



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 01:14 AM GMT

PDB ID : 2C4F  
Title : CRYSTAL STRUCTURE OF FACTOR VII:STF COMPLEXED WITH PD0297121  
Authors : Kohrt, J.T.; Zhang, E.  
Deposited on : 2005-10-18  
Resolution : 1.72 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

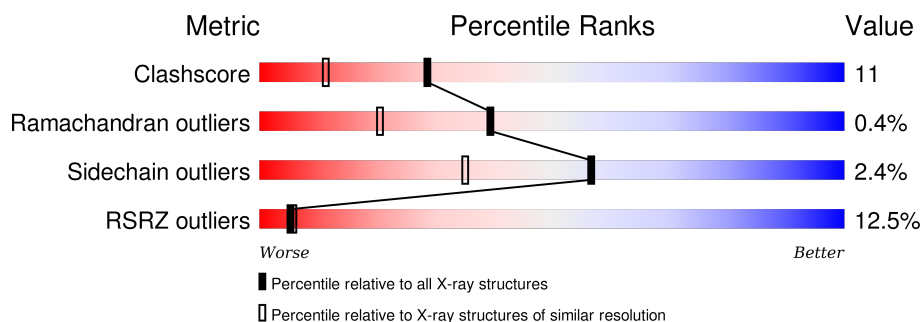
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.72 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	102246	4425 (1.74-1.70)
Ramachandran outliers	100387	4360 (1.74-1.70)
Sidechain outliers	100360	4360 (1.74-1.70)
RSRZ outliers	91569	4010 (1.74-1.70)

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	H	254	<div> <div>10%</div> <div>76%</div> <div>22%</div> <div>.</div> </div>
2	L	142	<div> <div>18%</div> <div>77%</div> <div>20%</div> <div>..</div> </div>
3	T	75	<div> <div>%</div> <div>87%</div> <div>13%</div> </div>
4	U	116	<div> <div>17%</div> <div>75%</div> <div>25%</div> </div>

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 5058 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 VII PRECURSOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	254	Total	C	N	O	S	0	0	0
			1974	1253	351	357	13			

- Molecule 2 is a protein called COAGULATION FACTOR VII PRECURSOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	139	Total	C	N	O	S	0	0	0
			1117	673	185	244	15			

- Molecule 3 is a protein called TISSUE FACTOR PRECURSOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	T	75	Total	C	N	O	S	0	0	0
			615	397	95	121	2			

- Molecule 4 is a protein called TISSUE FACTOR PRECURSOR.

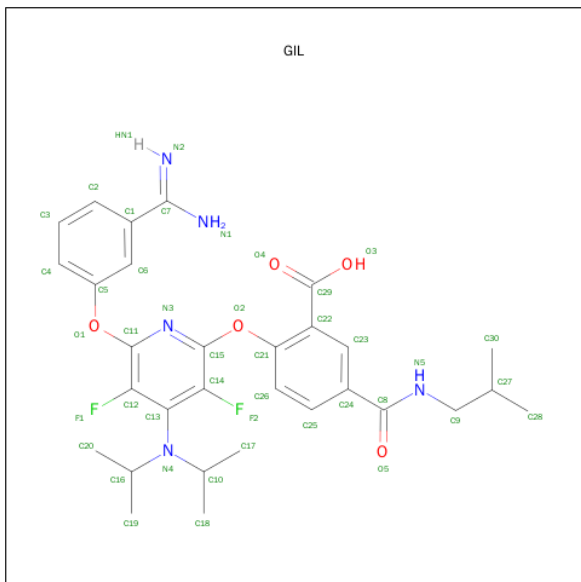
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	U	116	Total	C	N	O	S	0	0	0
			940	593	155	189	3			

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	H	1	Total	Ca	0	0
			1	1		
5	A	3	Total	Ca	0	0
			3	3		
5	L	1	Total	Ca	0	0
			1	1		

