



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 12, 2025 – 12:12 PM JST

PDB ID : 9WBD / pdb\_00009wbd  
Title : Crystal structure of HLA-B\*07:02 in complex with SPR epitope and Q04 TCR  
Authors : Ping, Y.; Daichao, W.  
Deposited on : 2025-08-13  
Resolution : 2.75 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4-5-2 with Phenix2.0  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.46

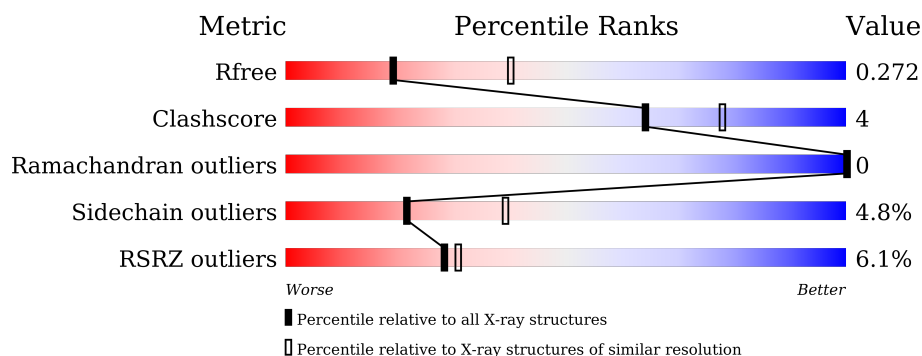
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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))
$R_{free}$	164625	1606 (2.78-2.74)
Clashscore	180529	1689 (2.78-2.74)
Ramachandran outliers	177936	1665 (2.78-2.74)
Sidechain outliers	177891	1665 (2.78-2.74)
RSRZ outliers	164620	1606 (2.78-2.74)

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 of 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	206	
1	F	206	
1	K	206	
1	P	206	
2	B	243	
2	G	243	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	L	243	
2	Q	243	
3	C	276	
3	H	276	
3	M	276	
3	R	276	
4	D	100	
4	I	100	
4	N	100	
4	S	100	
5	E	9	
5	J	9	
5	O	9	
5	T	9	

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 25365 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 T cell receptor alpha variable 25,T cell receptor alpha chain constant.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	191	Total	C	N	O	S	0	0	0
			1447	918	238	284	7			
1	F	198	Total	C	N	O	S	0	0	0
			1515	959	248	300	8			
1	K	192	Total	C	N	O	S	0	0	0
			1453	916	241	288	8			
1	P	204	Total	C	N	O	S	0	0	0
			1582	993	261	320	8			

There are 92 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP A0A0B4J276
A	93	TYR	-	linker	UNP A0A0B4J276
A	94	THR	-	linker	UNP A0A0B4J276
A	95	SER	-	linker	UNP A0A0B4J276
A	96	GLY	-	linker	UNP A0A0B4J276
A	97	THR	-	linker	UNP A0A0B4J276
A	98	TYR	-	linker	UNP A0A0B4J276
A	99	LYS	-	linker	UNP A0A0B4J276
A	100	TYR	-	linker	UNP A0A0B4J276
A	101	ILE	-	linker	UNP A0A0B4J276
A	102	PHE	-	linker	UNP A0A0B4J276
A	103	GLY	-	linker	UNP A0A0B4J276
A	104	THR	-	linker	UNP A0A0B4J276
A	105	GLY	-	linker	UNP A0A0B4J276
A	106	THR	-	linker	UNP A0A0B4J276
A	107	ARG	-	linker	UNP A0A0B4J276
A	108	LEU	-	linker	UNP A0A0B4J276
A	109	LYS	-	linker	UNP A0A0B4J276
A	110	VAL	-	linker	UNP A0A0B4J276
A	111	LEU	-	linker	UNP A0A0B4J276

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	112	ALA	-	linker	UNP A0A0B4J276
A	113	ASN	-	linker	UNP A0A0B4J276
A	160	CYS	THR	conflict	UNP P01848
F	1	MET	-	initiating methionine	UNP A0A0B4J276
F	93	TYR	-	linker	UNP A0A0B4J276
F	94	THR	-	linker	UNP A0A0B4J276
F	95	SER	-	linker	UNP A0A0B4J276
F	96	GLY	-	linker	UNP A0A0B4J276
F	97	THR	-	linker	UNP A0A0B4J276
F	98	TYR	-	linker	UNP A0A0B4J276
F	99	LYS	-	linker	UNP A0A0B4J276
F	100	TYR	-	linker	UNP A0A0B4J276
F	101	ILE	-	linker	UNP A0A0B4J276
F	102	PHE	-	linker	UNP A0A0B4J276
F	103	GLY	-	linker	UNP A0A0B4J276
F	104	THR	-	linker	UNP A0A0B4J276
F	105	GLY	-	linker	UNP A0A0B4J276
F	106	THR	-	linker	UNP A0A0B4J276
F	107	ARG	-	linker	UNP A0A0B4J276
F	108	LEU	-	linker	UNP A0A0B4J276
F	109	LYS	-	linker	UNP A0A0B4J276
F	110	VAL	-	linker	UNP A0A0B4J276
F	111	LEU	-	linker	UNP A0A0B4J276
F	112	ALA	-	linker	UNP A0A0B4J276
F	113	ASN	-	linker	UNP A0A0B4J276
F	160	CYS	THR	conflict	UNP P01848
K	1	MET	-	initiating methionine	UNP A0A0B4J276
K	93	TYR	-	linker	UNP A0A0B4J276
K	94	THR	-	linker	UNP A0A0B4J276
K	95	SER	-	linker	UNP A0A0B4J276
K	96	GLY	-	linker	UNP A0A0B4J276
K	97	THR	-	linker	UNP A0A0B4J276
K	98	TYR	-	linker	UNP A0A0B4J276
K	99	LYS	-	linker	UNP A0A0B4J276
K	100	TYR	-	linker	UNP A0A0B4J276
K	101	ILE	-	linker	UNP A0A0B4J276
K	102	PHE	-	linker	UNP A0A0B4J276
K	103	GLY	-	linker	UNP A0A0B4J276
K	104	THR	-	linker	UNP A0A0B4J276
K	105	GLY	-	linker	UNP A0A0B4J276
K	106	THR	-	linker	UNP A0A0B4J276
K	107	ARG	-	linker	UNP A0A0B4J276

