Full wwPDB X-ray Structure Validation Report (i)

Jun 12, 2016 – 09:05 AM EDT

PDB ID	:	5EOK
Title	:	Human Plasma Coagulation Factor XI in complex with peptide P39
Authors	:	Wong, S.S.; Ostergaard, S.; Hall, G.; Li, C.; Williams, P.M.; Stennicke, H.; Emsley, J.
Deposited on	:	2015-11-10
Resolution	:	2.80 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

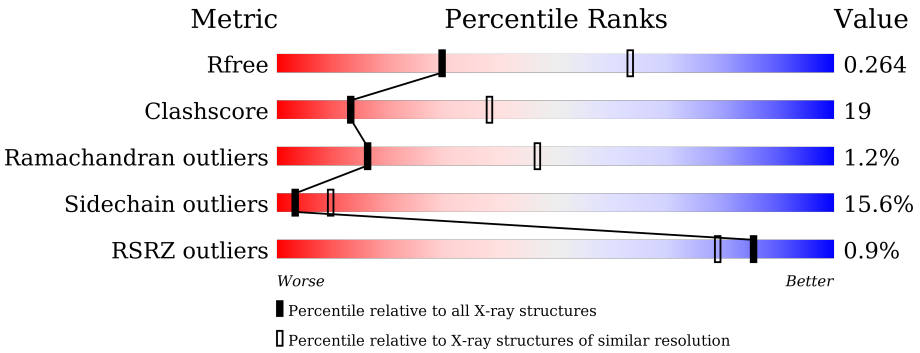
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027674
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027674

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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 >=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 <=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	606	% 57% 31% 7%
2	K	9	11% 89% 11%

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	NAG	A	701	X	-	-	-
3	NAG	A	703	X	-	-	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4785 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.

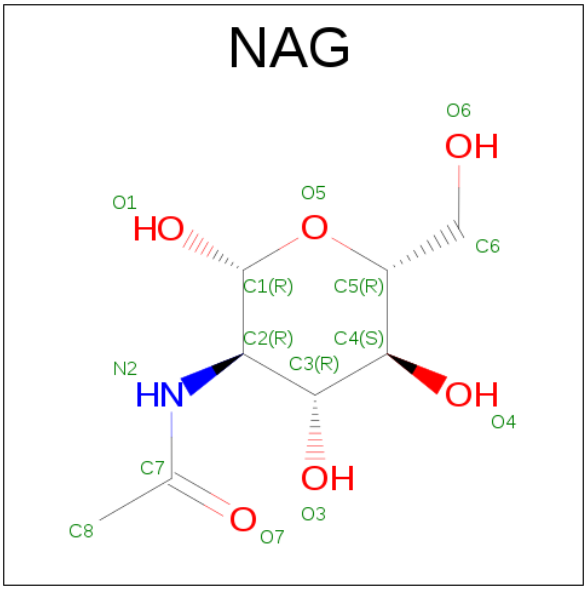
- Molecule 1 is a protein called Coagulation factor XI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	589	Total	C	N	O	S	100	0	0
			4625	2915	801	869	40			

- Molecule 2 is a protein called P39.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	K	9	Total	C	N	O	0	0	0
			78	52	11	15			

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



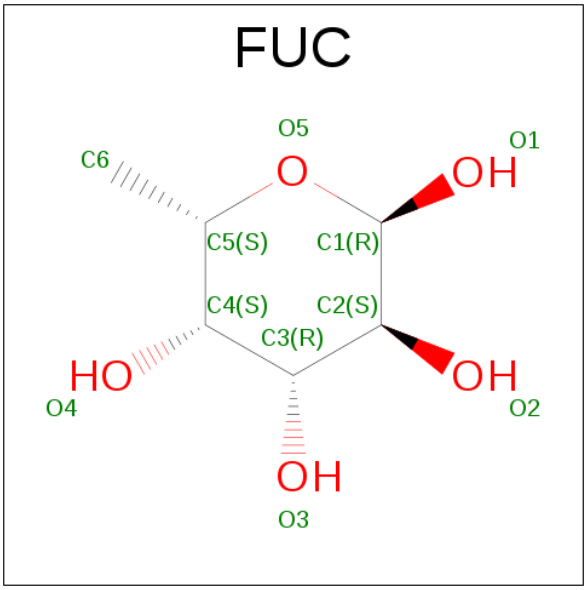
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

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

- Molecule 4 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula: $C_6H_{12}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 10 6 4	0	0

