



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 21, 2017 – 12:00 AM EST

PDB ID : 2CWG  
Title : CRYSTALLOGRAPHIC REFINEMENT AND STRUCTURE ANALYSIS  
OF THE COMPLEX OF WHEAT GERM AGGLUTININ WITH A BIVA-  
LENT SIALOGLYCOPEPTIDE FROM GLYCOPHORIN A  
Authors : Wright, C.S.  
Deposited on : unknown  
Resolution : 2.00 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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-20030345

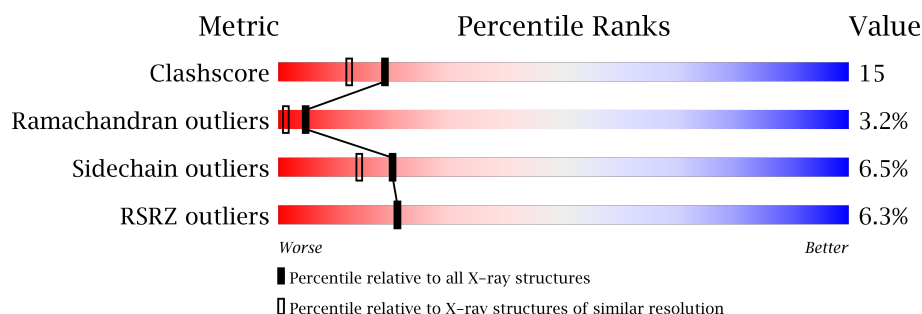
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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 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	171	<div> <div>78%</div> <div>18%</div> <div>• •</div> </div>
1	B	171	<div> <div>7%</div> <div>71%</div> <div>23%</div> <div>5%</div> <div>•</div> </div>
2	D	8	<div> <div>63%</div> <div>50%</div> <div>38%</div> <div>13%</div> </div>
2	E	8	<div> <div>38%</div> <div>25%</div> <div>13%</div> <div>63%</div> </div>

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
5	NDG	B	174	X	-	-	-
5	NDG	D	3	X	-	-	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 2935 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.

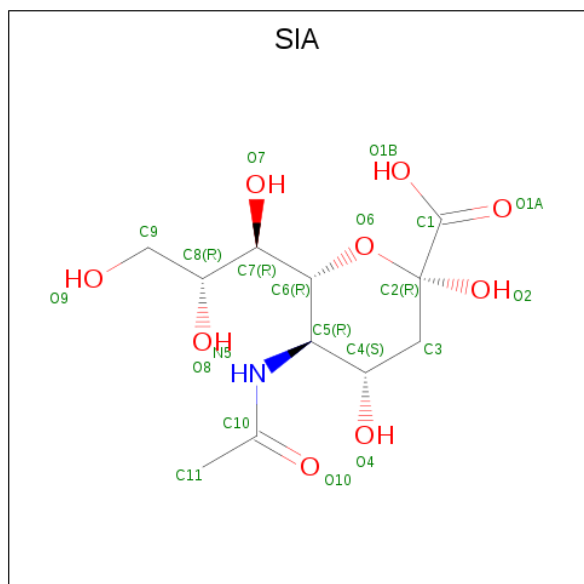
- Molecule 1 is a protein called AGGLUTININ ISOLECTIN 1 (WGA1).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	171	Total	C	N	O	S	0	0	0
			1167	681	214	238	34			
1	B	171	Total	C	N	O	S	0	0	0
			1167	681	214	238	34			

- Molecule 2 is a protein called T5 SIALOGLYCOPEPTIDE OF GLYCOPHORIN A.