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
K	108	LEU	-	linker	UNP A0A0B4J276
K	109	LYS	-	linker	UNP A0A0B4J276
K	110	VAL	-	linker	UNP A0A0B4J276
K	111	LEU	-	linker	UNP A0A0B4J276
K	112	ALA	-	linker	UNP A0A0B4J276
K	113	ASN	-	linker	UNP A0A0B4J276
K	160	CYS	THR	conflict	UNP P01848
P	1	MET	-	initiating methionine	UNP A0A0B4J276
P	93	TYR	-	linker	UNP A0A0B4J276
P	94	THR	-	linker	UNP A0A0B4J276
P	95	SER	-	linker	UNP A0A0B4J276
P	96	GLY	-	linker	UNP A0A0B4J276
P	97	THR	-	linker	UNP A0A0B4J276
P	98	TYR	-	linker	UNP A0A0B4J276
P	99	LYS	-	linker	UNP A0A0B4J276
P	100	TYR	-	linker	UNP A0A0B4J276
P	101	ILE	-	linker	UNP A0A0B4J276
P	102	PHE	-	linker	UNP A0A0B4J276
P	103	GLY	-	linker	UNP A0A0B4J276
P	104	THR	-	linker	UNP A0A0B4J276
P	105	GLY	-	linker	UNP A0A0B4J276
P	106	THR	-	linker	UNP A0A0B4J276
P	107	ARG	-	linker	UNP A0A0B4J276
P	108	LEU	-	linker	UNP A0A0B4J276
P	109	LYS	-	linker	UNP A0A0B4J276
P	110	VAL	-	linker	UNP A0A0B4J276
P	111	LEU	-	linker	UNP A0A0B4J276
P	112	ALA	-	linker	UNP A0A0B4J276
P	113	ASN	-	linker	UNP A0A0B4J276
P	160	CYS	THR	conflict	UNP P01848

- Molecule 2 is a protein called TCR beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	242	Total	C	N	O	S	0	0	0
			1877	1183	329	360	5			
2	G	241	Total	C	N	O	S	0	0	0
			1851	1170	323	353	5			
2	L	241	Total	C	N	O	S	0	0	0
			1860	1172	329	354	5			
2	Q	242	Total	C	N	O	S	0	0	0
			1867	1177	328	357	5			

- Molecule 3 is a protein called HLA class I histocompatibility antigen, B alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	271	Total	C	N	O	S	0	0	0
			2106	1320	377	403	6			
3	H	273	Total	C	N	O	S	0	0	0
			2160	1351	386	417	6			
3	M	240	Total	C	N	O	S	0	0	0
			1917	1201	342	368	6			
3	R	267	Total	C	N	O	S	0	0	0
			2067	1294	367	400	6			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	MET	-	initiating methionine	UNP P01889
H	0	MET	-	initiating methionine	UNP P01889
M	0	MET	-	initiating methionine	UNP P01889
R	0	MET	-	initiating methionine	UNP P01889

- Molecule 4 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	100	Total	C	N	O	S	0	0	0
			813	518	138	153	4			
4	I	100	Total	C	N	O	S	0	0	0
			813	517	136	156	4			
4	N	100	Total	C	N	O	S	0	0	0
			809	515	137	153	4			
4	S	100	Total	C	N	O	S	0	0	0
			813	515	135	159	4			

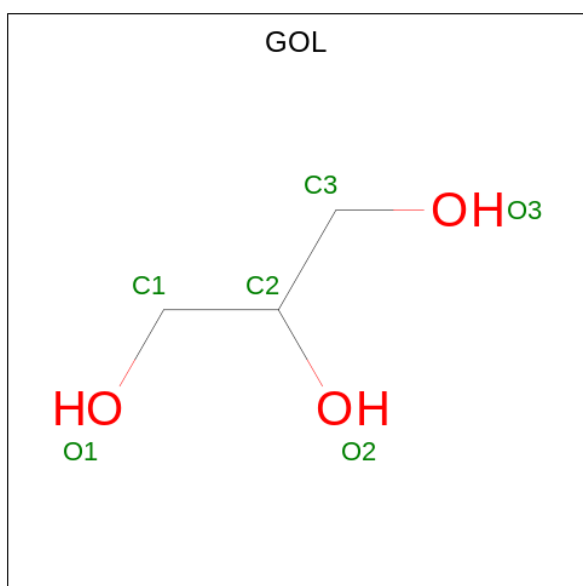
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	initiating methionine	UNP P61769
I	0	MET	-	initiating methionine	UNP P61769
N	0	MET	-	initiating methionine	UNP P61769
S	0	MET	-	initiating methionine	UNP P61769

