



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 9, 2026 – 04:13 PM EDT

PDB ID : 9Z51 / pdb_00009z51
Title : HPV_Jin TCR in complex with HLA-A*02:01 presenting HPV16 E7 9mer peptide (11-YMLDLQPET-19)
Authors : Palowitch, G.M.; Dulberger, C.L.
Deposited on : 2025-11-11
Resolution : 3.13 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

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	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

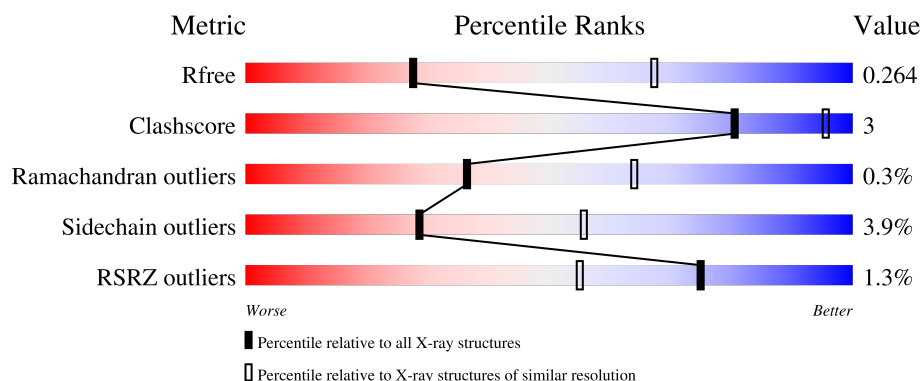
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.13 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}	180053	2351 (3.18-3.10)
Clashscore	190562	2452 (3.18-3.10)
Ramachandran outliers	187476	2324 (3.18-3.10)
Sidechain outliers	187428	2324 (3.18-3.10)
RSRZ outliers	180081	2351 (3.18-3.10)

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 ≥ 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	C	303	85% 10%
1	G	303	85% 5% 10%
1	L	303	83% 6% 12%
1	S	303	3% 83% 13%
2	D	100	94% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	100	 83% 9% 8%
2	M	100	 95% 5%
2	T	100	 76% 13% 9%
3	A	200	 88% 8% ..
3	E	200	 88% 8% ..
3	J	200	 90% 7% .
3	Q	200	 84% 11% ..
4	B	243	 91% 9%
4	F	243	 94% 6%
4	K	243	 92% 7% .
4	R	243	 91% 9%
5	I	9	 100%
5	N	9	 89% 11%
5	P	9	 89% 11%
5	U	9	 89% 11%

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 25248 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 MHC class I antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	272	Total	C	N	O	S	0	0	0
			2164	1358	389	408	9			
1	G	274	Total	C	N	O	S	0	0	0
			2175	1365	393	408	9			
1	L	268	Total	C	N	O	S	0	0	0
			2158	1353	393	403	9			
1	S	263	Total	C	N	O	S	0	0	0
			2069	1301	375	384	9			