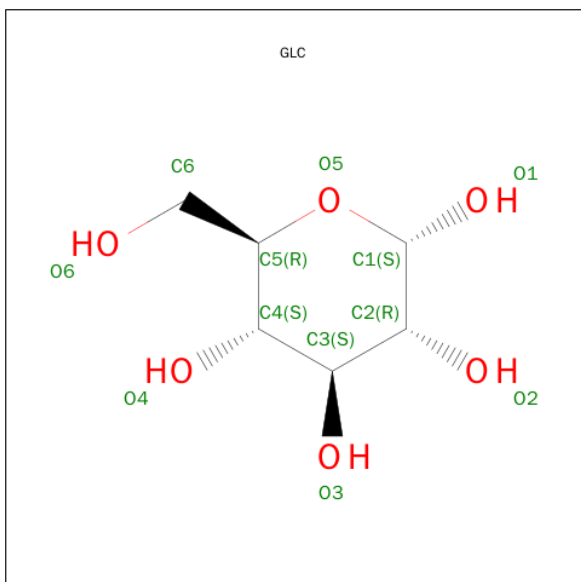
- Molecule 6 is 2-{[6-{3-[AMINO(IMINO)METHYL]PHENOXY}-4-(DIISOPROPYLAMINO

)-3,5-DIFLUOROPYRIDIN-2-YL]OXY}-5-[(ISOBUTYLAMINO)CARBONYL]BENZOIC ACID (three-letter code: GIL) (formula:  $C_{30}H_{35}F_2N_5O_5$ ).



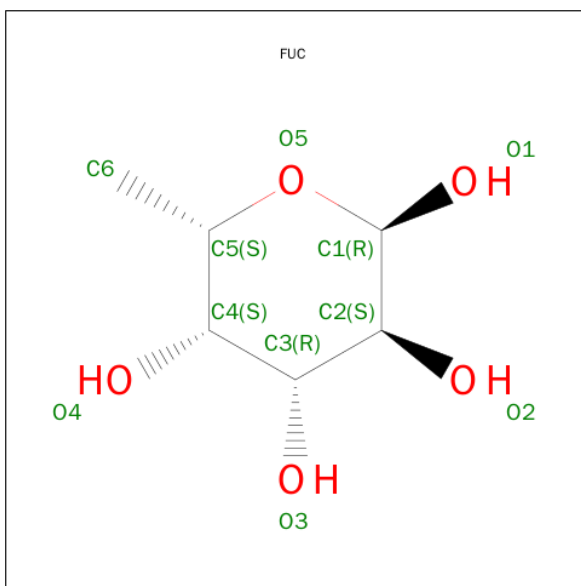
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	H	1	Total	C	F	N	O	0	0
			42	30	2	5	5		

- Molecule 7 is SUGAR (ALPHA-D-GLUCOSE) (three-letter code: GLC) (formula:  $C_6H_{12}O_6$ ).



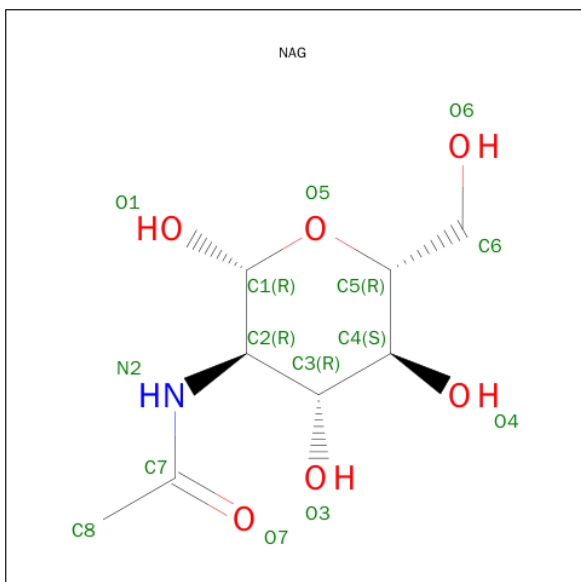
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	L	1	Total	C	O	0	0
			12	6	6		

- Molecule 8 is SUGAR (ALPHA-L-FUCOSE) (three-letter code: FUC) (formula:  $C_6H_{12}O_5$ ).



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

- Molecule 9 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	U	1	Total	C	N	O	0	0
			14	8	1	5		

