



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 12:20 AM GMT

PDB ID : 1ONQ  
Title : Crystal Structure of CD1a in Complex with a Sulfatide  
Authors : Zajonc, D.M.; Elsliger, M.A.; Teyton, L.; Wilson, I.A.  
Deposited on : 2003-02-28  
Resolution : 2.15 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

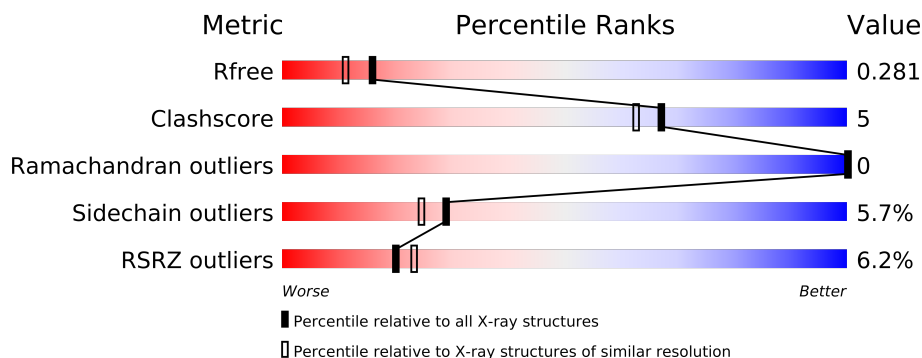
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.15 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))
$R_{free}$	66092	1094 (2.18-2.14)
Clashscore	79885	1299 (2.18-2.14)
Ramachandran outliers	78287	1272 (2.18-2.14)
Sidechain outliers	78261	1272 (2.18-2.14)
RSRZ outliers	66119	1094 (2.18-2.14)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	283	
1	C	283	
2	B	99	
2	D	99	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	FUC	C	532	-	X
7	SLF	A	601	-	X
7	SLF	C	602	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 6791 atoms, of which 0 are hydrogen and 0 are deuterium.

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 T-cell surface glycoprotein CD1a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	274	Total	C	N	O	S	0	0	0
			2228	1428	393	399	8			
1	C	273	Total	C	N	O	S	0	0	0
			2218	1422	390	398	8			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	278	HIS	-	EXPRESSION TAG	UNP P06126
A	279	HIS	-	EXPRESSION TAG	UNP P06126
A	280	HIS	-	EXPRESSION TAG	UNP P06126
A	281	HIS	-	EXPRESSION TAG	UNP P06126
A	282	HIS	-	EXPRESSION TAG	UNP P06126
A	283	HIS	-	EXPRESSION TAG	UNP P06126
C	278	HIS	-	EXPRESSION TAG	UNP P06126
C	279	HIS	-	EXPRESSION TAG	UNP P06126
C	280	HIS	-	EXPRESSION TAG	UNP P06126
C	281	HIS	-	EXPRESSION TAG	UNP P06126
C	282	HIS	-	EXPRESSION TAG	UNP P06126
C	283	HIS	-	EXPRESSION TAG	UNP P06126

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			829	528	140	158	3			
2	D	99	Total	C	N	O	S	0	0	0
			829	528	140	158	3			

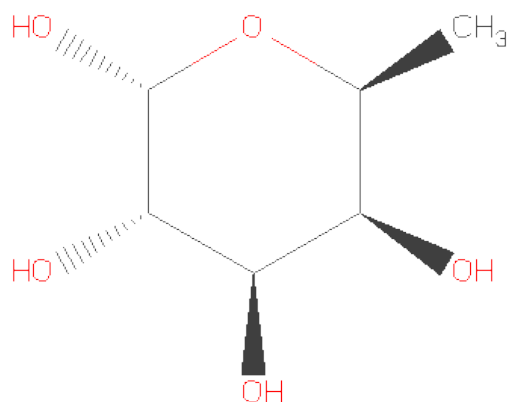
- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	C	2	Total	C	N	O	0	0
			28	16	2	10		