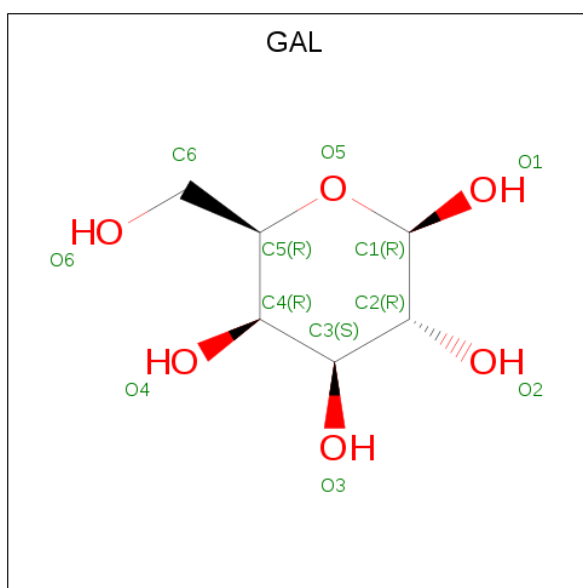
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	8	Total	C	N	O	0	0	0
			63	38	11	14			
2	E	3	Total	C	N	O	0	0	0
			19	12	3	4			

- Molecule 3 is O-SIALIC ACID (three-letter code: SIA) (formula:  $C_{11}H_{19}NO_9$ ).



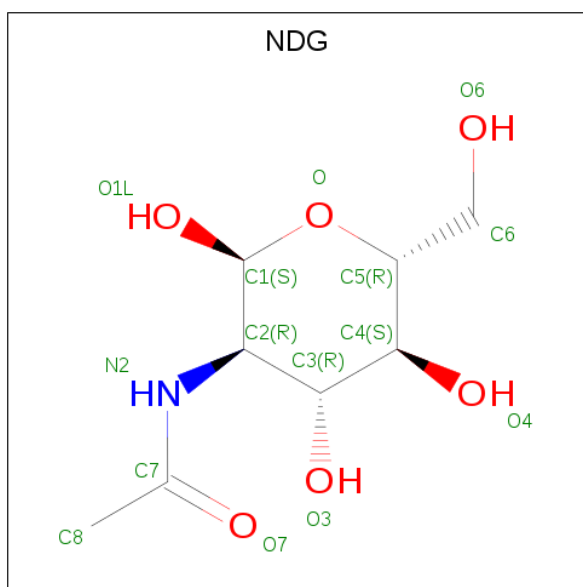
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	D	1	Total	C	N	O	0	0
			20	11	1	8		
3	D	1	Total	C	N	O	0	0
			21	11	1	9		
3	B	1	Total	C	N	O	0	0
			20	11	1	8		
3	B	1	Total	C	N	O	0	0
			21	11	1	9		

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			12	6	6		
4	B	1	Total	C	O	0	0
			12	6	6		

- Molecule 5 is 2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE (three-letter code: NDG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	D	1	Total	C	N	O	0	0
			12	8	1	3		
5	B	1	Total	C	N	O	0	0
			12	8	1	3		

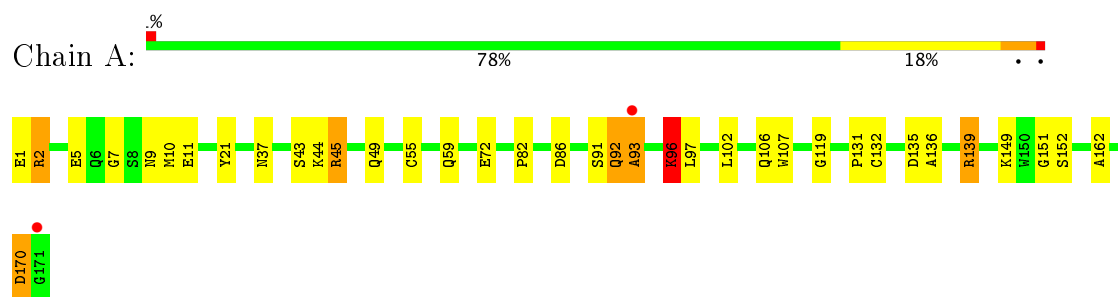
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	187	Total	O	0	0
			187	187		
6	D	24	Total	O	0	0
			24	24		
6	B	176	Total	O	0	0
			176	176		
6	E	2	Total	O	0	0
			2	2		

