W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Aug 9, 2020 – 05:23 AM BST

PDB ID	:	6D7G
Title	:	Structure of 5F3 TCR in complex with HLA-A2/MART-1
Authors	:	Roy, S.; Adams, E.J.
Deposited on	:	2018-04-24
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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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	:	2.13.1
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)	:	2.13.1

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.
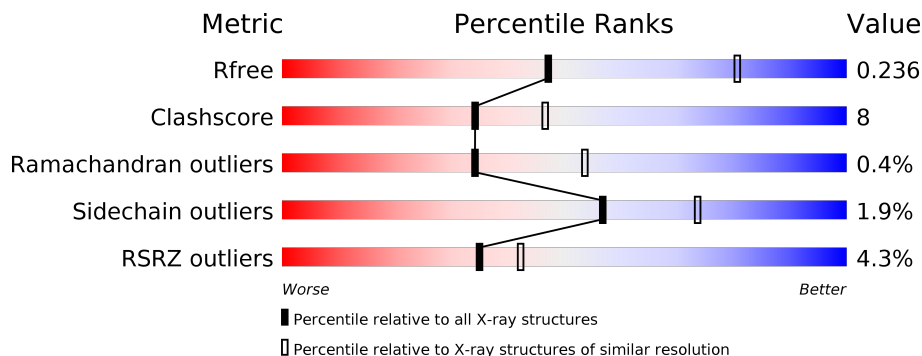
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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 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	A	427	% 76% 14% 10%
2	D	229	5% 68% 18% 14%
3	E	258	8% 59% 20% • 19%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6022 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 MART1 PEPTIDE-BETA-2-MICROGLOBULIN-HLA-A*02 CHIMERA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	386	3023	1907	535	571	10	0	0	0

There are 43 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	86	GLY	-	linker	PDB ?
A	87	GLY	-	linker	PDB ?
A	88	GLY	-	linker	PDB ?
A	89	GLY	-	linker	PDB ?
A	90	SER	-	linker	PDB ?
A	91	GLY	-	linker	PDB ?
A	92	GLY	-	linker	PDB ?
A	93	GLY	-	linker	PDB ?
A	94	GLY	-	linker	PDB ?
A	95	SER	-	linker	PDB ?
A	96	GLY	-	linker	PDB ?
A	97	GLY	-	linker	PDB ?
A	98	GLY	-	linker	PDB ?
A	99	GLY	-	linker	PDB ?
A	100	SER	-	linker	PDB ?
A	981	GLY	-	linker	UNP P61769
A	982	GLY	-	linker	UNP P61769
A	983	GLY	-	linker	UNP P61769
A	984	GLY	-	linker	UNP P61769
A	985	SER	-	linker	UNP P61769
A	986	GLY	-	linker	UNP P61769
A	987	GLY	-	linker	UNP P61769
A	988	GLY	-	linker	UNP P61769
A	989	GLY	-	linker	UNP P61769
A	990	SER	-	linker	UNP P61769
A	991	GLY	-	linker	UNP P61769

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	992	GLY	-	linker	UNP P61769
A	993	GLY	-	linker	UNP P61769
A	994	GLY	-	linker	UNP P61769
A	995	SER	-	linker	UNP P61769
A	996	GLY	-	linker	UNP P61769
A	997	GLY	-	linker	UNP P61769
A	998	GLY	-	linker	UNP P61769
A	999	GLY	-	linker	UNP P61769
A	1000	SER	-	linker	UNP P61769
A	1276	HIS	-	expression tag	UNP P01892
A	1277	HIS	-	expression tag	UNP P01892
A	1278	HIS	-	expression tag	UNP P01892
A	1279	HIS	-	expression tag	UNP P01892
A	1280	HIS	-	expression tag	UNP P01892
A	1281	HIS	-	expression tag	UNP P01892
A	1282	HIS	-	expression tag	UNP P01892
A	1283	HIS	-	expression tag	UNP P01892

- Molecule 2 is a protein called TRA@ protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	198	1489	949	252	281	7	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-2	ALA	-	expression tag	UNP Q6PJ56
D	-1	ASP	-	expression tag	UNP Q6PJ56
D	0	PRO	-	expression tag	UNP Q6PJ56
D	96	LEU	-	linker	UNP Q6PJ56
D	97	GLY	-	linker	UNP Q6PJ56
D	98	TRP	-	linker	UNP Q6PJ56
D	99	ASP	-	linker	UNP Q6PJ56