- Molecule 5 is a protein called Nucleoprotein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	E	9	Total	C	N	O	0	0	0
			94	67	13	14			
5	J	9	Total	C	N	O	0	0	0
			94	67	13	14			
5	O	9	Total	C	N	O	0	0	0
			94	67	13	14			
5	T	9	Total	C	N	O	0	0	0
			94	67	13	14			

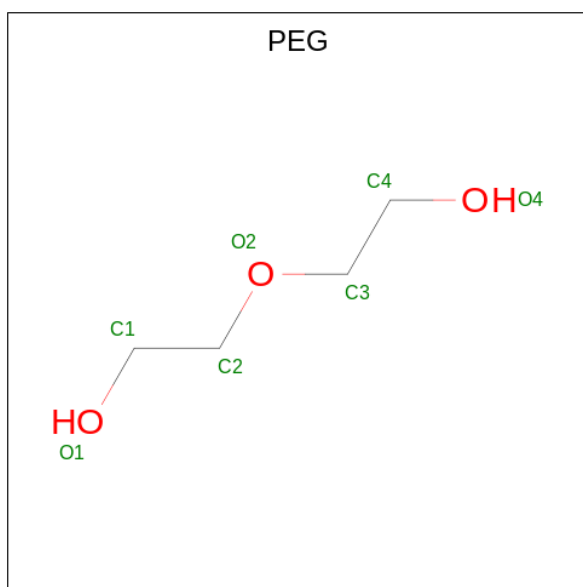
- Molecule 6 is GLYCEROL (CCD ID: GOL) (formula:  $C_3H_8O_3$ ).



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

- Molecule 7 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula:  $C_4H_{10}O_3$ ).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	G	1	Total	C	O	0	0
			7	4	3		

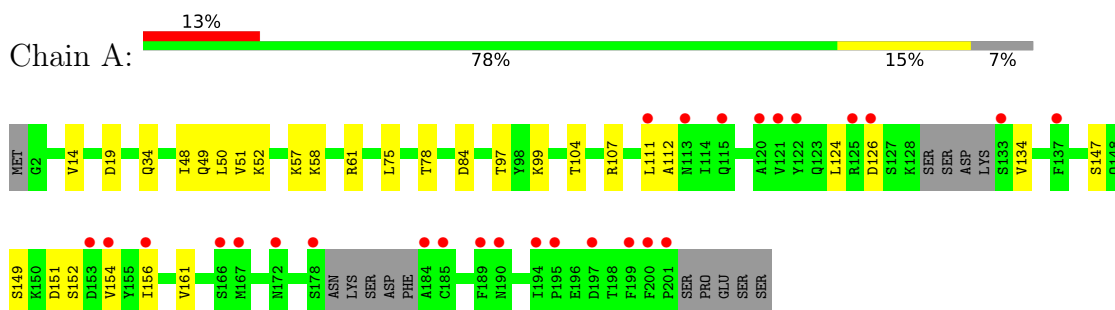
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	C	4	Total	O	0	0
			4	4		
8	H	4	Total	O	0	0
			4	4		
8	M	2	Total	O	0	0
			2	2		
8	R	4	Total	O	0	0
			4	4		

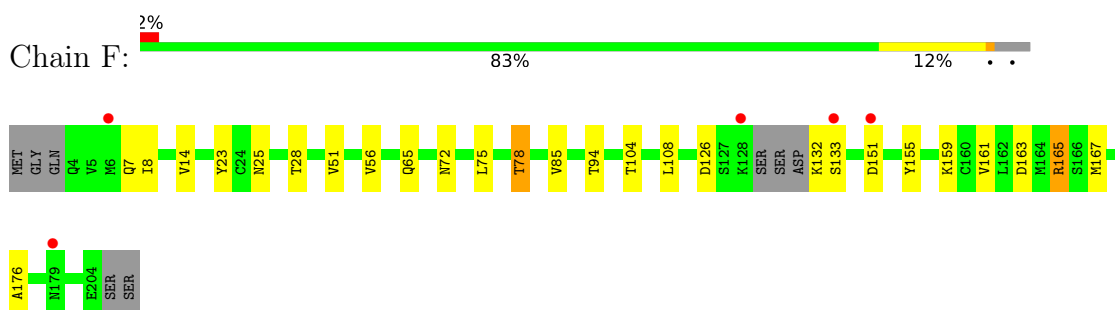
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

