



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 02:11 AM GMT

PDB ID : 2FZ3  
Title : The role of T cell receptor alpha genes in directing human MHC restriction  
Authors : Miles, J.J.; Borg, N.A.  
Deposited on : 2006-02-09  
Resolution : 1.90 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  
<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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

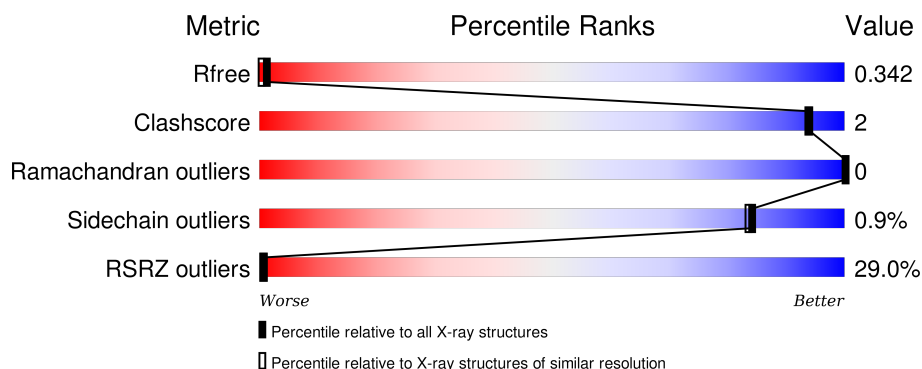
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	91344	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

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  $\geq 3$ , 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	276	<div> <div>26%</div> <div>96%</div> </div>
2	B	99	<div> <div>37%</div> <div>97%</div> </div>
3	C	11	<div> <div>27%</div> <div>100%</div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3376 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, B-35 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	S	0	0	0
			2257	1405	414	431	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	156	ARG	LEU	SEE REMARK 999	UNP P30474

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			829	528	140	158	3			

- Molecule 3 is a protein called 11-mer peptide from Epstein-Barr nuclear antigen 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	11	Total	C	N	O	0	0	0
			87	56	13	18			

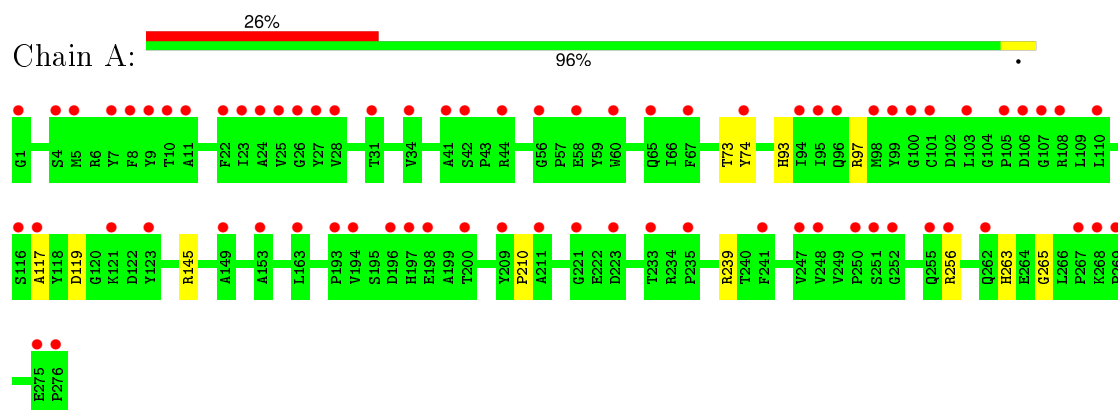
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	153	Total	O	0	0
			153	153		
4	B	42	Total	O	0	0
			42	42		
4	C	8	Total	O	0	0
			8	8		

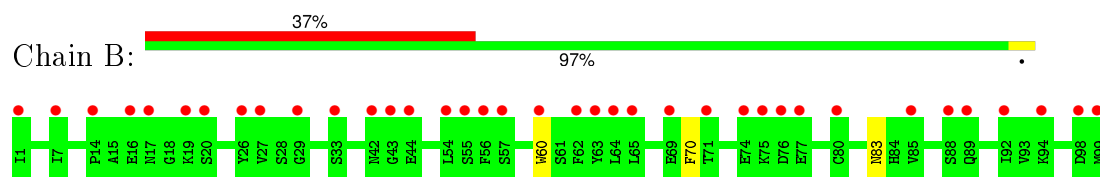
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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.

