W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Aug 3, 2023 – 01:10 AM EDT

PDB ID	:	1FO0
Title	:	MURINE ALLOREACTIVE SCFV TCR-PEPTIDE-MHC CLASS I MOLECULE COMPLEX
Authors	:	Reiser, J.B.; Darnault, C.; Guimezanes, A.; Gregoire, C.; Mosser, T.; Schmitt-Verhulst, A.-M.; Fontecilla-Camps, J.C.; Malissen, B.; Housset, D.; Mazza, G.
Deposited on	:	2000-08-24
Resolution	:	2.50 Å(reported)

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

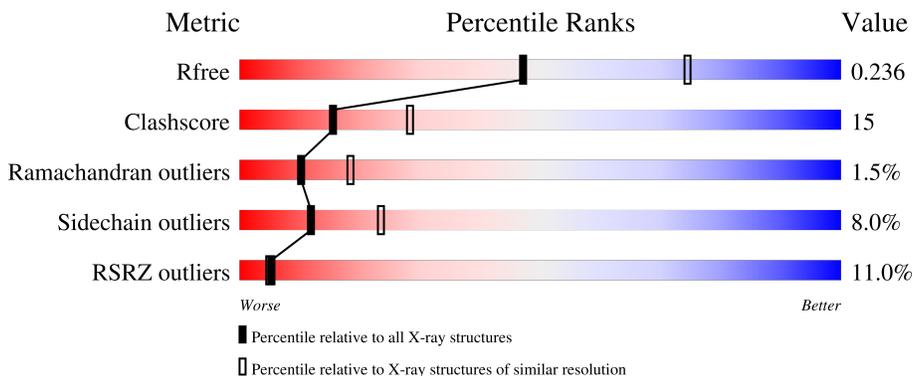|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.34
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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 2.50 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	Percentile Ranks	Value
Rfree		0.236
Clashscore		15
Ramachandran outliers		1.5%
Sidechain outliers		8.0%
RSRZ outliers		11.0%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	H	276	14% / 61% / 36% · ·
2	L	99	8% / 64% / 34% ·
3	P	8	75% / 25%
4	A	116	9% / 63% / 33% · · ·
5	B	112	8% / 65% / 27% / 7% ·

WORLDWIDE PDB
PROTEIN DATA BANK

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 5139 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 PROTEIN (ALLOGENEIC H-2KB MHC CLASS I MOLECULE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	276	Total 2248	C 1418	N 395	O 425	S 10	0	0	0

- Molecule 2 is a protein called PROTEIN (BETA-2 MICROGLOBULIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	99	Total 821	C 524	N 138	O 152	S 7	0	0	0

- Molecule 3 is a protein called NATURALLY PROCESSED OCTAPEPTIDE PBM1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	P	8	Total 70	C 46	N 10	O 14	0	0	0

- Molecule 4 is a protein called PROTEIN (BM3.3 T CELL RECEPTOR ALPHA-CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	115	Total 899	C 574	N 143	O 177	S 5	0	0	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	27	GLN	ARG	conflict	GB 201157
A	95	GLY	-	insertion	GB 201157
A	96	ASP	-	insertion	GB 201157
A	99	GLY	-	insertion	GB 201157
A	100	SER	-	insertion	GB 201157
A	101	GLY	ASN	conflict	GB 201157
A	102	ASN	GLU	conflict	GB 201157
A	104	LEU	ILE	conflict	GB 201157

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	105	ILE	THR	conflict	GB 201157
A	108	THR	ALA	conflict	GB 201157
A	111	LEU	LYS	conflict	GB 201157
A	113	SER	THR	conflict	GB 201157
A	114	VAL	ILE	conflict	GB 201157

- Molecule 5 is a protein called PROTEIN (BM3.3 T CELL RECEPTOR BETA-CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	B	112	910	575	165	165	5	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	96	ALA	GLY	conflict	GB 554307
B	97	ASP	GLY	conflict	GB 554307
B	98	ARG	THR	conflict	GB 554307
B	99	VAL	GLY	conflict	GB 554307
B	?	-	ALA	deletion	GB 554307
B	106	LEU	GLN	conflict	GB 554307
B	110	GLU	PRO	conflict	GB 554307
B	112	SER	THR	conflict	GB 554307
B	115	ILE	LEU	conflict	GB 554307
B	116	VAL	LEU	conflict	GB 554307

- Molecule 6 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	H	71	Total 71	O 71	0	0
6	L	27	Total 27	O 27	0	0
6	P	6	Total 6	O 6	0	0
6	A	42	Total 42	O 42	0	0
6	B	45	Total 45	O 45	0	0