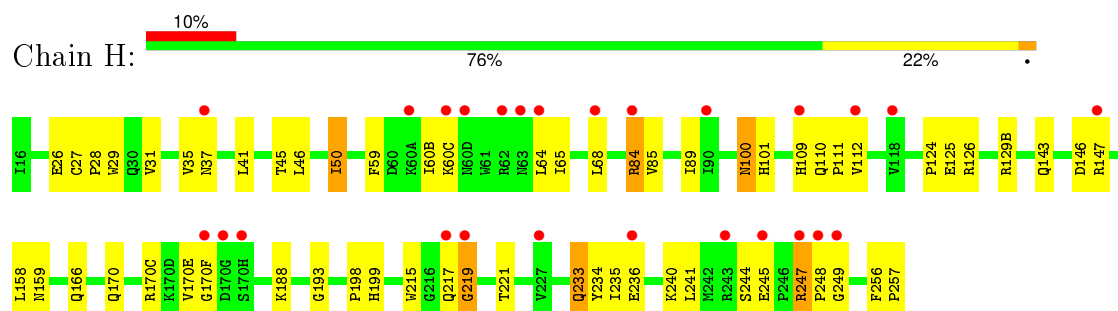
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	H	149	Total 149	O 149	0	0
10	L	59	Total 59	O 59	0	0
10	T	64	Total 64	O 64	0	0
10	U	56	Total 56	O 56	0	0

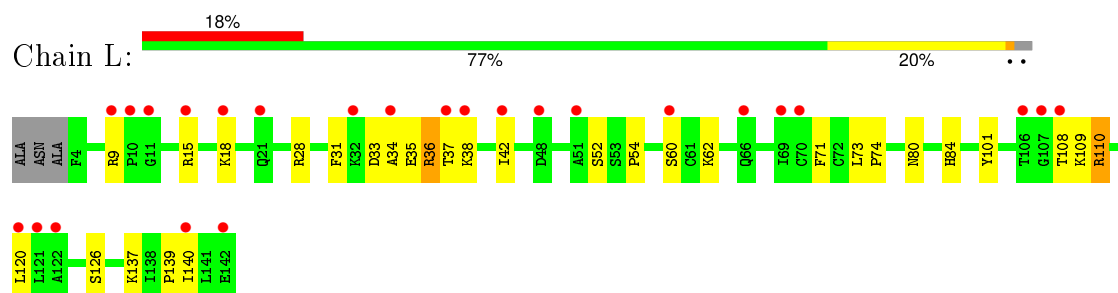
### 3 Residue-property plots [i](#)

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.

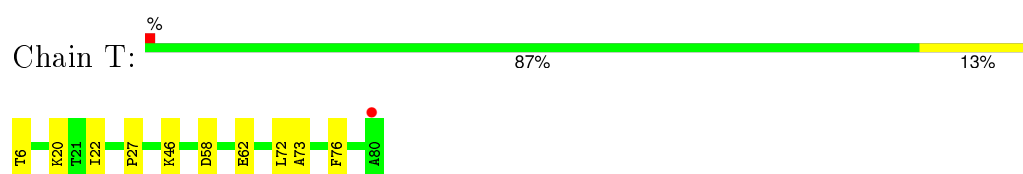
#### • Molecule 1: COAGULATION FACTOR VII PRECURSOR



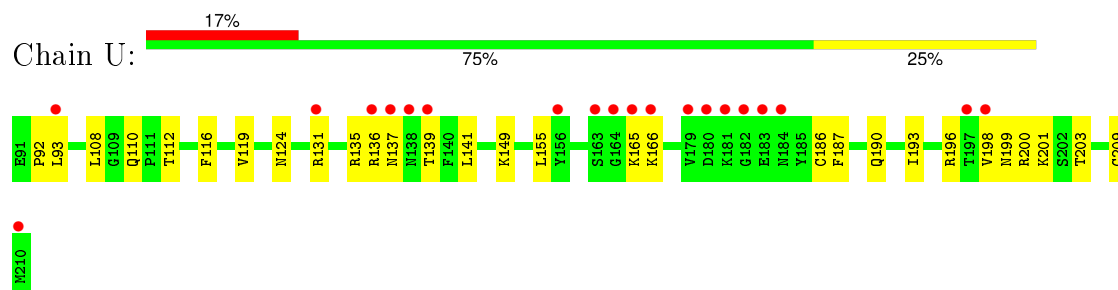
#### • Molecule 2: COAGULATION FACTOR VII PRECURSOR



#### • Molecule 3: TISSUE FACTOR PRECURSOR



#### • Molecule 4: TISSUE FACTOR PRECURSOR



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.99 Å 81.30 Å 125.86 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.72 19.54 – 1.73	Depositor EDS
% Data completeness (in resolution range)	88.8 (50.00-1.72) 92.0 (19.54-1.73)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.98 (at 1.72 Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.229 , (Not available) 0.228 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	19.9	Xtriage
Anisotropy	0.370	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 51.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 74805 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5058	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.63% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, CA, GLC, FUC, GIL, CGU

