wwPDB X-ray Structure Validation Summary Report (i)

Oct 25, 2022 – 04:03 PM EDT

PDB ID	:	7RK7
Title	:	The complex between TIL 1383i TCR and human Class I MHC HLA-A2 with the bound Tyrosinase(369-377)(N371D) nonameric peptide
Authors	:	Singh, N.K.; Davancaze, L.M.; Arbuiso, A.; Weiss, L.I.; Keller, G.L.J.; Baker, B.M.
Deposited on	:	2021-07-22
Resolution	:	2.54 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

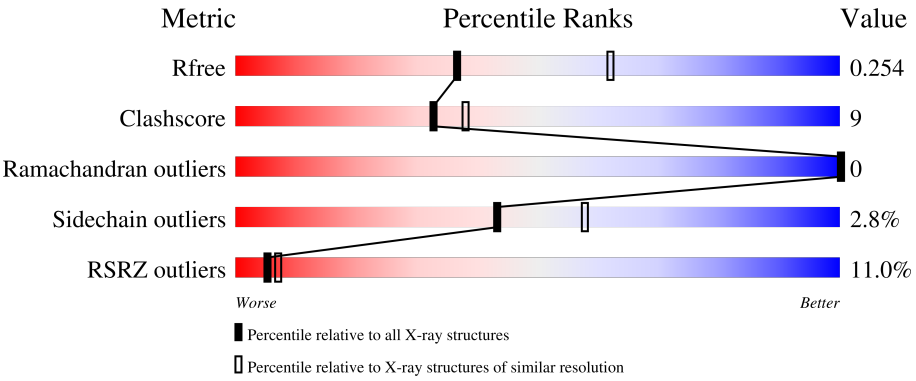
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 2.54 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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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 <=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	275	19% 64% 21% · 13%
2	B	100	27% 60% 29% · 10%
3	C	9	67% 22% 11%
4	D	214	% 77% 11% 12%
5	E	253	2% 77% 20% ·

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6295 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, A alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	238	Total	C	N	O	S	0	0	0
			1944	1217	352	367	8			

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	62	GLY	GLN	conflict	UNP P04439
A	66	LYS	ASN	conflict	UNP P04439
A	70	HIS	GLN	conflict	UNP P04439
A	74	HIS	ASP	conflict	UNP P04439
A	95	VAL	ILE	conflict	UNP P04439
A	97	ARG	ILE	conflict	UNP P04439
A	107	TRP	GLY	conflict	UNP P04439
A	114	HIS	ARG	conflict	UNP P04439
A	116	TYR	ASP	conflict	UNP P04439
A	127	LYS	ASN	conflict	UNP P04439
A	142	THR	ILE	conflict	UNP P04439
A	145	HIS	ARG	conflict	UNP P04439
A	152	VAL	GLU	conflict	UNP P04439
A	161	GLU	ASP	conflict	UNP P04439
A	184	ALA	PRO	conflict	UNP P04439
A	193	ALA	PRO	conflict	UNP P04439
A	194	VAL	ILE	conflict	UNP P04439
A	207	SER	GLY	conflict	UNP P04439
A	253	GLN	GLU	conflict	UNP P04439

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	90	Total	C	N	O	S	0	0	0
			754	485	130	136	3			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called Tyrosinase peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	S	0	0	0
			70	42	10	16	2			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	3	ASP	ASN	engineered mutation	UNP P14679

- Molecule 4 is a protein called TIL1383i (h3T) T cell receptor alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	189	Total	C	N	O	S	0	0	0
			1471	923	246	294	8			

- Molecule 5 is a protein called TIL1383i (h3T) T cell receptor beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	247	Total	C	N	O	S	0	0	0
			1974	1244	340	383	7			

- Molecule 6 is water.