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	278	HIS	-	EXPRESSION TAG	UNP P06126
A	279	HIS	-	EXPRESSION TAG	UNP P06126
A	280	HIS	-	EXPRESSION TAG	UNP P06126
A	281	HIS	-	EXPRESSION TAG	UNP P06126
A	282	HIS	-	EXPRESSION TAG	UNP P06126
A	283	HIS	-	EXPRESSION TAG	UNP P06126
A	278	HIS	-	EXPRESSION TAG	UNP P06126
A	279	HIS	-	EXPRESSION TAG	UNP P06126
A	280	HIS	-	EXPRESSION TAG	UNP P06126
A	281	HIS	-	EXPRESSION TAG	UNP P06126
A	282	HIS	-	EXPRESSION TAG	UNP P06126
A	283	HIS	-	EXPRESSION TAG	UNP P06126
C	278	HIS	-	EXPRESSION TAG	UNP P06126
C	279	HIS	-	EXPRESSION TAG	UNP P06126
C	280	HIS	-	EXPRESSION TAG	UNP P06126
C	281	HIS	-	EXPRESSION TAG	UNP P06126
C	282	HIS	-	EXPRESSION TAG	UNP P06126
C	283	HIS	-	EXPRESSION TAG	UNP P06126

- Molecule 4 is SUGAR (FUCOSE) (three-letter code: FUC) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>5</sub>).



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

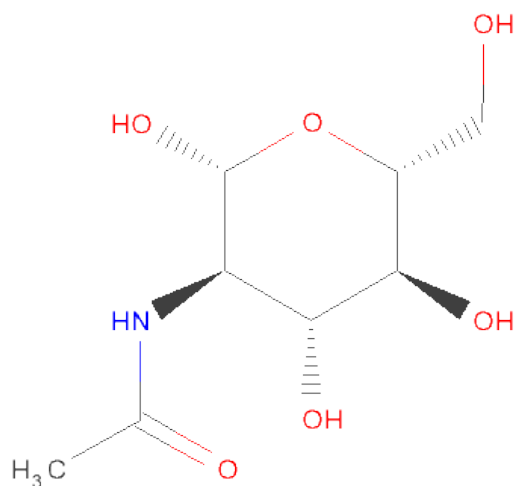
- Molecule 5 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	2	Total	C	N	O	0	0
			24	14	1	9		

There are 6 discrepancies between the modelled and reference sequences:

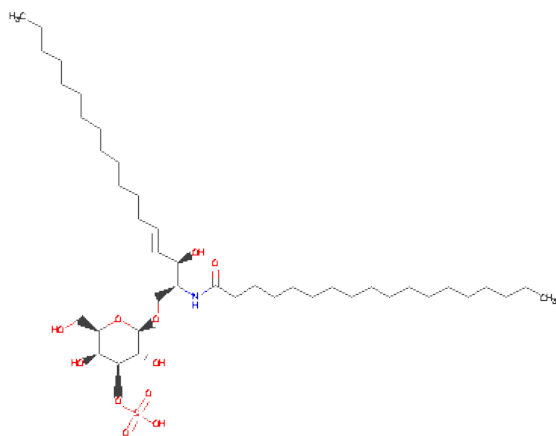
Chain	Residue	Modelled	Actual	Comment	Reference
A	278	HIS	-	EXPRESSION TAG	UNP P06126
A	279	HIS	-	EXPRESSION TAG	UNP P06126
A	280	HIS	-	EXPRESSION TAG	UNP P06126
A	281	HIS	-	EXPRESSION TAG	UNP P06126
A	282	HIS	-	EXPRESSION TAG	UNP P06126
A	283	HIS	-	EXPRESSION TAG	UNP P06126

- Molecule 6 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 7 is SULFATE-3-D-GALACTOSYL-BETA-1-1-N-STEAROYL-D-SPHINGOSIN E (three-letter code: SLF) (formula: C<sub>42</sub>H<sub>81</sub>NO<sub>11</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	S	0	0
			55	42	1	11	1		
7	C	1	Total	C	N	O		0	0
			40	36	1	3			

- Molecule 8 is water.