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	H	0.27	0/2024	0.65	1/2755 (0.0%)
2	L	0.26	0/1011	0.53	0/1349
3	T	0.29	0/631	0.68	1/860 (0.1%)
4	U	0.27	0/958	0.63	0/1299
All	All	0.27	0/4624	0.63	2/6263 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	199	HIS	N-CA-C	-6.55	93.30	111.00
3	T	20	LYS	N-CA-C	-5.07	97.31	111.00

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	H	1974	0	1950	48	0
2	L	1117	0	970	34	0
3	T	615	0	598	9	0
4	U	940	0	911	25	0
5	A	3	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	H	1	0	0	0	0
5	L	1	0	0	0	0
6	H	42	0	33	3	0
7	L	12	0	12	2	0
8	L	11	0	12	4	0
9	U	14	0	13	1	0
10	H	149	0	0	3	0
10	L	59	0	0	3	0
10	T	64	0	0	2	0
10	U	56	0	0	4	0
All	All	5058	0	4499	106	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 11.

All (106) 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:H:68:LEU:HD12	1:H:112:VAL:HG11	1.37	1.04
2:L:60:SER:HB3	8:L:1145:FUC:O1	1.59	1.01
2:L:74:PRO:HA	4:U:135:ARG:HH12	1.38	0.87
2:L:31:PHE:HB2	2:L:37:THR:HG22	1.61	0.81
2:L:60:SER:HB3	8:L:1145:FUC:C1	2.10	0.81
2:L:60:SER:OG	2:L:71:PHE:HB2	1.83	0.78
2:L:52:SER:OG	7:L:1144:GLC:H1	1.83	0.78
1:H:129(B):ARG:HD3	10:L:2040:HOH:O	1.82	0.77
1:H:126:ARG:HA	1:H:235:ILE:HD12	1.66	0.77
1:H:50:ILE:HG22	1:H:111:PRO:HB3	1.68	0.76
4:U:136:ARG:O	4:U:139:THR:HG22	1.87	0.74
3:T:6:THR:HG22	4:U:93:LEU:HD22	1.71	0.73
1:H:143:GLN:NE2	1:H:146:ASP:O	2.22	0.72
2:L:71:PHE:CZ	4:U:131:ARG:HG3	2.27	0.70
2:L:52:SER:OG	7:L:1144:GLC:C1	2.40	0.69
4:U:112:THR:HG23	10:U:2022:HOH:O	1.92	0.69
2:L:38:LYS:O	2:L:42:ILE:HG13	1.93	0.68
2:L:62:LYS:HE3	10:L:2016:HOH:O	1.92	0.68
2:L:74:PRO:HA	4:U:135:ARG:NH1	2.10	0.67
2:L:33:ASP:O	2:L:37:THR:HG23	1.96	0.66
1:H:35:VAL:HG13	1:H:41:LEU:HD22	1.75	0.66
1:H:126:ARG:CA	1:H:235:ILE:HD12	2.26	0.66
1:H:217:GLN:O	1:H:219:GLY:O	2.12	0.66

