



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2018 – 04:42 am GMT

PDB ID : 1OOK  
Title : Crystal Structure of the Complex of Platelet Receptor GPIb-alpha and Human alpha-Thrombin  
Authors : Varughese, K.I.; Celikel, R.; Ruggeri, Z.M.  
Deposited on : 2003-03-03  
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

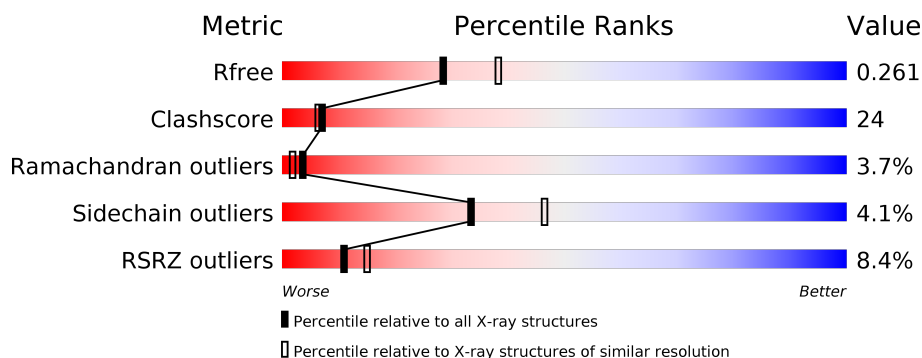
# 1 Overall quality at a glance

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))
$R_{free}$	111664	4477 (2.30-2.30)
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	36	<div> <div>17%</div> <div> <div></div> <div>64%</div> <div>33%</div> <div>•</div> </div> </div>
2	B	259	<div> <div>70%</div> <div>27%</div> <div>•</div> </div>
3	P	3	<div> <div>100%</div> </div>
4	G	290	<div> <div>14%</div> <div> <div>51%</div> <div>41%</div> <div>6%</div> <div>•</div> </div> </div>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 5111 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 Human Alpha Thrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	36	Total	C	N	O	S	0	0	0
			287	177	48	61	1			

- Molecule 2 is a protein called Human Alpha Thrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	259	Total	C	N	O	S	0	0	0
			2093	1334	370	375	14			

- Molecule 3 is a protein called PHE-PRO-ARG-Chloromethylketone.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	P	3	Total	C	N	O	0	0	0
			30	20	6	4			

- Molecule 4 is a protein called Platelet glycoprotein Ib alpha chain precursor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	G	284	Total	C	N	O	S	0	0	0
			2233	1427	363	433	10			

There are 3 discrepancies between the modelled and reference sequences:

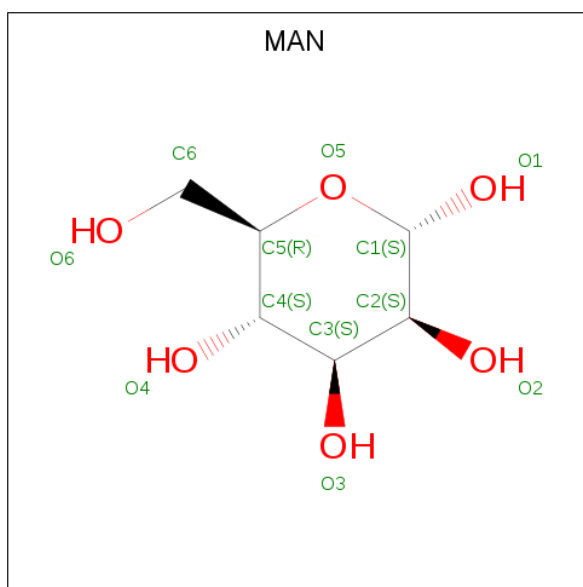
Chain	Residue	Modelled	Actual	Comment	Reference
G	65	ALA	CYS	ENGINEERED	UNP P07359
G	276	TYS	TYR	MODIFIED RESIDUE	UNP P07359
G	279	TYS	TYR	MODIFIED RESIDUE	UNP P07359

- Molecule 5 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 6 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula:  $C_6H_{12}O_6$ ).



