



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:31 PM GMT

PDB ID : 4P2Q  
Title : Crystal structure of the 5cc7 TCR in complex with 5c2/I-Ek  
Authors : Birnbaum, M.E.; Ozkan, E.; Garcia, K.C.  
Deposited on : 2014-03-04  
Resolution : 3.30 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](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

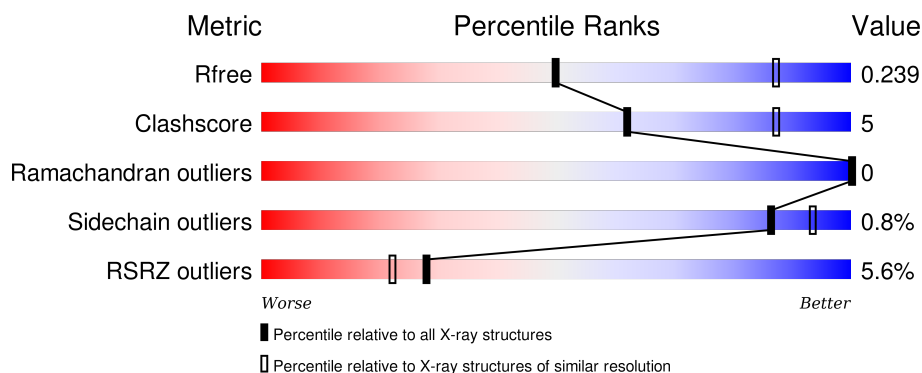
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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	2060 (3.40-3.20)
Clashscore	102246	1058 (3.38-3.22)
Ramachandran outliers	100387	1038 (3.38-3.22)
Sidechain outliers	100360	1037 (3.38-3.22)
RSRZ outliers	91569	2070 (3.40-3.20)

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	204	<div> <div>0%</div> <div> <div></div> <div>70%</div> <div>18%</div> <div>12%</div> </div> </div>
1	F	204	<div> <div>0%</div> <div> <div></div> <div>74%</div> <div>14%</div> <div>12%</div> </div> </div>
1	K	204	<div> <div></div> <div> <div></div> <div>73%</div> <div>15%</div> <div>12%</div> </div> </div>
1	P	204	<div> <div>4%</div> <div> <div></div> <div>73%</div> <div>15%</div> <div>12%</div> </div> </div>
2	B	212	<div> <div>5%</div> <div> <div></div> <div>67%</div> <div>14%</div> <div>19%</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	212	
2	L	212	
2	Q	212	
3	C	14	
3	H	14	
3	M	14	
3	R	14	
4	D	205	
4	I	205	
4	N	205	
4	S	205	
5	E	266	
5	J	266	
5	O	266	
5	T	266	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 25667 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 H-2 class II histocompatibility antigen, E-K alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	180	Total	C	N	O	S	0	0	0
			1473	951	243	275	4			
1	F	180	Total	C	N	O	S	0	0	0
			1473	951	243	275	4			
1	K	180	Total	C	N	O	S	0	0	0
			1469	948	242	275	4			
1	P	179	Total	C	N	O	S	0	0	0
			1461	942	241	274	4			

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	ALA	-	expression tag	UNP P04224
A	-1	ASP	-	expression tag	UNP P04224
A	0	PRO	-	expression tag	UNP P04224
A	192	SER	-	expression tag	UNP P04224
A	193	ARG	-	expression tag	UNP P04224
A	194	GLY	-	expression tag	UNP P04224
A	195	GLY	-	expression tag	UNP P04224
A	196	LEU	-	expression tag	UNP P04224
A	197	GLU	-	expression tag	UNP P04224
A	198	VAL	-	expression tag	UNP P04224
A	199	LEU	-	expression tag	UNP P04224
A	200	PHE	-	expression tag	UNP P04224
A	201	GLN	-	expression tag	UNP P04224
F	-2	ALA	-	expression tag	UNP P04224
F	-1	ASP	-	expression tag	UNP P04224
F	0	PRO	-	expression tag	UNP P04224
F	192	SER	-	expression tag	UNP P04224
F	193	ARG	-	expression tag	UNP P04224
F	194	GLY	-	expression tag	UNP P04224
F	195	GLY	-	expression tag	UNP P04224
F	196	LEU	-	expression tag	UNP P04224