- Molecule 3 is a protein called T-CELL RECEPTOR GAMMA VARIABLE 8,T-CELL RE-CEPTOR GAMMA-2 CHAIN C REGION.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	E	208	1493	952	254	281	6	0	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	O	0	0
			1	1		
5	D	1	Total	O	0	0
			1	1		
5	E	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.
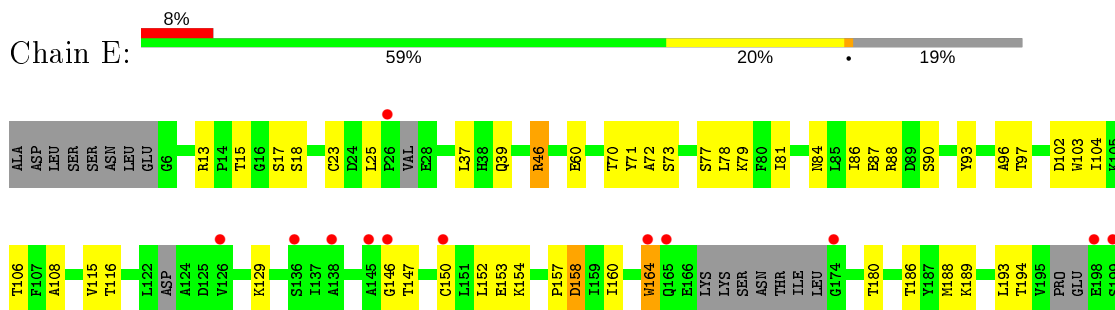
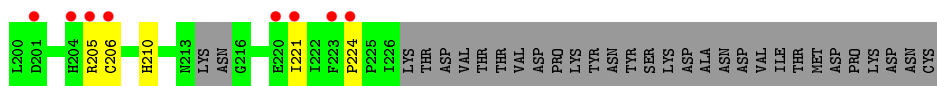- Molecule 1: MART1 PEPTIDE-BETA-2-MICROGLOBULIN-HLA-A*02 CHIMERA
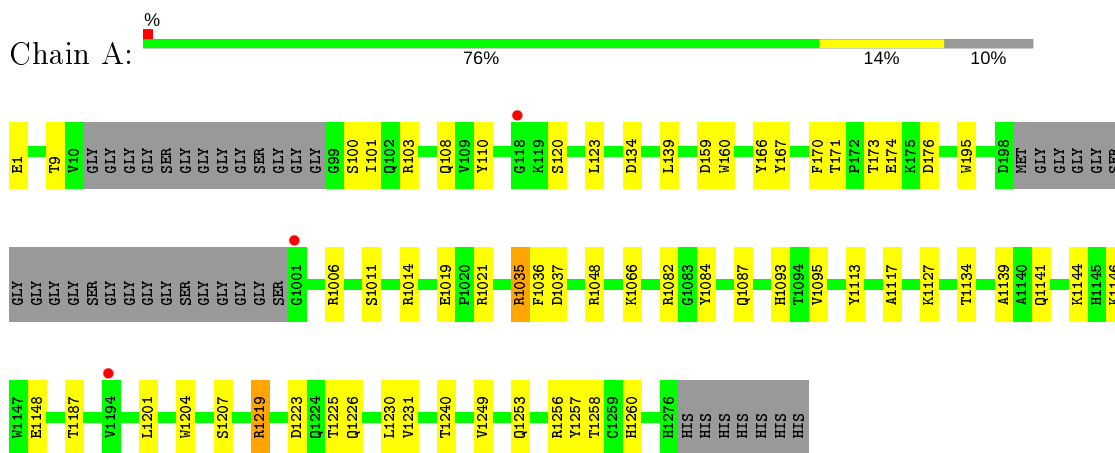
- Molecule 2: TRA@ protein
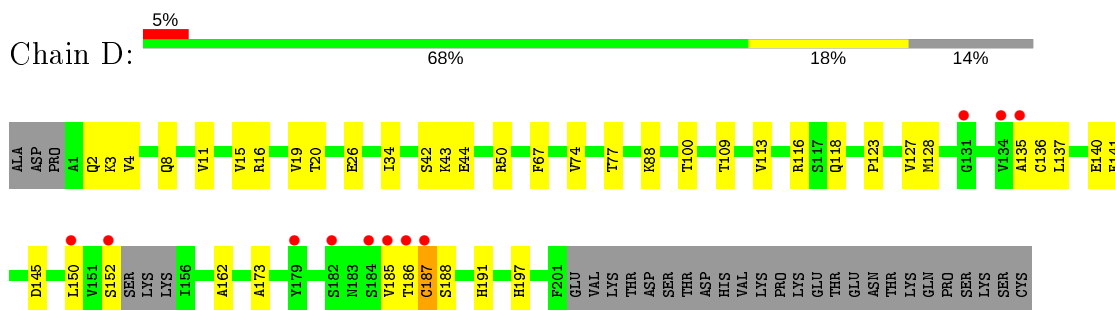
- Molecule 3: T-CELL RECEPTOR GAMMA VARIABLE 8,T-CELL RECEPTOR GAMMA-2 CHAIN C REGION