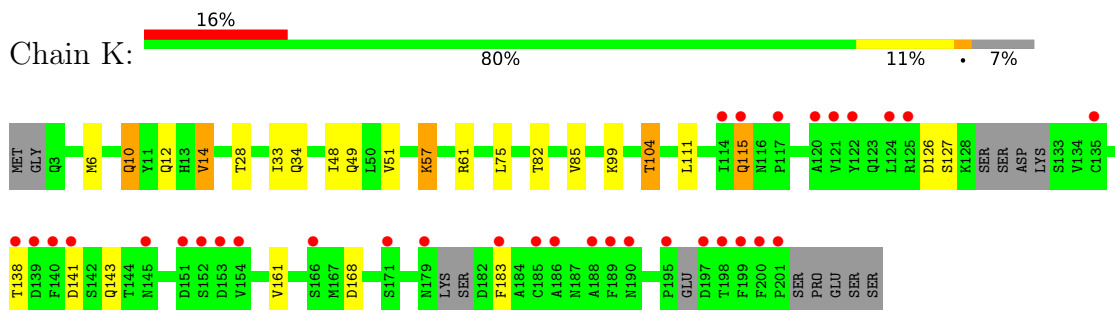
- Molecule 1: T cell receptor alpha variable 25,T cell receptor alpha chain constant



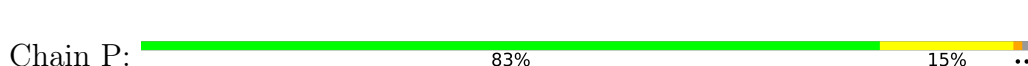
- Molecule 1: T cell receptor alpha variable 25,T cell receptor alpha chain constant

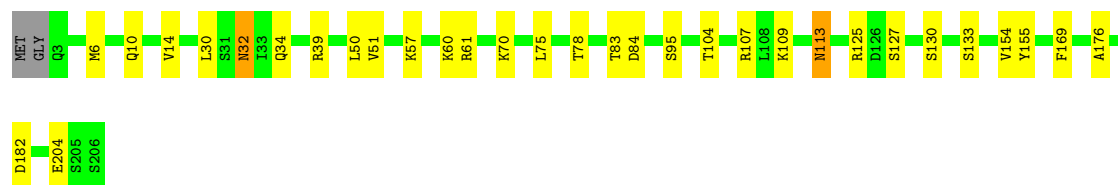


- Molecule 1: T cell receptor alpha variable 25,T cell receptor alpha chain constant

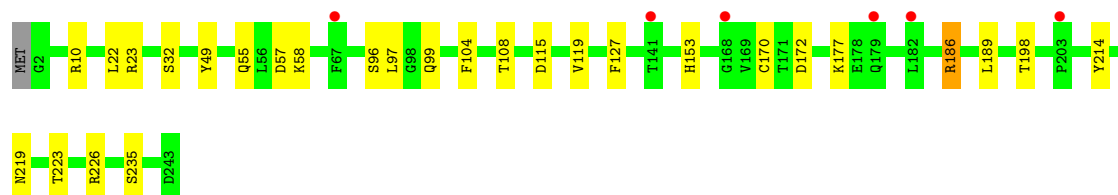
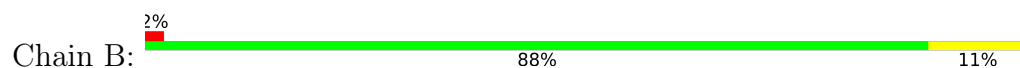


- Molecule 1: T cell receptor alpha variable 25,T cell receptor alpha chain constant

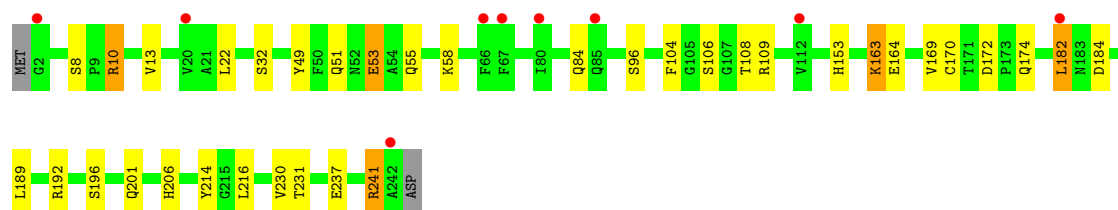
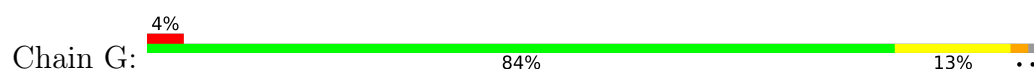




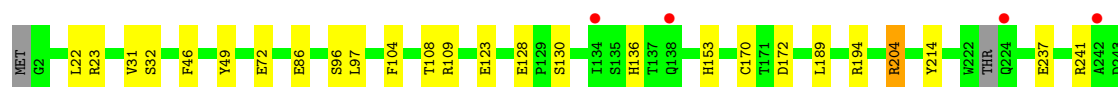
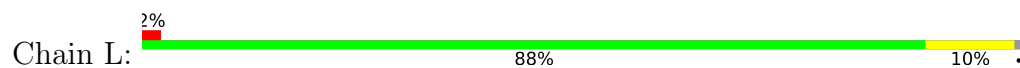
- Molecule 2: TCR beta



