Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2025 – 10:15 AM EDT

PDB ID	:	8U8C
Title	:	Crystal structure of the TREX-2 complex in complex with the N-terminal motif of Sub2
Authors	:	Xie, Y.; Ren, Y.
Deposited on	:	2023-09-17
Resolution	:	2.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

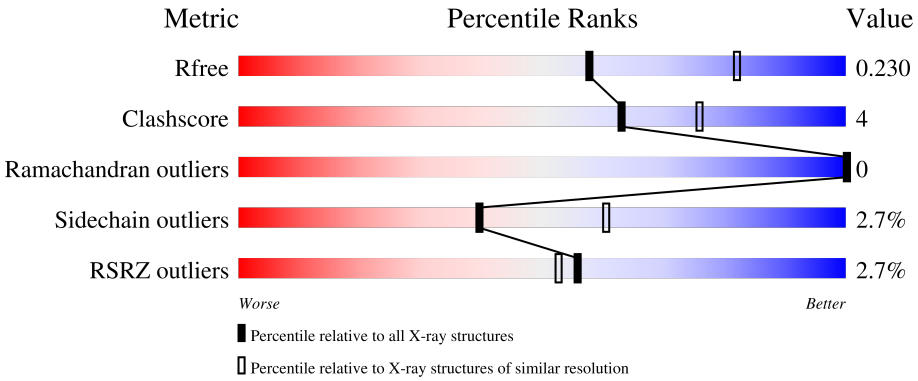
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.4

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 2.40 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}	164625	4642 (2.40-2.40)
Clashscore	180529	5218 (2.40-2.40)
Ramachandran outliers	177936	5158 (2.40-2.40)
Sidechain outliers	177891	5159 (2.40-2.40)
RSRZ outliers	164620	4642 (2.40-2.40)

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% 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	497	2% 80% 10% 10%
2	B	455	2% 87% 12%
3	C	89	4% 35% 8% 57%
4	D	55	5% 20% 80%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7881 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 Nuclear mRNA export protein SAC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	449	Total	C	N	O	S	0	0	0
			3697	2357	642	678	20			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	55	GLY	-	expression tag	UNP P46674
A	56	ALA	-	expression tag	UNP P46674
A	57	MET	-	expression tag	UNP P46674
A	58	GLY	-	expression tag	UNP P46674
A	59	SER	-	expression tag	UNP P46674

- Molecule 2 is a protein called Nuclear mRNA export protein THP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	450	Total	C	N	O	S	0	0	0
			3676	2366	640	653	17			

- Molecule 3 is a protein called 26S proteasome complex subunit SEM1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	38	Total	C	N	O	0	0	0
			338	210	52	76			

- Molecule 4 is a protein called ATP-dependent RNA helicase SUB2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	D	11	Total	C	N	O	0	0	0
			95	57	14	24			

- Molecule 5 is water.

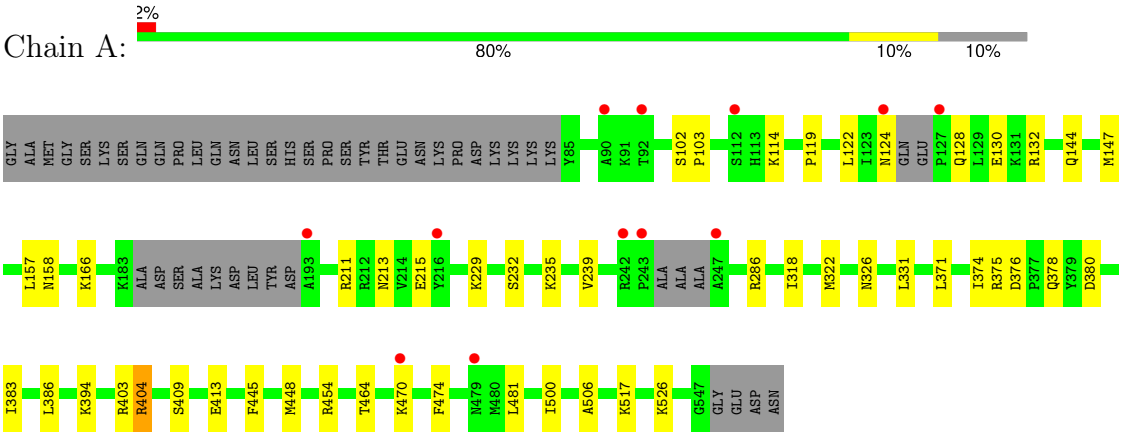
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	48	Total O 48 48	0	0
5	B	26	Total O 26 26	0	0
5	C	1	Total O 1 1	0	0

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 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.

