wwPDB X-ray Structure Validation Summary Report (i)

Jul 14, 2016 – 08:44 PM EDT

PDB ID	:	4XMM
Title	:	Structure of the yeast coat nucleoporin complex, space group C2
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	:	7.38 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 (i) symbol.

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

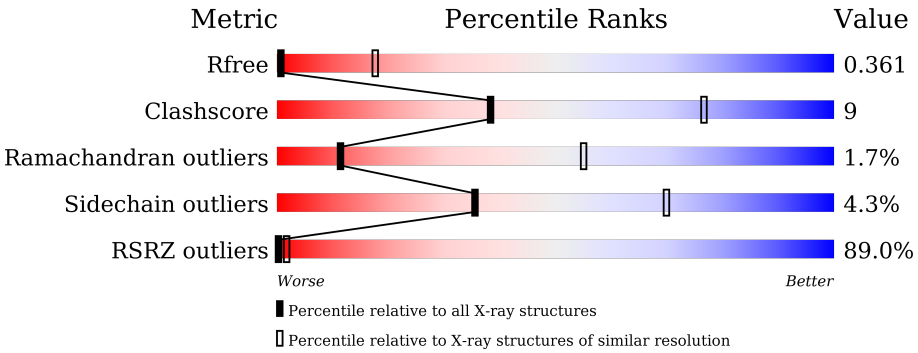
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027790
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027790

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 7.38 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	1014 (9.50-3.66)
Clashscore	102246	1063 (10.00-3.70)
Ramachandran outliers	100387	1035 (9.50-3.66)
Sidechain outliers	100360	1005 (9.50-3.66)
RSRZ outliers	91569	1013 (9.50-3.66)

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 >=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 <=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	92% 64% 25% • 8%
2	B	652	70% 56% 19% • 22%
3	C	349	83% 61% 26% • 12%
4	D	715	81% 66% 19% • 13%
5	E	1045	69% 71% 14% • 14%
6	F	454	82% 68% 21% • 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	H	271	
8	L	217	

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 26139 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			

- Molecule 2 is a protein called Nucleoporin NUP145.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	511	Total	C	N	O	S	0	0	0
			3805	2417	648	730	10			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	61	MET	-	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 SEH1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	307	Total	C	N	O	S	0	0	0
			2438	1543	422	462	11			

- Molecule 4 is a protein called Nucleoporin NUP85.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	620	Total	C	N	O	S	0	0	0
			4535	2884	753	877	21			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	30	MET	-	initiating methionine	UNP P46673
D	31	GLY	-	expression tag	UNP P46673
D	32	SER	-	expression tag	UNP P46673
D	33	SER	-	expression tag	UNP P46673
D	34	HIS	-	expression tag	UNP P46673
D	35	HIS	-	expression tag	UNP P46673
D	36	HIS	-	expression tag	UNP P46673
D	37	HIS	-	expression tag	UNP P46673
D	38	HIS	-	expression tag	UNP P46673
D	39	HIS	-	expression tag	UNP P46673
D	40	SER	-	expression tag	UNP P46673
D	41	ASP	-	expression tag	UNP P46673
D	42	GLN	-	expression tag	UNP P46673
D	43	PRO	-	expression tag	UNP P46673

- Molecule 5 is a protein called Nucleoporin NUP120.

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

There are 9 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
E	1	THR	-	expression tag	UNP P35729

- Molecule 6 is a protein called Nucleoporin NUP84.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	419	Total	C	N	O	S	0	0	0
			3404	2178	557	657	12			

There are 3 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

- Molecule 7 is a protein called Antibody 57 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	217	Total	C	N	O	S	0	0	0
			1576	988	267	315	6			

- Molecule 8 is a protein called Antibody 57 light chain.

