wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 06:57 PM EDT

PDB ID	:	2WP8
Title	:	yeast rrp44 nuclease
Authors	:	Basquin, J.; Bonneau, F.; Ebert, J.; Lorentzen, E.; Conti, E.
Deposited on	:	2009-08-03
Resolution	:	3.00 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:

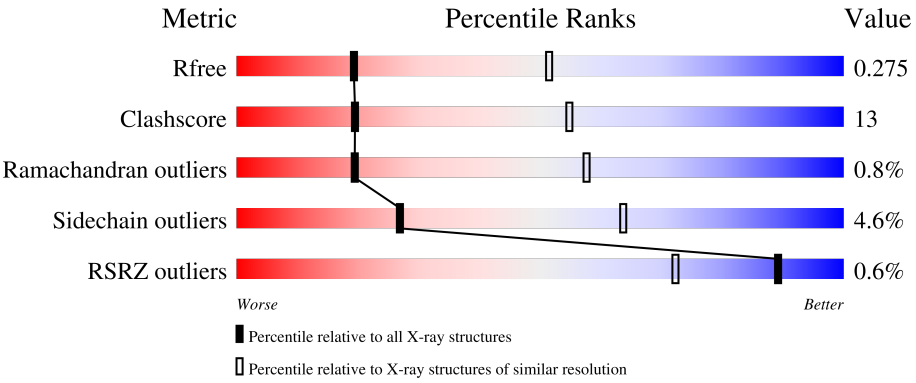
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	305	62% 20% • • 15%
2	B	246	63% 22% • 12%
3	J	977	% 71% 21% • 6%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 10816 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 EXOSOME COMPLEX COMPONENT RRP45.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	260	Total	C	N	O	S	0	0	0
			1946	1232	329	370	15			

- Molecule 2 is a protein called EXOSOME COMPLEX COMPONENT SKI6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	217	Total	C	N	O	S	0	0	0
			1647	1031	285	323	8			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	79	ASN	THR	conflict	UNP P46948

- Molecule 3 is a protein called EXOSOME COMPLEX EXONUCLEASE DIS3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	918	Total	C	N	O	S	0	0	0
			7216	4574	1250	1359	33			

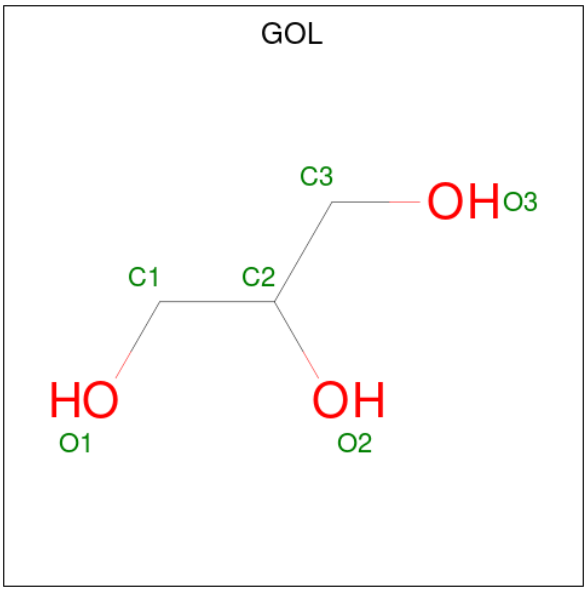
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	551	ASN	ASP	engineered mutation	UNP Q08162

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Cl	0	0
			1	1		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