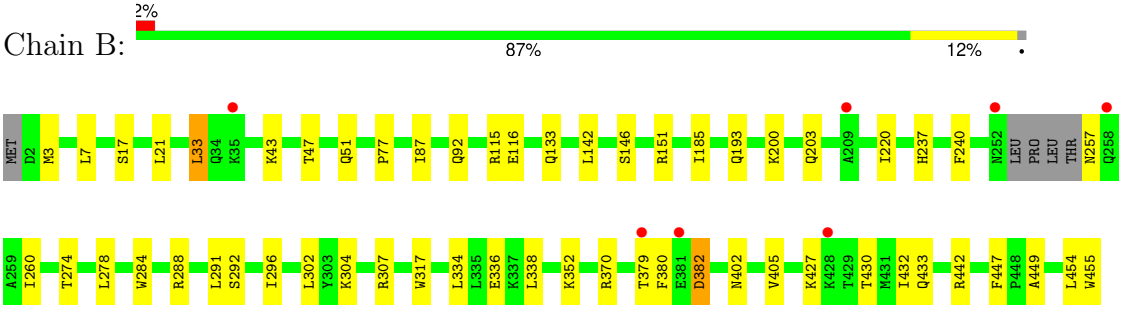
• Molecule 1: Nuclear mRNA export protein SAC3

Chain A:



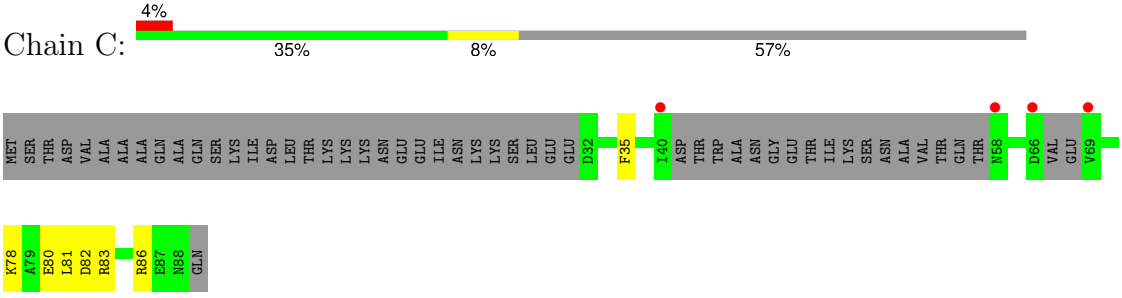
• Molecule 2: Nuclear mRNA export protein THP1

Chain B:



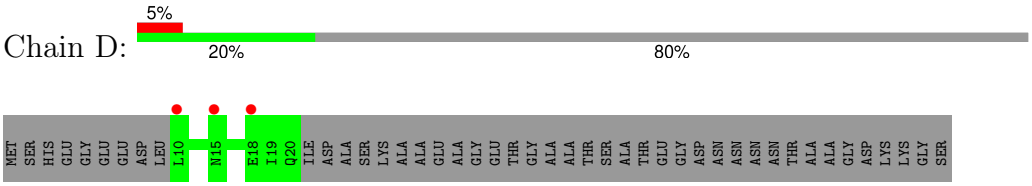
• Molecule 3: 26S proteasome complex subunit SEM1

Chain C:



• Molecule 4: ATP-dependent RNA helicase SUB2

Chain D:


5%
20%
80%
MET SER HIS GLU GLY GLU GLU ASP LEU L10 N15 E18 I19 Q20 ILE ASP ALA SER LYS ALA ALA GLU ALA GLY GLU THR GLY ALA ALA THR SER ALA THR GLU GLY ASP ASN ASN ASN ASN THR ALA ALA GLY ASP LYS LYS GLY SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	78.86Å 87.00Å 169.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.36 – 2.40 29.36 – 2.40	Depositor EDS
% Data completeness (in resolution range)	82.5 (29.36-2.40) 89.0 (29.36-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.70 (at 2.39Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.191 , 0.231 0.191 , 0.230	Depositor DCC
R_{free} test set	44458 reflections (4.65%)	wwPDB-VP
Wilson B-factor (Å^2)	32.7	Xtriage
Anisotropy	0.018	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 31.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7881	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.51% 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	0.26	0/3773	0.46	0/5089
2	B	0.24	0/3764	0.44	0/5111
3	C	0.23	0/344	0.41	0/463
4	D	0.23	0/95	0.36	0/127
All	All	0.25	0/7976	0.45	0/10790

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

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	3697	0	3723	26	0
2	B	3676	0	3705	36	0
3	C	338	0	279	4	0
4	D	95	0	79	0	0
5	A	48	0	0	0	0
5	B	26	0	0	0	0
5	C	1	0	0	0	0
All	All	7881	0	7786	60	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 4.