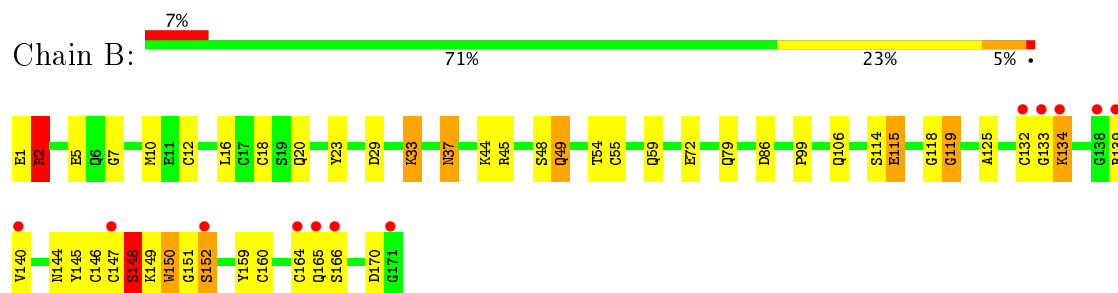
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

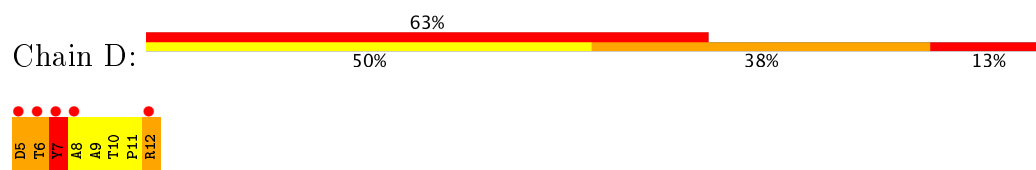
- Molecule 1: AGGLUTININ ISOLECTIN 1 (WGA1)



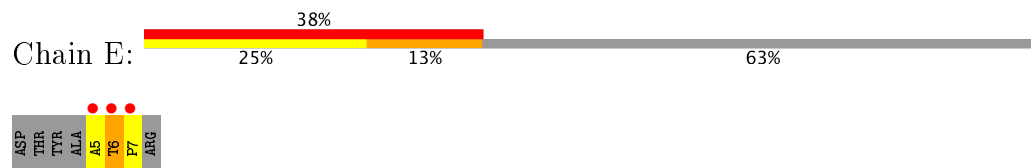
- Molecule 1: AGGLUTININ ISOLECTIN 1 (WGA1)



- Molecule 2: T5 SIALOGLYCOPEPTIDE OF GLYCOPHORIN A



- Molecule 2: T5 SIALOGLYCOPEPTIDE OF GLYCOPHORIN A



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.00 Å 50.40 Å 63.40 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.00 27.53 – 1.99	Depositor EDS
% Data completeness (in resolution range)	(Not available) (10.00-2.00) 91.9 (27.53-1.99)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.74 (at 1.99 Å)	Xtriage
Refinement program	PROLSQ, X-PLOR	Depositor
R, $R_{free}$	0.171 , (Not available) 0.161 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	26.4	Xtriage
Anisotropy	0.254	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 172.5	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	0.96	EDS
Total number of atoms	2935	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.24% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SIA, GAL, PCA, NDG

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	1.07	2/1181 (0.2%)	1.82	17/1580 (1.1%)
1	B	1.09	3/1181 (0.3%)	1.88	19/1580 (1.2%)
2	D	1.10	0/64	1.88	2/86 (2.3%)
2	E	1.28	0/19	1.81	0/26
All	All	1.08	5/2445 (0.2%)	1.85	38/3272 (1.2%)

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	B	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	5	GLU	CD-OE2	8.28	1.34	1.25
1	B	72	GLU	CD-OE2	6.71	1.33	1.25
1	B	115	GLU	CD-OE1	-6.03	1.19	1.25
1	A	11	GLU	CD-OE2	5.53	1.31	1.25
1	B	5	GLU	CD-OE1	-5.31	1.19	1.25

