



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:46 PM GMT

PDB ID : 4MNG  
Title : Structure of the DP10.7 TCR with CD1d-sulfatide  
Authors : Luoma, A.M.; Adams, E.J.  
Deposited on : 2013-09-10  
Resolution : 3.01 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

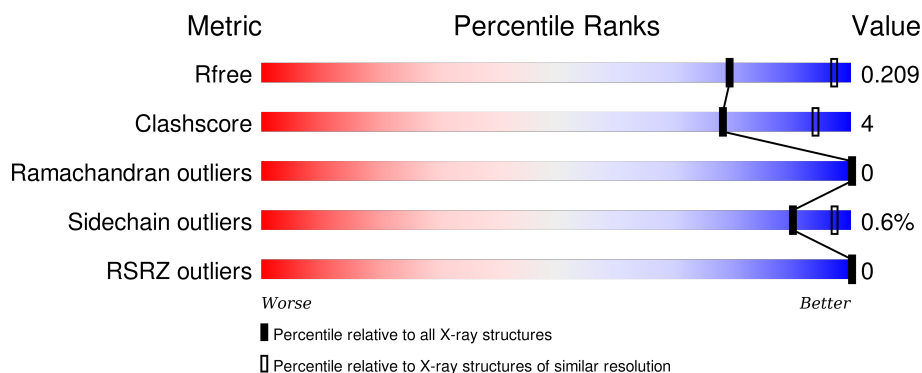
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.01 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	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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	B	99	<div> <div>95%</div> <div>5%</div> </div>
1	D	99	<div> <div>92%</div> <div>8%</div> </div>
2	A	281	<div> <div>87%</div> <div>10%</div> <div>•</div> </div>
2	C	281	<div> <div>88%</div> <div>10%</div> <div>•</div> </div>
3	E	262	<div> <div>77%</div> <div>10%</div> <div>13%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	262	 77%9%13%

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
4	NAG	A	303	-	-	-	X
4	NAG	C	303	-	-	-	X
5	CIS	A	304	-	-	-	X
5	CIS	C	304	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 9775 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 Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	99	Total	C	N	O	S	0	0	0
			798	509	135	147	7			
1	D	99	Total	C	N	O	S	0	0	0
			806	515	137	147	7			

- Molecule 2 is a protein called Antigen-presenting glycoprotein CD1d, Cd1d1 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	273	Total	C	N	O	S	0	0	0
			2177	1389	381	399	8			
2	C	273	Total	C	N	O	S	0	0	0
			2177	1387	380	402	8			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ALA	-	EXPRESSION TAG	UNP P15813
A	1	ASP	-	EXPRESSION TAG	UNP P15813
A	2	PRO	-	EXPRESSION TAG	UNP P15813
A	278	SER	-	EXPRESSION TAG	UNP Q7TMK5
A	279	GLY	-	EXPRESSION TAG	UNP Q7TMK5
C	0	ALA	-	EXPRESSION TAG	UNP P15813
C	1	ASP	-	EXPRESSION TAG	UNP P15813
C	2	PRO	-	EXPRESSION TAG	UNP P15813
C	278	SER	-	EXPRESSION TAG	UNP Q7TMK5
C	279	GLY	-	EXPRESSION TAG	UNP Q7TMK5

- Molecule 3 is a protein called TRA@ protein, Ti antigen CD3-associated protein gamma chain V-J-C region.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	228	Total	C	N	O	S	0	0	0
			1796	1153	299	337	7			
3	F	228	Total	C	N	O	S	0	0	0
			1797	1152	298	340	7			