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

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

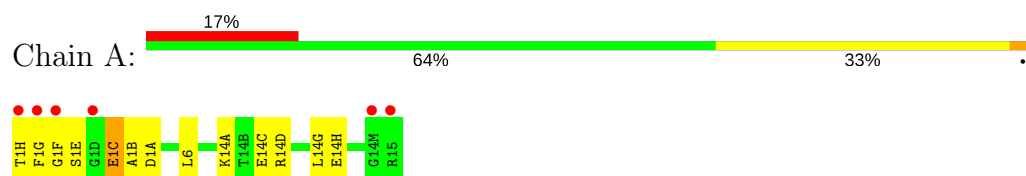
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	33	Total	O	0	0
			33	33		
8	B	239	Total	O	0	0
			239	239		
8	P	1	Total	O	0	0
			1	1		
8	G	119	Total	O	0	0
			119	119		

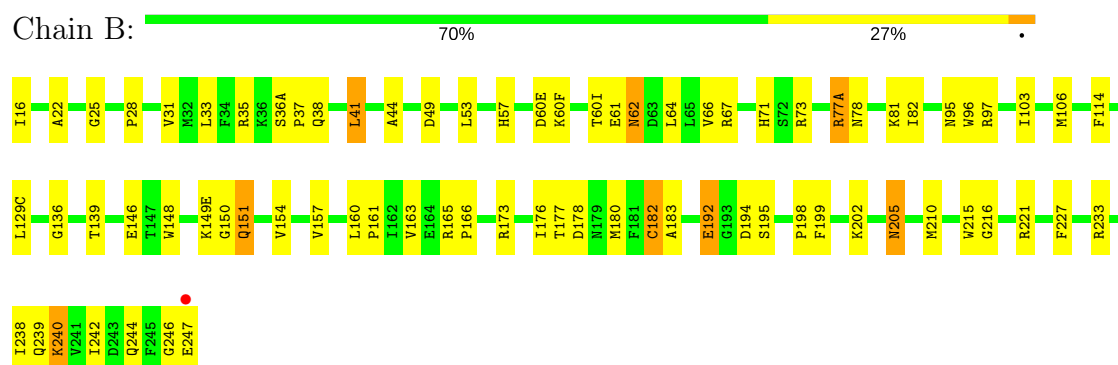
### 3 Residue-property plots [i](#)

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.

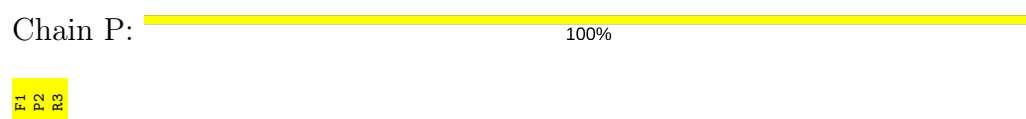
#### • Molecule 1: Human Alpha Thrombin



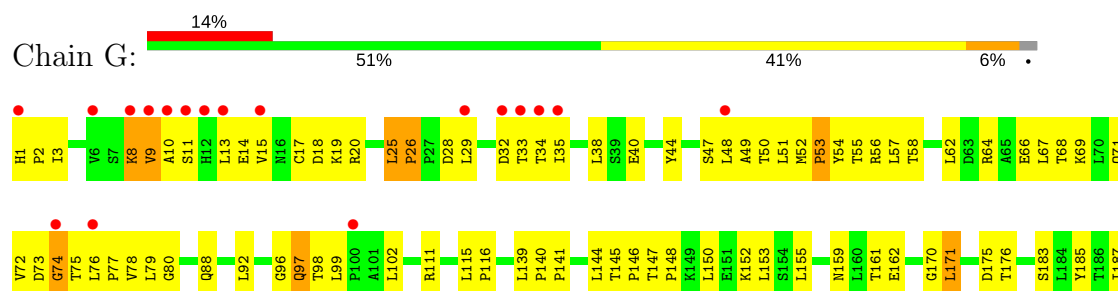
#### • Molecule 2: Human Alpha Thrombin

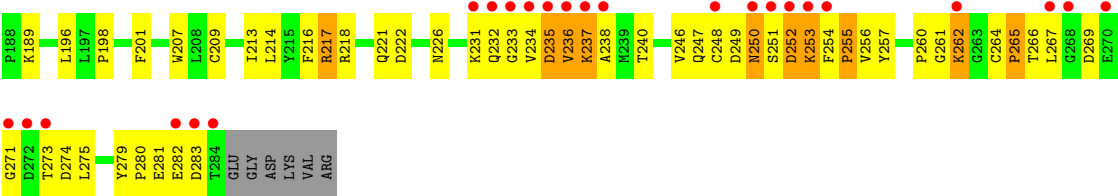


#### • Molecule 3: PHE-PRO-ARG-Chloromethylketone



#### • Molecule 4: Platelet glycoprotein Ib alpha chain precursor





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.65Å 67.65Å 328.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 2.30 47.84 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.2 (15.00-2.30) 98.3 (47.84-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.66 (at 2.29Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.212 , 0.260 0.215 , 0.261	Depositor DCC
$R_{free}$ test set	1722 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.7	Xtriage
Anisotropy	0.023	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 63.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5111	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.89% 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: CL, MAN, NAG, TYS