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	197	GLU	-	expression tag	UNP P04224
F	198	VAL	-	expression tag	UNP P04224
F	199	LEU	-	expression tag	UNP P04224
F	200	PHE	-	expression tag	UNP P04224
F	201	GLN	-	expression tag	UNP P04224
K	-2	ALA	-	expression tag	UNP P04224
K	-1	ASP	-	expression tag	UNP P04224
K	0	PRO	-	expression tag	UNP P04224
K	192	SER	-	expression tag	UNP P04224
K	193	ARG	-	expression tag	UNP P04224
K	194	GLY	-	expression tag	UNP P04224
K	195	GLY	-	expression tag	UNP P04224
K	196	LEU	-	expression tag	UNP P04224
K	197	GLU	-	expression tag	UNP P04224
K	198	VAL	-	expression tag	UNP P04224
K	199	LEU	-	expression tag	UNP P04224
K	200	PHE	-	expression tag	UNP P04224
K	201	GLN	-	expression tag	UNP P04224
P	-2	ALA	-	expression tag	UNP P04224
P	-1	ASP	-	expression tag	UNP P04224
P	0	PRO	-	expression tag	UNP P04224
P	192	SER	-	expression tag	UNP P04224
P	193	ARG	-	expression tag	UNP P04224
P	194	GLY	-	expression tag	UNP P04224
P	195	GLY	-	expression tag	UNP P04224
P	196	LEU	-	expression tag	UNP P04224
P	197	GLU	-	expression tag	UNP P04224
P	198	VAL	-	expression tag	UNP P04224
P	199	LEU	-	expression tag	UNP P04224
P	200	PHE	-	expression tag	UNP P04224
P	201	GLN	-	expression tag	UNP P04224

- Molecule 2 is a protein called MHC class II E-beta-k.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	171	Total	C	N	O	S	0	0	0
			1387	885	237	259	6			
2	G	171	Total	C	N	O	S	0	0	0
			1382	883	237	256	6			
2	L	171	Total	C	N	O	S	0	0	0
			1382	883	237	256	6			
2	Q	171	Total	C	N	O	S	0	0	0
			1382	883	237	256	6			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	GLY	-	expression tag	UNP Q31163
B	-2	SER	-	expression tag	UNP Q31163
B	-1	GLY	-	expression tag	UNP Q31163
B	0	GLY	-	expression tag	UNP Q31163
B	1	GLY	-	expression tag	UNP Q31163
B	2	GLY	-	expression tag	UNP Q31163
B	199	SER	-	expression tag	UNP Q31163
B	200	ARG	-	expression tag	UNP Q31163
B	201	GLY	-	expression tag	UNP Q31163
B	202	GLY	-	expression tag	UNP Q31163
B	203	LEU	-	expression tag	UNP Q31163
B	204	GLU	-	expression tag	UNP Q31163
B	205	VAL	-	expression tag	UNP Q31163
B	206	LEU	-	expression tag	UNP Q31163
B	207	PHE	-	expression tag	UNP Q31163
B	208	GLN	-	expression tag	UNP Q31163
G	-3	GLY	-	expression tag	UNP Q31163
G	-2	SER	-	expression tag	UNP Q31163
G	-1	GLY	-	expression tag	UNP Q31163
G	0	GLY	-	expression tag	UNP Q31163
G	1	GLY	-	expression tag	UNP Q31163
G	2	GLY	-	expression tag	UNP Q31163
G	199	SER	-	expression tag	UNP Q31163
G	200	ARG	-	expression tag	UNP Q31163
G	201	GLY	-	expression tag	UNP Q31163
G	202	GLY	-	expression tag	UNP Q31163
G	203	LEU	-	expression tag	UNP Q31163
G	204	GLU	-	expression tag	UNP Q31163
G	205	VAL	-	expression tag	UNP Q31163
G	206	LEU	-	expression tag	UNP Q31163
G	207	PHE	-	expression tag	UNP Q31163
G	208	GLN	-	expression tag	UNP Q31163
L	-3	GLY	-	expression tag	UNP Q31163
L	-2	SER	-	expression tag	UNP Q31163
L	-1	GLY	-	expression tag	UNP Q31163
L	0	GLY	-	expression tag	UNP Q31163
L	1	GLY	-	expression tag	UNP Q31163
L	2	GLY	-	expression tag	UNP Q31163
L	199	SER	-	expression tag	UNP Q31163
L	200	ARG	-	expression tag	UNP Q31163
L	201	GLY	-	expression tag	UNP Q31163
L	202	GLY	-	expression tag	UNP Q31163

