wwPDB X-ray Structure Validation Summary Report (i)

Aug 22, 2023 – 12:11 AM EDT

PDB ID	:	2P1Y
Title	:	1.B2.D9, a bispecific alpha/beta TCR
Authors	:	McBeth, C.; Pizarro, J.C.; Strong, R.K.
Deposited on	:	2007-03-06
Resolution	:	2.42 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

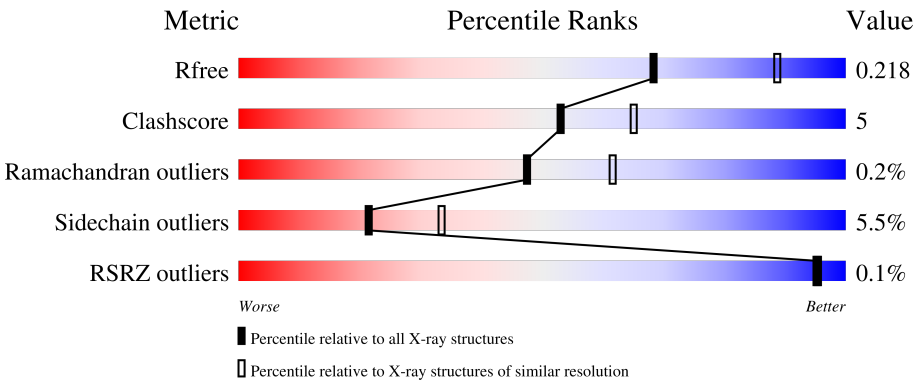
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.42 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	4647 (2.44-2.40)
Clashscore	141614	5161 (2.44-2.40)
Ramachandran outliers	138981	5073 (2.44-2.40)
Sidechain outliers	138945	5074 (2.44-2.40)
RSRZ outliers	127900	4543 (2.44-2.40)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	238	66% 28% • •
1	C	238	68% 25% • •
1	E	238	70% 21% 9%
1	G	238	74% 16% • 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7017 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 bispecific alpha/beta TCR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	231	1779	1124	304	346	5	0	0	0
			Total	C	N	O	S			
1	C	231	1771	1116	304	346	5	0	0	0
			Total	C	N	O	S			
1	E	217	1675	1064	283	323	5	0	0	0
			Total	C	N	O	S			
1	G	217	1690	1073	288	324	5	0	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	25	25	25	0	0
			Total	O		
2	C	15	15	15	0	0
			Total	O		
2	E	30	30	30	0	0
			Total	O		
2	G	32	32	32	0	0

3 Residue-property plots (i)

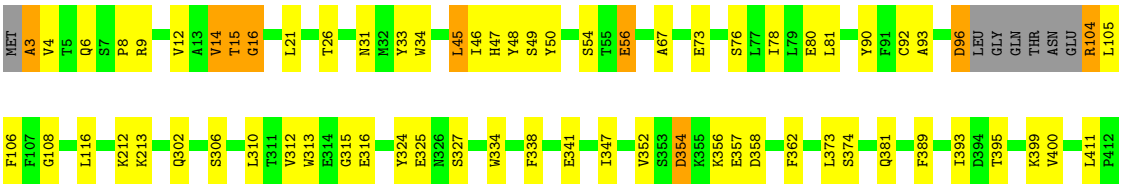
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: bispecific alpha/beta TCR

Chain A:

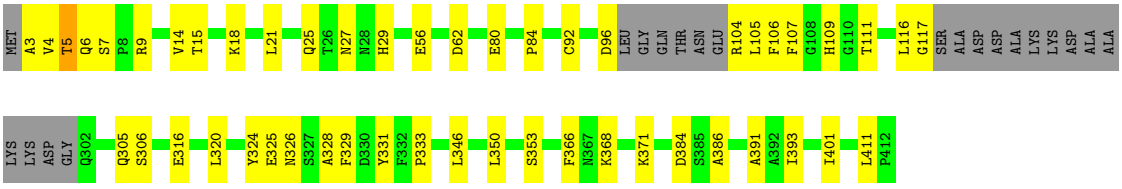
• Molecule 1: bispecific alpha/beta TCR

Chain C:



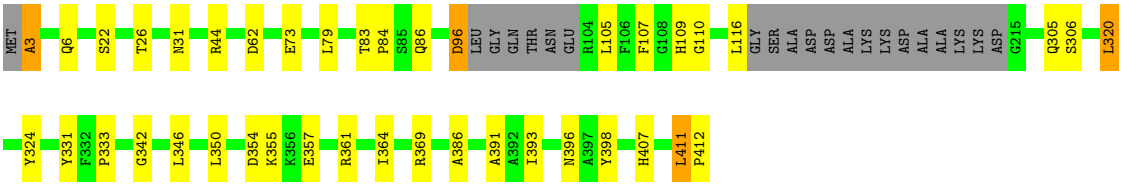
• Molecule 1: bispecific alpha/beta TCR

Chain E:



• Molecule 1: bispecific alpha/beta TCR

Chain G:


MET
A3
Q6
S22
T26
N31
R44
D62
E73
L79
T83
P84
S85
Q86
D96
LEU
GLY
GLN
THR
ASN
GLU
R104
L105
F106
F107
G108
H109
G110
L116
GLY
SER
ALA
ASP
ASP
ALA
LYS
LYS
ASP
ALA
ALA
LYS
LYS
ASP
G215
Q305
S306
L320
Y324
Y331
F332
P333
G342
L346
L350
D354
K355
K356
E357
R361
I364
R369
A386
A391
A392
I393
N396
A397
Y398
H407
L411
P412

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	71.68Å 105.48Å 81.53Å 90.00° 116.04° 90.00°	Depositor
Resolution (Å)	40.39 – 2.42 40.37 – 2.42	Depositor EDS
% Data completeness (in resolution range)	99.4 (40.39-2.42) 99.2 (40.37-2.42)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.33 (at 2.42Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.181 , 0.254 0.189 , 0.218	Depositor DCC
R_{free} test set	2073 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	31.4	Xtriage
Anisotropy	0.064	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 3.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.488 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7017	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.78% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	1.71	27/1821 (1.5%)	1.29	10/2465 (0.4%)
1	C	1.78	35/1814 (1.9%)	1.31	13/2459 (0.5%)
1	E	1.49	8/1717 (0.5%)	1.22	5/2329 (0.2%)
1	G	1.43	6/1733 (0.3%)	1.22	13/2350 (0.6%)
All	All	1.61	76/7085 (1.1%)	1.26	41/9603 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	C	0	2
1	E	0	2
1	G	0	1
All	All	0	6

The worst 5 of 76 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	73	GLU	CG-CD	11.55	1.69	1.51
1	A	73	GLU	CG-CD	9.81	1.66	1.51
1	A	34	TRP	CB-CG	8.53	1.65	1.50
1	E	324	TYR	CE2-CZ	8.24	1.49	1.38
1	A	390	CYS	CB-SG	8.19	1.96	1.82

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	354	ASP	CB-CG-OD1	10.71	127.94	118.30
1	A	306	SER	N-CA-CB	-8.39	97.91	110.50
1	A	358	ASP	CB-CG-OD2	8.24	125.72	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	358	ASP	CB-CG-OD2	8.22	125.70	118.30
1	E	384	ASP	CB-CG-OD1	8.06	125.56	118.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	104	ARG	Peptide
1	C	116	LEU	Peptide
1	C	16	GLY	Peptide
1	E	305	GLN	Peptide
1	E	5	THR	Peptide

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	1779	0	1697	24	0
1	C	1771	0	1671	16	0
1	E	1675	0	1590	21	0
1	G	1690	0	1608	13	0
2	A	25	0	0	0	0
2	C	15	0	0	1	0
2	E	30	0	0	0	0
2	G	32	0	0	1	0
All	All	7017	0	6566	73	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 73 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:C:104:ARG:O	1:C:104:ARG:HG3	1.64	0.98
1:E:25:GLN:HE22	1:E:29:HIS:H	1.12	0.94

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:16:GLY:HA3	1:G:3:ALA:N	1.88	0.88
1:A:15:THR:HG22	1:A:116:LEU:O	1.79	0.82
1:G:109:HIS:CD2	1:G:342:GLY:H	2.02	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	225/238 (94%)	216 (96%)	9 (4%)	0	100	100
1	C	227/238 (95%)	216 (95%)	11 (5%)	0	100	100
1	E	211/238 (89%)	202 (96%)	8 (4%)	1 (0%)	29	40
1	G	211/238 (89%)	198 (94%)	12 (6%)	1 (0%)	29	40
All	All	874/952 (92%)	832 (95%)	40 (5%)	2 (0%)	47	61

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	84	PRO
1	E	84	PRO

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	190/199 (96%)	177 (93%)	13 (7%)	16	24
1	C	188/199 (94%)	173 (92%)	15 (8%)	12	18
1	E	180/199 (90%)	174 (97%)	6 (3%)	38	56
1	G	182/199 (92%)	175 (96%)	7 (4%)	33	50
All	All	740/796 (93%)	699 (94%)	41 (6%)	21	33

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	7	SER
1	G	96	ASP
1	E	96	ASP
1	E	350	LEU
1	G	116	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	407	HIS
1	G	376	HIS
1	E	326	ASN
1	E	109	HIS
1	G	109	HIS

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 monosaccharides 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	118:SER	C	203:ALA	N	2.56

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	231/238 (97%)	-0.61	1 (0%)	92	91	11, 23, 54, 67	0
1	C	231/238 (97%)	-0.65	0	100	100	11, 23, 55, 64	0
1	E	217/238 (91%)	-0.66	0	100	100	12, 28, 50, 59	0
1	G	217/238 (91%)	-0.65	0	100	100	12, 28, 51, 58	0
All	All	896/952 (94%)	-0.64	1 (0%)	95	95	11, 26, 53, 67	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	203	ALA	3.0

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 monosaccharides 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.