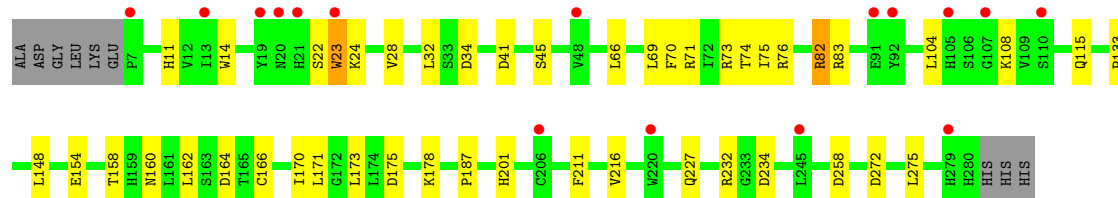
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	216	Total 216	O 216	0	0
8	B	82	Total 82	O 82	0	0
8	C	103	Total 103	O 103	0	0
8	D	49	Total 49	O 49	0	0

### 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.

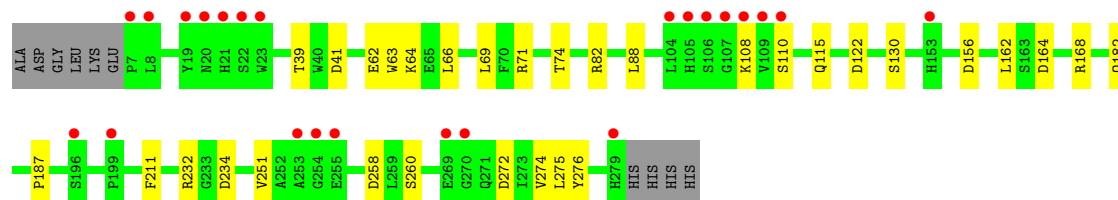
- Molecule 1: T-cell surface glycoprotein CD1a

Chain A: 



- Molecule 1: T-cell surface glycoprotein CD1a

Chain C: 



- Molecule 2: Beta-2-microglobulin

Chain B: 



- Molecule 2: Beta-2-microglobulin

Chain D: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.41Å 42.71Å 204.22Å 90.00° 90.78° 90.00°	Depositor
Resolution (Å)	40.00 – 2.15 39.40 – 2.15	Depositor EDS
% Data completeness (in resolution range)	92.0 (40.00-2.15) 92.0 (39.40-2.15)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.50 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.1.24/TLS	Depositor
R, $R_{free}$	0.227 , 0.276 0.234 , 0.281	Depositor DCC
$R_{free}$ test set	2236 reflections (4.82%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.9	Xtriage
Anisotropy	0.361	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 28.5	EDS
Estimated twinning fraction	0.029 for h,-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 48579 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	6791	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.92% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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: SLF, 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	0.62	0/2304	0.82	7/3136 (0.2%)
1	C	0.50	0/2293	0.74	7/3121 (0.2%)
2	B	0.68	1/852 (0.1%)	0.87	4/1152 (0.3%)
2	D	0.50	0/852	0.76	3/1152 (0.3%)
All	All	0.57	1/6301 (0.0%)	0.79	21/8561 (0.2%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
5	A	1	0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	44	GLU	CD-OE2	5.25	1.31	1.25

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	34	ASP	CB-CG-OD2	6.90	124.51	118.30
2	B	38	ASP	CB-CG-OD2	6.26	123.93	118.30
1	C	41	ASP	CB-CG-OD2	6.22	123.90	118.30
1	A	164	ASP	CB-CG-OD2	6.11	123.80	118.30
1	C	156	ASP	CB-CG-OD2	6.07	123.76	118.30
1	C	234	ASP	CB-CG-OD2	6.01	123.71	118.30
2	B	34	ASP	CB-CG-OD2	5.99	123.69	118.30
1	A	258	ASP	CB-CG-OD2	5.99	123.69	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	272	ASP	CB-CG-OD2	5.94	123.65	118.30
1	A	175	ASP	CB-CG-OD2	5.91	123.62	118.30
1	C	272	ASP	CB-CG-OD2	5.83	123.54	118.30
1	A	41	ASP	CB-CG-OD2	5.73	123.45	118.30
1	A	234	ASP	CB-CG-OD2	5.58	123.32	118.30
1	C	164	ASP	CB-CG-OD2	5.34	123.10	118.30
2	B	59	ASP	CB-CG-OD2	5.28	123.06	118.30
2	D	34	ASP	CB-CG-OD2	5.25	123.02	118.30
2	B	76	ASP	CB-CG-OD2	5.20	122.98	118.30
2	D	76	ASP	CB-CG-OD2	5.11	122.89	118.30
2	D	98	ASP	CB-CG-OD2	5.04	122.84	118.30
1	C	122	ASP	CB-CG-OD2	5.03	122.83	118.30
1	C	258	ASP	CB-CG-OD2	5.00	122.81	118.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	511	NAG	C1

There are no planarity outliers.