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
L	203	LEU	-	expression tag	UNP Q31163
L	204	GLU	-	expression tag	UNP Q31163
L	205	VAL	-	expression tag	UNP Q31163
L	206	LEU	-	expression tag	UNP Q31163
L	207	PHE	-	expression tag	UNP Q31163
L	208	GLN	-	expression tag	UNP Q31163
Q	-3	GLY	-	expression tag	UNP Q31163
Q	-2	SER	-	expression tag	UNP Q31163
Q	-1	GLY	-	expression tag	UNP Q31163
Q	0	GLY	-	expression tag	UNP Q31163
Q	1	GLY	-	expression tag	UNP Q31163
Q	2	GLY	-	expression tag	UNP Q31163
Q	199	SER	-	expression tag	UNP Q31163
Q	200	ARG	-	expression tag	UNP Q31163
Q	201	GLY	-	expression tag	UNP Q31163
Q	202	GLY	-	expression tag	UNP Q31163
Q	203	LEU	-	expression tag	UNP Q31163
Q	204	GLU	-	expression tag	UNP Q31163
Q	205	VAL	-	expression tag	UNP Q31163
Q	206	LEU	-	expression tag	UNP Q31163
Q	207	PHE	-	expression tag	UNP Q31163
Q	208	GLN	-	expression tag	UNP Q31163

- Molecule 3 is a protein called 5c2 peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	14	Total	C	N	O	0	0	0
			104	67	18	19			
3	H	14	Total	C	N	O	0	0	0
			104	67	18	19			
3	M	13	Total	C	N	O	0	0	0
			100	65	17	18			
3	R	13	Total	C	N	O	0	0	0
			100	65	17	18			

- Molecule 4 is a protein called 5cc7 T-cell receptor alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	196	Total	C	N	O	S	0	0	0
			1523	942	264	309	8			
4	I	196	Total	C	N	O	S	0	0	0
			1517	939	261	309	8			

*Continued on next page...*

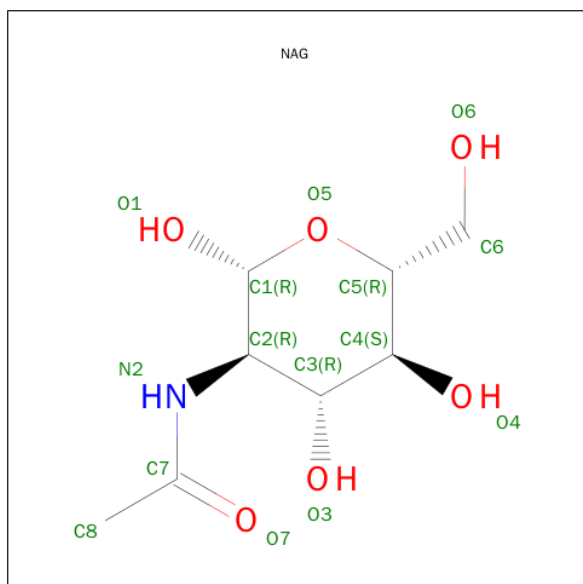
*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	N	196	Total	C	N	O	S	0	0	0
			1523	942	264	309	8			
4	S	196	Total	C	N	O	S	0	0	0
			1523	942	264	309	8			

- Molecule 5 is a protein called 5cc7 T-cell receptor beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	245	Total	C	N	O	S	0	0	0
			1929	1212	336	372	9			
5	J	245	Total	C	N	O	S	0	0	0
			1929	1212	336	372	9			
5	O	242	Total	C	N	O	S	0	0	0
			1920	1208	334	369	9			
5	T	242	Total	C	N	O	S	0	0	0
			1916	1205	333	369	9			

