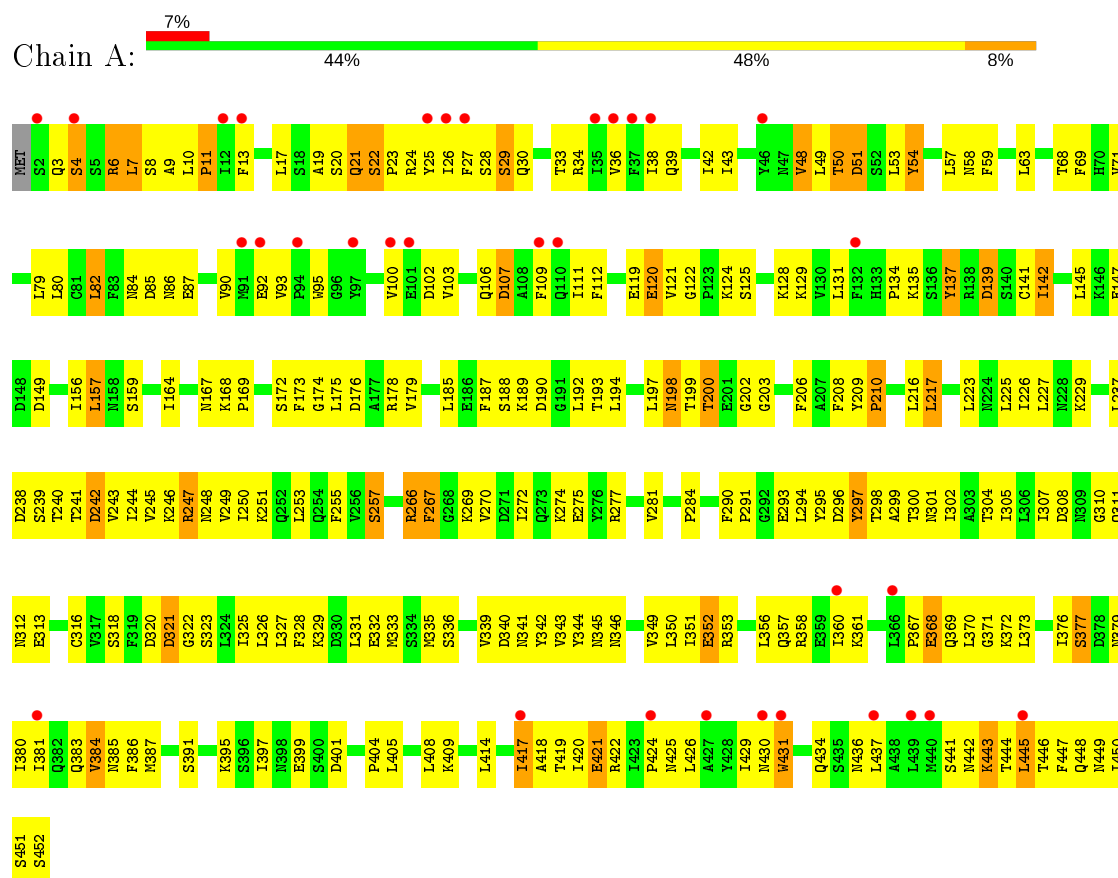
There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	729	GLY	-	EXPRESSION TAG	UNP Q6PFD9
C	730	PRO	-	EXPRESSION TAG	UNP Q6PFD9
C	731	HIS	-	EXPRESSION TAG	UNP Q6PFD9
F	729	GLY	-	EXPRESSION TAG	UNP Q6PFD9
F	730	PRO	-	EXPRESSION TAG	UNP Q6PFD9
F	731	HIS	-	EXPRESSION TAG	UNP Q6PFD9
I	729	GLY	-	EXPRESSION TAG	UNP Q6PFD9
I	730	PRO	-	EXPRESSION TAG	UNP Q6PFD9
I	731	HIS	-	EXPRESSION TAG	UNP Q6PFD9

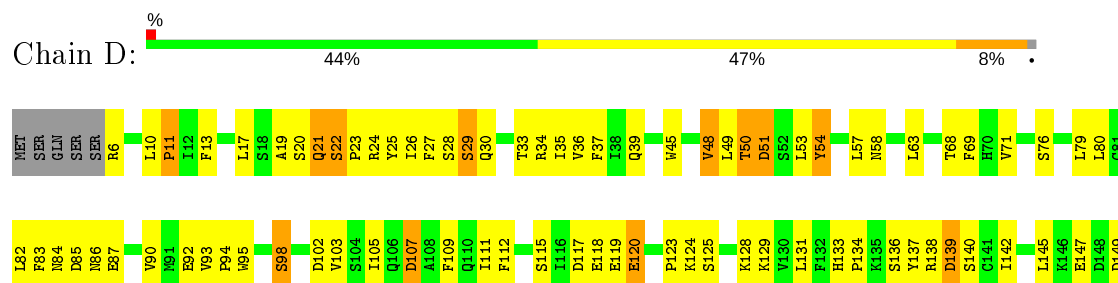
3 Residue-property plots [i](#)

