WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 6, 2024 – 10:52 AM EDT

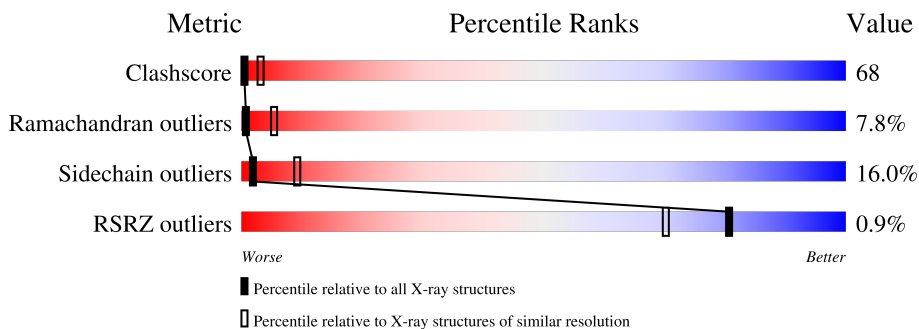|		
PDB ID	:	1LNL
Title	:	Structure of deoxygenated hemocyanin from Rapana thomasiana
Authors	:	Perbandt, M.; Guthoehrlein, E.W.; Rypniewski, W.; Idakieva, K.; Stoeva, S.; Voelter, W.; Genov, N.; Betzel, C.
Deposited on	:	2002-05-03
Resolution	:	3.30 Å(reported)

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 3.30 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	Percentile Ranks	Value
Clashscore		68
Ramachandran outliers		7.8%
Sidechain outliers		16.0%
RSRZ outliers		0.9%

Worse   *Better*

▊ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	180529	1128 (3.32-3.28)
Ramachandran outliers	177936	1125 (3.32-3.28)
Sidechain outliers	177891	1124 (3.32-3.28)
RSRZ outliers	164620	1085 (3.32-3.28)

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	408	20% 52% 24% 5%
1	B	408	21% 48% 26% 6%
1	C	408	23% 47% 24% 6%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CU	B	5503	-	-	X	-

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 9984 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 hemocyanin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	408	Total	C	N	O	S	0	0	0
			3310	2115	558	625	12			
1	B	408	Total	C	N	O	S	0	0	0
			3310	2115	558	625	12			
1	C	408	Total	C	N	O	S	3	0	0
			3310	2115	558	625	12			

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
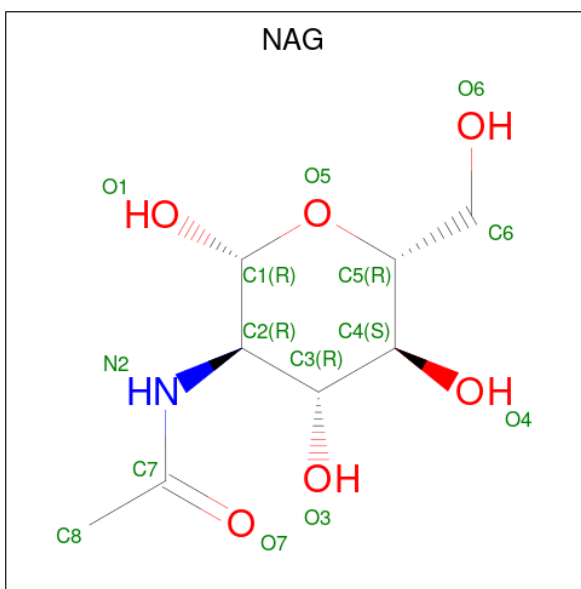
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			15	8	1	6		
2	C	1	Total	C	N	O	0	0
			15	8	1	6		
2	C	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 3 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total 2	Cu 2	0	0
3	B	4	Total 4	Cu 4	0	0
3	C	3	Total 3	Cu 3	0	0