All (38) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	2	ARG	NE-CZ-NH1	15.68	128.14	120.30
1	A	139	ARG	NE-CZ-NH1	13.84	127.22	120.30
1	A	2	ARG	NE-CZ-NH2	13.79	127.19	120.30
1	B	45	ARG	NE-CZ-NH1	-13.49	113.55	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	45	ARG	NE-CZ-NH2	11.90	126.25	120.30
1	A	2	ARG	NE-CZ-NH1	-11.78	114.41	120.30
1	A	86	ASP	CB-CG-OD2	-9.82	109.46	118.30
1	B	106	GLN	CA-CB-CG	8.38	131.85	113.40
1	A	170	ASP	CB-CG-OD2	7.96	125.47	118.30
1	B	150	TRP	CA-CB-CG	7.88	128.67	113.70
1	B	2	ARG	NH1-CZ-NH2	-7.83	110.78	119.40
1	A	135	ASP	CB-CG-OD2	7.75	125.28	118.30
1	A	86	ASP	CB-CG-OD1	7.57	125.11	118.30
1	B	86	ASP	CB-CG-OD2	-7.53	111.52	118.30
1	B	72	GLU	OE1-CD-OE2	-7.27	114.58	123.30
1	B	148	SER	C-N-CA	7.17	139.63	121.70
1	B	86	ASP	CB-CG-OD1	6.51	124.16	118.30
1	B	72	GLU	CG-CD-OE1	6.46	131.22	118.30
1	A	139	ARG	CD-NE-CZ	6.43	132.60	123.60
1	B	149	LYS	CA-CB-CG	6.34	127.35	113.40
1	A	136	ALA	N-CA-CB	5.99	118.49	110.10
1	A	72	GLU	CG-CD-OE1	5.98	130.25	118.30
2	D	5	ASP	O-C-N	5.80	131.99	122.70
1	A	45	ARG	CD-NE-CZ	-5.78	115.51	123.60
1	B	2	ARG	CB-CA-C	5.73	121.85	110.40
1	A	43	SER	N-CA-CB	5.64	118.96	110.50
1	A	139	ARG	NE-CZ-NH2	-5.63	117.48	120.30
1	B	23	TYR	CB-CG-CD1	-5.56	117.66	121.00
1	B	20	GLN	CA-CB-CG	5.54	125.60	113.40
1	B	114	SER	CA-CB-OG	-5.51	96.33	111.20
1	B	115	GLU	CG-CD-OE1	5.46	129.23	118.30
1	A	92	GLN	CB-CG-CD	5.40	125.63	111.60
1	B	54	THR	CA-CB-CG2	5.39	119.94	112.40
2	D	12	ARG	NE-CZ-NH2	5.31	122.96	120.30
1	A	102	LEU	CA-CB-CG	5.27	127.42	115.30
1	A	96	LYS	CA-CB-CG	-5.18	102.01	113.40
1	B	29	ASP	CB-CG-OD1	5.04	122.84	118.30
1	A	72	GLU	CG-CD-OE2	-5.01	108.28	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	2	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	1167	0	999	26	0
1	B	1167	0	998	33	0
2	D	63	0	55	9	0
2	E	19	0	18	9	0
3	B	41	0	34	0	0
3	D	41	0	34	1	0
4	B	12	0	10	3	0
4	D	12	0	10	0	0
5	B	12	0	8	5	0
5	D	12	0	8	1	0
6	A	187	0	0	6	0
6	B	176	0	0	8	0
6	D	24	0	0	0	0
6	E	2	0	0	0	0
All	All	2935	0	2174	70	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 15.