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 62	Depositor
Cell constants a, b, c, α, β, γ	123.95Å 123.95Å 154.16Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.30 – 2.75 48.30 – 2.75	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.30-2.75) 100.0 (48.30-2.75)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.80 (at 2.77Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, R_{free}	0.211 , 0.237 0.211 , 0.236	Depositor DCC
R_{free} test set	1806 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å^2)	67.2	Xtriage
Anisotropy	0.088	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 45.0	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.039 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6022	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.13% 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: NAG

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	0.61	0/3109	0.65	1/4242 (0.0%)
2	D	0.72	1/1522 (0.1%)	0.92	3/2072 (0.1%)
3	E	0.54	0/1527	0.66	0/2092
All	All	0.62	1/6158 (0.0%)	0.73	4/8406 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	187	CYS	CB-SG	-5.86	1.72	1.81

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	137	LEU	CA-CB-CG	8.65	135.20	115.30
2	D	113	VAL	C-N-CA	6.66	138.35	121.70
1	A	1230	LEU	CB-CG-CD2	-5.68	101.34	111.00
2	D	137	LEU	CB-CG-CD2	5.12	119.70	111.00

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	3023	0	2741	37	0
2	D	1489	0	1414	27	0
3	E	1493	0	1287	36	0
4	D	14	0	13	0	0
5	A	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
All	All	6022	0	5455	96	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 96 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:E:147:THR:HA	3:E:193:LEU:O	1.72	0.90
2:D:42:SER:OG	2:D:44:GLU:OE1	1.98	0.81
3:E:25:LEU:HB2	3:E:108:ALA:HB2	1.65	0.79
3:E:71:TYR:HE2	3:E:81:ILE:HD11	1.50	0.77
1:A:1219:ARG:NH2	1:A:1257:TYR:OH	2.17	0.77

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	380/427 (89%)	370 (97%)	9 (2%)	1 (0%)	41	60
2	D	194/229 (85%)	181 (93%)	12 (6%)	1 (0%)	29	47
3	E	196/258 (76%)	178 (91%)	17 (9%)	1 (0%)	29	47
All	All	770/914 (84%)	729 (95%)	38 (5%)	3 (0%)	34	53

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	101	ILE
3	E	224	PRO
2	D	140	GLU

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	301/347 (87%)	295 (98%)	6 (2%)	55	72
2	D	155/206 (75%)	155 (100%)	0	100	100
3	E	136/235 (58%)	131 (96%)	5 (4%)	34	54
All	All	592/788 (75%)	581 (98%)	11 (2%)	57	73

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1141	GLN
1	A	1219	ARG
3	E	90	SER
1	A	1082	ARG
3	E	84	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	82	GLN
2	D	133	ASN

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)

1 ligand 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	$\#\|Z\| > 2$
4	NAG	D	301	2	14,14,15	0.20	0	17,19,21	0.54	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
4	NAG	D	301	2	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	301	NAG	O5-C5-C6-O6

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 (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	A	386/427 (90%)	0.07	3 (0%)	86 90	31, 56, 93, 121	0
2	D	198/229 (86%)	0.30	11 (5%)	24 29	41, 57, 127, 154	0
3	E	208/258 (80%)	0.47	20 (9%)	8 9	39, 81, 155, 175	0
All	All	792/914 (86%)	0.23	34 (4%)	35 42	31, 59, 131, 175	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	E	145	ALA	5.9
3	E	220	GLU	4.4
2	D	179	TYR	3.8
3	E	126	VAL	3.7
3	E	204	HIS	3.6

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)

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	NAG	D	301	14/15	0.73	0.27	53,74,104,104	0

6.5 Other polymers (i)

There are no such residues in this entry.