



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2018 – 01:38 pm GMT

PDB ID : 1JS8  
Title : Structure of a Functional Unit from Octopus Hemocyanin  
Authors : Cuff, M.E.; Miller, K.I.; van Holde, K.E.; Hendrickson, W.A.  
Deposited on : 2001-08-16  
Resolution : 2.30 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

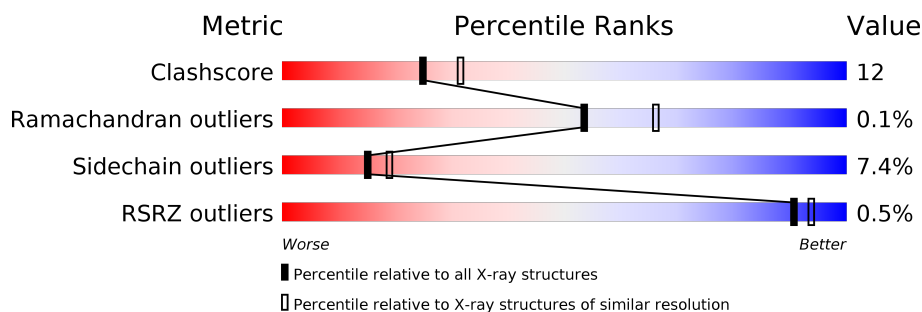
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	122126	5072 (2.30-2.30)
Ramachandran outliers	120053	5022 (2.30-2.30)
Sidechain outliers	120020	5021 (2.30-2.30)
RSRZ outliers	108989	4374 (2.30-2.30)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	394	
1	B	394	

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
4	BMA	A	896	-	-	-	X
4	BMA	B	995	-	-	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	MAN	A	992	-	-	X	X

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 6593 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.

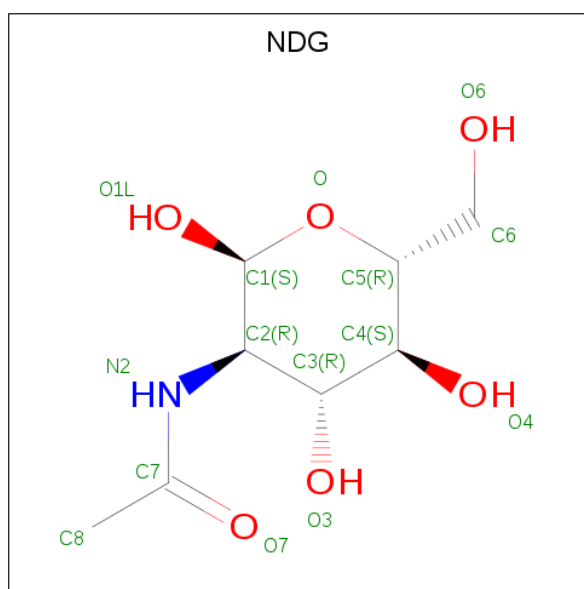
- Molecule 1 is a protein called Hemocyanin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	382	Total	C	N	O	S	0	0	0
			3067	1959	521	573	14			
1	B	382	Total	C	N	O	S	0	0	0
			3067	1959	521	573	14			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2843	LEU	PRO	see remark 999	? O61363
B	2843	LEU	PRO	see remark 999	? O61363

- Molecule 2 is 2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE (three-letter code: NDG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



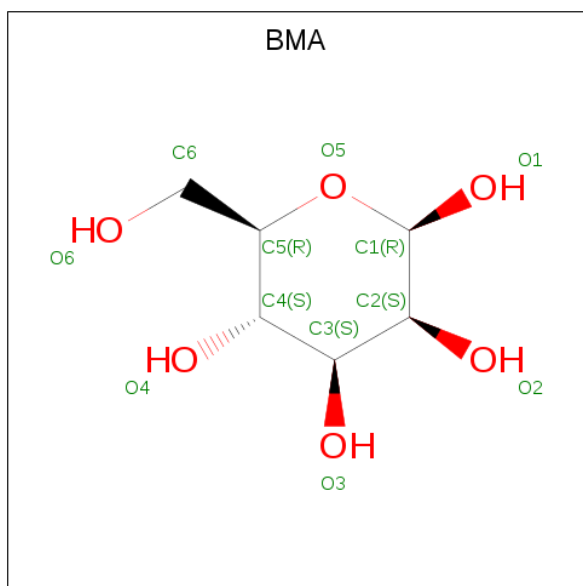
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



