



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 10, 2019 – 04:55 PM EST

PDB ID : 6U3O  
Title : JR51 DQ2-p.aeru-alpha2a complex  
Authors : Petersen, J.; Rossjohn, J.  
Deposited on : 2019-08-22  
Resolution : 2.74 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : 2.4  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.4

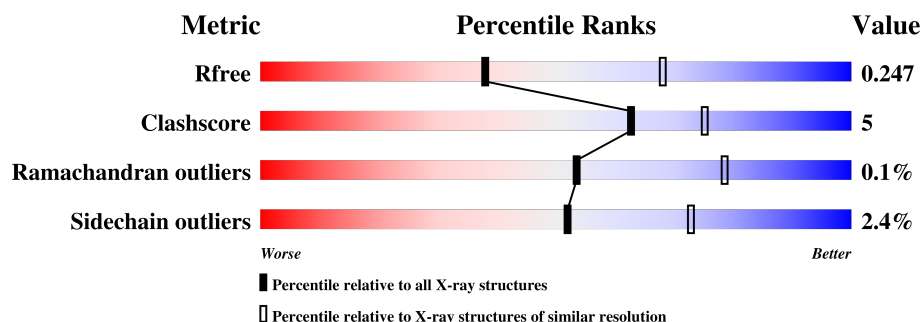
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.74 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))
$R_{free}$	111664	1033 (2.76-2.72)
Clashscore	122126	1084 (2.76-2.72)
Ramachandran outliers	120053	1064 (2.76-2.72)
Sidechain outliers	120020	1065 (2.76-2.72)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	202	81% 14% . .
1	G	202	81% 15% .
2	B	244	89% 10% ..
2	H	244	90% 9% .
3	C	191	85% 10% 5%
3	E	191	82% 12% . 5%
4	D	206	73% 14% . 12%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	F	206	
5	I	20	
5	J	20	

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12888 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 T-CELL RECEPTOR, JR5.1 alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	194	Total	C	N	O	S	0	0	0
			1480	924	254	292	10			
1	A	194	Total	C	N	O	S	0	0	0
			1490	931	254	295	10			

- Molecule 2 is a protein called T-CELL RECEPTOR, JR5.1 beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	242	Total	C	N	O	S	0	0	0
			1877	1182	328	362	5			
2	B	242	Total	C	N	O	S	0	0	0
			1877	1182	328	362	5			

- Molecule 3 is a protein called MHC class II HLA-DQ-alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	181	Total	C	N	O	S	0	0	0
			1445	931	236	276	2			
3	E	181	Total	C	N	O	S	0	0	0
			1445	931	236	276	2			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	44	SER	CYS	conflict	UNP O19705
C	182	THR	-	expression tag	UNP O19705
C	183	SER	-	expression tag	UNP O19705
C	184	GLY	-	expression tag	UNP O19705
C	185	ASP	-	expression tag	UNP O19705
C	186	ASP	-	expression tag	UNP O19705
C	187	ASP	-	expression tag	UNP O19705
C	188	ASP	-	expression tag	UNP O19705

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	189	LYS	-	expression tag	UNP O19705
E	44	SER	CYS	conflict	UNP O19705
E	182	THR	-	expression tag	UNP O19705
E	183	SER	-	expression tag	UNP O19705
E	184	GLY	-	expression tag	UNP O19705
E	185	ASP	-	expression tag	UNP O19705
E	186	ASP	-	expression tag	UNP O19705
E	187	ASP	-	expression tag	UNP O19705
E	188	ASP	-	expression tag	UNP O19705
E	189	LYS	-	expression tag	UNP O19705

- Molecule 4 is a protein called MHC class II HLA-DQ-beta-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	181	Total	C	N	O	S	0	0	0
			1480	935	264	274	7			
4	F	181	Total	C	N	O	S	0	0	0
			1480	935	264	274	7			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-5	GLY	-	expression tag	UNP O19712
D	-4	GLY	-	expression tag	UNP O19712
D	-3	SER	-	expression tag	UNP O19712
D	-2	GLY	-	expression tag	UNP O19712
D	-1	ALA	-	expression tag	UNP O19712
D	0	SER	-	expression tag	UNP O19712
D	193	THR	-	expression tag	UNP O19712
D	194	GLY	-	expression tag	UNP O19712
D	195	GLY	-	expression tag	UNP O19712
D	196	ASP	-	expression tag	UNP O19712
D	197	ASP	-	expression tag	UNP O19712
D	198	ASP	-	expression tag	UNP O19712
D	199	ASP	-	expression tag	UNP O19712
D	200	LYS	-	expression tag	UNP O19712
F	-5	GLY	-	expression tag	UNP O19712
F	-4	GLY	-	expression tag	UNP O19712
F	-3	SER	-	expression tag	UNP O19712
F	-2	GLY	-	expression tag	UNP O19712
F	-1	ALA	-	expression tag	UNP O19712
F	0	SER	-	expression tag	UNP O19712