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:219:GLY:O	1:H:221:THR:HG23	1.95	0.65
4:U:149:LYS:HD2	10:U:2040:HOH:O	1.96	0.64
1:H:125:GLU:HG3	2:L:101:TYR:CZ	2.34	0.63
1:H:124:PRO:O	1:H:235:ILE:HD13	1.98	0.62
1:H:35:VAL:CG1	1:H:41:LEU:HD22	2.29	0.62
1:H:46:LEU:HD13	1:H:68:LEU:HD11	1.82	0.61
3:T:46:LYS:HD2	10:T:2032:HOH:O	2.02	0.60
1:H:147:ARG:O	1:H:147:ARG:HG2	2.01	0.60
2:L:18:LYS:HE3	10:U:2054:HOH:O	2.03	0.58
1:H:64:LEU:HB3	1:H:85:VAL:HB	1.85	0.58
2:L:31:PHE:HB2	2:L:37:THR:CG2	2.31	0.58
2:L:34:ALA:O	2:L:38:LYS:HG2	2.04	0.58
1:H:26:GLU:HG3	2:L:140:ILE:HD11	1.87	0.57
1:H:166:GLN:O	1:H:170:GLN:HG3	2.05	0.57
1:H:45:THR:OG1	1:H:198:PRO:HB3	2.04	0.57
2:L:28:ARG:HA	2:L:37:THR:HG21	1.88	0.56
4:U:119:VAL:HG23	4:U:119:VAL:O	2.05	0.56
3:T:46:LYS:HD3	3:T:62:GLU:CD	2.28	0.55
1:H:170(E):VAL:HG21	1:H:217:GLN:OE1	2.07	0.54
1:H:240:LYS:HD2	10:H:2144:HOH:O	2.08	0.53
4:U:198:VAL:HG22	10:U:2049:HOH:O	2.09	0.52
1:H:64:LEU:HD23	1:H:85:VAL:HG11	1.91	0.52
1:H:126:ARG:NH2	1:H:233:GLN:NE2	2.58	0.51
2:L:28:ARG:HA	2:L:37:THR:CG2	2.40	0.51
1:H:60(C):LYS:HG2	10:H:2029:HOH:O	2.10	0.51
1:H:84:ARG:N	1:H:84:ARG:HD2	2.25	0.51
3:T:76:PHE:CD1	4:U:92:PRO:HG2	2.46	0.50
3:T:72:LEU:HD13	3:T:73:ALA:N	2.27	0.50
1:H:126:ARG:HH22	1:H:233:GLN:NE2	2.11	0.49
4:U:166:LYS:HD2	4:U:166:LYS:N	2.28	0.49
1:H:193:GLY:HA3	6:H:1259:GIL:H282	1.95	0.48
4:U:155:LEU:HD11	4:U:187:PHE:HB3	1.95	0.48
4:U:165:LYS:C	4:U:166:LYS:HD2	2.35	0.48
1:H:236:GLU:HG3	10:H:2143:HOH:O	2.14	0.48
1:H:215:TRP:HB3	6:H:1259:GIL:C18	2.44	0.47
1:H:215:TRP:HB3	6:H:1259:GIL:H182	1.96	0.47
2:L:137:LYS:O	2:L:139:PRO:HD3	2.15	0.47
2:L:73:LEU:HD21	8:L:1145:FUC:H4	1.97	0.46
3:T:22:ILE:HD13	3:T:58:ASP:HA	1.96	0.46
4:U:136:ARG:HE	4:U:141:LEU:HD21	1.79	0.46
4:U:196:ARG:O	4:U:200:ARG:NH1	2.48	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:U:190:GLN:OE1	4:U:201:LYS:HE2	2.16	0.46
1:H:158:LEU:HD11	1:H:188:LYS:HD3	1.98	0.46
1:H:170(C):ARG:HG2	1:H:170(C):ARG:HH11	1.79	0.46
1:H:89:ILE:HG21	1:H:241:LEU:HD13	1.98	0.46
2:L:33:ASP:CG	2:L:36:ARG:HD3	2.36	0.46
4:U:136:ARG:HG2	4:U:136:ARG:HH11	1.81	0.46
1:H:35:VAL:HG23	1:H:35:VAL:O	2.16	0.46
1:H:219:GLY:C	1:H:221:THR:HG23	2.37	0.45
3:T:76:PHE:HD1	4:U:92:PRO:HG2	1.80	0.45
4:U:136:ARG:O	4:U:137:ASN:HB2	2.15	0.45
3:T:46:LYS:HD3	3:T:62:GLU:OE2	2.16	0.45
2:L:60:SER:CB	8:L:1145:FUC:O1	2.47	0.45
1:H:256:PHE:HA	1:H:257:PRO:C	2.37	0.45
2:L:71:PHE:CE2	4:U:131:ARG:HG3	2.51	0.45
1:H:35:VAL:HG12	1:H:64:LEU:CD1	2.47	0.43
4:U:110:GLN:HB2	4:U:203:THR:CG2	2.48	0.43
2:L:84:HIS:HE1	10:L:2028:HOH:O	2.00	0.43
1:H:125:GLU:HG3	2:L:101:TYR:CE2	2.53	0.43
2:L:110:ARG:C	2:L:110:ARG:HD2	2.39	0.43
4:U:186:CYS:HA	4:U:209:CYS:HA	2.01	0.43
3:T:46:LYS:NZ	10:T:2031:HOH:O	2.53	0.42
1:H:247:ARG:HB3	1:H:248:PRO:HD2	2.01	0.42
2:L:108:THR:OG1	2:L:109:LYS:N	2.52	0.42
2:L:33:ASP:HB3	2:L:36:ARG:CG	2.50	0.42
1:H:241:LEU:HA	1:H:244:SER:OG	2.19	0.42
4:U:124:ASN:CG	9:U:1211:NAG:H82	2.39	0.42
1:H:27:CYS:N	1:H:28:PRO:CD	2.82	0.42
2:L:9:ARG:CZ	2:L:15:ARG:HD3	2.50	0.42
1:H:65:ILE:HD11	1:H:84:ARG:HH12	1.84	0.42
2:L:120:LEU:HD12	2:L:126:SER:O	2.20	0.42
4:U:108:LEU:HD11	4:U:193:ILE:HG13	2.01	0.41
1:H:101:HIS:HA	1:H:234:TYR:OH	2.20	0.41
1:H:59:PHE:HA	1:H:60(B):ILE:HG12	2.02	0.41
2:L:108:THR:HG23	2:L:109:LYS:N	2.35	0.41
2:L:35:CGU:O	2:L:38:LYS:HG2	2.20	0.41
1:H:245:GLU:O	1:H:247:ARG:NE	2.53	0.41
1:H:31:VAL:HG22	1:H:68:LEU:HD21	2.03	0.41
1:H:248:PRO:HG2	1:H:249:GLY:H	1.86	0.41
4:U:116:PHE:HA	4:U:124:ASN:O	2.21	0.41
2:L:54:PRO:O	2:L:80:ASN:HB3	2.21	0.40
1:H:109:HIS:ND1	1:H:110:GLN:HG2	2.37	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:100:ASN:HD22	1:H:100:ASN:HA	1.61	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	H	252/254 (99%)	238 (94%)	12 (5%)	2 (1%)	24	7
2	L	127/142 (89%)	122 (96%)	5 (4%)	0	100	100
3	T	73/75 (97%)	72 (99%)	1 (1%)	0	100	100
4	U	112/116 (97%)	107 (96%)	5 (4%)	0	100	100
All	All	564/587 (96%)	539 (96%)	23 (4%)	2 (0%)	39	20