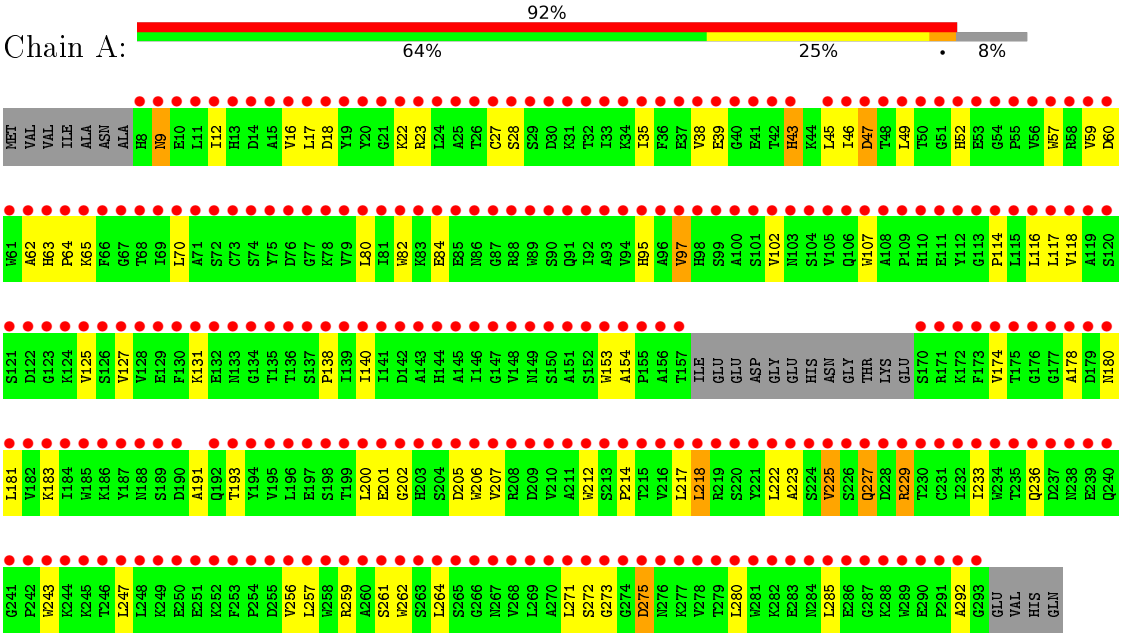
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	L	210	Total	C	N	O	S	0	0	0
			1599	996	270	327	6			

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 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.

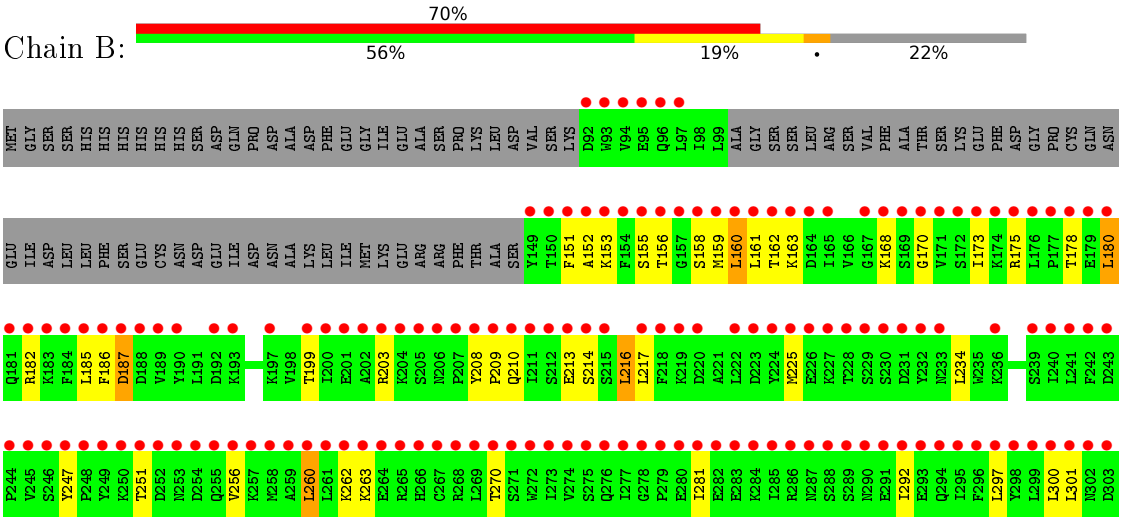
- Molecule 1: Protein transport protein SEC13