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	0.34	0/290	0.56	0/384
2	B	0.35	0/2148	0.65	0/2903
3	P	0.63	0/31	0.76	0/39
4	G	0.31	0/2248	0.58	0/3068
All	All	0.34	0/4717	0.61	0/6394

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [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	287	0	278	14	0
2	B	2093	0	2063	76	0
3	P	30	0	31	13	0
4	G	2233	0	2224	147	0
5	B	28	0	24	1	0
5	G	14	0	13	0	0
6	B	33	0	28	1	0
7	B	1	0	0	1	0
8	A	33	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	B	239	0	0	11	0
8	G	119	0	0	4	0
8	P	1	0	0	2	0
All	All	5111	0	4661	229	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 24.

All (229) 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)
4:G:25:LEU:HD23	4:G:25:LEU:H	1.13	1.11
2:B:149(E):LYS:HG3	4:G:283:ASP:HB3	1.47	0.95
4:G:49:ALA:HB2	4:G:73:ASP:HB2	1.47	0.95
4:G:273:THR:HG22	4:G:274:ASP:H	1.35	0.92
3:P:1:PHE:HB3	3:P:2:PRO:HD3	1.52	0.92
4:G:58:THR:HG22	4:G:78:VAL:HG12	1.54	0.89
2:B:165:ARG:HH11	2:B:165:ARG:HB3	1.35	0.88
4:G:17:CYS:HA	4:G:20:ARG:HE	1.38	0.88
2:B:165:ARG:NH1	2:B:165:ARG:HB3	1.89	0.86
2:B:57:HIS:NE2	3:P:3:ARG:OXT	2.09	0.85
7:B:400:CL:CL	8:P:299:HOH:O	2.34	0.82
4:G:14:GLU:HG3	4:G:35:ILE:HB	1.62	0.81
4:G:267:LEU:HD23	4:G:267:LEU:H	1.43	0.80
4:G:68:THR:HG22	4:G:88:GLN:HB2	1.62	0.79
2:B:202:LYS:HE2	2:B:205:ASN:ND2	1.99	0.78
1:A:1(H):THR:HA	8:B:587:HOH:O	1.83	0.78
2:B:150:GLY:H	4:G:281:GLU:HG3	1.49	0.77
1:A:1(F):GLY:HA2	8:A:48:HOH:O	1.84	0.76
4:G:25:LEU:HD11	4:G:50:THR:HG22	1.67	0.75
4:G:25:LEU:HD23	4:G:25:LEU:N	1.98	0.75
4:G:48:LEU:HD12	4:G:72:VAL:HA	1.67	0.75
4:G:25:LEU:CD2	4:G:25:LEU:H	1.93	0.73
3:P:1:PHE:CB	3:P:2:PRO:HD3	2.19	0.72
4:G:97:GLN:NE2	4:G:97:GLN:H	1.90	0.70
4:G:52:MET:HB3	4:G:53:PRO:HD3	1.74	0.70
2:B:195:SER:OG	3:P:3:ARG:C	2.31	0.69
4:G:214:LEU:O	4:G:217:ARG:HB3	1.93	0.69
2:B:67:ARG:CZ	8:B:531:HOH:O	2.41	0.69
4:G:145:THR:HB	4:G:146:PRO:HD3	1.75	0.69
4:G:218:ARG:HD2	4:G:271:GLY:H	1.58	0.69

