



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2026 – 10:42 AM UTC

PDB ID : 1SJE / pdb_00001sje
Title : HLA-DR1 complexed with a 16 residue HIV capsid peptide bound in a hairpin conformation
Authors : Zavala-Ruiz, Z.; Strug, I.; Walker, B.D.; Norris, P.J.; Stern, L.J.
Deposited on : 2004-03-03
Resolution : 2.45 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 ⓘ 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

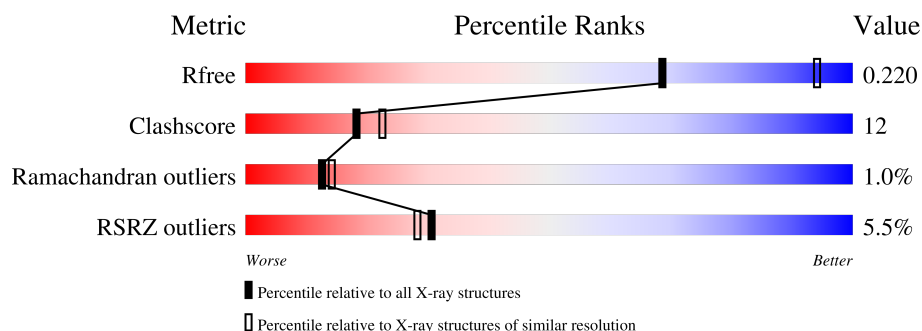
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.45 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}	180053	1190 (2.46-2.46)
Clashscore	190562	1229 (2.46-2.46)
Ramachandran outliers	187476	1218 (2.46-2.46)
RSRZ outliers	180081	1190 (2.46-2.46)

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 $\leq 5\%$. 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	180	4% 78% 22% .
2	B	190	6% 68% 30% .
3	C	16	12% 56% 31% 6% 6% .
4	D	239	5% 77% 19% ..

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5291 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 HLA class II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	180	Total	C	N	O	S	0	0	0
			1479	957	240	277	5			

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DRB1-1 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	190	Total	C	N	O	S	0	0	0
			1557	979	279	293	6			

- Molecule 3 is a protein called GAG polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	15	Total	C	N	O	S	0	0	0
			107	69	15	22	1			

- Molecule 4 is a protein called Enterotoxin type C-3.

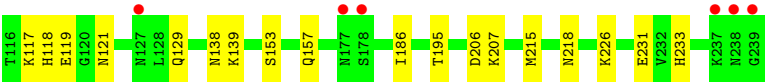
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	232	Total	C	N	O	S	0	0	0
			1899	1205	309	375	10			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	43	SER	LYS	engineered mutation	UNP P0A0L5
D	45	PHE	LEU	engineered mutation	UNP P0A0L5
D	46	LYS	ALA	engineered mutation	UNP P0A0L5
D	47	TRP	HIS	engineered mutation	UNP P0A0L5

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	84	Total 84	O 84	0	0
5	B	66	Total 66	O 66	0	0
5	C	8	Total 8	O 8	0	0
5	D	91	Total 91	O 91	0	0



4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	172.75Å 172.75Å 121.40Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.45 20.00 – 2.45	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.00-2.45) 99.8 (20.00-2.45)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.10 (at 2.44Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.196 , 0.223 0.196 , 0.220	Depositor DCC
R_{free} test set	4984 reflections (10.03%)	wwPDB-VP
Wilson B-factor (Å ²)	38.5	Xtriage
Anisotropy	0.189	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 47.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.033 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5291	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.44% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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.

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	0/1524	0.95	5/2077 (0.2%)
2	B	0.37	0/1597	0.87	2/2168 (0.1%)
3	C	0.44	0/109	1.33	2/147 (1.4%)
4	D	0.39	0/1941	0.87	2/2611 (0.1%)
All	All	0.39	0/5171	0.91	11/7003 (0.2%)

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	88	GLU	N-CA-C	-6.44	98.74	109.24
4	D	7	MET	N-CA-C	-6.29	102.21	110.39
1	A	85	VAL	CA-C-N	6.07	124.09	119.66
1	A	85	VAL	C-N-CA	6.07	124.09	119.66
4	D	24	MET	N-CA-C	-5.77	104.44	112.45

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	1479	0	1412	28	0
2	B	1557	0	1488	62	0
3	C	107	0	105	5	0
4	D	1899	0	1829	38	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	84	0	0	2	0
5	B	66	0	0	0	0
5	C	8	0	0	0	0
5	D	91	0	0	0	0
All	All	5291	0	4834	119	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 12.

The worst 5 of 119 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:150:ASN:HD22	2:B:154:THR:HG22	1.16	1.06
2:B:105:LYS:HE3	2:B:105:LYS:H	1.29	0.97
4:D:98:LYS:HE3	4:D:98:LYS:H	1.35	0.91
4:D:117:LYS:HE2	4:D:119:GLU:HB3	1.50	0.91
2:B:108:PRO:HG2	2:B:109:LEU:HD23	1.51	0.90

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	A	178/180 (99%)	172 (97%)	6 (3%)	0	100	100
2	B	188/190 (99%)	178 (95%)	6 (3%)	4 (2%)	5	4
3	C	13/16 (81%)	13 (100%)	0	0	100	100
4	D	228/239 (95%)	215 (94%)	11 (5%)	2 (1%)	14	17
All	All	607/625 (97%)	578 (95%)	23 (4%)	6 (1%)	12	14

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	110	GLN
2	B	111	HIS
4	D	57	LYS
2	B	108	PRO
4	D	95	PHE

5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

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 oligosaccharides 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, 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(Å ²)	Q < 0.9
1	A	180/180 (100%)	-0.17	7 (3%)	43 43	21, 34, 85, 170	0
2	B	190/190 (100%)	-0.04	12 (6%)	26 23	22, 40, 93, 157	0
3	C	15/16 (93%)	0.85	2 (13%)	7 6	24, 32, 125, 200	0
4	D	232/239 (97%)	-0.20	13 (5%)	30 27	23, 37, 85, 150	0
All	All	617/625 (98%)	-0.12	34 (5%)	30 28	21, 38, 89, 200	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	48	THR	9.6
4	D	56	LYS	5.0
1	A	101	GLU	4.9
4	D	239	GLY	4.4
1	A	182	ALA	4.4

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 oligosaccharides 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.