



# Full wwPDB X-ray Structure Validation Report

Mar 12, 2014 – 05:34 PM GMT

PDB ID : 4D8P  
Title : Structural and functional studies of the trans-encoded HLA-DQ2.3 (DQA1\*03:01/DQB1\*02:01)molecule  
Authors : Kim, C.-Y.; Hotta, K.; Mathews, I.I.; Chen, X.  
Deposited on : 2012-01-11  
Resolution : 3.05 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>

---

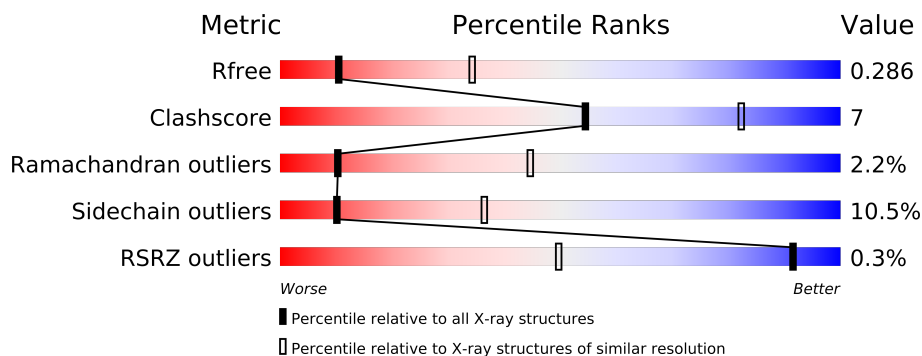
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1323  
EDS : trunk22714  
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) : trunk22714

# 1 Overall quality at a glance

The reported resolution of this entry is 3.05 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	2079 (3.12-3.00)
Clashscore	79885	2629 (3.12-3.00)
Ramachandran outliers	78287	2536 (3.12-3.00)
Sidechain outliers	78261	2539 (3.12-3.00)
RSRZ outliers	66119	2081 (3.12-3.00)

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. 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	213	
1	C	213	
2	B	250	
2	D	250	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6120 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 HLA-DQA1 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	184	Total	C	N	O	S	0	0	0
			1474	950	238	284	2			
1	C	183	Total	C	N	O	S	0	0	0
			1473	948	240	283	2			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	192	GLY	-	EXPRESSION TAG	UNP Q30063
A	193	GLY	-	EXPRESSION TAG	UNP Q30063
A	194	GLY	-	EXPRESSION TAG	UNP Q30063
A	195	SER	-	EXPRESSION TAG	UNP Q30063
A	196	GLY	-	EXPRESSION TAG	UNP Q30063
A	197	GLY	-	EXPRESSION TAG	UNP Q30063
A	198	GLY	-	EXPRESSION TAG	UNP Q30063
A	199	SER	-	EXPRESSION TAG	UNP Q30063
A	200	CYS	-	EXPRESSION TAG	UNP Q30063
A	201	SER	-	EXPRESSION TAG	UNP Q30063
A	202	ARG	-	EXPRESSION TAG	UNP Q30063
A	203	GLY	-	EXPRESSION TAG	UNP Q30063
A	204	GLY	-	EXPRESSION TAG	UNP Q30063
A	205	LEU	-	EXPRESSION TAG	UNP Q30063
A	206	GLU	-	EXPRESSION TAG	UNP Q30063
A	207	VAL	-	EXPRESSION TAG	UNP Q30063
A	208	LEU	-	EXPRESSION TAG	UNP Q30063
A	209	PHE	-	EXPRESSION TAG	UNP Q30063
A	210	GLN	-	EXPRESSION TAG	UNP Q30063
C	192	GLY	-	EXPRESSION TAG	UNP Q30063
C	193	GLY	-	EXPRESSION TAG	UNP Q30063
C	194	GLY	-	EXPRESSION TAG	UNP Q30063
C	195	SER	-	EXPRESSION TAG	UNP Q30063
C	196	GLY	-	EXPRESSION TAG	UNP Q30063
C	197	GLY	-	EXPRESSION TAG	UNP Q30063

