



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 06:55 am BST

PDB ID : 3TBS  
Title : CRYSTAL STRUCTURE OF THE MURINE CLASS I MAJOR HISTOCOMPATIBILITY COMPLEX H-2DB IN COMPLEX WITH LCMV-DERIVED GP33 ALTERED PEPTIDE ligand (V3P,Y4A)  
Authors : Duru, A.D.; Allerbring, E.B.; Uchtenhagen, H.; Mazumdar, P.A.; Badia-Martinez, D.; Madhurantakam, C.; Sandalova, T.; Nygren, P.; Achour, A.  
Deposited on : 2011-08-08  
Resolution : 2.49 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

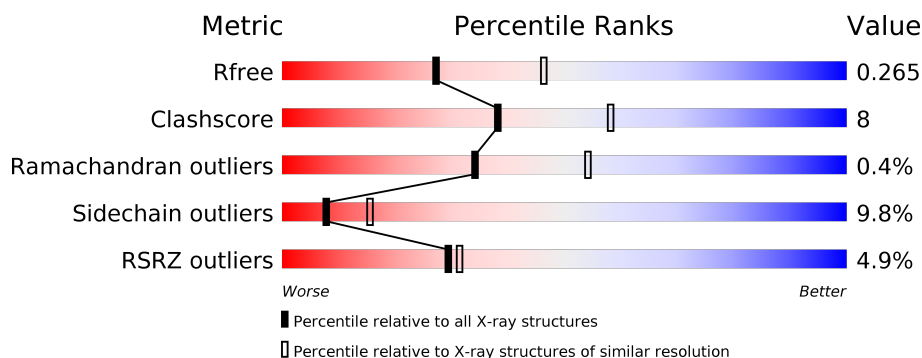
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.49 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))
$R_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 respectively. 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	276	<div> <div>7%</div> <div>78%</div> <div>17%</div> <div>••</div> </div>
1	D	276	<div> <div>7%</div> <div>76%</div> <div>19%</div> <div>••</div> </div>
2	B	99	<div> <div>%</div> <div>86%</div> <div>13%</div> <div>•</div> </div>
2	E	99	<div> <div>80%</div> <div>20%</div> </div>
3	C	9	<div> <div>89%</div> <div>11%</div> </div>
3	F	9	<div> <div>89%</div> <div>11%</div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6389 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 H-2 class I histocompatibility antigen, D-B alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	272	Total	C	N	O	S	0	0	0
			2229	1409	394	417	9			
1	D	271	Total	C	N	O	S	0	0	0
			2217	1402	392	414	9			

- 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
			821	524	138	152	7			
2	E	99	Total	C	N	O	S	0	0	0
			821	524	138	152	7			

- Molecule 3 is a protein called GLYCOPROTEIN G1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	S	0	0	0
			66	42	11	12	1			
3	F	9	Total	C	N	O	S	0	0	0
			66	42	11	12	1			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
C	4	ALA	TYR	ENGINEERED MUTATION	UNP P07399
C	9	MET	CYS	ENGINEERED MUTATION	UNP P07399
F	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
F	4	ALA	TYR	ENGINEERED MUTATION	UNP P07399
F	9	MET	CYS	ENGINEERED MUTATION	UNP P07399

