



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:19 PM GMT

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
<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

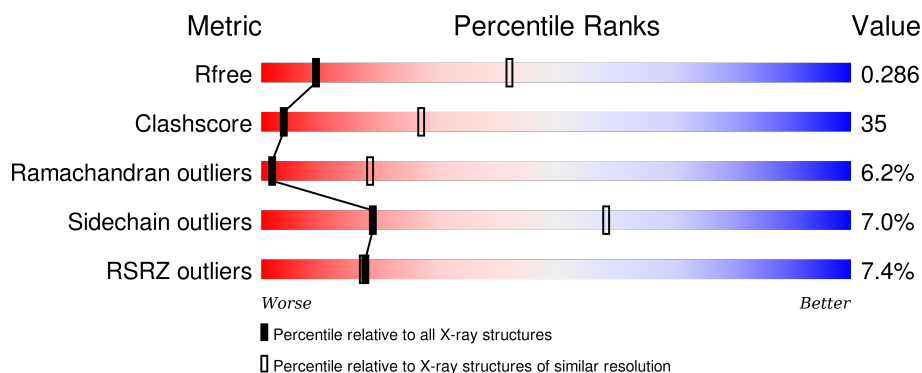
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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}	91344	1476 (3.50-3.30)
Clashscore	102246	1611 (3.50-3.30)
Ramachandran outliers	100387	1571 (3.50-3.30)
Sidechain outliers	100360	1571 (3.50-3.30)
RSRZ outliers	91569	1485 (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. 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>8%</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>7%</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>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	39	<div><div></div><div>5%</div><div>38%</div><div>26%</div><div>8%</div><div>28%</div></div>
3	C	152	<div><div></div><div>13%</div><div>38%</div><div>51%</div><div>8%</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>
3	I	152	<div><div></div><div>27%</div><div>39%</div><div>49%</div><div>8%</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			

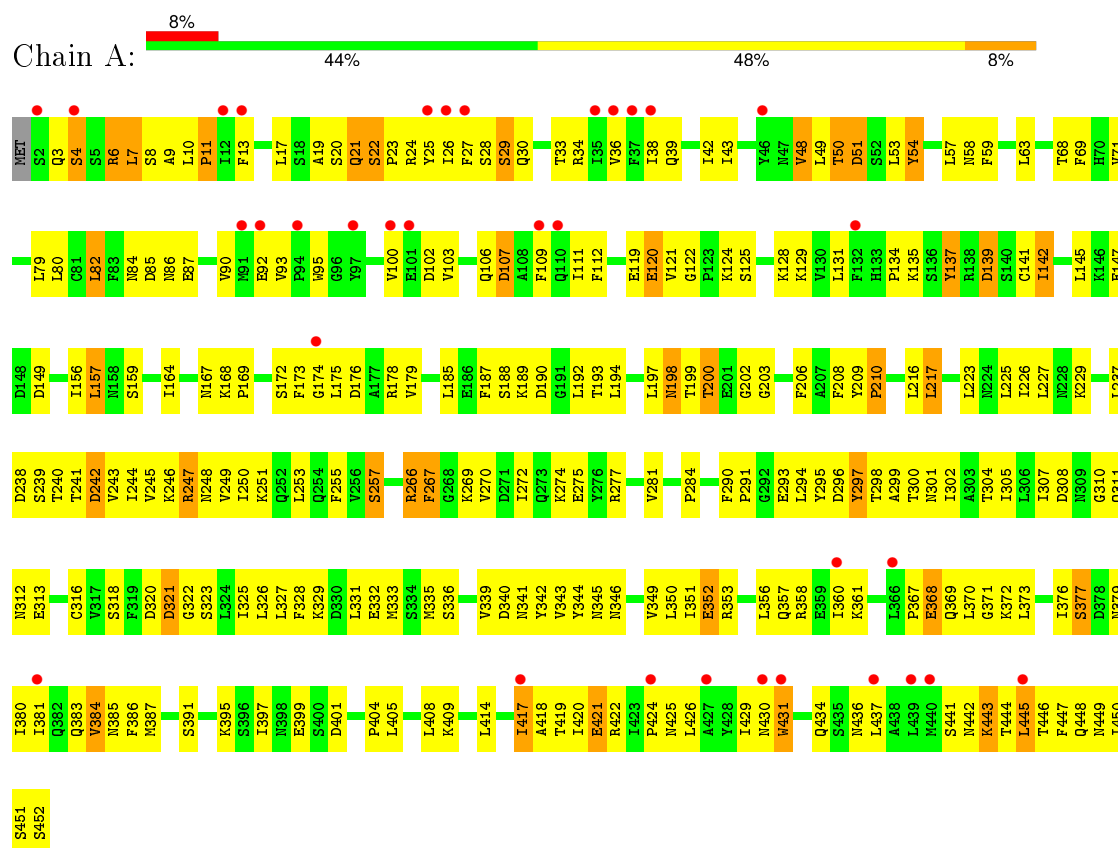
There are 9 discrepancies between the modelled and reference sequences:

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

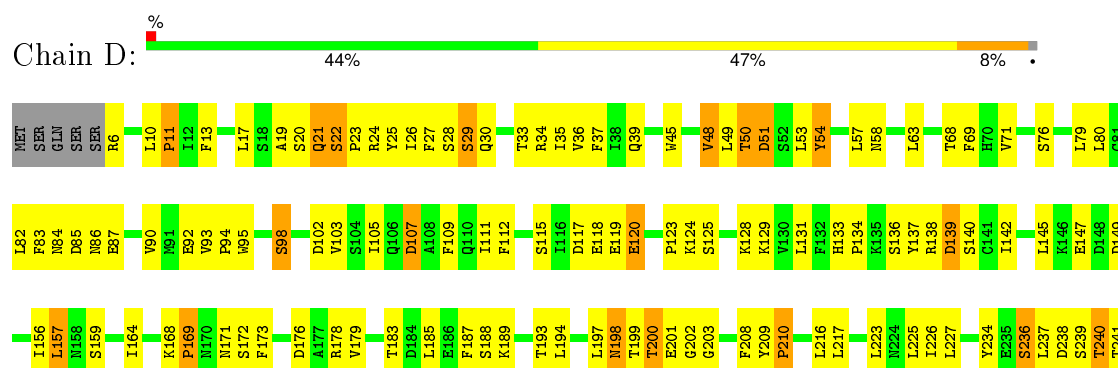
3 Residue-property plots

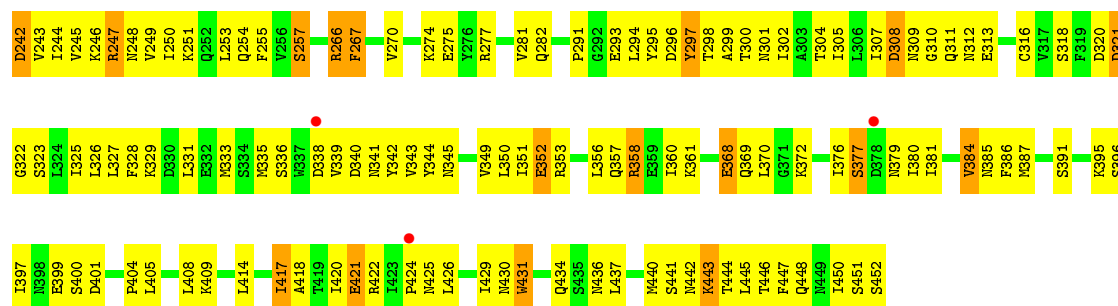
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