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:G:25:LEU:HD21	4:G:50:THR:HG21	1.76	0.68
4:G:13:LEU:HB2	4:G:33:THR:HA	1.75	0.68
4:G:3:ILE:N	4:G:3:ILE:HD12	2.09	0.68
4:G:1:HIS:HB3	4:G:2:PRO:HD3	1.76	0.67
2:B:150:GLY:N	4:G:281:GLU:HG3	2.11	0.66
4:G:273:THR:HG22	4:G:274:ASP:N	2.07	0.66
2:B:139:THR:HG22	2:B:157:VAL:HG22	1.78	0.66
2:B:35:ARG:HB2	2:B:41:LEU:HD22	1.76	0.66
2:B:77(A):ARG:HG3	2:B:78:ASN:N	2.11	0.65
1:A:1(E):SER:HB2	2:B:49:ASP:OD1	1.96	0.65
2:B:216:GLY:O	3:P:1:PHE:HA	1.96	0.65
1:A:14(D):ARG:O	1:A:14(H):GLU:HG3	1.97	0.64
2:B:148:TRP:HA	4:G:279:TYS:O2	1.98	0.64
2:B:60(E):ASP:HB3	8:B:705:HOH:O	1.98	0.63
1:A:1(C):GLU:H	1:A:1(C):GLU:CD	2.02	0.62
4:G:175:ASP:HA	4:G:198:PRO:HD2	1.80	0.62
2:B:151:GLN:HG2	4:G:282:GLU:HG3	1.81	0.62
2:B:146:GLU:OE2	2:B:221:ARG:HD2	1.99	0.62
4:G:97:GLN:H	4:G:97:GLN:CD	2.03	0.62
2:B:60(I):THR:HG22	2:B:61:GLU:N	2.13	0.61
2:B:77(A):ARG:HG3	2:B:78:ASN:H	1.64	0.61
1:A:1(C):GLU:C	1:A:1(A):ASP:H	2.03	0.61
2:B:22:ALA:HB2	2:B:157:VAL:HG23	1.83	0.60
1:A:14(C):GLU:O	1:A:14(G):LEU:HD23	2.01	0.60
4:G:267:LEU:H	4:G:267:LEU:CD2	2.13	0.60
2:B:151:GLN:HG2	4:G:282:GLU:CG	2.32	0.59
4:G:139:LEU:HD11	4:G:155:LEU:HD13	1.83	0.59
4:G:146:PRO:C	4:G:148:PRO:HD3	2.23	0.58
4:G:19:LYS:HE3	4:G:40:GLU:OE1	2.03	0.58
3:P:1:PHE:HB3	3:P:2:PRO:CD	2.30	0.58
4:G:17:CYS:HA	4:G:20:ARG:NE	2.13	0.58
2:B:73:ARG:NH1	4:G:282:GLU:OE2	2.36	0.58
4:G:17:CYS:HB2	4:G:38:LEU:HD23	1.86	0.57
4:G:75:THR:HG22	4:G:77:PRO:HD3	1.87	0.57
2:B:22:ALA:HB2	2:B:157:VAL:CG2	2.34	0.57
4:G:13:LEU:HB2	4:G:32:ASP:O	2.04	0.57
4:G:80:GLY:O	4:G:102:LEU:HD12	2.06	0.56
4:G:253:LYS:C	4:G:255:PRO:HD3	2.25	0.56
2:B:178:ASP:O	2:B:233:ARG:HD3	2.05	0.56
2:B:31:VAL:CG1	2:B:66:VAL:HG13	2.35	0.56
4:G:25:LEU:HD11	4:G:50:THR:CG2	2.33	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:G:40:GLU:HA	4:G:64:ARG:O	2.06	0.56
2:B:62:ASN:HB2	8:B:641:HOH:O	2.05	0.55
2:B:37:PRO:HD3	4:G:273:THR:HG21	1.88	0.55
2:B:192:GLU:HG3	8:B:650:HOH:O	2.05	0.55
2:B:81:LYS:HD3	2:B:82:ILE:N	2.21	0.55
2:B:238:ILE:O	2:B:242:ILE:HG12	2.07	0.55
2:B:67:ARG:NE	8:B:531:HOH:O	2.40	0.55
4:G:150:LEU:HB3	4:G:171:LEU:HD12	1.89	0.55
4:G:267:LEU:HD23	4:G:267:LEU:N	2.20	0.55
4:G:15:VAL:HG11	4:G:29:LEU:HD13	1.89	0.54
4:G:252:ASP:O	4:G:254:PHE:N	2.41	0.54
4:G:14:GLU:HG3	4:G:35:ILE:CB	2.35	0.53
4:G:74:GLY:O	4:G:98:THR:HG23	2.09	0.53
4:G:232:GLN:HG3	4:G:235:ASP:OD1	2.08	0.53
4:G:141:PRO:HA	8:G:645:HOH:O	2.08	0.53
4:G:237:LYS:HA	4:G:237:LYS:HE2	1.90	0.53
4:G:75:THR:O	4:G:76:LEU:HD23	2.08	0.53
2:B:165:ARG:HB2	2:B:166:PRO:HD3	1.91	0.53
4:G:47:SER:HA	4:G:71:GLN:HB2	1.91	0.53
4:G:8:LYS:HB3	4:G:13:LEU:HD13	1.91	0.53
1:A:14(A):LYS:HB2	8:B:624:HOH:O	2.08	0.52
4:G:251:SER:O	4:G:252:ASP:C	2.48	0.52
4:G:79:LEU:HD23	4:G:99:LEU:HD13	1.92	0.52
4:G:10:ALA:O	4:G:11:SER:HB3	2.10	0.52
4:G:53:PRO:HB2	4:G:54:TYR:CD1	2.44	0.52
1:A:1(C):GLU:O	1:A:1(B):ALA:HB3	2.10	0.52
