wwPDB X-ray Structure Validation Summary Report (i)

Nov 6, 2023 – 08:23 PM EST

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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

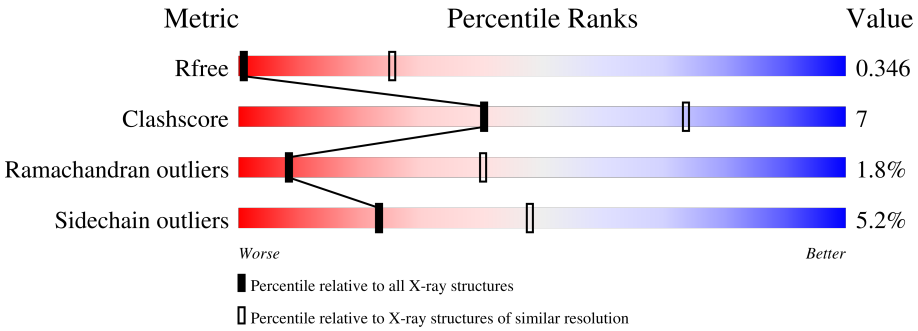
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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}	130704	1004 (10.00-3.90)
Clashscore	141614	1069 (10.00-3.90)
Ramachandran outliers	138981	1002 (10.00-3.90)
Sidechain outliers	138945	1002 (10.00-3.86)

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 of 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	297	66% 24% • 8%
2	B	652	58% 17% • 23%
3	F	454	70% 19% • 8%
4	E	1045	73% 12% • 14%
5	D	685	21% 79%
6	L	217	81% 13% • •
7	H	267	68% 10% • 19%

2 Entry composition (i)

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.

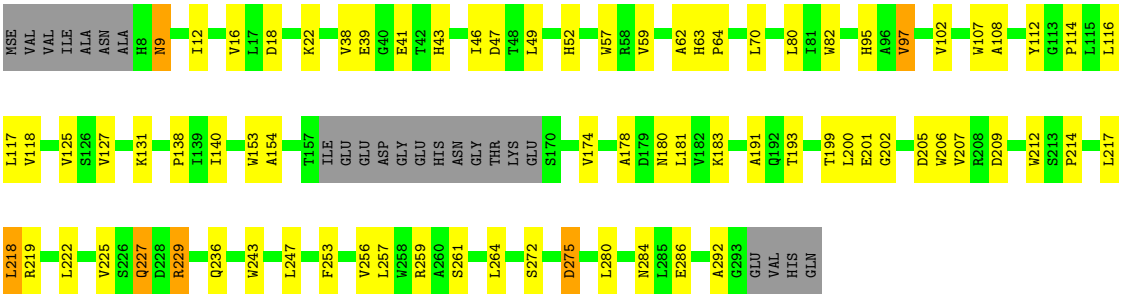
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	215	Total	C	N	O	S	0	0	0
			1560	977	265	312	6			

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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. 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. 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.

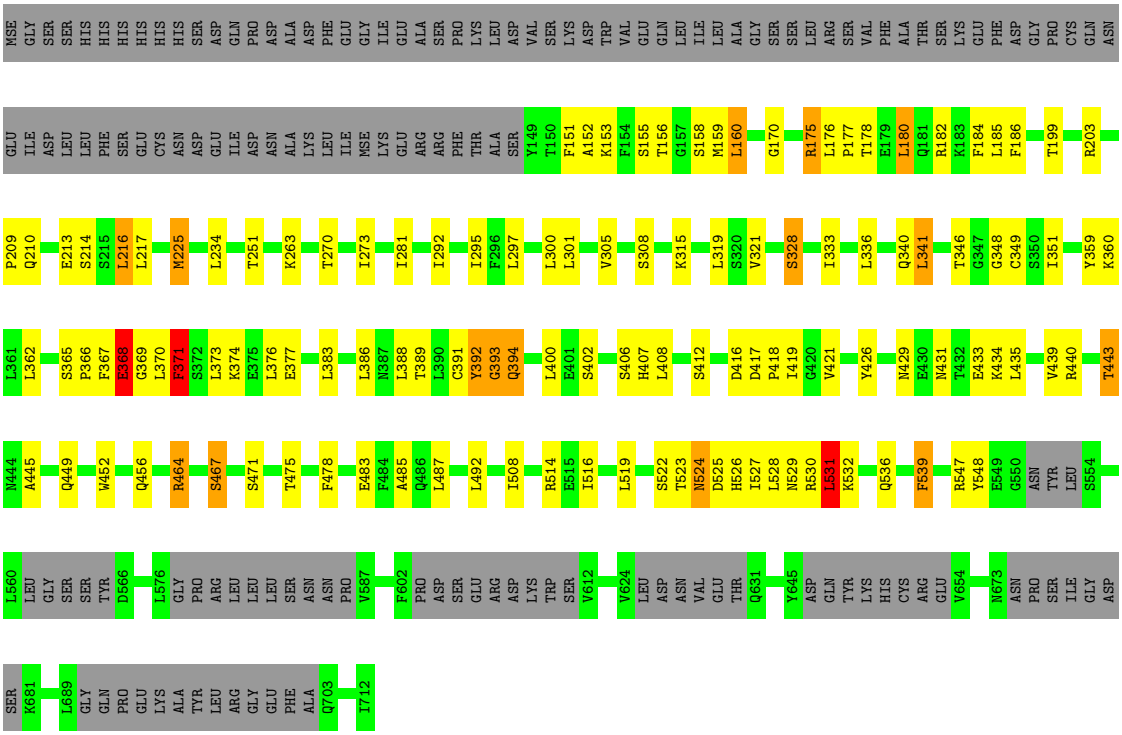
• Molecule 1: Protein transport protein SEC13