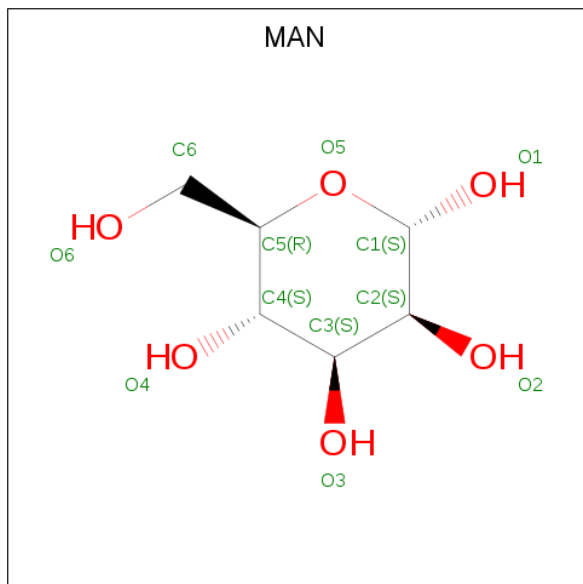
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is BETA-D-MANNOSE (three-letter code: BMA) (formula:  $C_6H_{12}O_6$ ).



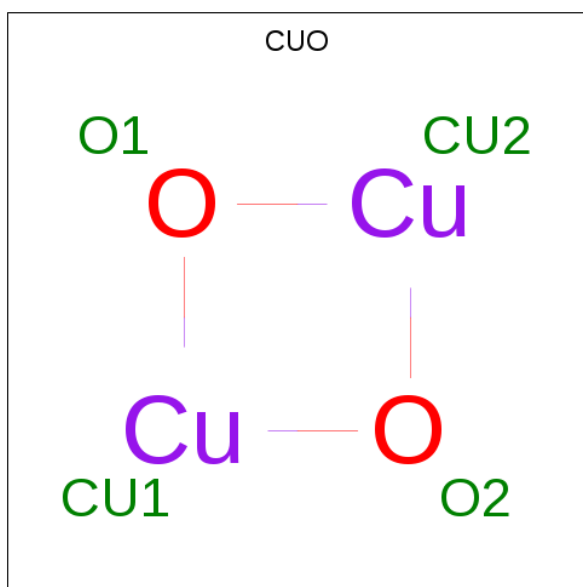
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	B	1	Total	C	O	0	0
			11	6	5		

- Molecule 5 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			11	6	5		
5	A	1	Total	C	O	0	0
			11	6	5		
5	A	1	Total	C	O	0	0
			11	6	5		
5	A	1	Total	C	O	0	0
			11	6	5		
5	B	1	Total	C	O	0	0
			11	6	5		
5	B	1	Total	C	O	0	0
			11	6	5		

- Molecule 6 is CU2-O2 CLUSTER (three-letter code: CUO) (formula: Cu<sub>2</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	Cu	O	0	0
			4	2	2		
6	B	1	Total	Cu	O	0	0
			4	2	2		

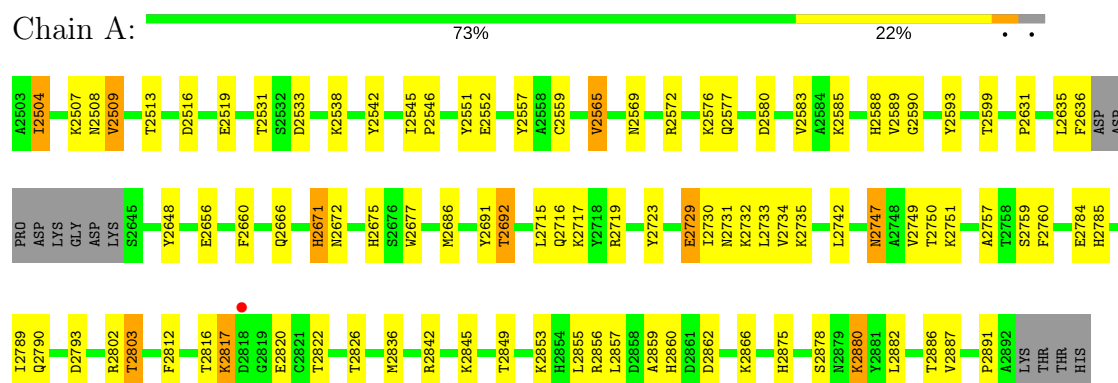
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	175	Total	O	0	0
			175	175		
7	B	121	Total	O	0	0
			121	121		