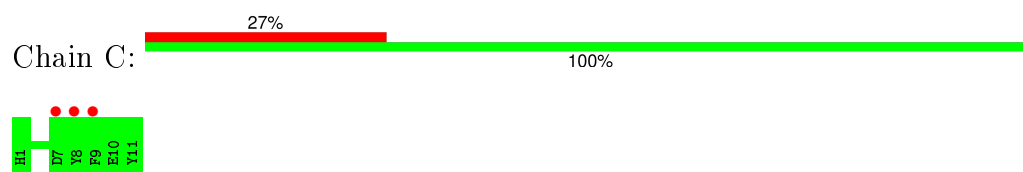
- Molecule 1: HLA class I histocompatibility antigen, B-35 alpha chain



- Molecule 2: Beta-2-microglobulin



- Molecule 3: 11-mer peptide from Epstein-Barr nuclear antigen 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.16Å 82.09Å 109.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.06 – 1.90 34.07 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.6 (34.06-1.90) 98.6 (34.07-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.14 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.232 , 0.275 0.313 , 0.342	Depositor DCC
$R_{free}$ test set	1132 reflections (3.18%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.1	Xtriage
Anisotropy	0.466	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 36.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 36693 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	3376	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.48% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.48	2/2320 (0.1%)	0.65	4/3153 (0.1%)
2	B	0.36	0/852	0.48	0/1152
3	C	0.46	0/90	0.43	0/122
All	All	0.45	2/3262 (0.1%)	0.61	4/4427 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	256	ARG	CZ-NH1	7.01	1.42	1.33
1	A	256	ARG	CZ-NH2	6.94	1.42	1.33

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	256	ARG	NE-CZ-NH2	-17.20	111.70	120.30
1	A	256	ARG	NE-CZ-NH1	6.00	123.30	120.30
1	A	256	ARG	CD-NE-CZ	-5.64	115.70	123.60
1	A	256	ARG	NH1-CZ-NH2	5.07	124.98	119.40

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	A	2257	0	2119	11	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	829	0	794	1	0
3	C	87	0	69	0	0
4	A	153	0	0	5	0
4	B	42	0	0	0	0
4	C	8	0	0	0	0
All	All	3376	0	2982	11	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 2.

All (11) 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)
1:A:93:HIS:HD2	1:A:119:ASP:OD2	1.80	0.65
1:A:263:HIS:CD2	1:A:265:GLY:H	2.17	0.63
1:A:210:PRO:O	1:A:263:HIS:HE1	1.87	0.56
1:A:263:HIS:HD2	1:A:265:GLY:H	1.56	0.54
1:A:93:HIS:HE1	4:A:307:HOH:O	1.90	0.54
1:A:145:ARG:NH1	4:A:312:HOH:O	2.45	0.50
1:A:73:THR:HG22	4:A:385:HOH:O	2.14	0.47
1:A:117:ALA:HB2	2:B:60:TRP:CE2	2.54	0.43
1:A:74:TYR:OH	1:A:97:ARG:HD3	2.19	0.42
1:A:73:THR:CG2	4:A:385:HOH:O	2.68	0.42
1:A:145:ARG:HG3	4:A:338:HOH:O	2.21	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	274/276 (99%)	269 (98%)	5 (2%)	0	<a href="#">100</a> <a href="#">100</a>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	97/99 (98%)	95 (98%)	2 (2%)	0	100	100
3	C	9/11 (82%)	9 (100%)	0	0	100	100
All	All	380/386 (98%)	373 (98%)	7 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	234/234 (100%)	233 (100%)	1 (0%)	93	94
2	B	94/94 (100%)	92 (98%)	2 (2%)	61	55
3	C	8/9 (89%)	8 (100%)	0	100	100
All	All	336/337 (100%)	333 (99%)	3 (1%)	84	83