All (70) 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:B:174:NDG:C1	2:E:6:THR:OG1	2.12	0.98
2:D:6:THR:O	2:D:7:TYR:HB2	1.65	0.94
1:A:2:ARG:O	1:B:10:MET:HE3	1.71	0.91
1:B:1:PCA:HB3	2:E:7:PRO:O	1.72	0.88
2:E:6:THR:N	2:E:7:PRO:HD3	1.92	0.83
2:D:5:ASP:O	2:D:6:THR:HG23	1.82	0.79
5:B:174:NDG:H8C1	6:B:866:HOH:O	1.90	0.71
1:B:48:SER:H	1:B:79:GLN:HE22	1.39	0.71
1:B:1:PCA:CB	2:E:7:PRO:O	2.39	0.69
1:A:2:ARG:NH2	1:A:21:TYR:O	2.26	0.68
5:B:174:NDG:C8	6:B:866:HOH:O	2.40	0.67
1:A:170:ASP:O	6:A:865:HOH:O	2.12	0.67
1:A:162:ALA:HB3	6:A:877:HOH:O	1.95	0.66

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:44:LYS:HG3	6:B:839:HOH:O	1.96	0.65
2:D:7:TYR:O	2:D:8:ALA:HB3	1.99	0.63
2:E:6:THR:N	2:E:7:PRO:CD	2.65	0.60
1:B:7:GLY:O	1:B:10:MET:HB2	2.02	0.60
1:A:96:LYS:HG2	1:A:97:LEU:N	2.16	0.60
1:A:132:CYS:HA	1:A:139:ARG:HG3	1.84	0.58
1:A:1:PCA:HB2	1:A:2:ARG:HG2	1.85	0.58
2:D:5:ASP:HB3	1:B:115:GLU:HG3	1.88	0.56
1:A:7:GLY:O	1:A:10:MET:HB2	2.05	0.56
2:D:12:ARG:HB3	5:D:3:NDG:H8C3	1.88	0.55
1:B:49:GLN:H	1:B:49:GLN:HE21	1.53	0.55
1:A:2:ARG:O	1:B:10:MET:CE	2.50	0.54
4:B:173:GAL:HO2	2:E:5:ALA:N	2.05	0.54
1:B:133:GLY:N	1:B:139:ARG:O	2.35	0.52
1:A:2:ARG:C	1:B:10:MET:CE	2.78	0.52
1:B:132:CYS:SG	1:B:151:GLY:HA2	2.50	0.51
1:B:132:CYS:SG	1:B:151:GLY:C	2.90	0.50
1:B:18:CYS:HB2	1:B:37:ASN:ND2	2.27	0.50
1:B:49:GLN:H	1:B:49:GLN:NE2	2.10	0.50
1:B:132:CYS:SG	1:B:152:SER:N	2.85	0.49
2:D:9:ALA:HB1	3:D:1:SIA:O9	2.13	0.49
1:A:131:PRO:HA	1:A:151:GLY:O	2.12	0.48
1:B:144:ASN:C	1:B:145:TYR:O	2.47	0.48
1:B:33:LYS:HA	1:B:33:LYS:HD3	1.68	0.47
1:A:2:ARG:C	1:B:10:MET:HE3	2.34	0.47
1:A:91:SER:C	1:A:93:ALA:H	2.18	0.46
1:B:147:CYS:SG	1:B:151:GLY:HA2	2.55	0.46
2:D:6:THR:O	2:D:7:TYR:CB	2.50	0.46
1:B:160:CYS:O	1:B:170:ASP:HB2	2.16	0.46
1:B:134:LYS:N	6:B:669:HOH:O	2.49	0.46
1:A:55:CYS:HB3	1:A:59:GLN:HB2	1.98	0.45
1:A:82:PRO:HA	6:A:879:HOH:O	2.16	0.45
1:B:99:PRO:HG3	6:B:867:HOH:O	2.17	0.45
4:B:173:GAL:O2	2:E:5:ALA:N	2.50	0.44
2:D:7:TYR:O	2:D:8:ALA:CB	2.64	0.44
2:E:5:ALA:C	2:E:7:PRO:HD3	2.37	0.44
5:B:174:NDG:H5	2:E:6:THR:HG23	1.99	0.44
1:B:44:LYS:HD2	6:B:839:HOH:O	2.17	0.44
1:A:10:MET:CG	1:B:10:MET:HG2	2.48	0.44
1:B:55:CYS:HB3	1:B:59:GLN:HB2	1.99	0.43
1:A:10:MET:HG3	1:B:10:MET:HG2	2.00	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:12:CYS:HB2	6:B:599:HOH:O	2.18	0.43
1:A:106:GLN:HG2	6:A:540:HOH:O	2.18	0.43
1:A:131:PRO:O	1:A:139:ARG:HD3	2.18	0.43
1:B:118:GLY:O	1:B:119:GLY:C	2.58	0.43
1:A:45:ARG:NH2	1:A:49:GLN:HG2	2.34	0.42
1:B:147:CYS:O	1:B:148:SER:O	2.37	0.42
1:A:2:ARG:C	1:B:10:MET:HE1	2.39	0.42
1:B:146:CYS:HB3	1:B:159:TYR:O	2.19	0.41
2:D:10:THR:HA	2:D:11:PRO:HD3	1.89	0.41
1:A:106:GLN:HG3	1:A:107:TRP:CE3	2.56	0.41
1:A:132:CYS:HA	1:A:139:ARG:CG	2.49	0.41
1:A:106:GLN:HB2	6:A:670:HOH:O	2.20	0.41
1:A:2:ARG:HB3	1:A:2:ARG:HE	1.71	0.41
1:A:9:ASN:ND2	6:A:546:HOH:O	2.51	0.40
4:B:173:GAL:H1	5:B:174:NDG:H4	1.68	0.40
1:B:37:ASN:ND2	6:B:599:HOH:O	2.49	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	169/171 (99%)	159 (94%)	7 (4%)	3 (2%)	10	4
1	B	169/171 (99%)	154 (91%)	10 (6%)	5 (3%)	5	1
2	D	6/8 (75%)	3 (50%)	1 (17%)	2 (33%)	0	0
2	E	1/8 (12%)	0	0	1 (100%)	0	0
All	All	345/358 (96%)	316 (92%)	18 (5%)	11 (3%)	5	1