There are 82 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	?	-	SER	DELETION	UNP Q6PJ56
E	?	-	PHE	DELETION	UNP Q6PJ56
E	?	-	LEU	DELETION	UNP Q6PJ56
E	97	SER	PHE	CONFLICT	UNP Q6PJ56
E	98	TYR	ARG	CONFLICT	UNP Q6PJ56
E	99	TRP	GLY	CONFLICT	UNP Q6PJ56
E	100	GLY	ASN	CONFLICT	UNP Q6PJ56
E	102	PRO	HIS	CONFLICT	UNP Q6PJ56
E	103	ARG	TYR	CONFLICT	UNP Q6PJ56
E	105	THR	ASP	CONFLICT	UNP Q6PJ56
E	106	ARG	LYS	CONFLICT	UNP Q6PJ56
E	107	VAL	LEU	CONFLICT	UNP Q6PJ56
E	120	GLY	-	LINKER	UNP 4MNG
E	121	GLY	-	LINKER	UNP 4MNG
E	122	GLY	-	LINKER	UNP 4MNG
E	123	GLY	-	LINKER	UNP 4MNG
E	124	SER	-	LINKER	UNP 4MNG
E	125	GLY	-	LINKER	UNP 4MNG
E	126	GLY	-	LINKER	UNP 4MNG
E	127	GLY	-	LINKER	UNP 4MNG
E	128	GLY	-	LINKER	UNP 4MNG
E	129	SER	-	LINKER	UNP 4MNG
E	130	GLY	-	LINKER	UNP 4MNG
E	131	GLY	-	LINKER	UNP 4MNG
E	132	GLY	-	LINKER	UNP 4MNG
E	133	GLY	-	LINKER	UNP 4MNG
E	134	SER	-	LINKER	UNP 4MNG
E	135	GLY	-	LINKER	UNP 4MNG
E	136	GLY	-	LINKER	UNP 4MNG
E	137	GLY	-	LINKER	UNP 4MNG
E	138	GLY	-	LINKER	UNP 4MNG
E	139	SER	-	LINKER	UNP 4MNG
E	251	LYS	THR	ENGINEERED MUTATION	UNP 4MNG
E	253	ILE	VAL	ENGINEERED MUTATION	UNP 4MNG
E	254	ILE	VAL	ENGINEERED MUTATION	UNP 4MNG

*Continued on next page...*

*Continued from previous page...*

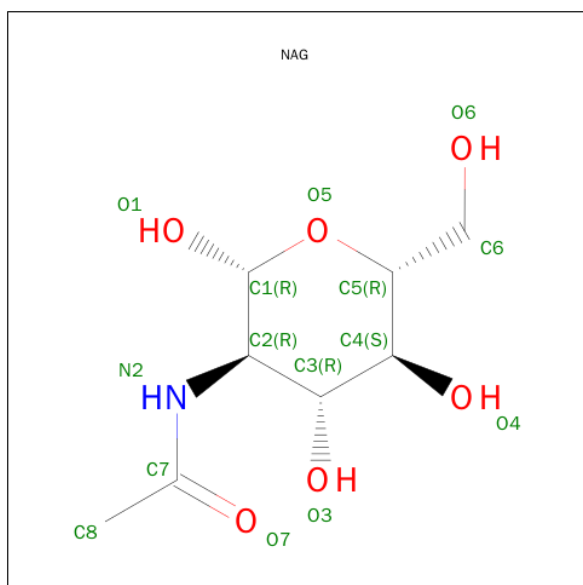
Chain	Residue	Modelled	Actual	Comment	Reference
E	257	ALA	-	EXPRESSION TAG	UNP 4MNG
E	258	ALA	-	EXPRESSION TAG	UNP 4MNG
E	259	SER	-	EXPRESSION TAG	UNP 4MNG
E	260	GLY	-	EXPRESSION TAG	UNP 4MNG
E	261	ALA	-	EXPRESSION TAG	UNP 4MNG
E	262	ASP	-	EXPRESSION TAG	UNP 4MNG
F	?	-	SER	DELETION	UNP Q6PJ56
F	?	-	PHE	DELETION	UNP Q6PJ56
F	?	-	LEU	DELETION	UNP Q6PJ56
F	97	SER	PHE	CONFLICT	UNP Q6PJ56
F	98	TYR	ARG	CONFLICT	UNP Q6PJ56
F	99	TRP	GLY	CONFLICT	UNP Q6PJ56
F	100	GLY	ASN	CONFLICT	UNP Q6PJ56
F	102	PRO	HIS	CONFLICT	UNP Q6PJ56
F	103	ARG	TYR	CONFLICT	UNP Q6PJ56
F	105	THR	ASP	CONFLICT	UNP Q6PJ56
F	106	ARG	LYS	CONFLICT	UNP Q6PJ56
F	107	VAL	LEU	CONFLICT	UNP Q6PJ56
F	120	GLY	-	LINKER	UNP 4MNG
F	121	GLY	-	LINKER	UNP 4MNG
F	122	GLY	-	LINKER	UNP 4MNG
F	123	GLY	-	LINKER	UNP 4MNG
F	124	SER	-	LINKER	UNP 4MNG
F	125	GLY	-	LINKER	UNP 4MNG
F	126	GLY	-	LINKER	UNP 4MNG
F	127	GLY	-	LINKER	UNP 4MNG
F	128	GLY	-	LINKER	UNP 4MNG
F	129	SER	-	LINKER	UNP 4MNG
F	130	GLY	-	LINKER	UNP 4MNG
F	131	GLY	-	LINKER	UNP 4MNG
F	132	GLY	-	LINKER	UNP 4MNG
F	133	GLY	-	LINKER	UNP 4MNG
F	134	SER	-	LINKER	UNP 4MNG
F	135	GLY	-	LINKER	UNP 4MNG
F	136	GLY	-	LINKER	UNP 4MNG
F	137	GLY	-	LINKER	UNP 4MNG
F	138	GLY	-	LINKER	UNP 4MNG
F	139	SER	-	LINKER	UNP 4MNG
F	251	LYS	THR	ENGINEERED MUTATION	UNP 4MNG
F	253	ILE	VAL	ENGINEERED MUTATION	UNP 4MNG
F	254	ILE	VAL	ENGINEERED MUTATION	UNP 4MNG
F	257	ALA	-	EXPRESSION TAG	UNP 4MNG