3 Residue-property plots ⓘ

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.
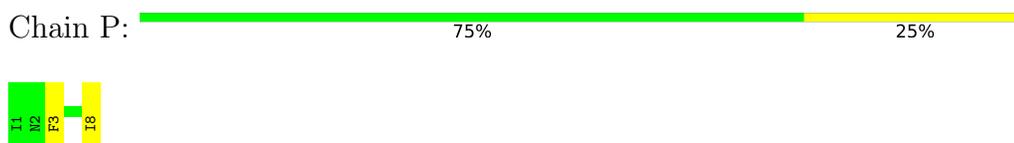
- Molecule 1: PROTEIN (ALLOGENEIC H-2KB MHC CLASS I MOLECULE)
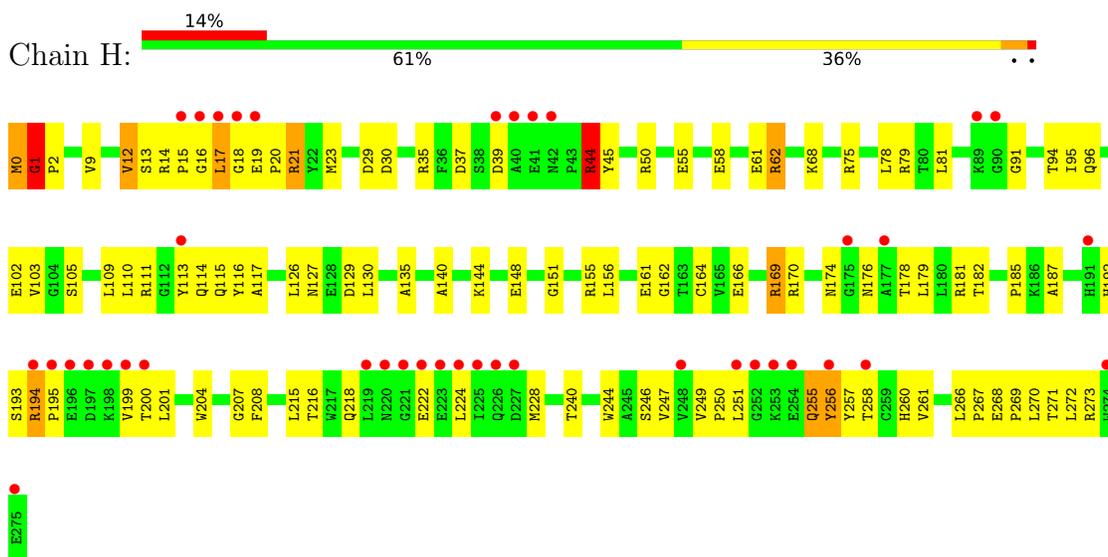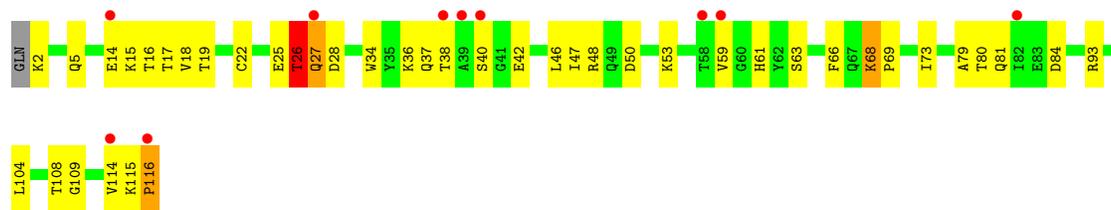

- Molecule 2: PROTEIN (BETA-2 MICROGLOBULIN)
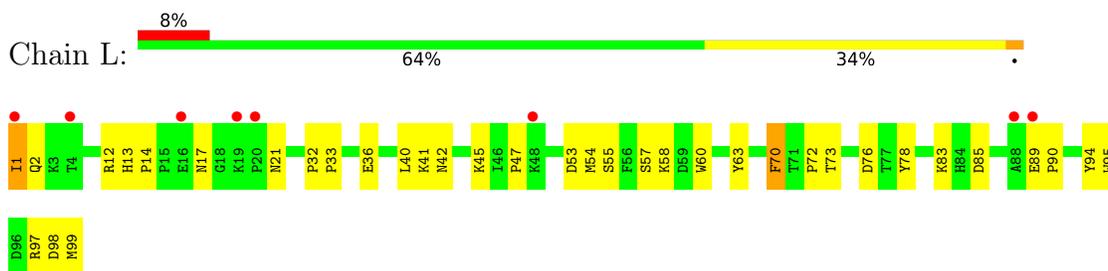
- Molecule 3: NATURALLY PROCESSED OCTAPEPTIDE PBM1

- Molecule 4: PROTEIN (BM3.3 T CELL RECEPTOR ALPHA-CHAIN)

• Molecule 5: PROTEIN (BM3.3 T CELL RECEPTOR BETA-CHAIN)