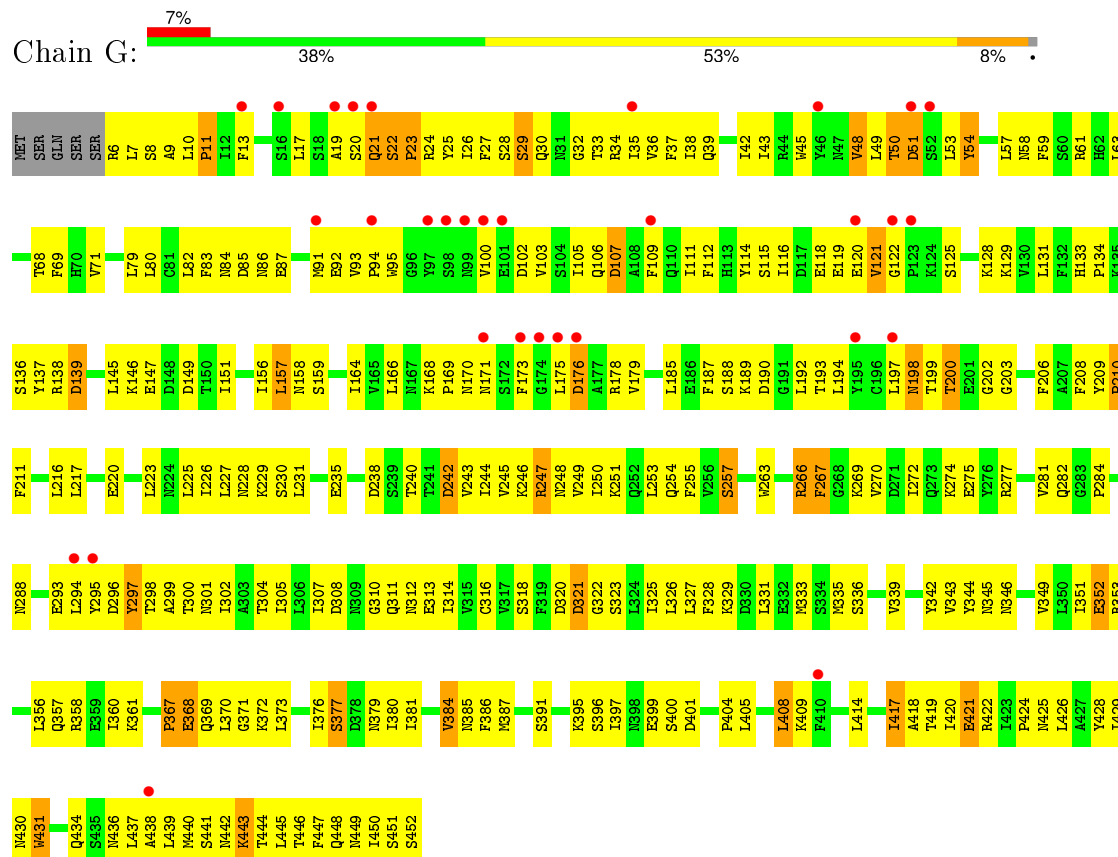


• Molecule 1: Nucleoporin NUP82

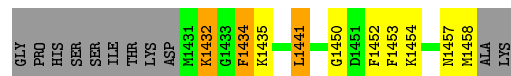




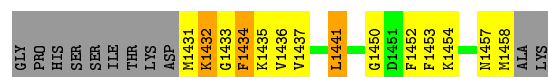
• Molecule 1: Nucleoporin NUP82



• Molecule 2: Nucleoporin NUP159



• Molecule 2: Nucleoporin NUP159



Chain H:

Amino Acid Type	Percentage
Red	5%
Green	38%
Yellow	26%
Orange	8%
Grey	28%

Sequence: GLY, PRO, HIS, SER, SER, SER, ILE, THR, LYS, ASP, M1431, K1432, G1433, F1434, K1435, V1436, V1437, L1441, T1449, G1450, D1451, F1452, F1453, K1454, M1457, M1458, ALA, LYS

Chain C:

13% 38% 51% 8%

GLY PRO HIS ALA I734 V735 W736 K737 T738 V739 G740 Y741 T742 W743 I744 P745 S746 D749 L750 A751 K752 I753 I754 N755 E756 K757 G758 E759 I761 V762 S763 T766 G768 R769 Y772 G773 S774 I775 Y776 F777 E778 G779 D780 V781 N782 L783 T784 N785 L786 N787 L788 D789 I790 V791 V792 H793 I794 R795 T796 K797 E798 V799 I800 V801 Y802 V803 D804 D805 N806 Q807 K808 P809 P810 E813 G814 L815 R816 R817 K818 A819 E820 V821 T822 L823 V826 W827 P828 R831 T832 S833 R834 T837 R838 S839 P840 D841 R842 L843 I846 N847 Y848 E849 G850 R851 L852 V855 P856

Chain F:

6% 40% 48% 8% ..