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	198	GLY	-	EXPRESSION TAG	UNP Q30063
C	199	SER	-	EXPRESSION TAG	UNP Q30063
C	200	CYS	-	EXPRESSION TAG	UNP Q30063
C	201	SER	-	EXPRESSION TAG	UNP Q30063
C	202	ARG	-	EXPRESSION TAG	UNP Q30063
C	203	GLY	-	EXPRESSION TAG	UNP Q30063
C	204	GLY	-	EXPRESSION TAG	UNP Q30063
C	205	LEU	-	EXPRESSION TAG	UNP Q30063
C	206	GLU	-	EXPRESSION TAG	UNP Q30063
C	207	VAL	-	EXPRESSION TAG	UNP Q30063
C	208	LEU	-	EXPRESSION TAG	UNP Q30063
C	209	PHE	-	EXPRESSION TAG	UNP Q30063
C	210	GLN	-	EXPRESSION TAG	UNP Q30063

- Molecule 2 is a protein called Peptide from Gamma-gliadin, linker, MHC class II antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	193	Total	C	N	O	S	0	0	0
			1576	998	278	293	7			
2	D	193	Total	C	N	O	S	0	0	0
			1579	998	281	293	7			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-33	ARG	-	EXPRESSION TAG	UNP Q94G94
B	-32	ASP	-	EXPRESSION TAG	UNP Q94G94
B	-31	SER	-	EXPRESSION TAG	UNP Q94G94
B	-30	GLY	-	EXPRESSION TAG	UNP Q94G94
B	-26	GLU	GLN	ENGINEERED MUTATION	UNP Q94G94
B	-23	GLU	GLN	ENGINEERED MUTATION	UNP Q94G94
B	-14	GLY	-	LINKER	UNP Q94G94
B	-13	ALA	-	LINKER	UNP Q94G94
B	-12	GLY	-	LINKER	UNP Q94G94
B	-11	SER	-	LINKER	UNP Q94G94
B	-10	LEU	-	LINKER	UNP Q94G94
B	-9	VAL	-	LINKER	UNP Q94G94
B	-8	PRO	-	LINKER	UNP Q94G94
B	-7	ARG	-	LINKER	UNP Q94G94
B	-6	GLY	-	LINKER	UNP Q94G94
B	-5	SER	-	LINKER	UNP Q94G94
B	-4	GLY	-	LINKER	UNP Q94G94