*Continued on next page...*

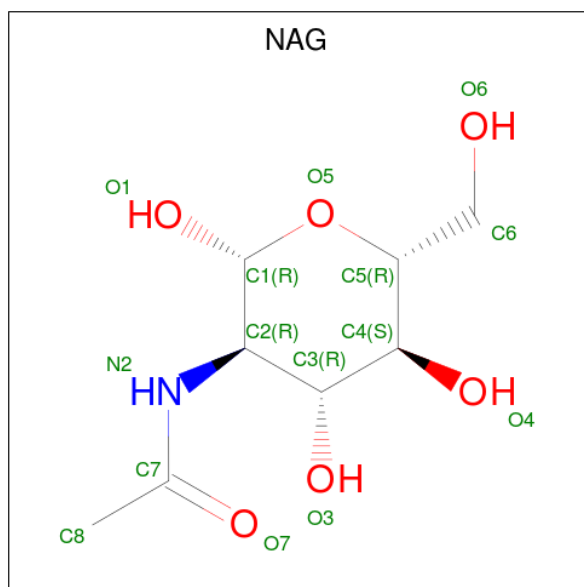
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	193	THR	-	expression tag	UNP O19712
F	194	GLY	-	expression tag	UNP O19712
F	195	GLY	-	expression tag	UNP O19712
F	196	ASP	-	expression tag	UNP O19712
F	197	ASP	-	expression tag	UNP O19712
F	198	ASP	-	expression tag	UNP O19712
F	199	ASP	-	expression tag	UNP O19712
F	200	LYS	-	expression tag	UNP O19712

- Molecule 5 is a protein called Peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	I	13	Total	C	N	O	0	0	0
			96	61	14	21			
5	J	13	Total	C	N	O	0	0	0
			96	61	14	21			

- Molecule 6 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	C	1	Total	C	N	O	0	0
			14	8	1	5		
6	C	1	Total	C	N	O	0	0
			14	8	1	5		
6	E	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	E	1	Total	C	N	O	0	0
			14	8	1	5		


- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	G	9	Total	O	0	0
			9	9		
7	H	6	Total	O	0	0
			6	6		
7	A	12	Total	O	0	0
			12	12		
7	B	2	Total	O	0	0
			2	2		
7	C	10	Total	O	0	0
			10	10		
7	D	9	Total	O	0	0
			9	9		
7	E	11	Total	O	0	0
			11	11		
7	F	6	Total	O	0	0
			6	6		
7	I	1	Total	O	0	0
			1	1		





- Molecule 3: MHC class II HLA-DQ-alpha chain

Chain E:  82% 12% 5%



- Molecule 4: MHC class II HLA-DQ-beta-1

Chain D:  73% 14% 12%



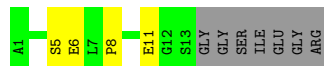
- Molecule 4: MHC class II HLA-DQ-beta-1

Chain F:  69% 18% 12%



- Molecule 5: Peptide

Chain I:  45% 20% 35%



- Molecule 5: Peptide

Chain J:  55% 10% 35%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.06Å 157.65Å 106.10Å 90.00° 96.53° 90.00°	Depositor
Resolution (Å)	47.03 – 2.74 48.12 – 2.74	Depositor EDS
% Data completeness (in resolution range)	99.3 (47.03-2.74) 99.3 (48.12-2.74)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.57 (at 2.73Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	(Not available) , (Not available) 0.214 , 0.247	Depositor DCC
$R_{free}$ test set	2000 reflections (3.42%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	61.6	Xtriage
Anisotropy	0.648	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 52.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	12888	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.13% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG

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	0.25	0/1523	0.47	0/2074
1	G	0.25	0/1512	0.49	0/2060
2	B	0.25	0/1926	0.46	0/2624
2	H	0.25	0/1926	0.45	0/2624
3	C	0.25	0/1487	0.44	0/2031
3	E	0.25	0/1487	0.44	0/2031
4	D	0.23	0/1513	0.43	0/2056
4	F	0.24	0/1513	0.43	0/2056
5	I	0.22	0/98	0.37	0/134
5	J	0.24	0/98	0.39	0/134
All	All	0.25	0/13083	0.45	0/17824