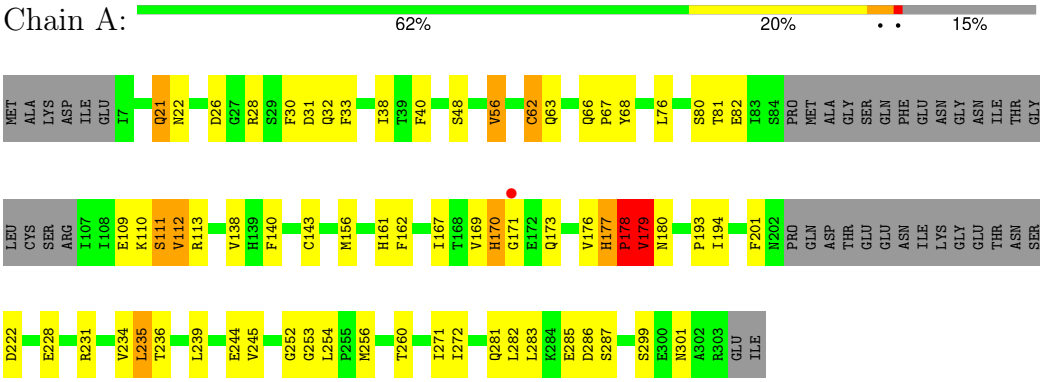
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	J	1	6	3	3	0	0

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.

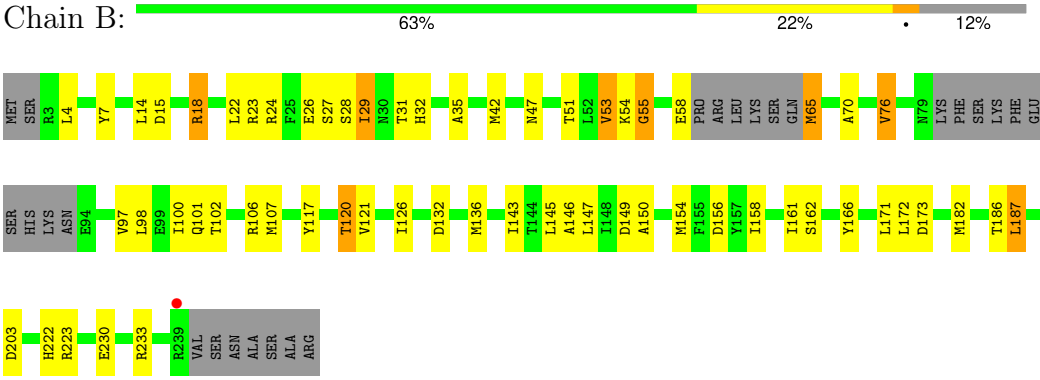
• Molecule 1: EXOSOME COMPLEX COMPONENT RRP45

Chain A:



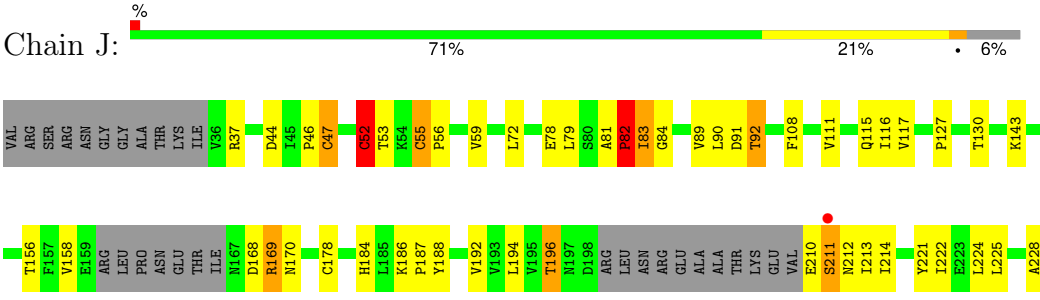
• Molecule 2: EXOSOME COMPLEX COMPONENT SKI6