4:G:34:THR:O	4:G:57:LEU:HD12	2.10	0.52
2:B:77(A):ARG:CG	2:B:78:ASN:N	2.72	0.51
4:G:248:CYS:O	4:G:251:SER:N	2.37	0.51
2:B:246:GLY:O	2:B:247:GLU:HB3	2.09	0.51
4:G:20:ARG:HH11	4:G:20:ARG:HG2	1.75	0.51
3:P:1:PHE:CB	3:P:2:PRO:CD	2.87	0.51
2:B:239:GLN:HG2	8:B:544:HOH:O	2.10	0.51
4:G:209:CYS:HA	4:G:213:ILE:HG21	1.92	0.51
2:B:246:GLY:O	2:B:247:GLU:CB	2.58	0.51
4:G:88:GLN:OE1	4:G:111:ARG:HD2	2.11	0.51
4:G:273:THR:CG2	4:G:274:ASP:H	2.10	0.51
4:G:67:LEU:HD23	4:G:68:THR:N	2.25	0.51
4:G:161:THR:HA	4:G:183:SER:O	2.09	0.50
2:B:77(A):ARG:HG3	2:B:78:ASN:ND2	2.26	0.50
2:B:151:GLN:CG	4:G:282:GLU:HG3	2.41	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:53:LEU:HD11	2:B:103:ILE:HD11	1.94	0.50
4:G:236:VAL:O	4:G:238:ALA:N	2.37	0.50
2:B:195:SER:CB	3:P:3:ARG:O	2.60	0.49
2:B:33:LEU:HD22	2:B:64:LEU:HD13	1.94	0.49
4:G:171:LEU:HD22	4:G:171:LEU:N	2.27	0.49
4:G:8:LYS:O	4:G:9:VAL:HB	2.13	0.49
4:G:251:SER:OG	4:G:254:PHE:HB2	2.13	0.49
4:G:97:GLN:NE2	4:G:97:GLN:N	2.60	0.49
4:G:147:THR:N	4:G:148:PRO:HD3	2.28	0.49
2:B:31:VAL:HB	2:B:44:ALA:HB3	1.95	0.48
4:G:34:THR:HA	4:G:56:ARG:O	2.13	0.48
2:B:60(I):THR:CG2	2:B:61:GLU:N	2.76	0.48
4:G:13:LEU:CB	4:G:33:THR:HA	2.43	0.48
4:G:49:ALA:HB2	4:G:73:ASP:CB	2.31	0.48
3:P:1:PHE:CG	3:P:2:PRO:HD3	2.47	0.48
4:G:273:THR:O	4:G:274:ASP:HB3	2.13	0.48
4:G:218:ARG:O	4:G:221:GLN:HB2	2.13	0.48
2:B:195:SER:HB2	3:P:3:ARG:O	2.14	0.48
4:G:254:PHE:CD1	4:G:254:PHE:N	2.82	0.48
4:G:29:LEU:HB2	4:G:54:TYR:HE2	1.79	0.48
4:G:58:THR:HA	4:G:79:LEU:HA	1.96	0.48
4:G:9:VAL:C	4:G:11:SER:H	2.16	0.48
2:B:95:ASN:CG	8:B:589:HOH:O	2.51	0.47
4:G:67:LEU:HD23	4:G:69:LYS:N	2.28	0.47
4:G:78:VAL:HG12	4:G:78:VAL:O	2.14	0.47
2:B:163:VAL:HG23	2:B:183:ALA:HA	1.96	0.47
4:G:247:GLN:NE2	4:G:252:ASP:O	2.48	0.47
2:B:136:GLY:HA3	2:B:199:PHE:CZ	2.49	0.47
2:B:150:GLY:CA	4:G:281:GLU:HG3	2.45	0.47
4:G:170:GLY:O	4:G:171:LEU:C	2.53	0.47
2:B:165:ARG:CB	2:B:165:ARG:HH11	2.17	0.47
4:G:171:LEU:HD22	4:G:171:LEU:H	1.79	0.47
4:G:75:THR:HG22	4:G:77:PRO:CD	2.45	0.46
4:G:217:ARG:HE	4:G:257:TYR:HA	1.80	0.46
2:B:150:GLY:HA3	4:G:281:GLU:HG3	1.98	0.46
4:G:249:ASP:C	4:G:251:SER:N	2.69	0.46
2:B:77(A):ARG:NH1	4:G:175:ASP:OD2	2.48	0.46
4:G:144:LEU:HD11	4:G:153:LEU:HD22	1.97	0.46
4:G:249:ASP:O	4:G:249:ASP:OD1	2.34	0.46
2:B:60(I):THR:HG22	2:B:61:GLU:H	1.81	0.45
1:A:1(G):PHE:HB2	8:A:382:HOH:O	2.17	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:G:279:TYS:N	4:G:280:PRO:HD3	2.30	0.45
1:A:6:LEU:HD12	2:B:25:GLY:HA3	1.98	0.45
2:B:71:HIS:CD2	2:B:154:VAL:HG22	2.51	0.45
2:B:41:LEU:N	2:B:41:LEU:HD23	2.31	0.45
4:G:159:ASN:HB3	8:G:658:HOH:O	2.17	0.45
4:G:217:ARG:HD2	4:G:256:VAL:O	2.17	0.45
4:G:68:THR:HG22	4:G:88:GLN:CB	2.42	0.45
2:B:176:ILE:HG22	2:B:177:THR:N	2.32	0.45
4:G:53:PRO:O	4:G:54:TYR:C	2.56	0.44
2:B:31:VAL:HG13	2:B:66:VAL:HG13	1.98	0.44
4:G:44:TYR:HE2	4:G:66:GLU:OE2	2.01	0.44
4:G:49:ALA:CB	4:G:73:ASP:HB2	2.34	0.44
1:A:1(C):GLU:HB2	8:A:348:HOH:O	2.18	0.44