3 Residue-property plots ⓘ

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: hemocyanin

- Molecule 1: hemocyanin

• Molecule 1: hemocyanin

Chain C:
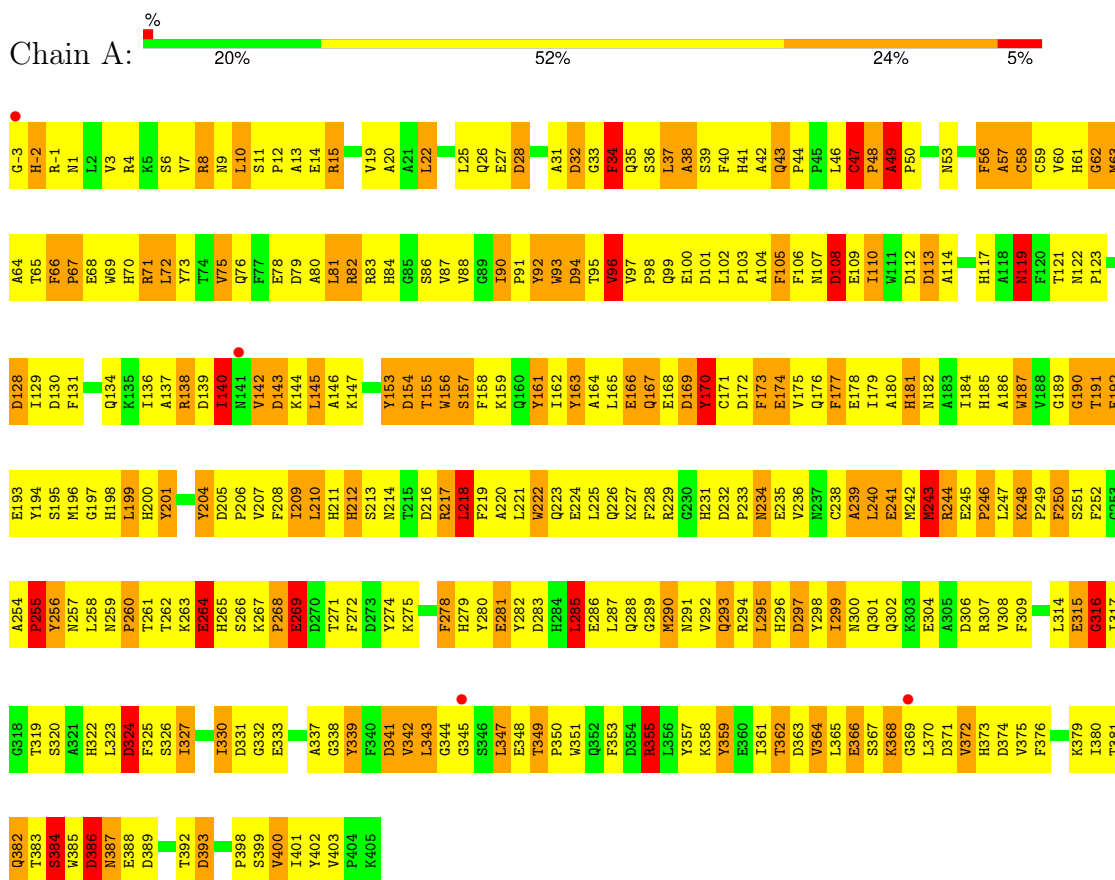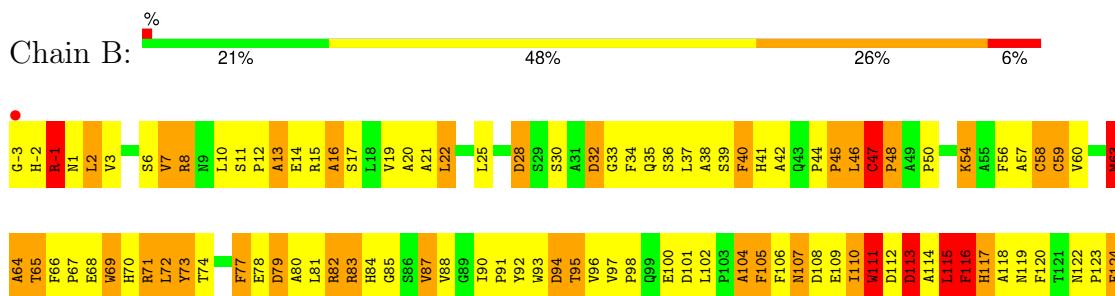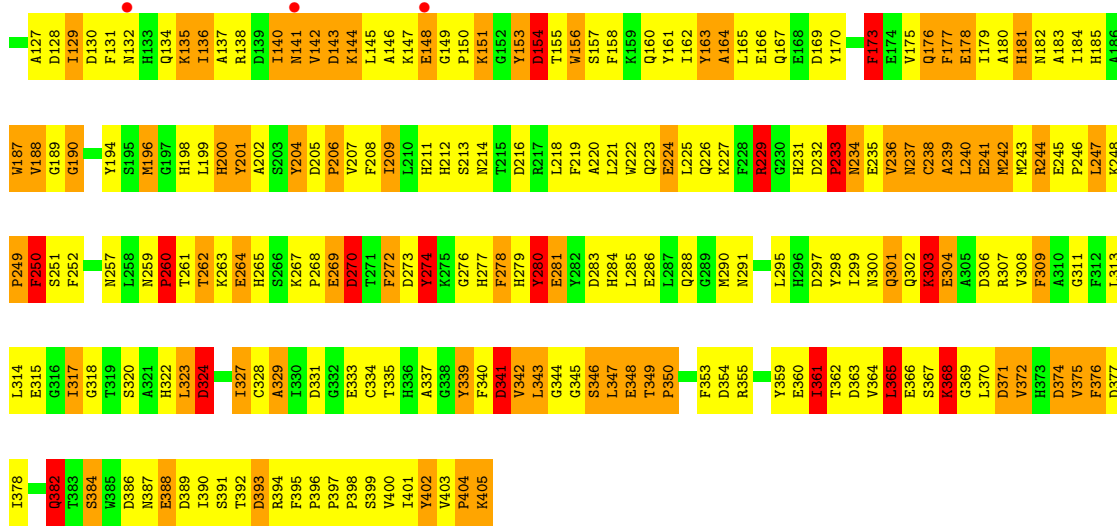
4 Data and refinement statistics ⓘ