*Continued on next page...*

Continued from previous page...

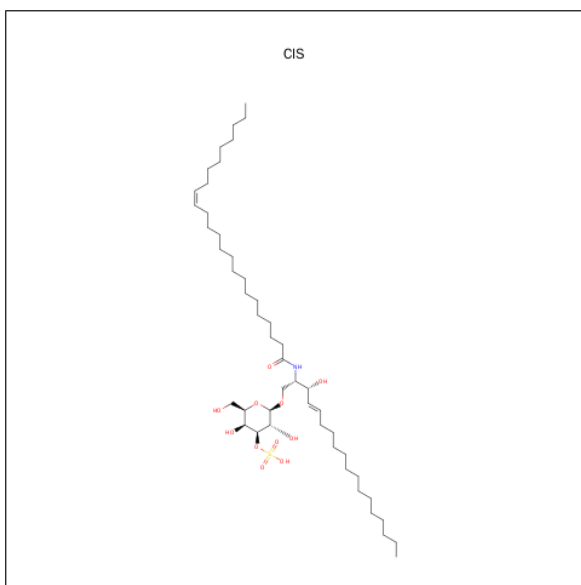
Chain	Residue	Modelled	Actual	Comment	Reference
F	258	ALA	-	EXPRESSION TAG	UNP 4MNG
F	259	SER	-	EXPRESSION TAG	UNP 4MNG
F	260	GLY	-	EXPRESSION TAG	UNP 4MNG
F	261	ALA	-	EXPRESSION TAG	UNP 4MNG
F	262	ASP	-	EXPRESSION TAG	UNP 4MNG

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 5 is (15Z)-N-((1S,2R,3E)-2-HYDROXY-1-{[(3-O-SULFO-BETA-D-GALACTOPYRANOSYL)OXY]METHYL}HEPTADEC-3-ENYL)TETRACOS-15-ENAMIDE (three-letter code: CIS) (formula:  $C_{48}H_{91}NO_{11}S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	S	0	0
			61	48	1	11	1		
5	C	1	Total	C	N	O	S	0	0
			61	48	1	11	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	7	Total	O	0	0
			7	7		
6	C	4	Total	O	0	0
			4	4		
6	E	2	Total	O	0	0
			2	2		
6	F	5	Total	O	0	0
			5	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: Beta-2-microglobulin