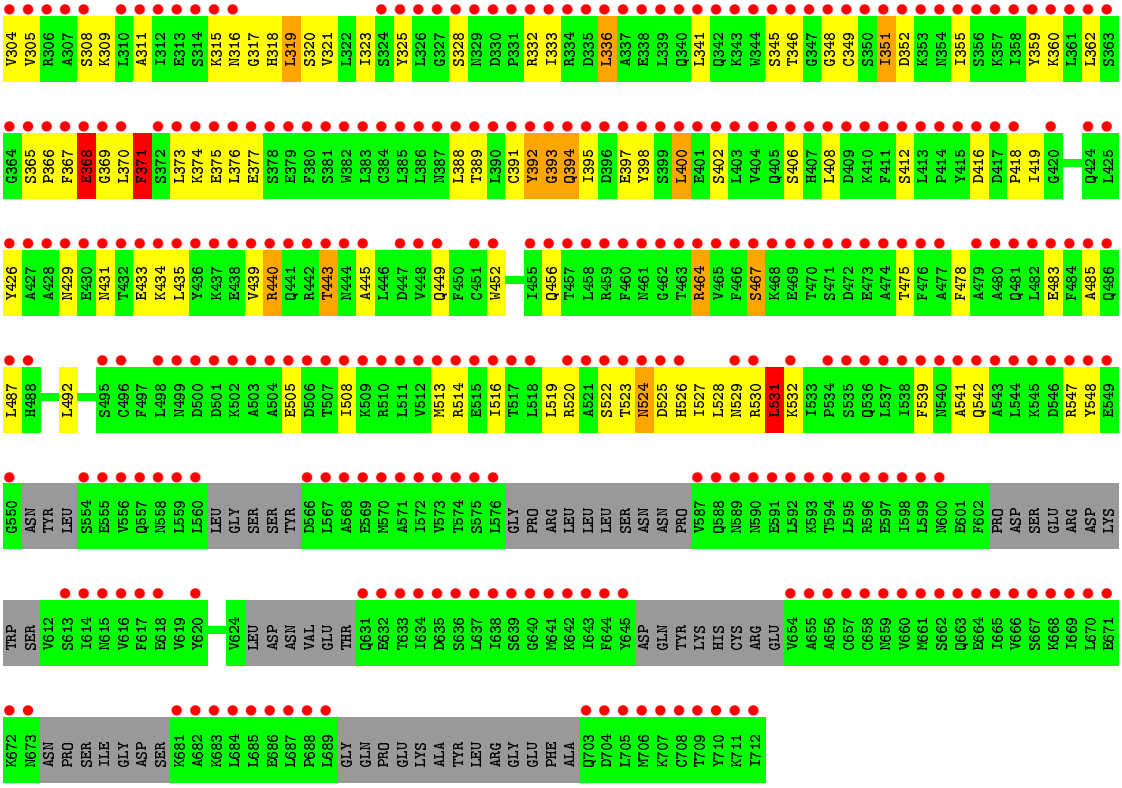
Chain A:



- Molecule 2: Nucleoporin NUP145

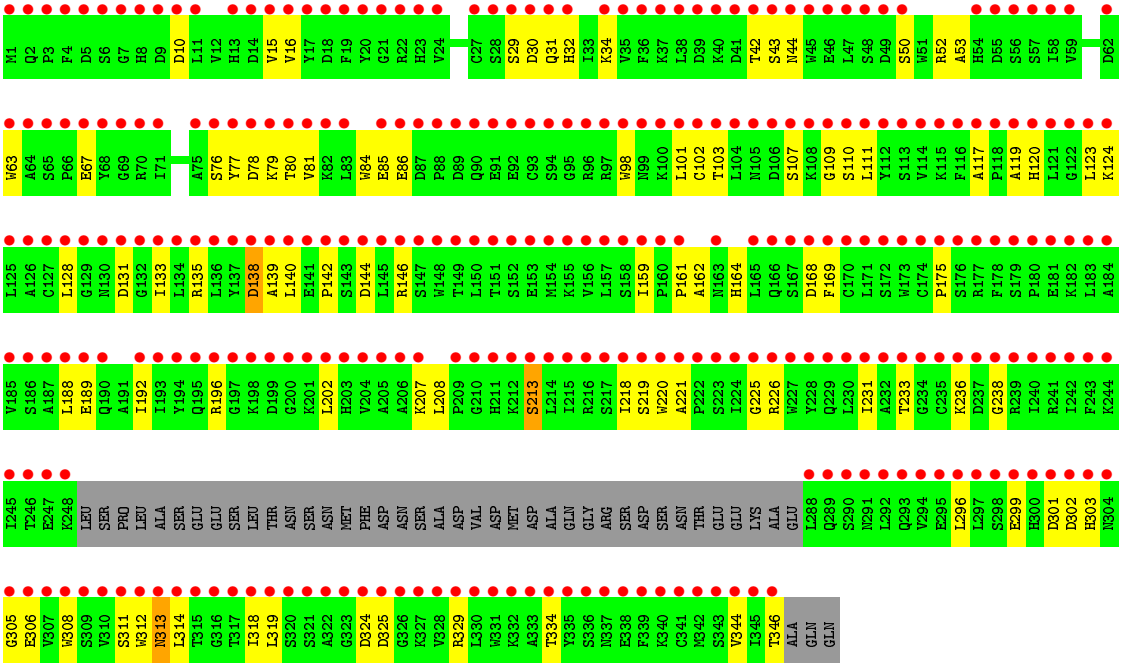
Chain B:





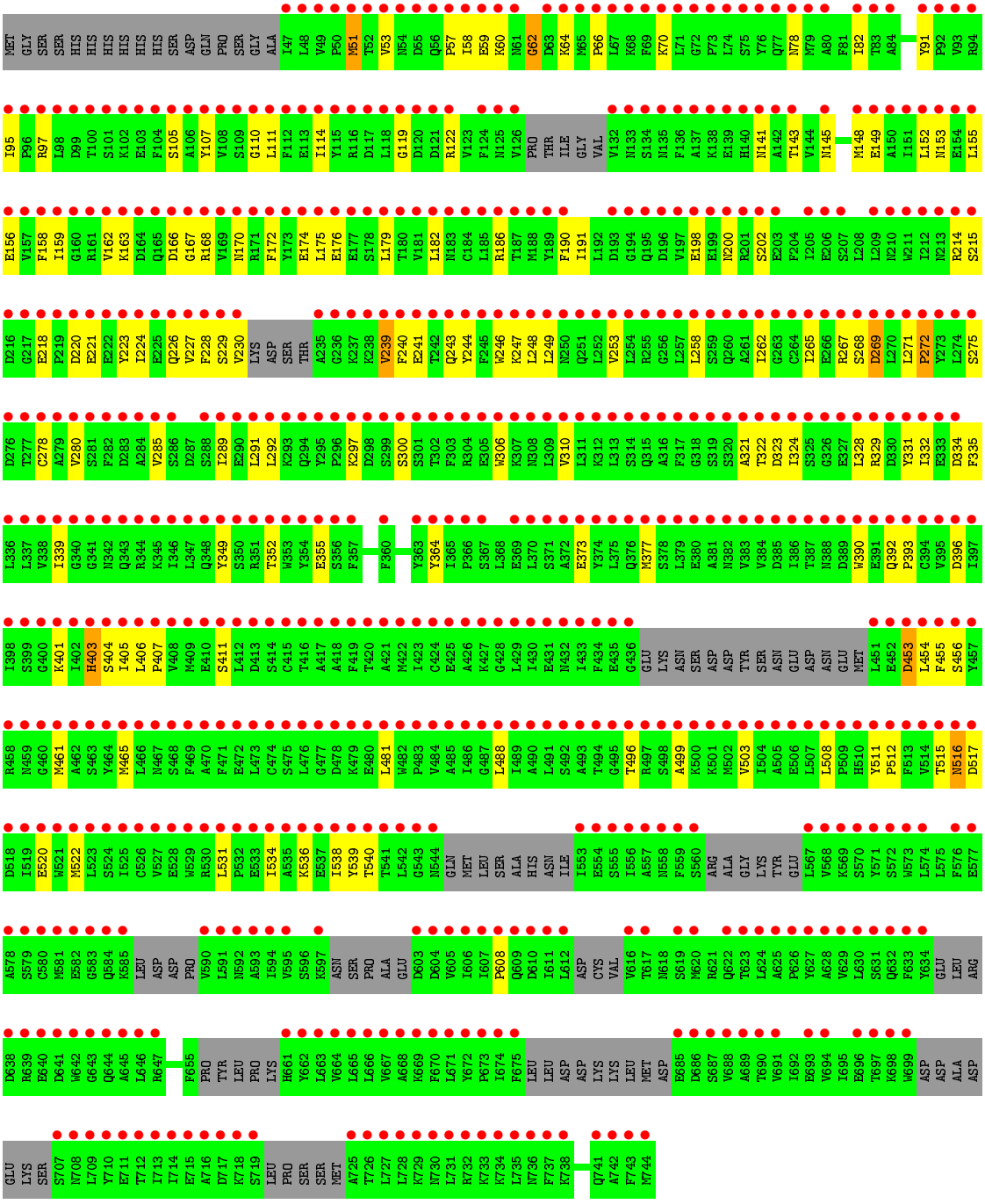
• Molecule 3: Nucleoporin SEH1

Chain C:



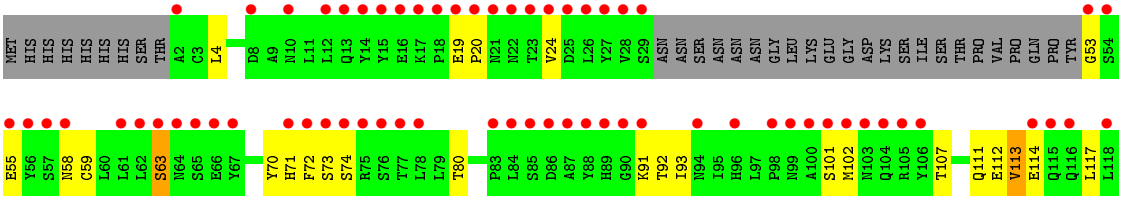
• Molecule 4: Nucleoporin NUP85

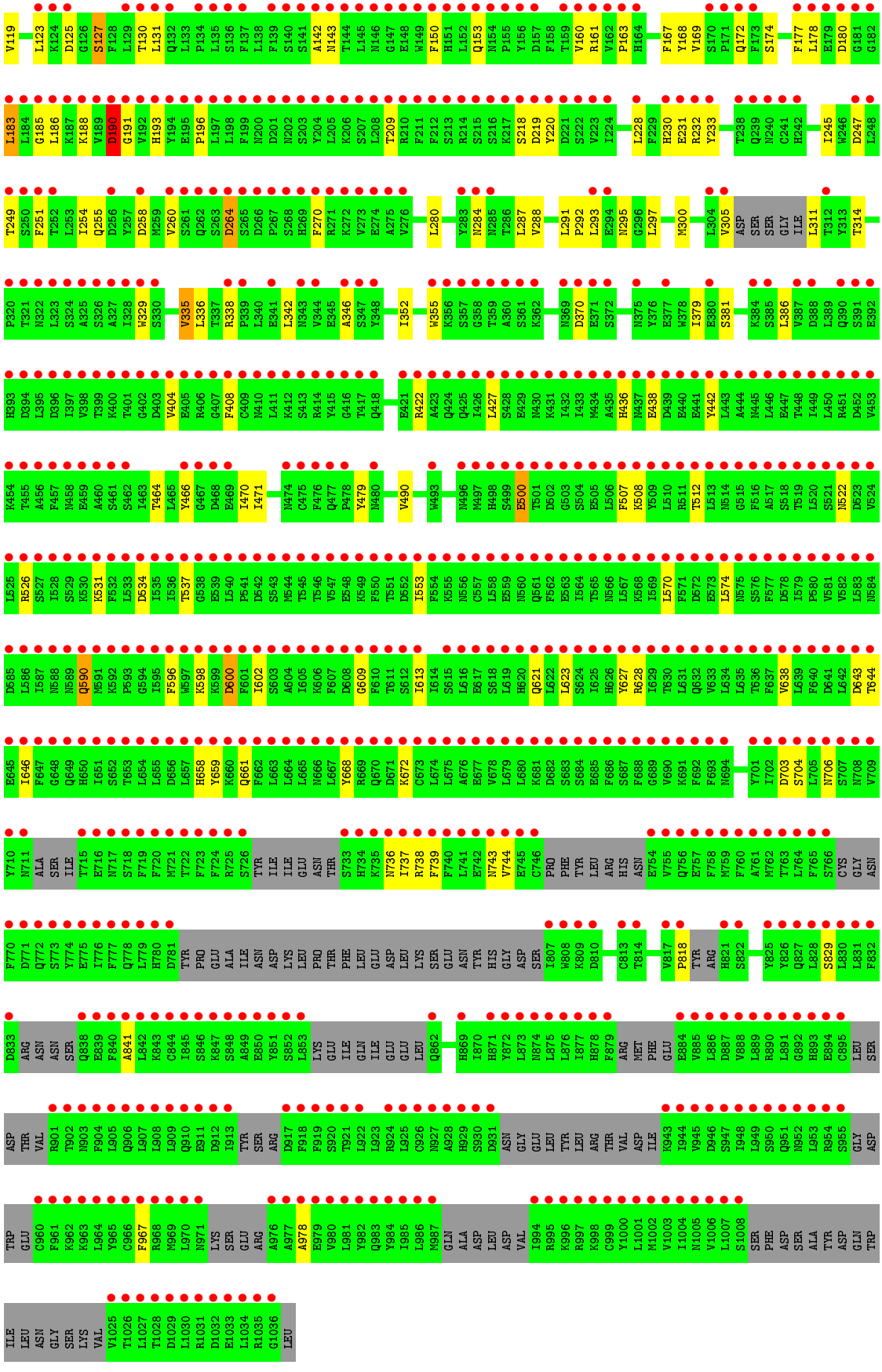
Chain D:



• Molecule 5: Nucleoporin NUP120

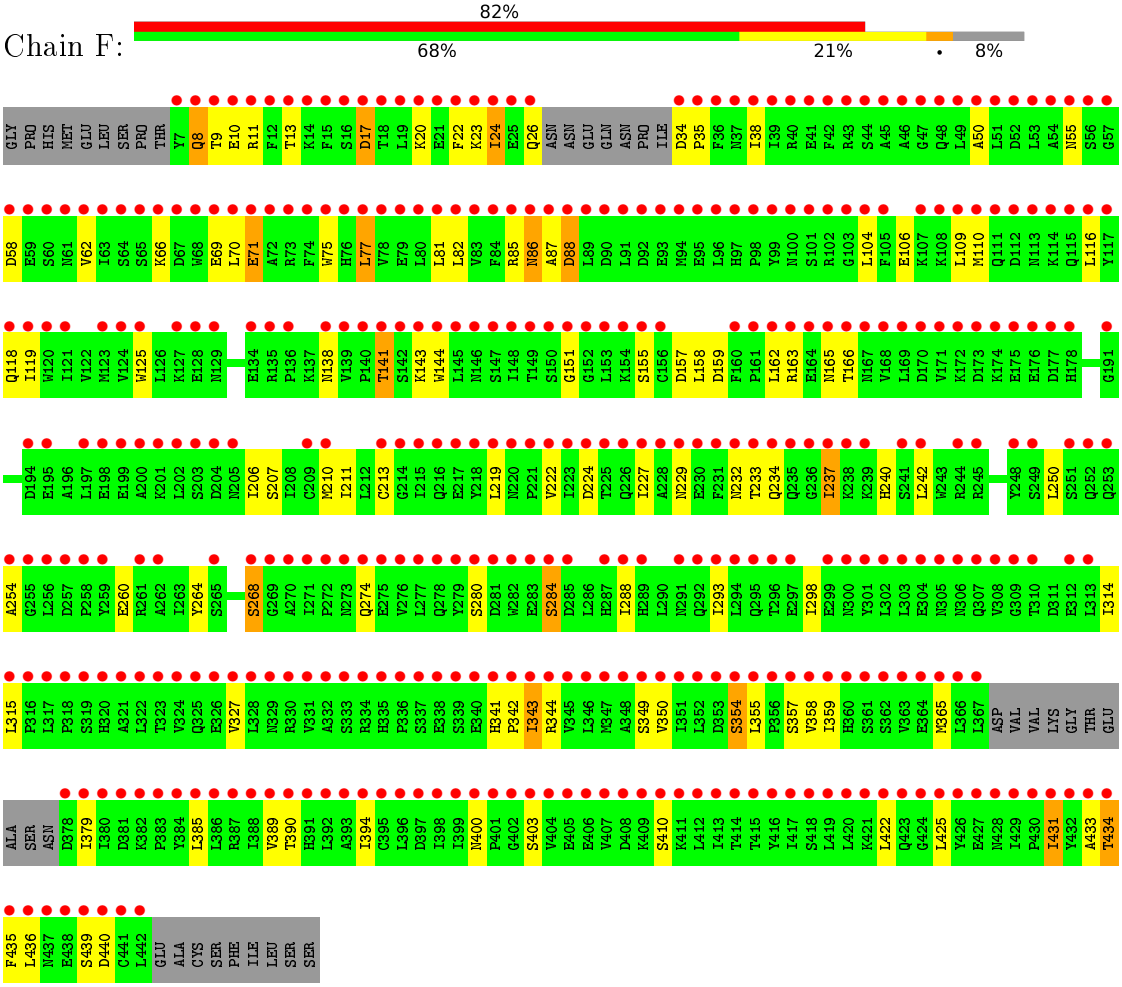
Chain E:





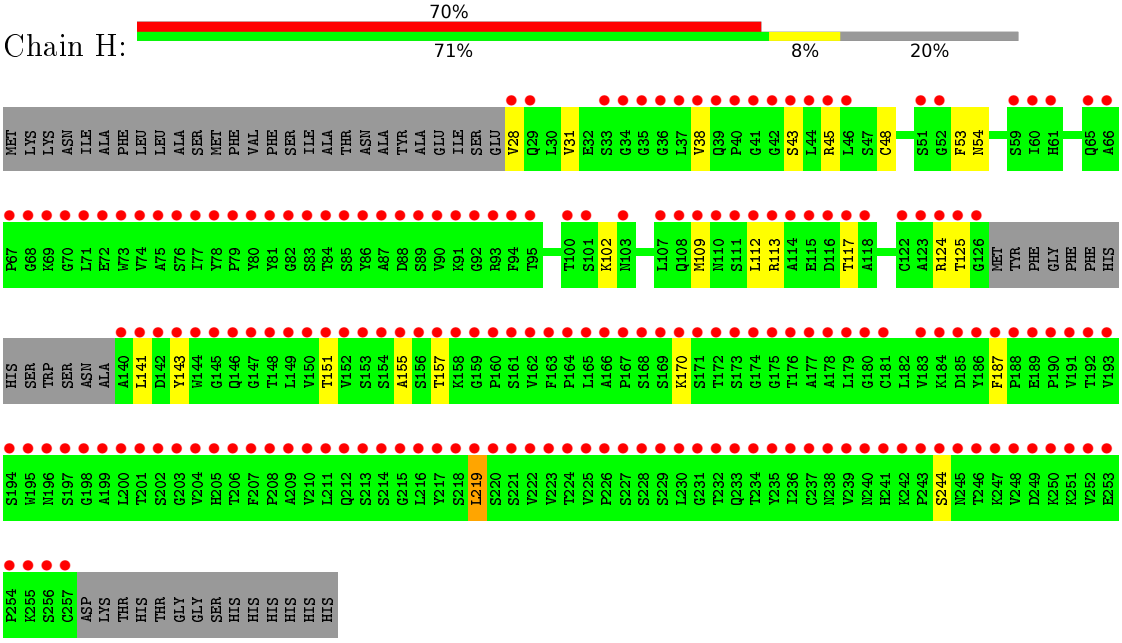
- Molecule 6: Nucleoporin NUP84

Chain F:



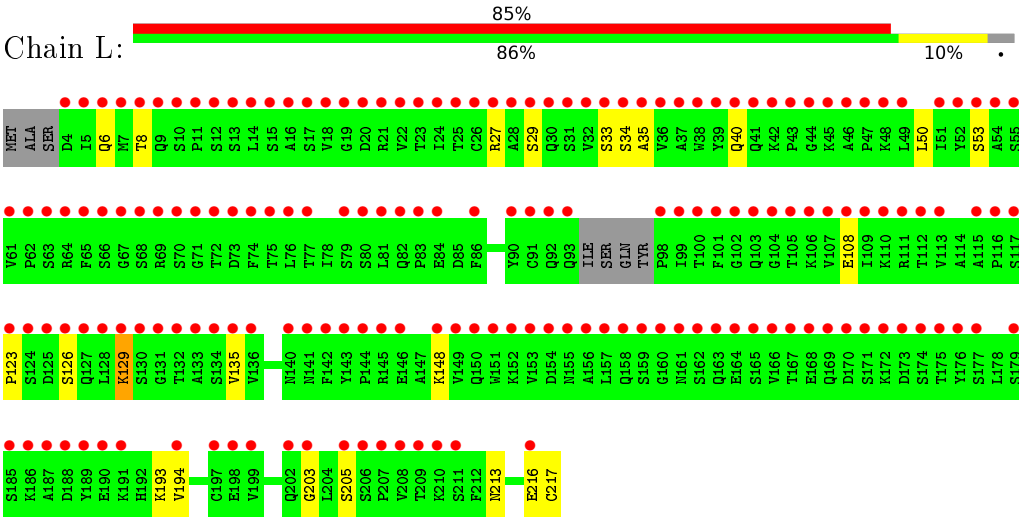
- Molecule 7: Antibody 57 heavy chain

Chain H:



• Molecule 8: Antibody 57 light chain

Chain L:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	210.65Å 186.30Å 199.57Å 90.00° 100.85° 90.00°	Depositor
Resolution (Å)	67.52 – 7.38 68.01 – 7.38	Depositor EDS
% Data completeness (in resolution range)	99.5 (67.52-7.38) 99.7 (68.01-7.38)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.75 (at 7.41Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1809)	Depositor
R, R_{free}	0.330 , 0.353 0.328 , 0.361	Depositor DCC
R_{free} test set	1021 reflections (9.98%)	DCC
Wilson B-factor (Å^2)	734.6	Xtriage
Anisotropy	0.358	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.50 , -10.0	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	0.75	EDS
Total number of atoms	26139	wwPDB-VP
Average B, all atoms (Å^2)	716.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.02% 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	0.32	0/2220	0.62	0/3028
2	B	0.37	0/3860	0.66	2/5224 (0.0%)
3	C	0.28	0/2499	0.64	0/3388
4	D	0.30	0/4602	0.58	2/6246 (0.0%)
5	E	0.33	0/6730	0.55	1/9158 (0.0%)
6	F	0.35	0/3472	0.64	2/4714 (0.0%)
7	H	0.31	0/1610	0.62	1/2194 (0.0%)
8	L	0.29	0/1631	0.60	0/2210
All	All	0.32	0/26624	0.60	8/36162 (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
2	B	0	2
5	E	0	2
6	F	0	1
All	All	0	5