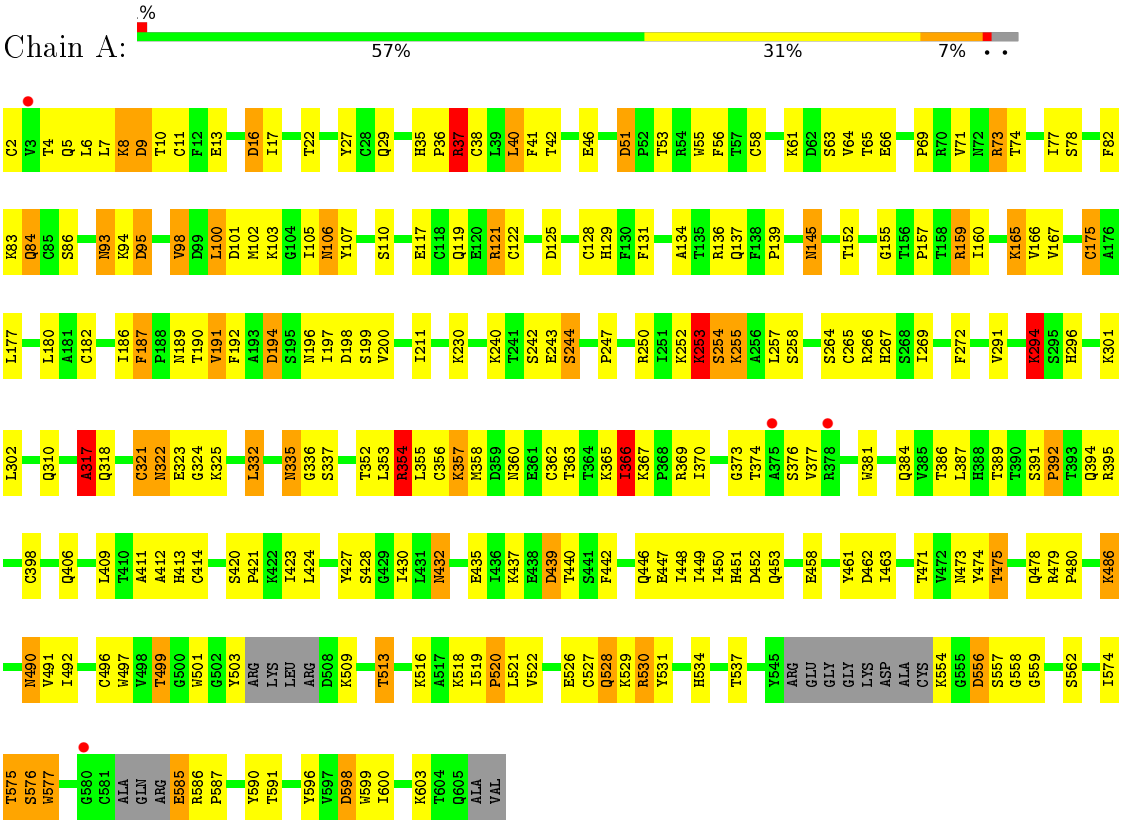
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	2	Total O 2 2	0	0

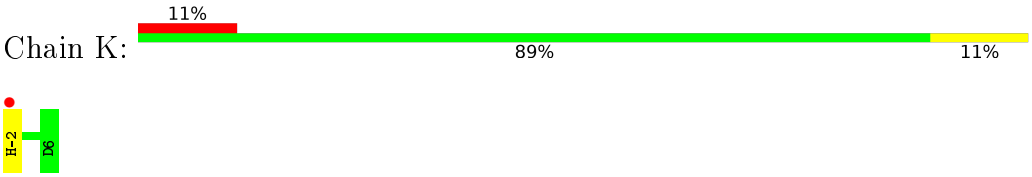
3 Residue-property plots (i)

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



- Molecule 2: P39



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	81.33Å 81.33Å 252.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.47 – 2.80 42.87 – 2.80	Depositor EDS
% Data completeness (in resolution range)	95.1 (47.47-2.80) 95.1 (42.87-2.80)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.32 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.201 , 0.264 0.200 , 0.264	Depositor DCC
R_{free} test set	1065 reflections (5.42%)	DCC
Wilson B-factor (Å^2)	78.7	Xtriage
Anisotropy	0.108	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 44.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4785	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.89% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: NAG, FUC

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	0.99	20/4733 (0.4%)	1.03	25/6410 (0.4%)
2	K	0.69	0/82	0.73	0/113
All	All	0.98	20/4815 (0.4%)	1.03	25/6523 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	5