All (60) 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)
2:B:92:GLN:HE21	2:B:151:ARG:HH22	1.38	0.71
2:B:133:GLN:N	2:B:133:GLN:OE1	2.29	0.65
1:A:404:ARG:HD2	2:B:336:GLU:HB2	1.81	0.63
3:C:78:LYS:HA	3:C:81:LEU:HB2	1.82	0.60
2:B:203:GLN:HG3	2:B:455:TRP:CD2	2.37	0.59
2:B:3:MET:HG2	2:B:33:LEU:HG	1.87	0.56
2:B:115:ARG:HD2	2:B:116:GLU:OE2	2.06	0.56
2:B:292:SER:O	2:B:296:ILE:HD12	2.06	0.55
1:A:376:ASP:O	1:A:403:ARG:NH2	2.41	0.54
2:B:7:LEU:HD22	2:B:33:LEU:HD11	1.89	0.54
1:A:464:THR:HG21	2:B:405:VAL:HG22	1.91	0.53
2:B:257:ASN:HB2	2:B:260:ILE:HG13	1.91	0.52
2:B:21:LEU:HD13	2:B:87:ILE:HG23	1.93	0.51
1:A:371:LEU:HA	1:A:374:ILE:HD12	1.93	0.50
1:A:318:ILE:O	1:A:322:MET:HG2	2.12	0.49
1:A:383:ILE:HD12	1:A:386:LEU:HD12	1.94	0.49
1:A:448:MET:O	1:A:526:LYS:HE3	2.13	0.49
1:A:331:LEU:HD11	1:A:448:MET:SD	2.53	0.49
1:A:413:GLU:OE2	1:A:517:LYS:NZ	2.42	0.48
1:A:239:VAL:HG12	1:A:286:ARG:CZ	2.43	0.48
2:B:146:SER:HB3	2:B:185:ILE:HG21	1.94	0.48
3:C:82:ASP:O	3:C:86:ARG:HG3	2.14	0.48
2:B:240:PHE:HB2	2:B:274:THR:HG22	1.97	0.47
1:A:119:PRO:HG2	1:A:122:LEU:HD12	1.96	0.47
1:A:374:ILE:HG23	1:A:375:ARG:HG3	1.97	0.47
1:A:474:PHE:HE1	1:A:506:ALA:HB2	1.78	0.47
2:B:43:LYS:HE3	2:B:47:THR:OG1	2.15	0.47
2:B:382:ASP:N	2:B:382:ASP:OD1	2.48	0.47
2:B:203:GLN:HG3	2:B:455:TRP:CE3	2.50	0.46
1:A:378:GLN:OE1	2:B:193:GLN:HG3	2.15	0.46
2:B:304:LYS:HB3	2:B:304:LYS:HE3	1.69	0.46
1:A:331:LEU:HD21	1:A:526:LYS:HE2	1.99	0.45
1:A:454:ARG:HG2	1:A:481:LEU:HB3	1.97	0.45
1:A:474:PHE:HZ	1:A:500:ILE:HG12	1.82	0.44
2:B:220:ILE:HG22	3:C:35:PHE:HE2	1.82	0.44
2:B:142:LEU:HD23	2:B:142:LEU:HA	1.85	0.44
1:A:404:ARG:HD3	1:A:404:ARG:HA	1.54	0.44
2:B:379:THR:HG22	2:B:380:PHE:H	1.82	0.43
1:A:211:ARG:O	1:A:215:GLU:HG3	2.18	0.43
2:B:288:ARG:NH1	2:B:291:LEU:O	2.51	0.43
2:B:379:THR:HG22	2:B:380:PHE:N	2.34	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:33:LEU:HD23	2:B:33:LEU:HA	1.82	0.43
2:B:302:LEU:HB2	2:B:317:TRP:CE2	2.54	0.43
1:A:144:GLN:HA	1:A:147:MET:HE2	2.01	0.43
1:A:232:SER:HB3	1:A:235:LYS:HB2	2.00	0.43
1:A:128:GLN:NE2	1:A:130:GLU:HG3	2.34	0.43
2:B:334:LEU:O	2:B:338:LEU:HB2	2.19	0.43
1:A:409:SER:HB3	2:B:447:PHE:HZ	1.84	0.42
2:B:284:TRP:CD2	2:B:296:ILE:HG23	2.55	0.42
2:B:442:ARG:HA	2:B:442:ARG:HD2	1.88	0.42
1:A:157:LEU:HD13	1:A:213:ASN:HB3	2.01	0.42
1:A:380:ASP:OD1	1:A:404:ARG:NH2	2.42	0.42
2:B:352:LYS:HG3	2:B:432:ILE:HD13	2.02	0.42
2:B:92:GLN:HE21	2:B:151:ARG:NH2	2.10	0.41
2:B:237:HIS:CG	2:B:449:ALA:HB2	2.56	0.41
2:B:237:HIS:HA	2:B:278:LEU:HD11	2.03	0.41
1:A:103:PRO:HB3	2:B:77:PRO:HB2	2.02	0.41
2:B:302:LEU:HD22	2:B:317:TRP:CZ3	2.56	0.40
3:C:80:GLU:OE2	3:C:83:ARG:NH2	2.55	0.40
2:B:92:GLN:NE2	2:B:151:ARG:HH22	2.13	0.40