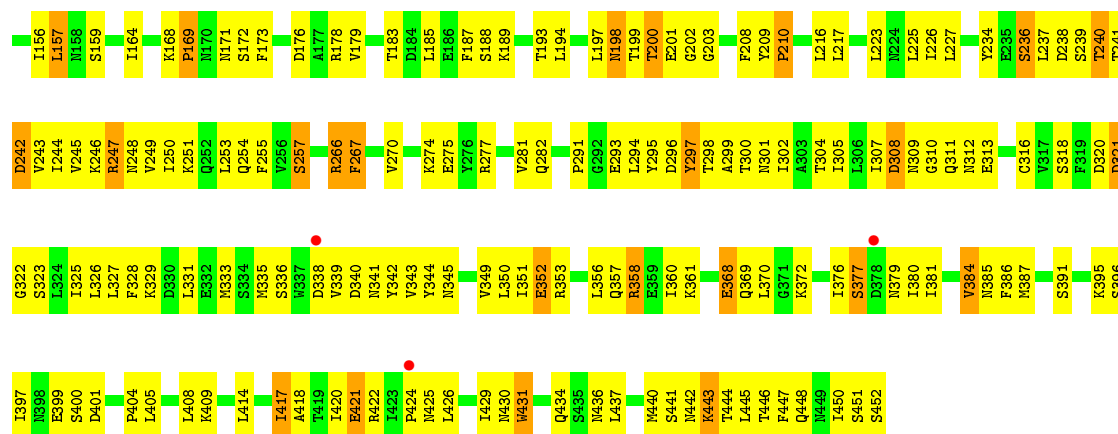
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Nucleoporin NUP82