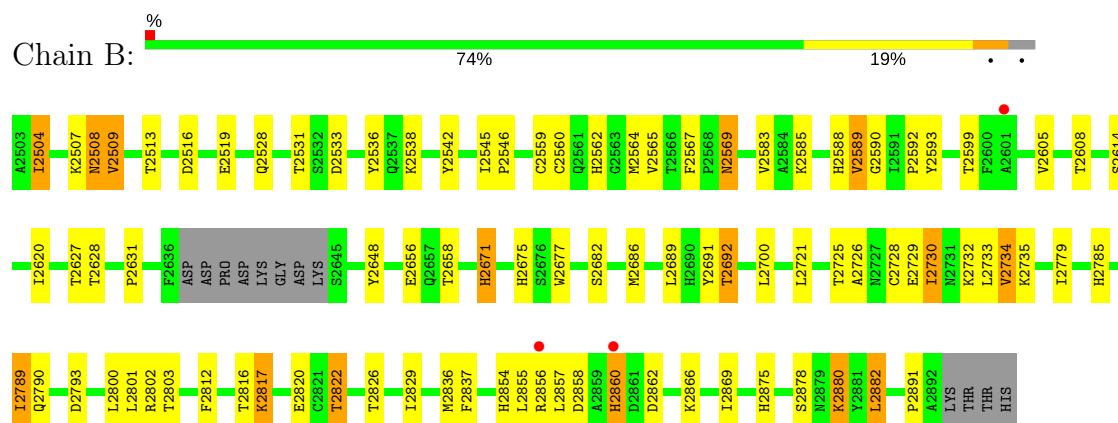
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.86Å 168.39Å 58.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.30 17.95 – 2.30	Depositor EDS
% Data completeness (in resolution range)	(Not available) (10.00-2.30) 94.6 (17.95-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.03 (at 2.30Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, $R_{free}$	0.202 , 0.262 0.190 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.5	Xtriage
Anisotropy	0.475	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 65.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6593	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.62% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CUO, BMA, NAG, NDG, MAN

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	0.64	0/3157	0.79	2/4293 (0.0%)
1	B	0.60	0/3157	0.78	0/4293
All	All	0.62	0/6314	0.79	2/8586 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	2842	ARG	NE-CZ-NH2	-5.52	117.54	120.30
1	A	2802	ARG	NE-CZ-NH2	5.47	123.04	120.30

There are no chirality outliers.

There are no planarity outliers.

### 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	3067	0	2906	69	0
1	B	3067	0	2906	71	0
2	A	14	0	12	0	0
3	A	14	0	12	0	0
3	B	28	0	25	5	0
4	A	22	0	18	0	0
4	B	11	0	10	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	44	0	38	6	0
5	B	22	0	18	1	0
6	A	4	0	0	0	0
6	B	4	0	0	0	0
7	A	175	0	0	5	0
7	B	121	0	0	7	0
All	All	6593	0	5945	140	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 12.