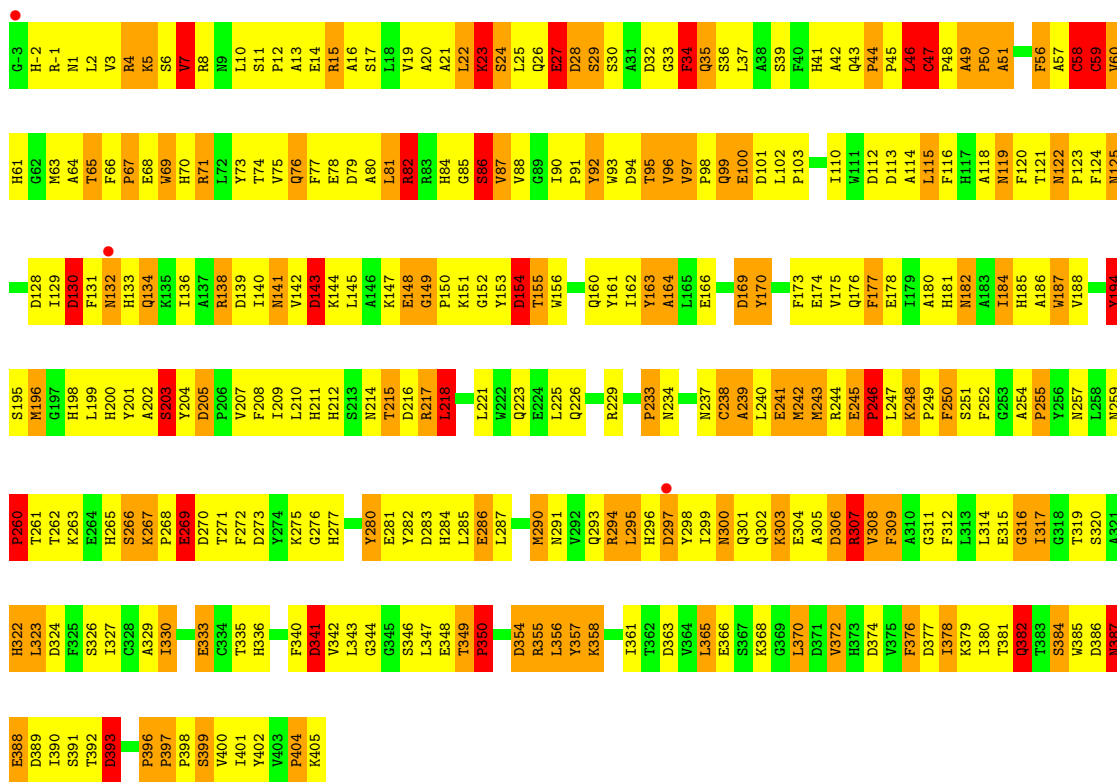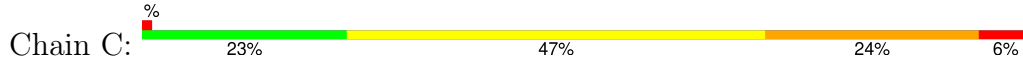| Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	105.49Å 105.49Å 374.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 3.30 15.00 – 3.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (15.00-3.30) 95.8 (15.00-3.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.47 (at 3.31Å)	Xtriage
Refinement program	REFMAC 5.0	Depositor
R, R_{free}	0.248 , 0.288 0.246 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	79.8	Xtriage
Anisotropy	0.078	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 22.8	EDS
L-test for twinning[2]	$< |L| > = 0.44, < L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	9984	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.44% 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: CU, NAG

