



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 12, 2016 – 08:48 AM EDT

PDB ID : 5EOD  
Title : Human Plasma Coagulation FXI with peptide LP2  
Authors : Wong, S.S.; Ostergaard, S.; Hall, G.; Li, C.; Williams, P.M.; Stennicke, H.; Emsley, J.  
Deposited on : 2015-11-10  
Resolution : 3.10 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 : 1.7.1 (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

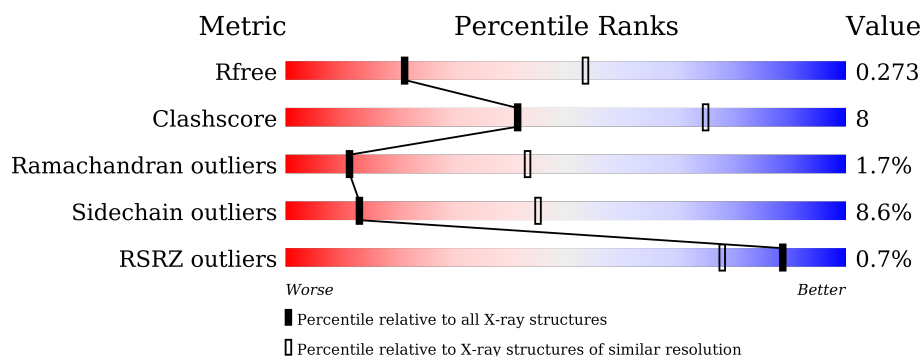
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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	606	<div> <div style="width: 100%; height: 10px; background-color: red;"></div> <div style="display: flex; justify-content: space-between; align-items: center;"> <span>%</span> <div style="width: 70%; height: 10px; background-color: green;"></div> <div style="width: 22%; height: 10px; background-color: yellow;"></div> <div style="width: 8%; height: 10px; background-color: orange;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <div style="display: flex; justify-content: space-between; align-items: center;"> <span>70%</span> <span>22%</span> <span>...</span> </div> </div>
2	B	6	<div> <div style="width: 83%; height: 10px; background-color: green;"></div> <div style="width: 17%; height: 10px; background-color: yellow;"></div> </div> <div style="display: flex; justify-content: space-between; align-items: center;"> <span>83%</span> <span>17%</span> </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
3	NAG	A	901	X	-	-	-
3	NAG	A	902	X	-	-	-

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4698 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.

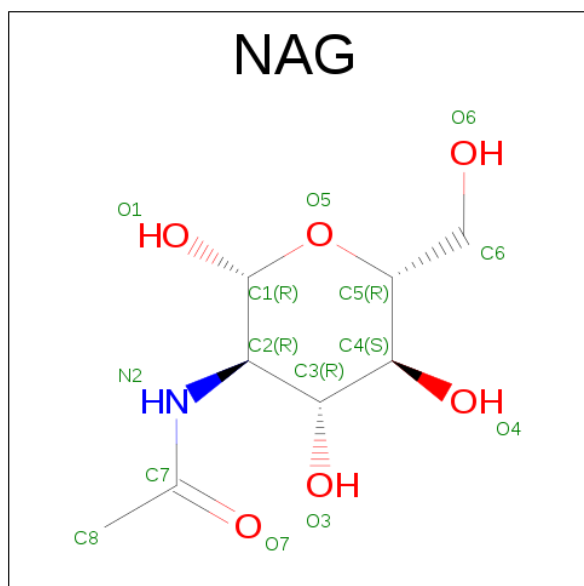
- Molecule 1 is a protein called Coagulation factor XI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	586	Total	C	N	O	S	77	0	0
			4603	2903	797	864	39			

- Molecule 2 is a protein called LP2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	6	Total	C	N	O	0	0	0
			53	37	6	10			