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	321	CYS	C-N	-22.42	0.82	1.34
1	A	323	GLU	CA-C	-16.24	1.10	1.52
1	A	253	LYS	C-N	-12.06	1.06	1.34
1	A	317	ALA	C-N	-11.88	1.06	1.34
1	A	366	ILE	C-N	-11.77	1.06	1.34
1	A	78	SER	CA-CB	-10.83	1.36	1.52
1	A	253	LYS	CA-C	-10.78	1.25	1.52
1	A	78	SER	C-N	-10.71	1.13	1.33
1	A	253	LYS	CG-CD	-9.26	1.21	1.52
1	A	294	LYS	C-N	-8.80	1.13	1.34
1	A	324	GLY	C-N	8.66	1.53	1.34
1	A	323	GLU	CB-CG	-8.51	1.35	1.52
1	A	253	LYS	CA-CB	8.29	1.72	1.53
1	A	373	GLY	C-N	-6.47	1.19	1.34
1	A	78	SER	CA-C	-6.35	1.36	1.52
1	A	252	LYS	C-N	6.35	1.48	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	323	GLU	CA-CB	6.13	1.67	1.53
1	A	294	LYS	CA-C	-6.07	1.37	1.52
1	A	577	TRP	CD2-CE2	5.92	1.48	1.41
1	A	381	TRP	CD2-CE2	5.09	1.47	1.41

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	317	ALA	O-C-N	-19.34	91.75	122.70
1	A	323	GLU	CB-CA-C	17.04	144.48	110.40
1	A	366	ILE	O-C-N	-15.54	97.84	122.70
1	A	321	CYS	O-C-N	-13.30	101.42	122.70
1	A	253	LYS	CB-CA-C	-10.69	89.02	110.40
1	A	78	SER	CB-CA-C	10.12	129.34	110.10
1	A	323	GLU	CA-CB-CG	9.67	134.67	113.40
1	A	321	CYS	C-N-CA	9.47	145.38	121.70
1	A	317	ALA	CA-C-N	8.66	136.26	117.20
1	A	317	ALA	C-N-CA	8.48	142.91	121.70
1	A	366	ILE	CA-C-N	8.24	135.32	117.20
1	A	324	GLY	C-N-CA	-7.85	102.08	121.70
1	A	253	LYS	CA-C-N	-7.14	101.49	117.20
1	A	294	LYS	CB-CG-CD	7.03	129.89	111.60
1	A	324	GLY	O-C-N	6.47	133.05	122.70
1	A	78	SER	CA-C-N	6.39	128.98	116.20
1	A	321	CYS	CA-C-N	6.38	131.23	117.20
1	A	324	GLY	CA-C-N	-6.33	103.28	117.20
1	A	322	ASN	N-CA-CB	-6.10	99.61	110.60
1	A	373	GLY	C-N-CA	5.75	136.07	121.70
1	A	9	ASP	N-CA-C	-5.72	95.55	111.00
1	A	37	ARG	NE-CZ-NH1	5.56	123.08	120.30
1	A	253	LYS	CG-CD-CE	-5.16	96.44	111.90
1	A	253	LYS	CA-C-O	5.15	130.92	120.10
1	A	121	ARG	NE-CZ-NH2	-5.07	117.77	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	323	GLU	CA

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	253	LYS	Mainchain
1	A	317	ALA	Mainchain,Peptide
1	A	321	CYS	Mainchain
1	A	366	ILE	Mainchain

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4625	0	4490	173	0
2	K	78	0	67	0	0
3	A	70	0	63	2	0
4	A	10	0	10	1	0
5	A	2	0	0	0	0
All	All	4785	0	4630	174	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 19.

All (174) 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:360:ASN:ND2	1:A:362:CYS:SG	2.30	1.02
1:A:103:LYS:HG3	1:A:159:ARG:HB3	1.50	0.93
1:A:93:ASN:HD21	1:A:95:ASP:HB2	1.36	0.90
1:A:189:ASN:HD22	1:A:254:SER:HA	1.39	0.88
1:A:8:LYS:HE2	1:A:9:ASP:HB2	1.57	0.87
1:A:389:THR:O	1:A:394:GLN:HA	1.77	0.84
1:A:37:ARG:HG3	1:A:37:ARG:HH11	1.43	0.81
1:A:462:ASP:OD2	1:A:576:SER:HB2	1.80	0.81
1:A:432:ASN:HD22	1:A:513:THR:HG23	1.46	0.80
1:A:406:GLN:HG3	1:A:471:THR:HG22	1.63	0.78
1:A:129:HIS:CD2	1:A:155:GLY:HA2	2.22	0.74
1:A:530:ARG:NH2	1:A:585:GLU:O	2.20	0.74
1:A:474:TYR:CE1	1:A:480:PRO:HD3	2.22	0.74
1:A:596:TYR:O	1:A:600:ILE:HG13	1.88	0.74
1:A:93:ASN:HD22	1:A:93:ASN:C	1.91	0.73