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	D	1	Total	C	O	0	0
			6	3	3		

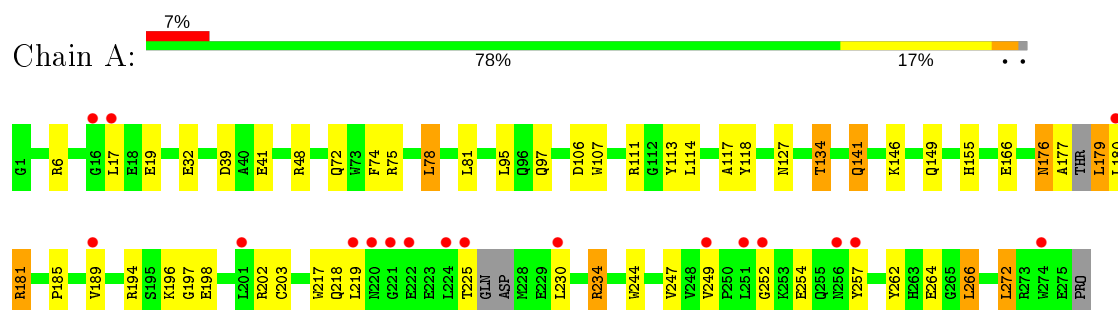
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	46	Total 46	O 46	0	0
6	B	27	Total 27	O 27	0	0
6	C	5	Total 5	O 5	0	0
6	D	41	Total 41	O 41	0	0
6	E	25	Total 25	O 25	0	0
6	F	3	Total 3	O 3	0	0

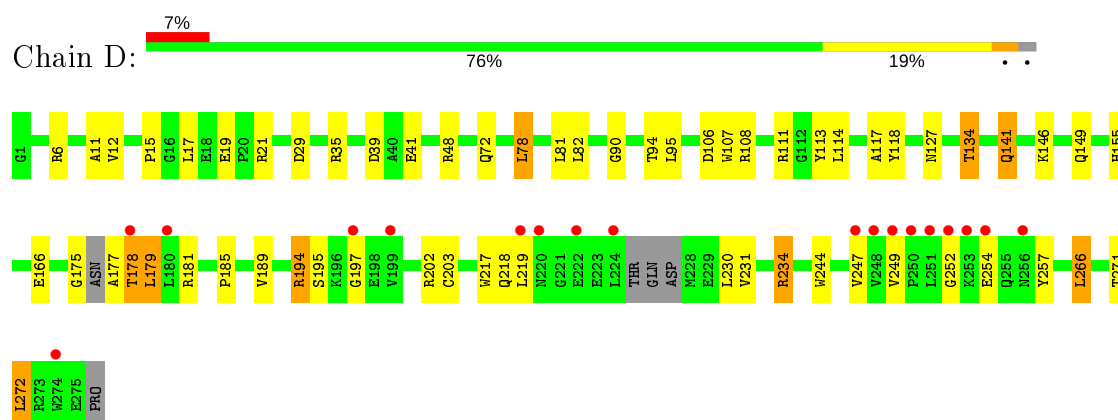
### 3 Residue-property plots [i](#)

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

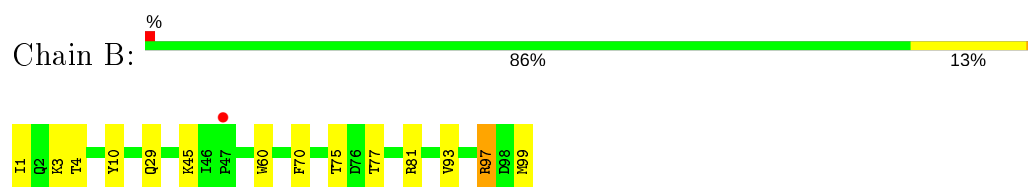
- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain



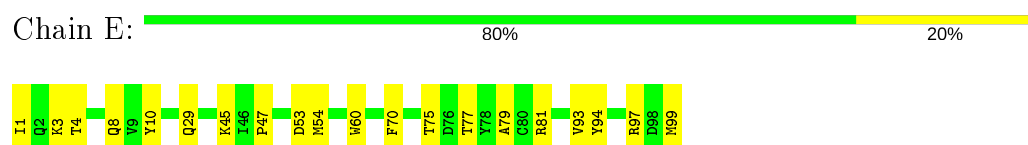
- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain




- Molecule 2: Beta-2-microglobulin



- Molecule 2: Beta-2-microglobulin




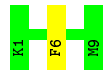
- Molecule 3: GLYCOPROTEIN G1

Chain C:  89% 11%



- Molecule 3: GLYCOPROTEIN G1

Chain F:  89% 11%



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.84Å 125.12Å 101.32Å 90.00° 128.15° 90.00°	Depositor
Resolution (Å)	49.21 – 2.49 49.21 – 2.49	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.21-2.49) 99.6 (49.21-2.49)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.04 (at 2.48Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.230 , 0.270 0.233 , 0.265	Depositor DCC
$R_{free}$ test set	2006 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.3	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 49.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.027 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6389	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.39% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: GOL, SO4