All (140) 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:B:2880:LYS:HB3	1:B:2880:LYS:NZ	1.72	1.01
1:B:2880:LYS:HB3	1:B:2880:LYS:HZ2	1.29	0.93
1:A:2666:GLN:HE22	1:A:2845:LYS:H	1.28	0.80
1:B:2878:SER:OG	1:B:2880:LYS:HB2	1.84	0.76
1:A:2880:LYS:HG2	5:A:992:MAN:O2	1.86	0.74
1:B:2675:HIS:HD1	1:B:2692:THR:HG21	1.53	0.73
1:A:2519:GLU:HG3	1:A:2588:HIS:CE1	2.24	0.73
1:A:2675:HIS:ND1	1:A:2692:THR:HG21	2.03	0.72
1:A:2675:HIS:HD1	1:A:2692:THR:HG21	1.54	0.72
1:B:2533:ASP:HA	1:B:2538:LYS:HG2	1.72	0.72
1:B:2675:HIS:ND1	1:B:2692:THR:HG21	2.06	0.71
1:A:2675:HIS:CE1	1:A:2692:THR:HG21	2.25	0.71
1:B:2686:MET:HA	1:B:2692:THR:HG23	1.74	0.69
1:B:2519:GLU:HG3	1:B:2588:HIS:CE1	2.28	0.69
1:B:2565:VAL:HG11	1:B:2729:GLU:HG3	1.74	0.68
5:A:992:MAN:C5	3:B:991:NAG:H61	2.24	0.68
1:B:2565:VAL:CG1	1:B:2729:GLU:HG3	2.25	0.67
1:B:2816:THR:HG22	1:B:2857:LEU:HD21	1.77	0.67
1:B:2675:HIS:CE1	1:B:2692:THR:HG21	2.30	0.67
1:A:2686:MET:HA	1:A:2692:THR:HG23	1.76	0.66
1:B:2569:ASN:H	1:B:2569:ASN:HD22	1.42	0.66
1:A:2886:THR:HG21	7:B:537:HOH:O	1.95	0.65
1:A:2716:GLN:HE21	1:A:2719:ARG:HH11	1.45	0.65
1:A:2533:ASP:HA	1:A:2538:LYS:HG2	1.77	0.65
1:A:2729:GLU:HG3	7:A:580:HOH:O	1.97	0.64
1:A:2716:GLN:NE2	1:A:2719:ARG:HH11	1.96	0.64
1:A:2859:ALA:HB1	1:B:2730:ILE:HD12	1.80	0.64