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	7	TYR
1	B	140	VAL
1	B	148	SER
1	A	92	GLN
1	A	119	GLY
2	D	6	THR
1	B	119	GLY
1	B	125	ALA
1	A	93	ALA
1	B	164	CYS
2	E	6	THR

### 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	119/119 (100%)	114 (96%)	5 (4%)	34	30
1	B	119/119 (100%)	109 (92%)	10 (8%)	13	8
2	D	6/6 (100%)	5 (83%)	1 (17%)	2	1
2	E	2/6 (33%)	2 (100%)	0	100	100
All	All	246/250 (98%)	230 (94%)	16 (6%)	20	14

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	37	ASN
1	A	44	LYS
1	A	96	LYS
1	A	149	LYS
1	A	152	SER
2	D	7	TYR
1	B	2	ARG
1	B	16	LEU
1	B	33	LYS
1	B	37	ASN
1	B	49	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	134	LYS
1	B	150	TRP
1	B	152	SER
1	B	165	GLN
1	B	166	SER

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

Mol	Chain	Res	Type
1	A	9	ASN
1	A	37	ASN
1	A	63	GLN
1	A	106	GLN
1	B	37	ASN
1	B	49	GLN
1	B	59	GLN
1	B	79	GLN
1	B	165	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
1	PCA	A	1	1	8,8,9	1.02	0	9,10,12	2.43	4 (44%)
1	PCA	B	1	1	8,8,9	0.73	0	9,10,12	2.03	2 (22%)

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
1	PCA	A	1	1	-	0/0/11/13	0/1/1/1
1	PCA	B	1	1	-	0/0/11/13	0/1/1/1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1	PCA	CB-CA-C	-4.80	106.10	112.70
1	A	1	PCA	O-C-CA	-3.92	116.00	125.15
1	B	1	PCA	O-C-CA	-3.49	117.02	125.15
1	A	1	PCA	OE-CD-CG	-2.82	121.67	126.86
1	A	1	PCA	OE-CD-N	2.12	130.12	124.75
1	B	1	PCA	CB-CA-C	4.60	119.04	112.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1	PCA	1	0
1	B	1	PCA	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