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	0.60	1/2293 (0.0%)	0.69	1/3111 (0.0%)
1	D	0.61	1/2281 (0.0%)	0.70	3/3095 (0.1%)
2	B	0.66	0/847	0.66	0/1148
2	E	0.62	0/847	0.64	0/1148
3	C	0.53	0/67	0.63	0/88
3	F	0.68	0/67	0.90	0/88
All	All	0.62	2/6402 (0.0%)	0.69	4/8678 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	181	ARG	CG-CD	-10.95	1.24	1.51
1	A	181	ARG	CG-CD	-7.84	1.32	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	181	ARG	CB-CG-CD	6.60	128.76	111.60
1	D	181	ARG	CB-CG-CD	5.90	126.93	111.60
1	D	181	ARG	NE-CZ-NH1	-5.86	117.37	120.30
1	D	181	ARG	CG-CD-NE	5.26	122.86	111.80

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	A	2229	0	2097	41	0
1	D	2217	0	2079	44	0
2	B	821	0	796	9	0
2	E	821	0	796	13	0
3	C	66	0	68	1	0
3	F	66	0	68	1	0
4	A	5	0	0	0	0
4	D	5	0	0	1	0
5	A	6	0	8	1	0
5	D	6	0	8	0	0
6	A	46	0	0	2	0
6	B	27	0	0	5	0
6	C	5	0	0	0	0
6	D	41	0	0	1	0
6	E	25	0	0	0	0
6	F	3	0	0	0	0
All	All	6389	0	5920	92	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 (92) 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:179:LEU:HD23	1:A:179:LEU:C	1.62	1.16
1:A:141:GLN:HA	1:A:141:GLN:HE21	1.08	1.16
1:A:179:LEU:C	1:A:179:LEU:CD2	2.30	0.95
1:A:179:LEU:HD23	1:A:179:LEU:O	1.65	0.93
5:A:340:GOL:HO1	5:A:340:GOL:HO3	1.18	0.90
1:D:127:ASN:OD1	1:D:134:THR:HG22	1.75	0.86
1:A:141:GLN:HA	1:A:141:GLN:NE2	1.90	0.85
1:A:141:GLN:CA	1:A:141:GLN:HE21	1.88	0.79
2:B:99:MET:CE	6:B:114:HOH:O	2.32	0.78
1:A:176:ASN:O	1:A:177:ALA:HB3	1.84	0.77
1:A:176:ASN:O	1:A:177:ALA:CB	2.30	0.77
1:A:176:ASN:HD22	1:A:176:ASN:N	1.86	0.74
2:B:97:ARG:HG2	6:B:115:HOH:O	1.89	0.71
1:A:127:ASN:OD1	1:A:134:THR:HG22	1.91	0.70
1:D:155:HIS:HB3	3:F:6:PHE:CZ	2.28	0.68