GLY PRO HIS ALA GLY I734 I735 I736 I737 K738 T738 I741 I742 I743 I744 I745 S746 D749 L750 A751 K752 I753 I754 N755 E756 K757 G758 E759 G760 I761 I762 T766 I767 G768 I769 Y772 I775 I776 F777 E778 G779 D780 I781 I782 L783 I784 N785 L786 N787 L788 D789 D790 I791 V792 I793 I794 R795 R796 R797 K797 E798 I799 I800 Y801 Y802 V803 D804 D805 N806 Q807 K808 P809 P810 E813 G814 L815 R816 K817 K818 A819 E820 V821 T822 L823 V826 R827 P828 T829 D830 K831 T832 S833 R834 I837 D841 R842 L843 A844 D845 I846 I847 Y848 E849 G850 R851 L852 V855 Q859 K864 R865

Chain I:

27%

39%

49%

8%

• •

GLY

PRO

HIS

ALA

GLY

Y734

Y737

K738

Y741

Y742

I744

P745

S746

D749

L750

A751

K752

I753

Y754

N755

E756

K757

G758

C760

I761

V762

S763

D764

F765

I766

I767

G768

R769

Y772

I775

Y776

F777

E778

G779

D780

Y781

N782

L783

Y784

N785

L786

N787

L788

D789

D790

Y791

Y792

F793

I794

R795

K796

K797

E798

Y799

I800

Y801

Y802

Y803

D804

D805

N806

Q807

K808

P809

P810

E813

G814

L815

R816

R817

K818

R819

A819

V820

T821

T822

L823

V826

W827

P828

K831

T832

S833

R834

I837

K838

S839

P840

D841

R842

L843

I846

N847

Y848

E849

G850

R851

L852

V855

Q859

Y864

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	CNS1.2	Depositor
R, R_{free}	0.249 , 0.285 0.250 , 0.286	Depositor DCC
R_{free} test set	3437 reflections (9.99%)	DCC
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
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 36635 reflections	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, 0.375 respectively for untwinned datasets, and 0.333, 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%)	2	19
1	D	445/452 (98%)	351 (79%)	65 (15%)	29 (6%)	1	17
1	G	445/452 (98%)	358 (80%)	62 (14%)	25 (6%)	2	21
2	B	26/39 (67%)	21 (81%)	5 (19%)	0	100	100
2	E	26/39 (67%)	21 (81%)	4 (15%)	1 (4%)	4	32
2	H	26/39 (67%)	20 (77%)	6 (23%)	0	100	100
3	C	145/152 (95%)	110 (76%)	24 (17%)	11 (8%)	1	13
3	F	145/152 (95%)	110 (76%)	24 (17%)	11 (8%)	1	13

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	13
All	All	1852/1929 (96%)	1459 (79%)	278 (15%)	115 (6%)	2	18

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%)	18	58
1	D	418/423 (99%)	391 (94%)	27 (6%)	21	61
1	G	418/423 (99%)	389 (93%)	29 (7%)	19	59
2	B	24/33 (73%)	21 (88%)	3 (12%)	6	27
2	E	24/33 (73%)	22 (92%)	2 (8%)	14	49
2	H	24/33 (73%)	21 (88%)	3 (12%)	6	27
3	C	130/132 (98%)	121 (93%)	9 (7%)	19	59
3	F	130/132 (98%)	121 (93%)	9 (7%)	19	59
3	I	130/132 (98%)	121 (93%)	9 (7%)	19	59
All	All	1720/1764 (98%)	1599 (93%)	121 (7%)	19	58

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%)	0.31	34 (7%) 17 17	60, 104, 150, 170	0
1	D	447/452 (98%)	0.35	3 (0%) 89 85	61, 99, 145, 170	0
1	G	447/452 (98%)	0.33	31 (6%) 20 19	63, 104, 148, 170	0
2	B	28/39 (71%)	-0.01	0 100 100	68, 90, 136, 136	0
2	E	28/39 (71%)	0.05	0 100 100	65, 87, 127, 136	0
2	H	28/39 (71%)	0.48	2 (7%) 19 18	69, 88, 126, 137	0
3	C	147/152 (96%)	0.74	19 (12%) 5 4	95, 141, 190, 199	0
3	F	147/152 (96%)	0.44	9 (6%) 25 23	93, 140, 190, 198	0
3	I	147/152 (96%)	1.21	41 (27%) 1 1	96, 141, 190, 199	0
All	All	1870/1929 (96%)	0.43	139 (7%) 17 17	60, 111, 169, 199	0

The worst 5 of 139 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	880	PHE	10.8
1	G	101	GLU	6.5
2	H	1458	MET	6.0
2	H	1457	ASN	5.4
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 [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.