Chain B: 




- Molecule 1: Beta-2-microglobulin

Chain D: 




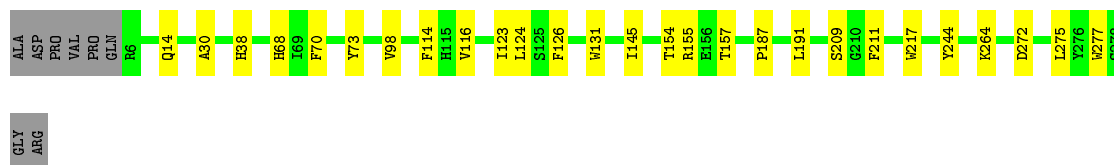
- Molecule 2: Antigen-presenting glycoprotein CD1d, Cd1d1 protein

Chain A: 




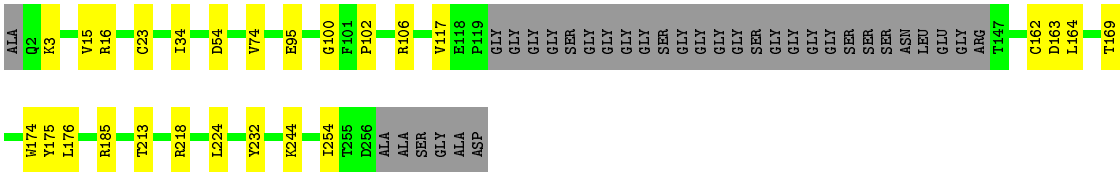
- Molecule 2: Antigen-presenting glycoprotein CD1d, Cd1d1 protein

Chain C: 



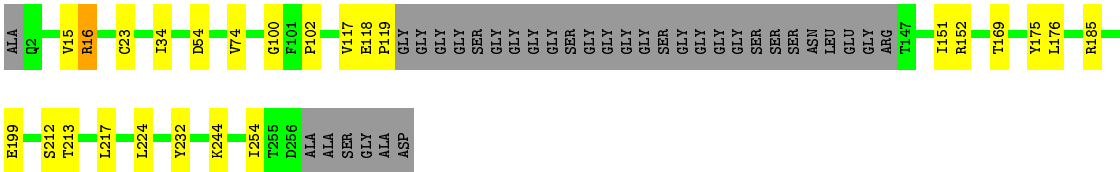
- Molecule 3: TRA@ protein, Ti antigen CD3-associated protein gamma chain V-J-C region

Chain E: 



- Molecule 3: TRA@ protein, Ti antigen CD3-associated protein gamma chain V-J-C region

Chain F:  77% 9% 13%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	229.33 Å   229.33 Å   52.65 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	49.65 – 3.01 49.65 – 3.01	Depositor EDS
% Data completeness (in resolution range)	99.0 (49.65-3.01) 99.0 (49.65-3.01)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.38 (at 3.01 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, $R_{free}$	0.178 , 0.207 0.184 , 0.209	Depositor DCC
$R_{free}$ test set	3093 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	48.1	Xtriage
Anisotropy	0.394	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 15.8	EDS
Estimated twinning fraction	0.458 for -h,-k,l 0.031 for h,-h-k,-l 0.030 for -k,-h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 61091 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9775	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.92% 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: CIS, 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	B	0.21	0/824	0.40	0/1123
1	D	0.21	0/832	0.39	0/1131
2	A	0.21	0/2242	0.38	0/3054
2	C	0.21	0/2242	0.38	0/3053
3	E	0.22	0/1839	0.40	0/2493
3	F	0.21	0/1840	0.39	0/2495
All	All	0.21	0/9819	0.39	0/13349

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	798	0	751	2	0
1	D	806	0	773	4	0
2	A	2177	0	2058	16	0
2	C	2177	0	2059	19	0
3	E	1796	0	1745	15	0
3	F	1797	0	1741	14	0
4	A	42	0	39	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	42	0	39	1	0
5	A	61	0	91	2	0
5	C	61	0	91	7	0
6	A	7	0	0	0	0
6	C	4	0	0	0	0
6	E	2	0	0	0	0
6	F	5	0	0	0	0
All	All	9775	0	9387	69	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 4.