- 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...*

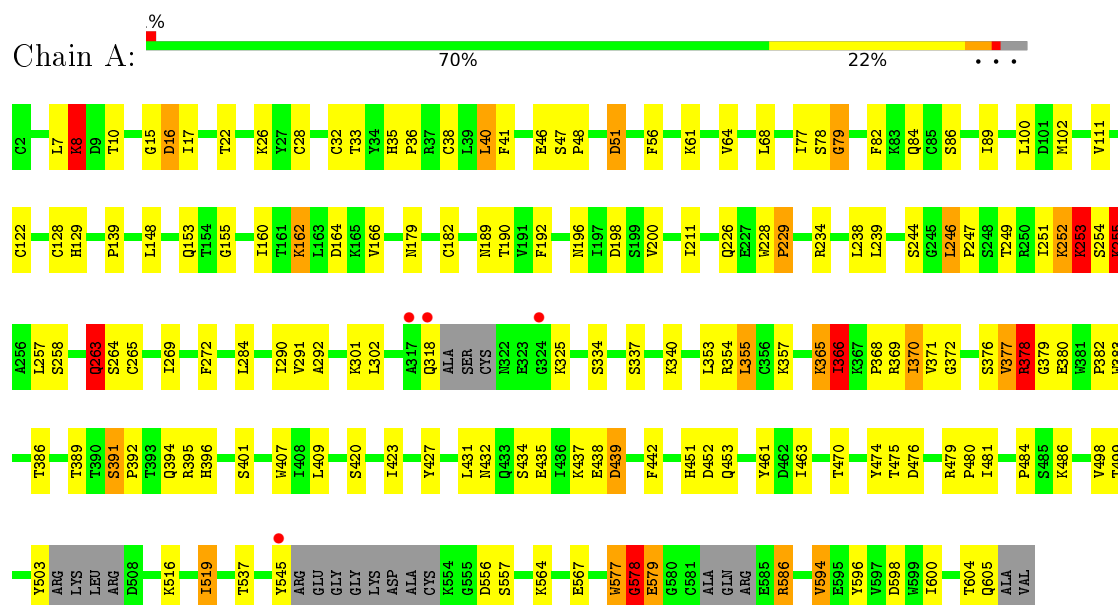
*Continued from previous page...*

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

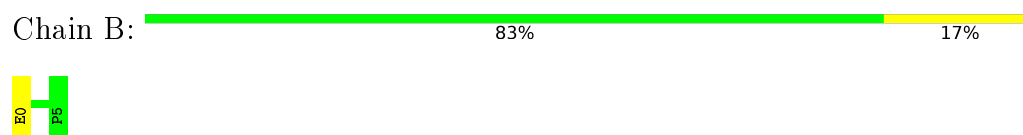
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: LP2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.06 Å 81.06 Å 253.08 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.40 – 3.10 47.41 – 3.10	Depositor EDS
% Data completeness (in resolution range)	97.8 (47.40-3.10) 97.8 (47.41-3.10)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.34 (at 3.12 Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, $R_{free}$	0.189 , 0.274 0.189 , 0.273	Depositor DCC
$R_{free}$ test set	804 reflections (5.36%)	DCC
Wilson B-factor (Å <sup>2</sup> )	82.4	Xtriage
Anisotropy	0.007	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 56.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4698	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.19% 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, 0.375 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

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	1.92	22/4709 (0.5%)	1.54	43/6373 (0.7%)
2	B	0.70	0/56	0.70	0/76
All	All	1.91	22/4765 (0.5%)	1.53	43/6449 (0.7%)

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	10

All (22) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	253	LYS	CA-C	63.02	3.16	1.52
1	A	579	GLU	C-N	-56.28	0.31	1.33
1	A	253	LYS	C-N	-41.37	0.38	1.34
1	A	253	LYS	CG-CD	-32.12	0.43	1.52
1	A	8	LYS	C-N	-30.79	0.63	1.34
1	A	252	LYS	CA-CB	-29.57	0.88	1.53
1	A	8	LYS	CA-CB	-27.37	0.93	1.53
1	A	378	ARG	C-N	-26.14	0.85	1.33
1	A	378	ARG	CA-CB	-25.12	0.98	1.53
1	A	366	ILE	C-N	-23.66	0.79	1.34
1	A	263	GLN	CG-CD	20.70	1.98	1.51
1	A	368	PRO	C-N	17.31	1.73	1.34
1	A	371	VAL	C-N	10.86	1.52	1.33
1	A	257	LEU	CB-CG	-9.21	1.25	1.52
1	A	318	GLN	CA-C	7.81	1.73	1.52
1	A	378	ARG	CA-C	7.40	1.72	1.52