There are 108 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1	MET	-	initiating methionine	UNP Q8WLS4
C	278	GLY	-	expression tag	UNP Q8WLS4
C	279	SER	-	expression tag	UNP Q8WLS4
C	280	GLY	-	expression tag	UNP Q8WLS4
C	281	GLY	-	expression tag	UNP Q8WLS4
C	282	SER	-	expression tag	UNP Q8WLS4
C	283	GLY	-	expression tag	UNP Q8WLS4
C	284	GLY	-	expression tag	UNP Q8WLS4
C	285	SER	-	expression tag	UNP Q8WLS4
C	286	ALA	-	expression tag	UNP Q8WLS4
C	287	GLY	-	expression tag	UNP Q8WLS4
C	288	GLY	-	expression tag	UNP Q8WLS4
C	289	GLY	-	expression tag	UNP Q8WLS4
C	290	LEU	-	expression tag	UNP Q8WLS4
C	291	ASN	-	expression tag	UNP Q8WLS4
C	292	ASP	-	expression tag	UNP Q8WLS4
C	293	ILE	-	expression tag	UNP Q8WLS4
C	294	PHE	-	expression tag	UNP Q8WLS4
C	295	GLU	-	expression tag	UNP Q8WLS4
C	296	ALA	-	expression tag	UNP Q8WLS4
C	297	GLN	-	expression tag	UNP Q8WLS4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	298	LYS	-	expression tag	UNP Q8WLS4
C	299	ILE	-	expression tag	UNP Q8WLS4
C	300	GLU	-	expression tag	UNP Q8WLS4
C	301	TRP	-	expression tag	UNP Q8WLS4
C	302	HIS	-	expression tag	UNP Q8WLS4
C	303	GLU	-	expression tag	UNP Q8WLS4
G	1	MET	-	initiating methionine	UNP Q8WLS4
G	278	GLY	-	expression tag	UNP Q8WLS4
G	279	SER	-	expression tag	UNP Q8WLS4
G	280	GLY	-	expression tag	UNP Q8WLS4
G	281	GLY	-	expression tag	UNP Q8WLS4
G	282	SER	-	expression tag	UNP Q8WLS4
G	283	GLY	-	expression tag	UNP Q8WLS4
G	284	GLY	-	expression tag	UNP Q8WLS4
G	285	SER	-	expression tag	UNP Q8WLS4
G	286	ALA	-	expression tag	UNP Q8WLS4
G	287	GLY	-	expression tag	UNP Q8WLS4
G	288	GLY	-	expression tag	UNP Q8WLS4
G	289	GLY	-	expression tag	UNP Q8WLS4
G	290	LEU	-	expression tag	UNP Q8WLS4
G	291	ASN	-	expression tag	UNP Q8WLS4
G	292	ASP	-	expression tag	UNP Q8WLS4
G	293	ILE	-	expression tag	UNP Q8WLS4
G	294	PHE	-	expression tag	UNP Q8WLS4
G	295	GLU	-	expression tag	UNP Q8WLS4
G	296	ALA	-	expression tag	UNP Q8WLS4
G	297	GLN	-	expression tag	UNP Q8WLS4
G	298	LYS	-	expression tag	UNP Q8WLS4
G	299	ILE	-	expression tag	UNP Q8WLS4
G	300	GLU	-	expression tag	UNP Q8WLS4
G	301	TRP	-	expression tag	UNP Q8WLS4
G	302	HIS	-	expression tag	UNP Q8WLS4
G	303	GLU	-	expression tag	UNP Q8WLS4
L	1	MET	-	initiating methionine	UNP Q8WLS4
L	278	GLY	-	expression tag	UNP Q8WLS4
L	279	SER	-	expression tag	UNP Q8WLS4
L	280	GLY	-	expression tag	UNP Q8WLS4
L	281	GLY	-	expression tag	UNP Q8WLS4
L	282	SER	-	expression tag	UNP Q8WLS4
L	283	GLY	-	expression tag	UNP Q8WLS4
L	284	GLY	-	expression tag	UNP Q8WLS4
L	285	SER	-	expression tag	UNP Q8WLS4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	286	ALA	-	expression tag	UNP Q8WLS4
L	287	GLY	-	expression tag	UNP Q8WLS4
L	288	GLY	-	expression tag	UNP Q8WLS4
L	289	GLY	-	expression tag	UNP Q8WLS4
L	290	LEU	-	expression tag	UNP Q8WLS4
L	291	ASN	-	expression tag	UNP Q8WLS4
L	292	ASP	-	expression tag	UNP Q8WLS4
L	293	ILE	-	expression tag	UNP Q8WLS4
L	294	PHE	-	expression tag	UNP Q8WLS4
L	295	GLU	-	expression tag	UNP Q8WLS4
L	296	ALA	-	expression tag	UNP Q8WLS4
L	297	GLN	-	expression tag	UNP Q8WLS4
L	298	LYS	-	expression tag	UNP Q8WLS4
L	299	ILE	-	expression tag	UNP Q8WLS4
L	300	GLU	-	expression tag	UNP Q8WLS4
L	301	TRP	-	expression tag	UNP Q8WLS4
L	302	HIS	-	expression tag	UNP Q8WLS4
L	303	GLU	-	expression tag	UNP Q8WLS4
S	1	MET	-	initiating methionine	UNP Q8WLS4
S	278	GLY	-	expression tag	UNP Q8WLS4
S	279	SER	-	expression tag	UNP Q8WLS4
S	280	GLY	-	expression tag	UNP Q8WLS4
S	281	GLY	-	expression tag	UNP Q8WLS4
S	282	SER	-	expression tag	UNP Q8WLS4
S	283	GLY	-	expression tag	UNP Q8WLS4
S	284	GLY	-	expression tag	UNP Q8WLS4
S	285	SER	-	expression tag	UNP Q8WLS4
S	286	ALA	-	expression tag	UNP Q8WLS4
S	287	GLY	-	expression tag	UNP Q8WLS4
S	288	GLY	-	expression tag	UNP Q8WLS4
S	289	GLY	-	expression tag	UNP Q8WLS4
S	290	LEU	-	expression tag	UNP Q8WLS4
S	291	ASN	-	expression tag	UNP Q8WLS4
S	292	ASP	-	expression tag	UNP Q8WLS4
S	293	ILE	-	expression tag	UNP Q8WLS4
S	294	PHE	-	expression tag	UNP Q8WLS4
S	295	GLU	-	expression tag	UNP Q8WLS4
S	296	ALA	-	expression tag	UNP Q8WLS4
S	297	GLN	-	expression tag	UNP Q8WLS4
S	298	LYS	-	expression tag	UNP Q8WLS4
S	299	ILE	-	expression tag	UNP Q8WLS4
S	300	GLU	-	expression tag	UNP Q8WLS4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
S	301	TRP	-	expression tag	UNP Q8WLS4
S	302	HIS	-	expression tag	UNP Q8WLS4
S	303	GLU	-	expression tag	UNP Q8WLS4