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:576:SER:HG	1:A:577:TRP:HD1	1.35	0.72
1:A:496:CYS:SG	1:A:521:LEU:HD11	2.29	0.72
1:A:530:ARG:HH21	1:A:586:ARG:HA	1.54	0.71
1:A:106:ASN:ND2	1:A:106:ASN:H	1.89	0.71
1:A:189:ASN:ND2	1:A:254:SER:HA	2.05	0.71
1:A:137:GLN:HB2	1:A:165:LYS:HD3	1.72	0.71
1:A:187:PHE:CD2	1:A:187:PHE:N	2.58	0.70
1:A:93:ASN:ND2	1:A:95:ASP:H	1.90	0.70
1:A:269:ILE:HG13	1:A:269:ILE:O	1.92	0.69
1:A:355:LEU:HA	1:A:358:MET:HE1	1.75	0.69
1:A:317:ALA:HB1	1:A:322:ASN:HB3	1.76	0.68
1:A:40:LEU:HD12	1:A:40:LEU:N	2.09	0.67
1:A:406:GLN:CG	1:A:471:THR:HG22	2.24	0.67
1:A:486:LYS:HE3	1:A:598:ASP:OD1	1.94	0.66
1:A:355:LEU:HA	1:A:358:MET:CE	2.25	0.66
1:A:575:THR:HA	1:A:590:TYR:CD1	2.30	0.66
1:A:187:PHE:N	1:A:187:PHE:HD2	1.92	0.66
1:A:102:MET:HE2	1:A:157:PRO:HG3	1.78	0.66
1:A:182:CYS:SG	1:A:264:SER:HB3	2.36	0.65
1:A:13:GLU:HB3	1:A:69:PRO:HG2	1.78	0.65
1:A:9:ASP:OD2	1:A:74:THR:HA	1.97	0.65
1:A:530:ARG:HE	1:A:586:ARG:HG3	1.60	0.65
1:A:106:ASN:HD22	1:A:106:ASN:H	1.45	0.65
1:A:17:ILE:HD12	1:A:61:LYS:HA	1.80	0.64
1:A:119:GLN:HA	1:A:131:PHE:CE1	2.33	0.63
1:A:192:PHE:O	1:A:240:LYS:HE2	1.99	0.62
1:A:8:LYS:CE	1:A:9:ASP:HB2	2.28	0.62
1:A:374:THR:O	1:A:374:THR:HG22	1.99	0.62
1:A:521:LEU:H	1:A:521:LEU:HD12	1.65	0.62
1:A:125:ASP:HB2	4:A:704:FUC:H62	1.81	0.62
1:A:432:ASN:ND2	1:A:513:THR:HG23	2.13	0.62
1:A:134:ALA:HA	1:A:166:VAL:HG23	1.81	0.61
1:A:37:ARG:CG	1:A:37:ARG:HH11	2.11	0.61
1:A:522:VAL:HG23	1:A:526:GLU:HB3	1.82	0.61
1:A:93:ASN:ND2	1:A:93:ASN:C	2.52	0.61
1:A:411:ALA:O	1:A:414:CYS:HB2	2.01	0.60
1:A:575:THR:HG22	1:A:590:TYR:HE1	1.66	0.60
1:A:291:VAL:HG21	1:A:302:LEU:HD13	1.84	0.60
1:A:272:PHE:O	1:A:355:LEU:HB3	2.01	0.60
1:A:332:LEU:HD23	1:A:332:LEU:C	2.22	0.60
1:A:93:ASN:HD22	1:A:94:LYS:N	1.97	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:53:THR:HA	1:A:159:ARG:HH11	1.67	0.60
1:A:447:GLU:HG2	1:A:448:ILE:N	2.16	0.60
1:A:391:SER:HA	1:A:392:PRO:C	2.22	0.59
1:A:101:ASP:OD2	1:A:103:LYS:HE3	2.02	0.59
1:A:102:MET:HE3	1:A:160:ILE:HD11	1.83	0.59
1:A:474:TYR:CZ	1:A:480:PRO:HD3	2.37	0.59
1:A:93:ASN:HD21	1:A:95:ASP:CB	2.12	0.59
1:A:294:LYS:HG3	1:A:325:LYS:HD3	1.85	0.58
1:A:6:LEU:O	1:A:7:LEU:HD23	2.03	0.57
1:A:196:ASN:OD1	1:A:240:LYS:NZ	2.27	0.56
1:A:102:MET:HG2	1:A:160:ILE:HD13	1.86	0.56
1:A:449:ILE:HG21	1:A:603:LYS:HD2	1.87	0.56
1:A:107:TYR:OH	1:A:125:ASP:OD1	2.16	0.56
