



# Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 11:16 am BST

PDB ID : 6U07  
Title : Computational Stabilization of T Cell Receptor Constant Domains  
Authors : Froning, K.; Maguire, J.; Sereno, A.; Huang, F.; Chang, S.; Weichert, K.; Frommelt, A.J.; Dong, J.; Wu, X.; Austin, H.; Conner, E.M.; Fitchett, J.R.; Heng, A.R.; Balasubramaniam, D.; Hilgers, M.T.; Kuhlman, B.; Demarest, S.J.  
Deposited on : 2019-08-13  
Resolution : 1.76 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

<https://www.wwpdb.org/validation/2017/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) : 1.13  
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

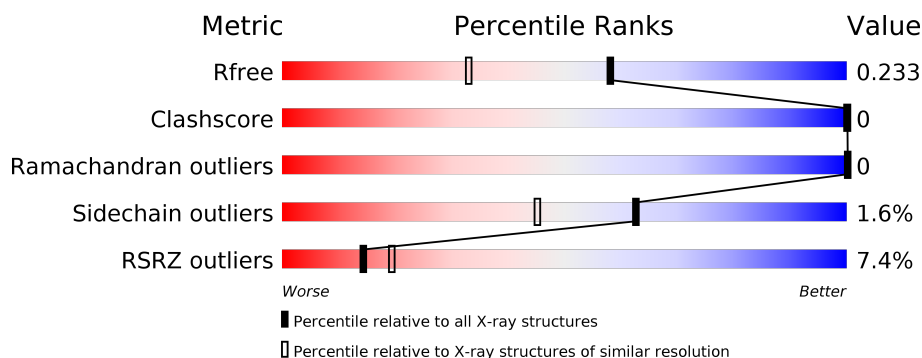
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.76 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	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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 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  $\leq 5\%$ . 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	106	
2	B	131	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 1904 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 Stabilized T cell receptor constant domain (Calpha).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	88	Total	C	N	O	S	0	1	0
			698	436	112	145	5			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	108	HIS	-	expression tag	UNP Q2YD82
A	109	HIS	-	expression tag	UNP Q2YD82
A	110	HIS	-	expression tag	UNP Q2YD82
A	111	HIS	-	expression tag	UNP Q2YD82
A	112	HIS	-	expression tag	UNP Q2YD82
A	113	HIS	-	expression tag	UNP Q2YD82
A	114	HIS	-	expression tag	UNP Q2YD82
A	115	HIS	-	expression tag	UNP Q2YD82
A	116	GLY	-	expression tag	UNP Q2YD82
A	117	SER	-	expression tag	UNP Q2YD82
A	139	PHE	SER	conflict	UNP Q2YD82
A	150	ILE	THR	conflict	UNP Q2YD82
A	166	CYS	THR	conflict	UNP Q2YD82
A	190	THR	ALA	conflict	UNP Q2YD82

- Molecule 2 is a protein called Stabilized T cell receptor constant domain (Cbeta).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	129	Total	C	N	O	S	0	0	0
			1028	653	182	190	3			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	134	LYS	GLU	conflict	UNP K7N5M4
B	139	ARG	HIS	conflict	UNP K7N5M4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	155	PRO	ASP	conflict	UNP K7N5M4
B	170	ASP	SER	conflict	UNP K7N5M4
B	205	ASP	ASN	conflict	UNP K7N5M4
B	247	CYS	-	expression tag	UNP K7N5M4

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Mg 1 1	0	0
3	A	2	Total Mg 2 2	0	0

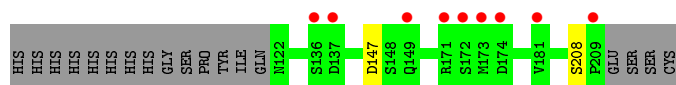
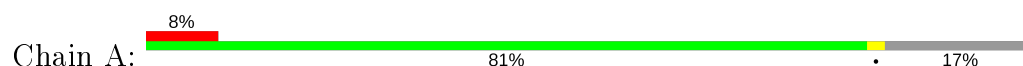
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	71	Total O 71 71	0	0
4	B	104	Total O 104 104	0	0

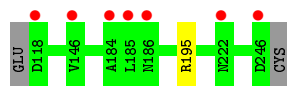
### 3 Residue-property plots [i](#)

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: Stabilized T cell receptor constant domain (Calpha)



- Molecule 2: Stabilized T cell receptor constant domain (Cbeta)



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	96.85Å 59.85Å 61.35Å 90.00° 110.24° 90.00°	Depositor
Resolution (Å)	29.33 – 1.76 29.33 – 1.76	Depositor EDS
% Data completeness (in resolution range)	99.1 (29.33-1.76) 99.2 (29.33-1.76)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.45 (at 1.76Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.202 , 0.226 0.211 , 0.233	Depositor DCC
$R_{free}$ test set	1642 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.6	Xtriage
Anisotropy	0.087	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 56.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	1904	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.51% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG

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	0.40	0/714	0.62	0/969
2	B	0.40	0/1059	0.61	0/1447
All	All	0.40	0/1773	0.62	0/2416

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

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	698	0	641	0	0
2	B	1028	0	976	0	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
4	A	71	0	0	0	0
4	B	104	0	0	0	0
All	All	1904	0	1617	0	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 0.

There are no clashes within the asymmetric unit.

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	87/106 (82%)	87 (100%)	0	0	100	100
2	B	127/131 (97%)	126 (99%)	1 (1%)	0	100	100
All	All	214/237 (90%)	213 (100%)	1 (0%)	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	83/100 (83%)	81 (98%)	2 (2%)	49	26
2	B	110/115 (96%)	109 (99%)	1 (1%)	78	67
All	All	193/215 (90%)	190 (98%)	3 (2%)	62	45

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

Mol	Chain	Res	Type
1	A	147	ASP
1	A	208	SER
2	B	195	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no



such sidechains identified.

### 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](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

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.

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, 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	88/106 (83%)	0.63	9 (10%) 6 9	18, 28, 49, 61	0
2	B	129/131 (98%)	0.12	7 (5%) 25 32	19, 26, 49, 61	0
All	All	217/237 (91%)	0.33	16 (7%) 14 19	18, 27, 49, 61	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	136	SER	4.6
2	B	184	ALA	4.0
1	A	209	PRO	3.9
1	A	171	ARG	3.7
1	A	137	ASP	3.3
2	B	185	LEU	2.9
1	A	149	GLN	2.7
2	B	186	ASN	2.4
1	A	181	VAL	2.3
2	B	246	ASP	2.3
1	A	172	SER	2.3
2	B	222	ASN	2.3
2	B	146	VAL	2.2
1	A	174	ASP	2.1
2	B	118	ASP	2.1
1	A	173	MET	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](#)

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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
3	MG	B	301	1/1	0.97	0.23	44,44,44,44	0
3	MG	A	301	1/1	0.98	0.06	23,23,23,23	1
3	MG	A	302	1/1	0.99	0.05	32,32,32,32	1

### 6.5 Other polymers [i](#)

There are no such residues in this entry.