All (69) 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)
3:F:169:THR:HB	3:F:213:THR:HB	1.76	0.68
2:A:155:ARG:NH1	3:E:54:ASP:OD2	2.28	0.67
2:C:217:TRP:HB3	2:C:264:LYS:HB2	1.79	0.64
3:E:169:THR:HB	3:E:213:THR:HB	1.79	0.62
2:C:30:ALA:HB3	2:C:38:HIS:HB2	1.81	0.62
2:A:14:GLN:HB3	2:A:98:VAL:HB	1.82	0.62
2:A:217:TRP:HB3	2:A:264:LYS:HB2	1.82	0.61
2:C:264:LYS:NZ	2:C:272:ASP:OD2	2.34	0.60
2:A:30:ALA:HB3	2:A:38:HIS:HB2	1.82	0.59
1:D:83:LYS:HG2	1:D:90:PRO:HB3	1.84	0.59
2:C:14:GLN:HB3	2:C:98:VAL:HB	1.84	0.58
3:E:3:LYS:NZ	3:E:95:GLU:OE1	2.36	0.58
2:C:14:GLN:HB2	5:C:304:CIS:H292	1.86	0.58
3:F:151:ILE:O	3:F:152:ARG:NH1	2.35	0.58
2:C:116:VAL:HG11	5:C:304:CIS:H112	1.86	0.58
2:C:187:PRO:HB3	2:C:211:PHE:HB3	1.85	0.58
2:C:154:THR:HA	5:C:304:CIS:H202	1.86	0.57
2:C:155:ARG:NH1	3:F:54:ASP:OD2	2.38	0.57
2:C:30:ALA:HB2	5:C:304:CIS:H352	1.86	0.57
4:C:302:NAG:H83	4:C:302:NAG:H3	1.86	0.56
2:A:187:PRO:HB3	2:A:211:PHE:HB3	1.87	0.56
2:A:222:ARG:HH22	2:A:255:GLU:HG2	1.71	0.55
2:C:70:PHE:HE1	5:C:304:CIS:H261	1.73	0.54
1:D:36:GLU:HB3	1:D:83:LYS:HB2	1.91	0.53
2:C:191:LEU:HD13	2:C:275:LEU:HD23	1.90	0.53
3:F:224:LEU:HB3	3:F:254:ILE:HD13	1.91	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:122:ASP:HB3	2:A:134:THR:HG21	1.92	0.52
2:A:209:SER:HB2	2:A:244:TYR:HD1	1.76	0.51
2:C:114:PHE:HB3	2:C:126:PHE:HB3	1.93	0.50
1:B:37:ILE:HG12	1:B:82:VAL:HG22	1.93	0.50
2:A:114:PHE:HB3	2:A:126:PHE:HB3	1.93	0.50
1:D:37:ILE:HG12	1:D:82:VAL:HG22	1.93	0.49
3:E:224:LEU:HB3	3:E:254:ILE:HD13	1.94	0.49
2:C:209:SER:HB2	2:C:244:TYR:HD1	1.78	0.47
3:F:34:ILE:HG21	3:F:74:VAL:HG11	1.97	0.47
2:C:73:TYR:CD1	5:C:304:CIS:H232	2.51	0.46
3:E:175:TYR:OH	3:E:244:LYS:NZ	2.48	0.46
2:C:123:ILE:HG13	2:C:124:LEU:H	1.80	0.46
2:A:123:ILE:HG13	2:A:124:LEU:H	1.80	0.46