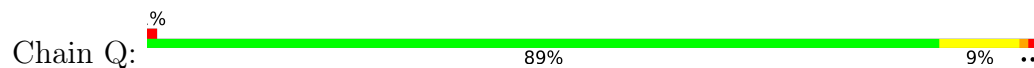
- Molecule 2: TCR beta



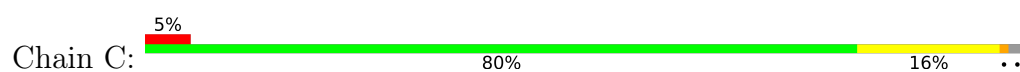
- Molecule 2: TCR beta

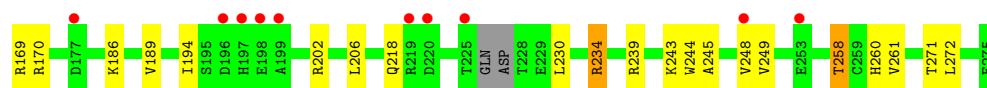


- Molecule 2: TCR beta

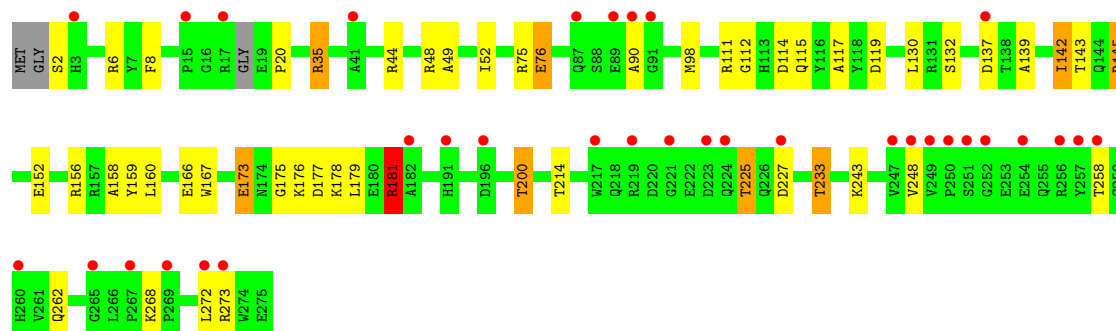
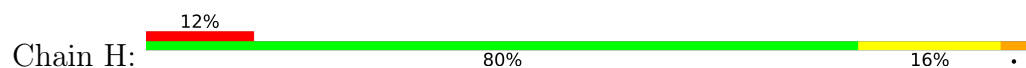


- Molecule 3: HLA class I histocompatibility antigen, B alpha chain

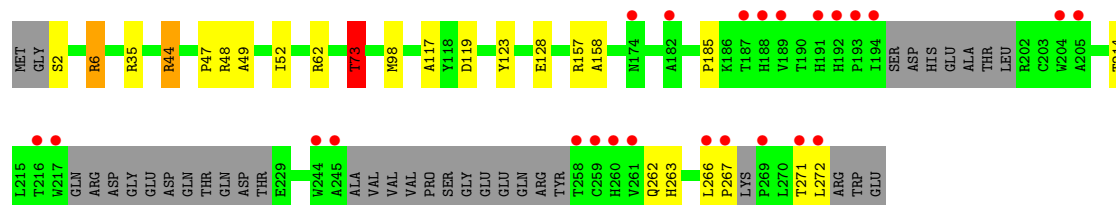
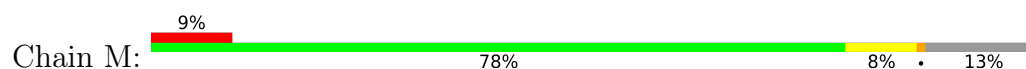




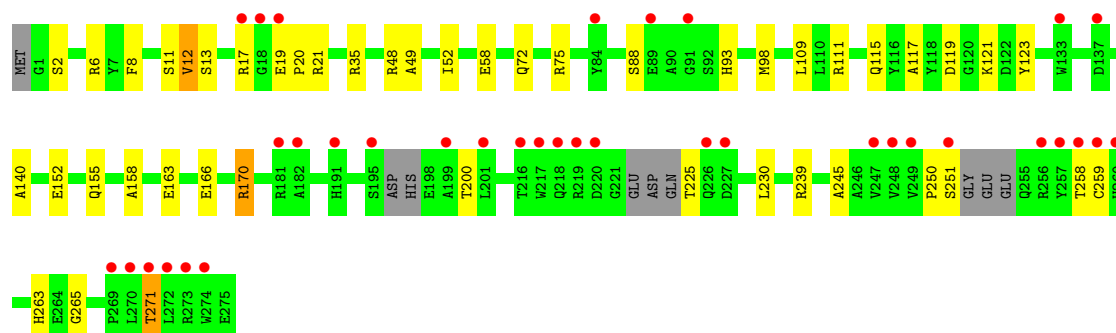
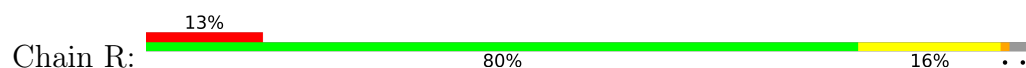
- Molecule 3: HLA class I histocompatibility antigen, B alpha chain



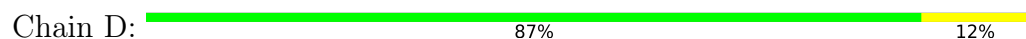
- Molecule 3: HLA class I histocompatibility antigen, B alpha chain