8 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	$\# Z  > 2$
3	SIA	B	172	4	17,20,21	1.31	3 (17%)	19,28,31	1.37	2 (10%)
4	GAL	B	173	3,5	12,12,12	1.31	1 (8%)	17,17,17	2.43	4 (23%)
5	NDG	B	174	3,4	11,12,15	0.99	1 (9%)	14,16,21	2.56	5 (35%)
3	SIA	B	175	5	18,21,21	1.81	5 (27%)	19,31,31	1.69	5 (26%)
3	SIA	D	1	4	17,20,21	1.32	3 (17%)	19,28,31	1.69	5 (26%)
4	GAL	D	2	3,5	12,12,12	0.79	0	17,17,17	1.24	2 (11%)
5	NDG	D	3	3,4	11,12,15	1.56	3 (27%)	14,16,21	2.41	7 (50%)
3	SIA	D	4	5	18,21,21	1.50	4 (22%)	19,31,31	1.65	5 (26%)

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
3	SIA	B	172	4	-	0/14/34/38	0/1/1/1
4	GAL	B	173	3,5	-	0/2/22/22	0/1/1/1
5	NDG	B	174	3,4	1/1/4/7	0/4/17/26	0/1/1/1
3	SIA	B	175	5	-	0/14/38/38	0/1/1/1
3	SIA	D	1	4	-	0/14/34/38	0/1/1/1
4	GAL	D	2	3,5	-	0/2/22/22	0/1/1/1
5	NDG	D	3	3,4	1/1/4/7	0/4/17/26	0/1/1/1
3	SIA	D	4	5	-	0/14/38/38	0/1/1/1

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	173	GAL	O3-C3	-3.49	1.34	1.43
3	B	175	SIA	O6-C2	-3.08	1.39	1.43
3	D	4	SIA	C4-C5	-2.91	1.50	1.53
3	D	1	SIA	C3-C4	-2.58	1.48	1.52
5	D	3	NDG	O-C1	-2.48	1.39	1.43
3	D	1	SIA	O6-C6	-2.27	1.40	1.43
3	D	4	SIA	O6-C2	-2.25	1.40	1.43
3	B	172	SIA	O6-C6	-2.24	1.40	1.43
3	D	4	SIA	C3-C4	-2.19	1.49	1.53
3	B	175	SIA	C3-C4	-2.09	1.49	1.53
5	D	3	NDG	C3-C2	-2.06	1.49	1.52

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	174	NDG	C2-N2	2.09	1.50	1.46
3	D	1	SIA	O10-C10	2.27	1.28	1.23
3	B	172	SIA	O10-C10	2.31	1.28	1.23
3	B	175	SIA	C3-C2	2.47	1.54	1.51
3	B	175	SIA	C6-C5	2.58	1.57	1.53
5	D	3	NDG	C1-C2	3.22	1.55	1.51
3	D	4	SIA	C3-C2	3.23	1.55	1.51
3	B	172	SIA	C3-C2	3.42	1.58	1.52
3	B	175	SIA	O2-C2	4.98	1.45	1.39