Chain A:



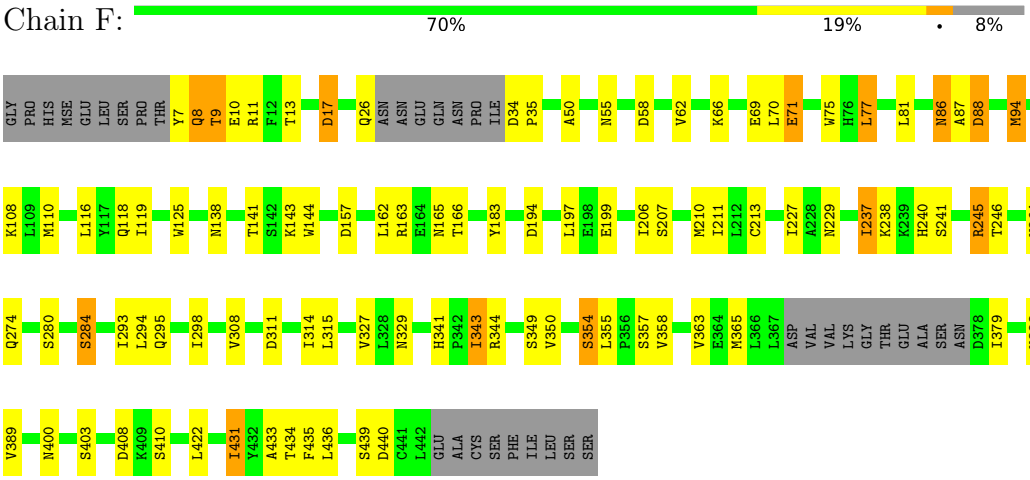
• Molecule 2: Nucleoporin NUP145

Chain B:



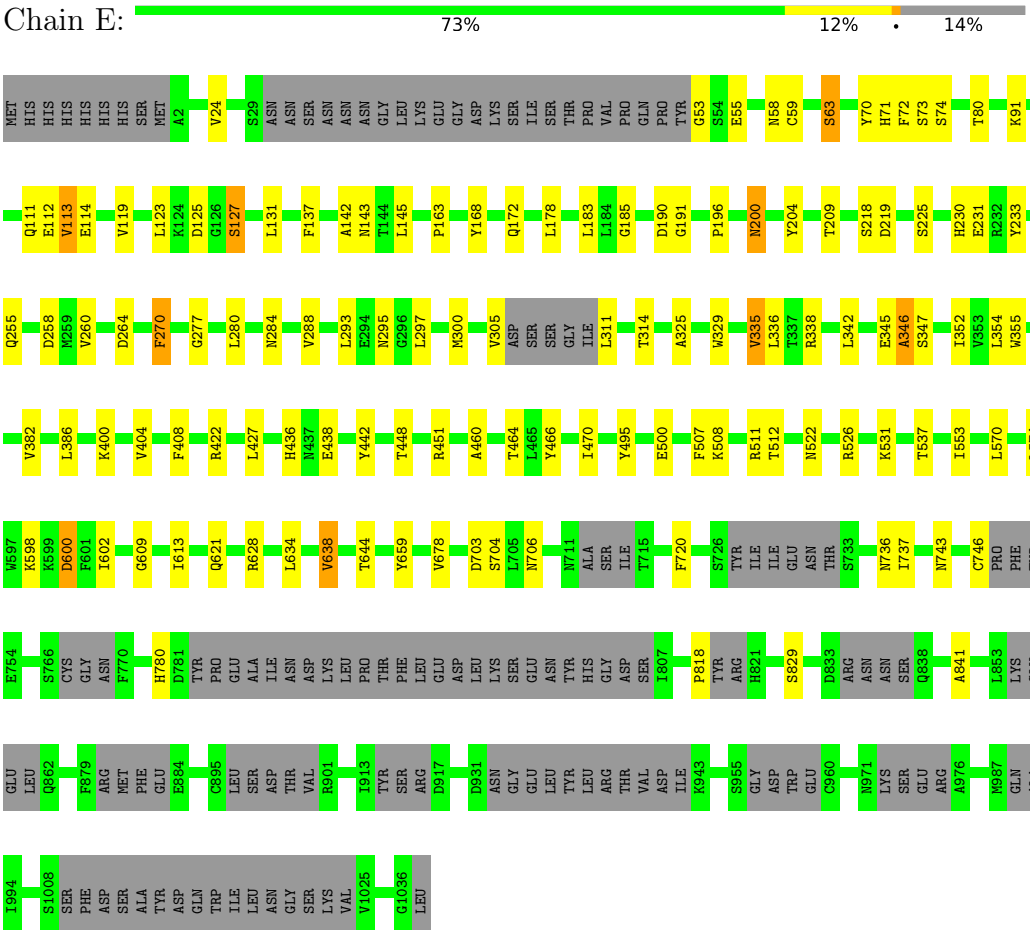
• Molecule 3: Nucleoporin NUP84

Chain F:



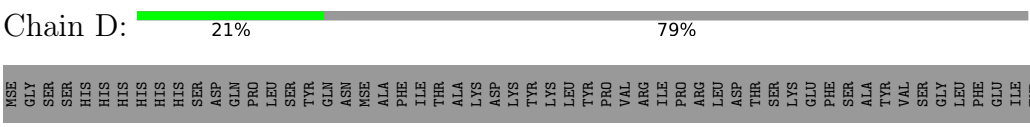
• Molecule 4: Nucleoporin NUP120

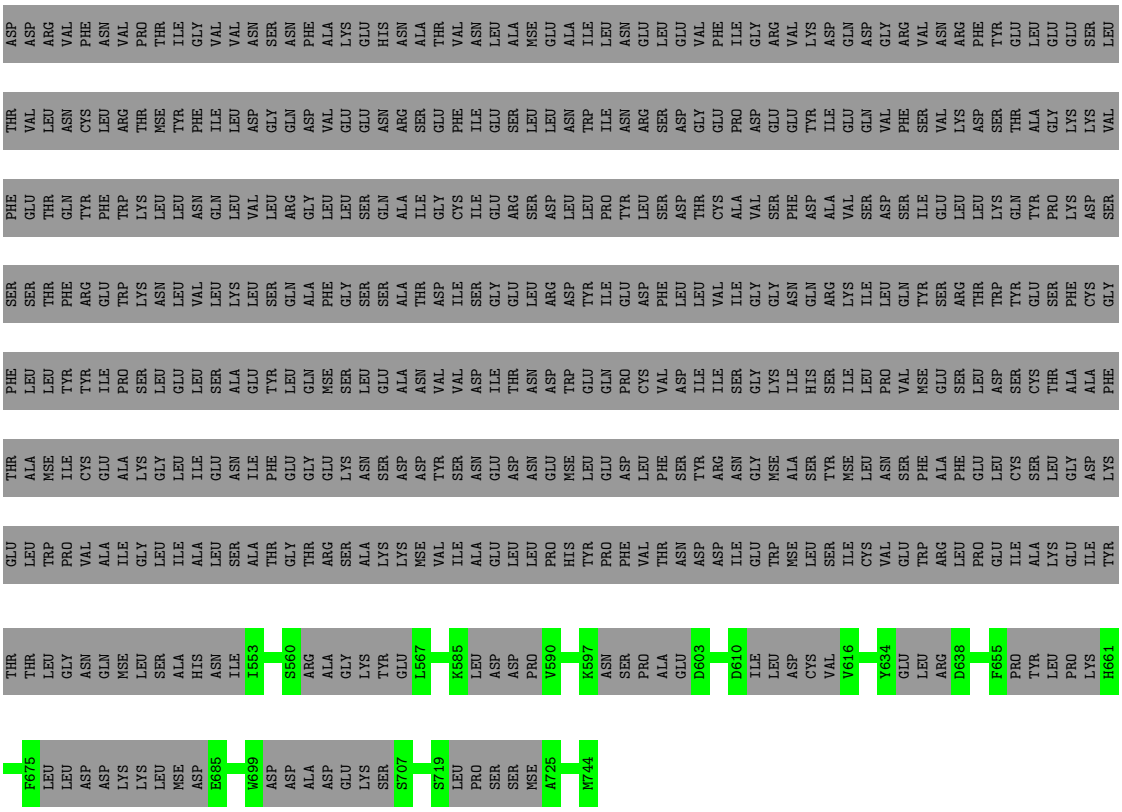
Chain E:



• Molecule 5: Nucleoporin NUP85

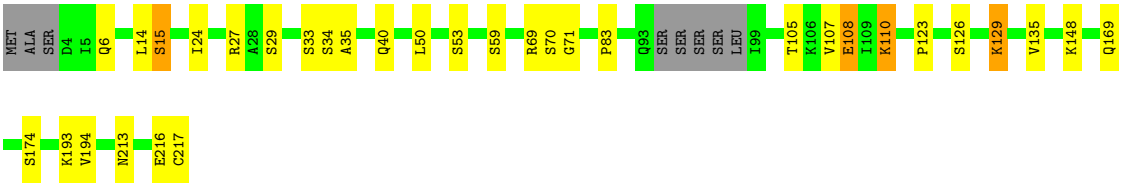
Chain D:





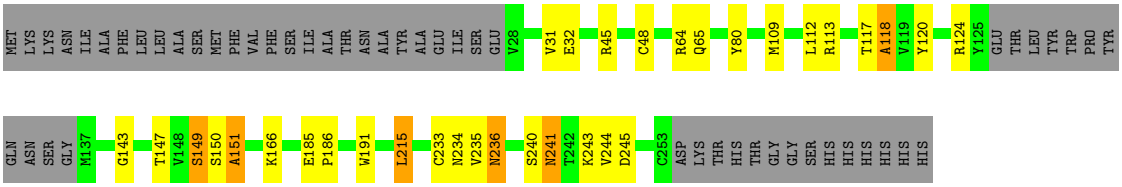
• Molecule 6: Antibody 87 light chain

Chain L:



• Molecule 7: Antibody 87 heavy chain

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	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
$< I/\sigma(I) >$ [1]	1.94 (at 7.37Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1809)	Depositor
R, R_{free}	0.318 , 0.347 0.319 , 0.346	Depositor DCC
R_{free} test set	1201 reflections (9.98%)	wwPDB-VP
Wilson B-factor (Å^2)	597.7	Xtriage
Anisotropy	0.157	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 705.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.69	EDS
Total number of atoms	19816	wwPDB-VP
Average B, all atoms (Å^2)	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.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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(^{o})	Ideal(^{o})
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 (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	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 (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	270/297 (91%)	227 (84%)	37 (14%)	6 (2%)	6 35
2	B	485/652 (74%)	447 (92%)	28 (6%)	10 (2%)	7 36
3	F	413/454 (91%)	369 (89%)	33 (8%)	11 (3%)	5 31
4	E	858/1045 (82%)	797 (93%)	51 (6%)	10 (1%)	13 50
5	D	123/685 (18%)	123 (100%)	0	0	100 100
6	L	205/217 (94%)	195 (95%)	7 (3%)	3 (2%)	10 46
7	H	211/267 (79%)	198 (94%)	8 (4%)	5 (2%)	6 33
All	All	2565/3617 (71%)	2356 (92%)	164 (6%)	45 (2%)	8 40

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 (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	233/251 (93%)	224 (96%)	9 (4%)	32 56
2	B	367/584 (63%)	332 (90%)	35 (10%)	8 27

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	F	387/410 (94%)	365 (94%)	22 (6%)	20 45
4	E	639/980 (65%)	615 (96%)	24 (4%)	33 57
6	L	183/191 (96%)	177 (97%)	6 (3%)	38 61
7	H	173/223 (78%)	166 (96%)	7 (4%)	31 55
All	All	1982/2639 (75%)	1879 (95%)	103 (5%)	23 48

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

Mol	Chain	Res	Type
3	F	268	SER
4	E	200	ASN
7	H	166	LYS
3	F	284	SER
3	F	431	ILE

Sometimes 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	620	HIS
6	L	213	ASN
4	E	706	ASN
4	E	255	GLN
4	E	575	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 monosaccharides 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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.