

wwPDB X-ray Structure Validation Summary Report (i)

Mar 20, 2024 – 12:11 PM EDT

PDB ID	:	8EO8
Title	:	Cross-reactive 3180 TCR recognition of HLA-B*35:01-NP8 epitope from 2005
		H1N1 influenza strain
Authors	:	Littler, D.R.; Rossjohn, J.
Deposited on	:	2022-10-02
Resolution	:	2.30 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 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	:	2.36
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)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 2.30 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}	130704	5042 (2.30-2.30)			
Clashscore	141614	5643 (2.30-2.30)			
Ramachandran outliers	138981	5575 (2.30-2.30)			
Sidechain outliers	138945	5575 (2.30-2.30)			
RSRZ outliers	127900	4938 (2.30-2.30)			

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	Δ	276	9%	0%
1	11	210	6%	9% •
1	F	276	91%	9%
1	K	276	82% 9%	8%
			12%	
1	Р	276	89%	11%
2	В	100	90%	8% ••



Mol	Chain	Length	Quality of chain	
2	G	100	% • 88%	11% •
2	L	100	91%	9%
2	Q	100	88%	10% ••
3	С	9	89%	11%
3	Н	9	100%	
3	М	9	100%	
3	R	9	100%	
4	D	206	92%	7%
4	Ι	206	91%	8%
4	Ν	206	87%	9% •
4	S	206	6%	9% ••
5	Е	246	9%	19% ·
5	J	246	88%	10% •
5	Ο	246	88%	10% ••
5	Т	246	<u>4%</u> 87%	9% ••

Continued from previous page...



2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 28288 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.

Mol	Chain	Residues		Ate	oms			ZeroOcc	AltConf	Trace
1	Δ	275	Total	С	Ν	0	\mathbf{S}	0	1	0
	A	273	2258	1408	411	431	8	0	1	0
1	Б	275	Total	С	Ν	0	S	0	1	0
		275	2261	1409	414	431	7		1	
1	K	253	Total	С	Ν	0	S	0	0	0
			2073	1302	377	387	7	0	0	0
1 P	275	Total	С	Ν	0	S	0	0	0	
		2250	1403	410	430	$\overline{7}$	0	0	0	

• Molecule 1 is a protein called MHC class I antigen.

• Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
9	В	00	Total	С	Ν	Ο	\mathbf{S}	0	0	0
	D	99	829	528	140	158	3	0	0	0
0	С	00	Total	С	Ν	0	S	0	2	0
	G	99	848	540	145	160	3			0
0	т	100	Total	С	Ν	0	S	0	0	0
			837	533	141	159	4	0	0	U
0	2 0	00	Total	С	Ν	0	S	0	0	0
$2 \qquad Q$	99	829	528	140	158	3	0	0	0	

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
В	0	MET	-	initiating methionine	UNP P61769
G	0	MET	-	initiating methionine	UNP P61769
L	0	MET	-	initiating methionine	UNP P61769
Q	0	MET	-	initiating methionine	UNP P61769

• Molecule 3 is a protein called Nucleoprotein NP8 epitope.



Mol	Chain	Residues		Ate	oms			ZeroOcc	AltConf	Trace
2 0	С	0	Total	С	Ν	Ο	\mathbf{S}	0	0	0
0	U	9	72	48	10	13	1	0	0	0
3	3 H	0	Total	С	Ν	Ο	S	0	0	0
0		9	72	48	10	13	1	0	0	
2	М	9	Total	С	Ν	Ο	S	0	0	0
0	5 M		72	48	10	13	1	0	0	U
3 R	D	9	Total	С	Ν	Ο	S	0	0	0
	n		72	48	10	13	1	0	0	U

• Molecule 4 is a protein called 3180 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	л	206	Total	С	Ν	0	S	0	0	0
4	D	200	1597	995	267	328	7	0	0	0
4	т	206	Total	С	Ν	0	S	0	0	0
4	1	200	1597	995	267	328	$\overline{7}$		0	0
4	N	199	Total	С	Ν	0	S	0	0	0
4	4 IN		1539	956	260	316	$\overline{7}$	0	0	
4	4 S	199	Total	С	Ν	0	S	0	0	0
4			1539	956	260	316	7	0	U	U

