



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 20, 2026 – 12:11 AM UTC

PDB ID : 4XK2 / pdb_00004xk2
Title : Crystal structure of aldo-keto reductase from Polaromonas sp. JS666
Authors : Gasiorowska, O.A.; Handing, K.B.; Shabalin, I.G.; Sroka, P.; Hillerich, B.S.; Bonanno, J.; Seidel, R.; Almo, S.C.; Minor, W.; New York Structural Genomics Research Consortium (NYSGRG)
Deposited on : 2015-01-09
Resolution : 1.90 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 : 4-5-2 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

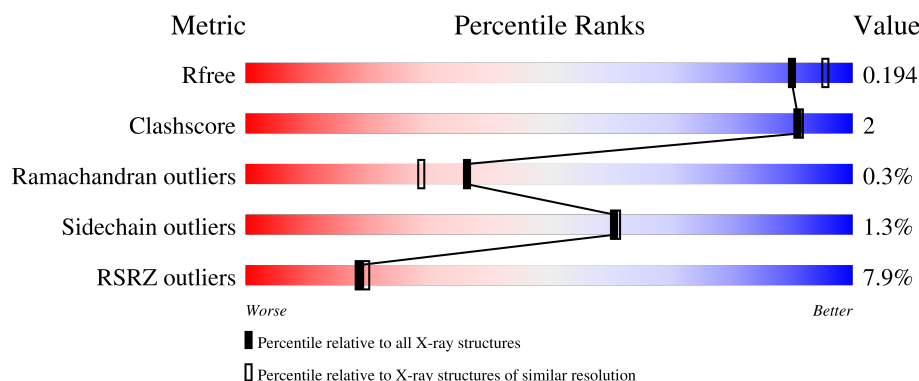
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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}	180053	7789 (1.90-1.90)
Clashscore	190562	8410 (1.90-1.90)
Ramachandran outliers	187476	8333 (1.90-1.90)
Sidechain outliers	187428	8333 (1.90-1.90)
RSRZ outliers	180081	7790 (1.90-1.90)

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 $\leq 5\%$. 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	344	6% 88% • 8%
1	B	344	8% 87% • 10%
1	C	344	7% 88% • 8%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8217 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 Aldo/keto reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	317	Total	C	N	O	S	0	8	0
			2485	1585	432	461	7			
1	B	309	Total	C	N	O	S	0	1	0
			2333	1494	398	435	6			
1	C	316	Total	C	N	O	S	0	3	0
			2427	1552	420	448	7			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP Q124A1
B	0	SER	-	expression tag	UNP Q124A1
C	0	SER	-	expression tag	UNP Q124A1

- Molecule 2 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Cl	0	0
			2	2		
2	B	1	Total	Cl	0	0
			1	1		
2	C	2	Total	Cl	0	0
			2	2		

- Molecule 3 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Na	0	0
			1	1		
3	B	1	Total	Na	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	1	Total 1	Na 1	0	0

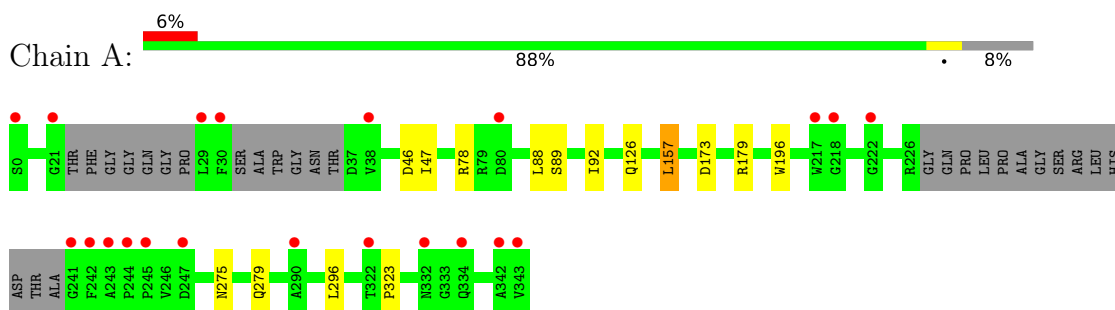
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	423	Total 427	O 427	0	15
4	B	192	Total 192	O 192	0	0
4	C	343	Total 345	O 345	0	6

