



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 20, 2023 – 04:13 pm BST

PDB ID : 7ZT3  
Title : Structure of E8 TCR in complex in human MR1 K43A  
Authors : Karuppiah, V.; Srikannathasan, V.; Robinson, R.A.  
Deposited on : 2022-05-09  
Resolution : 2.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The 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 : 2.33  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.33

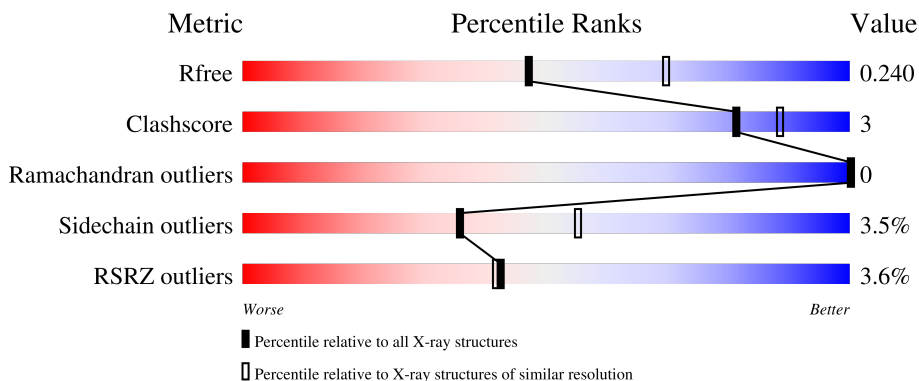
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	290	 3% 82% 7% 10%
2	B	100	 2% 92% 5%
3	D	205	 7% 75% 13% 11%
4	E	262	 1% 85% 6% 8%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6548 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 Major histocompatibility complex class I-related gene protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	260	2158	1383	372	392	11	0	1	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q95460
A	43	ALA	LYS	conflict	UNP Q95460
A	261	SER	CYS	conflict	UNP Q95460
A	271	GLY	-	expression tag	UNP Q95460
A	272	SER	-	expression tag	UNP Q95460
A	273	GLY	-	expression tag	UNP Q95460
A	274	GLY	-	expression tag	UNP Q95460
A	275	GLY	-	expression tag	UNP Q95460
A	276	LEU	-	expression tag	UNP Q95460
A	277	ASN	-	expression tag	UNP Q95460
A	278	ASP	-	expression tag	UNP Q95460
A	279	ILE	-	expression tag	UNP Q95460
A	280	PHE	-	expression tag	UNP Q95460
A	281	GLU	-	expression tag	UNP Q95460
A	282	ALA	-	expression tag	UNP Q95460
A	283	GLN	-	expression tag	UNP Q95460
A	284	LYS	-	expression tag	UNP Q95460
A	285	ILE	-	expression tag	UNP Q95460
A	286	GLU	-	expression tag	UNP Q95460
A	287	TRP	-	expression tag	UNP Q95460
A	288	HIS	-	expression tag	UNP Q95460
A	289	GLU	-	expression tag	UNP Q95460

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	97	817	521	138	155	3	0	1	0

There is a discrepancy between the modelled and reference sequences:

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

- Molecule 3 is a protein called TCR alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	182	1425	899	232	285	9	0	0	0

- Molecule 4 is a protein called TCR beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	E	240	1917	1204	332	372	9	0	1	0

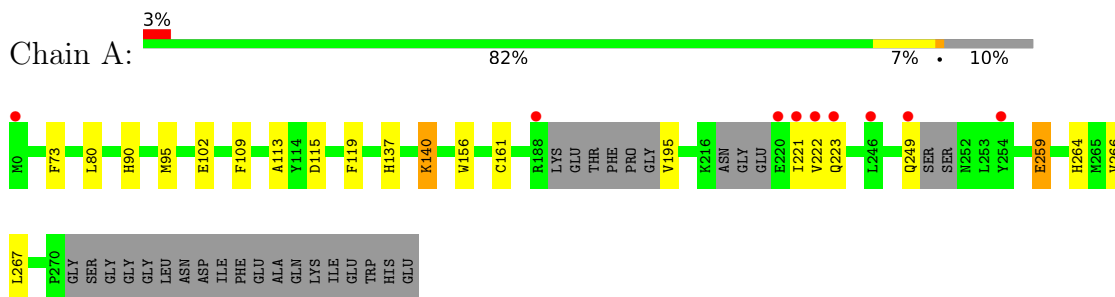
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	88	Total	O	0	0
			88	88		
5	B	24	Total	O	0	0
			24	24		
5	D	37	Total	O	0	0
			37	37		
5	E	82	Total	O	0	0
			82	82		

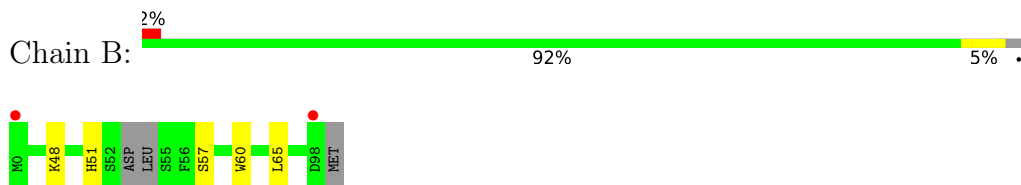
### 3 Residue-property plots [i](#)

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.