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	252	LYS	C-N	7.14	1.50	1.34
1	A	255	LYS	CE-NZ	6.92	1.66	1.49
1	A	252	LYS	CA-C	6.69	1.70	1.52
1	A	578	GLY	C-N	-6.42	1.19	1.34
1	A	8	LYS	CA-C	5.78	1.68	1.52
1	A	318	GLN	C-O	5.57	1.33	1.23

All (43) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	8	LYS	O-C-N	-33.29	69.44	122.70
1	A	579	GLU	CA-C-N	-32.18	51.85	116.20
1	A	578	GLY	O-C-N	-28.21	77.57	122.70
1	A	263	GLN	CG-CD-OE1	-27.97	65.66	121.60
1	A	579	GLU	C-N-CA	-23.88	72.16	122.30
1	A	253	LYS	N-CA-C	-23.55	47.42	111.00
1	A	263	GLN	CB-CG-CD	23.18	171.87	111.60
1	A	366	ILE	O-C-N	-22.95	85.98	122.70
1	A	368	PRO	C-N-CA	18.16	167.09	121.70
1	A	8	LYS	CB-CA-C	17.36	145.11	110.40
1	A	252	LYS	CB-CA-C	16.89	144.18	110.40
1	A	253	LYS	CB-CG-CD	16.79	155.25	111.60
1	A	253	LYS	CB-CA-C	-16.71	76.99	110.40
1	A	252	LYS	C-N-CA	-16.44	80.59	121.70
1	A	378	ARG	O-C-N	-14.64	98.31	123.20
1	A	253	LYS	CA-C-N	-14.49	85.31	117.20
1	A	366	ILE	CA-C-N	14.45	149.00	117.20
1	A	263	GLN	CG-CD-NE2	14.20	150.79	116.70
1	A	8	LYS	CA-C-O	-13.52	91.70	120.10
1	A	8	LYS	C-N-CA	12.80	153.70	121.70
1	A	255	LYS	CD-CE-NZ	-11.78	84.61	111.70
1	A	378	ARG	C-N-CA	11.50	146.45	122.30
1	A	370	ILE	N-CA-C	11.45	141.90	111.00
1	A	8	LYS	CA-C-N	11.44	142.36	117.20
1	A	252	LYS	N-CA-C	-11.30	80.48	111.00
1	A	579	GLU	O-C-N	-11.06	104.40	123.20
1	A	578	GLY	CA-C-N	10.81	140.98	117.20
1	A	253	LYS	C-N-CA	-10.64	95.10	121.70
1	A	253	LYS	O-C-N	10.24	139.08	122.70
1	A	378	ARG	CB-CA-C	9.42	129.23	110.40
1	A	371	VAL	O-C-N	8.88	138.30	123.20
1	A	378	ARG	CA-C-N	7.18	130.56	116.20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	368	PRO	O-C-N	-7.15	111.26	122.70
1	A	578	GLY	C-N-CA	6.67	138.39	121.70
1	A	263	GLN	CA-CB-CG	-6.63	98.81	113.40
1	A	371	VAL	CA-C-N	-6.46	103.28	116.20
1	A	257	LEU	CB-CG-CD2	-6.05	100.71	111.00
1	A	252	LYS	O-C-N	-5.75	113.50	122.70
1	A	355	LEU	CA-CB-CG	5.64	128.28	115.30
1	A	40	LEU	CA-CB-CG	5.61	128.21	115.30
1	A	8	LYS	N-CA-C	-5.59	95.91	111.00
1	A	318	GLN	CA-C-O	-5.54	108.47	120.10
1	A	366	ILE	C-N-CA	5.02	134.24	121.70

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	8	LYS	CA

All (10) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	252	LYS	Mainchain
1	A	253	LYS	Mainchain
1	A	263	GLN	Sidechain
1	A	366	ILE	Mainchain
1	A	378	ARG	Mainchain
1	A	578	GLY	Mainchain
1	A	579	GLU	Mainchain,Peptide
1	A	8	LYS	Mainchain,Peptide

## 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	4603	0	4466	77	0
2	B	53	0	44	0	0
3	A	42	0	39	0	0
All	All	4698	0	4549	77	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 8.