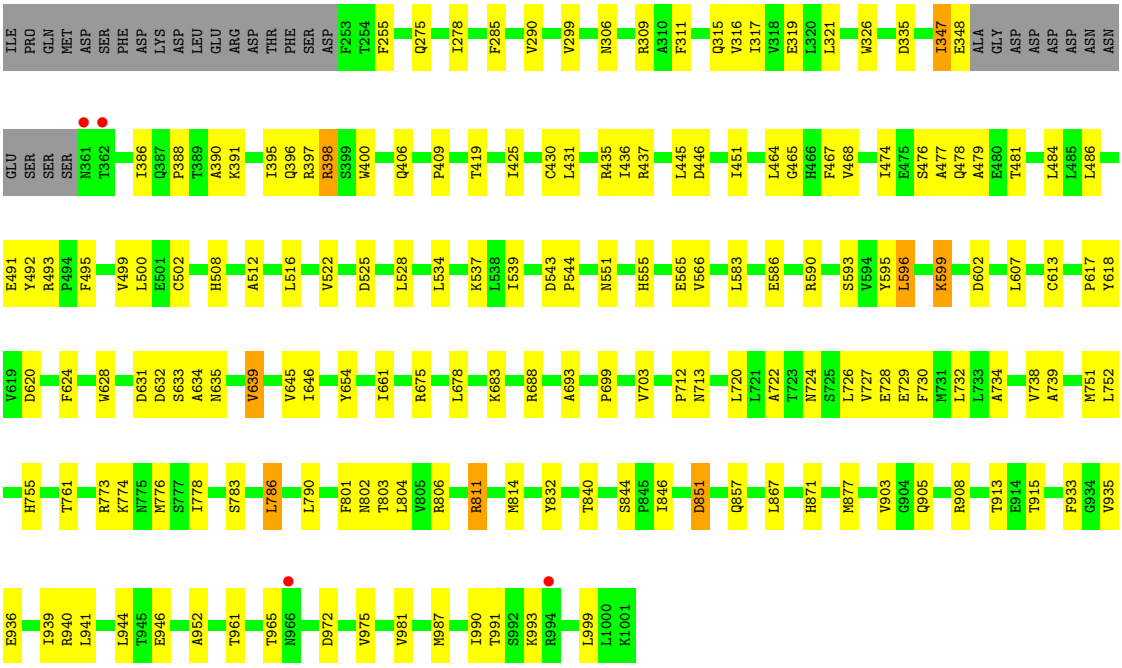
Chain B:



• Molecule 3: EXOSOME COMPLEX EXONUCLEASE DIS3

Chain J:




ILE PRO GLN MET ASP SER PHE ASP LYS ASP LEU GLU ARG ASP THR PHE SER ASP F253 T254 F255 Q275 I278 F285 V290 V299 N306 R309 A310 F311 Q315 V316 I317 V318 E319 L320 L321 W326 D335 I347 E348 ALA GLY ASP ASP ASP ASP ASN ASN
GLU SER SER SER N361 T362 I386 Q387 P388 T389 A390 K391 I395 Q396 R397 R398 S399 W400 Q406 P409 T419 I425 C430 L431 R435 I436 R437 L445 D446 I451 L464 G465 H466 F467 V468 I474 E475 S476 A477 Q478 A479 E480 T481 L484 L485 L486
E491 Y492 R493 P494 F495 V499 L500 E501 C502 H508 A512 L516 V522 D525 L528 L534 K537 L538 I539 D543 P544 N551 H555 E565 V566 L583 E586 R590 S593 V594 Y595 L596 K599 D602 L607 C613 P617 Y618
V619 D620 F624 W628 D631 D632 S633 A634 N635 V639 V645 I646 Y654 I661 R675 L678 K683 R688 A693 P699 V703 P712 N713 L720 L721 A722 T723 N724 S725 L726 V727 E728 E729 F730 M731 L732 L733 A734 V738 A739 M751 L752
H755 T761 R773 K774 N775 M776 S777 I778 S783 L786 L790 P801 N802 T803 L804 V805 R806 R811 M814 Y832 T840 S844 P845 I846 D851 Q857 L867 H871 M877 V903 G904 Q905 R908 T913 E914 T915 F933 G934 V935
E936 I939 R940 L941 L944 T945 E946 A952 T961 T965 N966 D972 V975 V981 M987 I990 T991 S992 K993 R994 L999 L1000 K1001

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	110.90Å 125.68Å 139.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 3.00 19.99 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.99-3.00) 100.0 (19.99-3.00)	Depositor EDS
R_{merge}	0.01	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.60 (at 2.98Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, R_{free}	0.207 , 0.273 0.211 , 0.275	Depositor DCC
R_{free} test set	1185 reflections (3.00%)	wwPDB-VP
Wilson B-factor (Å^2)	54.6	Xtriage
Anisotropy	0.094	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 34.7	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	0.92	EDS
Total number of atoms	10816	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.78% 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

Bond lengths and bond angles in the following residue types are not validated in this section: CL, GOL