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:2686:MET:HA	1:B:2692:THR:CG2	2.30	0.61
1:A:2504:ILE:HD11	1:A:2590:GLY:HA2	1.83	0.61
5:A:992:MAN:H5	3:B:991:NAG:H61	1.81	0.60
1:A:2793:ASP:OD1	1:A:2891:PRO:HA	2.03	0.59
1:A:2545:ILE:HD11	1:A:2691:TYR:HE1	1.66	0.59
1:A:2803:THR:HG21	7:A:564:HOH:O	2.02	0.57
1:B:2880:LYS:HZ3	1:B:2880:LYS:HB3	1.68	0.57
1:A:2504:ILE:HD12	1:A:2583:VAL:HG21	1.85	0.57
1:B:2592:PRO:HG2	1:B:2700:LEU:HD22	1.87	0.57
1:A:2878:SER:OG	1:A:2880:LYS:HB2	2.06	0.56
1:A:2542:TYR:O	1:A:2559:CYS:HB2	2.05	0.56
1:B:2569:ASN:ND2	1:B:2569:ASN:H	2.03	0.55
1:A:2545:ILE:HD11	1:A:2691:TYR:CE1	2.40	0.55
1:A:2686:MET:HA	1:A:2692:THR:CG2	2.37	0.55
1:A:2817:LYS:HA	1:A:2817:LYS:HE2	1.88	0.54
1:B:2513:THR:O	1:B:2516:ASP:HB2	2.07	0.54
1:B:2817:LYS:HE3	1:B:2817:LYS:HA	1.88	0.54
1:B:2562:HIS:CE1	1:B:2689:LEU:HD11	2.43	0.54
1:B:2812:PHE:HA	1:B:2866:LYS:O	2.07	0.54
1:B:2656:GLU:HG3	1:B:2790:GLN:NE2	2.24	0.53
1:B:2793:ASP:OD1	1:B:2891:PRO:HA	2.08	0.53
1:B:2732:LYS:HB3	7:B:664:HOH:O	2.08	0.53
1:A:2565:VAL:HG23	1:A:2660:PHE:CE2	2.44	0.52
5:A:992:MAN:O5	3:B:991:NAG:H61	2.10	0.52
1:B:2789:ILE:HG22	1:B:2790:GLN:HG3	1.92	0.51
1:A:2557:TYR:HE2	1:A:2836:MET:CE	2.24	0.51
1:B:2631:PRO:HA	1:B:2677:TRP:O	2.11	0.50
1:B:2565:VAL:HG11	1:B:2728:CYS:SG	2.52	0.50
1:B:2508:ASN:HD22	1:B:2509:VAL:N	2.10	0.50
1:A:2671:HIS:CE1	1:A:2672:ASN:OD1	2.65	0.49
1:A:2816:THR:HG22	1:A:2857:LEU:HD21	1.94	0.49
1:B:2560:CYS:SG	1:B:2689:LEU:HD11	2.52	0.49
1:B:2605:VAL:HA	1:B:2608:THR:OG1	2.12	0.49
1:A:2715:LEU:O	1:A:2719:ARG:HG3	2.12	0.49
1:A:2648:TYR:OH	1:A:2785:HIS:CE1	2.66	0.49
1:B:2545:ILE:HD11	1:B:2691:TYR:CE1	2.47	0.49
1:B:2648:TYR:OH	1:B:2785:HIS:NE2	2.45	0.49
1:B:2802:ARG:NH1	7:B:512:HOH:O	2.45	0.49
1:B:2822:THR:HG21	1:B:2855:LEU:HG	1.95	0.48
1:A:2812:PHE:HA	1:A:2866:LYS:O	2.12	0.48
1:A:2734:VAL:HG11	1:B:2860:HIS:O	2.13	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:2565:VAL:HG23	1:A:2660:PHE:HE2	1.79	0.48
1:A:2656:GLU:HG3	1:A:2790:GLN:NE2	2.28	0.48
1:B:2620:ILE:HD11	1:B:2627:THR:HG22	1.96	0.48
1:A:2729:GLU:OE2	1:A:2732:LYS:HE3	2.14	0.48
1:B:2542:TYR:O	1:B:2559:CYS:HB2	2.14	0.48
5:A:992:MAN:C5	3:B:991:NAG:C6	2.90	0.48
1:B:2509:VAL:HG22	1:B:2593:TYR:O	2.13	0.47
1:B:2816:THR:OG1	1:B:2820:GLU:HB3	2.15	0.47
1:B:2822:THR:CG2	1:B:2855:LEU:HG	2.45	0.47
5:A:992:MAN:C6	3:B:991:NAG:O6	2.63	0.47
1:A:2717:LYS:HD2	1:A:2723:TYR:CZ	2.49	0.47
1:B:2628:THR:HB	1:B:2682:SER:HB2	1.97	0.47
1:A:2513:THR:O	1:A:2516:ASP:HB2	2.16	0.46
1:A:2557:TYR:CE2	1:A:2836:MET:CE	2.99	0.46
1:B:2822:THR:HG22	1:B:2855:LEU:HD11	1.96	0.46