All (77) 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:189:ASN:HD22	1:A:254:SER:HA	1.34	0.91
1:A:378:ARG:CA	1:A:379:GLY:N	2.36	0.87
1:A:391:SER:HB3	1:A:423:ILE:HD11	1.64	0.80
1:A:40:LEU:HD11	1:A:68:LEU:HD12	1.77	0.66
1:A:189:ASN:ND2	1:A:255:LYS:H	1.95	0.65
1:A:474:TYR:CE1	1:A:480:PRO:HD3	2.33	0.63
1:A:40:LEU:HD12	1:A:61:LYS:HB2	1.78	0.63
1:A:179:ASN:OD1	1:A:486:LYS:HE2	1.99	0.62
1:A:129:HIS:ND1	1:A:155:GLY:HA2	2.14	0.62
1:A:200:VAL:HG21	1:A:211:ILE:HD13	1.83	0.61
1:A:564:LYS:HE2	1:A:567:GLU:HA	1.83	0.60
1:A:407:TRP:CE2	1:A:604:THR:HG22	2.35	0.60
1:A:41:PHE:CE2	1:A:79:GLY:HA2	2.37	0.60
1:A:389:THR:O	1:A:394:GLN:HA	2.02	0.59
1:A:253:LYS:H	1:A:254:SER:N	2.00	0.59
1:A:189:ASN:ND2	1:A:254:SER:HA	2.12	0.59
1:A:484:PRO:HB2	1:A:594:VAL:HG13	1.85	0.59
1:A:386:THR:OG1	1:A:396:HIS:HD2	1.87	0.57
1:A:292:ALA:HB1	1:A:325:LYS:HE2	1.85	0.57
1:A:604:THR:O	1:A:605:GLN:HB2	2.03	0.56
1:A:432:ASN:HB3	1:A:434:SER:HB3	1.87	0.56
1:A:376:SER:HA	1:A:380:GLU:OE2	2.06	0.56
1:A:451:HIS:HE1	1:A:461:TYR:HD1	1.53	0.55
1:A:427:TYR:CE2	1:A:442:PHE:HB3	2.44	0.53
1:A:182:CYS:SG	1:A:264:SER:HB3	2.49	0.53
1:A:46:GLU:OE2	1:A:162:LYS:HD3	2.10	0.52
1:A:41:PHE:CZ	1:A:79:GLY:HA2	2.46	0.51
1:A:246:LEU:HD23	1:A:247:PRO:HD2	1.94	0.50
1:A:451:HIS:CE1	1:A:461:TYR:HD1	2.30	0.49
1:A:226:GLN:O	1:A:234:ARG:HG2	2.12	0.49
1:A:196:ASN:OD1	1:A:238:LEU:HD13	2.12	0.49
1:A:198:ASP:HB3	1:A:239:LEU:HD12	1.93	0.49
1:A:179:ASN:OD1	1:A:486:LYS:CE	2.61	0.49
1:A:498:VAL:HG22	1:A:519:ILE:HD13	1.96	0.48
1:A:253:LYS:N	1:A:254:SER:N	2.59	0.47
1:A:420:SER:O	1:A:423:ILE:HG22	2.15	0.47
1:A:452:ASP:O	1:A:453:GLN:HB2	2.14	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:265:CYS:O	1:A:269:ILE:HG12	2.13	0.47
1:A:51:ASP:N	1:A:51:ASP:OD1	2.42	0.47
1:A:22:THR:HG21	1:A:28:CYS:HA	1.97	0.47
1:A:192:PHE:CE2	1:A:258:SER:HB3	2.50	0.47
1:A:41:PHE:O	1:A:79:GLY:HA3	2.15	0.46
1:A:437:LYS:HB2	1:A:439:ASP:OD1	2.15	0.46
1:A:33:THR:HG21	1:A:82:PHE:HB2	1.96	0.46
1:A:272:PHE:HB3	1:A:355:LEU:HB3	1.98	0.46
1:A:392:PRO:HD2	1:A:395:ARG:HH11	1.80	0.45
1:A:228:TRP:HA	1:A:229:PRO:HD2	1.68	0.45
1:A:377:VAL:HA	1:A:431:LEU:HD21	1.99	0.45
1:A:409:LEU:HD11	1:A:463:ILE:HD11	1.98	0.45
1:A:438:GLU:CD	1:A:438:GLU:H	2.20	0.45
1:A:382:PRO:HB2	1:A:479:ARG:H	1.81	0.44
1:A:7:LEU:HG	1:A:78:SER:O	2.18	0.43
1:A:15:GLY:O	1:A:16:ASP:C	2.57	0.43
1:A:47:SER:HA	1:A:48:PRO:HD2	1.74	0.43
1:A:189:ASN:HD22	1:A:254:SER:CA	2.17	0.42
1:A:386:THR:OG1	1:A:396:HIS:CD2	2.71	0.42
1:A:383:TRP:CG	1:A:481:ILE:HB	2.53	0.42
1:A:35:HIS:CE1	1:A:36:PRO:HD2	2.54	0.42
1:A:577:TRP:HB2	1:A:578:GLY:H	1.63	0.42
1:A:596:TYR:O	1:A:600:ILE:HG13	2.20	0.42
1:A:190:THR:HA	1:A:251:ILE:O	2.20	0.41
1:A:284:LEU:HD12	1:A:340:LYS:HB3	2.02	0.41
1:A:377:VAL:HA	1:A:431:LEU:CD2	2.50	0.41
1:A:432:ASN:O	1:A:435:GLU:HB3	2.19	0.41
1:A:189:ASN:HD22	1:A:255:LYS:H	1.65	0.41
1:A:391:SER:HB3	1:A:423:ILE:CD1	2.44	0.41
1:A:56:PHE:CZ	1:A:100:LEU:HD11	2.55	0.41
1:A:291:VAL:HG21	1:A:302:LEU:HD13	2.03	0.41
1:A:35:HIS:CG	1:A:36:PRO:HD2	2.56	0.41
1:A:102:MET:HG2	1:A:160:ILE:HG12	2.03	0.41
1:A:122:CYS:O	1:A:128:CYS:HB3	2.21	0.41
1:A:353:LEU:C	1:A:355:LEU:N	2.73	0.41
1:A:365:LYS:HD2	1:A:366:ILE:H	1.86	0.41
1:A:377:VAL:HG22	1:A:378:ARG:H	1.86	0.40
1:A:357:LYS:HA	1:A:357:LYS:HD3	1.78	0.40
1:A:586:ARG:HE	1:A:586:ARG:N	2.19	0.40
1:A:32:CYS:O	1:A:38:CYS:HB3	2.21	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	574/606 (95%)	516 (90%)	48 (8%)	10 (2%)	11	43
2	B	4/6 (67%)	4 (100%)	0	0	100	100
All	All	578/612 (94%)	520 (90%)	48 (8%)	10 (2%)	11	43