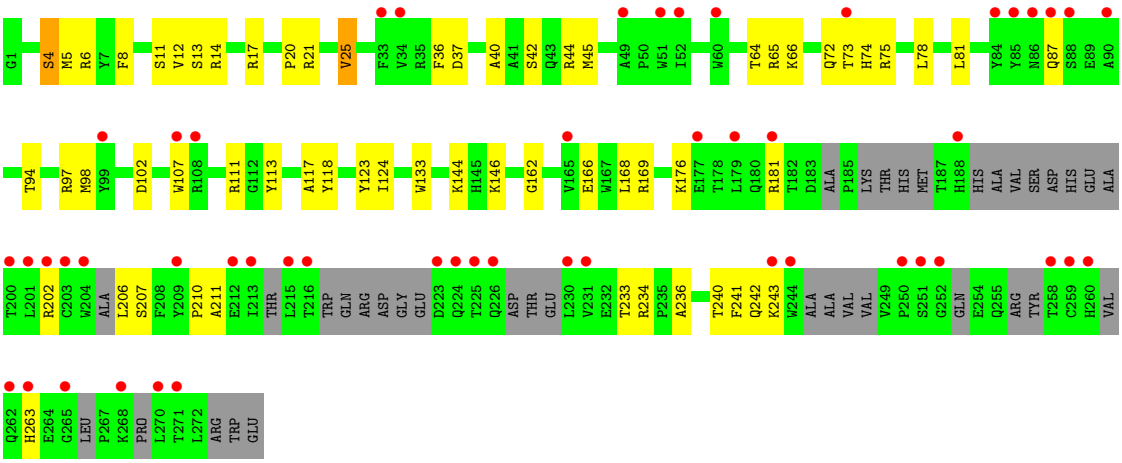
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	7	Total	O	0	0
			7	7		
6	B	1	Total	O	0	0
			1	1		
6	C	1	Total	O	0	0
			1	1		
6	D	34	Total	O	0	0
			34	34		
6	E	39	Total	O	0	0
			39	39		

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.

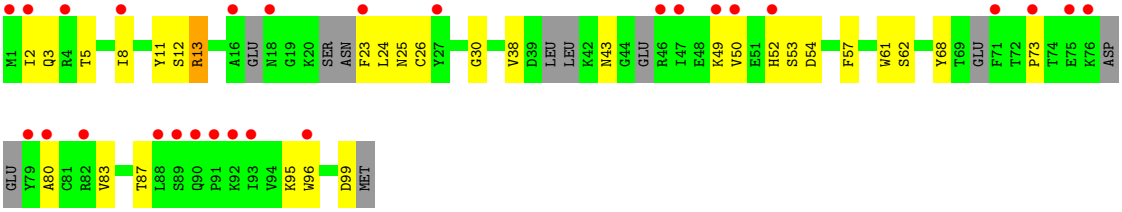
• Molecule 1: HLA class I histocompatibility antigen, A alpha chain

Chain A:



• Molecule 2: Beta-2-microglobulin

Chain B:

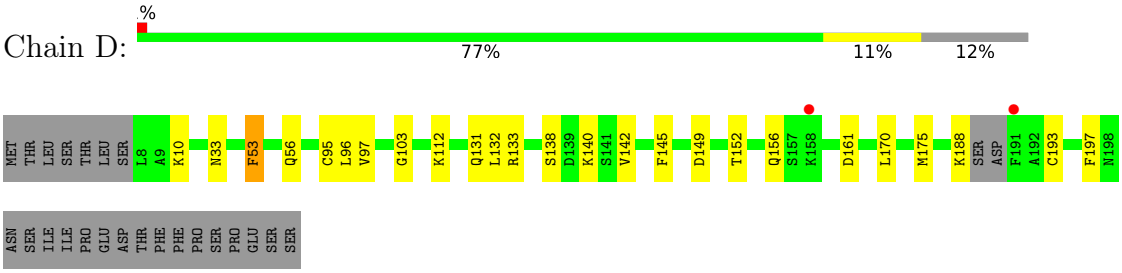


• Molecule 3: Tyrosinase peptide

Chain C:

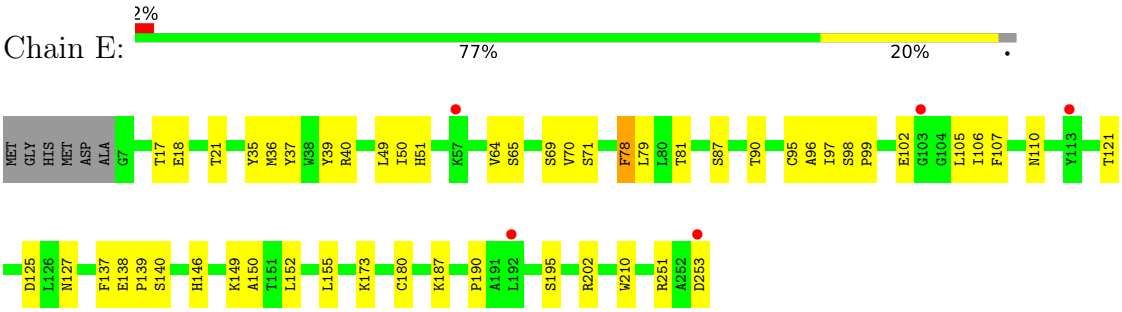
• Molecule 4: TIL1383i (h3T) T cell receptor alpha chain