1:B:2789:ILE:HD13	7:B:552:HOH:O	2.14	0.46
1:A:2880:LYS:HB3	1:A:2880:LYS:NZ	2.31	0.46
1:B:2858:ASP:OD1	1:B:2860:HIS:CE1	2.69	0.46
1:B:2779:ILE:HD13	7:B:501:HOH:O	2.14	0.46
1:B:2507:LYS:HE3	1:B:2589:VAL:HG21	1.97	0.45
1:B:2658:THR:HB	1:B:2721:LEU:HD13	1.98	0.45
1:A:2747:ASN:ND2	1:A:2750:THR:H	2.15	0.45
1:B:2504:ILE:HD12	1:B:2583:VAL:HG21	1.97	0.45
1:B:2802:ARG:HD2	7:B:529:HOH:O	2.16	0.45
1:A:2551:TYR:HE2	7:A:657:HOH:O	1.99	0.45
1:A:2569:ASN:HD22	1:A:2569:ASN:H	1.64	0.45
1:A:2509:VAL:HG22	1:A:2593:TYR:O	2.17	0.44
1:B:2817:LYS:CE	1:B:2817:LYS:HA	2.48	0.44
1:A:2735:LYS:NZ	1:B:2862:ASP:OD1	2.48	0.44
1:B:2726:ALA:HB3	1:B:2733:LEU:HD11	1.99	0.44
1:A:2747:ASN:C	1:A:2747:ASN:HD22	2.20	0.44
1:A:2822:THR:HG22	1:A:2855:LEU:HD11	1.99	0.44
1:A:2545:ILE:HA	1:A:2546:PRO:C	2.38	0.44
1:A:2569:ASN:H	1:A:2569:ASN:ND2	2.15	0.44
1:B:2545:ILE:HD11	1:B:2691:TYR:HE1	1.82	0.44
1:B:2567:PHE:CZ	1:B:2671:HIS:CE1	3.06	0.44
1:B:2878:SER:C	1:B:2880:LYS:H	2.21	0.43
1:B:2686:MET:CA	1:B:2692:THR:HG23	2.47	0.43
1:A:2849:THR:O	1:A:2853:LYS:HG2	2.19	0.43
1:A:2577:GLN:HE21	1:A:2750:THR:HG21	1.83	0.43
1:A:2731:ASN:O	1:A:2734:VAL:HG22	2.18	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:2717:LYS:HA	7:A:682:HOH:O	2.19	0.43
1:B:2729:GLU:O	1:B:2732:LYS:HB2	2.19	0.42
1:A:2742:LEU:O	1:A:2751:LYS:HE3	2.19	0.42
1:A:2860:HIS:O	1:B:2734:VAL:HG21	2.19	0.42
1:A:2686:MET:O	1:A:2692:THR:HG23	2.19	0.42
1:B:2801:LEU:HD22	1:B:2829:ILE:HD11	2.00	0.42
1:A:2572:ARG:NH1	1:A:2760:PHE:O	2.52	0.42
1:B:2837:PHE:HD1	5:B:994:MAN:HO6	1.63	0.42
1:A:2733:LEU:O	1:A:2757:ALA:N	2.52	0.42
1:B:2528:GLN:HA	1:B:2536:TYR:HB2	2.00	0.42
1:A:2576:LYS:HZ2	1:A:2580:ASP:CG	2.22	0.42
1:A:2631:PRO:HA	1:A:2677:TRP:O	2.20	0.42
1:B:2585:LYS:HD3	1:B:2585:LYS:HA	1.83	0.42
1:A:2635:LEU:HD23	1:A:2636:PHE:CZ	2.55	0.42
1:A:2747:ASN:ND2	1:A:2749:VAL:H	2.18	0.42
1:A:2816:THR:OG1	1:A:2820:GLU:HB3	2.19	0.42
1:B:2504:ILE:HD11	1:B:2590:GLY:HA2	2.02	0.41
1:B:2656:GLU:HG3	1:B:2790:GLN:HE21	1.85	0.41
1:A:2886:THR:HG22	1:A:2887:VAL:N	2.35	0.41
1:B:2869:ILE:HD11	1:B:2882:LEU:HB2	2.02	0.41
1:A:2565:VAL:HG12	7:A:593:HOH:O	2.21	0.41
1:A:2675:HIS:HD1	1:A:2692:THR:CG2	2.28	0.41
1:A:2504:ILE:HD12	1:A:2583:VAL:CG2	2.50	0.41
1:A:2585:LYS:HD3	1:A:2585:LYS:HA	1.92	0.41
1:B:2628:THR:CB	1:B:2682:SER:HB2	2.51	0.41
1:B:2565:VAL:HG13	1:B:2729:GLU:HG3	2.02	0.41
1:A:2862:ASP:OD1	1:B:2735:LYS:HE3	2.21	0.41
1:B:2803:THR:HG21	7:B:548:HOH:O	2.20	0.40
1:B:2545:ILE:HA	1:B:2546:PRO:C	2.40	0.40
1:A:2577:GLN:NE2	1:A:2750:THR:OG1	2.49	0.40
1:A:2880:LYS:HB3	1:A:2880:LYS:HE3	1.90	0.40