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	8	LYS
1	A	229	PRO
1	A	369	ARG
1	A	370	ILE
1	A	372	GLY
1	A	378	ARG
1	A	139	PRO
1	A	354	ARG
1	A	16	ASP
1	A	79	GLY

### 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	519/534 (97%)	475 (92%)	44 (8%)	13	45
2	B	6/6 (100%)	5 (83%)	1 (17%)	3	11
All	All	525/540 (97%)	480 (91%)	45 (9%)	13	45

All (45) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	THR
1	A	17	ILE
1	A	26	LYS
1	A	51	ASP
1	A	64	VAL
1	A	77	ILE
1	A	84	GLN
1	A	86	SER
1	A	89	ILE
1	A	111	VAL
1	A	148	LEU
1	A	153	GLN
1	A	162	LYS
1	A	164	ASP
1	A	166	VAL
1	A	244	SER
1	A	246	LEU
1	A	249	THR
1	A	255	LYS
1	A	263	GLN
1	A	290	ILE
1	A	301	LYS
1	A	334	SER
1	A	337	SER
1	A	365	LYS
1	A	377	VAL
1	A	391	SER
1	A	401	SER
1	A	439	ASP
1	A	470	THR
1	A	475	THR
1	A	476	ASP
1	A	499	THR
1	A	503	TYR
1	A	516	LYS
1	A	519	ILE
1	A	537	THR
1	A	545	TYR
1	A	556	ASP
1	A	557	SER
1	A	577	TRP
1	A	586	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	594	VAL
1	A	598	ASP
2	B	0	GLU

