



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:24 PM GMT

PDB ID : 4XMN  
Title : Structure of the yeast coat nucleoporin complex, space group P212121  
Authors : Stuwe, T.; Correia, A.R.; Lin, D.H.; Paduch, M.; Lu, V.T.; Kossiakoff, A.A.;  
Hoelz, A.  
Deposited on : 2015-01-14  
Resolution : 7.60 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

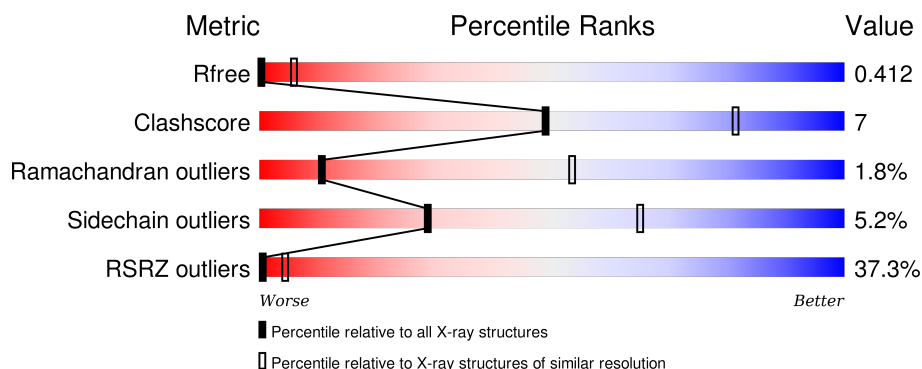
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 7.60 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	1015 (11.50-3.66)
Clashscore	102246	1064 (11.50-3.70)
Ramachandran outliers	100387	1036 (11.50-3.66)
Sidechain outliers	100360	1006 (11.50-3.66)
RSRZ outliers	91569	1014 (11.50-3.66)

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	297	
2	B	652	
3	F	454	
4	E	1045	
5	D	685	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	L	217	<div><div></div><div>41%</div><div>81%</div><div>13%</div><div></div><div></div></div>
7	H	267	<div><div></div><div>41%</div><div>68%</div><div>10%</div><div></div><div>19%</div></div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 19816 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 Protein transport protein SEC13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	274	Total	C	N	O	S	0	0	0
			2160	1379	369	409	3			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	initiating methionine	UNP Q04491

- Molecule 2 is a protein called Nucleoporin NUP145.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	503	Total	C	N	O	S	Se	0	0	0
			3765	2393	640	722	6	4			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	61	MSE	-	initiating methionine	UNP P49687
B	62	GLY	-	expression tag	UNP P49687
B	63	SER	-	expression tag	UNP P49687
B	64	SER	-	expression tag	UNP P49687
B	65	HIS	-	expression tag	UNP P49687
B	66	HIS	-	expression tag	UNP P49687
B	67	HIS	-	expression tag	UNP P49687
B	68	HIS	-	expression tag	UNP P49687
B	69	HIS	-	expression tag	UNP P49687
B	70	HIS	-	expression tag	UNP P49687
B	71	SER	-	expression tag	UNP P49687
B	72	ASP	-	expression tag	UNP P49687
B	73	GLN	-	expression tag	UNP P49687
B	74	PRO	-	expression tag	UNP P49687

- Molecule 3 is a protein called Nucleoporin NUP84.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	F	419	Total	C	N	O	S	Se	0	0	0
			3404	2178	557	657	5	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-2	GLY	-	expression tag	UNP P52891
F	-1	PRO	-	expression tag	UNP P52891
F	0	HIS	-	expression tag	UNP P52891
F	1	MSE	-	expression tag	UNP P52891

- Molecule 4 is a protein called Nucleoporin NUP120.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	896	Total	C	N	O	S	0	0	0
			6622	4232	1099	1275	16			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	-7	MET	-	initiating methionine	UNP P35729
E	-6	HIS	-	expression tag	UNP P35729
E	-5	HIS	-	expression tag	UNP P35729
E	-4	HIS	-	expression tag	UNP P35729
E	-3	HIS	-	expression tag	UNP P35729
E	-2	HIS	-	expression tag	UNP P35729
E	-1	HIS	-	expression tag	UNP P35729
E	0	SER	-	expression tag	UNP P35729

- Molecule 5 is a protein called Nucleoporin NUP85.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	D	143	Total	C	N	O	0	0	0
			713	427	143	143			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	60	MSE	-	initiating methionine	UNP P46673
D	61	GLY	-	expression tag	UNP P46673
D	62	SER	-	expression tag	UNP P46673