4:G:25:LEU:HD21	4:G:50:THR:CG2	2.47	0.44
2:B:114:PHE:HB3	8:B:592:HOH:O	2.16	0.44
2:B:41:LEU:H	2:B:41:LEU:HD23	1.83	0.44
4:G:249:ASP:C	4:G:251:SER:H	2.19	0.44
5:B:502:NAG:O3	6:B:503:MAN:C1	2.66	0.43
4:G:140:PRO:HA	4:G:141:PRO:HD3	1.86	0.43
4:G:38:LEU:HB2	4:G:62:LEU:HD23	1.99	0.43
4:G:3:ILE:CD1	4:G:3:ILE:N	2.79	0.43
4:G:54:TYR:CD1	4:G:54:TYR:N	2.86	0.43
2:B:36(A):SER:HA	2:B:37:PRO:HA	1.91	0.43
4:G:9:VAL:C	4:G:11:SER:N	2.72	0.43
4:G:201:PHE:CZ	4:G:237:LYS:HG3	2.54	0.43
2:B:163:VAL:HB	2:B:182:CYS:SG	2.59	0.43
4:G:214:LEU:HA	4:G:217:ARG:HB3	1.99	0.43
4:G:250:ASN:HA	4:G:250:ASN:HD22	1.61	0.43
4:G:254:PHE:O	4:G:255:PRO:C	2.56	0.43
3:P:3:ARG:C	8:P:299:HOH:O	2.56	0.43
4:G:76:LEU:N	4:G:77:PRO:HD3	2.34	0.43
2:B:16:ILE:N	2:B:194:ASP:OD2	2.52	0.42
4:G:152:LYS:HG2	4:G:176:THR:HB	2.01	0.42
4:G:218:ARG:HA	4:G:221:GLN:OE1	2.19	0.42
4:G:231:LYS:HB2	4:G:237:LYS:NZ	2.34	0.42
4:G:25:LEU:HA	4:G:26:PRO:HD3	1.86	0.42
1:A:1(F):GLY:O	1:A:1(E):SER:HB2	2.19	0.42
2:B:240:LYS:O	2:B:244:GLN:HB2	2.19	0.42
4:G:10:ALA:O	4:G:11:SER:CB	2.66	0.42
2:B:67:ARG:NH2	8:B:531:HOH:O	2.51	0.42
4:G:262:LYS:HB2	8:G:694:HOH:O	2.20	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:G:246:VAL:HG12	4:G:256:VAL:CG2	2.49	0.42
4:G:260:PRO:C	4:G:262:LYS:H	2.23	0.42
2:B:57:HIS:O	2:B:60(F):LYS:HE3	2.19	0.42
4:G:222:ASP:HA	8:G:633:HOH:O	2.20	0.42
2:B:33:LEU:HD21	2:B:106:MET:HE1	2.01	0.42
4:G:49:ALA:HB2	4:G:74:GLY:H	1.84	0.42
4:G:51:LEU:O	4:G:52:MET:C	2.57	0.42
4:G:73:ASP:O	4:G:74:GLY:O	2.37	0.42
4:G:64:ARG:HG3	4:G:64:ARG:HH11	1.84	0.42
1:A:1(C):GLU:C	1:A:1(A):ASP:N	2.71	0.42
4:G:187:ILE:HB	4:G:216:PHE:HB2	2.01	0.42
4:G:201:PHE:N	4:G:201:PHE:CD1	2.87	0.42
2:B:67:ARG:NH1	4:G:196:LEU:HD23	2.35	0.41
2:B:35:ARG:O	2:B:38:GLN:HA	2.19	0.41
4:G:281:GLU:O	4:G:281:GLU:HG3	2.20	0.41
4:G:9:VAL:O	4:G:10:ALA:HB3	2.20	0.41
4:G:254:PHE:N	4:G:255:PRO:HD3	2.34	0.41
2:B:216:GLY:O	3:P:1:PHE:CA	2.67	0.41
4:G:18:ASP:O	4:G:20:ARG:HG3	2.20	0.41
4:G:14:GLU:HG3	4:G:35:ILE:CG2	2.51	0.41
2:B:160:LEU:HA	2:B:161:PRO:HD3	1.84	0.41
2:B:25:GLY:O	2:B:28:PRO:HD3	2.20	0.41
4:G:273:THR:HA	4:G:275:LEU:HD13	2.02	0.41
4:G:47:SER:HB2	4:G:71:GLN:NE2	2.35	0.41
2:B:96:TRP:CZ3	2:B:97:ARG:HG3	2.56	0.41
4:G:264:CYS:HA	4:G:265:PRO:HD3	1.94	0.41
4:G:218:ARG:HB3	4:G:271:GLY:CA	2.51	0.41
4:G:8:LYS:HB3	4:G:13:LEU:CD1	2.50	0.40
4:G:185:TYR:HA	4:G:207:TRP:HA	2.02	0.40
4:G:233:GLY:C	4:G:235:ASP:H	2.24	0.40
4:G:92:LEU:HD22	4:G:115:LEU:HD22	2.04	0.40
2:B:215:TRP:HB2	3:P:1:PHE:O	2.21	0.40
2:B:176:ILE:HD12	2:B:227:PHE:CE2	2.56	0.40
2:B:38:GLN:HE22	4:G:226:ASN:HD21	1.69	0.40
4:G:261:GLY:O	4:G:264:CYS:SG	2.80	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	34/36 (94%)	27 (79%)	7 (21%)	0	100	100
2	B	257/259 (99%)	246 (96%)	10 (4%)	1 (0%)	36	45
3	P	1/3 (33%)	1 (100%)	0	0	100	100
4	G	280/290 (97%)	220 (79%)	40 (14%)	20 (7%)	1	0
All	All	572/588 (97%)	494 (86%)	57 (10%)	21 (4%)	4	2