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	100	Total	C	N	O	S	0	0	0
			802	512	136	150	4			
2	H	92	Total	C	N	O	S	0	0	0
			707	457	114	133	3			
2	M	100	Total	C	N	O	S	0	0	0
			792	506	132	150	4			
2	T	91	Total	C	N	O	S	0	0	0
			677	433	110	130	4			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	20	MET	-	initiating methionine	UNP P61769
H	20	MET	-	initiating methionine	UNP P61769
M	20	MET	-	initiating methionine	UNP P61769
T	20	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called HPV_Jin TCR alpha chain (TRAV1-2*03).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	195	Total	C	N	O	S	0	0	0
			1444	918	229	288	9			
3	E	195	Total	C	N	O	S	0	0	0
			1448	919	228	292	9			
3	J	193	Total	C	N	O	S	0	0	0
			1422	904	223	286	9			
3	Q	194	Total	C	N	O	S	0	0	0
			1427	908	226	284	9			

- Molecule 4 is a protein called HPV_Jin TCR beta chain (TRBV5-6*01).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	243	Total	C	N	O	S	0	0	0
			1879	1194	328	352	5			

Continued on next page...

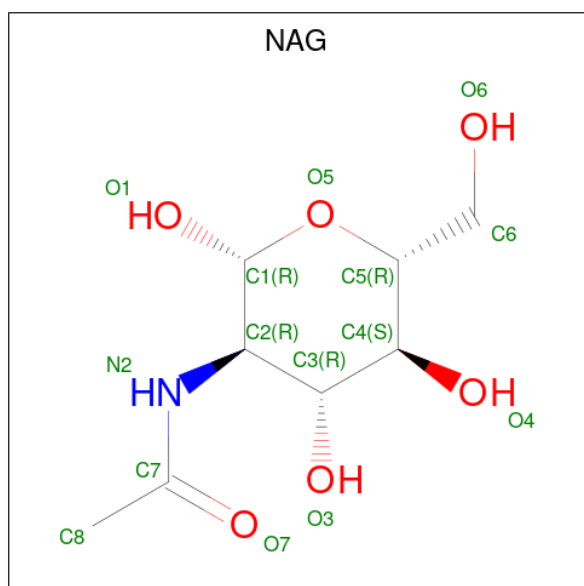
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	243	Total	C	N	O	S	0	0	0
			1860	1184	326	345	5			
4	K	243	Total	C	N	O	S	0	0	0
			1888	1198	333	352	5			
4	R	243	Total	C	N	O	S	0	0	0
			1877	1191	327	354	5			

- Molecule 5 is a protein called Protein E7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	P	9	Total	C	N	O	S	0	0	0
			76	49	10	16	1			
5	I	9	Total	C	N	O	S	0	0	0
			76	49	10	16	1			
5	N	9	Total	C	N	O	S	0	0	0
			76	49	10	16	1			
5	U	9	Total	C	N	O	S	0	0	0
			76	49	10	16	1			

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula: $C_8H_{15}NO_6$).