3:E:100:GLY:O	3:E:102:PRO:HD3	2.16	0.46
3:F:15:VAL:O	3:F:16:ARG:HB2	2.16	0.45
1:B:79:ALA:HB2	1:B:94:TYR:CD2	2.52	0.45
3:F:100:GLY:O	3:F:102:PRO:HD3	2.17	0.45
3:F:212:SER:O	3:F:217:LEU:HB2	2.16	0.45
3:E:15:VAL:HG22	3:E:117:VAL:HG12	1.99	0.45
3:E:15:VAL:O	3:E:16:ARG:HB2	2.18	0.44
3:F:185:ARG:O	3:F:199:GLU:HG3	2.18	0.44
2:A:131:TRP:HB2	2:A:145:ILE:HG23	2.00	0.44
3:F:15:VAL:HG22	3:F:117:VAL:HG12	1.99	0.44
2:C:157:THR:HB	5:C:304:CIS:H201	1.98	0.44
3:F:118:GLU:HA	3:F:119:PRO:HD3	1.88	0.44
1:D:79:ALA:HB2	1:D:94:TYR:CD2	2.53	0.43
3:F:175:TYR:OH	3:F:244:LYS:NZ	2.50	0.43
2:C:191:LEU:HD23	2:C:277:TRP:HE3	1.83	0.43
3:E:162:CYS:HB2	3:E:174:TRP:CH2	2.54	0.43
3:E:163:ASP:OD1	3:E:163:ASP:N	2.50	0.43
2:A:118:PHE:HB2	2:A:123:ILE:HD13	2.00	0.42
2:C:131:TRP:HB2	2:C:145:ILE:HG23	2.01	0.42
3:E:176:LEU:HD13	3:E:232:TYR:CZ	2.54	0.42
2:A:263:VAL:HB	2:A:273:ILE:HB	2.01	0.42
2:A:116:VAL:HG11	5:A:304:CIS:H112	2.02	0.41
3:E:23:CYS:HB3	3:E:74:VAL:O	2.20	0.41
3:E:34:ILE:HG21	3:E:74:VAL:HG11	2.00	0.41
2:A:14:GLN:HB2	5:A:304:CIS:H292	2.02	0.41
3:F:23:CYS:HB3	3:F:74:VAL:O	2.20	0.41
3:F:176:LEU:HD13	3:F:232:TYR:CZ	2.56	0.41
3:E:164:LEU:HD12	3:E:164:LEU:HA	1.92	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:10:LEU:O	2:A:101:GLY:HA3	2.21	0.41
3:E:106:ARG:HB3	3:E:185:ARG:CZ	2.51	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	B	97/99 (98%)	96 (99%)	1 (1%)	0	100	100
1	D	97/99 (98%)	96 (99%)	1 (1%)	0	100	100
2	A	271/281 (96%)	268 (99%)	3 (1%)	0	100	100
2	C	271/281 (96%)	268 (99%)	3 (1%)	0	100	100
3	E	224/262 (86%)	218 (97%)	6 (3%)	0	100	100
3	F	224/262 (86%)	217 (97%)	7 (3%)	0	100	100
All	All	1184/1284 (92%)	1163 (98%)	21 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	B	88/93 (95%)	87 (99%)	1 (1%)	80	94