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:179:LEU:HD23	1:A:180:LEU:N	2.10	0.66
1:D:35:ARG:NH2	2:E:54:MET:O	2.29	0.65
1:A:249:VAL:HG22	1:A:257:TYR:CE2	2.33	0.63
1:D:127:ASN:OD1	1:D:134:THR:CG2	2.47	0.63
1:A:155:HIS:HB3	3:C:6:PHE:CZ	2.34	0.63
1:A:176:ASN:N	1:A:176:ASN:ND2	2.48	0.62
2:B:99:MET:HE1	6:B:114:HOH:O	1.98	0.61
1:D:234:ARG:HD3	2:E:10:TYR:CE2	2.36	0.61
1:D:141:GLN:NE2	1:D:141:GLN:HA	2.16	0.61
1:D:234:ARG:HD3	2:E:10:TYR:CZ	2.36	0.61
1:A:179:LEU:CD2	1:A:180:LEU:N	2.64	0.58
1:D:141:GLN:CA	1:D:141:GLN:HE21	2.16	0.58
1:A:252:GLY:N	1:A:254:GLU:OE1	2.37	0.58
1:A:177:ALA:C	1:A:179:LEU:N	2.56	0.58
1:A:117:ALA:HB2	2:B:60:TRP:CE2	2.39	0.57
1:D:90:GLY:N	4:D:339:SO4:O3	2.28	0.57
1:D:249:VAL:HG22	1:D:257:TYR:CE2	2.39	0.57
1:A:234:ARG:HD3	2:B:10:TYR:CE2	2.41	0.56
1:D:141:GLN:HE21	1:D:141:GLN:HA	1.71	0.55
1:D:252:GLY:N	1:D:254:GLU:OE1	2.40	0.55
1:A:234:ARG:HD3	2:B:10:TYR:CZ	2.44	0.53
1:D:175:GLY:O	1:D:177:ALA:HB3	2.10	0.52
1:D:234:ARG:NH2	2:E:99:MET:O	2.33	0.51
1:A:19:GLU:HA	1:A:19:GLU:OE1	2.10	0.51
1:D:141:GLN:NE2	1:D:141:GLN:CA	2.73	0.51
1:D:48:ARG:NH2	2:E:53:ASP:OD2	2.45	0.50
1:D:185:PRO:HD2	1:D:266:LEU:HD13	1.93	0.50
1:D:35:ARG:CZ	2:E:53:ASP:CB	2.90	0.50
2:B:99:MET:HE2	6:B:114:HOH:O	2.02	0.50
1:D:117:ALA:HB2	2:E:60:TRP:CE2	2.48	0.49
1:D:29:ASP:HB2	1:D:179:LEU:HD21	1.94	0.49
1:D:219:LEU:O	1:D:219:LEU:HG	2.12	0.49
1:A:219:LEU:HG	1:A:219:LEU:O	2.13	0.49
1:D:35:ARG:NH2	2:E:53:ASP:HB3	2.27	0.49
1:A:78:LEU:CD1	1:A:95:LEU:HB2	2.44	0.48
1:D:141:GLN:N	1:D:141:GLN:HE21	2.11	0.48
1:D:35:ARG:CZ	2:E:53:ASP:HB3	2.42	0.48
1:D:111:ARG:HG2	1:D:113:TYR:CE1	2.49	0.48
1:D:19:GLU:OE1	1:D:19:GLU:HA	2.14	0.47
1:A:185:PRO:HD2	1:A:266:LEU:HD13	1.96	0.47
1:D:179:LEU:HA	1:D:179:LEU:HD23	1.73	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:111:ARG:HG2	1:A:113:TYR:CE1	2.50	0.46
1:D:81:LEU:HD13	1:D:118:TYR:CD1	2.50	0.46
1:A:176:ASN:ND2	1:A:176:ASN:H	2.13	0.46
1:A:78:LEU:HD13	1:A:95:LEU:HB2	1.96	0.46
1:A:127:ASN:OD1	1:A:134:THR:CG2	2.62	0.46
1:D:6:ARG:NH2	1:D:113:TYR:CE1	2.84	0.45
1:A:264:GLU:HG2	6:A:353:HOH:O	2.16	0.45
1:D:178:THR:OG1	1:D:178:THR:O	2.34	0.45
1:D:189:VAL:HG23	1:D:272:LEU:HD12	1.99	0.45
1:A:75:ARG:CZ	6:A:348:HOH:O	2.64	0.44
2:B:3:LYS:HE3	6:B:125:HOH:O	2.17	0.44
1:D:271:THR:C	1:D:272:LEU:HD23	2.38	0.44
1:D:78:LEU:CD1	1:D:95:LEU:HB2	2.48	0.44
1:A:81:LEU:HD13	1:A:118:TYR:CD1	2.53	0.43
1:D:12:VAL:HG22	1:D:94:THR:HG22	1.98	0.43
1:A:202:ARG:HD3	1:A:244:TRP:CE3	2.54	0.43
1:A:106:ASP:O	1:A:107:TRP:HB2	2.18	0.43
6:D:349:HOH:O	2:E:3:LYS:HE3	2.19	0.43
1:A:196:LYS:O	1:A:198:GLU:N	2.52	0.42
1:D:179:LEU:O	1:D:179:LEU:HD22	2.19	0.42
1:A:6:ARG:NH2	1:A:113:TYR:CE1	2.88	0.42
1:A:203:CYS:HB2	1:A:217:TRP:CZ2	2.54	0.42
1:A:74:PHE:HZ	1:A:97:GLN:HE21	1.67	0.42
1:D:202:ARG:HD3	1:D:244:TRP:CE3	2.55	0.42
1:D:194:ARG:HG3	1:D:195:SER:N	2.34	0.42
1:D:249:VAL:HG22	1:D:257:TYR:CZ	2.55	0.42
1:A:189:VAL:HG23	1:A:272:LEU:HD12	2.01	0.41
1:A:262:TYR:CG	1:D:108:ARG:HD3	2.55	0.41
1:D:35:ARG:CZ	2:E:53:ASP:HB2	2.50	0.41
2:E:79:ALA:HB2	2:E:94:TYR:CD1	2.55	0.41
1:D:231:VAL:HB	2:E:8:GLN:HE22	1.86	0.41
1:A:32:GLU:OE2	1:A:48:ARG:HD2	2.20	0.41
1:A:234:ARG:NH2	2:B:99:MET:O	2.38	0.41
1:D:11:ALA:HA	1:D:21:ARG:O	2.21	0.41
1:D:106:ASP:O	1:D:107:TRP:HB2	2.21	0.40
1:D:203:CYS:HB2	1:D:217:TRP:CZ2	2.56	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	266/276 (96%)	259 (97%)	6 (2%)	1 (0%)	34	54
1	D	265/276 (96%)	257 (97%)	7 (3%)	1 (0%)	34	54
2	B	97/99 (98%)	94 (97%)	3 (3%)	0	100	100
2	E	97/99 (98%)	93 (96%)	3 (3%)	1 (1%)	15	28
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	F	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
All	All	739/768 (96%)	715 (97%)	21 (3%)	3 (0%)	34	54