There are no symmetry-related clashes.

## 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	378/394 (96%)	368 (97%)	9 (2%)	1 (0%)	43	53
1	B	378/394 (96%)	364 (96%)	14 (4%)	0	100	100
All	All	756/788 (96%)	732 (97%)	23 (3%)	1 (0%)	53	66

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2730	ILE

### 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	332/343 (97%)	309 (93%)	23 (7%)	17	22
1	B	332/343 (97%)	306 (92%)	26 (8%)	14	17
All	All	664/686 (97%)	615 (93%)	49 (7%)	15	19

All (49) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2504	ILE
1	A	2507	LYS
1	A	2508	ASN
1	A	2509	VAL
1	A	2531	THR
1	A	2552	GLU
1	A	2565	VAL
1	A	2589	VAL
1	A	2599	THR
1	A	2671	HIS
1	A	2692	THR
1	A	2729	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	2747	ASN
1	A	2759	SER
1	A	2784	GLU
1	A	2789	ILE
1	A	2803	THR
1	A	2817	LYS
1	A	2826	THR
1	A	2856	ARG
1	A	2875	HIS
1	A	2880	LYS
1	A	2882	LEU
1	B	2504	ILE
1	B	2508	ASN
1	B	2509	VAL
1	B	2531	THR
1	B	2564	MET
1	B	2569	ASN
1	B	2589	VAL
1	B	2599	THR
1	B	2614	SER
1	B	2671	HIS
1	B	2692	THR
1	B	2725	THR
1	B	2730	ILE
1	B	2734	VAL
1	B	2789	ILE
1	B	2800	LEU
1	B	2817	LYS
1	B	2822	THR
1	B	2826	THR
1	B	2836	MET
1	B	2854	HIS
1	B	2856	ARG
1	B	2860	HIS
1	B	2875	HIS
1	B	2880	LYS
1	B	2882	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2508	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	2569	ASN
1	A	2577	GLN
1	A	2666	GLN
1	A	2704	ASN
1	A	2713	GLN
1	A	2716	GLN
1	A	2724	ASN
1	A	2747	ASN
1	A	2773	ASN
1	A	2775	HIS
1	A	2785	HIS
1	A	2790	GLN
1	B	2508	ASN
1	B	2569	ASN
1	B	2577	GLN
1	B	2704	ASN
1	B	2724	ASN
1	B	2773	ASN
1	B	2775	HIS
1	B	2790	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

15 ligands are modelled in this entry.

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
6	CUO	A	888	1	0,4,4	0.00	-	0,4,4	0.00	-
2	NDG	A	890	3	14,14,15	1.44	2 (14%)	17,19,21	0.88	1 (5%)
3	NAG	A	891	2,4	14,14,15	0.68	0	17,19,21	1.01	2 (11%)
4	BMA	A	892	3,5,4	11,11,12	1.18	2 (18%)	15,15,17	1.86	3 (20%)
5	MAN	A	893	5,4	11,11,12	0.71	0	15,15,17	0.85	0
5	MAN	A	894	5	11,11,12	0.63	0	15,15,17	0.95	1 (6%)
5	MAN	A	895	5	11,11,12	0.63	0	15,15,17	0.81	0
4	BMA	A	896	4	11,11,12	1.30	2 (18%)	15,15,17	1.43	2 (13%)
5	MAN	A	992	-	11,11,12	1.55	3 (27%)	15,15,17	1.30	2 (13%)
3	NAG	B	990	1,3	14,14,15	1.10	2 (14%)	17,19,21	1.58	4 (23%)
3	NAG	B	991	3	14,14,15	1.35	2 (14%)	17,19,21	1.40	3 (17%)
5	MAN	B	993	5,4	11,11,12	0.43	0	15,15,17	1.04	0
5	MAN	B	994	5	11,11,12	0.90	1 (9%)	15,15,17	0.87	1 (6%)
4	BMA	B	995	5	11,11,12	1.02	0	15,15,17	2.05	3 (20%)
6	CUO	B	999	1	0,4,4	0.00	-	0,4,4	0.00	-

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
6	CUO	A	888	1	-	0/0/4/4	0/1/1/1
2	NDG	A	890	3	-	0/6/23/26	0/1/1/1
3	NAG	A	891	2,4	-	0/6/23/26	0/1/1/1
4	BMA	A	892	3,5,4	-	0/2/19/22	0/1/1/1
5	MAN	A	893	5,4	-	0/2/19/22	0/1/1/1
5	MAN	A	894	5	-	0/2/19/22	0/1/1/1
5	MAN	A	895	5	-	0/2/19/22	0/1/1/1
4	BMA	A	896	4	-	0/2/19/22	0/1/1/1
5	MAN	A	992	-	-	0/2/19/22	0/1/1/1
3	NAG	B	990	1,3	-	0/6/23/26	0/1/1/1
3	NAG	B	991	3	-	0/6/23/26	0/1/1/1
5	MAN	B	993	5,4	-	0/2/19/22	0/1/1/1
5	MAN	B	994	5	-	0/2/19/22	0/1/1/1
4	BMA	B	995	5	-	0/2/19/22	0/1/1/1
6	CUO	B	999	1	-	0/0/4/4	0/1/1/1