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	0.72	0/1973	0.82	3/2668 (0.1%)
2	B	0.68	0/1665	0.81	2/2254 (0.1%)
3	J	0.67	7/7362 (0.1%)	0.81	4/10007 (0.0%)
All	All	0.68	7/11000 (0.1%)	0.81	9/14929 (0.1%)

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	3
3	J	0	3
All	All	0	6

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	J	82	PRO	C-N	-15.57	0.98	1.34
3	J	430	CYS	CB-SG	-7.98	1.68	1.82
3	J	52	CYS	CB-SG	7.58	1.95	1.82
3	J	55	CYS	CB-SG	7.04	1.94	1.82
3	J	502	CYS	CB-SG	-6.43	1.71	1.82

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	J	82	PRO	O-C-N	-23.90	84.46	122.70
3	J	82	PRO	CA-C-N	17.47	155.63	117.20
3	J	82	PRO	C-N-CA	12.62	153.25	121.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	18	ARG	NE-CZ-NH2	-6.36	117.12	120.30
2	B	18	ARG	NE-CZ-NH1	5.41	123.01	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	178	PRO	Peptide
1	A	21	GLN	Peptide
1	A	252	GLY	Peptide
3	J	168	ASP	Peptide
3	J	82	PRO	Peptide,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	1946	0	1883	56	0
2	B	1647	0	1616	59	0
3	J	7216	0	7084	164	0
4	B	1	0	0	0	0
5	J	6	0	8	1	0
All	All	10816	0	10591	274	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 13.

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)
1:A:66:GLN:O	1:A:67:PRO:N	1.76	1.19
3:J:210:GLU:CB	3:J:211:SER:HA	1.76	1.15
3:J:210:GLU:CB	3:J:211:SER:CA	2.29	1.10
3:J:52:CYS:SG	3:J:184:HIS:NE2	2.24	1.09
2:B:31:THR:HG22	2:B:32:HIS:CD2	1.90	1.06

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	252/305 (83%)	231 (92%)	15 (6%)	6 (2%)	6 29
2	B	211/246 (86%)	195 (92%)	15 (7%)	1 (0%)	29 68
3	J	908/977 (93%)	862 (95%)	42 (5%)	4 (0%)	34 72
All	All	1371/1528 (90%)	1288 (94%)	72 (5%)	11 (1%)	19 57

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	111	SER
1	A	178	PRO
1	A	179	VAL
3	J	212	ASN
2	B	55	GLY

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	204/266 (77%)	190 (93%)	14 (7%)	15 48
2	B	175/218 (80%)	164 (94%)	11 (6%)	18 51
3	J	789/881 (90%)	760 (96%)	29 (4%)	34 70
All	All	1168/1365 (86%)	1114 (95%)	54 (5%)	27 64

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	196	THR
3	J	522	VAL
3	J	913	THR
3	J	230	ASP
3	J	398	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
3	J	857	GLN
3	J	910	ASN
3	J	923	ASN
3	J	905	GLN
3	J	396	GLN

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)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	GOL	J	2002	-	5,5,5	0.58	0	5,5,5	0.92	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GOL	J	2002	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	J	2002	GOL	O1-C1-C2-C3
5	J	2002	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	J	2002	GOL	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	J	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	J	82:PRO	C	83:ILE	N	0.98

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	260/305 (85%)	-0.42	1 (0%)	92	79	18, 29, 49, 55	0
2	B	217/246 (88%)	-0.55	1 (0%)	91	75	17, 25, 43, 66	0
3	J	918/977 (93%)	-0.44	6 (0%)	87	69	14, 29, 45, 61	0
All	All	1395/1528 (91%)	-0.45	8 (0%)	89	72	14, 28, 45, 66	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	J	966	ASN	6.8
3	J	361	ASN	4.0
3	J	362	THR	3.3
3	J	211	SER	2.6
3	J	994	ARG	2.4

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 monosaccharides in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	GOL	J	2002	6/6	0.91	0.26	33,36,38,39	0
4	CL	B	1240	1/1	0.95	0.10	44,44,44,44	0

6.5 Other polymers (i)

There are no such residues in this entry.