Chain D:



• Molecule 5: TIL1383i (h3T) T cell receptor beta chain

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	114.74Å 74.43Å 131.06Å 90.00° 93.49° 90.00°	Depositor
Resolution (Å)	37.21 – 2.54 37.21 – 2.54	Depositor EDS
% Data completeness (in resolution range)	99.5 (37.21-2.54) 99.5 (37.21-2.54)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.08 (at 2.54Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.211 , 0.255 0.210 , 0.254	Depositor DCC
R_{free} test set	2006 reflections (5.51%)	wwPDB-VP
Wilson B-factor (Å^2)	56.5	Xtriage
Anisotropy	0.416	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 51.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6295	wwPDB-VP
Average B, all atoms (Å^2)	89.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.04% 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

5.1 Standard geometry

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	0.30	0/1987	0.57	0/2671
2	B	0.46	1/771 (0.1%)	0.80	2/1034 (0.2%)
3	C	0.24	0/70	0.54	0/91
4	D	0.33	1/1500 (0.1%)	0.58	0/2034
5	E	0.28	0/2031	0.57	0/2767
All	All	0.32	2/6359 (0.0%)	0.60	2/8597 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	53	PHE	C-N	-5.79	1.20	1.34
2	B	13	ARG	C-N	-5.22	1.22	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	43	ASN	O-C-N	5.69	132.88	123.20
2	B	49	LYS	N-CA-C	-5.39	96.45	111.00

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	A	1944	0	1811	47	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	754	0	727	23	0
3	C	70	0	65	3	0
4	D	1471	0	1424	16	3
5	E	1974	0	1857	37	0
6	A	7	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	D	34	0	0	1	0
6	E	39	0	0	0	0
All	All	6295	0	5884	105	3

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 9.

The worst 5 of 105 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:B:8:ILE:HD11	2:B:26:CYS:SG	1.75	1.24
1:A:234:ARG:NH2	2:B:11:TYR:HB3	1.65	1.12
5:E:69:SER:OG	5:E:81:THR:HB	1.51	1.08
1:A:234:ARG:HH22	2:B:11:TYR:HB3	1.39	0.82
1:A:234:ARG:NH2	2:B:11:TYR:CB	2.48	0.77

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:161:ASP:OD1	4:D:188:LYS:NZ[2_455]	1.14	1.06
4:D:161:ASP:CG	4:D:188:LYS:NZ[2_455]	1.96	0.24
4:D:161:ASP:OD1	4:D:188:LYS:CE[2_455]	2.15	0.05

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	211/275 (77%)	200 (95%)	11 (5%)	0	100	100
2	B	76/100 (76%)	76 (100%)	0	0	100	100
3	C	7/9 (78%)	7 (100%)	0	0	100	100
4	D	185/214 (86%)	179 (97%)	6 (3%)	0	100	100
5	E	245/253 (97%)	244 (100%)	1 (0%)	0	100	100
All	All	724/851 (85%)	706 (98%)	18 (2%)	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	201/231 (87%)	195 (97%)	6 (3%)	41	55
2	B	85/95 (90%)	82 (96%)	3 (4%)	36	49
3	C	8/8 (100%)	7 (88%)	1 (12%)	4	4
4	D	169/194 (87%)	166 (98%)	3 (2%)	59	74
5	E	218/222 (98%)	212 (97%)	6 (3%)	43	58
All	All	681/750 (91%)	662 (97%)	19 (3%)	43	58

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	87	SER
5	E	152	LEU
5	E	202	ARG
5	E	146	HIS
2	B	13	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	263	HIS
2	B	52	HIS
5	E	51	HIS
1	A	87	GLN
1	A	74	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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^{th} 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	238/275 (86%)	1.22	51 (21%) 0 0	50, 109, 164, 192	0
2	B	90/100 (90%)	1.70	27 (30%) 0 0	87, 136, 176, 189	0
3	C	9/9 (100%)	1.18	0 100 100	57, 66, 81, 93	0
4	D	189/214 (88%)	0.16	2 (1%) 80 85	37, 57, 129, 174	0
5	E	247/253 (97%)	0.11	5 (2%) 65 72	40, 66, 103, 144	0
All	All	773/851 (90%)	0.66	85 (10%) 5 7	37, 81, 157, 192	0

The worst 5 of 85 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	230	LEU	10.5
1	A	216	THR	8.5
1	A	271	THR	7.8
4	D	191	PHE	7.4
2	B	71	PHE	7.2

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.