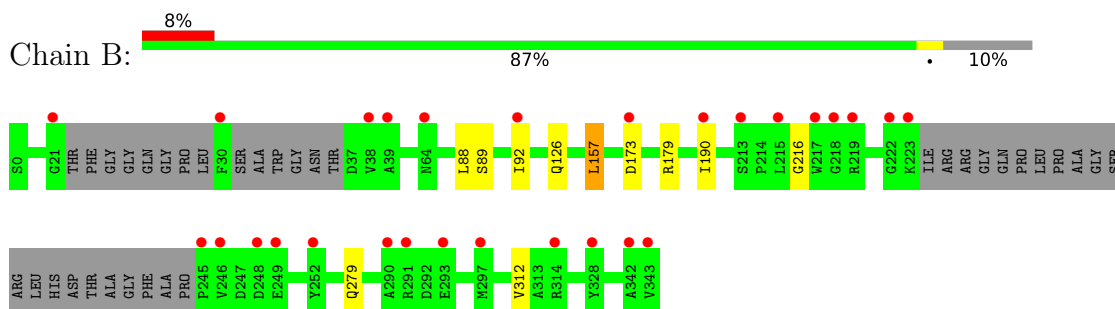
3 Residue-property plots [i](#)

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.

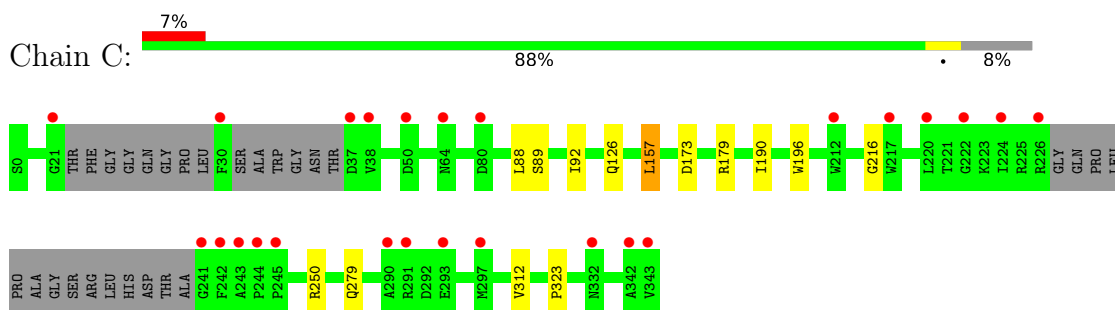
- Molecule 1: Aldo/keto reductase



- Molecule 1: Aldo/keto reductase



- Molecule 1: Aldo/keto reductase



4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α , β , γ	105.11Å 105.11Å 479.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.90 50.00 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (50.00-1.90) 99.9 (50.00-1.90)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.06 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.161 , 0.185 0.171 , 0.194	Depositor DCC
R_{free} test set	5304 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	27.2	Xtriage
Anisotropy	0.279	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 51.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	8217	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.05% 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 [i](#)

5.1 Standard geometry [i](#)

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

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	0.70	0/2554	0.78	0/3470
1	B	0.58	0/2383	0.74	0/3242
1	C	0.64	0/2487	0.78	0/3382
All	All	0.64	0/7424	0.77	0/10094

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 [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	A	2485	0	2455	10	0
1	B	2333	0	2288	6	0
1	C	2427	0	2395	8	0
2	A	2	0	0	0	0
2	B	1	0	0	0	0
2	C	2	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	427	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	192	0	0	0	0
4	C	345	0	0	0	0
All	All	8217	0	7138	24	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 (24) 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:173[B]:ASP:OD2	1:A:179:ARG:NH1	2.33	0.62
1:C:173:ASP:OD2	1:C:179:ARG:NH1	2.34	0.61
1:B:173:ASP:OD2	1:B:179:ARG:NH1	2.36	0.58
1:A:196[A]:TRP:CZ3	1:A:323:PRO:HB3	2.38	0.58
1:A:173[A]:ASP:OD1	4:A:905[A]:HOH:O	2.17	0.58
1:C:196[A]:TRP:CZ3	1:C:323:PRO:HB3	2.38	0.58
1:B:157:LEU:C	1:B:157:LEU:HD22	2.34	0.53
1:C:157:LEU:C	1:C:157:LEU:HD22	2.35	0.52
1:A:47[B]:ILE:HG22	1:A:296:LEU:HD23	1.93	0.51
1:A:47[B]:ILE:HG22	1:A:296:LEU:CD2	2.42	0.50
1:A:173[B]:ASP:CG	1:A:179:ARG:NH1	2.71	0.48
1:C:173:ASP:CG	1:C:179:ARG:NH1	2.72	0.47
1:C:190:ILE:O	1:C:250:ARG:NH1	2.48	0.47
1:A:92:ILE:HG13	1:A:92:ILE:O	2.14	0.47
1:A:157:LEU:C	1:A:157:LEU:HD22	2.39	0.47
1:C:92:ILE:O	1:C:92:ILE:HG13	2.15	0.46
1:B:173:ASP:CG	1:B:179:ARG:NH1	2.74	0.46
1:A:46:ASP:OD1	1:A:78:ARG:NH2	2.38	0.46
1:B:92:ILE:O	1:B:92:ILE:HG13	2.16	0.45
1:B:279:GLN:OE1	1:B:312:VAL:HG21	2.19	0.43
1:C:190:ILE:HD11	1:C:216:GLY:HA3	2.02	0.41
1:C:279:GLN:OE1	1:C:312:VAL:HG21	2.20	0.41
1:A:275:ASN:O	1:A:279:GLN:HG2	2.20	0.41
1:B:190:ILE:HD11	1:B:216:GLY:HA3	2.02	0.40