- Molecule 6 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	B	1	Total	C	N	O	0	0
			14	8	1	5		
6	G	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*



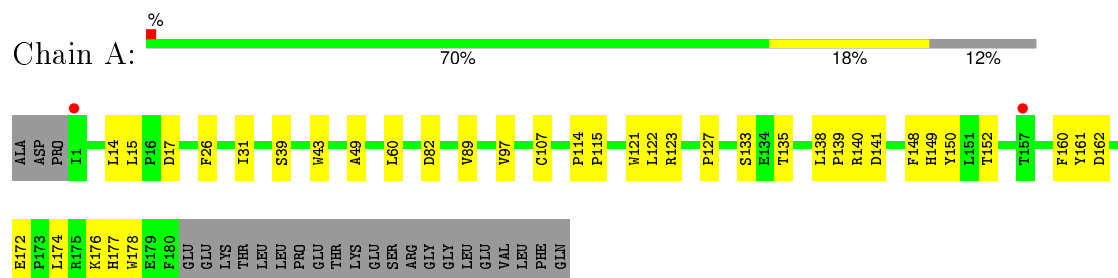
*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	K	1	Total	C	N	O	0	0
			14	8	1	5		
6	L	1	Total	C	N	O	0	0
			14	8	1	5		

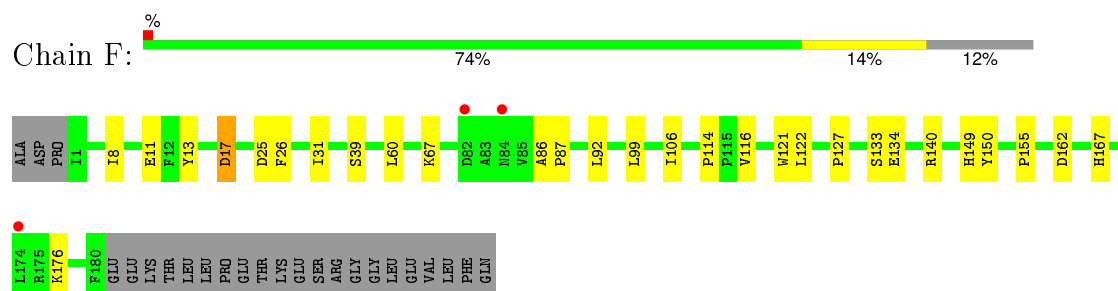
### 3 Residue-property plots

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.

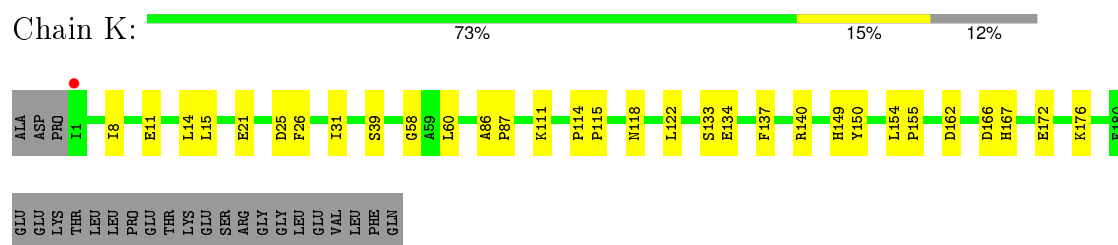
- Molecule 1: H-2 class II histocompatibility antigen, E-K alpha chain



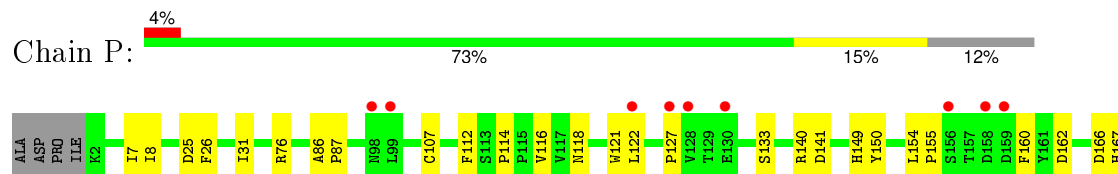
- Molecule 1: H-2 class II histocompatibility antigen, E-K alpha chain