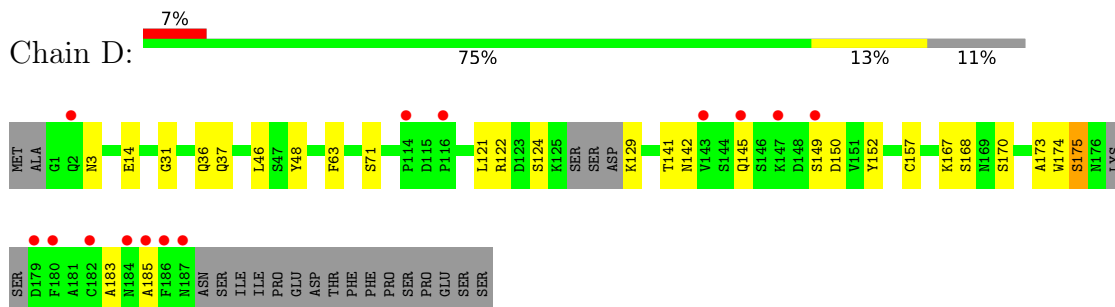
- Molecule 1: Major histocompatibility complex class I-related gene protein



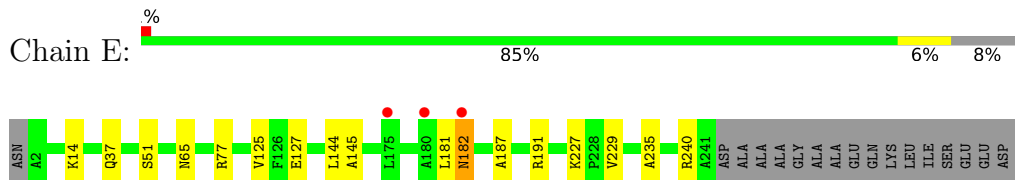
- Molecule 2: Beta-2-microglobulin



- Molecule 3: TCR alpha



- Molecule 4: TCR beta



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.24Å 111.38Å 143.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.03 – 2.40 47.83 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.8 (44.03-2.40) 99.8 (47.83-2.40)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.35 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.191 , 0.236 0.205 , 0.240	Depositor DCC
$R_{free}$ test set	2216 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	52.9	Xtrriage
Anisotropy	0.103	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 33.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6548	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.00% 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	A	0.65	0/2226	0.82	0/3026
2	B	0.64	0/842	0.86	0/1138
3	D	0.63	0/1453	0.85	0/1963
4	E	0.68	0/1970	0.84	0/2685
All	All	0.65	0/6491	0.84	0/8812

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 [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2158	0	2049	12	0
2	B	817	0	782	2	0
3	D	1425	0	1349	12	0
4	E	1917	0	1807	10	0
5	A	88	0	0	0	0
5	B	24	0	0	0	0
5	D	37	0	0	0	0
5	E	82	0	0	1	0
All	All	6548	0	5987	32	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.