*Continued on next page...*

*Continued from previous page...*

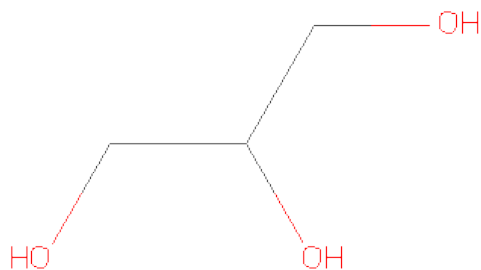
Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	GLY	-	LINKER	UNP Q94G94
B	-2	GLY	-	LINKER	UNP Q94G94
B	-1	GLY	-	LINKER	UNP Q94G94
B	0	SER	-	LINKER	UNP Q94G94
B	199	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
B	200	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
B	201	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
B	202	SER	-	EXPRESSION TAG	UNP Q5Y7D3
B	203	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
B	204	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
B	205	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
B	206	SER	-	EXPRESSION TAG	UNP Q5Y7D3
B	207	CYS	-	EXPRESSION TAG	UNP Q5Y7D3
B	208	SER	-	EXPRESSION TAG	UNP Q5Y7D3
B	209	ARG	-	EXPRESSION TAG	UNP Q5Y7D3
B	210	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
B	211	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
B	212	LEU	-	EXPRESSION TAG	UNP Q5Y7D3
B	213	GLU	-	EXPRESSION TAG	UNP Q5Y7D3
B	214	VAL	-	EXPRESSION TAG	UNP Q5Y7D3
B	215	LEU	-	EXPRESSION TAG	UNP Q5Y7D3
B	216	PHE	-	EXPRESSION TAG	UNP Q5Y7D3
B	217	GLN	-	EXPRESSION TAG	UNP Q5Y7D3
D	-33	ARG	-	EXPRESSION TAG	UNP Q94G94
D	-32	ASP	-	EXPRESSION TAG	UNP Q94G94
D	-31	SER	-	EXPRESSION TAG	UNP Q94G94
D	-30	GLY	-	EXPRESSION TAG	UNP Q94G94
D	-26	GLU	GLN	ENGINEERED MUTATION	UNP Q94G94
D	-23	GLU	GLN	ENGINEERED MUTATION	UNP Q94G94
D	-14	GLY	-	LINKER	UNP Q94G94
D	-13	ALA	-	LINKER	UNP Q94G94
D	-12	GLY	-	LINKER	UNP Q94G94
D	-11	SER	-	LINKER	UNP Q94G94
D	-10	LEU	-	LINKER	UNP Q94G94
D	-9	VAL	-	LINKER	UNP Q94G94
D	-8	PRO	-	LINKER	UNP Q94G94
D	-7	ARG	-	LINKER	UNP Q94G94
D	-6	GLY	-	LINKER	UNP Q94G94
D	-5	SER	-	LINKER	UNP Q94G94
D	-4	GLY	-	LINKER	UNP Q94G94
D	-3	GLY	-	LINKER	UNP Q94G94
D	-2	GLY	-	LINKER	UNP Q94G94

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	GLY	-	LINKER	UNP Q94G94
D	0	SER	-	LINKER	UNP Q94G94
D	199	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
D	200	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
D	201	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
D	202	SER	-	EXPRESSION TAG	UNP Q5Y7D3
D	203	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
D	204	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
D	205	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
D	206	SER	-	EXPRESSION TAG	UNP Q5Y7D3
D	207	CYS	-	EXPRESSION TAG	UNP Q5Y7D3
D	208	SER	-	EXPRESSION TAG	UNP Q5Y7D3
D	209	ARG	-	EXPRESSION TAG	UNP Q5Y7D3
D	210	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
D	211	GLY	-	EXPRESSION TAG	UNP Q5Y7D3
D	212	LEU	-	EXPRESSION TAG	UNP Q5Y7D3
D	213	GLU	-	EXPRESSION TAG	UNP Q5Y7D3
D	214	VAL	-	EXPRESSION TAG	UNP Q5Y7D3
D	215	LEU	-	EXPRESSION TAG	UNP Q5Y7D3
D	216	PHE	-	EXPRESSION TAG	UNP Q5Y7D3
D	217	GLN	-	EXPRESSION TAG	UNP Q5Y7D3

- Molecule 3 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
3	A	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is water.

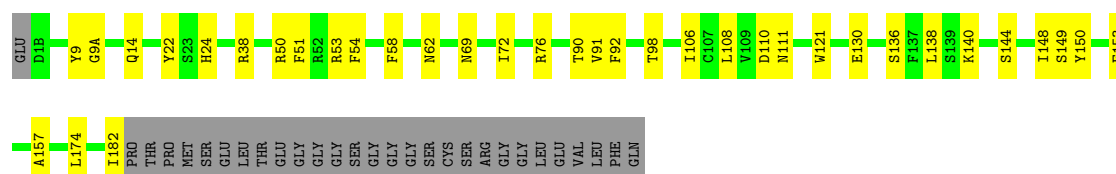
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	O	0	0
			1	1		
4	B	3	Total	O	0	0
			3	3		
4	C	1	Total	O	0	0
			1	1		
4	D	1	Total	O	0	0
			1	1		