- Molecule 1: H-2 class II histocompatibility antigen, E-K alpha chain

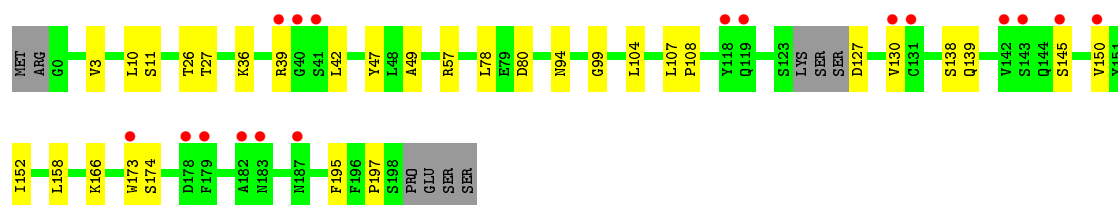


- Molecule 1: H-2 class II histocompatibility antigen, E-K alpha chain



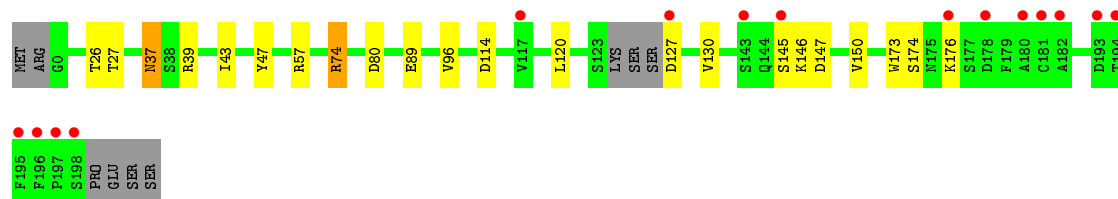






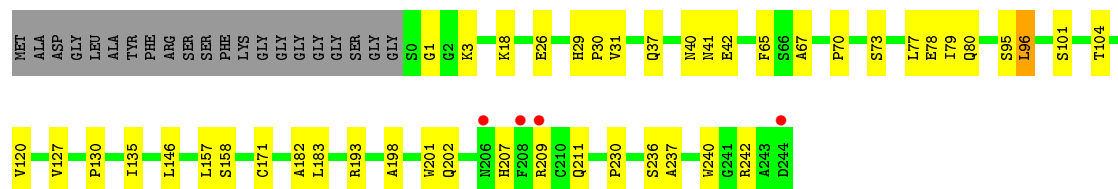
- Molecule 4: 5cc7 T-cell receptor alpha chain

Chain S: 7% 85% 10%



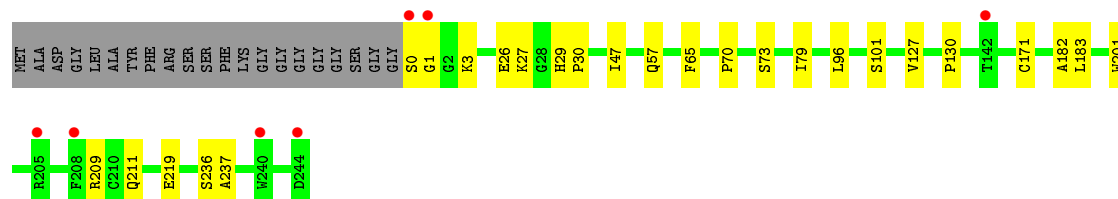
- Molecule 5: 5cc7 T-cell receptor beta chain