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2228	0	2103	30	0
1	C	2218	0	2097	12	0
2	B	829	0	794	5	1
2	D	829	0	794	4	0
3	A	56	0	50	4	0
3	C	28	0	25	4	0
4	A	10	0	10	2	0
4	C	10	0	10	1	0
5	A	24	0	22	1	0
6	C	14	0	13	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	55	0	81	16	0
7	C	40	0	69	3	0
8	A	216	0	0	10	1
8	B	82	0	0	1	0
8	C	103	0	0	0	0
8	D	49	0	0	1	0
All	All	6791	0	6068	62	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

All (62) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
7:A:601:SLF:H312	7:A:601:SLF:HO9	1.40	0.86
7:A:601:SLF:O9	7:A:601:SLF:H312	1.77	0.85
3:A:501:NAG:O6	4:A:531:FUC:C1	2.26	0.83
1:A:73:ARG:CD	7:A:601:SLF:O4	2.36	0.74
3:C:501:NAG:O6	4:C:532:FUC:C1	2.38	0.72
2:B:50:GLU:OE2	8:B:137:HOH:O	2.09	0.71
1:A:73:ARG:NE	7:A:601:SLF:O4	2.28	0.65
1:A:11:HIS:ND1	8:A:751:HOH:O	2.29	0.64
1:A:73:ARG:NH1	7:A:601:SLF:H1A	2.13	0.63
2:D:31:HIS:ND1	2:D:32:PRO:HA	2.14	0.62
1:A:216:VAL:O	1:A:232:ARG:NH2	2.34	0.60
1:A:23:TRP:CZ3	5:A:512:FUC:H61	2.38	0.59
7:A:601:SLF:H211	7:A:601:SLF:HN	1.68	0.59
1:C:168:ARG:HD2	3:C:501:NAG:H83	1.84	0.59
1:A:32:LEU:HD13	1:A:173:LEU:HD23	1.86	0.57
1:A:187:PRO:HB3	1:A:211:PHE:HB3	1.85	0.57
1:C:74:THR:HG22	7:C:602:SLF:H221	1.85	0.57
1:C:62:GLU:O	1:C:66:LEU:HD23	2.05	0.57
7:A:601:SLF:O4	7:A:601:SLF:H2A	2.06	0.55
7:A:601:SLF:N	7:A:601:SLF:H211	2.22	0.55
1:A:158:THR:HG23	7:A:601:SLF:O1	2.09	0.53
1:A:45:SER:OG	1:A:71:ARG:HD3	2.09	0.53
1:C:260:SER:HB3	1:C:276:TYR:CD2	2.44	0.53
1:C:187:PRO:HB3	1:C:211:PHE:HB3	1.91	0.52
2:B:91:LYS:C	2:B:92:ILE:HD12	2.32	0.50
1:A:73:ARG:CG	7:A:601:SLF:O4	2.60	0.50
1:A:133:PRO:HB2	8:A:817:HOH:O	2.12	0.50
1:A:11:HIS:CE1	8:A:751:HOH:O	2.64	0.50
1:C:168:ARG:HH11	3:C:501:NAG:H81	1.77	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:274:VAL:HG12	1:C:276:TYR:CE1	2.47	0.49
1:A:148:LEU:CD2	7:A:601:SLF:O9	2.60	0.49
1:A:70:PHE:O	1:A:74:THR:HG23	2.12	0.48
2:D:96:ASP:HB3	2:D:99:MET:HB2	1.94	0.48
1:A:73:ARG:HD3	7:A:601:SLF:O4	2.13	0.48
1:A:69:LEU:HD11	1:A:73:ARG:HD2	1.96	0.48
1:A:154:GLU:HG2	7:A:601:SLF:O7	2.14	0.48
3:A:501:NAG:HO6	4:A:531:FUC:C1	2.26	0.48
2:B:33:SER:HB2	2:B:54:LEU:HD21	1.96	0.48
1:C:74:THR:HG22	7:C:602:SLF:C22	2.43	0.48
1:C:274:VAL:HG12	1:C:276:TYR:CZ	2.49	0.47
1:A:14:TRP:HD1	1:A:28:VAL:HG22	1.81	0.46
1:A:160:ASN:HB2	8:A:761:HOH:O	2.15	0.46
1:A:70:PHE:CD1	7:A:601:SLF:H201	2.50	0.46
1:A:104:LEU:HD13	1:A:170:ILE:HD13	1.98	0.46
7:A:601:SLF:C	7:A:601:SLF:O4	2.64	0.46
1:C:168:ARG:HH11	3:C:501:NAG:C8	2.29	0.46
3:A:521:NAG:H83	8:A:715:HOH:O	2.16	0.45
1:A:115:GLN:NE2	8:A:687:HOH:O	2.49	0.45
1:A:171:LEU:HD12	3:A:502:NAG:H82	1.98	0.45
1:A:71:ARG:O	1:A:75:ILE:HD12	2.17	0.45
1:A:227:GLN:NE2	8:A:680:HOH:O	2.25	0.44
1:C:274:VAL:CG1	1:C:276:TYR:CZ	3.01	0.44
1:A:82:ARG:NH1	8:A:730:HOH:O	2.51	0.43
1:A:14:TRP:CD1	1:A:28:VAL:HG22	2.55	0.42
2:D:58:LYS:HE3	8:D:142:HOH:O	2.20	0.42
2:D:37:VAL:HB	2:D:66:TYR:CZ	2.55	0.41
2:B:21:ASN:HB3	2:B:70:PHE:CE1	2.55	0.41
7:A:601:SLF:C21	7:A:601:SLF:HN	2.34	0.40
2:B:92:ILE:HD12	2:B:92:ILE:N	2.37	0.40
1:A:201:HIS:CD2	8:A:749:HOH:O	2.74	0.40
1:A:201:HIS:HD2	8:A:749:HOH:O	2.04	0.40
1:C:63:TRP:CZ3	7:C:602:SLF:H131	2.57	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:16:GLU:OE2	8:A:739:HOH:O[2_556]	2.15	0.05