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	151	HIS
1	A	189	ASN
1	A	226	GLN
1	A	235	ASN
1	A	343	HIS
1	A	396	HIS
1	A	406	GLN
1	A	451	HIS
1	A	473	ASN

### 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](#)

3 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	901	1	14,14,15	1.00	1 (7%)	15,19,21	3.55	5 (33%)
3	NAG	A	902	1	14,14,15	1.45	2 (14%)	15,19,21	4.38	8 (53%)
3	NAG	A	903	1	14,14,15	1.06	1 (7%)	15,19,21	2.21	4 (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	NAG	A	901	1	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	A	902	1	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	A	903	1	-	0/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	902	NAG	C3-C2	2.14	1.57	1.52
3	A	901	NAG	C1-C2	2.44	1.55	1.52
3	A	903	NAG	C1-C2	3.22	1.57	1.52
3	A	902	NAG	C1-C2	4.30	1.58	1.52

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	902	NAG	O5-C5-C4	-3.35	104.59	110.13
3	A	902	NAG	O7-C7-N2	-2.65	116.45	121.84
3	A	902	NAG	O7-C7-C8	-2.31	117.81	122.07
3	A	902	NAG	O3-C3-C2	2.02	113.69	109.37
3	A	903	NAG	O5-C5-C4	2.08	113.58	110.13
3	A	903	NAG	C8-C7-N2	2.47	120.84	116.10
3	A	901	NAG	C3-C4-C5	2.73	115.09	110.23
3	A	902	NAG	C1-O5-C5	3.38	117.10	112.14
3	A	901	NAG	C4-C3-C2	4.70	118.64	111.34
3	A	902	NAG	C4-C3-C2	4.80	118.79	111.34
3	A	901	NAG	O5-C5-C4	4.92	118.29	110.13
3	A	903	NAG	C2-N2-C7	5.01	129.62	123.11
3	A	902	NAG	C8-C7-N2	5.03	125.73	116.10
3	A	903	NAG	C1-O5-C5	5.14	119.70	112.14
3	A	901	NAG	C2-N2-C7	5.83	130.69	123.11
3	A	901	NAG	C1-O5-C5	9.29	125.80	112.14

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	902	NAG	C2-N2-C7	13.86	141.13	123.11

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	902	NAG	C1
3	A	901	NAG	C1

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	8

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	367:LYS	C	368:PRO	N	2.39
1	A	368:PRO	C	369:ARG	N	1.73
1	A	578:GLY	C	579:GLU	N	1.19
1	A	378:ARG	C	379:GLY	N	0.86
1	A	366:ILE	C	367:LYS	N	0.79
1	A	8:LYS	C	9:ASP	N	0.63
1	A	253:LYS	C	254:SER	N	0.38
1	A	579:GLU	C	580:GLY	N	0.31

## 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	581/606 (95%)	-0.13	4 (0%) 89 78	42, 74, 110, 158	7 (1%)
2	B	6/6 (100%)	-0.19	0 100 100	91, 103, 107, 107	0
All	All	587/612 (95%)	-0.13	4 (0%) 89 78	42, 74, 110, 158	7 (1%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	318	GLN	4.0
1	A	324	GLY	3.2
1	A	317	ALA	3.0
1	A	545	TYR	2.0

### 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. 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( $\text{\AA}^2$ )	Q<0.9
3	NAG	A	902	14/15	0.84	0.15	-	66,82,96,99	0
3	NAG	A	901	14/15	0.84	0.23	-	96,121,130,146	0
3	NAG	A	903	14/15	0.83	0.30	-	100,119,134,137	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.