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	1.80	53/3415 (1.6%)	1.61	69/4644 (1.5%)
1	B	2.00	91/3415 (2.7%)	1.77	79/4644 (1.7%)
1	C	1.83	64/3415 (1.9%)	1.58	62/4644 (1.3%)
All	All	1.88	208/10245 (2.0%)	1.66	210/13932 (1.5%)

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	C	0	1

The worst 5 of 208 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	333	GLU	CD-OE1	11.16	1.38	1.25
1	B	238	CYS	CB-SG	-10.44	1.64	1.82
1	C	27	GLU	CG-CD	-10.42	1.36	1.51
1	C	69	TRP	CB-CG	-9.38	1.33	1.50
1	B	137	ALA	CA-CB	-9.33	1.32	1.52

The worst 5 of 210 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	143	ASP	CB-CG-OD2	18.04	134.54	118.30
1	B	47	CYS	CA-CB-SG	-16.17	84.89	114.00
1	B	394	ARG	NE-CZ-NH1	13.36	126.98	120.30
1	A	218	LEU	CB-CG-CD2	-13.07	88.79	111.00
1	B	32	ASP	CB-CG-OD1	12.88	129.89	118.30

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	82	ARG	Sidechain

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	3310	0	3075	455	2
1	B	3310	0	3074	400	0
1	C	3310	0	3074	452	0
2	A	15	0	15	3	0
2	C	30	0	30	1	0
3	A	2	0	0	0	0
3	B	4	0	0	0	2
3	C	3	0	0	0	0
All	All	9984	0	9268	1305	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 68.

The worst 5 of 1305 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:290:MET:SD	1:A:290:MET:CE	2.02	1.47
1:B:290:MET:SD	1:B:290:MET:CE	2.05	1.44
1:C:59:CYS:SG	1:C:61:HIS:CE1	2.18	1.36
1:C:242:MET:CE	1:C:242:MET:SD	2.14	1.36
1:C:-2:HIS:CD2	1:C:-1:ARG:HE	1.57	1.22

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)
1:A:-2:HIS:CA	3:B:5503:CU:CU[6_456]	1.67	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:-2:HIS:O	3:B:5503:CU:CU[6_456]	1.69	0.51

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	404/408 (99%)	280 (69%)	86 (21%)	38 (9%)	0	3
1	B	404/408 (99%)	311 (77%)	70 (17%)	23 (6%)	1	9
1	C	404/408 (99%)	287 (71%)	83 (20%)	34 (8%)	0	4
All	All	1212/1224 (99%)	878 (72%)	239 (20%)	95 (8%)	1	5

5 of 95 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	-2	HIS
1	A	34	PHE
1	A	57	ALA
1	A	58	CYS
1	A	105	PHE

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	352/352 (100%)	306 (87%)	46 (13%)	3	14