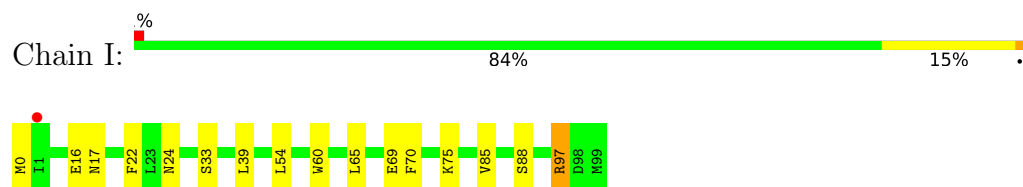
- Molecule 3: HLA class I histocompatibility antigen, B alpha chain



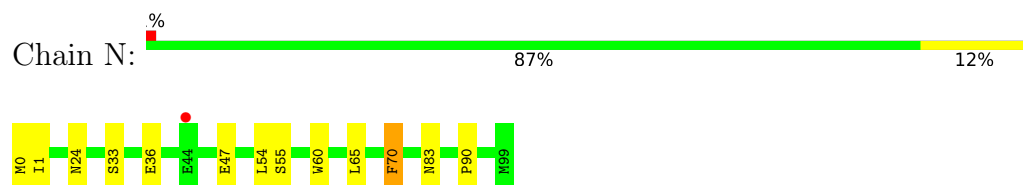
- Molecule 4: Beta-2-microglobulin



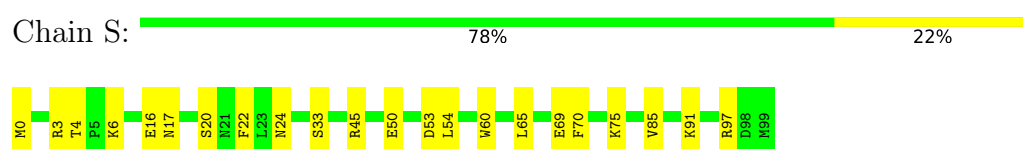
## • Molecule 4: Beta-2-microglobulin



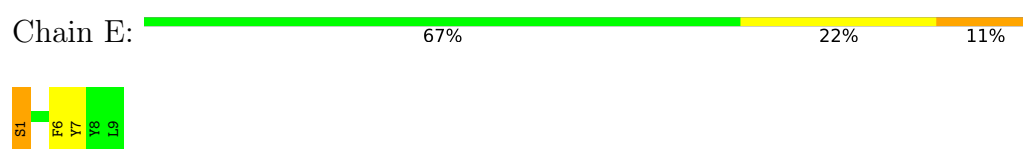
## • Molecule 4: Beta-2-microglobulin



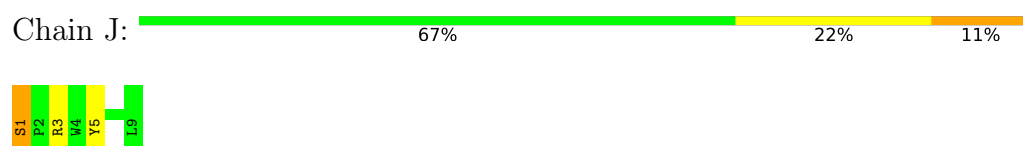
## • Molecule 4: Beta-2-microglobulin



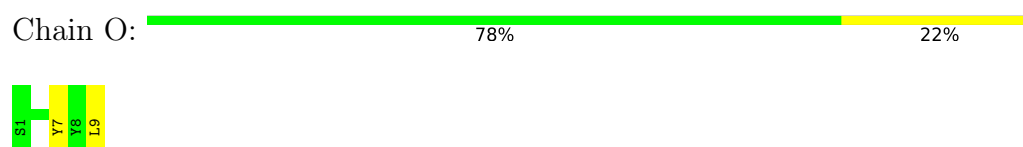
## • Molecule 5: Nucleoprotein



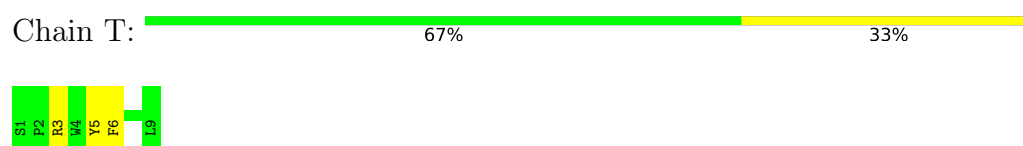
## • Molecule 5: Nucleoprotein



## • Molecule 5: Nucleoprotein



## • Molecule 5: Nucleoprotein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.79Å 102.56Å 120.70Å 84.14° 68.60° 81.10°	Depositor
Resolution (Å)	48.52 – 2.75 48.52 – 2.75	Depositor EDS
% Data completeness (in resolution range)	98.6 (48.52-2.75) 98.8 (48.52-2.75)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.80 (at 2.73Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, $R_{free}$	0.218 , 0.270 0.220 , 0.272	Depositor DCC
$R_{free}$ test set	5025 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.0	Xtriage
Anisotropy	0.055	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 44.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.006 for -h,-k,-h+l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	25365	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.02% 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, PEG