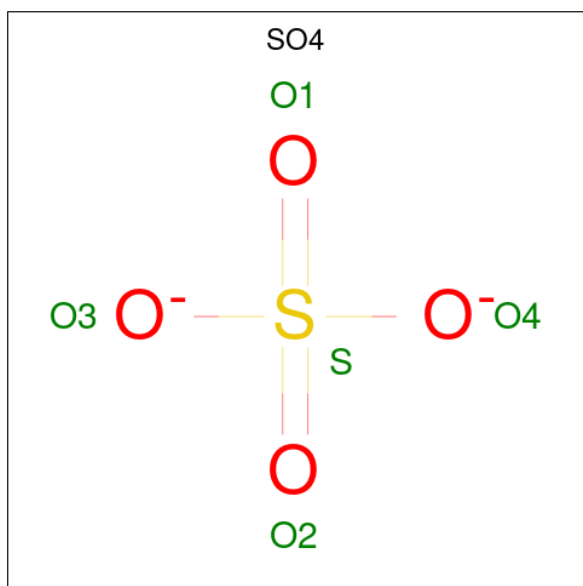
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	B	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	E	1	Total	C	N	O	0	0
			14	8	1	5		
6	F	1	Total	C	N	O	0	0
			14	8	1	5		
6	J	1	Total	C	N	O	0	0
			14	8	1	5		
6	K	1	Total	C	N	O	0	0
			14	8	1	5		
6	Q	1	Total	C	N	O	0	0
			14	8	1	5		
6	Q	1	Total	C	N	O	0	0
			14	8	1	5		
6	R	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 7 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	O	S	0	0
			5	4	1		
7	E	1	Total	O	S	0	0
			5	4	1		
7	J	1	Total	O	S	0	0
			5	4	1		
7	Q	1	Total	O	S	0	0
			5	4	1		

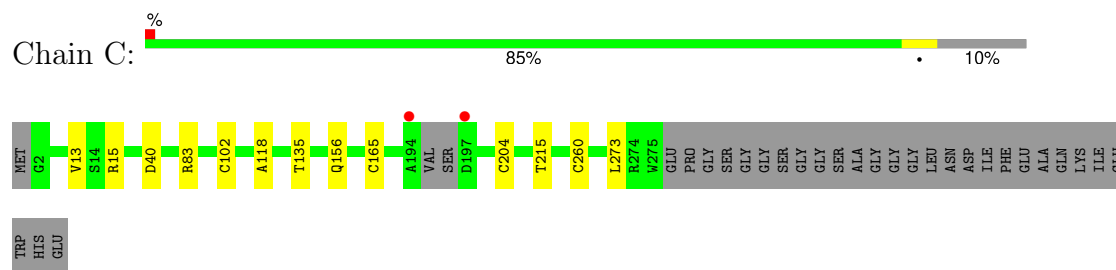
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	C	1	Total 1	O 1	0	0
8	B	1	Total 1	O 1	0	0
8	G	1	Total 1	O 1	0	0
8	L	3	Total 3	O 3	0	0
8	J	2	Total 2	O 2	0	0
8	S	1	Total 1	O 1	0	0

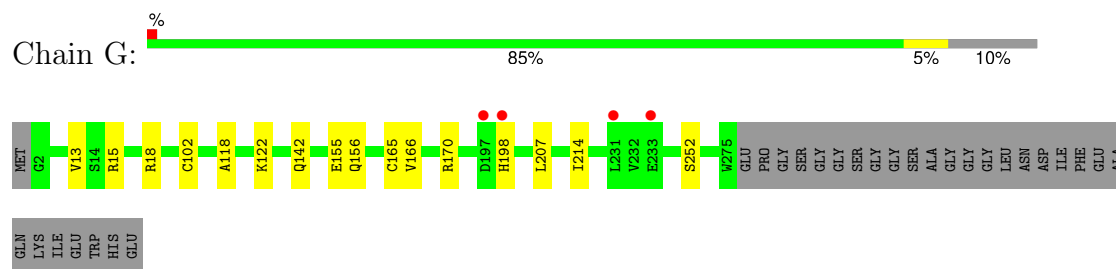
3 Residue-property plots [i](#)

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.