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	219	GLY
1	H	170(F)	GLY

### 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	H	216/216 (100%)	208 (96%)	8 (4%)	41	17

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	L	113/114 (99%)	111 (98%)	2 (2%)	66	48
3	T	70/70 (100%)	69 (99%)	1 (1%)	74	58
4	U	109/109 (100%)	108 (99%)	1 (1%)	84	75
All	All	508/509 (100%)	496 (98%)	12 (2%)	57	35

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	29	TRP
1	H	37	ASN
1	H	50	ILE
1	H	84	ARG
1	H	100	ASN
1	H	159	ASN
1	H	233	GLN
1	H	247	ARG
2	L	36	ARG
2	L	110	ARG
3	T	27	PRO
4	U	199	ASN

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

Mol	Chain	Res	Type
1	H	37	ASN
1	H	63	ASN
1	H	100	ASN
1	H	110	GLN
1	H	143	GLN
1	H	159	ASN
1	H	202	HIS
1	H	233	GLN
1	H	239	GLN
2	L	80	ASN
2	L	84	HIS
3	T	31	ASN
4	U	171	ASN
4	U	199	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

10 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	$\# Z  > 2$
2	CGU	L	14	2,5	4,11,12	0.49	0	4,14,16	2.04	2 (50%)
2	CGU	L	16	2,5	4,11,12	0.44	0	4,14,16	2.10	2 (50%)
2	CGU	L	19	2,5	4,11,12	0.49	0	4,14,16	1.97	2 (50%)
2	CGU	L	20	2	4,11,12	0.46	0	4,14,16	1.96	1 (25%)
2	CGU	L	25	2,5	4,11,12	0.56	0	4,14,16	2.04	2 (50%)
2	CGU	L	26	2,5	4,11,12	0.43	0	4,14,16	2.11	1 (25%)
2	CGU	L	29	2,5	4,11,12	0.46	0	4,14,16	1.89	1 (25%)
2	CGU	L	35	2	4,11,12	0.46	0	4,14,16	2.07	2 (50%)
2	CGU	L	6	2	4,11,12	0.51	0	4,14,16	1.91	1 (25%)
2	CGU	L	7	2	4,11,12	0.47	0	4,14,16	2.09	2 (50%)