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	0/1479	1.14	4/2011 (0.2%)
1	F	0.60	0/1550	1.15	6/2109 (0.3%)
1	K	0.57	0/1483	1.11	7/2016 (0.3%)
1	P	0.61	0/1618	1.19	6/2197 (0.3%)
2	B	0.57	0/1928	1.06	5/2626 (0.2%)
2	G	0.58	0/1902	1.10	3/2595 (0.1%)
2	L	0.57	0/1910	1.09	7/2602 (0.3%)
2	Q	0.60	0/1918	1.10	5/2615 (0.2%)
3	C	0.58	0/2165	1.12	5/2956 (0.2%)
3	H	0.56	0/2220	1.11	9/3026 (0.3%)
3	M	0.56	0/1968	1.13	5/2674 (0.2%)
3	R	0.58	0/2124	1.12	5/2898 (0.2%)
4	D	0.60	0/836	1.12	3/1135 (0.3%)
4	I	0.58	0/836	1.04	2/1136 (0.2%)
4	N	0.56	0/832	0.99	1/1131 (0.1%)
4	S	0.58	0/836	1.15	5/1138 (0.4%)
5	E	0.57	0/100	0.96	0/135
5	J	0.54	0/100	1.06	0/135
5	O	0.67	0/100	1.08	0/135
5	T	0.67	0/100	1.15	1/135 (0.7%)
All	All	0.58	0/26005	1.11	79/35405 (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
1	A	0	1
1	F	0	1
1	K	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	2
2	G	0	5
2	L	0	1
2	Q	0	4
3	C	0	5
3	H	0	4
3	M	0	2
3	R	0	2
4	I	0	1
4	S	0	1
All	All	0	30

There are no bond length outliers.

The worst 5 of 79 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L	123	GLU	CB-CG-CD	9.13	128.12	112.60
3	M	73	THR	CA-CB-OG1	-8.74	96.49	109.60
1	P	204	GLU	CB-CA-C	-8.00	97.25	110.85
3	M	44	ARG	CA-CB-CG	7.82	129.73	114.10
4	S	75	LYS	CB-CA-C	7.77	122.12	109.07

There are no chirality outliers.

5 of 30 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	107	ARG	Sidechain
2	B	186	ARG	Sidechain
2	B	23	ARG	Sidechain
3	C	17	ARG	Sidechain
3	C	6	ARG	Sidechain

## 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.

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1447	0	1337	13	0
1	F	1515	0	1403	13	0
1	K	1453	0	1325	11	0
1	P	1582	0	1476	14	0
2	B	1877	0	1756	13	0
2	G	1851	0	1722	21	0
2	L	1860	0	1725	15	0
2	Q	1867	0	1736	14	0
3	C	2106	0	1882	17	0
3	H	2160	0	1955	23	0
3	M	1917	0	1743	11	0
3	R	2067	0	1836	20	0
4	D	813	0	758	4	0
4	I	813	0	749	7	0
4	N	809	0	747	6	0
4	S	813	0	737	6	0
5	E	94	0	84	3	0
5	J	94	0	84	3	0
5	O	94	0	84	2	0
5	T	94	0	84	2	0
6	B	6	0	8	0	0
6	L	6	0	8	0	0
6	Q	6	0	8	0	0
7	G	7	0	10	0	0
8	C	4	0	0	0	0
8	H	4	0	0	0	0
8	M	2	0	0	0	0
8	R	4	0	0	0	0
All	All	25365	0	23257	184	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.

The worst 5 of 184 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:F:25:ASN:HD22	1:F:72:ASN:ND2	1.72	0.88
2:Q:10:ARG:HH11	2:Q:10:ARG:HG3	1.39	0.87
1:F:25:ASN:HD22	1:F:72:ASN:HD21	1.24	0.85

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:R:119:ASP:HB3	4:S:0:MET:HB3	1.64	0.77
2:Q:10:ARG:HG3	2:Q:10:ARG:NH1	2.01	0.73

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	A	185/206 (90%)	178 (96%)	7 (4%)	0	100	100
1	F	194/206 (94%)	189 (97%)	5 (3%)	0	100	100
1	K	184/206 (89%)	179 (97%)	5 (3%)	0	100	100
1	P	202/206 (98%)	194 (96%)	8 (4%)	0	100	100
2	B	240/243 (99%)	235 (98%)	5 (2%)	0	100	100
2	G	239/243 (98%)	233 (98%)	6 (2%)	0	100	100
2	L	237/243 (98%)	231 (98%)	6 (2%)	0	100	100
2	Q	240/243 (99%)	234 (98%)	6 (2%)	0	100	100
3	C	265/276 (96%)	257 (97%)	8 (3%)	0	100	100
3	H	269/276 (98%)	260 (97%)	9 (3%)	0	100	100
3	M	230/276 (83%)	225 (98%)	5 (2%)	0	100	100
3	R	259/276 (94%)	247 (95%)	12 (5%)	0	100	100
4	D	98/100 (98%)	96 (98%)	2 (2%)	0	100	100
4	I	98/100 (98%)	97 (99%)	1 (1%)	0	100	100
4	N	98/100 (98%)	97 (99%)	1 (1%)	0	100	100
4	S	98/100 (98%)	97 (99%)	1 (1%)	0	100	100
5	E	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
5	J	7/9 (78%)	7 (100%)	0	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	O	7/9 (78%)	7 (100%)	0	0	100	100
5	T	7/9 (78%)	5 (71%)	2 (29%)	0	100	100
All	All	3164/3336 (95%)	3074 (97%)	90 (3%)	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	154/185 (83%)	146 (95%)	8 (5%)	19	35
1	F	165/185 (89%)	157 (95%)	8 (5%)	21	39
1	K	154/185 (83%)	145 (94%)	9 (6%)	17	31
1	P	177/185 (96%)	168 (95%)	9 (5%)	20	36
2	B	198/207 (96%)	194 (98%)	4 (2%)	50	70
2	G	193/207 (93%)	186 (96%)	7 (4%)	30	51
2	L	193/207 (93%)	189 (98%)	4 (2%)	48	69
2	Q	195/207 (94%)	189 (97%)	6 (3%)	35	56
3	C	203/234 (87%)	188 (93%)	15 (7%)	11	20
3	H	213/234 (91%)	196 (92%)	17 (8%)	10	18
3	M	190/234 (81%)	183 (96%)	7 (4%)	29	50
3	R	198/234 (85%)	185 (93%)	13 (7%)	14	25
4	D	89/95 (94%)	85 (96%)	4 (4%)	23	42
4	I	89/95 (94%)	86 (97%)	3 (3%)	32	54
4	N	88/95 (93%)	84 (96%)	4 (4%)	23	42
4	S	89/95 (94%)	83 (93%)	6 (7%)	13	25
5	E	9/9 (100%)	8 (89%)	1 (11%)	5	8
5	J	9/9 (100%)	8 (89%)	1 (11%)	5	8
5	O	9/9 (100%)	9 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	T	9/9 (100%)	9 (100%)	0	100	100
All	All	2624/2920 (90%)	2498 (95%)	126 (5%)	21	39