1:A:191:VAL:CG1	1:A:253:LYS:HG2	2.36	0.56
1:A:110:SER:HB2	1:A:121:ARG:NH1	2.22	0.55
1:A:102:MET:HG2	1:A:160:ILE:CD1	2.37	0.54
1:A:310:GLN:OE1	1:A:336:GLY:HA2	2.08	0.54
1:A:136:ARG:HB3	1:A:145:ASN:HD21	1.73	0.54
1:A:335:ASN:ND2	1:A:337:SER:H	2.07	0.53
1:A:559:GLY:O	1:A:575:THR:HG23	2.08	0.53
1:A:194:ASP:OD2	1:A:247:PRO:HA	2.08	0.53
1:A:398:CYS:HB2	1:A:557:SER:O	2.08	0.53
1:A:427:TYR:CE2	1:A:442:PHE:HB3	2.45	0.52
1:A:189:ASN:ND2	1:A:255:LYS:H	2.08	0.52
1:A:22:THR:HG22	1:A:27:TYR:HD2	1.75	0.51
1:A:335:ASN:HD22	1:A:335:ASN:C	2.14	0.51
1:A:17:ILE:HD12	1:A:61:LYS:CA	2.40	0.51
1:A:450:ILE:O	1:A:451:HIS:C	2.46	0.50
1:A:360:ASN:O	1:A:363:THR:OG1	2.28	0.50
1:A:9:ASP:O	1:A:73:ARG:O	2.29	0.50
1:A:159:ARG:NH2	1:A:160:ILE:H	2.09	0.50
1:A:437:LYS:HB2	1:A:439:ASP:OD2	2.11	0.50
1:A:528:GLN:HA	1:A:528:GLN:NE2	2.27	0.50
1:A:451:HIS:HD2	1:A:599:TRP:CE2	2.30	0.49
1:A:451:HIS:CE1	1:A:461:TYR:CD2	3.01	0.49
1:A:519:ILE:HB	1:A:520:PRO:CD	2.42	0.49
1:A:51:ASP:N	1:A:51:ASP:OD2	2.45	0.49
1:A:386:THR:CG2	1:A:427:TYR:HB2	2.42	0.49
1:A:265:CYS:O	1:A:267:HIS:O	2.31	0.49
1:A:65:THR:O	1:A:66:GLU:HB2	2.12	0.48
1:A:106:ASN:ND2	1:A:106:ASN:N	2.59	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:119:GLN:HG3	1:A:131:PHE:HE1	1.77	0.48
1:A:491:VAL:CG1	1:A:492:ILE:N	2.76	0.48
1:A:452:ASP:OD1	1:A:453:GLN:N	2.46	0.48
1:A:129:HIS:CD2	1:A:152:THR:H	2.32	0.48
1:A:157:PRO:HG3	1:A:160:ILE:HD11	1.96	0.48
1:A:296:HIS:H	1:A:296:HIS:CD2	2.31	0.48
1:A:37:ARG:NH1	1:A:37:ARG:CG	2.75	0.48
1:A:591:THR:CG2	1:A:596:TYR:CE2	2.97	0.48
1:A:353:LEU:C	1:A:355:LEU:N	2.68	0.47
1:A:192:PHE:HE2	1:A:250:ARG:HG2	1.80	0.47
1:A:29:GLN:NE2	1:A:82:PHE:CZ	2.82	0.47
3:A:702:NAG:H82	3:A:702:NAG:O3	2.15	0.47
1:A:40:LEU:HD12	1:A:40:LEU:H	1.77	0.47
1:A:38:CYS:O	1:A:83:LYS:NZ	2.47	0.47
1:A:243:GLU:O	1:A:244:SER:C	2.53	0.47
1:A:335:ASN:C	1:A:335:ASN:ND2	2.68	0.47
1:A:40:LEU:N	1:A:40:LEU:CD1	2.76	0.47
1:A:432:ASN:O	1:A:435:GLU:HG2	2.15	0.47
1:A:110:SER:HB2	1:A:121:ARG:HH12	1.78	0.47
1:A:352:THR:OG1	1:A:354:ARG:HB3	2.15	0.47
1:A:101:ASP:OD2	1:A:103:LYS:CE	2.63	0.46
1:A:521:LEU:HD12	1:A:521:LEU:N	2.31	0.46
1:A:575:THR:HG22	1:A:590:TYR:CE1	2.50	0.46
1:A:93:ASN:HD22	1:A:95:ASP:H	1.60	0.46
1:A:93:ASN:ND2	1:A:95:ASP:N	2.61	0.46
1:A:122:CYS:O	1:A:128:CYS:HB3	2.15	0.46
1:A:22:THR:OG1	1:A:58:CYS:HB2	2.16	0.45
