



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:45 PM GMT

PDB ID : 4MNH
Title : Structure of the DP10.7 TCR
Authors : Luoma, A.M.; Adams, E.J.
Deposited on : 2013-09-10
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
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

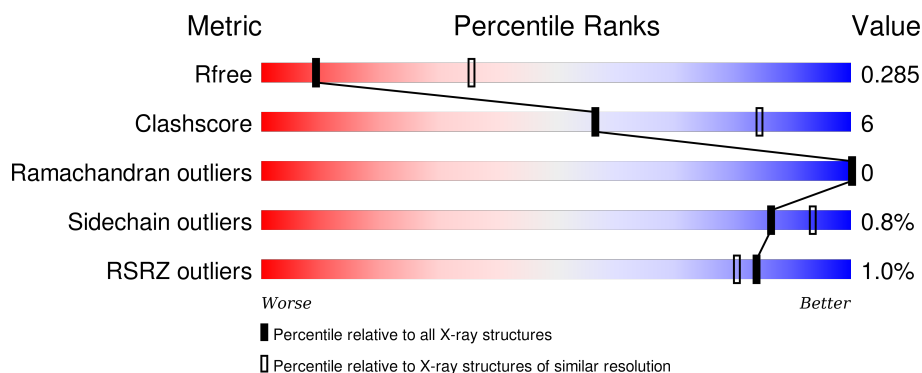
1 Overall quality at a glance

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 ≥ 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	259	<div> <div>2%</div> <div>73% 16% 11%</div> </div>
1	C	259	<div> <div>2%</div> <div>77% 14% 9%</div> </div>
2	B	226	<div> <div>72% 15% 13%</div> </div>
2	D	226	<div> <div>2%</div> <div>77% 11% 12%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6440 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 T-cell receptor gamma chain V region PT-gamma-1/2, Human nkt tcr beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	231	Total	C	N	O	S	0	0	0
			1751	1118	301	326	6			
1	C	236	Total	C	N	O	S	0	0	0
			1803	1147	308	342	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	PRO	-	EXPRESSION TAG	UNP 4MNH
A	-1	ASP	-	EXPRESSION TAG	UNP 4MNH
A	0	LEU	-	EXPRESSION TAG	UNP 4MNH
C	-2	PRO	-	EXPRESSION TAG	UNP 4MNH
C	-1	ASP	-	EXPRESSION TAG	UNP 4MNH
C	0	LEU	-	EXPRESSION TAG	UNP 4MNH

- Molecule 2 is a protein called TRA@ protein, Human nkt tcr alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	197	Total	C	N	O	S	0	0	0
			1449	927	243	270	9			
2	D	199	Total	C	N	O	S	0	0	0
			1437	917	240	271	9			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	ALA	-	EXPRESSION TAG	UNP Q6PJ56
B	-1	ASP	-	EXPRESSION TAG	UNP Q6PJ56
B	0	PRO	-	EXPRESSION TAG	UNP Q6PJ56
B	96	PRO	-	LINKER	UNP Q6PJ56
B	98	TYR	-	LINKER	UNP Q6PJ56

