



# wwPDB X-ray Structure Validation Summary Report

Jul 19, 2022 – 01:05 pm BST

PDB ID : 7NDQ  
Title : Gag:02 TCR in complex with HLA-E.  
Authors : Pengelly, R.J.; Robinson, R.A.  
Deposited on : 2021-02-02  
Resolution : 2.55 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 : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.29  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.29

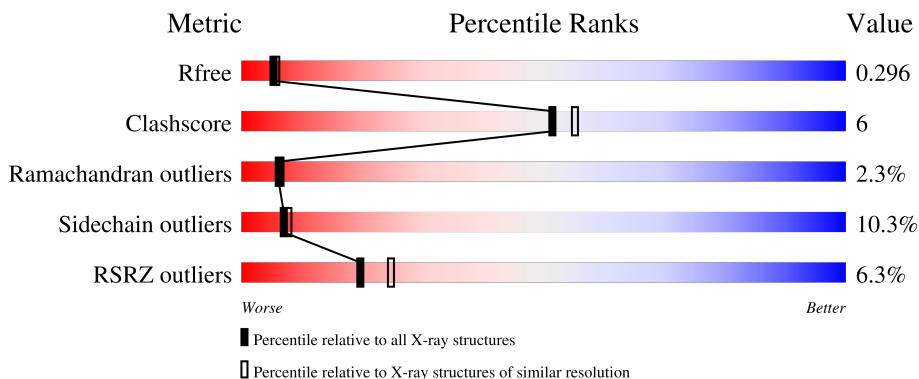
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 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	1279 (2.58-2.54)
Clashscore	141614	1327 (2.58-2.54)
Ramachandran outliers	138981	1312 (2.58-2.54)
Sidechain outliers	138945	1312 (2.58-2.54)
RSRZ outliers	127900	1269 (2.58-2.54)

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	AAA	277	
2	BBB	100	
3	CCC	9	
4	DDD	199	
5	EEE	244	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 6573 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 HLA class I histocompatibility antigen, alpha chain E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	276	2253	1408	403	435	7	0	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	0	MET	-	initiating methionine	UNP P13747

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	BBB	100	837	533	141	159	4	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called Gag6V.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	CCC	9	74	48	12	13	1	0	0	0

- Molecule 4 is a protein called TCR Gag:02-alpha,TCR Gag:02-alpha,TCR Gag:02-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	DDD	185	1452	908	240	296	8	0	2	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
DDD	0	MET	-	initiating methionine	UNP A0A0B4J268
DDD	108	SER	-	linker	UNP A0A0B4J268
DDD	109	SER	-	linker	UNP A0A0B4J268
DDD	110	PHE	-	linker	UNP A0A0B4J268
DDD	129	ASN	-	linker	UNP A0A075B6U7
DDD	176	CYS	THR	engineered mutation	UNP P0DSE1

- Molecule 5 is a protein called T cell receptor beta variable 7-9,M1-specific T cell receptor beta chain,T cell receptor beta variable 7-9,M1-specific T cell receptor beta chain,T cell receptor beta variable 7-9,M1-specific T cell receptor beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	EEE	241	1944	1219	347	372	6	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
EEE	0	MET	-	initiating methionine	UNP P04435
EEE	109	GLY	-	linker	UNP P04435
EEE	110	ARG	-	linker	UNP P04435
EEE	113	GLU	-	linker	UNP P04435
EEE	132	ASN	LYS	engineered mutation	UNP K7N5M4
EEE	133	LYS	ASN	engineered mutation	UNP K7N5M4
EEE	217	ASP	ASN	engineered mutation	UNP K7N5M4

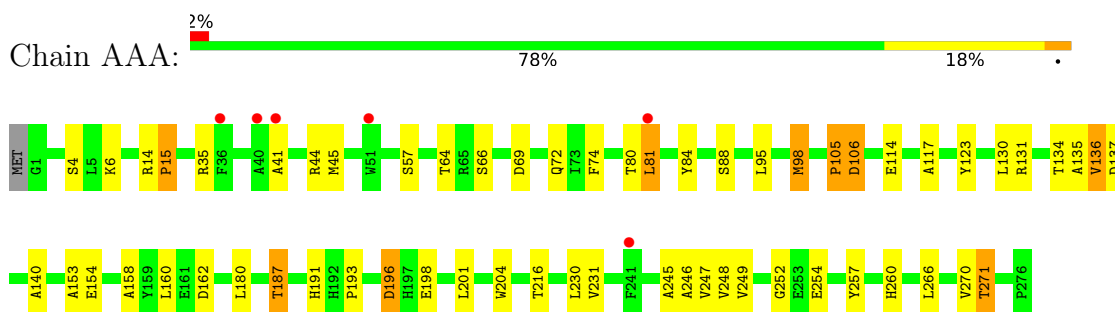
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	AAA	8	Total	O	0	0
			8	8		
6	BBB	3	Total	O	0	0
			3	3		
6	DDD	1	Total	O	0	0
			1	1		
6	EEE	1	Total	O	0	0
			1	1		