All (35) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	174	NDG	O4-C4-C5	-6.58	96.72	110.00
5	B	174	NDG	O-C5-C4	-3.93	102.91	109.99
5	D	3	NDG	O4-C4-C5	-3.73	102.47	110.00
3	D	1	SIA	O9-C9-C8	-3.65	103.06	111.11
3	D	4	SIA	C8-C7-C6	-3.25	106.72	113.04
3	D	4	SIA	O6-C6-C7	-3.22	102.72	107.41
5	B	174	NDG	C6-C5-C4	-3.20	107.04	112.76
4	D	2	GAL	O2-C2-C1	-2.97	103.58	109.75
3	B	175	SIA	O6-C6-C7	-2.91	103.18	107.41
5	D	3	NDG	O7-C7-C8	-2.87	116.83	122.06
5	B	174	NDG	C1-C2-N2	-2.77	106.76	110.64
3	D	4	SIA	O9-C9-C8	-2.76	105.02	111.11
3	B	175	SIA	C11-C10-N5	-2.64	111.33	116.11
3	D	4	SIA	C6-C5-N5	-2.47	106.66	111.00
3	D	1	SIA	C9-C8-C7	-2.44	106.96	112.41
3	D	1	SIA	O8-C8-C9	-2.43	103.58	109.21
4	D	2	GAL	C1-C2-C3	-2.34	106.42	110.65
5	B	174	NDG	C2-N2-C7	-2.33	119.73	122.89
3	B	175	SIA	O2-C2-O6	-2.30	105.03	109.88
3	B	175	SIA	O8-C8-C7	-2.22	103.57	109.09
3	D	1	SIA	O7-C7-C6	2.20	114.37	109.46
3	B	172	SIA	O7-C7-C8	2.23	114.33	108.82
3	B	175	SIA	C4-C5-N5	2.25	115.04	110.40
3	D	4	SIA	C5-N5-C10	2.27	128.93	123.19
3	D	1	SIA	O4-C4-C3	2.53	116.14	110.02
5	D	3	NDG	C1-C2-N2	2.72	114.46	110.64
5	D	3	NDG	C1-O-C5	2.74	118.45	112.39
5	D	3	NDG	C2-N2-C7	2.89	126.81	122.89
5	D	3	NDG	O4-C4-C3	3.00	117.27	110.02

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	172	SIA	C5-N5-C10	3.07	130.97	123.19
4	B	173	GAL	O5-C1-C2	3.57	115.96	110.04
4	B	173	GAL	C6-C5-C4	3.61	121.45	113.00
5	D	3	NDG	C3-C2-N2	4.11	116.44	110.36
4	B	173	GAL	O3-C3-C2	4.18	119.46	110.36
4	B	173	GAL	O3-C3-C4	6.35	124.17	110.36

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	D	3	NDG	C4
5	B	174	NDG	C4

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	173	GAL	3	0
5	B	174	NDG	5	0
3	D	1	SIA	1	0
5	D	3	NDG	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

### 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<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	170/171 (99%)	-0.17	2 (1%) 79 78	11, 22, 40, 55	0
1	B	170/171 (99%)	0.18	12 (7%) 17 17	13, 25, 56, 67	0
2	D	8/8 (100%)	3.33	5 (62%) 0 1	51, 74, 87, 92	0
2	E	3/8 (37%)	4.85	3 (100%) 0 0	99, 99, 99, 100	0
All	All	351/358 (98%)	0.12	22 (6%) 21 21	11, 24, 56, 100	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	6	THR	7.1
2	D	7	TYR	7.0
2	D	8	ALA	5.6
1	B	133	GLY	5.3
2	D	6	THR	5.1
1	B	140	VAL	5.0
2	D	12	ARG	4.7
2	E	5	ALA	4.5
1	B	166	SER	4.0
1	B	171	GLY	3.7
1	B	147	CYS	3.7
1	B	132	CYS	3.6
1	B	138	GLY	3.3
2	E	7	PRO	3.0
1	A	93	ALA	2.9
1	B	165	GLN	2.9
2	D	5	ASP	2.6
1	A	171	GLY	2.5
1	B	164	CYS	2.2
1	B	139	ARG	2.2
1	B	152	SER	2.2
1	B	134	LYS	2.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	PCA	A	1	8/9	0.92	0.12	-	35,36,37,38	0
1	PCA	B	1	8/9	0.96	0.08	-	29,32,34,34	0

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	NDG	B	174	12/15	0.74	0.26	0.28	62,77,86,86	0
3	SIA	D	1	20/21	0.95	0.11	-0.05	16,21,28,30	0
3	SIA	B	175	21/21	0.92	0.10	-0.29	26,42,46,52	0
3	SIA	D	4	21/21	0.97	0.10	-0.71	13,20,26,36	0
5	NDG	D	3	12/15	0.97	0.08	-1.02	21,24,25,27	0
4	GAL	D	2	12/12	0.94	0.09	-	20,29,36,40	0
4	GAL	B	173	12/12	0.53	0.47	-	86,96,99,99	0
3	SIA	B	172	20/21	0.48	0.58	-	99,99,99,99	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.