## 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	272/283 (96%)	264 (97%)	8 (3%)	0	100	100
1	C	271/283 (96%)	264 (97%)	7 (3%)	0	100	100
2	B	97/99 (98%)	96 (99%)	1 (1%)	0	100	100
2	D	97/99 (98%)	96 (99%)	1 (1%)	0	100	100
All	All	737/764 (96%)	720 (98%)	17 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	239/246 (97%)	227 (95%)	12 (5%)	34	30
1	C	238/246 (97%)	223 (94%)	15 (6%)	25	20
2	B	94/94 (100%)	86 (92%)	8 (8%)	15	10
2	D	94/94 (100%)	91 (97%)	3 (3%)	51	53
All	All	665/680 (98%)	627 (94%)	38 (6%)	29	25

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	22	SER
1	A	23	TRP
1	A	24	LYS
1	A	66	LEU
1	A	76	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	82	ARG
1	A	83	ARG
1	A	108	LYS
1	A	162	LEU
1	A	166	CYS
1	A	178	LYS
1	A	275	LEU
2	B	4	THR
2	B	16	GLU
2	B	45	ARG
2	B	70	PHE
2	B	75	LYS
2	B	88	SER
2	B	89	GLN
2	B	99	MET
1	C	39	THR
1	C	64	LYS
1	C	69	LEU
1	C	71	ARG
1	C	82	ARG
1	C	88	LEU
1	C	108	LYS
1	C	110	SER
1	C	115	GLN
1	C	130	SER
1	C	162	LEU
1	C	182	GLN
1	C	232	ARG
1	C	251	VAL
1	C	275	LEU
2	D	19	LYS
2	D	45	ARG
2	D	70	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	11	HIS
1	A	25	GLN
1	A	184	GLN
1	A	228	GLN
1	C	115	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	143	HIS
1	C	150	GLN
1	C	180	HIS
1	C	184	GLN
1	C	193	HIS
1	C	225	GLN
1	C	227	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.5 Carbohydrates ⓘ

