



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 18, 2017 – 02:03 PM EDT

PDB ID : 5JZI
Title : Crystal structure of 1406 TCR bound to HLA-A2 with HCV 1406-1415 antigen peptide
Authors : Wang, Y.; Piepenbrink, K.H.; Baker, B.M.
Deposited on : 2016-05-16
Resolution : 2.50 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029077
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20029077

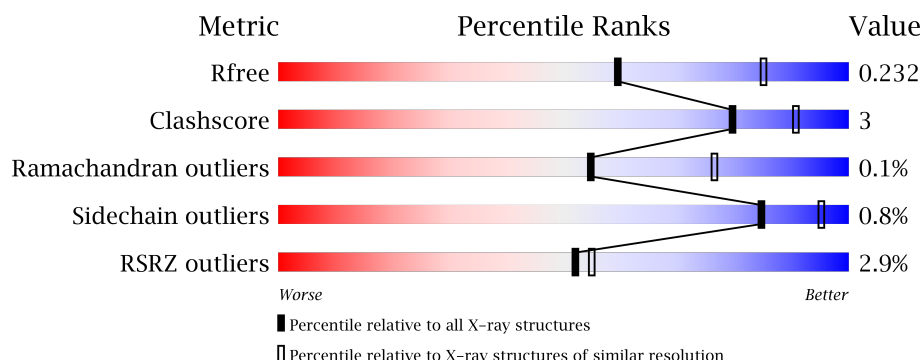
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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}	100719	3846 (2.50-2.50)
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (2.50-2.50)

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. 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	275	<div> <div>7%</div> <div>87%</div> <div>7%</div> <div>5%</div> </div>
1	F	275	<div> <div>6%</div> <div>83%</div> <div>12%</div> <div>5%</div> </div>
2	B	100	<div> <div>%</div> <div>88%</div> <div>11%</div> <div>.</div> </div>
2	G	100	<div> <div>9%</div> <div>86%</div> <div>12%</div> <div>.</div> </div>
3	C	10	<div> <div>80%</div> <div>20%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	10	 100%
4	D	211	 88%9%
4	I	211	 91%7%
5	E	245	 91%7%
5	J	245	 91%7%

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 25825 atoms, of which 12500 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-2 alpha chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	260	Total	C	H	N	O	S	0	0	0
			4082	1326	1962	385	400	9			
1	F	262	Total	C	H	N	O	S	0	0	0
			4161	1349	2002	396	406	8			

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	99	Total	C	H	N	O	S	0	0	0
			1618	524	793	140	157	4			
2	G	98	Total	C	H	N	O	S	0	0	0
			1607	521	788	139	155	4			

There are 2 discrepancies between the modelled and reference sequences:

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

- Molecule 3 is a protein called KLV peptide.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	10	Total	C	H	N	O		0	0	0
			155	46	85	12	12				
3	H	10	Total	C	H	N	O		0	0	0
			155	46	85	12	12				

- Molecule 4 is a protein called HCV1406 TCR alpha chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	D	207	Total	C	H	N	O	S	0	0	0
			3207	1033	1560	268	334	12			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	I	207	Total	C	H	N	O	S	0	0	0
			3208	1034	1560	268	334	12			

- Molecule 5 is a protein called HCV1406 TCR beta chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
5	E	240	Total	C	H	N	O	S	0	0	0
			3762	1215	1831	336	371	9			
5	J	241	Total	C	H	N	O	S	0	0	0
			3773	1219	1834	337	374	9			

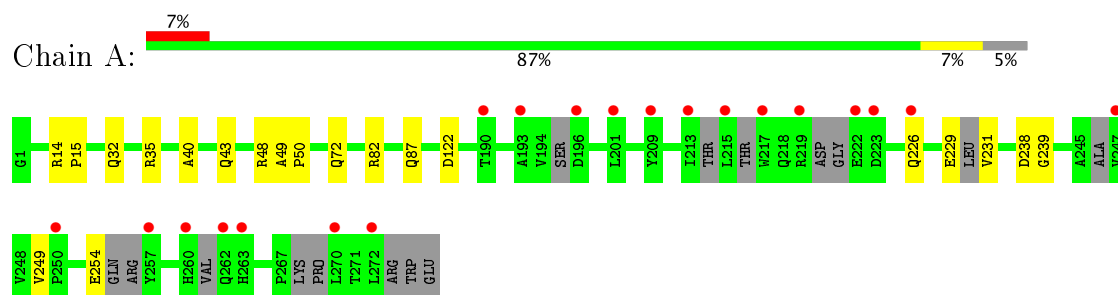
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	14	Total	O	0	0
			14	14		
6	B	4	Total	O	0	0
			4	4		
6	F	11	Total	O	0	0
			11	11		
6	G	2	Total	O	0	0
			2	2		
6	H	1	Total	O	0	0
			1	1		
6	D	19	Total	O	0	0
			19	19		
6	E	20	Total	O	0	0
			20	20		
6	I	10	Total	O	0	0
			10	10		
6	J	16	Total	O	0	0
			16	16		