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	197	GLY
1	D	197	GLY
2	E	47	PRO

### 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	229/234 (98%)	207 (90%)	22 (10%)	8	16
1	D	226/234 (97%)	204 (90%)	22 (10%)	8	16
2	B	94/94 (100%)	84 (89%)	10 (11%)	6	13
2	E	94/94 (100%)	84 (89%)	10 (11%)	6	13

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	6/6 (100%)	6 (100%)	0	100	100
3	F	6/6 (100%)	6 (100%)	0	100	100
All	All	655/668 (98%)	591 (90%)	64 (10%)	8	15

All (64) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	17	LEU
1	A	39	ASP
1	A	41	GLU
1	A	72	GLN
1	A	78	LEU
1	A	114	LEU
1	A	134	THR
1	A	141	GLN
1	A	146	LYS
1	A	149	GLN
1	A	166	GLU
1	A	176	ASN
1	A	179	LEU
1	A	181	ARG
1	A	194	ARG
1	A	218	GLN
1	A	225	THR
1	A	230	LEU
1	A	234	ARG
1	A	247	VAL
1	A	266	LEU
1	A	272	LEU
2	B	1	ILE
2	B	4	THR
2	B	29	GLN
2	B	45	LYS
2	B	70	PHE
2	B	75	THR
2	B	77	THR
2	B	81	ARG
2	B	93	VAL
2	B	97	ARG
1	D	15	PRO
1	D	17	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	39	ASP
1	D	41	GLU
1	D	72	GLN
1	D	78	LEU
1	D	82	LEU
1	D	114	LEU
1	D	134	THR
1	D	141	GLN
1	D	146	LYS
1	D	149	GLN
1	D	166	GLU
1	D	178	THR
1	D	179	LEU
1	D	194	ARG
1	D	218	GLN
1	D	230	LEU
1	D	234	ARG
1	D	247	VAL
1	D	266	LEU
1	D	272	LEU
2	E	1	ILE
2	E	4	THR
2	E	29	GLN
2	E	45	LYS
2	E	70	PHE
2	E	75	THR
2	E	77	THR
2	E	81	ARG
2	E	93	VAL
2	E	97	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (19) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	72	GLN
1	A	97	GLN
1	A	141	GLN
1	A	149	GLN
1	A	176	ASN
1	A	192	HIS
1	A	218	GLN
2	B	38	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	C	5	ASN
1	D	30	ASN
1	D	72	GLN
1	D	97	GLN
1	D	141	GLN
1	D	149	GLN
1	D	192	HIS
2	E	8	GLN
2	E	34	HIS
2	E	38	GLN
3	F	5	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](#)

