W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 12, 2020 – 11:13 pm BST

PDB ID	:	4L9L
Title	:	Crystal structure of a human Valpha7.2/Vbeta13.2 MAIT TCR in complex with bovine MR1
Authors	:	Lopez-Sagaseta, J.; Adams, E.J.
Deposited on	:	2013-06-18
Resolution	:	3.40 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

1   Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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.
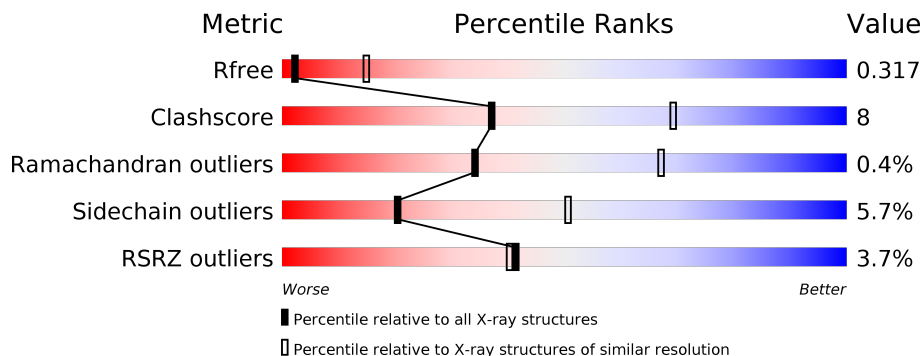
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	392	3% 66% 16% · 16%
2	A	208	3% 82% 12% 6%
3	B	252	4% 70% 20% 5% 5%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6041 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 Beta-2-microglobulin, MHC class I-related protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	C	329	2684	1723	466	482	13	0	0	0

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	99	GLY	-	LINKER	UNP P01888
C	100	GLY	-	LINKER	UNP P01888
C	101	GLY	-	LINKER	UNP P01888
C	102	GLY	-	LINKER	UNP P01888
C	103	SER	-	LINKER	UNP P01888
C	104	GLY	-	LINKER	UNP P01888
C	105	GLY	-	LINKER	UNP P01888
C	106	SER	-	LINKER	UNP P01888
C	107	GLY	-	LINKER	UNP P01888
C	108	SER	-	LINKER	UNP P01888
C	109	GLY	-	LINKER	UNP P01888
C	110	GLY	-	LINKER	UNP P01888
C	111	GLY	-	LINKER	UNP P01888
C	112	GLY	-	LINKER	UNP P01888
C	113	SER	-	LINKER	UNP P01888

- Molecule 2 is a protein called Human MAIT TCR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	196	1492	945	239	300	8	0	0	0

- Molecule 3 is a protein called Human MAIT TCR beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	239	Total 1854	C 1178	N 317	O 350	S 9	0	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0

- Molecule 5 is water.

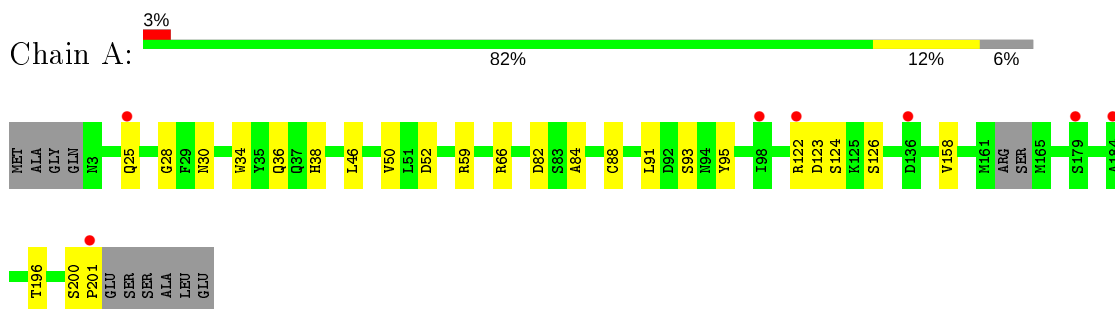| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	1	Total 1	O 1	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