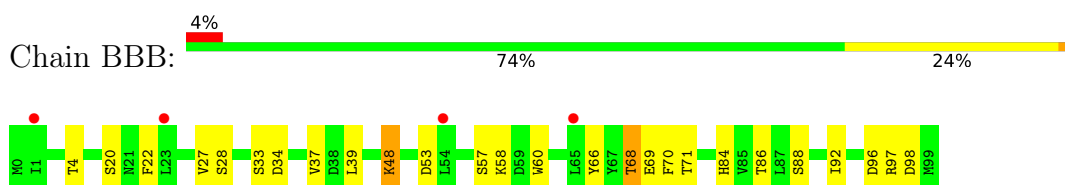
### 3 Residue-property plots

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.

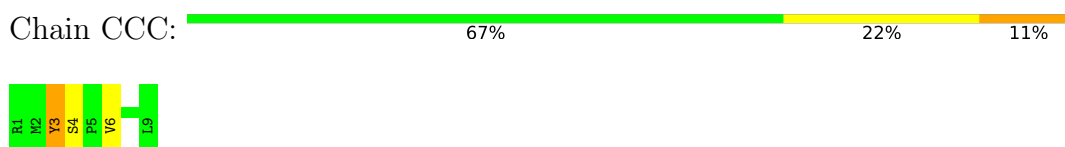
- Molecule 1: HLA class I histocompatibility antigen, alpha chain E



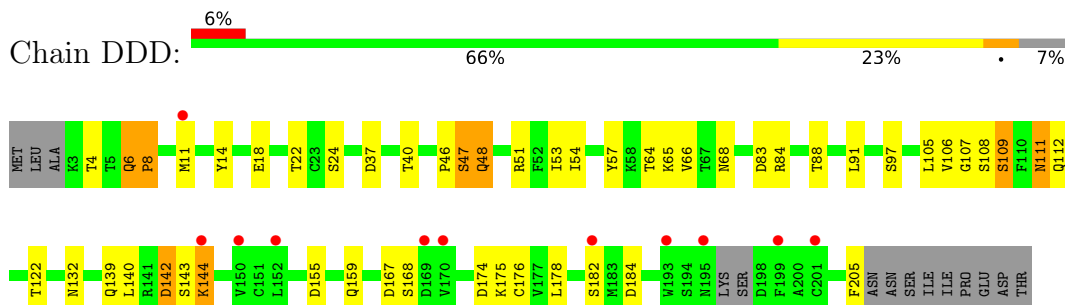
- Molecule 2: Beta-2-microglobulin



- Molecule 3: Gag6V




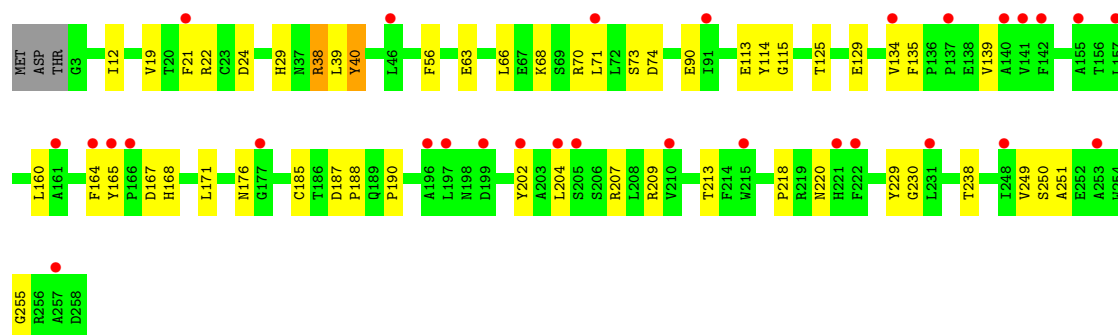
- Molecule 4: TCR Gag:02-alpha,TCR Gag:02-alpha,TCR Gag:02-alpha



- Molecule 5: T cell receptor beta variable 7-9,M1-specific T cell receptor beta chain,T cell receptor beta variable 7-9,M1-specific T cell receptor beta chain,T cell receptor beta variable 7-9,M1-specific

## T cell receptor beta chain

Chain EEE:  12% 78% 20% ..