### 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.

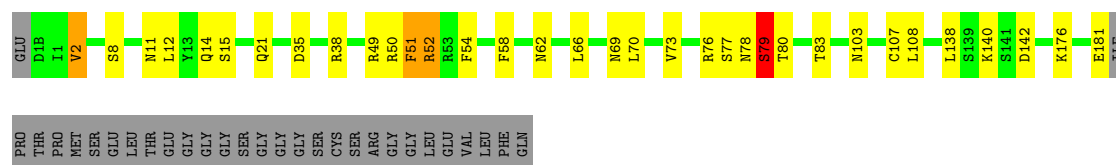
#### • Molecule 1: HLA-DQA1 protein

Chain A: 



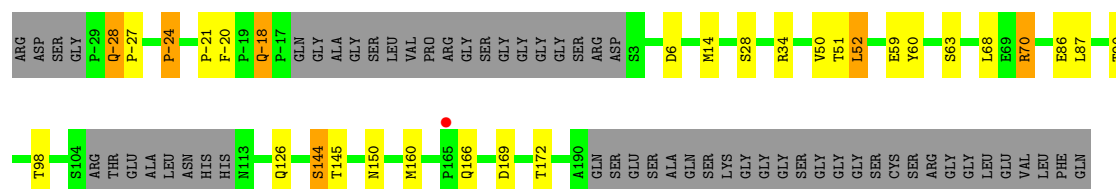
#### • Molecule 1: HLA-DQA1 protein

Chain C: 



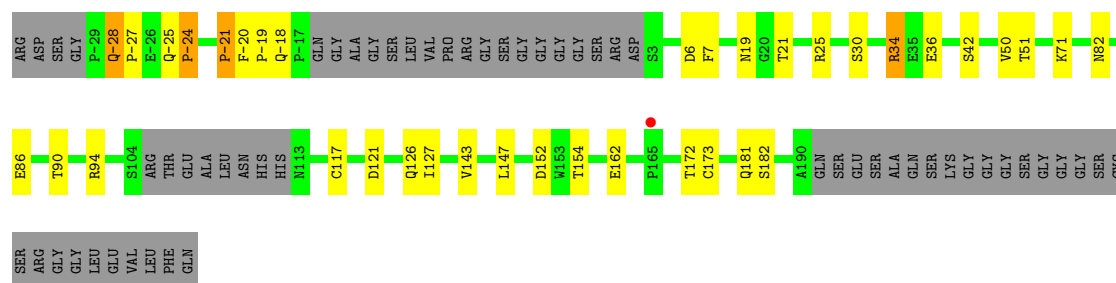
#### • Molecule 2: Peptide from Gamma-gliadin, linker, MHC class II antigen

Chain B: 



#### • Molecule 2: Peptide from Gamma-gliadin, linker, MHC class II antigen

Chain D: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.89Å 114.92Å 138.03Å 90.00° 103.36° 90.00°	Depositor
Resolution (Å)	29.31 – 3.05 29.31 – 3.05	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.31-3.05) 99.9 (29.31-3.05)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.57 (at 3.05Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.210 , 0.283 0.216 , 0.286	Depositor DCC
$R_{free}$ test set	1086 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	84.7	Xtriage
Anisotropy	0.166	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 53.0	EDS
Estimated twinning fraction	0.012 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	2 of 21712 reflections (0.009%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6120	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.01% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL

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	0.58	0/1518	0.68	0/2076
1	C	0.55	0/1517	0.67	0/2074
2	B	0.55	0/1615	0.70	1/2197 (0.0%)
2	D	0.52	0/1617	0.67	0/2198
All	All	0.55	0/6267	0.68	1/8545 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	87	LEU	CA-CB-CG	5.01	126.82	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	52	ARG	Peptide

## 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	1474	0	0	15	0
1	C	1473	0	0	11	0
2	B	1576	0	97	20	0
2	D	1579	0	88	12	0
3	A	6	0	8	0	0
3	B	6	0	8	0	0
4	A	1	0	0	0	0
4	B	3	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
All	All	6120	0	201	43	0

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 7.