• Molecule 5 is a protein called 3180 beta chain.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
Б	F	246	Total	С	Ν	0	\mathbf{S}	0	0	0
0		240	1942	1222	335	376	9	0	0	0
5	т	246	Total	С	Ν	0	S	0	2	0
0	1	240	1961	1233	340	378	10	0		
5	0	946	Total	С	Ν	0	S	0	0	0
0	0	240	1942	1222	335	376	9		0	0
Б	F T	244	Total	С	Ν	0	S	0	0	0
	244	1928	1215	333	371	9	0	0	0	

• Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	А	151	Total O 151 151	0	0
6	В	77	Total O 77 77	0	0
6	С	6	Total O 6 6	0	0



Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	D	70	Total O 70 70	0	0
6	Е	89	Total O 89 89	0	0
6	F	190	Total O 190 190	0	0
6	G	86	Total O 86 86	0	0
6	Н	7	Total O 7 7	0	0
6	Ι	95	Total O 95 95	0	0
6	J	104	Total O 104 104	0	0
6	К	189	Total O 189 189	0	0
6	L	84	Total O 84 84	0	0
6	М	8	Total O 8 8	0	0
6	Ν	96	Total O 96 96	0	0
6	О	114	Total O 114 114	0	0
6	Р	162	Total O 162 162	0	0
6	Q	67	Total O 67 67	0	0
6	R	9	Total O 9 9	0	0
6	S	81	Total O 81 81	0	0
6	Т	85	Total O 85 85	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: MHC class I antigen















4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	168.09Å 79.47Å 172.17Å	Deperitor
a, b, c, α , β , γ	90.00° 95.53° 90.00°	Depositor
$\mathbf{P}_{\text{assolution}}(\hat{\mathbf{A}})$	48.95 - 2.30	Depositor
Resolution (A)	48.95 - 2.30	EDS
% Data completeness	99.9 (48.95-2.30)	Depositor
(in resolution range)	$100.0 \ (48.95 - 2.30)$	EDS
R _{merge}	(Not available)	Depositor
R _{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$2.42 (at 2.29 \text{\AA})$	Xtriage
Refinement program	BUSTER, PHENIX 1.10.1_2155	Depositor
D D.	0.181 , 0.219	Depositor
Π, Π_{free}	0.186 , 0.222	DCC
R_{free} test set	10122 reflections $(5.03%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	33.7	Xtriage
Anisotropy	0.165	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.31 , 52.4	EDS
L-test for twinning ²	$< L > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.000 for l,-k,h	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	28288	wwPDB-VP
Average B, all atoms $(Å^2)$	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The analyses of the Patterson function reveals a significant off-origin peak that is 34.21 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.0722e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



¹Intensities estimated from amplitudes.

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	Bo	ond angles
	Unam	RMSZ	# Z > 5	RMSZ	# Z > 5
1	А	0.53	0/2321	0.67	0/3155
1	F	0.53	0/2324	0.66	0/3159
1	K	0.51	0/2130	0.66	0/2893
1	Р	0.49	0/2313	0.65	0/3145
2	В	0.52	0/852	0.69	0/1152
2	G	0.55	0/871	0.69	0/1177
2	L	0.51	0/860	0.68	0/1162
2	Q	0.51	0/852	0.67	0/1152
3	С	0.62	0/73	0.65	0/96
3	Н	0.62	0/73	0.65	0/96
3	М	0.66	0/73	0.69	0/96
3	R	0.68	0/73	0.65	0/96
4	D	0.48	0/1631	0.68	0/2210
4	Ι	0.50	0/1631	0.69	0/2210
4	N	0.50	0/1569	0.69	0/2125
4	S	0.50	0/1569	0.70	0/2125
5	Е	0.52	0/1993	0.73	0/2712
5	J	0.54	0/2012	0.69	0/2736
5	0	0.53	0/1993	0.72	1/2712~(0.0%)
5	Т	0.54	0/1979	0.73	1/2694~(0.0%)
All	All	0.52	0/27192	0.69	2/36903~(0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
5	0	190	LEU	CA-CB-CG	6.93	131.24	115.30
5	Т	188	GLN	C-N-CD	5.10	139.12	128.40