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	89.39Å 89.39Å 293.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	76.44 – 2.55 76.33 – 2.55	Depositor EDS
% Data completeness (in resolution range)	100.0 (76.44-2.55) 99.8 (76.33-2.55)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.98 (at 2.55Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.240 , 0.296 0.246 , 0.296	Depositor DCC
$R_{free}$ test set	1948 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	78.1	Xtriage
Anisotropy	0.451	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6573	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.57% 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](#)

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	AAA	0.69	0/2320	0.85	0/3154
2	BBB	0.66	0/860	0.85	0/1162
3	CCC	0.67	0/75	1.01	0/99
4	DDD	0.71	0/1487	0.92	2/2017 (0.1%)
5	EEE	0.70	0/1995	0.85	0/2708
All	All	0.70	0/6737	0.87	2/9140 (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	AAA	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	DDD	6	GLN	C-N-CD	-10.81	96.81	120.60
4	DDD	8	PRO	CA-N-CD	-9.85	97.71	111.50

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	136	VAL	Peptide



## 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	AAA	2253	0	2092	26	0
2	BBB	837	0	803	11	0
3	CCC	74	0	81	2	0
4	DDD	1452	0	1387	21	0
5	EEE	1944	0	1846	23	0
6	AAA	8	0	0	0	0
6	BBB	3	0	0	0	0
6	DDD	1	0	0	0	0
6	EEE	1	0	0	0	0
All	All	6573	0	6209	71	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 6.

The worst 5 of 71 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:BBB:4:THR:HA	2:BBB:86:THR:HG21	1.54	0.87
4:DDD:8:PRO:O	4:DDD:122:THR:OG1	2.07	0.72
4:DDD:37:ASP:O	4:DDD:84:ARG:NH2	2.21	0.72
2:BBB:39:LEU:HD12	2:BBB:68:THR:HG22	1.70	0.72
1:AAA:260:HIS:ND1	1:AAA:271:THR:HG22	2.07	0.69

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	AAA	275/277 (99%)	242 (88%)	29 (10%)	4 (2%)	10	13
2	BBB	98/100 (98%)	85 (87%)	8 (8%)	5 (5%)	2	1
3	CCC	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
4	DDD	183/199 (92%)	151 (82%)	25 (14%)	7 (4%)	3	2
5	EEE	239/244 (98%)	216 (90%)	21 (9%)	2 (1%)	19	27
All	All	802/829 (97%)	700 (87%)	84 (10%)	18 (2%)	6	7

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	BBB	33	SER
4	DDD	143	SER
1	AAA	105	PRO
2	BBB	20	SER
2	BBB	58	LYS

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	AAA	238/238 (100%)	215 (90%)	23 (10%)	8	9
2	BBB	95/95 (100%)	85 (90%)	10 (10%)	7	7
3	CCC	9/9 (100%)	7 (78%)	2 (22%)	1	1
4	DDD	169/180 (94%)	148 (88%)	21 (12%)	4	5
5	EEE	211/214 (99%)	191 (90%)	20 (10%)	8	10
All	All	722/736 (98%)	646 (90%)	76 (10%)	7	7

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	EEE	21	PHE
5	EEE	160	LEU
5	EEE	24	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
5	EEE	71	LEU
5	EEE	250	SER

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

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	AAA	276/277 (99%)	0.53	6 (2%) 62 70	59, 86, 117, 134	0
2	BBB	100/100 (100%)	0.60	4 (4%) 38 47	64, 86, 103, 119	0
3	CCC	9/9 (100%)	0.95	0 100 100	71, 77, 87, 88	0
4	DDD	185/199 (92%)	0.42	11 (5%) 22 27	77, 118, 150, 172	0
5	EEE	241/244 (98%)	0.93	30 (12%) 4 6	67, 94, 133, 144	0
All	All	811/829 (97%)	0.64	51 (6%) 20 25	59, 93, 140, 172	0

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	EEE	71	LEU	6.4
4	DDD	150	VAL	5.2
1	AAA	41	ALA	4.6
4	DDD	144	LYS	4.5
1	AAA	40	ALA	4.5

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.