Chain E: 2% 75% 17% 8%



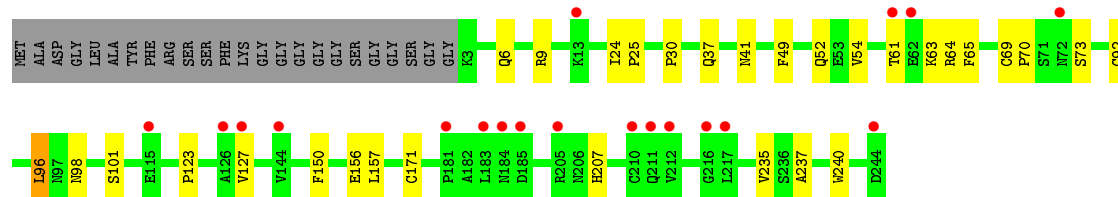
- Molecule 5: 5cc7 T-cell receptor beta chain

Chain J: 3% 82% 10% 8%



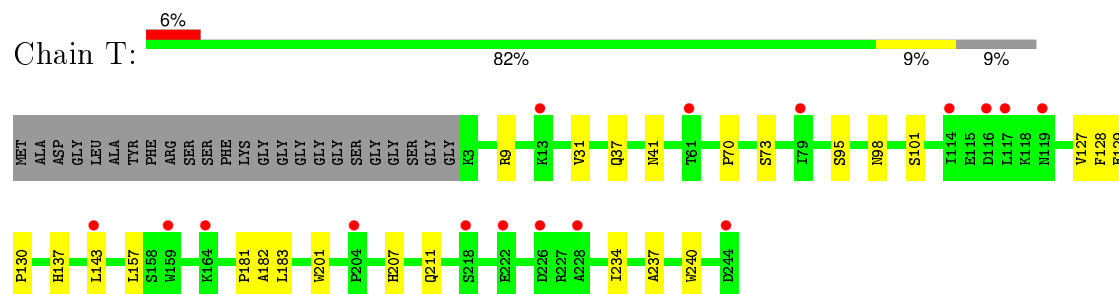
- Molecule 5: 5cc7 T-cell receptor beta chain

Chain O: 7% 79% 11% 9%



- Molecule 5: 5cc7 T-cell receptor beta chain

Chain T:



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	262.90Å 102.21Å 214.11Å 90.00° 95.04° 90.00°	Depositor
Resolution (Å)	39.63 – 3.30 39.67 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.1 (39.63-3.30) 99.3 (39.67-3.30)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.53 (at 3.32Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, $R_{free}$	0.188 , 0.236 0.195 , 0.239	Depositor DCC
$R_{free}$ test set	4237 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	103.1	Xtriage
Anisotropy	0.399	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 78.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	2 of 84866 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	25667	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	124.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.11% 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

### 5.1 Standard geometry

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	0.35	0/1516	0.56	0/2062
1	F	0.31	0/1516	0.50	0/2062
1	K	0.31	0/1512	0.51	0/2058
1	P	0.26	0/1504	0.46	0/2047
2	B	0.38	0/1424	0.51	0/1943
2	G	0.35	0/1419	0.53	0/1937
2	L	0.31	0/1419	0.47	0/1937
2	Q	0.28	0/1419	0.46	0/1937
3	C	0.38	0/106	0.47	0/139
3	H	0.33	0/106	0.45	0/139
3	M	0.38	0/102	0.37	0/134
3	R	0.28	0/102	0.40	0/134
4	D	0.33	0/1551	0.51	0/2098
4	I	0.30	0/1545	0.48	0/2091
4	N	0.32	0/1551	0.50	0/2098
4	S	0.30	0/1551	0.47	0/2098
5	E	0.30	0/1978	0.48	0/2691
5	J	0.28	0/1978	0.46	0/2691
5	O	0.25	0/1969	0.43	0/2678
5	T	0.25	0/1965	0.44	0/2674
All	All	0.30	0/26233	0.48	0/35648

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	1473	0	1402	21	0
1	F	1473	0	1403	19	0
1	K	1469	0	1391	21	0
1	P	1461	0	1378	19	0
2	B	1387	0	1279	19	0
2	G	1382	0	1272	20	0
2	L	1382	0	1272	22	0
2	Q	1382	0	1273	17	0
3	C	104	0	99	1	0
3	H	104	0	99	1	0
3	M	100	0	96	4	0
3	R	100	0	96	1	0
4	D	1523	0	1441	17	0
4	I	1517	0	1430	15	0
4	N	1523	0	1441	24	0
4	S	1523	0	1441	16	0
5	E	1929	0	1816	27	0
5	J	1929	0	1816	17	0
5	O	1920	0	1816	20	0
5	T	1916	0	1805	19	0
6	A	14	0	13	0	0
6	B	14	0	13	0	0
6	G	14	0	13	0	0
6	K	14	0	13	0	0
6	L	14	0	13	0	0
All	All	25667	0	24131	269	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 5.