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	890	NDG	C1-C2	-3.36	1.47	1.52
2	A	890	NDG	C2-N2	-2.36	1.42	1.46
4	A	892	BMA	O3-C3	2.04	1.48	1.43
4	A	892	BMA	C2-C3	2.07	1.55	1.52
3	B	990	NAG	O5-C5	2.18	1.48	1.43
5	B	994	MAN	O5-C5	2.20	1.48	1.43
5	A	992	MAN	O5-C1	2.20	1.47	1.43
4	A	896	BMA	C2-C3	2.25	1.55	1.52
3	B	991	NAG	C8-C7	2.36	1.55	1.50
4	A	896	BMA	O5-C5	2.52	1.48	1.43
3	B	991	NAG	C3-C2	2.54	1.58	1.52
5	A	992	MAN	O5-C5	2.57	1.48	1.43
3	B	990	NAG	C4-C5	2.67	1.58	1.53
5	A	992	MAN	C2-C3	2.81	1.56	1.52

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	995	BMA	C1-C2-C3	-4.27	104.25	109.66
4	A	892	BMA	C1-C2-C3	-3.89	104.73	109.66
5	A	992	MAN	C2-C3-C4	-3.27	105.19	110.87
3	B	990	NAG	O4-C4-C3	-3.22	102.82	110.34
3	B	990	NAG	O5-C1-C2	-2.64	107.88	111.52
3	B	991	NAG	C4-C3-C2	-2.62	107.19	111.02
3	B	991	NAG	C1-O5-C5	-2.61	108.59	112.19
2	A	890	NDG	C1-O-C5	-2.53	108.71	112.19
3	B	990	NAG	C8-C7-N2	-2.42	111.86	116.10
3	A	891	NAG	O5-C1-C2	-2.39	108.23	111.52
5	A	992	MAN	C3-C4-C5	-2.38	105.98	110.24
4	A	896	BMA	C3-C4-C5	-2.20	106.30	110.24
3	A	891	NAG	C4-C3-C2	-2.11	107.93	111.02
4	A	892	BMA	O2-C2-C3	2.05	114.19	110.19
5	B	994	MAN	C1-O5-C5	2.14	115.14	112.19
5	A	894	MAN	C1-O5-C5	2.83	116.08	112.19
3	B	991	NAG	C6-C5-C4	3.05	120.19	112.99
3	B	990	NAG	C4-C3-C2	3.13	115.60	111.02
4	B	995	BMA	C3-C4-C5	3.27	116.08	110.24
4	A	896	BMA	C1-C2-C3	3.83	114.50	109.66
4	A	892	BMA	O3-C3-C2	4.89	119.14	110.04
4	B	995	BMA	O5-C1-C2	5.18	118.86	110.78

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	992	MAN	6	0
3	B	991	NAG	5	0
5	B	994	MAN	1	0

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	382/394 (96%)	-0.39	1 (0%) 93 96	4, 17, 38, 51	0
1	B	382/394 (96%)	-0.26	3 (0%) 86 89	5, 19, 40, 53	0
All	All	764/788 (96%)	-0.33	4 (0%) 90 94	4, 18, 39, 53	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2860	HIS	4.1
1	A	2818	ASP	2.3
1	B	2856	ARG	2.2
1	B	2601	ALA	2.1

### 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 carbohydrates in this entry.

### 6.4 Ligands [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
5	MAN	A	992	11/12	0.46	0.74	34,36,39,40	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	B	991	14/15	0.62	0.32	45,52,54,58	0
4	BMA	A	896	11/12	0.63	0.45	37,39,44,45	0
4	BMA	B	995	11/12	0.72	0.42	62,65,68,72	0
5	MAN	B	993	11/12	0.83	0.17	32,37,46,56	0
3	NAG	B	990	14/15	0.84	0.23	41,47,52,52	0
5	MAN	B	994	11/12	0.85	0.21	39,43,48,60	0
4	BMA	A	892	11/12	0.88	0.12	31,33,39,41	0
5	MAN	A	895	11/12	0.91	0.19	40,44,49,50	0
5	MAN	A	894	11/12	0.92	0.12	19,27,35,39	0
3	NAG	A	891	14/15	0.92	0.12	26,30,33,36	0
2	NDG	A	890	14/15	0.92	0.13	18,21,27,31	0
5	MAN	A	893	11/12	0.93	0.11	16,21,28,35	0
6	CUO	B	999	4/4	0.98	0.07	11,15,17,22	0
6	CUO	A	888	4/4	0.99	0.07	9,10,11,12	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.