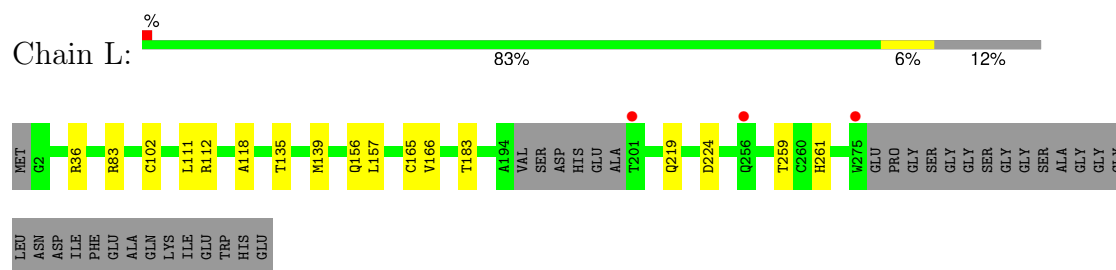
- Molecule 1: MHC class I antigen



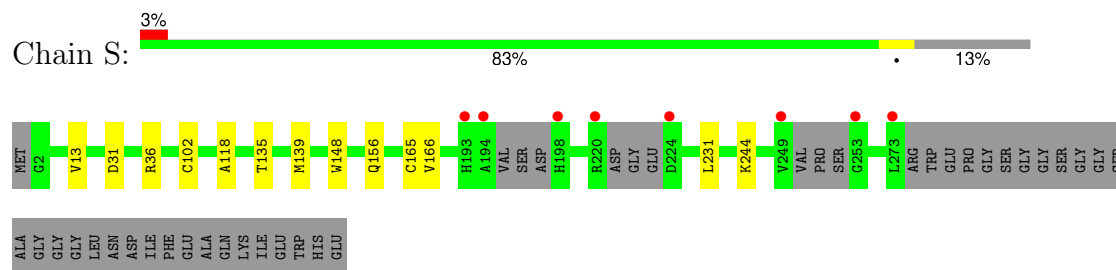
- Molecule 1: MHC class I antigen



- Molecule 1: MHC class I antigen



- Molecule 1: MHC class I antigen



- Molecule 2: Beta-2-microglobulin

Chain D:  94% 6%



- Molecule 2: Beta-2-microglobulin

Chain H:  83% 9% 8%



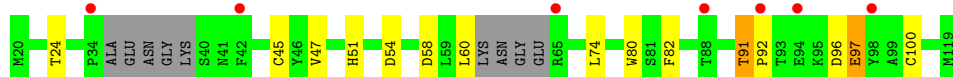
- Molecule 2: Beta-2-microglobulin

Chain M:  95% 5% 2%



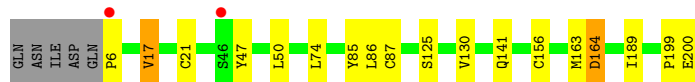
- Molecule 2: Beta-2-microglobulin

Chain T:  76% 13% 9% 7%



- Molecule 3: HPV_Jin TCR alpha chain (TRAV1-2*03)

Chain A:  88% 8% 2%



- Molecule 3: HPV_Jin TCR alpha chain (TRAV1-2*03)

Chain E:  88% 8% 2%

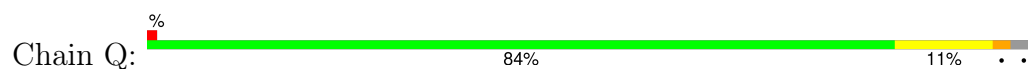


- Molecule 3: HPV_Jin TCR alpha chain (TRAV1-2*03)

Chain J:  90% 7% 2%



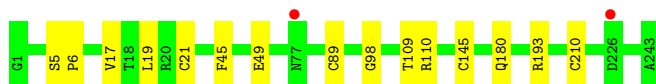
- Molecule 3: HPV_Jin TCR alpha chain (TRAV1-2*03)



- Molecule 4: HPV_Jin TCR beta chain (TRBV5-6*01)



- Molecule 4: HPV_Jin TCR beta chain (TRBV5-6*01)



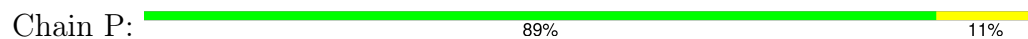
- Molecule 4: HPV_Jin TCR beta chain (TRBV5-6*01)



- Molecule 4: HPV_Jin TCR beta chain (TRBV5-6*01)



- Molecule 5: Protein E7



- Molecule 5: Protein E7

Chain I:  100%

There are no outlier residues recorded for this chain.