The worst 5 of 269 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:K:133:SER:HB2	1:K:150:TYR:HB2	1.63	0.79
2:L:10:TYR:HB3	2:L:31:PHE:HB2	1.64	0.79

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:10:TYR:HB3	2:B:31:PHE:HB2	1.64	0.79
2:L:48:ARG:NH1	1:P:141:ASP:OD2	2.17	0.77
2:L:21:THR:O	2:L:80:ARG:NH1	2.16	0.77

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	178/204 (87%)	173 (97%)	5 (3%)	0	100	100
1	F	178/204 (87%)	174 (98%)	4 (2%)	0	100	100
1	K	178/204 (87%)	174 (98%)	4 (2%)	0	100	100
1	P	177/204 (87%)	171 (97%)	6 (3%)	0	100	100
2	B	165/212 (78%)	157 (95%)	8 (5%)	0	100	100
2	G	165/212 (78%)	158 (96%)	7 (4%)	0	100	100
2	L	165/212 (78%)	157 (95%)	8 (5%)	0	100	100
2	Q	165/212 (78%)	158 (96%)	7 (4%)	0	100	100
3	C	12/14 (86%)	11 (92%)	1 (8%)	0	100	100
3	H	12/14 (86%)	11 (92%)	1 (8%)	0	100	100
3	M	11/14 (79%)	11 (100%)	0	0	100	100
3	R	11/14 (79%)	11 (100%)	0	0	100	100
4	D	192/205 (94%)	183 (95%)	9 (5%)	0	100	100
4	I	192/205 (94%)	184 (96%)	8 (4%)	0	100	100
4	N	192/205 (94%)	184 (96%)	8 (4%)	0	100	100
4	S	192/205 (94%)	184 (96%)	8 (4%)	0	100	100
5	E	243/266 (91%)	229 (94%)	14 (6%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	J	243/266 (91%)	230 (95%)	13 (5%)	0	100	100
5	O	240/266 (90%)	229 (95%)	11 (5%)	0	100	100
5	T	240/266 (90%)	230 (96%)	10 (4%)	0	100	100
All	All	3151/3604 (87%)	3019 (96%)	132 (4%)	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	161/184 (88%)	158 (98%)	3 (2%)	65	85
1	F	161/184 (88%)	160 (99%)	1 (1%)	90	95
1	K	160/184 (87%)	159 (99%)	1 (1%)	90	95
1	P	159/184 (86%)	158 (99%)	1 (1%)	90	95
2	B	150/191 (78%)	148 (99%)	2 (1%)	76	89
2	G	148/191 (78%)	147 (99%)	1 (1%)	88	94
2	L	148/191 (78%)	147 (99%)	1 (1%)	88	94
2	Q	148/191 (78%)	147 (99%)	1 (1%)	88	94
3	C	9/9 (100%)	9 (100%)	0	100	100
3	H	9/9 (100%)	9 (100%)	0	100	100
3	M	9/9 (100%)	9 (100%)	0	100	100
3	R	9/9 (100%)	9 (100%)	0	100	100
4	D	171/181 (94%)	169 (99%)	2 (1%)	78	90
4	I	170/181 (94%)	168 (99%)	2 (1%)	78	90
4	N	171/181 (94%)	170 (99%)	1 (1%)	90	95
4	S	171/181 (94%)	168 (98%)	3 (2%)	66	85
5	E	209/226 (92%)	206 (99%)	3 (1%)	74	88
5	J	209/226 (92%)	209 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	O	210/226 (93%)	209 (100%)	1 (0%)	92	95
5	T	209/226 (92%)	209 (100%)	0	100	100
All	All	2791/3164 (88%)	2768 (99%)	23 (1%)	86	93