1:A:175:CYS:O	1:A:177:LEU:N	2.47	0.45
1:A:192:PHE:CE1	1:A:258:SER:HB3	2.52	0.45
1:A:412:ALA:O	1:A:414:CYS:N	2.48	0.45
1:A:187:PHE:H	1:A:187:PHE:HD2	1.60	0.45
1:A:2:CYS:HB3	1:A:84:GLN:HB2	1.99	0.45
1:A:503:TYR:N	1:A:556:ASP:HB3	2.32	0.45
1:A:6:LEU:HD13	1:A:77:ILE:HG21	1.98	0.45
1:A:355:LEU:HA	1:A:358:MET:HE3	1.99	0.45
1:A:374:THR:CG2	1:A:374:THR:O	2.65	0.44
1:A:98:VAL:HA	1:A:167:VAL:HG12	1.99	0.44
1:A:384:GLN:HE21	1:A:499:THR:HB	1.82	0.44
1:A:129:HIS:HD2	1:A:152:THR:H	1.66	0.44
1:A:35:HIS:ND1	1:A:36:PRO:HD2	2.33	0.43
1:A:409:LEU:HD12	1:A:409:LEU:HA	1.57	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:56:PHE:CZ	1:A:100:LEU:HD11	2.53	0.43
1:A:194:ASP:OD1	1:A:242:SER:HB3	2.18	0.43
1:A:191:VAL:HG11	1:A:253:LYS:HG2	2.00	0.43
1:A:424:LEU:HD11	1:A:448:ILE:HD11	2.00	0.43
1:A:335:ASN:HD22	1:A:337:SER:H	1.64	0.43
1:A:16:ASP:N	1:A:16:ASP:OD1	2.51	0.43
1:A:357:LYS:HD2	1:A:357:LYS:HA	1.40	0.43
1:A:451:HIS:CE1	1:A:461:TYR:HD2	2.37	0.43
1:A:473:ASN:HD22	3:A:706:NAG:H62	1.84	0.43
1:A:437:LYS:N	1:A:440:THR:OG1	2.33	0.42
1:A:531:TYR:HE1	1:A:587:PRO:HG2	1.83	0.42
1:A:386:THR:O	1:A:386:THR:HG23	2.19	0.42
1:A:29:GLN:HG3	1:A:41:PHE:CE2	2.55	0.42
1:A:355:LEU:HD23	1:A:358:MET:HE1	2.02	0.41
1:A:35:HIS:CE1	1:A:36:PRO:HD2	2.55	0.41
1:A:430:ILE:HG22	1:A:501:TRP:CZ2	2.55	0.41
1:A:40:LEU:HD13	1:A:61:LYS:HB2	2.03	0.41
1:A:451:HIS:HB2	1:A:463:ILE:HG22	2.01	0.41
1:A:119:GLN:HA	1:A:131:PHE:HE1	1.83	0.41
1:A:197:ILE:O	1:A:198:ASP:HB2	2.21	0.41
1:A:412:ALA:C	1:A:414:CYS:N	2.74	0.41
1:A:412:ALA:C	1:A:414:CYS:H	2.24	0.41
1:A:447:GLU:HG2	1:A:448:ILE:H	1.85	0.41
1:A:558:GLY:HA2	1:A:574:ILE:CG2	2.51	0.41
1:A:420:SER:HA	1:A:421:PRO:HD3	1.95	0.41
1:A:428:SER:O	1:A:430:ILE:HD13	2.20	0.41
1:A:490:ASN:H	1:A:490:ASN:ND2	2.19	0.41
1:A:428:SER:O	1:A:430:ILE:CD1	2.68	0.41
1:A:389:THR:OG1	1:A:423:ILE:HD12	2.21	0.40
1:A:474:TYR:C	1:A:475:THR:CG2	2.90	0.40
1:A:497:TRP:CE2	1:A:518:LYS:HB2	2.56	0.40
1:A:200:VAL:HG21	1:A:211:ILE:HD13	2.03	0.40
1:A:159:ARG:HA	1:A:159:ARG:HH21	1.86	0.40
1:A:479:ARG:HG3	1:A:480:PRO:HD2	2.04	0.40
1:A:46:GLU:HA	1:A:55:TRP:CG	2.56	0.40
1:A:458:GLU:HA	1:A:577:TRP:CE2	2.57	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	581/606 (96%)	528 (91%)	46 (8%)	7 (1%)	16 47
2	K	7/9 (78%)	6 (86%)	1 (14%)	0	100 100
All	All	588/615 (96%)	534 (91%)	47 (8%)	7 (1%)	16 47