- Molecule 5: Protein E7

Chain N:  89% 11%



- Molecule 5: Protein E7

Chain U:  89% 11%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	113.02Å 203.81Å 206.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	102.11 – 3.13 102.11 – 3.13	Depositor EDS
% Data completeness (in resolution range)	100.0 (102.11-3.13) 100.0 (102.11-3.13)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.84 (at 3.13Å)	Xtriage
Refinement program	REFMAC 5.8.0430	Depositor
R, R_{free}	0.224 , 0.267 0.225 , 0.264	Depositor DCC
R_{free} test set	4235 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	76.8	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 48.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.002 for -h,l,k	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	25248	wwPDB-VP
Average B, all atoms (Å ²)	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.45% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, 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	C	0.45	0/2227	0.80	0/3031
1	G	0.45	0/2239	0.83	0/3049
1	L	0.45	0/2221	0.82	0/3019
1	S	0.45	0/2125	0.84	0/2889
2	D	0.46	0/825	0.73	0/1122
2	H	0.46	0/728	0.73	0/999
2	M	0.45	0/815	0.72	0/1111
2	T	0.47	0/694	0.76	0/953
3	A	0.47	0/1479	0.76	0/2020
3	E	0.47	0/1483	0.75	0/2026
3	J	0.47	0/1456	0.77	0/1990
3	Q	0.47	0/1462	0.76	0/1999
4	B	0.46	0/1936	0.75	0/2647
4	F	0.46	0/1917	0.75	0/2624
4	K	0.46	0/1945	0.75	0/2658
4	R	0.46	0/1934	0.76	0/2646
5	I	0.46	0/77	0.89	0/104
5	N	0.46	0/77	0.94	0/104
5	P	0.46	0/77	0.90	0/104
5	U	0.46	0/77	0.86	0/104
All	All	0.46	0/25794	0.78	0/35199

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 ⓘ

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	C	2164	0	1981	9	0
1	G	2175	0	1997	7	0
1	L	2158	0	1997	6	0
1	S	2069	0	1897	5	0
2	D	802	0	738	2	0
2	H	707	0	621	3	0
2	M	792	0	716	1	0
2	T	677	0	586	6	0
3	A	1444	0	1295	9	0
3	E	1448	0	1292	7	0
3	J	1422	0	1264	8	0
3	Q	1427	0	1272	9	0
4	B	1879	0	1733	15	0
4	F	1860	0	1702	7	0
4	K	1888	0	1746	12	0
4	R	1877	0	1718	10	0
5	I	76	0	74	0	0
5	N	76	0	74	1	0
5	P	76	0	74	0	0
5	U	76	0	74	1	0
6	A	14	0	13	0	0
6	B	14	0	13	0	0
6	E	14	0	13	0	0
6	F	14	0	13	0	0
6	J	14	0	13	0	0
6	K	14	0	13	0	0
6	Q	28	0	26	1	0
6	R	14	0	13	0	0
7	A	5	0	0	0	0
7	E	5	0	0	0	0
7	J	5	0	0	0	0
7	Q	5	0	0	0	0
8	B	1	0	0	0	0
8	C	1	0	0	0	0
8	G	1	0	0	0	0
8	J	2	0	0	0	0
8	L	3	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	S	1	0	0	0	0
All	All	25248	0	22968	98	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 3.

The worst 5 of 98 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)
2:H:45:CYS:HG	2:H:100:CYS:HG	1.02	1.01
1:G:102:CYS:HG	1:G:165:CYS:HG	0.97	0.96
3:A:21:CYS:HG	3:A:87:CYS:HG	0.97	0.96
1:C:102:CYS:HG	1:C:165:CYS:HG	1.01	0.88
1:S:102:CYS:HG	1:S:165:CYS:HG	1.14	0.83

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	C	268/303 (88%)	254 (95%)	14 (5%)	0	100	100
1	G	272/303 (90%)	254 (93%)	18 (7%)	0	100	100
1	L	264/303 (87%)	255 (97%)	9 (3%)	0	100	100
1	S	255/303 (84%)	236 (92%)	19 (8%)	0	100	100
2	D	98/100 (98%)	95 (97%)	3 (3%)	0	100	100
2	H	88/100 (88%)	84 (96%)	4 (4%)	0	100	100
2	M	98/100 (98%)	94 (96%)	4 (4%)	0	100	100
2	T	85/100 (85%)	73 (86%)	10 (12%)	2 (2%)	4	19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	A	193/200 (96%)	175 (91%)	16 (8%)	2 (1%)	12	38
3	E	193/200 (96%)	179 (93%)	12 (6%)	2 (1%)	12	38
3	J	191/200 (96%)	178 (93%)	12 (6%)	1 (0%)	24	54
3	Q	192/200 (96%)	179 (93%)	11 (6%)	2 (1%)	12	38
4	B	241/243 (99%)	224 (93%)	17 (7%)	0	100	100
4	F	241/243 (99%)	222 (92%)	19 (8%)	0	100	100
4	K	241/243 (99%)	230 (95%)	10 (4%)	1 (0%)	30	59
4	R	241/243 (99%)	229 (95%)	11 (5%)	1 (0%)	30	59
5	I	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
5	N	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
5	P	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
5	U	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
All	All	3189/3420 (93%)	2985 (94%)	193 (6%)	11 (0%)	36	64