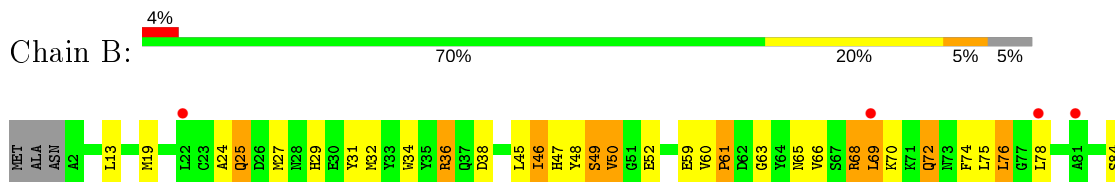- Molecule 1: Beta-2-microglobulin, MHC class I-related protein
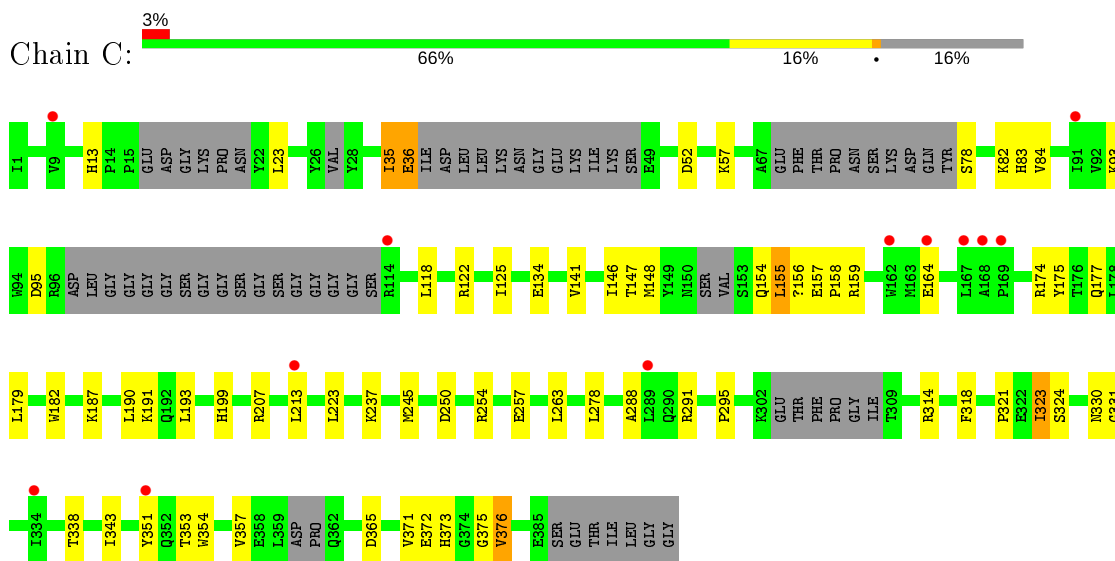
- Molecule 2: Human MAIT TCR alpha chain

- Molecule 3: Human MAIT TCR beta chain

Q85 T86 S87 V88 A92 S93 S94 P97 D98 I103 T112 V123 VAL GLU ASP L127 V137 F138 I145 L153 V154 D163 H164 V165 E166 L167 W170 T182 D183 P184 L193 N194 L200 S201 S202 R203 A208 T209 F210 N216 Q221 V222 Q223 W233

T234 Q235 A253 ASP SER ALA ALA LEU GLU
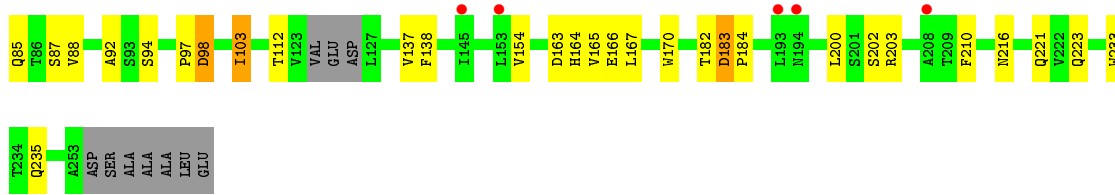
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	82.77Å 86.98Å 156.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.58 – 3.40 47.58 – 3.39	Depositor EDS
% Data completeness (in resolution range)	97.3 (47.58-3.40) 97.5 (47.58-3.39)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	6.58 (at 3.40Å)	Xtriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, R_{free}	0.254 , 0.312 0.257 , 0.317	Depositor DCC
R_{free} test set	806 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	40.0	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 66.7	EDS
L-test for twinning[2]	$<|L|> = 0.46, <L^2> = 0.29$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o,F_c correlation	0.76	EDS
Total number of atoms	6041	wwPDB-VP
Average B, all atoms (Å^2)	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.74% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<|L|>$, $<L^2>$ 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: SO4, KFP