8 carbohydrates 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	501	1,3	12,14,15	0.70	0	15,19,21	1.36	2 (13%)
3	NAG	A	502	3	12,14,15	0.92	0	15,19,21	2.13	5 (33%)
5	NAG	A	511	1,5	12,14,15	0.62	0	15,19,21	1.52	2 (13%)
5	FUC	A	512	5	9,10,11	0.77	0	10,14,16	0.53	0
3	NAG	A	521	1,3	12,14,15	0.68	0	15,19,21	1.21	1 (6%)
3	NAG	A	522	3	12,14,15	0.65	0	15,19,21	1.39	1 (6%)
3	NAG	C	501	1,3	12,14,15	0.71	0	15,19,21	1.43	3 (20%)
3	NAG	C	502	3	12,14,15	0.39	0	15,19,21	1.56	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	501	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	502	3	-	0/6/23/26	0/1/1/1
5	NAG	A	511	1,5	1/1/5/7	0/6/23/26	0/1/1/1
5	FUC	A	512	5	-	0/0/17/20	0/1/1/1
3	NAG	A	521	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	522	3	-	0/6/23/26	0/1/1/1
3	NAG	C	501	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	502	3	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	502	NAG	O5-C5-C6	5.03	112.26	106.98
5	A	511	NAG	O5-C5-C6	4.61	111.81	106.98
3	A	522	NAG	O5-C5-C6	4.16	111.35	106.98
3	C	502	NAG	O5-C5-C4	3.98	115.70	110.65
3	A	502	NAG	C3-C2-N2	-3.76	106.04	111.76
3	C	501	NAG	O7-C7-C8	-3.09	116.01	122.04
3	A	521	NAG	O5-C5-C6	2.86	109.98	106.98
5	A	511	NAG	O5-C5-C4	-2.61	107.34	110.65
3	C	501	NAG	C2-N2-C7	2.46	127.23	123.09
3	A	502	NAG	O3-C3-C2	2.43	114.18	109.09
3	C	502	NAG	C3-C4-C5	2.34	114.38	110.20
3	C	502	NAG	C6-C5-C4	-2.33	107.37	113.00
3	A	501	NAG	O3-C3-C4	-2.21	105.40	110.35
3	A	502	NAG	O7-C7-C8	-2.17	117.81	122.04
3	A	501	NAG	C3-C2-N2	2.09	114.95	111.76
3	A	502	NAG	O4-C4-C5	2.07	114.75	109.28
3	C	501	NAG	C8-C7-N2	2.04	120.10	116.11
3	C	502	NAG	O5-C5-C6	2.01	109.09	106.98

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	511	NAG	C1

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry

5 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$
4	FUC	A	531	-	9,10,11	0.78	0	10,14,16	0.57	0
7	SLF	A	601	-	55,55,55	1.07	3 (5%)	66,66,66	1.86	15 (22%)
6	NAG	C	511	1	12,14,15	0.67	0	15,19,21	1.23	2 (13%)
4	FUC	C	532	-	9,10,11	0.87	0	10,14,16	1.05	0
7	SLF	C	602	-	39,39,55	0.46	0	41,41,66	1.93	8 (19%)

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	FUC	A	531	-	-	0/0/17/20	0/1/1/1
7	SLF	A	601	-	-	0/52/72/72	0/1/1/1
6	NAG	C	511	1	-	1/6/23/26	0/1/1/1
4	FUC	C	532	-	-	0/0/17/20	0/1/1/1
7	SLF	C	602	-	1/1/3/12	0/42/42/72	0/0/0/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	601	SLF	O5-S	4.51	1.66	1.60
7	A	601	SLF	O-C36	4.33	1.48	1.40
7	A	601	SLF	C2-C3	2.04	1.53	1.49

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	601	SLF	O-C36-C37	6.21	116.09	108.18
7	C	602	SLF	C-C1-C2	6.14	126.41	112.28