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	239	ARG
2	B	70	PHE
2	B	83	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	70	ASN
1	A	80	ASN
1	A	93	HIS
1	A	263	HIS
2	B	83	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 carbohydrates 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	276/276 (100%)	1.47	72 (26%)  	23, 28, 37, 43	0
2	B	99/99 (100%)	1.79	37 (37%)  	25, 31, 37, 43	0
3	C	11/11 (100%)	1.61	3 (27%)  	23, 30, 35, 37	0
All	All	386/386 (100%)	1.55	112 (29%)  	23, 29, 37, 43	0

All (112) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	75	LYS	5.5
1	A	105	PRO	5.0
2	B	74	GLU	4.3
1	A	25	VAL	4.3
1	A	197	HIS	4.3
1	A	193	PRO	4.3
2	B	56	PHE	4.2
1	A	107	GLY	4.1
1	A	41	ALA	4.1
1	A	108	ARG	4.0
1	A	276	PRO	3.9
1	A	194	VAL	3.9
1	A	8	PHE	3.8
1	A	27	TYR	3.7
2	B	54	LEU	3.6
2	B	65	LEU	3.6
2	B	1	ILE	3.6
2	B	92	ILE	3.6
2	B	60	TRP	3.6
1	A	26	GLY	3.5
3	C	8	TYR	3.4
1	A	100	GLY	3.3
1	A	99	TYR	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	17	ASN	3.2
2	B	19	LYS	3.2
1	A	95	ILE	3.2
1	A	98	MET	3.1
1	A	268	LYS	3.1
2	B	55	SER	3.1
1	A	9	TYR	3.1
2	B	89	GLN	3.0
1	A	23	ILE	3.0
1	A	22	PHE	3.0
1	A	7	TYR	3.0
1	A	267	PRO	2.9
1	A	221	GLY	2.9
2	B	63	TYR	2.9
1	A	31	THR	2.9
1	A	56	GLY	2.9
2	B	76	ASP	2.9
2	B	16	GLU	2.8
1	A	67	PHE	2.8
2	B	69	GLU	2.8
1	A	11	ALA	2.8
3	C	7	ASP	2.7
1	A	1	GLY	2.7
1	A	28	VAL	2.7
2	B	26	TYR	2.7
2	B	71	THR	2.7
1	A	74	TYR	2.7
1	A	200	THR	2.6
1	A	247	VAL	2.6
1	A	94	ILE	2.6
1	A	223	ASP	2.6
1	A	34	VAL	2.6
1	A	103	LEU	2.6
2	B	42	ASN	2.6
1	A	10	THR	2.6
1	A	252	GLY	2.5
1	A	117	ALA	2.5
2	B	80	CYS	2.5
1	A	196	ASP	2.5
3	C	9	PHE	2.5
1	A	116	SER	2.5
1	A	251	SER	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	62	PHE	2.4
1	A	153	ALA	2.4
1	A	163	LEU	2.4
2	B	27	VAL	2.4
2	B	43	GLY	2.4
1	A	262	GLN	2.4
1	A	250	PRO	2.4
2	B	33	SER	2.4
1	A	58	GLU	2.4
1	A	106	ASP	2.4
2	B	94	LYS	2.3
2	B	64	LEU	2.3
1	A	275	GLU	2.3
2	B	44	GLU	2.3
1	A	241	PHE	2.3
2	B	99	MET	2.3
1	A	211	ALA	2.3
1	A	44	ARG	2.3
1	A	4	SER	2.3
1	A	96	GLN	2.3
2	B	88	SER	2.2
1	A	5	MET	2.2
1	A	256	ARG	2.2
2	B	57	SER	2.2
1	A	235	PRO	2.2
2	B	98	ASP	2.2
1	A	101	CYS	2.2
2	B	14	PRO	2.2
1	A	60	TRP	2.2
2	B	77	GLU	2.2
1	A	123	TYR	2.1
1	A	209	TYR	2.1
1	A	42	SER	2.1
1	A	65	GLN	2.1
2	B	20	SER	2.1
1	A	110	LEU	2.1
1	A	24	ALA	2.1
1	A	121	LYS	2.1
1	A	269	PRO	2.1
2	B	7	ILE	2.1
1	A	248	VAL	2.1
1	A	233	THR	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	149	ALA	2.0
1	A	198	GLU	2.0
2	B	29	GLY	2.0
2	B	85	VAL	2.0
1	A	255	GLN	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 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.