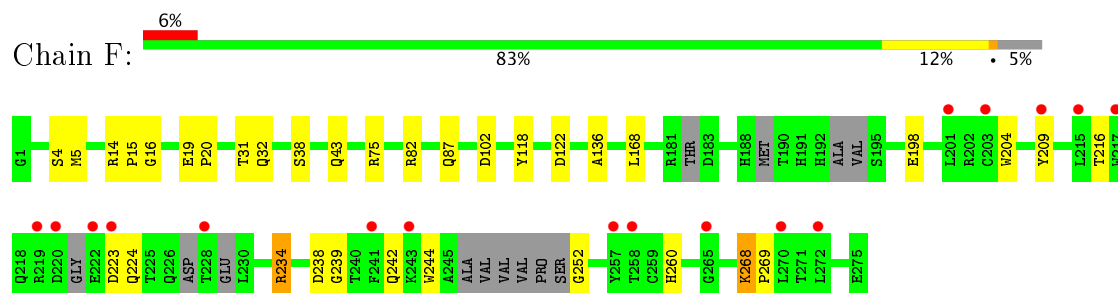
3 Residue-property plots [i](#)

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.

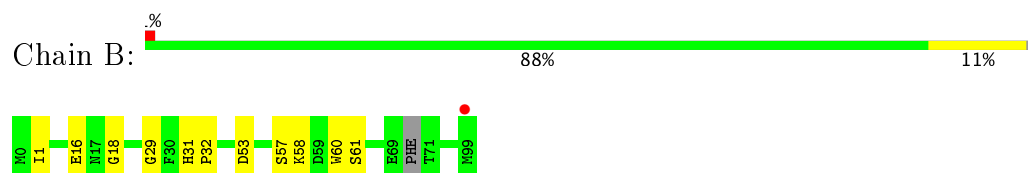
- Molecule 1: HLA class I histocompatibility antigen, A-2 alpha chain



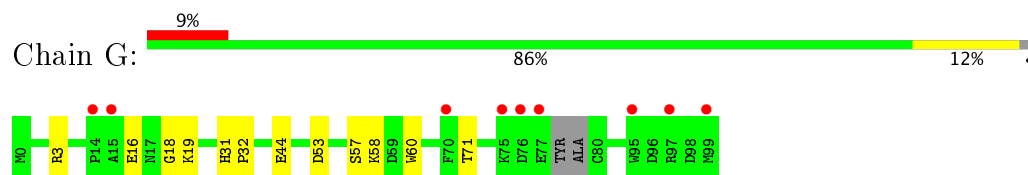
- Molecule 1: HLA class I histocompatibility antigen, A-2 alpha chain



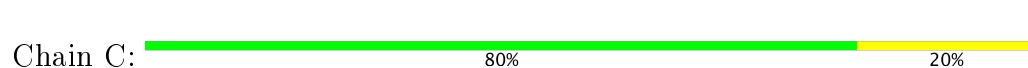
- Molecule 2: Beta-2-microglobulin

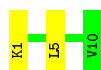


- Molecule 2: Beta-2-microglobulin



- Molecule 3: KLV peptide





- Molecule 3: KLV peptide

Chain H: 100%

There are no outlier residues recorded for this chain.

- Molecule 4: HCV1406 TCR alpha chain

Chain D: 88% 9%



- Molecule 4: HCV1406 TCR alpha chain

Chain I: 91% 7%



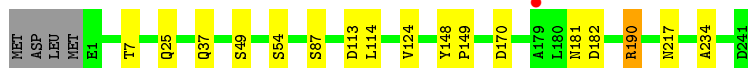
- Molecule 5: HCV1406 TCR beta chain

Chain E: 91% 7%



- Molecule 5: HCV1406 TCR beta chain

Chain J: 91% 7%



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	128.76 Å 128.76 Å 223.59 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	37.17 – 2.50 37.17 – 2.08	Depositor EDS
% Data completeness (in resolution range)	99.6 (37.17-2.50) 81.4 (37.17-2.08)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.01 (at 2.08 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.196 , 0.232 0.195 , 0.232	Depositor DCC
R_{free} test set	1386 reflections (1.86%)	DCC
Wilson B-factor (Å ²)	41.2	Xtriage
Anisotropy	0.032	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 28.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.36$, $\langle L^2 \rangle = 0.19$	Xtriage
Estimated twinning fraction	0.459 for -h,-k,l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	25825	wwPDB-VP
Average B, all atoms (Å ²)	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.09% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.27	0/2173	0.43	0/2934
1	F	0.26	0/2216	0.43	0/2994
2	B	0.26	0/846	0.41	0/1143
2	G	0.26	0/840	0.44	0/1134
3	C	0.29	0/69	0.44	0/91
3	H	0.33	0/69	0.50	0/91
4	D	0.27	0/1682	0.43	0/2274
4	I	0.26	0/1683	0.43	0/2276
5	E	0.27	0/1986	0.42	0/2704
5	J	0.28	0/1994	0.43	0/2715
All	All	0.27	0/13558	0.43	0/18356

