wwPDB/EMDataBank EM Map/Model Validation Summary Report (i)

Sep 7, 2019 – 10:31 PM EDT

PDB ID	:	6JXR
EMDB ID:	:	EMD-9895
Title	:	Structure of a protein
Authors	:	Dong, D.; Zheng, L.; Lin, J.; Zhu, Y.; Li, N.; Zhang, B.; Xie, S.; Zheng, J.; Wang, Y.; Gao, N.; Huang, Z.
Deposited on	:	2019-04-24
Resolution	:	3.70 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

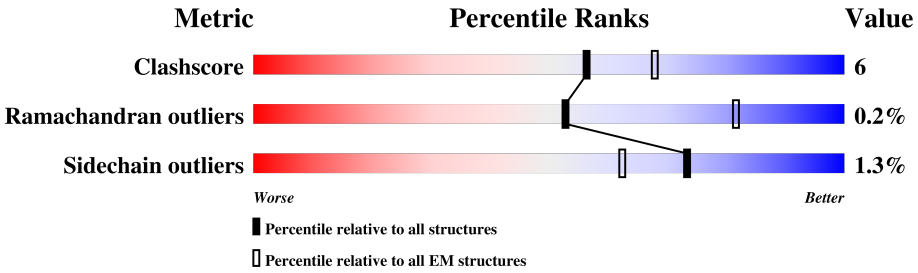
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.4

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.70 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)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 <=5%

Mol	Chain	Length	Quality of chain
1	a	164	21% · 78%
1	b	164	18% · 80%
2	d	171	63% 37%
3	e	207	59% 41%
3	f	207	59% 40%
4	g	182	62% · 37%
5	m	252	94% · ·
6	n	291	98% ··

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 8437 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 surface glycoprotein CD3 zeta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	a	36	Total	C	N	O	S	0	0
			286	197	42	46	1		
1	b	32	Total	C	N	O	S	0	0
			260	183	37	39	1		

- Molecule 2 is a protein called T-cell surface glycoprotein CD3 delta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	d	108	Total	C	N	O	S	0	0
			840	534	140	160	6		

- Molecule 3 is a protein called T-cell surface glycoprotein CD3 epsilon chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	f	124	Total	C	N	O	S	0	0
			983	622	158	195	8		
3	e	123	Total	C	N	O	S	0	0
			978	619	157	194	8		

- Molecule 4 is a protein called T-cell surface glycoprotein CD3 gamma chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	g	115	Total	C	N	O	S	0	0
			904	582	149	166	7		

- Molecule 5 is a protein called T cell receptor alpha variable 12-3,Possible J 11 gene segment,T cell receptor alpha constant.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	m	248	Total	C	N	O	S	0	0
			1940	1226	314	388	12		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
m	115	LYS	-	linker	UNP A0A0B4J271
m	133	ASP	-	linker	UNP A0N4Z6

- Molecule 6 is a protein called T cell receptor beta variable 6-5,M1-specific T cell receptor beta chain,T cell receptor beta constant 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	n	287	Total	C	N	O	S	0	0
			2246	1416	389	430	11		

There are 8 discrepancies between the modelled and reference sequences:

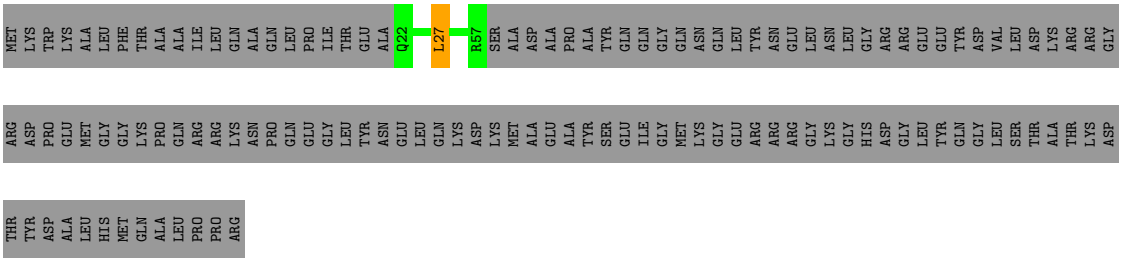
Chain	Residue	Modelled	Actual	Comment	Reference
n	113	ARG	-	linker	UNP A0A0K0K1A5
n	114	ARG	-	linker	UNP A0A0K0K1A5
n	115	ARG	-	linker	UNP A0A0K0K1A5
n	116	GLN	-	linker	UNP A0A0K0K1A5
n	117	GLY	-	linker	UNP A0A0K0K1A5
n	118	ALA	-	linker	UNP A0A0K0K1A5
n	119	SER	-	linker	UNP A0A0K0K1A5
n	120	GLY	-	linker	UNP A0A0K0K1A5

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. 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. 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.

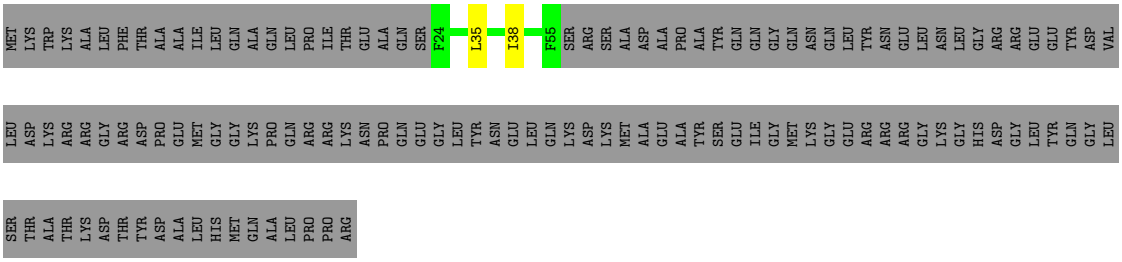
• Molecule 1: T-cell surface glycoprotein CD3 zeta chain