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	63	SER	-	expression tag	UNP P46673
D	64	HIS	-	expression tag	UNP P46673
D	65	HIS	-	expression tag	UNP P46673
D	66	HIS	-	expression tag	UNP P46673
D	67	HIS	-	expression tag	UNP P46673
D	68	HIS	-	expression tag	UNP P46673
D	69	HIS	-	expression tag	UNP P46673
D	70	SER	-	expression tag	UNP P46673
D	71	ASP	-	expression tag	UNP P46673
D	72	GLN	-	expression tag	UNP P46673
D	744	MSE	-	expression tag	UNP P46673

- Molecule 6 is a protein called Antibody 87 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	L	209	Total	C	N	O	S	0	0	0
			1592	991	269	326	6			

- Molecule 7 is a protein called Antibody 87 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	215	Total	C	N	O	S	0	0	0
			1560	977	265	312	6			

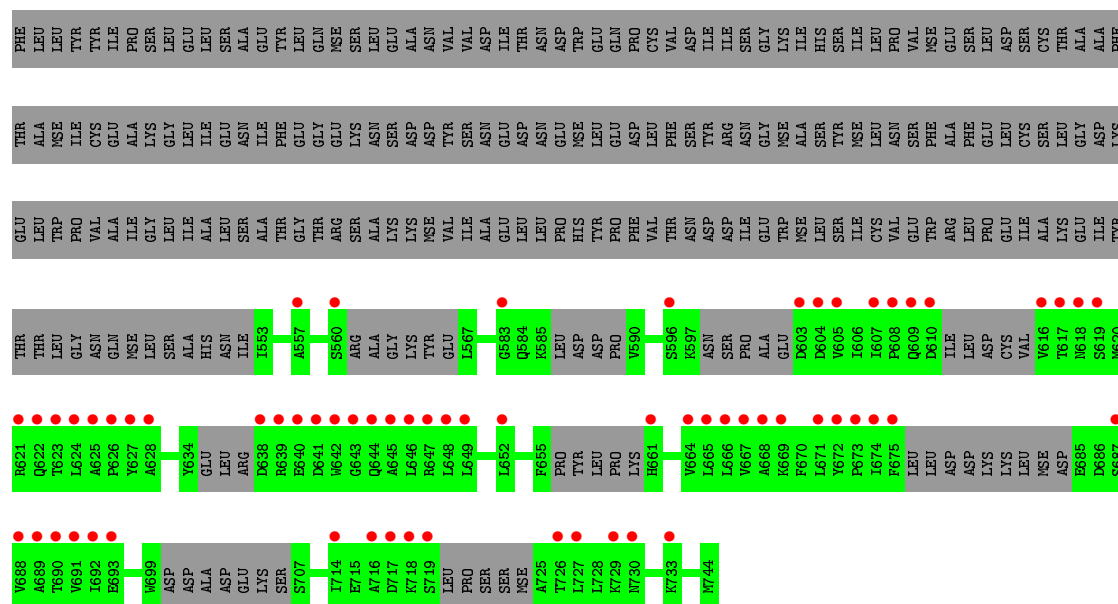




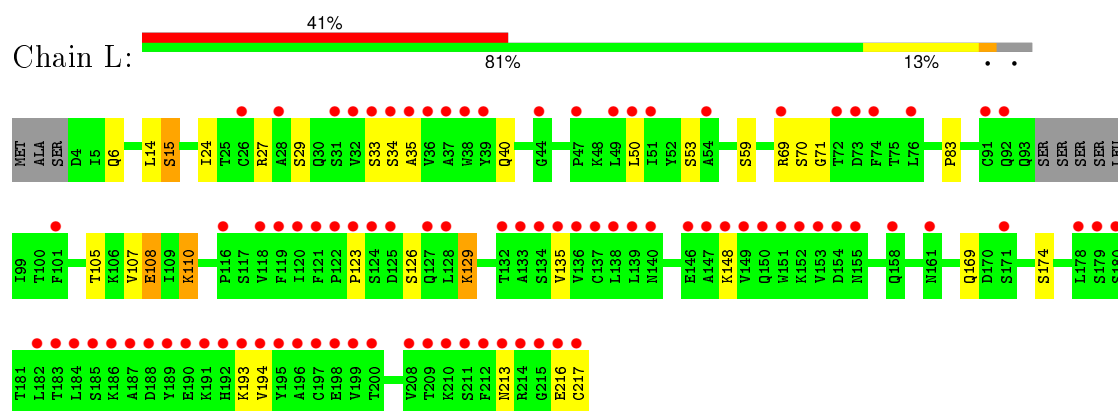




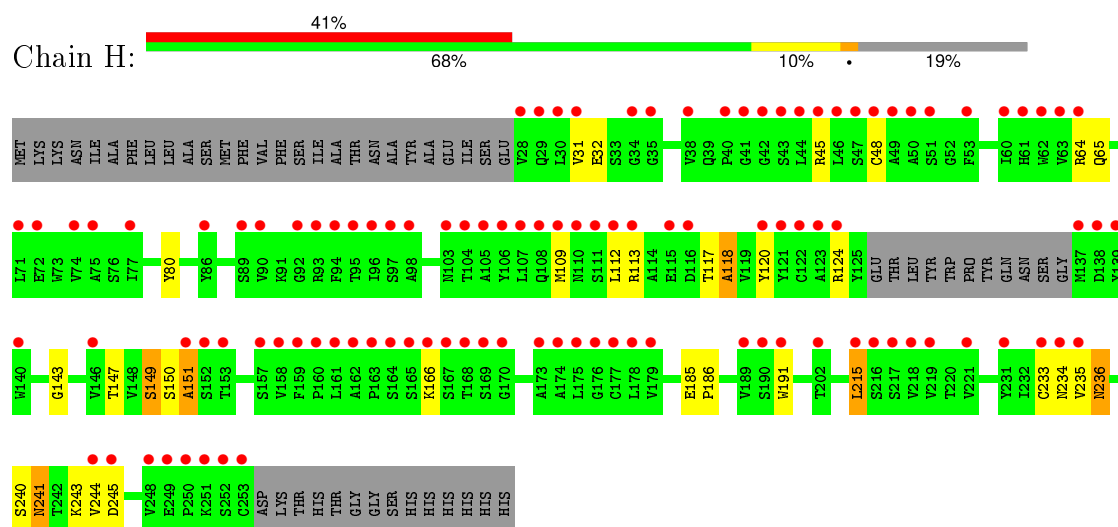
SER	PHE	THR	ASP	MSE
SER	GLU	VAL	ASP	GLY
THR	THR	LEU	ARG	SER
PHE	GLN	ASN	VAL	HIS
ARG	TYR	CYS	PHE	SER