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	99	TRP	-	LINKER	UNP Q6PJ56
B	100	GLY	-	LINKER	UNP Q6PJ56
B	?	-	LEU	DELETION	UNP Q6PJ56
B	103	ARG	-	LINKER	UNP Q6PJ56
B	104	THR	-	LINKER	UNP Q6PJ56
B	105	THR	PHE	CONFLICT	UNP Q6PJ56
B	107	VAL	GLY	CONFLICT	UNP Q6PJ56
B	108	ILE	ASN	CONFLICT	UNP Q6PJ56
B	110	GLY	HIS	CONFLICT	UNP Q6PJ56
B	111	LYS	TYR	CONFLICT	UNP Q6PJ56
B	112	GLY	-	LINKER	UNP Q6PJ56
B	113	THR	-	LINKER	UNP Q6PJ56
B	114	ARG	-	LINKER	UNP Q6PJ56
B	115	VAL	-	LINKER	UNP Q6PJ56
B	116	THR	-	LINKER	UNP Q6PJ56
B	117	VAL	-	LINKER	UNP Q6PJ56
B	118	GLU	-	LINKER	UNP Q6PJ56
B	119	PRO	-	LINKER	UNP Q6PJ56
B	120	ASN	ASP	CONFLICT	UNP K7N5M3
D	-2	ALA	-	EXPRESSION TAG	UNP Q6PJ56
D	-1	ASP	-	EXPRESSION TAG	UNP Q6PJ56
D	0	PRO	-	EXPRESSION TAG	UNP Q6PJ56
D	96	PRO	-	LINKER	UNP Q6PJ56
D	98	TYR	-	LINKER	UNP Q6PJ56
D	99	TRP	-	LINKER	UNP Q6PJ56
D	100	GLY	-	LINKER	UNP Q6PJ56
D	?	-	LEU	DELETION	UNP Q6PJ56
D	103	ARG	-	LINKER	UNP Q6PJ56
D	104	THR	-	LINKER	UNP Q6PJ56
D	105	THR	PHE	CONFLICT	UNP Q6PJ56
D	107	VAL	GLY	CONFLICT	UNP Q6PJ56
D	108	ILE	ASN	CONFLICT	UNP Q6PJ56
D	110	GLY	HIS	CONFLICT	UNP Q6PJ56
D	111	LYS	TYR	CONFLICT	UNP Q6PJ56
D	112	GLY	-	LINKER	UNP Q6PJ56
D	113	THR	-	LINKER	UNP Q6PJ56
D	114	ARG	-	LINKER	UNP Q6PJ56
D	115	VAL	-	LINKER	UNP Q6PJ56
D	116	THR	-	LINKER	UNP Q6PJ56
D	117	VAL	-	LINKER	UNP Q6PJ56
D	118	GLU	-	LINKER	UNP Q6PJ56
D	119	PRO	-	LINKER	UNP Q6PJ56

Continued on next page...

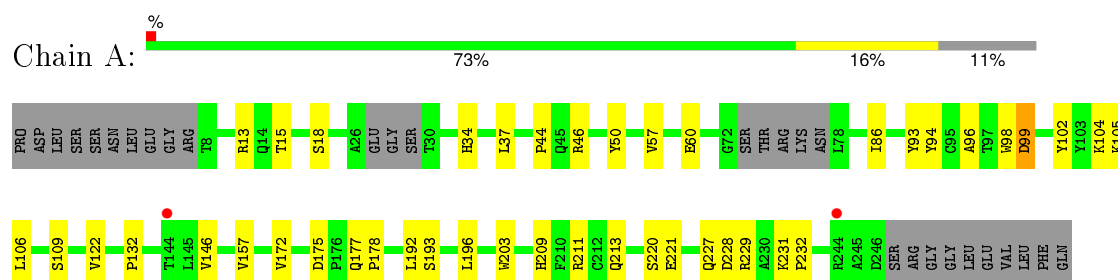
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	120	ASN	ASP	CONFLICT	UNP K7N5M3

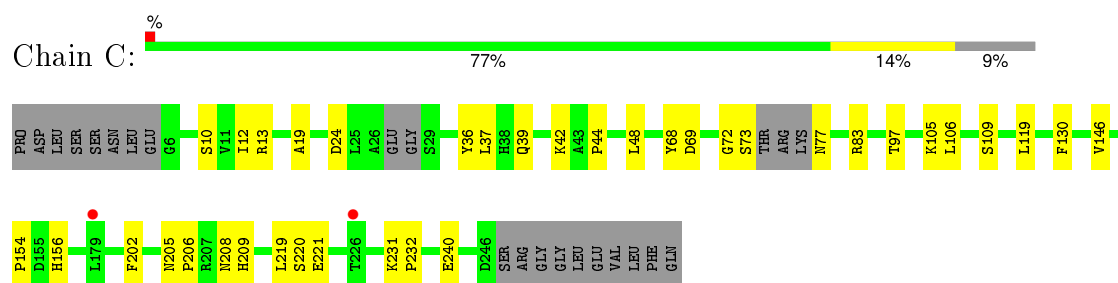
3 Residue-property plots [i](#)

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.

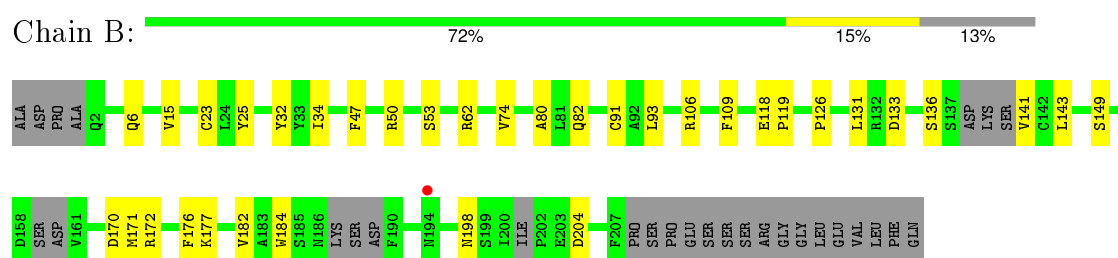
- Molecule 1: T-cell receptor gamma chain V region PT-gamma-1/2, Human nkt tcr beta chain