Chain a:



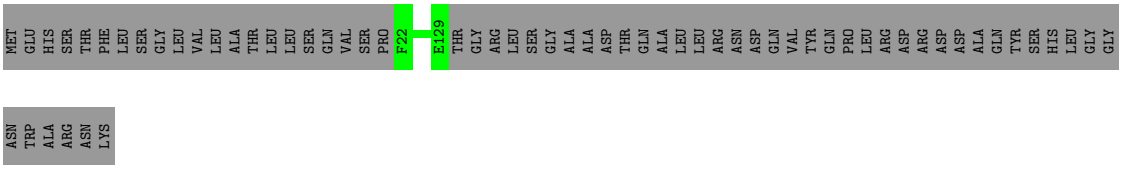
• Molecule 1: T-cell surface glycoprotein CD3 zeta chain

Chain b:



• Molecule 2: T-cell surface glycoprotein CD3 delta chain

Chain d:



• Molecule 3: T-cell surface glycoprotein CD3 epsilon chain

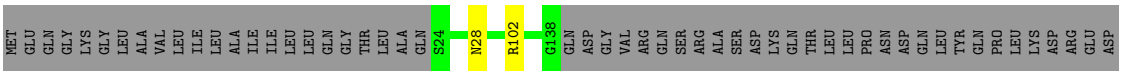
Chain f:

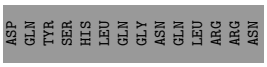
• Molecule 3: T-cell surface glycoprotein CD3 epsilon chain

Chain e:

• Molecule 4: T-cell surface glycoprotein CD3 gamma chain

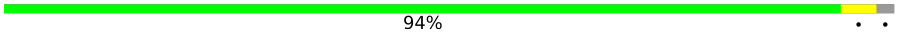
Chain g:





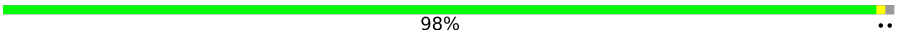
• Molecule 5: T cell receptor alpha variable 12-3,Possible J 11 gene segment,T cell receptor alpha constant

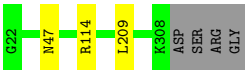
Chain m:



• Molecule 6: T cell receptor beta variable 6-5,M1-specific T cell receptor beta chain,T cell receptor beta constant 2

Chain n:





4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	197487	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	64.4	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

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	#$\|Z\|$ >2	RMSZ	#$\|Z\|$ >2
1	a	0.36	0/292	0.63	1/394 (0.3%)
1	b	0.35	0/266	0.74	0/359
2	d	0.43	0/853	0.66	0/1159
3	e	0.38	0/999	0.56	0/1356
3	f	0.40	0/1004	0.53	0/1363
4	g	0.39	0/922	0.59	0/1240
5	m	0.39	1/1982 (0.1%)	0.60	1/2685 (0.0%)
6	n	0.45	0/2301	0.59	1/3131 (0.0%)
All	All	0.41	1/8619 (0.0%)	0.59	3/11687 (0.0%)

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	b	0	1
5	m	0	2
All	All	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	m	190	LYS	CE-NZ	5.35	1.62	1.49

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	a	27	LEU	CA-CB-CG	5.73	128.47	115.30
6	n	209	LEU	CA-CB-CG	5.72	128.46	115.30
5	m	158	THR	C-N-CA	5.11	134.48	121.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	b	38	ILE	Peptide
5	m	159	ASP	Peptide
5	m	217	ASP	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	286	0	302	0	0
1	b	260	0	282	0	0
2	d	840	0	844	0	0
3	e	978	0	938	0	0
3	f	983	0	940	0	0
4	g	904	0	901	0	0
5	m	1940	0	1870	0	0
6	n	2246	0	2166	0	0
All	All	8437	0	8243	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 6.

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 PDB entries followed by that with respect to all EM entries.

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	34/164 (21%)	33 (97%)	1 (3%)	0	100	100
1	b	30/164 (18%)	25 (83%)	5 (17%)	0	100	100
2	d	106/171 (62%)	93 (88%)	13 (12%)	0	100	100
3	e	121/207 (58%)	113 (93%)	8 (7%)	0	100	100
3	f	122/207 (59%)	115 (94%)	7 (6%)	0	100	100
4	g	113/182 (62%)	104 (92%)	9 (8%)	0	100	100
5	m	246/252 (98%)	224 (91%)	20 (8%)	2 (1%)	21	63
6	n	285/291 (98%)	272 (95%)	13 (5%)	0	100	100
All	All	1057/1638 (64%)	979 (93%)	76 (7%)	2 (0%)	53	83

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	m	160	PHE
5	m	218	THR

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 PDB entries followed by that with respect to all EM entries.

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	31/135 (23%)	30 (97%)	1 (3%)	42	73
1	b	28/135 (21%)	27 (96%)	1 (4%)	38	71
2	d	95/147 (65%)	95 (100%)	0	100	100
3	e	111/177 (63%)	110 (99%)	1 (1%)	81	91
3	f	111/177 (63%)	110 (99%)	1 (1%)	81	91
4	g	97/155 (63%)	95 (98%)	2 (2%)	56	81
5	m	224/230 (97%)	220 (98%)	4 (2%)	62	84
6	n	244/249 (98%)	242 (99%)	2 (1%)	83	92
All	All	941/1405 (67%)	929 (99%)	12 (1%)	73	88

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	m	185	ARG
5	m	199	ASN
6	n	47	ASN
4	g	102	ARG
5	m	254	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
5	m	199	ASN
5	m	211	ASN
6	n	122	GLN
5	m	192	ASN
6	n	104	GLN

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.