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	17	ASP
4	I	37	ASN
4	S	74	ARG
2	G	173	CYS
4	I	127	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
4	D	2	GLN
5	T	37	GLN
5	T	72	ASN

### 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 ⓘ

5 ligands are 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	$\# Z  > 2$
6	NAG	A	901	1	14,14,15	0.42	0	15,19,21	0.55	0
6	NAG	B	901	2	14,14,15	0.49	0	15,19,21	0.29	0
6	NAG	G	901	2	14,14,15	0.36	0	15,19,21	0.77	1 (6%)
6	NAG	K	901	1	14,14,15	0.52	0	15,19,21	0.50	0
6	NAG	L	901	2	14,14,15	0.49	0	15,19,21	0.80	1 (6%)

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
6	NAG	A	901	1	-	0/6/23/26	0/1/1/1
6	NAG	B	901	2	-	0/6/23/26	0/1/1/1
6	NAG	G	901	2	-	0/6/23/26	0/1/1/1
6	NAG	K	901	1	-	0/6/23/26	0/1/1/1
6	NAG	L	901	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	901	NAG	C1-O5-C5	2.77	115.76	112.25
6	L	901	NAG	C1-O5-C5	2.85	115.87	112.25

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 ⓘ

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<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	180/204 (88%)	-0.02	2 (1%) 82 78	57, 83, 146, 185	0
1	F	180/204 (88%)	0.09	3 (1%) 73 67	69, 109, 182, 243	0
1	K	180/204 (88%)	-0.02	1 (0%) 90 88	77, 101, 163, 207	0
1	P	179/204 (87%)	0.29	9 (5%) 32 26	102, 144, 206, 234	0
2	B	171/212 (80%)	0.23	10 (5%) 26 21	60, 86, 182, 221	0
2	G	171/212 (80%)	0.34	11 (6%) 23 19	60, 94, 167, 207	0
2	L	171/212 (80%)	0.29	13 (7%) 17 14	70, 112, 206, 264	0
2	Q	171/212 (80%)	0.76	29 (16%) 2 2	87, 132, 222, 265	0
3	C	14/14 (100%)	-0.13	0 100 100	56, 69, 120, 125	0
3	H	14/14 (100%)	0.12	0 100 100	69, 88, 118, 141	0
3	M	13/14 (92%)	0.13	0 100 100	87, 101, 132, 146	0
3	R	13/14 (92%)	0.39	1 (7%) 16 13	88, 111, 144, 154	0
4	D	196/205 (95%)	0.16	9 (4%) 36 30	60, 109, 194, 259	0
4	I	196/205 (95%)	0.37	14 (7%) 19 15	72, 121, 194, 246	0
4	N	196/205 (95%)	0.35	17 (8%) 13 10	77, 114, 193, 226	0
4	S	196/205 (95%)	0.38	15 (7%) 16 13	80, 120, 211, 248	0
5	E	245/266 (92%)	0.03	4 (1%) 74 69	65, 113, 166, 201	0
5	J	245/266 (92%)	0.09	7 (2%) 55 49	71, 119, 172, 236	1 (0%)
5	O	242/266 (90%)	0.46	19 (7%) 15 12	93, 155, 212, 258	0
5	T	242/266 (90%)	0.45	16 (6%) 22 17	105, 150, 204, 231	0
All	All	3215/3604 (89%)	0.26	180 (5%) 28 22	56, 119, 197, 265	1 (0%)

The worst 5 of 180 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	J	0	SER	11.9
5	J	1	GLY	7.0
4	S	195	PHE	6.0
1	P	159	ASP	5.9
5	J	244	ASP	5.2

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	NAG	A	901	14/15	0.89	0.20	-0.19	98,120,127,127	0
6	NAG	K	901	14/15	0.87	0.18	-0.62	126,144,153,154	0
6	NAG	G	901	14/15	0.78	0.16	-	90,112,133,144	0
6	NAG	L	901	14/15	0.83	0.18	-	98,125,137,138	0
6	NAG	B	901	14/15	0.71	0.23	-	92,139,146,147	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.