Full wwPDB X-ray Structure Validation Report ⓘ

Dec 17, 2023 – 10:20 am GMT

PDB ID	:	2YEZ
Title	:	COMPLEX OF A B21 CHICKEN MHC CLASS I MOLECULE AND A 10MER CHICKEN PEPTIDE
Authors	:	Chappell, P.; Roversi, P.; Harrison, M.C.; Mears, L.E.; Kaufman, J.F.; Lea, S.M.
Deposited on	:	2011-03-31
Resolution	:	2.90 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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 ⓘ) were used in the production of this report:

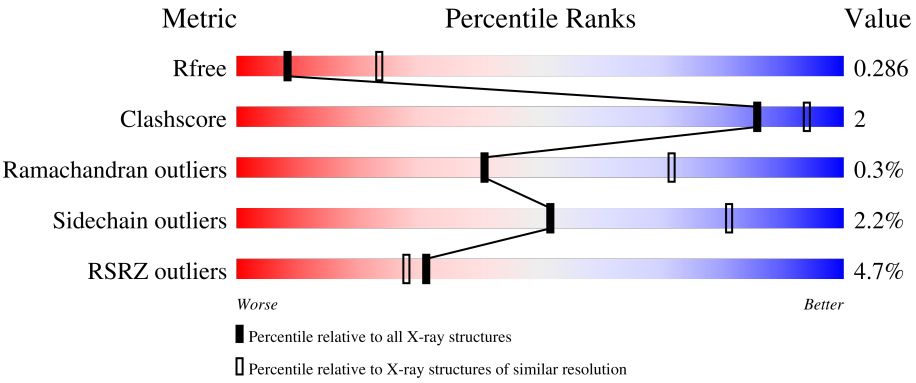
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	329	4% 79% 5% 16%
2	B	98	6% 92% 5%
3	C	10	90% 10%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3121 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 MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I GLYCOPROTEIN HAPLOTYPE B21.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	278	2237	1403	403	422	9	0	2	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	271	ARG	-	expression tag	UNP Q95601
A	272	SER	-	expression tag	UNP Q95601
A	273	GLY	-	expression tag	UNP Q95601
A	274	GLY	-	expression tag	UNP Q95601
A	275	GLY	-	expression tag	UNP Q95601
A	276	LEU	-	expression tag	UNP Q95601
A	277	ASN	-	expression tag	UNP Q95601
A	278	ASP	-	expression tag	UNP Q95601
A	279	ILE	-	expression tag	UNP Q95601
A	280	PHE	-	expression tag	UNP Q95601
A	281	GLU	-	expression tag	UNP Q95601
A	282	ALA	-	expression tag	UNP Q95601
A	283	GLN	-	expression tag	UNP Q95601
A	284	LYS	-	expression tag	UNP Q95601
A	285	ILE	-	expression tag	UNP Q95601
A	286	GLU	-	expression tag	UNP Q95601
A	287	TRP	-	expression tag	UNP Q95601
A	288	HIS	-	expression tag	UNP Q95601
A	289	GLU	-	expression tag	UNP Q95601
A	290	ASN	-	expression tag	UNP Q95601
A	291	SER	-	expression tag	UNP Q95601
A	292	SER	-	expression tag	UNP Q95601
A	293	SER	-	expression tag	UNP Q95601
A	294	VAL	-	expression tag	UNP Q95601
A	295	ASP	-	expression tag	UNP Q95601
A	296	LYS	-	expression tag	UNP Q95601

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	297	LEU	-	expression tag	UNP Q95601
A	298	ALA	-	expression tag	UNP Q95601
A	299	ALA	-	expression tag	UNP Q95601
A	300	ALA	-	expression tag	UNP Q95601
A	301	LEU	-	expression tag	UNP Q95601
A	302	GLU	-	expression tag	UNP Q95601
A	303	HIS	-	expression tag	UNP Q95601
A	304	HIS	-	expression tag	UNP Q95601
A	305	HIS	-	expression tag	UNP Q95601
A	306	HIS	-	expression tag	UNP Q95601
A	307	HIS	-	expression tag	UNP Q95601
A	308	HIS	-	expression tag	UNP Q95601

- Molecule 2 is a protein called BETA-2-MICROGLOBULIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	96	Total	C	N	O	S	0	0	0
			763	491	124	143	5			

- Molecule 3 is a protein called 14-3-3 PROTEIN THETA.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	10	Total	C	N	O	0	1	0
			91	60	12	19			

- Molecule 4 is water.

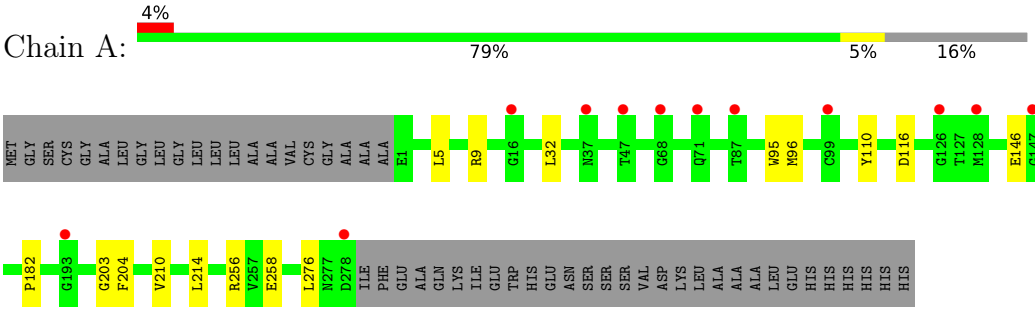
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	21	Total	O	0	0
			21	21		
4	B	8	Total	O	0	0
			8	8		
4	C	1	Total	O	0	0
			1	1		

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.