All (21) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	77(A)	ARG
4	G	9	VAL
4	G	28	ASP
4	G	189	LYS
4	G	234	VAL
4	G	240	THR
4	G	252	ASP
4	G	253	LYS
4	G	266	THR
4	G	74	GLY
4	G	237	LYS
4	G	262	LYS
4	G	53	PRO
4	G	55	THR
4	G	171	LEU
4	G	265	PRO
4	G	236	VAL
4	G	269	ASP
4	G	96	GLY
4	G	116	PRO
4	G	26	PRO



### 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	31/31 (100%)	30 (97%)	1 (3%)	42	58
2	B	225/225 (100%)	213 (95%)	12 (5%)	25	34
3	P	3/3 (100%)	3 (100%)	0	100	100
4	G	252/257 (98%)	244 (97%)	8 (3%)	42	58
All	All	511/516 (99%)	490 (96%)	21 (4%)	33	46

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1(C)	GLU
2	B	41	LEU
2	B	62	ASN
2	B	129(C)	LEU
2	B	151	GLN
2	B	173	ARG
2	B	180	MET
2	B	182	CYS
2	B	192	GLU
2	B	198	PRO
2	B	205	ASN
2	B	210	MET
2	B	240	LYS
4	G	8	LYS
4	G	25	LEU
4	G	97	GLN
4	G	162	GLU
4	G	217	ARG
4	G	235	ASP
4	G	250	ASN
4	G	255	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	62	ASN
2	B	78	ASN
2	B	205	ASN
4	G	71	GLN
4	G	97	GLN
4	G	226	ASN
4	G	247	GLN
4	G	250	ASN

### 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	$\# Z  > 2$
4	TYS	G	276	4	15,16,17	1.04	2 (13%)	19,22,24	1.18	2 (10%)
4	TYS	G	279	4	15,16,17	1.10	2 (13%)	19,22,24	1.01	1 (5%)

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
4	TYS	G	276	4	-	0/9/11/13	0/1/1/1
4	TYS	G	279	4	-	0/9/11/13	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	276	TYS	OH-CZ	-2.52	1.38	1.42
4	G	279	TYS	OH-CZ	-2.49	1.38	1.42
4	G	276	TYS	CA-C	2.07	1.53	1.50
4	G	279	TYS	CA-C	2.56	1.53	1.50