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	253	LYS
1	A	354	ARG
1	A	317	ALA
1	A	413	HIS
1	A	520	PRO
1	A	139	PRO
1	A	392	PRO

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	522/534 (98%)	440 (84%)	82 (16%)	3 9
2	K	9/9 (100%)	8 (89%)	1 (11%)	8 23
All	All	531/543 (98%)	448 (84%)	83 (16%)	3 9

All (83) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	THR
1	A	5	GLN
1	A	8	LYS
1	A	10	THR
1	A	11	CYS
1	A	16	ASP
1	A	37	ARG
1	A	40	LEU
1	A	42	THR
1	A	51	ASP
1	A	63	SER
1	A	64	VAL
1	A	71	VAL
1	A	73	ARG
1	A	84	GLN
1	A	86	SER
1	A	93	ASN
1	A	95	ASP
1	A	98	VAL
1	A	100	LEU
1	A	105	ILE
1	A	106	ASN
1	A	117	GLU
1	A	145	ASN
1	A	159	ARG
1	A	165	LYS
1	A	175	CYS
1	A	180	LEU
1	A	186	ILE
1	A	187	PHE
1	A	190	THR
1	A	191	VAL
1	A	194	ASP
1	A	199	SER
1	A	230	LYS
1	A	244	SER
1	A	253	LYS
1	A	254	SER
1	A	255	LYS
1	A	257	LEU
1	A	266	ARG
1	A	294	LYS
1	A	301	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	318	GLN
1	A	332	LEU
1	A	335	ASN
1	A	354	ARG
1	A	356	CYS
1	A	357	LYS
1	A	365	LYS
1	A	366	ILE
1	A	367	LYS
1	A	369	ARG
1	A	370	ILE
1	A	376	SER
1	A	377	VAL
1	A	387	LEU
1	A	395	ARG
1	A	432	ASN
1	A	439	ASP
1	A	446	GLN
1	A	475	THR
1	A	478	GLN
1	A	486	LYS
1	A	490	ASN
1	A	499	THR
1	A	509	LYS
1	A	513	THR
1	A	516	LYS
1	A	527	CYS
1	A	528	GLN
1	A	529	LYS
1	A	530	ARG
1	A	534	HIS
1	A	537	THR
1	A	554	LYS
1	A	556	ASP
1	A	562	SER
1	A	575	THR
1	A	576	SER
1	A	585	GLU
1	A	598	ASP
2	K	-2	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	29	GLN
1	A	93	ASN
1	A	106	ASN
1	A	129	HIS
1	A	145	ASN
1	A	189	ASN
1	A	296	HIS
1	A	305	ASN
1	A	335	ASN
1	A	384	GLN
1	A	446	GLN
1	A	451	HIS
1	A	490	ASN
1	A	515	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)

6 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	$\#\|Z\| > 2$
3	NAG	A	701	1	14,14,15	0.70	0	15,19,21	1.76	3 (20%)
3	NAG	A	702	1,3,4	14,14,15	0.66	0	15,19,21	1.69	3 (20%)
3	NAG	A	703	3	14,14,15	1.11	1 (7%)	15,19,21	2.88	4 (26%)
4	FUC	A	704	3	10,10,11	0.71	0	13,14,16	1.39	2 (15%)
3	NAG	A	705	1	14,14,15	0.80	1 (7%)	15,19,21	2.64	4 (26%)
3	NAG	A	706	1	14,14,15	0.66	0	15,19,21	2.68	3 (20%)

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	NAG	A	701	1	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	A	702	1,3,4	-	0/6/23/26	0/1/1/1
3	NAG	A	703	3	1/1/5/7	0/6/23/26	0/1/1/1
4	FUC	A	704	3	-	0/0/17/20	0/1/1/1
3	NAG	A	705	1	-	0/6/23/26	0/1/1/1
3	NAG	A	706	1	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	705	NAG	C1-C2	2.02	1.55	1.52
3	A	703	NAG	C7-N2	2.32	1.43	1.34

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	705	NAG	C4-C3-C2	-3.81	105.43	111.34
4	A	704	FUC	C1-C2-C3	-3.17	105.71	109.55
3	A	702	NAG	O7-C7-C8	-2.85	116.82	122.07
3	A	703	NAG	C4-C3-C2	-2.43	107.57	111.34
3	A	706	NAG	C4-C3-C2	-2.26	107.83	111.34
3	A	703	NAG	O7-C7-C8	-2.12	118.16	122.07
3	A	701	NAG	O7-C7-N2	2.13	126.18	121.84
3	A	702	NAG	O4-C4-C5	2.40	115.55	109.23
3	A	703	NAG	C3-C4-C5	2.44	114.58	110.23
3	A	705	NAG	O5-C5-C6	2.46	112.61	107.34
4	A	704	FUC	O2-C2-C1	2.65	114.54	109.23
3	A	705	NAG	O5-C5-C4	2.66	114.53	110.13
3	A	701	NAG	C1-O5-C5	2.75	116.19	112.14
3	A	702	NAG	C8-C7-N2	3.69	123.17	116.10
3	A	701	NAG	C2-N2-C7	4.78	129.32	123.11
3	A	706	NAG	C1-O5-C5	5.38	120.05	112.14
3	A	706	NAG	C2-N2-C7	7.88	133.36	123.11
3	A	705	NAG	C1-O5-C5	8.00	123.91	112.14
3	A	703	NAG	C2-N2-C7	9.89	135.97	123.11

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	701	NAG	C1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom
3	A	703	NAG	C1

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	702	NAG	1	0
4	A	704	FUC	1	0
3	A	706	NAG	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