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	441/497 (89%)	438 (99%)	3 (1%)	0	100	100
2	B	446/455 (98%)	443 (99%)	3 (1%)	0	100	100
3	C	32/89 (36%)	32 (100%)	0	0	100	100
4	D	9/55 (16%)	9 (100%)	0	0	100	100
All	All	928/1096 (85%)	922 (99%)	6 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	415/454 (91%)	403 (97%)	12 (3%)	37	58
2	B	414/419 (99%)	402 (97%)	12 (3%)	37	58
3	C	37/81 (46%)	37 (100%)	0	100	100
4	D	11/40 (28%)	11 (100%)	0	100	100
All	All	877/994 (88%)	853 (97%)	24 (3%)	40	60

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	102	SER
1	A	114	LYS
1	A	124	ASN
1	A	132	ARG
1	A	158	ASN
1	A	166	LYS
1	A	229	LYS
1	A	326	ASN
1	A	394	LYS
1	A	404	ARG
1	A	445	PHE
1	A	470	LYS
2	B	17	SER
2	B	33	LEU
2	B	51	GLN
2	B	200	LYS
2	B	307	ARG
2	B	370	ARG
2	B	382	ASP
2	B	402	ASN
2	B	427	LYS
2	B	430	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	433	GLN
2	B	454	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	128	GLN
1	A	326	ASN
1	A	332	GLN
1	A	382	ASN
1	A	487	GLN
1	A	502	ASN
1	A	537	GLN
2	B	92	GLN
4	D	20	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

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	449/497 (90%)	-0.17	12 (2%) 56 53	12, 34, 64, 82	1 (0%)
2	B	450/455 (98%)	-0.18	7 (1%) 70 67	26, 38, 56, 84	0
3	C	38/89 (42%)	0.80	4 (10%) 13 11	32, 59, 81, 86	0
4	D	11/55 (20%)	1.51	3 (27%) 2 2	53, 67, 78, 78	0
All	All	948/1096 (86%)	-0.12	26 (2%) 56 53	12, 37, 64, 86	1 (0%)

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	216	TYR	4.2
2	B	209	ALA	3.8
3	C	66	ASP	3.0
3	C	58	ASN	2.9
4	D	10	LEU	2.9
1	A	112	SER	2.9
3	C	40	ILE	2.8
1	A	243	PRO	2.8
1	A	479	ASN	2.5
2	B	379	THR	2.5
1	A	90	ALA	2.5
2	B	381	GLU	2.4
1	A	92	THR	2.4
3	C	69	VAL	2.4
1	A	247	ALA	2.2
4	D	18	GLU	2.2
1	A	193	ALA	2.2
1	A	127	PRO	2.2
1	A	470	LYS	2.1
2	B	252	ASN	2.1
1	A	242	ARG	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	124	ASN	2.0
4	D	15	ASN	2.0
2	B	35	LYS	2.0
2	B	428	LYS	2.0
2	B	258	GLN	2.0

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 monosaccharides 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.