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	90/93 (97%)	89 (99%)	1 (1%)	80	94
2	A	231/245 (94%)	230 (100%)	1 (0%)	93	98
2	C	233/245 (95%)	232 (100%)	1 (0%)	93	98
3	E	189/211 (90%)	188 (100%)	1 (0%)	92	98
3	F	190/211 (90%)	189 (100%)	1 (0%)	92	98
All	All	1021/1098 (93%)	1015 (99%)	6 (1%)	90	97

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	70	PHE
2	A	68	HIS
1	D	70	PHE
2	C	68	HIS
3	E	218	ARG
3	F	16	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

8 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	NAG	A	301	2	14,14,15	1.19	2 (14%)	15,19,21	1.58	3 (20%)
4	NAG	A	302	2	14,14,15	0.26	0	15,19,21	0.44	0
4	NAG	A	303	2	14,14,15	0.44	0	15,19,21	0.23	0
5	CIS	A	304	-	60,61,61	0.54	1 (1%)	66,72,72	0.70	0
4	NAG	C	301	2	14,14,15	1.29	1 (7%)	15,19,21	1.89	6 (40%)
4	NAG	C	302	2	14,14,15	0.33	0	15,19,21	1.41	1 (6%)
4	NAG	C	303	2	14,14,15	0.47	0	15,19,21	0.32	0
5	CIS	C	304	-	60,61,61	0.54	1 (1%)	66,72,72	0.72	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	301	2	-	0/6/23/26	0/1/1/1
4	NAG	A	302	2	-	0/6/23/26	0/1/1/1
4	NAG	A	303	2	-	0/6/23/26	0/1/1/1
5	CIS	A	304	-	-	0/58/78/78	0/1/1/1
4	NAG	C	301	2	-	0/6/23/26	0/1/1/1
4	NAG	C	302	2	-	0/6/23/26	0/1/1/1
4	NAG	C	303	2	-	0/6/23/26	0/1/1/1
5	CIS	C	304	-	-	0/58/78/78	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	301	NAG	O5-C1	-2.13	1.40	1.43
4	A	301	NAG	O5-C5	-2.06	1.38	1.43
4	A	301	NAG	O5-C1	-2.02	1.40	1.43
5	C	304	CIS	O10-S	2.59	1.63	1.50
5	A	304	CIS	O10-S	2.60	1.64	1.50

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
4	C	301	NAG	C2-N2-C7	-4.29	117.53	123.04
4	A	301	NAG	C2-N2-C7	-3.66	118.33	123.04
4	A	301	NAG	O3-C3-C2	-2.75	103.67	109.11
4	C	301	NAG	O6-C6-C5	-2.74	102.29	111.33
4	C	301	NAG	C6-C5-C4	-2.60	106.60	113.02
4	C	301	NAG	O3-C3-C2	-2.44	104.28	109.11
4	A	301	NAG	O6-C6-C5	-2.43	103.31	111.33
4	C	301	NAG	C3-C4-C5	-2.39	106.03	110.20
4	C	301	NAG	C3-C2-N2	-2.12	105.49	110.56
4	C	302	NAG	C2-N2-C7	4.84	129.26	123.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	304	CIS	2	0
4	C	302	NAG	1	0
5	C	304	CIS	7	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 [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	B	99/99 (100%)	-0.20	0 100 100	30, 57, 80, 100	0
1	D	99/99 (100%)	-0.20	0 100 100	30, 59, 81, 102	0
2	A	273/281 (97%)	-0.15	0 100 100	24, 44, 87, 117	0
2	C	273/281 (97%)	-0.15	0 100 100	26, 44, 86, 117	0
3	E	228/262 (87%)	-0.13	0 100 100	25, 38, 80, 125	0
3	F	228/262 (87%)	-0.17	0 100 100	25, 38, 83, 118	0
All	All	1200/1284 (93%)	-0.16	0 100 100	24, 43, 85, 125	0

There are no RSRZ outliers to report.

### 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
4	NAG	A	303	14/15	0.65	0.28	4.35	34,102,118,122	0
5	CIS	A	304	61/61	0.98	0.33	2.99	22,40,57,70	0
4	NAG	C	303	14/15	0.67	0.29	2.73	48,104,117,119	0
5	CIS	C	304	61/61	0.98	0.30	2.13	27,37,56,59	0
4	NAG	A	301	14/15	0.90	0.15	-	79,84,95,97	0
4	NAG	C	301	14/15	0.93	0.15	-	50,71,92,98	0
4	NAG	A	302	14/15	0.93	0.19	-	29,41,57,66	0
4	NAG	C	302	14/15	0.92	0.19	-	33,58,67,88	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.