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
2	CGU	L	14	2,5	-	0/4/14/16	0/0/0/0
2	CGU	L	16	2,5	-	0/4/14/16	0/0/0/0
2	CGU	L	19	2,5	-	0/4/14/16	0/0/0/0
2	CGU	L	20	2	-	0/4/14/16	0/0/0/0
2	CGU	L	25	2,5	-	0/4/14/16	0/0/0/0
2	CGU	L	26	2,5	-	0/4/14/16	0/0/0/0
2	CGU	L	29	2,5	-	0/4/14/16	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	CGU	L	35	2	-	0/4/14/16	0/0/0/0
2	CGU	L	6	2	-	0/4/14/16	0/0/0/0
2	CGU	L	7	2	-	0/4/14/16	0/0/0/0

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L	26	CGU	CB-CG-CD2	-3.80	105.07	112.83
2	L	16	CGU	CB-CG-CD2	-3.47	105.76	112.83
2	L	7	CGU	CB-CG-CD2	-3.32	106.06	112.83
2	L	35	CGU	CB-CG-CD2	-3.19	106.32	112.83
2	L	14	CGU	CB-CG-CD2	-3.16	106.39	112.83
2	L	29	CGU	CB-CG-CD2	-3.14	106.42	112.83
2	L	25	CGU	CB-CG-CD2	-3.11	106.49	112.83
2	L	19	CGU	CB-CG-CD2	-3.09	106.54	112.83
2	L	20	CGU	CB-CG-CD2	-3.07	106.56	112.83
2	L	6	CGU	CB-CG-CD2	-2.92	106.88	112.83
2	L	35	CGU	O-C-CA	-2.17	119.83	125.49
2	L	25	CGU	O-C-CA	-2.12	119.96	125.49
2	L	7	CGU	O-C-CA	-2.10	120.01	125.49
2	L	19	CGU	O-C-CA	-2.10	120.01	125.49
2	L	14	CGU	O-C-CA	-2.08	120.06	125.49
2	L	16	CGU	O-C-CA	-2.07	120.10	125.49

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	L	35	CGU	1	0

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.



## 5.6 Ligand geometry

Of 9 ligands modelled in this entry, 5 are monoatomic - leaving 4 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	$\# Z  > 2$
6	GIL	H	1259	-	40,44,44	1.68	9 (22%)	53,63,63	1.34	5 (9%)
7	GLC	L	1144	-	12,12,12	0.25	0	17,17,17	0.35	0
8	FUC	L	1145	-	11,11,11	0.30	0	16,16,16	0.43	0
9	NAG	U	1211	4	14,14,15	0.53	0	15,19,21	0.72	1 (6%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	GIL	H	1259	-	-	0/33/37/37	0/3/3/3
7	GLC	L	1144	-	-	0/2/22/22	0/1/1/1
8	FUC	L	1145	-	-	0/0/20/20	0/1/1/1
9	NAG	U	1211	4	-	0/6/23/26	0/1/1/1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	H	1259	GIL	C6-C1	2.12	1.42	1.39
6	H	1259	GIL	C2-C1	2.21	1.43	1.39
6	H	1259	GIL	C13-C14	2.37	1.43	1.39
6	H	1259	GIL	C11-N3	2.53	1.36	1.32
6	H	1259	GIL	C22-C21	2.75	1.45	1.40
6	H	1259	GIL	C23-C24	2.97	1.43	1.39
6	H	1259	GIL	C13-C12	3.02	1.44	1.39
6	H	1259	GIL	C15-N3	3.11	1.36	1.32
6	H	1259	GIL	C12-C11	4.32	1.43	1.38

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	H	1259	GIL	O2-C15-C14	-2.76	114.32	117.76
9	U	1211	NAG	C2-N2-C7	-2.08	120.36	123.04
6	H	1259	GIL	O2-C21-C26	-2.03	115.57	120.82
6	H	1259	GIL	C17-C10-N4	2.07	117.10	112.34
6	H	1259	GIL	O1-C11-C12	3.83	122.52	117.76
6	H	1259	GIL	C16-N4-C10	6.04	127.64	116.31

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	H	1259	GIL	3	0
7	L	1144	GLC	2	0
8	L	1145	FUC	4	0
9	U	1211	NAG	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	H	254/254 (100%)	0.66	26 (10%) 9 9	10, 19, 39, 61	1 (0%)
2	L	129/142 (90%)	1.06	25 (19%) 1 1	16, 27, 44, 51	0
3	T	75/75 (100%)	0.40	1 (1%) 79 83	12, 18, 25, 30	0
4	U	116/116 (100%)	1.19	20 (17%) 2 2	15, 23, 50, 55	1 (0%)
All	All	574/587 (97%)	0.82	72 (12%) 5 6	10, 22, 41, 61	2 (0%)