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	T	97	GLU
3	E	126	SER
3	Q	164	ASP
3	A	164	ASP
3	E	125	SER

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	C	215/249 (86%)	211 (98%)	4 (2%)	50	69
1	G	216/249 (87%)	210 (97%)	6 (3%)	38	62
1	L	217/249 (87%)	207 (95%)	10 (5%)	24	51
1	S	201/249 (81%)	195 (97%)	6 (3%)	36	61

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	86/95 (90%)	83 (96%)	3 (4%)	32	58
2	H	73/95 (77%)	69 (94%)	4 (6%)	19	46
2	M	84/95 (88%)	80 (95%)	4 (5%)	23	50
2	T	70/95 (74%)	65 (93%)	5 (7%)	13	37
3	A	151/178 (85%)	144 (95%)	7 (5%)	24	51
3	E	152/178 (85%)	142 (93%)	10 (7%)	15	40
3	J	148/178 (83%)	143 (97%)	5 (3%)	32	58
3	Q	148/178 (83%)	138 (93%)	10 (7%)	14	38
4	B	193/210 (92%)	188 (97%)	5 (3%)	40	64
4	F	187/210 (89%)	181 (97%)	6 (3%)	34	59
4	K	194/210 (92%)	188 (97%)	6 (3%)	35	60
4	R	192/210 (91%)	185 (96%)	7 (4%)	31	58
5	I	9/9 (100%)	9 (100%)	0	100	100
5	N	9/9 (100%)	9 (100%)	0	100	100
5	P	9/9 (100%)	8 (89%)	1 (11%)	6	21
5	U	9/9 (100%)	9 (100%)	0	100	100
All	All	2563/2964 (86%)	2464 (96%)	99 (4%)	28	56

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	M	87	TYR
1	S	13	VAL
2	M	103	ASN
4	K	21	CYS
1	S	139	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
3	J	35	GLN
4	K	119	ASN
4	R	66	GLN
3	J	48	ASN
4	K	39	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

13 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	NAG	F	301	-	14,14,15	0.39	0	17,19,21	0.47	0
6	NAG	A	301	-	14,14,15	0.37	0	17,19,21	1.25	2 (11%)
7	SO4	A	302	-	4,4,4	0.34	0	6,6,6	0.07	0
7	SO4	E	302	-	4,4,4	0.34	0	6,6,6	0.07	0
7	SO4	J	302	-	4,4,4	0.34	0	6,6,6	0.08	0
6	NAG	K	301	-	14,14,15	0.39	0	17,19,21	0.75	1 (5%)
6	NAG	E	301	-	14,14,15	0.40	0	17,19,21	0.69	1 (5%)
6	NAG	J	301	-	14,14,15	0.38	0	17,19,21	1.02	2 (11%)
6	NAG	Q	301	-	14,14,15	0.41	0	17,19,21	0.69	0
7	SO4	Q	303	-	4,4,4	0.34	0	6,6,6	0.08	0
6	NAG	R	301	-	14,14,15	0.41	0	17,19,21	0.71	0
6	NAG	Q	302	-	14,14,15	0.38	0	17,19,21	0.45	0
6	NAG	B	301	-	14,14,15	0.39	0	17,19,21	0.79	1 (5%)

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	NAG	F	301	-	-	2/6/23/26	0/1/1/1
6	NAG	A	301	-	-	4/6/23/26	0/1/1/1
6	NAG	K	301	-	-	3/6/23/26	0/1/1/1
6	NAG	E	301	-	-	2/6/23/26	0/1/1/1
6	NAG	J	301	-	-	3/6/23/26	0/1/1/1
6	NAG	Q	301	-	-	2/6/23/26	0/1/1/1
6	NAG	R	301	-	-	3/6/23/26	0/1/1/1
6	NAG	Q	302	-	-	3/6/23/26	0/1/1/1
6	NAG	B	301	-	-	5/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	301	NAG	O5-C1-C2	-4.02	105.07	111.29
6	J	301	NAG	O5-C1-C2	-3.15	106.42	111.29
6	A	301	NAG	C1-C2-N2	2.65	114.62	110.43
6	B	301	NAG	O5-C1-C2	-2.62	107.24	111.29
6	E	301	NAG	C1-O5-C5	2.55	115.61	112.19

There are no chirality outliers.