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	А	2258	0	2119	12	0
1	F	2261	0	2123	10	0
1	K	2073	0	1962	14	0
1	Р	2250	0	2111	11	0
2	В	829	0	794	3	0
2	G	848	0	818	6	0
2	L	837	0	803	2	0
2	Q	829	0	794	3	0
3	С	72	0	78	1	0
3	Н	72	0	78	0	0
3	М	72	0	78	0	0
3	R	72	0	78	0	0
4	D	1597	0	1522	9	0
4	Ι	1597	0	1522	7	0
4	N	1539	0	1474	8	0
4	S	1539	0	1474	5	0
5	Е	1942	0	1861	29	0
5	J	1961	0	1881	16	0
5	0	1942	0	1861	12	0
5	Т	1928	0	1852	18	0
6	А	151	0	0	1	0
6	В	77	0	0	0	0
6	С	6	0	0	0	0
6	D	70	0	0	0	0
6	Е	89	0	0	1	0
6	F	190	0	0	1	0
6	G	86	0	0	0	0
6	Н	7	0	0	0	0
6	Ι	95	0	0	0	0
6	J	104	0	0	2	0
6	K	189	0	0	1	0
6	L	84	0	0	0	0
6	М	8	0	0	0	0
6	N	96	0	0	0	0
6	0	114	0	0	0	0
6	Р	162	0	0	1	0
6	Q	67	0	0	0	0



001000	Show and from proceed as pagem							
Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes		
6	R	9	0	0	0	0		
6	S	81	0	0	0	0		
6	Т	85	0	0	0	0		
All	All	28288	0	25283	147	0		

Continued from previous page...

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 3.

The worst 5 of 147 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)
4:D:13:VAL:O	4:D:126:ILE:HG12	1.73	0.88
4:D:14:GLN:HA	4:D:126:ILE:HG13	1.58	0.83
2:L:48:LYS:O	2:L:68:THR:HG22	1.81	0.80
5:E:236:TRP:HZ2	5:E:243:PRO:HD3	1.46	0.79
5:J:3:VAL:CG1	5:J:26:LEU:HD12	2.15	0.76

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	Perce	ntiles
1	А	274/276~(99%)	261 (95%)	13 (5%)	0	100	100
1	F	274/276~(99%)	263~(96%)	11 (4%)	0	100	100
1	Κ	245/276~(89%)	238~(97%)	7 (3%)	0	100	100
1	Р	273/276~(99%)	264 (97%)	8 (3%)	1 (0%)	34	42
2	В	97/100~(97%)	97~(100%)	0	0	100	100
2	G	98/100~(98%)	96~(98%)	2 (2%)	0	100	100
2	L	98/100~(98%)	98 (100%)	0	0	100	100



8EO8	
------	--

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Perce	ntiles
2	Q	97/100~(97%)	97 (100%)	0	0	100	100
3	С	7/9~(78%)	7 (100%)	0	0	100	100
3	Н	7/9~(78%)	7 (100%)	0	0	100	100
3	М	7/9~(78%)	7 (100%)	0	0	100	100
3	R	7/9~(78%)	7 (100%)	0	0	100	100
4	D	204/206~(99%)	198 (97%)	6 (3%)	0	100	100
4	Ι	204/206~(99%)	198 (97%)	6 (3%)	0	100	100
4	Ν	197/206~(96%)	188 (95%)	9 (5%)	0	100	100
4	S	197/206~(96%)	189 (96%)	7 (4%)	1 (0%)	29	35
5	Е	244/246~(99%)	231 (95%)	11 (4%)	2 (1%)	19	23
5	J	246/246~(100%)	236 (96%)	10 (4%)	0	100	100
5	Ο	244/246~(99%)	237 (97%)	6 (2%)	1 (0%)	34	42
5	Т	242/246~(98%)	233 (96%)	7 (3%)	2 (1%)	19	23
All	All	3262/3348~(97%)	3152 (97%)	103 (3%)	7 (0%)	47	58

Continued from previous page...