GLU	PHE	LEU	ASN	HIS
TRP	TRP	ARG	VAL	HIS
LYS	TRP	THR	PRO	HIS
ASN	LEU	MSE	THR	HIS
LEU	LEU	TYR	ILE	HIS
VAL	ASN	PHE	GLY	SER
LEU	GLN	ILE	VAL	ASP
LYS	LEU	LEU	VAL	GLN
LEU	VAL	ASP	ASN	PRO
SER	LEU	GLY	SER	LEU
GLN	ARG	GLN	ASN	SER
ALA	GLY	ASP	PHE	TYR
PHE	LEU	VAL	ALA	GLN
GLY	LEU	GLU	LYS	ASN
GLY	LEU	GLU	GLU	MSE
SER	GLN	ASN	HIS	NSE
ALA	ALA	ARG	ASN	ALA
ALA	ILE	SER	ALA	PHE
THR	GLY	GLU	THR	ILE
ASP	CYS	PHE	VAL	THR
ILE	ILE	ILE	ASN	ALA
SER	GLU	GLU	LEU	LYS
GLY	ARG	SER	ALA	ASP
LEU	SER	LEU	MSE	LYS
ANG	ASP	LEU	GLU	TYR
ASP	LEU	ASN	ALA	LYS
TYR	LEU	TRP	ILE	LEU
ILE	PRO	ILE	LEU	TYR
GLU	TYR	ASN	ASN	PRO
ASP	LEU	ARG	GLI	VAL
PHE	SER	SER	LEU	ARG
LEU	ASP	ASP	LEU	ILE
LEU	THR	GLY	GLU	PRO
GLN	PHE	GLU	VAL	ARG
ASN	ASP	TYR	LYS	LEU
GLY	VAL	ILE	GLY	ASP
GLY	SER	GLU	ARG	THR
ASN	PHE	GLU	VAL	SER
GLN	ASP	TYR	LYS	LYS
THR	ILE	SER	VAL	GLU
TYR	SER	PHE	GLY	GLU
ARG	LEU	LYS	ASN	PHG
THR	LEU	ASP	ARG	PHE
TRP	LYS	SER	TYR	GLU
TYR	GLN	THR	GLU	ILE
GLU	TRN	ALA	LEU	TYR
SER	PRO	GLY	GLU	ASP
PHE	LYS	LYS	GLU	ASP
PHG	TYR	ASN	SER	GLY
CYS	SER	VAL	LEU	GLY