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	76.58Å 120.42Å 102.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	12.00 – 2.50 23.06 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (12.00-2.50) 98.7 (23.06-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	3.84 (at 2.50Å)	Xtriage
Refinement program	REFMAC 4.0	Depositor
R, R_{free}	0.225 , 0.276 0.216 , 0.236	Depositor DCC
R_{free} test set	3384 reflections (10.08%)	wwPDB-VP
Wilson B-factor (Å^2)	35.5	Xtriage
Anisotropy	0.095	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 60.8	EDS
L-test for twinning [2]	$< |L| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5139	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.48% 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 ⓘ

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	#$|Z|$>5	RMSZ	#$|Z|$>5
1	H	0.60	1/2309 (0.0%)	1.36	15/3135 (0.5%)
2	L	0.59	0/847	1.28	3/1148 (0.3%)
3	P	0.77	0/71	1.57	2/94 (2.1%)
4	A	0.62	0/916	1.27	4/1237 (0.3%)
5	B	0.69	0/929	1.45	7/1261 (0.6%)
All	All	0.62	1/5072 (0.0%)	1.35	31/6875 (0.5%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	1	GLY	N-CA	-5.51	1.37	1.46

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	H	44	ARG	NE-CZ-NH2	-14.60	113.00	120.30
5	B	86	ARG	CD-NE-CZ	12.41	140.98	123.60
5	B	98	ARG	NE-CZ-NH1	-7.84	116.38	120.30
1	H	1	GLY	N-CA-C	7.24	131.21	113.10
1	H	62	ARG	NE-CZ-NH1	-7.15	116.73	120.30

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	H	2248	0	2140	69	0
2	L	821	0	796	25	0
3	P	70	0	65	2	0
4	A	899	0	889	26	0
5	B	910	0	914	26	0
6	A	42	0	0	3	0
6	B	45	0	0	0	0
6	H	71	0	0	3	0
6	L	27	0	0	0	0
6	P	6	0	0	0	0
All	All	5139	0	4804	142	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 15.

The worst 5 of 142 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:L:1:ILE:HG22	2:L:2:GLN:H	1.23	1.02
4:A:50:ASP:HB2	4:A:53:LYS:HG3	1.57	0.85
1:H:19:GLU:HG3	1:H:20:PRO:HD2	1.59	0.84
5:B:37:GLN:HB2	5:B:43:LEU:HD13	1.64	0.79
4:A:68:LYS:HB3	4:A:69:PRO:HD3	1.63	0.78

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	H	274/276 (99%)	250 (91%)	18 (7%)	6 (2%)	6	10
2	L	97/99 (98%)	94 (97%)	2 (2%)	1 (1%)	15	28

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	P	6/8 (75%)	4 (67%)	2 (33%)	0	100	100
4	A	113/116 (97%)	102 (90%)	10 (9%)	1 (1%)	17	31
5	B	110/112 (98%)	105 (96%)	4 (4%)	1 (1%)	17	31
All	All	600/611 (98%)	555 (92%)	36 (6%)	9 (2%)	10	18

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	1	GLY
1	H	195	PRO
1	H	255	GLN
1	H	18	GLY
1	H	176	ASN

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	H	234/234 (100%)	214 (92%)	20 (8%)	10	21
2	L	94/94 (100%)	89 (95%)	5 (5%)	22	43
3	P	8/8 (100%)	8 (100%)	0	100	100
4	A	101/102 (99%)	95 (94%)	6 (6%)	19	37
5	B	100/100 (100%)	88 (88%)	12 (12%)	5	9
All	All	537/538 (100%)	494 (92%)	43 (8%)	12	23

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	63	SER
5	B	39	LEU
4	A	68	LYS
5	B	13	VAL
5	B	43	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	255	GLN
5	B	37	GLN
2	L	8	GLN
5	B	42	GLN
4	A	61	HIS

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 monosaccharides 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 ⓘ

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, 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(Å²)	Q<0.9
1	H	276/276 (100%)	1.01	40 (14%)	2	2	29, 50, 133, 167	35 (12%)
2	L	99/99 (100%)	0.50	8 (8%)	12	12	39, 63, 98, 122	6 (6%)
3	P	8/8 (100%)	0.27	0	100	100	31, 33, 39, 44	0
4	A	115/116 (99%)	0.28	10 (8%)	10	10	30, 50, 79, 85	8 (6%)
5	B	112/112 (100%)	0.32	9 (8%)	12	12	29, 47, 72, 90	7 (6%)
All	All	610/611 (99%)	0.65	67 (10%)	5	5	29, 52, 108, 167	56 (9%)

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	254	GLU	22.1
1	H	253	LYS	21.4
1	H	197	ASP	18.4
1	H	198	LYS	12.5
1	H	196	GLU	12.4

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

6.5　Other polymers ⓘ

There are no such residues in this entry.