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	C	602	SLF	C-C1-N	5.90	118.02	109.74
7	A	601	SLF	C41-C40-C39	-4.81	101.38	113.00
7	C	602	SLF	C1-C2-C3	4.34	117.73	111.84
7	A	601	SLF	C-C1-N	4.13	115.54	109.74
7	A	601	SLF	C2-C1-N	-4.01	102.56	110.25
7	A	601	SLF	O8-C40-C41	3.83	115.76	106.34
7	A	601	SLF	C-O-C36	3.74	121.27	113.81
7	A	601	SLF	O5-C38-C37	3.73	116.01	107.51
7	C	602	SLF	O1-C2-C3	-3.62	100.48	110.90
7	A	601	SLF	O-C-C1	3.53	116.58	107.93
7	C	602	SLF	C19-C18-N	-3.52	109.13	115.83
7	C	602	SLF	C1-N-C18	3.22	129.80	122.97
7	A	601	SLF	C37-C38-C39	-3.14	106.22	110.68
7	A	601	SLF	C38-O5-S	2.99	123.24	118.37
6	C	511	NAG	O5-C5-C6	2.78	109.89	106.98
7	A	601	SLF	O8-C40-C39	2.63	114.62	109.76
7	A	601	SLF	C20-C19-C18	-2.40	106.48	113.31
7	A	601	SLF	C1-N-C18	2.39	128.04	122.97
7	C	602	SLF	C2-C1-N	2.11	114.28	110.25
6	C	511	NAG	O5-C5-C4	-2.07	108.02	110.65
7	A	601	SLF	O8-C36-C37	-2.05	106.10	110.31
7	C	602	SLF	C5-C4-C3	-2.05	117.60	125.64
7	A	601	SLF	C36-C37-C38	2.01	113.69	110.03

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
7	C	602	SLF	C1

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	C	511	NAG	O7-C7-N2-C2

There are no ring outliers.

## 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	274/283 (96%)	0.59	16 (5%) 22 26	6, 14, 24, 29	0
1	C	273/283 (96%)	0.60	23 (8%) 11 12	7, 14, 21, 25	0
2	B	99/99 (100%)	0.53	1 (1%) 79 83	10, 15, 21, 27	0
2	D	99/99 (100%)	0.52	7 (7%) 16 18	9, 15, 19, 25	0
All	All	745/764 (97%)	0.58	47 (6%) 20 22	6, 15, 22, 29	0

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	109	VAL	5.7
1	A	23	TRP	5.1
1	C	107	GLY	4.9
1	C	19	TYR	4.8
1	C	21	HIS	4.8
1	C	23	TRP	4.5
1	A	21	HIS	4.2
1	C	105	HIS	4.0
2	D	18	GLY	3.8
2	D	99	MET	3.8
1	C	279	HIS	3.8
1	C	108	LYS	3.6
1	C	22	SER	3.6
1	C	196	SER	3.5
1	C	269	GLU	3.5
1	A	19	TYR	3.3
1	C	106	SER	3.2
1	C	20	ASN	3.2
2	D	75	LYS	3.1
1	A	13	ILE	3.1
1	C	199	PRO	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	73	THR	2.9
1	C	104	LEU	2.8
1	A	110	SER	2.8
1	C	254	GLY	2.8
1	C	153	HIS	2.6
1	A	20	ASN	2.6
1	C	270	GLY	2.6
1	A	91	GLU	2.4
1	A	245	LEU	2.4
2	D	43	GLY	2.4
1	C	255	GLU	2.3
1	A	279	HIS	2.2
1	A	107	GLY	2.2
1	A	220	TRP	2.2
2	D	89	GLN	2.2
2	D	44	GLU	2.2
1	A	105	HIS	2.2
1	A	48	VAL	2.1
1	A	92	TYR	2.1
1	C	8	LEU	2.1
1	C	110	SER	2.1
1	A	7	PRO	2.1
1	A	206	CYS	2.1
1	C	253	ALA	2.0
2	B	56	PHE	2.0
1	C	7	PRO	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	A	502	14/15	0.18	10.32	12,15,18,20	0
5	NAG	A	511	14/15	0.45	2.37	29,30,32,33	0
3	NAG	A	521	14/15	0.23	1.49	13,14,16,18	0
3	NAG	C	501	14/15	0.21	1.36	12,14,15,17	0
3	NAG	A	522	14/15	0.24	0.80	14,16,19,20	0
3	NAG	C	502	14/15	0.26	0.39	16,17,17,18	0
5	FUC	A	512	10/11	0.27	-0.08	27,27,28,29	0
3	NAG	A	501	14/15	0.09	-1.16	8,12,15,16	0

## 6.4 Ligands

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	FUC	C	532	10/11	0.34	9.78	17,18,19,19	0
7	SLF	C	602	40/55	0.30	5.53	18,25,30,32	0
7	SLF	A	601	55/55	0.26	2.33	18,32,44,46	0
4	FUC	A	531	10/11	0.16	0.77	21,22,23,23	0
6	NAG	C	511	14/15	0.31	-	16,17,18,18	0

## 6.5 Other polymers

There are no such residues in this entry.