• Molecule 6: Antibody 87 light chain



• Molecule 7: Antibody 87 heavy chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.13Å 179.95Å 441.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.61 – 7.60 49.61 – 7.59	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.61-7.60) 99.6 (49.61-7.59)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.94 (at 7.37Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1809)	Depositor
R, $R_{free}$	0.318 , 0.347 0.396 , 0.412	Depositor DCC
$R_{free}$ test set	1205 reflections (11.13%)	DCC
Wilson B-factor (Å <sup>2</sup> )	597.7	Xtriage
Anisotropy	0.157	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 705.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Outliers	0 of 12029 reflections	Xtriage
$F_o, F_c$ correlation	0.69	EDS
Total number of atoms	19816	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	536.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.34% 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	0.28	0/2220	0.58	0/3028
2	B	0.29	0/3813	0.58	0/5147
3	F	0.30	0/3465	0.59	0/4693
4	E	0.27	0/6730	0.49	1/9158 (0.0%)
5	D	0.18	0/700	0.34	0/958
6	L	0.31	0/1623	0.63	0/2199
7	H	0.31	0/1594	0.71	3/2172 (0.1%)
All	All	0.29	0/20145	0.56	4/27355 (0.0%)

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
2	B	0	3
All	All	0	4

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	H	215	LEU	CA-CB-CG	5.90	128.88	115.30
4	E	818	PRO	N-CA-CB	5.82	110.29	103.30
7	H	185	GLU	C-N-CD	-5.39	108.75	120.60
7	H	245	ASP	N-CA-C	-5.18	97.00	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	41	GLU	Mainchain
2	B	368	GLU	Peptide
2	B	371	PHE	Peptide
2	B	392	TYR	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	2160	0	2096	50	0
2	B	3765	0	3484	80	0
3	F	3404	0	3378	61	1
4	E	6622	0	5907	69	1
5	D	713	0	308	0	0
6	L	1592	0	1546	23	0
7	H	1560	0	1512	17	0
All	All	19816	0	18231	274	1

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 7.

The worst 5 of 274 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)
4:E:451:ARG:HH12	6:L:69:ARG:HD3	1.27	0.97
4:E:325:ALA:HA	6:L:59:SER:HB2	1.50	0.94
3:F:314:ILE:HG22	3:F:315:LEU:HG	1.61	0.82
1:A:18:ASP:OD2	2:B:548:TYR:OH	1.97	0.80
4:E:293:LEU:HD13	4:E:297:LEU:HD12	1.63	0.79

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:329:ASN:OD1	4:E:204:TYR:OH[2_454]	1.97	0.23

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	270/297 (91%)	227 (84%)	37 (14%)	6 (2%)	8	49
2	B	485/652 (74%)	447 (92%)	28 (6%)	10 (2%)	9	50
3	F	413/454 (91%)	369 (89%)	33 (8%)	11 (3%)	6	45
4	E	858/1045 (82%)	797 (93%)	51 (6%)	10 (1%)	16	61
5	D	123/685 (18%)	123 (100%)	0	0	100	100
6	L	205/217 (94%)	195 (95%)	7 (3%)	3 (2%)	13	57
7	H	211/267 (79%)	198 (94%)	8 (4%)	5 (2%)	7	47
All	All	2565/3617 (71%)	2356 (92%)	164 (6%)	45 (2%)	11	53

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	158	SER
2	B	368	GLU
2	B	371	PHE
2	B	394	GLN
3	F	9	THR

### 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	233/251 (93%)	224 (96%)	9 (4%)	39	72
2	B	367/584 (63%)	332 (90%)	35 (10%)	11	41

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	F	387/410 (94%)	365 (94%)	22 (6%)	25	62
4	E	639/980 (65%)	615 (96%)	24 (4%)	40	73
6	L	183/191 (96%)	177 (97%)	6 (3%)	45	76
7	H	173/223 (78%)	166 (96%)	7 (4%)	38	71
All	All	1982/2639 (75%)	1879 (95%)	103 (5%)	29	65

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	58	ASP
3	F	268	SER
6	L	148	LYS
3	F	71	GLU
3	F	118	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
4	E	255	GLN
4	E	295	ASN
4	E	620	HIS
3	F	97	HIS
3	F	335	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

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	274/297 (92%)	2.81	146 (53%) 0 4	579, 633, 655, 668	0
2	B	495/652 (75%)	1.80	168 (33%) 0 5	466, 557, 629, 666	0
3	F	412/454 (90%)	1.50	125 (30%) 1 5	455, 526, 597, 626	0
4	E	896/1045 (85%)	1.60	281 (31%) 1 5	424, 509, 574, 666	0
5	D	140/685 (20%)	2.33	65 (46%) 0 4	555, 634, 690, 699	0
6	L	209/217 (96%)	2.10	90 (43%) 0 4	422, 477, 527, 558	0
7	H	215/267 (80%)	2.47	109 (50%) 0 4	451, 495, 558, 613	0
All	All	2641/3617 (73%)	1.90	984 (37%) 0 4	422, 522, 644, 699	0

The worst 5 of 984 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	L	122	PRO	15.3
2	B	169	SER	14.3
1	A	104	SER	13.2
4	E	29	SER	12.9
6	L	123	PRO	12.7

### 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.