4 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	SO4	A	339	-	4,4,4	0.14	0	6,6,6	0.41	0
4	SO4	D	339	-	4,4,4	0.16	0	6,6,6	0.17	0
5	GOL	A	340	-	5,5,5	0.39	0	5,5,5	0.23	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	GOL	D	340	-	5,5,5	0.45	0	5,5,5	0.26	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
5	GOL	A	340	-	-	4/4/4/4	-
5	GOL	D	340	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	340	GOL	O1-C1-C2-C3
5	A	340	GOL	C1-C2-C3-O3
5	D	340	GOL	C1-C2-C3-O3
5	A	340	GOL	O1-C1-C2-O2
5	A	340	GOL	O2-C2-C3-O3
5	D	340	GOL	O2-C2-C3-O3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	339	SO4	1	0
5	A	340	GOL	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	A	272/276 (98%)	0.32	18 (6%) 18 19	28, 51, 94, 110	0
1	D	271/276 (98%)	0.39	18 (6%) 18 19	30, 55, 94, 114	0
2	B	99/99 (100%)	0.01	1 (1%) 82 84	29, 43, 65, 73	0
2	E	99/99 (100%)	-0.01	0 100 100	35, 46, 60, 66	0
3	C	9/9 (100%)	-0.05	0 100 100	36, 37, 42, 44	0
3	F	9/9 (100%)	0.03	0 100 100	31, 34, 44, 49	0
All	All	759/768 (98%)	0.26	37 (4%) 29 31	28, 49, 90, 114	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	220	ASN	6.3
1	A	219	LEU	4.9
1	D	220	ASN	4.9
1	D	249	VAL	4.6
1	D	180	LEU	4.5
1	D	251	LEU	4.5
1	A	249	VAL	4.4
1	A	251	LEU	4.0
1	A	221	GLY	3.7
1	D	219	LEU	3.5
1	D	250	PRO	3.4
1	A	180	LEU	3.3
1	A	222	GLU	3.3
1	A	252	GLY	3.2
1	D	252	GLY	3.2
1	D	199	VAL	2.9
1	A	256	ASN	2.9
1	D	222	GLU	2.8
1	D	254	GLU	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	230	LEU	2.7
1	A	16	GLY	2.7
1	D	197	GLY	2.7
1	D	253	LYS	2.6
1	D	178	THR	2.6
1	A	274	TRP	2.6
1	D	256	ASN	2.6
1	D	274	TRP	2.4
1	A	189	VAL	2.4
2	B	47	PRO	2.3
1	D	248	VAL	2.2
1	A	224	LEU	2.2
1	A	17	LEU	2.2
1	A	201	LEU	2.2
1	D	247	VAL	2.1
1	A	225	THR	2.1
1	A	257	TYR	2.1
1	D	224	LEU	2.1

## 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
5	GOL	D	340	6/6	0.38	0.33	80,83,83,84	0
5	GOL	A	340	6/6	0.94	0.10	63,63,64,64	0
4	SO4	D	339	5/5	0.96	0.13	73,73,74,74	0
4	SO4	A	339	5/5	0.98	0.10	54,55,58,58	0

## 6.5 Other polymers

There are no such residues in this entry.