All (32) 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)
3:D:37:GLN:HE22	4:E:37:GLN:HE22	1.26	0.81
3:D:14:GLU:OE2	3:D:167:LYS:HE2	1.90	0.71
1:A:90:HIS:HD2	1:A:115:ASP:OD2	1.79	0.66
3:D:150:ASP:O	3:D:175:SER:OG	2.20	0.60
1:A:259:GLU:HG3	1:A:264:HIS:CE1	2.37	0.59
1:A:95:MET:HG2	1:A:109:PHE:HE1	1.68	0.59
3:D:121:LEU:N	3:D:121:LEU:HD12	2.20	0.56
4:E:182:ASN:N	4:E:182:ASN:OD1	2.43	0.52
3:D:174:TRP:CD2	4:E:144:LEU:HD21	2.47	0.50
3:D:183:ALA:C	3:D:185:ALA:H	2.15	0.50
1:A:267:LEU:HD12	1:A:267:LEU:N	2.27	0.50
4:E:181:LEU:HD12	4:E:181:LEU:N	2.27	0.49
1:A:80:LEU:HD21	1:A:119:PHE:CZ	2.48	0.49
3:D:36:GLN:HB2	3:D:46:LEU:HD11	1.97	0.47
1:A:113:ALA:HB2	2:B:60:TRP:CE2	2.50	0.45
4:E:125:VAL:HG23	4:E:235:ALA:HB3	1.97	0.45
4:E:227:LYS:HG2	4:E:229:VAL:HG13	1.97	0.45
3:D:121:LEU:N	3:D:121:LEU:CD1	2.80	0.44
2:B:51:HIS:HA	2:B:65:LEU:O	2.18	0.43
3:D:124:SER:HB3	4:E:127:GLU:HG3	2.01	0.43
3:D:63:PHE:O	3:D:71:SER:HA	2.19	0.43
1:A:223:GLN:H	1:A:223:GLN:CD	2.23	0.42
4:E:145:ALA:O	4:E:187:ALA:HA	2.19	0.42
1:A:221:ILE:HG22	1:A:222:VAL:O	2.20	0.41
4:E:144:LEU:HD12	4:E:144:LEU:HA	1.93	0.41
4:E:14:LYS:NZ	5:E:302:HOH:O	2.52	0.41
1:A:156:TRP:CZ3	1:A:161:CYS:HB2	2.55	0.41
3:D:152:TYR:O	3:D:173:ALA:HA	2.21	0.41
1:A:137:HIS:CE1	1:A:140:LYS:HE3	2.56	0.40
1:A:266:VAL:C	1:A:267:LEU:HD12	2.42	0.40
3:D:31:GLY:HA3	3:D:48:TYR:CE1	2.56	0.40
1:A:102:GLU:OE1	1:A:102:GLU:N	2.54	0.40

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	253/290 (87%)	246 (97%)	7 (3%)	0	100	100
2	B	94/100 (94%)	92 (98%)	2 (2%)	0	100	100
3	D	176/205 (86%)	160 (91%)	16 (9%)	0	100	100
4	E	239/262 (91%)	231 (97%)	8 (3%)	0	100	100
All	All	762/857 (89%)	729 (96%)	33 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	A	231/254 (91%)	226 (98%)	5 (2%)	52	71
2	B	93/95 (98%)	91 (98%)	2 (2%)	52	71
3	D	157/181 (87%)	146 (93%)	11 (7%)	15	24
4	E	210/222 (95%)	204 (97%)	6 (3%)	42	62
All	All	691/752 (92%)	667 (96%)	24 (4%)	36	55

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	73	PHE
1	A	140	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	195	VAL
1	A	249	GLN
1	A	259	GLU
2	B	48	LYS
2	B	57	SER
3	D	3	ASN
3	D	122	ARG
3	D	129	LYS
3	D	141	THR
3	D	142	ASN
3	D	145	GLN
3	D	149	SER
3	D	157	CYS
3	D	168	SER
3	D	170	SER
3	D	175	SER
4	E	51	SER
4	E	65	ASN
4	E	77	ARG
4	E	182	ASN
4	E	191	ARG
4	E	240	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	70	GLN
1	A	90	HIS
1	A	93	GLN
1	A	137	HIS
1	A	141	GLN
1	A	223	GLN
2	B	17	ASN
3	D	37	GLN
3	D	112	GLN
3	D	113	ASN
3	D	140	GLN
4	E	17	GLN
4	E	105	ASN
4	E	137	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 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

### 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	260/290 (89%)	-0.08	9 (3%) 44 43	40, 55, 99, 120	0
2	B	97/100 (97%)	-0.07	2 (2%) 63 61	43, 64, 97, 120	0
3	D	182/205 (88%)	0.49	14 (7%) 13 12	39, 60, 112, 125	0
4	E	240/262 (91%)	-0.10	3 (1%) 77 75	39, 55, 83, 101	0
All	All	779/857 (90%)	0.05	28 (3%) 42 42	39, 56, 101, 125	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	145	GLN	4.9
1	A	0	MET	4.1
3	D	182	CYS	4.0
3	D	180	PHE	3.7
3	D	186	PHE	3.5
1	A	254	TYR	3.4
3	D	185	ALA	3.4
4	E	180	ALA	3.4
1	A	246	LEU	3.3
1	A	221	ILE	3.2
3	D	143	VAL	3.2
2	B	98	ASP	3.1
4	E	175	LEU	3.0
1	A	220	GLU	2.9
3	D	179	ASP	2.7
1	A	249	GLN	2.6
3	D	147	LYS	2.5
3	D	2	GLN	2.5
3	D	184	ASN	2.3
3	D	114	PRO	2.3
1	A	188	ARG	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	0	MET	2.3
3	D	149	SER	2.3
1	A	223	GLN	2.3
4	E	182	ASN	2.2
3	D	187	ASN	2.1
3	D	116	PRO	2.0
1	A	222	VAL	2.0

## 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.