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	276	TYS	OH-CZ-CE2	-2.22	114.46	118.69
4	G	279	TYS	OH-CZ-CE1	2.47	123.39	118.69
4	G	276	TYS	OH-CZ-CE1	3.25	124.88	118.69

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	G	279	TYS	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 1 is monoatomic - leaving 6 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	NAG	B	501	2,5	14,14,15	0.59	0	17,19,21	0.61	0
5	NAG	B	502	5,6	14,14,15	0.68	0	17,19,21	0.89	1 (5%)
6	MAN	B	503	5,6	11,11,12	0.68	0	15,15,17	0.69	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	MAN	B	504	6	11,11,12	0.49	0	15,15,17	0.60	1 (6%)
6	MAN	B	508	6	11,11,12	0.52	0	15,15,17	0.49	0
5	NAG	G	601	4	14,14,15	0.60	0	17,19,21	0.63	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	NAG	B	501	2,5	-	0/6/23/26	0/1/1/1
5	NAG	B	502	5,6	-	0/6/23/26	0/1/1/1
6	MAN	B	503	5,6	-	0/2/19/22	0/1/1/1
6	MAN	B	504	6	-	0/2/19/22	0/1/1/1
6	MAN	B	508	6	-	0/2/19/22	0/1/1/1
5	NAG	G	601	4	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	502	NAG	C2-N2-C7	-2.10	119.88	122.94
6	B	504	MAN	C1-O5-C5	2.13	115.11	112.19

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	502	NAG	1	0
6	B	503	MAN	1	0

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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	36/36 (100%)	0.88	6 (16%) ⓘ ⓘ	24, 45, 115, 119	0
2	B	259/259 (100%)	-0.05	1 (0%) ⓘ ⓘ	17, 31, 52, 86	0
3	P	3/3 (100%)	0.35	0 ⓘ ⓘ	23, 23, 24, 25	0
4	G	282/290 (97%)	0.88	42 (14%) ⓘ ⓘ	29, 56, 113, 124	0
All	All	580/588 (98%)	0.46	49 (8%) ⓘ ⓘ	17, 43, 101, 124	0

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	G	234	VAL	11.9
4	G	12	HIS	7.8
1	A	1(H)	THR	7.2
4	G	284	THR	7.0
4	G	11	SER	6.5
1	A	1(F)	GLY	6.2
4	G	251	SER	6.2
4	G	237	LYS	6.1
4	G	1	HIS	5.9
1	A	1(G)	PHE	5.8
4	G	271	GLY	5.7
4	G	267	LEU	5.4
4	G	233	GLY	5.3
1	A	15	ARG	5.2
4	G	272	ASP	5.2
1	A	14(M)	GLY	5.0
4	G	74	GLY	4.8
4	G	8	LYS	4.8
4	G	268	GLY	4.7
4	G	15	VAL	4.6
4	G	76	LEU	4.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	G	10	ALA	4.5
4	G	33	THR	4.3
4	G	254	PHE	4.2
4	G	236	VAL	4.2
4	G	9	VAL	4.1
4	G	235	ASP	4.0
4	G	282	GLU	3.9
4	G	270	GLU	3.8
4	G	238	ALA	3.4
4	G	6	VAL	3.3
1	A	1(D)	GLY	3.2
4	G	283	ASP	3.1
4	G	252	ASP	3.0
4	G	273	THR	3.0
4	G	253	LYS	3.0
4	G	232	GLN	2.8
4	G	250	ASN	2.8
4	G	34	THR	2.6
4	G	29	LEU	2.5
4	G	262	LYS	2.5
4	G	248	CYS	2.4
4	G	32	ASP	2.3
4	G	13	LEU	2.3
4	G	35	ILE	2.2
2	B	247	GLU	2.2
4	G	100	PRO	2.2
4	G	231	LYS	2.1
4	G	48	LEU	2.0

## 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. 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
4	TYS	G	276	16/17	0.92	0.16	41,50,75,76	0
4	TYS	G	279	16/17	0.95	0.16	30,42,56,59	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. 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( $\text{\AA}^2$ )	Q<0.9
6	MAN	B	503	11/12	0.74	0.22	83,89,92,96	0
5	NAG	G	601	14/15	0.77	0.20	66,78,82,82	0
5	NAG	B	502	14/15	0.79	0.17	69,80,85,86	0
6	MAN	B	508	11/12	0.81	0.38	93,99,102,102	0
6	MAN	B	504	11/12	0.85	0.19	74,78,82,85	0
5	NAG	B	501	14/15	0.88	0.13	61,68,79,79	0
7	CL	B	400	1/1	0.98	0.33	35,35,35,35	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.