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	2120	1962	1963	11	0
1	F	2159	2002	2002	21	1
2	B	825	793	793	7	0
2	G	819	788	788	8	0
3	C	70	85	85	1	0
3	H	70	85	85	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	1647	1560	1560	15	0
4	I	1648	1560	1560	9	0
5	E	1931	1831	1833	9	0
5	J	1939	1834	1837	7	1
6	A	14	0	0	1	0
6	B	4	0	0	0	0
6	D	19	0	0	3	0
6	E	20	0	0	1	0
6	F	11	0	0	0	0
6	G	2	0	0	0	0
6	H	1	0	0	0	0
6	I	10	0	0	0	0
6	J	16	0	0	0	0
All	All	13325	12500	12506	81	1

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.

The worst 5 of 81 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)
4:D:131:ASP:HB3	4:D:135:SER:HB3	1.70	0.74
5:E:107:ARG:NH2	5:E:150:ASP:OD2	2.21	0.73
4:I:112:ARG:NH1	4:I:113:LEU:O	2.24	0.71
5:E:188:SER:OG	5:E:190:ARG:NH1	2.24	0.70
1:F:122:ASP:OD1	2:G:60:TRP:NE1	2.25	0.69

All (1) 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)
1:F:136:ALA:O	5:J:217:ASN:ND2[6_555]	2.19	0.01

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	241/275 (88%)	226 (94%)	14 (6%)	1 (0%)	38	59
1	F	247/275 (90%)	234 (95%)	13 (5%)	0	100	100
2	B	95/100 (95%)	88 (93%)	7 (7%)	0	100	100
2	G	94/100 (94%)	89 (95%)	5 (5%)	0	100	100
3	C	8/10 (80%)	8 (100%)	0	0	100	100
3	H	8/10 (80%)	7 (88%)	1 (12%)	0	100	100
4	D	203/211 (96%)	190 (94%)	12 (6%)	1 (0%)	32	53
4	I	203/211 (96%)	190 (94%)	13 (6%)	0	100	100
5	E	238/245 (97%)	225 (94%)	13 (6%)	0	100	100
5	J	239/245 (98%)	226 (95%)	13 (5%)	0	100	100
All	All	1576/1682 (94%)	1483 (94%)	91 (6%)	2 (0%)	55	76

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	226	GLN
4	D	207	SER

5.3.2 Protein sidechains ⓘ

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	217/231 (94%)	216 (100%)	1 (0%)	91	97
1	F	221/231 (96%)	219 (99%)	2 (1%)	82	94
2	B	94/95 (99%)	94 (100%)	0	100	100
2	G	94/95 (99%)	93 (99%)	1 (1%)	78	92
3	C	7/7 (100%)	6 (86%)	1 (14%)	4	7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	H	7/7 (100%)	7 (100%)	0	100	100
4	D	188/192 (98%)	187 (100%)	1 (0%)	91	97
4	I	188/192 (98%)	188 (100%)	0	100	100
5	E	213/218 (98%)	210 (99%)	3 (1%)	71	90
5	J	214/218 (98%)	211 (99%)	3 (1%)	71	90
All	All	1443/1486 (97%)	1431 (99%)	12 (1%)	85	95

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	207	SER
5	E	7	THR
5	J	7	THR
2	G	44	GLU
5	E	61	SER

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

Mol	Chain	Res	Type
1	A	226	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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, 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	260/275 (94%)	-0.02	20 (7%) 14 14	35, 53, 126, 133	0
1	F	262/275 (95%)	-0.08	17 (6%) 20 20	41, 61, 127, 134	0
2	B	99/100 (99%)	-0.21	1 (1%) 82 83	50, 84, 107, 114	0
2	G	98/100 (98%)	0.20	9 (9%) 10 9	58, 97, 117, 125	0
3	C	10/10 (100%)	-0.65	0 100 100	37, 38, 41, 42	0
3	H	10/10 (100%)	-0.60	0 100 100	46, 48, 53, 57	0
4	D	207/211 (98%)	-0.55	0 100 100	31, 62, 83, 90	0
4	I	207/211 (98%)	-0.65	0 100 100	37, 53, 72, 85	0
5	E	240/245 (97%)	-0.57	0 100 100	29, 47, 66, 76	0
5	J	241/245 (98%)	-0.58	1 (0%) 92 92	32, 45, 62, 86	0
All	All	1634/1682 (97%)	-0.35	48 (2%) 52 55	29, 55, 118, 134	0

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	272	LEU	6.7
1	A	257	TYR	6.4
1	A	209	TYR	5.6
1	F	220	ASP	5.4
1	F	217	TRP	4.7

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 [i](#)

There are no carbohydrates 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.