All (72) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	U	137	ASN	8.5
4	U	181	LYS	7.8
1	H	248	PRO	7.6
4	U	182	GLY	7.3
2	L	9	ARG	6.6
4	U	138	ASN	6.6
1	H	170(G)	ASP	6.3
4	U	156	TYR	5.8
2	L	122	ALA	5.3
1	H	247	ARG	5.2
1	H	62	ARG	5.0
1	H	147	ARG	5.0
4	U	197	THR	5.0
2	L	106	THR	4.9
1	H	219	GLY	4.7
4	U	164	GLY	4.5
1	H	37	ASN	4.4
3	T	80	ALA	4.4
4	U	163	SER	4.4
2	L	107	GLY	4.4
4	U	136	ARG	4.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	L	15	ARG	4.1
4	U	139	THR	3.9
1	H	249	GLY	3.8
4	U	180	ASP	3.7
2	L	21	GLN	3.6
4	U	198	VAL	3.6
2	L	108	THR	3.6
2	L	42	ILE	3.4
4	U	165	LYS	3.4
1	H	109	HIS	3.1
1	H	84	ARG	3.1
4	U	183	GLU	3.0
1	H	64	LEU	3.0
4	U	184	ASN	3.0
1	H	170(F)	GLY	3.0
2	L	66	GLN	2.9
2	L	10	PRO	2.9
1	H	245	GLU	2.9
2	L	18	LYS	2.9
2	L	69	ILE	2.8
1	H	60(C)	LYS	2.8
2	L	32	LYS	2.8
2	L	140	ILE	2.8
1	H	60(D)	ASN	2.7
4	U	131	ARG	2.7
2	L	11	GLY	2.7
1	H	68	LEU	2.7
2	L	51	ALA	2.6
1	H	90	ILE	2.6
1	H	243	ARG	2.5
4	U	166	LYS	2.5
1	H	236	GLU	2.5
2	L	70	CYS	2.4
2	L	37	THR	2.4
4	U	179	VAL	2.4
4	U	93	LEU	2.4
2	L	34	ALA	2.3
1	H	217	GLN	2.3
1	H	118	VAL	2.3
2	L	121	LEU	2.3
2	L	60	SER	2.2
1	H	63	ASN	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	U	210	MET	2.2
2	L	38	LYS	2.2
1	H	112	VAL	2.2
1	H	227	VAL	2.2
2	L	48	ASP	2.1
1	H	170(H)	SER	2.1
1	H	60(A)	LYS	2.1
2	L	120	LEU	2.0
2	L	142	GLU	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. 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
2	CGU	L	7	12/13	0.81	0.20	-	30,33,36,38	0
2	CGU	L	19	12/13	0.75	0.33	-	46,51,54,54	0
2	CGU	L	6	12/13	0.81	0.23	-	26,34,37,38	0
2	CGU	L	14	12/13	0.71	0.29	-	37,46,48,49	0
2	CGU	L	25	12/13	0.87	0.20	-	25,26,28,29	0
2	CGU	L	29	12/13	0.90	0.13	-	25,27,30,30	0
2	CGU	L	26	12/13	0.85	0.20	-	23,29,33,35	0
2	CGU	L	16	12/13	0.86	0.17	-	34,35,37,37	0
2	CGU	L	35	12/13	0.42	0.47	-	39,51,53,54	0
2	CGU	L	20	12/13	0.72	0.34	-	44,50,52,53	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( $\text{\AA}^2$ )	Q<0.9
9	NAG	U	1211	14/15	0.81	0.21	1.69	24,31,41,41	0
8	FUC	L	1145	11/11	0.67	0.19	0.87	35,36,39,42	0
6	GIL	H	1259	42/42	0.88	0.14	0.46	14,25,34,36	0
5	CA	H	1258	1/1	0.98	0.13	0.14	29,29,29,29	0
5	CA	L	1143	1/1	0.97	0.05	-1.70	28,28,28,28	0
5	CA	A	4	1/1	0.86	0.14	-	85,85,85,85	0
5	CA	A	9	1/1	0.95	0.14	-	45,45,45,45	0
5	CA	A	3	1/1	0.96	0.16	-	46,46,46,46	0
7	GLC	L	1144	12/12	0.70	0.19	-	50,52,52,52	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.