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	0.29	0/2740	0.46	1/3717 (0.0%)
2	A	0.27	0/1526	0.46	0/2078
3	B	0.29	0/1908	0.51	0/2606
All	All	0.28	0/6174	0.48	1/8401 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	157	GLU	C-N-CD	5.83	140.65	128.40

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	C	2684	0	2463	36	0
2	A	1492	0	1374	16	0
3	B	1854	0	1730	45	0
4	A	5	0	0	1	0
4	B	5	0	0	0	0
5	C	1	0	0	0	0
All	All	6041	0	5567	91	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 8.

The worst 5 of 91 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:B:92:ALA:HB1	3:B:103:ILE:HD11	1.67	0.76
1:C:36:GLU:O	1:C:82:LYS:N	2.12	0.71
1:C:122:ARG:HB2	1:C:207:ARG:HB3	1.74	0.69
3:B:59:GLU:HG2	3:B:60:VAL:HG13	1.74	0.68
3:B:163:ASP:O	3:B:164:HIS:ND1	2.30	0.64

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	310/392 (79%)	300 (97%)	9 (3%)	1 (0%)	41	72
2	A	192/208 (92%)	183 (95%)	9 (5%)	0	100	100
3	B	235/252 (93%)	225 (96%)	8 (3%)	2 (1%)	17	49
All	All	737/852 (86%)	708 (96%)	26 (4%)	3 (0%)	34	67

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	61	PRO
3	B	69	LEU
1	C	376	VAL

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	C	273/341 (80%)	256 (94%)	17 (6%)	18	48
2	A	161/183 (88%)	158 (98%)	3 (2%)	57	78
3	B	195/213 (92%)	179 (92%)	16 (8%)	11	37
All	All	629/737 (85%)	593 (94%)	36 (6%)	20	50

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	323	ILE
3	B	25	GLN
3	B	137	VAL
2	A	50	VAL
3	B	36	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	373	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

1 non-standard protein/DNA/RNA residue is 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	$\#|Z| > 2$
1	KFP	C	156	1	22,23,24	2.15	6 (27%)	22,30,32	2.61	11 (50%)

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
1	KFP	C	156	1	-	0/10/11/13	0/2/2/2

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	156	KFP	CAF-C6	-5.81	1.39	1.51
1	C	156	KFP	C7-N8	4.91	1.39	1.31
1	C	156	KFP	C8A-N8	-3.36	1.32	1.37
1	C	156	KFP	C4-C4A	-2.65	1.36	1.41
1	C	156	KFP	CAF-NAL	-2.30	1.35	1.46

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	156	KFP	N1-C2-N3	-4.86	120.73	127.22
1	C	156	KFP	C7-C6-N5	-4.49	117.92	120.85
1	C	156	KFP	C4A-C4-N3	-3.95	118.03	123.43
1	C	156	KFP	C2-N1-C8A	3.90	119.81	115.36
1	C	156	KFP	CAF-NAL-CAJ	3.57	125.64	113.41

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	156	KFP	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

2 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
4	SO4	A	301	-	4,4,4	0.14	0	6,6,6	0.09	0
4	SO4	B	301	-	4,4,4	0.16	0	6,6,6	0.12	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	301	SO4	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 (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, 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	328/392 (83%)	0.55	12 (3%)	41	40	24, 97, 130, 152	0
2	A	196/208 (94%)	0.46	7 (3%)	42	42	53, 87, 141, 174	0
3	B	239/252 (94%)	0.56	9 (3%)	40	39	54, 92, 128, 155	0
All	All	763/852 (89%)	0.53	28 (3%)	41	40	24, 92, 132, 174	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	81	ALA	4.5
1	C	213	LEU	3.5
1	C	164	GLU	3.1
1	C	114	ARG	2.9
1	C	167	LEU	2.9

6.2 Non-standard residues in protein, DNA, RNA chains (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(Å^2)	Q<0.9
1	KFP	C	156	22/23	0.85	0.30	82,100,107,112	0

6.3 Carbohydrates (i)

There are no carbohydrates 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($Å^2$)	Q<0.9
4	SO4	A	301	5/5	0.84	0.27	115,116,118,118	0
4	SO4	B	301	5/5	0.91	0.14	86,87,92,92	0

6.5 Other polymers (i)

There are no such residues in this entry.