6 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	NAG	A	701	1	14,14,15	0.70	0	15,19,21	1.76	3 (20%)
3	NAG	A	702	1,3,4	14,14,15	0.66	0	15,19,21	1.69	3 (20%)
3	NAG	A	703	3	14,14,15	1.11	1 (7%)	15,19,21	2.88	4 (26%)
4	FUC	A	704	3	10,10,11	0.71	0	13,14,16	1.39	2 (15%)
3	NAG	A	705	1	14,14,15	0.80	1 (7%)	15,19,21	2.64	4 (26%)
3	NAG	A	706	1	14,14,15	0.66	0	15,19,21	2.68	3 (20%)

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	NAG	A	701	1	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	A	702	1,3,4	-	0/6/23/26	0/1/1/1
3	NAG	A	703	3	1/1/5/7	0/6/23/26	0/1/1/1
4	FUC	A	704	3	-	0/0/17/20	0/1/1/1
3	NAG	A	705	1	-	0/6/23/26	0/1/1/1
3	NAG	A	706	1	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	705	NAG	C1-C2	2.02	1.55	1.52
3	A	703	NAG	C7-N2	2.32	1.43	1.34

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	705	NAG	C4-C3-C2	-3.81	105.43	111.34
4	A	704	FUC	C1-C2-C3	-3.17	105.71	109.55
3	A	702	NAG	O7-C7-C8	-2.85	116.82	122.07
3	A	703	NAG	C4-C3-C2	-2.43	107.57	111.34
3	A	706	NAG	C4-C3-C2	-2.26	107.83	111.34
3	A	703	NAG	O7-C7-C8	-2.12	118.16	122.07
3	A	701	NAG	O7-C7-N2	2.13	126.18	121.84
3	A	702	NAG	O4-C4-C5	2.40	115.55	109.23
3	A	703	NAG	C3-C4-C5	2.44	114.58	110.23
3	A	705	NAG	O5-C5-C6	2.46	112.61	107.34
4	A	704	FUC	O2-C2-C1	2.65	114.54	109.23
3	A	705	NAG	O5-C5-C4	2.66	114.53	110.13
3	A	701	NAG	C1-O5-C5	2.75	116.19	112.14
3	A	702	NAG	C8-C7-N2	3.69	123.17	116.10
3	A	701	NAG	C2-N2-C7	4.78	129.32	123.11
3	A	706	NAG	C1-O5-C5	5.38	120.05	112.14
3	A	706	NAG	C2-N2-C7	7.88	133.36	123.11
3	A	705	NAG	C1-O5-C5	8.00	123.91	112.14
3	A	703	NAG	C2-N2-C7	9.89	135.97	123.11

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	701	NAG	C1
3	A	703	NAG	C1

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	702	NAG	1	0
4	A	704	FUC	1	0
3	A	706	NAG	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	7

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	373:GLY	C	374:THR	N	1.19
1	A	78:SER	C	79:GLY	N	1.13
1	A	294:LYS	C	295:SER	N	1.13
1	A	366:ILE	C	367:LYS	N	1.07
1	A	253:LYS	C	254:SER	N	1.06
1	A	317:ALA	C	318:GLN	N	1.06
1	A	321:CYS	C	322:ASN	N	0.82

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	577/606 (95%)	-0.10	4 (0%) 89 84	38, 69, 106, 155	4 (0%)
2	K	9/9 (100%)	0.63	1 (11%) 7 3	85, 94, 125, 127	0
All	All	586/615 (95%)	-0.09	5 (0%) 85 79	38, 70, 106, 155	4 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	580	GLY	4.4
2	K	-2	HIS	3.5
1	A	375	ALA	2.4
1	A	378	ARG	2.4
1	A	3	VAL	2.2

6.2 Non-standard residues in protein, DNA, RNA chains

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^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
4	FUC	A	704	10/11	0.93	0.20	1.00	72,82,89,103	0
3	NAG	A	702	14/15	0.95	0.20	0.29	86,93,103,108	0
3	NAG	A	701	14/15	0.75	0.23	-	108,120,139,141	0
3	NAG	A	706	14/15	0.78	0.26	-	99,128,133,136	0
3	NAG	A	703	14/15	0.83	0.40	-	102,126,135,138	0
3	NAG	A	705	14/15	0.80	0.32	-	88,113,125,127	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^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
4	FUC	A	704	10/11	0.93	0.20	1.00	72,82,89,103	0
3	NAG	A	702	14/15	0.95	0.20	0.29	86,93,103,108	0
3	NAG	A	701	14/15	0.75	0.23	-	108,120,139,141	0
3	NAG	A	706	14/15	0.78	0.26	-	99,128,133,136	0
3	NAG	A	703	14/15	0.83	0.40	-	102,126,135,138	0
3	NAG	A	705	14/15	0.80	0.32	-	88,113,125,127	0

6.5 Other polymers (i)

There are no such residues in this entry.