All (43) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:62:ASN:OD1	2:B:-21:PRO:HD2	1.64	0.97
1:A:110:ASP:OD1	1:A:111:ASN:N	2.22	0.72
2:B:-28:GLN:HB3	2:B:-27:PRO:HD3	1.70	0.71
2:B:-20:PHE:CE1	2:B:70:ARG:NH2	2.61	0.68
1:C:50:ARG:O	1:C:52:ARG:N	2.27	0.68
2:B:-20:PHE:HE1	2:B:70:ARG:NH2	1.97	0.61
1:A:14:GLN:NE2	2:B:6:ASP:OD1	2.34	0.61
1:A:58:PHE:CE2	2:B:-24:PRO:HG2	2.39	0.57
2:D:-28:GLN:HE21	2:D:-28:GLN:HA	1.70	0.57
1:A:62:ASN:OD1	2:B:-21:PRO:CD	2.49	0.54
1:C:58:PHE:CE2	2:D:-24:PRO:HG2	2.43	0.54
2:B:60:TYR:O	2:B:63:SER:OG	2.29	0.51
1:C:70:LEU:O	1:C:73:VAL:N	2.44	0.51
1:C:77:SER:O	1:C:79:SER:N	2.44	0.51
1:C:69:ASN:OD1	2:D:-19:PRO:HA	2.10	0.50
2:D:152:ASP:OD2	2:D:154:THR:OG1	2.29	0.50
2:B:-28:GLN:NE2	2:B:-28:GLN:HA	2.27	0.49

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:-28:GLN:HE21	2:B:-28:GLN:HA	1.77	0.49
2:B:70:ARG:NH1	2:B:70:ARG:CG	2.74	0.49
1:A:121:TRP:NE1	1:A:149:SER:OG	2.45	0.49
1:A:92:PHE:O	1:A:106:ILE:N	2.46	0.48
2:B:-18:GLN:O	2:B:-18:GLN:HG2	2.11	0.48
1:A:9:TYR:OH	2:B:86:GLU:OE1	2.31	0.48
2:D:6:ASP:OD2	2:D:7:PHE:N	2.48	0.47
2:D:36:GLU:O	2:D:50:VAL:CG2	2.63	0.47
1:A:92:PHE:CD2	1:A:92:PHE:N	2.84	0.46
1:A:51:PHE:O	2:B:-28:GLN:NE2	2.49	0.46
2:D:-25:GLN:O	2:D:82:ASN:ND2	2.49	0.46
2:B:59:GLU:OE1	2:D:162:GLU:CG	2.63	0.45
1:A:150:TYR:OH	2:B:150:ASN:O	2.34	0.45
1:A:54:PHE:CE1	2:B:-24:PRO:HG3	2.51	0.45
1:C:142:ASP:O	2:D:34:ARG:NH1	2.51	0.44
1:A:106:ILE:CG2	1:A:148:ILE:CG2	2.96	0.44
2:B:144:SER:OG	2:B:145:THR:O	2.35	0.44
1:C:62:ASN:ND2	2:D:-21:PRO:HD3	2.33	0.43
1:A:69:ASN:ND2	2:B:-20:PHE:O	2.52	0.42
1:C:51:PHE:O	1:C:52:ARG:NH1	2.53	0.42
2:D:-25:GLN:HA	2:D:-24:PRO:HD3	1.83	0.41
1:C:35:ASP:OD1	1:C:38:ARG:N	2.54	0.41
2:B:51:THR:O	2:B:52:LEU:C	2.58	0.41
1:A:22:TYR:OH	1:A:24:HIS:NE2	2.54	0.41
1:C:142:ASP:N	1:C:142:ASP:OD1	2.55	0.40
1:C:54:PHE:CE1	2:D:-24:PRO:HG3	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	182/213 (85%)	161 (88%)	17 (9%)	4 (2%)	10	44
1	C	181/213 (85%)	158 (87%)	19 (10%)	4 (2%)	10	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	187/250 (75%)	161 (86%)	23 (12%)	3 (2%)	14	55
2	D	187/250 (75%)	164 (88%)	18 (10%)	5 (3%)	8	38
All	All	737/926 (80%)	644 (87%)	77 (10%)	16 (2%)	10	44