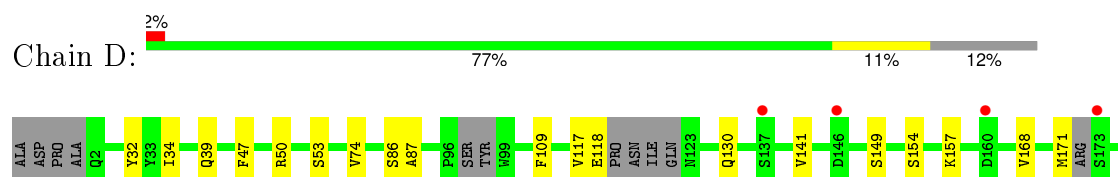
- Molecule 1: T-cell receptor gamma chain V region PT-gamma-1/2, Human nkt tcr beta chain



- Molecule 2: TRA@ protein, Human nkt tcr alpha chain



- Molecule 2: TRA@ protein, Human nkt tcr alpha chain





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	92.26 Å 60.84 Å 99.09 Å 90.00° 113.33° 90.00°	Depositor
Resolution (Å)	49.41 – 3.30 49.41 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.0 (49.41-3.30) 99.1 (49.41-3.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.10 (at 3.33 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.231 , 0.285 0.228 , 0.285	Depositor DCC
R_{free} test set	758 reflections (4.95%)	DCC
Wilson B-factor (Å ²)	72.1	Xtriage
Anisotropy	0.314	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 44.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 15343 reflections	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	6440	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.37% 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, 0.375 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	0.21	0/1799	0.40	0/2462
1	C	0.21	0/1851	0.39	0/2531
2	B	0.22	0/1480	0.40	0/2012
2	D	0.22	0/1467	0.38	0/2000
All	All	0.21	0/6597	0.39	0/9005

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	1751	0	1585	25	0
1	C	1803	0	1629	20	0
2	B	1449	0	1317	22	0
2	D	1437	0	1241	12	0
All	All	6440	0	5772	75	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 6.

The worst 5 of 75 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:32:TYR:H	2:B:53:SER:HB3	1.56	0.71
2:B:170:ASP:OD1	2:B:177:LYS:NZ	2.24	0.70
1:A:104:LYS:HE3	1:A:106:LEU:HD21	1.73	0.69
2:D:32:TYR:H	2:D:53:SER:HB2	1.58	0.67
2:B:141:VAL:HG23	2:B:184:TRP:HB3	1.76	0.67

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	225/259 (87%)	219 (97%)	6 (3%)	0	100	100
1	C	230/259 (89%)	226 (98%)	4 (2%)	0	100	100
2	B	187/226 (83%)	181 (97%)	6 (3%)	0	100	100
2	D	189/226 (84%)	187 (99%)	2 (1%)	0	100	100
All	All	831/970 (86%)	813 (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	171/226 (76%)	168 (98%)	3 (2%)	66	85
1	C	178/226 (79%)	177 (99%)	1 (1%)	90	95

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	141/201 (70%)	141 (100%)	0	100	100
2	D	132/201 (66%)	131 (99%)	1 (1%)	86	93
All	All	622/854 (73%)	617 (99%)	5 (1%)	86	93

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	99	ASP
1	A	157	VAL
1	A	192	LEU
1	C	119	LEU
2	D	168	VAL

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

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 ⓘ

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 [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, 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	231/259 (89%)	-0.03	2 (0%) 85 82	50, 74, 118, 148	0
1	C	236/259 (91%)	0.09	2 (0%) 87 84	49, 71, 103, 124	0
2	B	197/226 (87%)	0.01	1 (0%) 91 90	50, 80, 127, 147	0
2	D	199/226 (88%)	0.14	4 (2%) 68 62	53, 88, 128, 144	0
All	All	863/970 (88%)	0.05	9 (1%) 84 80	49, 78, 121, 148	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	137	SER	4.2
2	D	173	SER	3.9
1	A	144	THR	3.3
1	C	226	THR	2.7
2	D	160	ASP	2.4

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.