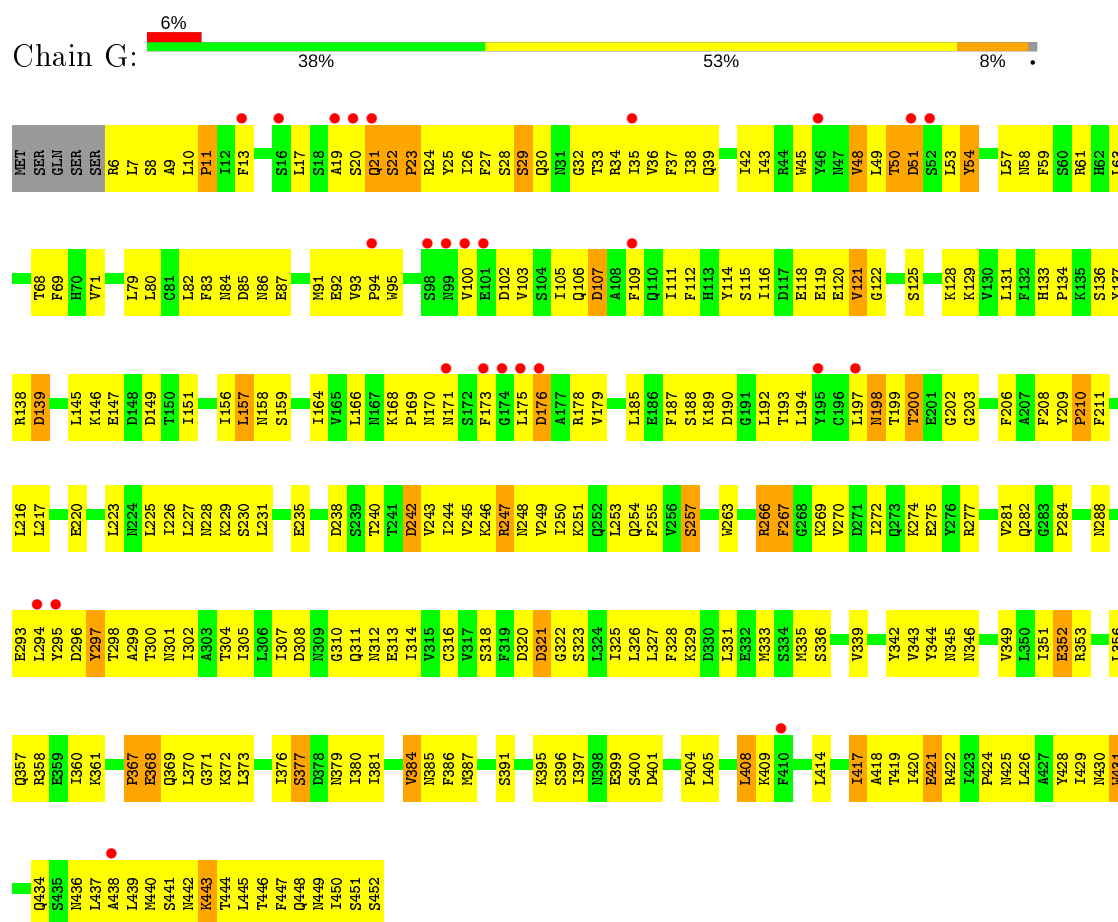


• Molecule 1: Nucleoporin NUP82



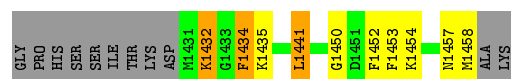


• Molecule 1: Nucleoporin NUP82



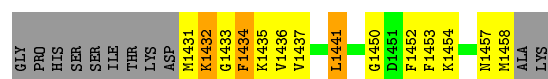
• Molecule 2: Nucleoporin NUP159

Chain B: 46% 18% 8% 28%



• Molecule 2: Nucleoporin NUP159

Chain E: 



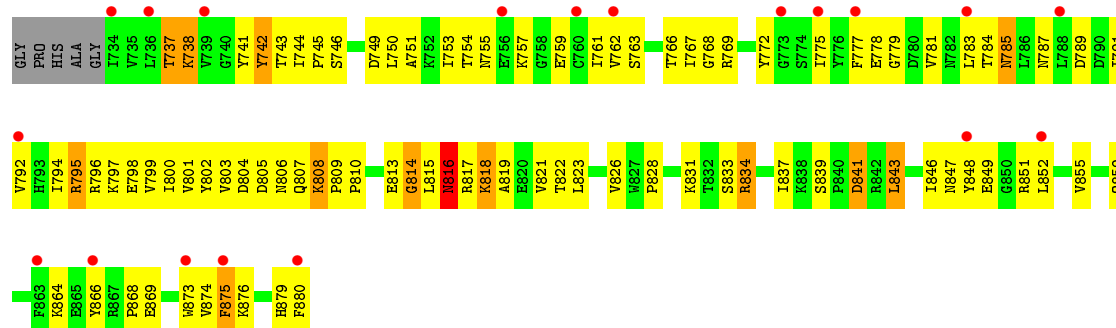
• Molecule 2: Nucleoporin NUP159

Chain H: 



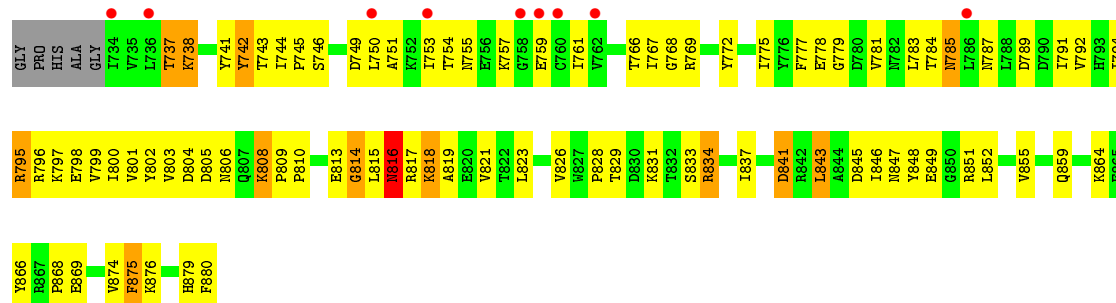
• Molecule 3: Nucleoporin 98

Chain C: 



• Molecule 3: Nucleoporin 98

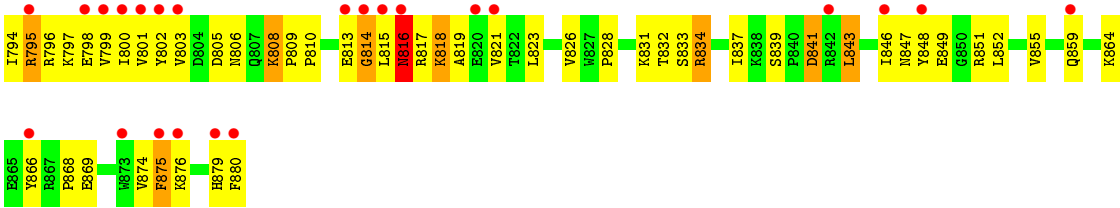
Chain F: 



• Molecule 3: Nucleoporin 98

Chain I: 