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	Е	48	GLN
4	S	145	SER
5	Т	63	SER
5	Т	45	PHE
5	Е	72	GLU

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain 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 side chain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percent	tiles
1	А	235/234~(100%)	223~(95%)	12~(5%)	24	33
1	F	235/234~(100%)	225~(96%)	10 (4%)	29	40
1	К	215/234~(92%)	211 (98%)	4 (2%)	57	73



8EO8

Mol	Chain	Analysed	Rotameric Outliers		Percentiles		
1	Р	234/234~(100%)	223~(95%)	11 (5%)		26	37
2	В	94/95~(99%)	87~(93%)	7 (7%)		13	17
2	G	96/95~(101%)	90~(94%)	6 (6%)		18	24
2	L	95/95~(100%)	90~(95%)	5 (5%)		22	31
2	Q	94/95~(99%)	87~(93%)	7 (7%)		13	17
3	С	8/8~(100%)	8 (100%)	0	1	.00	100
3	Н	8/8~(100%)	8 (100%)	0	1	.00	100
3	М	8/8~(100%)	8 (100%)	0	1	.00	100
3	R	8/8~(100%)	8 (100%)	0	1	.00	100
4	D	182/182~(100%)	175 (96%)	7 (4%)		33	47
4	Ι	182/182~(100%)	173~(95%)	9~(5%)		25	35
4	Ν	175/182~(96%)	165 (94%)	10 (6%)		20	28
4	S	175/182~(96%)	163~(93%)	12 (7%)		15	20
5	Ε	216/216~(100%)	196 (91%)	20 (9%)		9	10
5	J	218/216~(101%)	201~(92%)	17 (8%)		12	16
5	Ο	216/216~(100%)	200 (93%)	16 (7%)		13	17
5	Т	215/216~(100%)	197 (92%)	18 (8%)		11	13
All	All	2909/2940~(99%)	2738 (94%)	171 (6%)		19	27

Continued from previous page...

5 of 171 residues with a non-rotameric sidechain are listed below:

Mol	Chain	\mathbf{Res}	Type
5	0	93	SER
4	S	28	SER
5	0	190	LEU
1	Р	231	VAL
4	S	156	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	А	192	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)

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, 95^{th} 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	$\langle RSRZ \rangle$	#RSRZ>2	$OWAB(Å^2)$	Q<0.9
1	А	275/276~(99%)	0.27	25 (9%) 9 12	21, 38, 101, 121	0
1	F	275/276~(99%)	0.19	17 (6%) 20 26	17, 31, 87, 110	1 (0%)
1	Κ	253/276~(91%)	0.23	22 (8%) 10 14	18, 35, 98, 119	0
1	Р	275/276~(99%)	0.27	34 (12%) 4 5	18, 39, 105, 124	0
2	В	99/100 (99%)	-0.29	1 (1%) 82 86	29, 45, 68, 75	0
2	G	99/100~(99%)	-0.39	1 (1%) 82 86	23, 38, 59, 72	0
2	L	100/100~(100%)	-0.24	1 (1%) 82 86	23, 41, 64, 74	0
2	Q	99/100~(99%)	-0.00	4 (4%) 38 45	29, 49, 74, 88	0
3	С	9/9~(100%)	-0.26	0 100 100	24, 25, 31, 32	0
3	Н	9/9~(100%)	0.01	0 100 100	19, 20, 25, 28	0
3	М	9/9~(100%)	0.03	0 100 100	19, 21, 24, 28	0
3	R	9/9~(100%)	-0.13	0 100 100	20, 22, 25, 32	0
4	D	206/206~(100%)	0.11	13 (6%) 20 25	25, 58, 98, 112	0
4	Ι	206/206~(100%)	-0.02	7 (3%) 45 52	21, 50, 86, 105	0
4	Ν	199/206~(96%)	-0.06	6 (3%) 50 57	21, 47, 84, 102	0
4	S	199/206~(96%)	0.07	12 (6%) 21 28	21, 53, 97, 118	0
5	Е	246/246~(100%)	0.53	21 (8%) 10 14	21, 63, 109, 133	1 (0%)
5	J	246/246~(100%)	0.03	8 (3%) 46 53	19, 49, 88, 108	0
5	0	246/246~(100%)	0.01	11 (4%) 33 40	18, 45, 83, 109	2 (0%)
5	Т	244/246~(99%)	-0.01	10 (4%) 37 44	18, 52, 92, 113	1 (0%)
All	All	3303/3348 (98%)	0.10	193 (5%) 23 29	17, 45, 93, 133	5 (0%)

The worst 5 of 193 RSRZ outliers are listed below:



Mol	Chain	Res	Type	RSRZ
5	Ε	238	GLN	7.2
1	Κ	191	HIS	6.6
1	Р	197	HIS	6.4
5	Е	257	ASP	6.2
4	S	198	PHE	5.9

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.