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	H	160	LEU
3	R	35	ARG
1	K	28	THR
3	R	12	VAL
3	R	225	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 73 such sidechains are listed below:

Mol	Chain	Res	Type
1	P	116	ASN
4	S	13	HIS
2	Q	101	HIS
3	R	70	GLN
2	G	212	GLN

### 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 oligosaccharides 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	$\# Z  > 2$
6	GOL	L	301	-	5,5,5	0.26	0	5,5,5	0.39	0
6	GOL	B	301	-	5,5,5	0.15	0	5,5,5	0.38	0
6	GOL	Q	301	-	5,5,5	0.18	0	5,5,5	0.40	0
7	PEG	G	301	-	6,6,6	0.48	0	5,5,5	0.33	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
6	GOL	L	301	-	-	2/4/4/4	-
6	GOL	B	301	-	-	4/4/4/4	-
6	GOL	Q	301	-	-	4/4/4/4	-
7	PEG	G	301	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	301	GOL	C1-C2-C3-O3
6	L	301	GOL	O1-C1-C2-C3
6	Q	301	GOL	C1-C2-C3-O3
6	B	301	GOL	O1-C1-C2-O2
7	G	301	PEG	O2-C3-C4-O4

There are no ring outliers.

No monomer is involved in short contacts.

## 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	191/206 (92%)	0.57	27 (14%) 7 9	22, 48, 94, 114	0
1	F	198/206 (96%)	0.05	5 (2%) 58 59	20, 37, 70, 82	0
1	K	192/206 (93%)	0.55	33 (17%) 5 6	18, 46, 91, 110	0
1	P	204/206 (99%)	-0.24	0 100 100	19, 34, 57, 76	0
2	B	242/243 (99%)	0.19	6 (2%) 58 59	20, 45, 84, 106	0
2	G	241/243 (99%)	0.06	9 (3%) 45 47	16, 37, 75, 89	0
2	L	241/243 (99%)	0.00	4 (1%) 69 70	16, 38, 76, 128	0
2	Q	242/243 (99%)	-0.12	3 (1%) 76 78	18, 34, 67, 106	0
3	C	271/276 (98%)	0.35	14 (5%) 34 35	22, 45, 89, 115	0
3	H	273/276 (98%)	0.70	34 (12%) 9 12	26, 53, 95, 115	0
3	M	240/276 (86%)	0.42	24 (10%) 14 16	23, 46, 91, 125	0
3	R	267/276 (96%)	0.74	36 (13%) 8 10	21, 51, 94, 112	0
4	D	100/100 (100%)	-0.28	0 100 100	19, 33, 47, 65	0
4	I	100/100 (100%)	0.04	1 (1%) 79 82	27, 40, 66, 83	0
4	N	100/100 (100%)	0.22	1 (1%) 79 82	27, 47, 72, 87	0
4	S	100/100 (100%)	-0.28	0 100 100	22, 34, 56, 63	0
5	E	9/9 (100%)	-0.54	0 100 100	19, 30, 35, 40	0
5	J	9/9 (100%)	-0.33	0 100 100	29, 34, 38, 49	0
5	O	9/9 (100%)	-0.52	0 100 100	19, 24, 35, 35	0
5	T	9/9 (100%)	-0.23	0 100 100	26, 31, 39, 40	0
All	All	3238/3336 (97%)	0.23	197 (6%) 28 31	16, 41, 85, 128	0

The worst 5 of 197 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	K	120	ALA	4.9
3	C	220	ASP	4.8
3	R	220	ASP	4.6
3	R	227	ASP	4.4
3	R	272	LEU	4.4

## 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 oligosaccharides 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( $\text{\AA}^2$ )	Q<0.9
6	GOL	Q	301	6/6	0.73	0.21	39,52,53,54	0
6	GOL	L	301	6/6	0.75	0.19	38,40,44,47	0
6	GOL	B	301	6/6	0.79	0.19	40,45,47,50	0
7	PEG	G	301	7/7	0.89	0.13	37,46,46,50	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.