All (16) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	-24	PRO
1	C	51	PHE
1	C	79	SER
2	D	121	ASP
1	C	2	VAL
2	D	-24	PRO
2	D	19	ASN
1	A	136	SER
1	A	157	ALA
1	C	78	ASN
2	D	-27	PRO
1	A	38	ARG
2	B	169	ASP
2	D	-21	PRO
1	A	9(A)	GLY
2	B	-28	GLN

### 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	166/190 (87%)	151 (91%)	15 (9%)	14	47
1	C	166/190 (87%)	146 (88%)	20 (12%)	7	29
2	B	176/216 (82%)	161 (92%)	15 (8%)	15	51
2	D	176/216 (82%)	154 (88%)	22 (12%)	7	27
All	All	684/812 (84%)	612 (90%)	72 (10%)	10	36

All (72) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	50	ARG
1	A	53	ARG
1	A	72	ILE
1	A	76	ARG
1	A	90	THR
1	A	91	VAL
1	A	98	THR
1	A	108	LEU
1	A	130	GLU
1	A	138	LEU
1	A	140	LYS
1	A	144	SER
1	A	153	PHE
1	A	174	LEU
1	A	182	ILE
2	B	-18	GLN
2	B	14	MET
2	B	28	SER
2	B	34	ARG
2	B	50	VAL
2	B	52	LEU
2	B	68	LEU
2	B	70	ARG
2	B	90	THR
2	B	98	THR
2	B	126	GLN
2	B	144	SER
2	B	160	MET
2	B	166	GLN
2	B	172	THR
1	C	2	VAL
1	C	8	SER
1	C	11	ASN
1	C	12	LEU
1	C	14	GLN
1	C	15	SER
1	C	21	GLN
1	C	49	ARG
1	C	66	LEU
1	C	76	ARG
1	C	79	SER
1	C	80	THR
1	C	83	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	103	ASN
1	C	107	CYS
1	C	108	LEU
1	C	138	LEU
1	C	140	LYS
1	C	176	LYS
1	C	181	GLU
2	D	-28	GLN
2	D	-20	PHE
2	D	-18	GLN
2	D	21	THR
2	D	25	ARG
2	D	30	SER
2	D	34	ARG
2	D	42	SER
2	D	51	THR
2	D	71	LYS
2	D	86	GLU
2	D	90	THR
2	D	94	ARG
2	D	117	CYS
2	D	126	GLN
2	D	127	ILE
2	D	143	VAL
2	D	147	LEU
2	D	172	THR
2	D	173	CYS
2	D	181	GLN
2	D	182	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	-28	GLN
2	D	-28	GLN
2	D	-18	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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	GOL	A	301	-	5,5,5	0.34	0	5,5,5	1.13	1 (20%)
3	GOL	B	301	-	5,5,5	0.34	0	5,5,5	0.47	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
3	GOL	A	301	-	-	0/4/4/4	0/0/0/0
3	GOL	B	301	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	301	GOL	O2-C2-C1	2.03	117.86	108.44

There are no chirality outliers.

There are no torsion outliers.



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	184/213 (86%)	-0.22	0	100 100	50, 89, 112, 134	0
1	C	183/213 (85%)	-0.16	0	100 100	67, 93, 116, 125	0
2	B	193/250 (77%)	-0.15	1 (0%)	88 37	53, 94, 121, 141	0
2	D	193/250 (77%)	-0.21	1 (0%)	88 37	67, 89, 118, 131	0
All	All	753/926 (81%)	-0.18	2 (0%)	91 50	50, 92, 118, 141	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	165	PRO	2.2
2	D	165	PRO	2.1

### 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

There are no carbohydrates in this entry.

### 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
3	GOL	A	301	6/6	0.23	-	87,89,91,91	0
3	GOL	B	301	6/6	0.29	-	96,98,99,99	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.