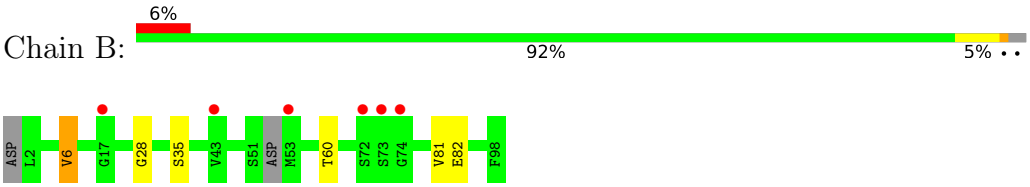
• Molecule 1: MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I GLYCOPROTEIN HAPLOTYPE B21

Chain A:



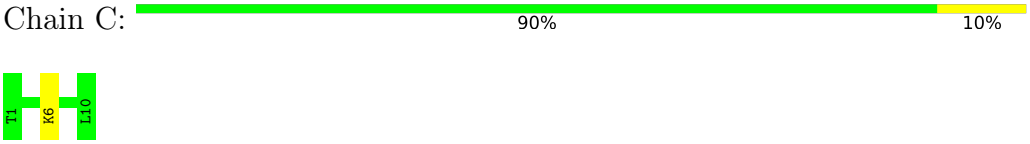
• Molecule 2: BETA-2-MICROGLOBULIN

Chain B:



• Molecule 3: 14-3-3 PROTEIN THETA

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	60.57Å 68.98Å 94.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.03 – 2.90 41.03 – 2.45	Depositor EDS
% Data completeness (in resolution range)	96.4 (41.03-2.90) 85.0 (41.03-2.45)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.86 (at 2.45Å)	Xtriage
Refinement program	BUSTER 2.9.5	Depositor
R, R_{free}	0.251 , 0.273 0.263 , 0.286	Depositor DCC
R_{free} test set	660 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å^2)	29.3	Xtriage
Anisotropy	0.758	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 20.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	3121	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.15% 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	A	0.41	0/2303	0.54	0/3128
2	B	0.40	0/787	0.52	0/1067
3	C	0.42	0/96	0.54	0/128
All	All	0.41	0/3186	0.54	0/4323

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

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	2237	0	2122	10	0
2	B	763	0	725	3	0
3	C	91	0	90	0	0
4	A	21	0	0	0	0
4	B	8	0	0	0	0
4	C	1	0	0	0	0
All	All	3121	0	2937	13	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 2.

All (13) 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:A:96[A]:MET:SD	1:A:110[A]:TYR:CE1	2.36	1.17
1:A:96[A]:MET:SD	1:A:110[A]:TYR:CD1	2.40	1.13
1:A:96[A]:MET:SD	1:A:110[A]:TYR:HE1	2.09	0.70
1:A:96[A]:MET:SD	1:A:110[A]:TYR:HD1	2.14	0.67
1:A:96[A]:MET:CE	1:A:110[A]:TYR:CD1	2.81	0.64
2:B:28:GLY:HA2	2:B:60:THR:HB	1.85	0.59
2:B:35:SER:HB3	2:B:82:GLU:HB2	1.88	0.56
1:A:5:LEU:HB2	1:A:165:LEU:HD13	1.88	0.54
1:A:9:ARG:HB2	1:A:95:TRP:HB2	1.90	0.54
1:A:210:VAL:HB	1:A:258:GLU:HB2	1.95	0.49
1:A:182:PRO:HB3	1:A:204:PHE:HB3	1.97	0.47
2:B:6:VAL:HG11	2:B:81:VAL:HG21	1.99	0.43
1:A:214:LEU:HD11	1:A:256:ARG:HB2	2.02	0.41

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	278/329 (84%)	268 (96%)	9 (3%)	1 (0%)	34 66
2	B	92/98 (94%)	88 (96%)	4 (4%)	0	100 100
3	C	9/10 (90%)	8 (89%)	1 (11%)	0	100 100
All	All	379/437 (87%)	364 (96%)	14 (4%)	1 (0%)	41 71

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	203	GLY

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	229/263 (87%)	224 (98%)	5 (2%)	52 81
2	B	84/86 (98%)	83 (99%)	1 (1%)	71 91
3	C	11/10 (110%)	10 (91%)	1 (9%)	9 28
All	All	324/359 (90%)	317 (98%)	7 (2%)	52 81

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	32	LEU
1	A	116	ASP
1	A	146	GLU
1	A	151	GLU
1	A	276	LEU
2	B	6	VAL
3	C	6	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	277	ASN
2	B	23	ASN

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

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, 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	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	278/329 (84%)	0.42	12 (4%) 35 31	10, 28, 56, 80	0
2	B	96/98 (97%)	0.52	6 (6%) 20 16	16, 32, 52, 75	0
3	C	10/10 (100%)	0.60	0 100 100	22, 28, 34, 38	0
All	All	384/437 (87%)	0.45	18 (4%) 31 28	10, 29, 55, 80	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	16	GLY	5.2
1	A	147	GLY	3.5
1	A	126	GLY	3.2
1	A	128	MET	2.9
2	B	74	GLY	2.9
1	A	278	ASP	2.8
1	A	99	CYS	2.7
2	B	53	MET	2.6
2	B	73	SER	2.5
1	A	37	ASN	2.4
2	B	72	SER	2.4
1	A	47	THR	2.3
1	A	87	THR	2.2
1	A	71	GLN	2.2
1	A	68	GLY	2.2
2	B	17	GLY	2.2
2	B	43	VAL	2.1
1	A	193	GLY	2.0

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 (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.