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	1490	0	1381	18	0
1	G	1480	0	1369	14	0
2	B	1877	0	1776	17	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	H	1877	0	1776	12	2
3	C	1445	0	1397	13	0
3	E	1445	0	1397	12	1
4	D	1480	0	1443	18	0
4	F	1480	0	1443	24	0
5	I	96	0	92	3	0
5	J	96	0	92	1	0
6	C	28	0	25	0	0
6	E	28	0	25	0	0
7	A	12	0	0	0	0
7	B	2	0	0	0	0
7	C	10	0	0	0	0
7	D	9	0	0	0	0
7	E	11	0	0	0	0
7	F	6	0	0	0	0
7	G	9	0	0	0	0
7	H	6	0	0	0	0
7	I	1	0	0	0	0
All	All	12888	0	12216	117	2

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 117 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:B:109:ARG:NH2	3:C:62:ASN:OD1	2.18	0.76
2:H:109:ARG:NH2	3:E:62:ASN:OD1	2.22	0.73
3:E:96:PRO:HD3	4:F:120:THR:HG21	1.74	0.70
4:D:10:GLN:HB2	4:D:31:ILE:HB	1.73	0.69
3:C:96:PRO:HD3	4:D:120:THR:HG21	1.74	0.69

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:210:SER:OG	3:E:159:GLU:OE1[1_556]	2.10	0.10
2:H:22:ARG:NH2	2:B:79:GLU:OE2[2_458]	2.19	0.01

## 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	A	190/202 (94%)	177 (93%)	12 (6%)	1 (0%)	31	53
1	G	190/202 (94%)	175 (92%)	15 (8%)	0	100	100
2	B	240/244 (98%)	232 (97%)	8 (3%)	0	100	100
2	H	240/244 (98%)	232 (97%)	8 (3%)	0	100	100
3	C	179/191 (94%)	177 (99%)	2 (1%)	0	100	100
3	E	179/191 (94%)	177 (99%)	2 (1%)	0	100	100
4	D	177/206 (86%)	171 (97%)	6 (3%)	0	100	100
4	F	177/206 (86%)	171 (97%)	6 (3%)	0	100	100
5	I	11/20 (55%)	11 (100%)	0	0	100	100
5	J	11/20 (55%)	11 (100%)	0	0	100	100
All	All	1594/1726 (92%)	1534 (96%)	59 (4%)	1 (0%)	53	78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	207	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	A	162/181 (90%)	157 (97%)	5 (3%)	43	66
1	G	159/181 (88%)	151 (95%)	8 (5%)	27	47

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	200/209 (96%)	195 (98%)	5 (2%)	50	72
2	H	200/209 (96%)	198 (99%)	2 (1%)	78	87
3	C	165/174 (95%)	164 (99%)	1 (1%)	87	92
3	E	165/174 (95%)	161 (98%)	4 (2%)	52	73
4	D	164/184 (89%)	161 (98%)	3 (2%)	62	79
4	F	164/184 (89%)	160 (98%)	4 (2%)	52	73
5	I	11/15 (73%)	10 (91%)	1 (9%)	10	20
5	J	11/15 (73%)	10 (91%)	1 (9%)	10	20
All	All	1401/1526 (92%)	1367 (98%)	34 (2%)	52	73

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	79	GLU
2	B	216	ASN
4	F	51	THR
2	B	156	LEU
1	G	177	VAL

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

Mol	Chain	Res	Type
2	B	216	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry

4 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
6	NAG	C	201	3,6	14,14,15	0.29	0	17,19,21	0.41	0
6	NAG	C	202	6	14,14,15	0.42	0	17,19,21	0.42	0
6	NAG	E	201	3,6	14,14,15	0.33	0	17,19,21	0.39	0
6	NAG	E	202	6	14,14,15	0.41	0	17,19,21	0.44	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	C	201	3,6	-	0/6/23/26	0/1/1/1
6	NAG	C	202	6	-	2/6/23/26	0/1/1/1
6	NAG	E	201	3,6	-	2/6/23/26	0/1/1/1
6	NAG	E	202	6	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	C	202	NAG	O5-C5-C6-O6
6	C	202	NAG	C4-C5-C6-O6
6	E	201	NAG	O5-C5-C6-O6
6	E	201	NAG	C4-C5-C6-O6
6	E	202	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.