5 of 27 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	E	301	NAG	C8-C7-N2-C2
6	E	301	NAG	O7-C7-N2-C2
6	J	301	NAG	C8-C7-N2-C2
6	J	301	NAG	O7-C7-N2-C2
6	Q	301	NAG	C8-C7-N2-C2

There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	Q	301	NAG	1	0
6	Q	302	NAG	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, 95th 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 ²)	Q<0.9
1	C	272/303 (89%)	0.02	2 (0%) 84 68	54, 76, 115, 132	0
1	G	274/303 (90%)	0.16	4 (1%) 72 51	59, 87, 127, 142	0
1	L	268/303 (88%)	-0.01	3 (1%) 78 59	53, 74, 119, 132	0
1	S	263/303 (86%)	0.31	8 (3%) 52 31	59, 89, 135, 146	0
2	D	100/100 (100%)	0.08	0 100 100	56, 79, 110, 120	0
2	H	92/100 (92%)	0.41	1 (1%) 78 59	68, 101, 131, 144	0
2	M	100/100 (100%)	0.06	2 (2%) 65 43	59, 86, 114, 118	0
2	T	91/100 (91%)	0.86	7 (7%) 19 10	70, 109, 147, 168	0
3	A	195/200 (97%)	0.03	2 (1%) 79 61	55, 72, 96, 108	0
3	E	195/200 (97%)	-0.03	2 (1%) 79 61	53, 74, 99, 123	0
3	J	193/200 (96%)	0.11	3 (1%) 70 49	54, 74, 97, 119	0
3	Q	194/200 (97%)	0.24	2 (1%) 79 61	57, 86, 115, 129	0
4	B	243/243 (100%)	-0.16	1 (0%) 88 76	50, 66, 82, 102	0
4	F	243/243 (100%)	0.03	2 (0%) 82 65	54, 74, 98, 117	0
4	K	243/243 (100%)	-0.15	1 (0%) 88 76	52, 66, 85, 106	0
4	R	243/243 (100%)	0.04	1 (0%) 88 76	61, 81, 111, 120	0
5	I	9/9 (100%)	-0.08	0 100 100	63, 65, 73, 74	0
5	N	9/9 (100%)	-0.08	0 100 100	60, 63, 71, 76	0
5	P	9/9 (100%)	-0.39	0 100 100	57, 60, 65, 71	0
5	U	9/9 (100%)	-0.13	0 100 100	62, 67, 75, 76	0
All	All	3245/3420 (94%)	0.08	41 (1%) 75 55	50, 77, 119, 168	0

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	275	TRP	3.5
2	T	98	TYR	3.4
3	E	6	PRO	3.3
1	S	198	HIS	3.3
3	Q	6	PRO	3.2

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, 95th 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	NAG	Q	302	14/15	0.52	0.16	156,163,164,166	0
6	NAG	A	301	14/15	0.60	0.16	100,106,108,108	0
6	NAG	R	301	14/15	0.72	0.16	120,123,126,126	0
6	NAG	J	301	14/15	0.73	0.15	122,130,132,133	0
7	SO4	Q	303	5/5	0.78	0.12	114,116,117,119	0
6	NAG	B	301	14/15	0.79	0.15	94,96,98,98	0
6	NAG	F	301	14/15	0.79	0.14	98,102,105,105	0
6	NAG	Q	301	14/15	0.80	0.14	98,103,105,105	0
6	NAG	K	301	14/15	0.82	0.12	90,91,95,96	0
7	SO4	E	302	5/5	0.84	0.11	106,106,108,108	0
7	SO4	J	302	5/5	0.85	0.10	106,107,108,110	0
7	SO4	A	302	5/5	0.86	0.09	104,106,107,108	0
6	NAG	E	301	14/15	0.87	0.12	105,107,109,109	0

6.5 Other polymers [i](#)

There are no such residues in this entry.