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	106.91Å 115.85Å 118.49Å 90.00° 111.06° 90.00°	Depositor
Resolution (Å)	20.00 – 3.40 49.90 – 3.40	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-3.40) 92.5 (49.90-3.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.76 (at 3.40Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.249 , 0.285 0.250 , 0.286	Depositor DCC
R_{free} test set	3659 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å ²)	96.5	Xtriage
Anisotropy	0.058	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 41.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	15021	wwPDB-VP
Average B, all atoms (Å ²)	113.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.52% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.48	0/3701	0.70	0/5025
1	D	0.49	0/3674	0.69	0/4989
1	G	0.48	0/3674	0.69	0/4989
2	B	0.60	0/223	0.69	0/293
2	E	0.60	0/223	0.71	0/293
2	H	0.57	0/223	0.66	0/293
3	C	0.35	0/1199	0.59	0/1620
3	F	0.36	0/1199	0.59	0/1620
3	I	0.36	0/1199	0.58	0/1620
All	All	0.46	0/15315	0.67	0/20742

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	3628	0	3594	251	0
1	D	3601	0	3571	239	0
1	G	3601	0	3571	272	0
2	B	221	0	235	12	0
2	E	221	0	235	17	0
2	H	221	0	235	17	0
3	C	1176	0	1170	105	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	F	1176	0	1170	102	0
3	I	1176	0	1170	98	0
All	All	15021	0	14951	1053	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 35.

The worst 5 of 1053 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:G:57:LEU:HD22	1:G:100:VAL:HG11	1.36	1.02
1:A:121:VAL:HG12	1:A:122:GLY:H	1.32	0.95
1:A:6:ARG:HB2	1:A:431:TRP:HH2	1.30	0.93
1:D:142:ILE:HG23	1:D:156:ILE:HD11	1.53	0.89
1:D:381:ILE:HD11	1:D:424:PRO:HG3	1.55	0.89

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	449/452 (99%)	358 (80%)	64 (14%)	27 (6%)	1	10
1	D	445/452 (98%)	351 (79%)	65 (15%)	29 (6%)	1	8
1	G	445/452 (98%)	358 (80%)	62 (14%)	25 (6%)	2	12
2	B	26/39 (67%)	21 (81%)	5 (19%)	0	100	100
2	E	26/39 (67%)	21 (81%)	4 (15%)	1 (4%)	3	19
2	H	26/39 (67%)	20 (77%)	6 (23%)	0	100	100
3	C	145/152 (95%)	110 (76%)	24 (17%)	11 (8%)	1	6
3	F	145/152 (95%)	110 (76%)	24 (17%)	11 (8%)	1	6

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	I	145/152 (95%)	110 (76%)	24 (17%)	11 (8%)	1	6
All	All	1852/1929 (96%)	1459 (79%)	278 (15%)	115 (6%)	1	10

5 of 115 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	120	GLU
1	A	266	ARG
1	A	311	GLN
1	A	377	SER
3	C	738	LYS

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	422/423 (100%)	392 (93%)	30 (7%)	14	44
1	D	418/423 (99%)	391 (94%)	27 (6%)	17	46
1	G	418/423 (99%)	389 (93%)	29 (7%)	15	45
2	B	24/33 (73%)	21 (88%)	3 (12%)	4	17
2	E	24/33 (73%)	22 (92%)	2 (8%)	11	36
2	H	24/33 (73%)	21 (88%)	3 (12%)	4	17
3	C	130/132 (98%)	121 (93%)	9 (7%)	15	45
3	F	130/132 (98%)	121 (93%)	9 (7%)	15	45
3	I	130/132 (98%)	121 (93%)	9 (7%)	15	45
All	All	1720/1764 (98%)	1599 (93%)	121 (7%)	15	44

5 of 121 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	242	ASP
1	D	431	TRP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	H	1441	LEU
1	D	247	ARG
1	D	352	GLU

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

Mol	Chain	Res	Type
1	G	436	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 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, 95th 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 ²)	Q<0.9
1	A	451/452 (99%)	0.28	33 (7%) 15 17	60, 104, 150, 170	0
1	D	447/452 (98%)	0.32	3 (0%) 87 87	61, 99, 145, 170	0
1	G	447/452 (98%)	0.30	26 (5%) 23 24	63, 104, 148, 170	0
2	B	28/39 (71%)	-0.04	0 100 100	68, 90, 136, 136	0
2	E	28/39 (71%)	0.02	0 100 100	65, 87, 127, 136	0
2	H	28/39 (71%)	0.45	2 (7%) 16 18	69, 88, 126, 137	0
3	C	147/152 (96%)	0.70	19 (12%) 3 4	95, 141, 190, 199	0
3	F	147/152 (96%)	0.40	9 (6%) 21 22	93, 140, 190, 198	0
3	I	147/152 (96%)	1.15	37 (25%) 0 0	96, 141, 190, 199	0
All	All	1870/1929 (96%)	0.40	129 (6%) 16 18	60, 111, 169, 199	0

The worst 5 of 129 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	880	PHE	10.5
1	G	101	GLU	6.4
2	H	1458	MET	5.8
2	H	1457	ASN	5.3
1	G	171	ASN	5.3

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.