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	352/352 (100%)	289 (82%)	63 (18%)	1	7
1	C	352/352 (100%)	292 (83%)	60 (17%)	1	8
All	All	1056/1056 (100%)	887 (84%)	169 (16%)	2	9

5 of 169 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	27	GLU
1	C	241	GLU
1	C	47	CYS
1	C	120	PHE
1	C	297	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	226	GLN
1	C	-2	HIS
1	C	265	HIS
1	B	234	ASN
1	B	301	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 ⓘ

Of 12 ligands modelled in this entry, 9 are monoatomic - leaving 3 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	$\#|Z| > 2$
2	NAG	C	650	-	15,15,15	1.89	4 (26%)	21,21,21	1.82	6 (28%)
2	NAG	A	600	-	15,15,15	1.71	3 (20%)	21,21,21	1.96	6 (28%)
2	NAG	C	651	-	15,15,15	1.80	4 (26%)	21,21,21	4.28	11 (52%)

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
2	NAG	C	650	-	-	2/6/26/26	0/1/1/1
2	NAG	A	600	-	-	3/6/26/26	0/1/1/1
2	NAG	C	651	-	-	4/6/26/26	0/1/1/1

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	650	NAG	O7-C7	4.74	1.33	1.23
2	A	600	NAG	C3-C2	3.62	1.59	1.53
2	C	650	NAG	C8-C7	3.14	1.57	1.50
2	C	651	NAG	C8-C7	2.92	1.56	1.50
2	C	651	NAG	O1-C1	2.74	1.48	1.39

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	651	NAG	C1-C2-C3	-13.28	92.43	110.54
2	C	651	NAG	O1-C1-C2	8.11	126.06	109.22
2	C	651	NAG	C1-O5-C5	-7.08	99.96	113.65
2	C	651	NAG	O5-C5-C6	4.81	118.36	106.44
2	A	600	NAG	C1-C2-N2	-4.63	105.37	110.73

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	600	NAG	C8-C7-N2-C2
2	A	600	NAG	O7-C7-N2-C2
2	C	650	NAG	C8-C7-N2-C2
2	C	650	NAG	O7-C7-N2-C2
2	C	651	NAG	C8-C7-N2-C2

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	650	NAG	1	0
2	A	600	NAG	3	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1
1	C	1
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	-1:ARG	C	1:ASN	N	5.08
1	C	-1:ARG	C	1:ASN	N	3.99
1	B	-1:ARG	C	1:ASN	N	2.64

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, 95th 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	408/408 (100%)	-0.16	4 (0%)	79	68	43, 43, 43, 78	0
1	B	408/408 (100%)	-0.24	4 (0%)	79	68	43, 43, 43, 78	0
1	C	408/408 (100%)	-0.06	3 (0%)	84	75	31, 43, 43, 78	1 (0%)
All	All	1224/1224 (100%)	-0.15	11 (0%)	81	70	31, 43, 43, 78	1 (0%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	-3	GLY	7.3
1	B	141	ASN	4.6
1	B	132	ASN	3.1
1	C	132	ASN	3.0
1	B	148	GLU	2.8

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, 95th 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
2	NAG	C	651	15/15	0.57	0.23	77,77,77,77	0
2	NAG	A	600	15/15	0.63	0.30	77,77,77,77	0
2	NAG	C	650	15/15	0.72	0.14	77,77,77,77	0
3	CU	A	5001	1/1	0.97	0.07	77,77,77,77	0
3	CU	B	5503	1/1	0.97	0.07	77,77,77,77	0
3	CU	B	5506	1/1	0.97	0.07	77,77,77,77	0
3	CU	A	5012	1/1	0.98	0.04	77,77,77,77	0
3	CU	C	5509	1/1	0.98	0.05	77,77,77,77	0
3	CU	B	5004	1/1	0.99	0.08	77,77,77,77	0
3	CU	C	5007	1/1	0.99	0.03	77,77,77,77	0
3	CU	C	5018	1/1	0.99	0.04	77,77,77,77	0
3	CU	B	5015	1/1	0.99	0.06	77,77,77,77	0

6.5   Other polymers ⓘ

There are no such residues in this entry.