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	167	GLY	N-CA-C	-6.40	97.09	113.10
4	D	608	PRO	N-CA-CB	6.10	110.62	103.30
2	B	187	ASP	CB-CG-OD2	6.04	123.73	118.30
7	H	219	LEU	CA-CB-CG	6.04	129.19	115.30
5	E	818	PRO	N-CA-CB	5.78	110.23	103.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	371	PHE	Peptide
2	B	392	TYR	Peptide
5	E	190	ASP	Mainchain
5	E	264	ASP	Sidechain
6	F	151	GLY	Mainchain

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	60	0
2	B	3805	0	3499	106	0
3	C	2438	0	2378	56	0
4	D	4535	0	4073	104	0
5	E	6622	0	5907	80	0
6	F	3404	0	3378	77	2
7	H	1576	0	1532	13	0
8	L	1599	0	1554	9	2
All	All	26139	0	24417	462	2

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 9.

The worst 5 of 462 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)
5:E:537:THR:HG22	5:E:743:ASN:HA	1.54	0.86
5:E:293:LEU:HD13	5:E:297:LEU:HD12	1.59	0.85
4:D:517:ASP:OD1	7:H:54:ASN:N	2.10	0.85
4:D:159:ILE:HG12	4:D:175:LEU:HB3	1.59	0.84
4:D:156:GLU:OE2	4:D:214:ARG:NH1	2.13	0.81

All (2) 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)
6:F:233:THR:OG1	8:L:203:GLY:O[4_455]	2.05	0.15
6:F:232:ASN:O	8:L:205:SER:N[4_455]	2.08	0.12

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%)	228 (84%)	36 (13%)	6 (2%)	8	49
2	B	491/652 (75%)	451 (92%)	30 (6%)	10 (2%)	9	51
3	C	303/349 (87%)	266 (88%)	31 (10%)	6 (2%)	9	51
4	D	592/715 (83%)	532 (90%)	50 (8%)	10 (2%)	11	55
5	E	858/1045 (82%)	798 (93%)	49 (6%)	11 (1%)	15	60
6	F	413/454 (91%)	368 (89%)	34 (8%)	11 (3%)	6	45
7	H	213/271 (79%)	206 (97%)	7 (3%)	0	100	100
8	L	206/217 (95%)	198 (96%)	6 (3%)	2 (1%)	19	65
All	All	3346/4000 (84%)	3047 (91%)	243 (7%)	56 (2%)	11	55

5 of 56 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	C	43	SER

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	233/252 (92%)	224 (96%)	9 (4%)	39	72
2	B	367/594 (62%)	332 (90%)	35 (10%)	11	41
3	C	269/305 (88%)	261 (97%)	8 (3%)	48	77
4	D	424/642 (66%)	414 (98%)	10 (2%)	57	82
5	E	639/980 (65%)	616 (96%)	23 (4%)	42	74
6	F	387/418 (93%)	367 (95%)	20 (5%)	29	65
7	H	175/224 (78%)	170 (97%)	5 (3%)	50	78
8	L	184/191 (96%)	179 (97%)	5 (3%)	52	79
All	All	2678/3606 (74%)	2563 (96%)	115 (4%)	35	70

5 of 115 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	51	MET
5	E	127	SER
7	H	113	ARG
4	D	143	THR
4	D	488	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	313	ASN
4	D	467	ASN
5	E	575	ASN
2	B	542	GLN
5	E	588	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

6.1 Protein, DNA and RNA chains

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^{th} 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	274/297 (92%)	15.07	272 (99%)	0	0	629, 712, 763, 800	0
2	B	511/652 (78%)	9.33	459 (89%)	0	1	374, 655, 746, 784	0
3	C	307/349 (87%)	10.24	290 (94%)	0	1	631, 756, 814, 905	0
4	D	620/715 (86%)	9.98	581 (93%)	0	1	388, 746, 989, 1000	0
5	E	896/1045 (85%)	8.87	726 (81%)	0	2	422, 725, 816, 883	0
6	F	419/454 (92%)	12.15	373 (89%)	0	2	545, 699, 803, 877	0
7	H	217/271 (80%)	8.87	189 (87%)	0	2	528, 677, 821, 884	0
8	L	210/217 (96%)	8.38	185 (88%)	0	2	526, 676, 756, 791	0
All	All	3454/4000 (86%)	10.12	3075 (89%)	0	2	374, 711, 890, 1000	0

The worst 5 of 3075 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	14	ASP	75.7
5	E	519	THR	62.9
1	A	27	CYS	59.4
1	A	100	ALA	58.7
2	B	153	LYS	56.3

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

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

6.3 Carbohydrates

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.