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	317/344 (92%)	309 (98%)	7 (2%)	1 (0%)	36	29
1	B	303/344 (88%)	296 (98%)	6 (2%)	1 (0%)	36	29
1	C	312/344 (91%)	305 (98%)	6 (2%)	1 (0%)	36	29
All	All	932/1032 (90%)	910 (98%)	19 (2%)	3 (0%)	36	29

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	89	SER
1	B	89	SER
1	C	89	SER

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	252/270 (93%)	249 (99%)	3 (1%)	63	63
1	B	231/270 (86%)	228 (99%)	3 (1%)	61	61
1	C	244/270 (90%)	241 (99%)	3 (1%)	63	63
All	All	727/810 (90%)	718 (99%)	9 (1%)	61	63

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	88	LEU
1	A	126	GLN
1	A	157	LEU
1	B	88	LEU
1	B	126	GLN
1	B	157	LEU
1	C	88	LEU
1	C	126	GLN
1	C	157	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	97	ASN
1	A	104	HIS
1	A	159	ASN
1	B	105	HIS
1	B	159	ASN
1	C	97	ASN
1	C	104	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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(Å ²)	Q<0.9
1	A	317/344 (92%)	0.15	21 (6%)	24 26	14, 31, 59, 92	8 (2%)
1	B	309/344 (89%)	0.85	28 (9%)	15 15	32, 54, 84, 101	1 (0%)
1	C	316/344 (91%)	0.39	25 (7%)	18 20	17, 38, 72, 92	3 (0%)
All	All	942/1032 (91%)	0.46	74 (7%)	18 20	14, 43, 74, 101	12 (1%)

All (74) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	29	LEU	6.5
1	C	244	PRO	5.6
1	A	244	PRO	5.3
1	C	217	TRP	5.2
1	B	30	PHE	4.8
1	C	30	PHE	4.4
1	C	242	PHE	4.4
1	A	242	PHE	4.2
1	C	342	ALA	3.9
1	C	241	GLY	3.9
1	B	328	TYR	3.7
1	B	223	LYS	3.7
1	B	38	VAL	3.6
1	B	245	PRO	3.6
1	B	217	TRP	3.5
1	A	241	GLY	3.5
1	B	291	ARG	3.5
1	C	245	PRO	3.5
1	A	30	PHE	3.5
1	A	342	ALA	3.4
1	C	243	ALA	3.3
1	A	217	TRP	3.2
1	A	38	VAL	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	243	ALA	3.1
1	B	252	TYR	3.1
1	C	37	ASP	3.0
1	A	245	PRO	3.0
1	B	249	GLU	3.0
1	B	343	VAL	3.0
1	C	220	LEU	2.9
1	A	218	GLY	2.8
1	C	21	GLY	2.8
1	C	291	ARG	2.8
1	B	219	ARG	2.7
1	B	290	ALA	2.7
1	C	222	GLY	2.7
1	A	343	VAL	2.7
1	B	218	GLY	2.6
1	B	314	ARG	2.6
1	B	92	ILE	2.6
1	C	343	VAL	2.6
1	B	246	VAL	2.5
1	B	222	GLY	2.5
1	C	212	TRP	2.5
1	B	342	ALA	2.5
1	C	226	ARG	2.4
1	A	21	GLY	2.4
1	A	334	GLN	2.3
1	C	224	ILE	2.3
1	C	38	VAL	2.3
1	A	247	ASP	2.3
1	B	215	LEU	2.3
1	B	190	ILE	2.3
1	C	297	MET	2.3
1	C	293	GLU	2.3
1	A	332	ASN	2.3
1	C	64	ASN	2.3
1	B	293	GLU	2.2
1	A	0	SER	2.2
1	B	39	ALA	2.2
1	B	297	MET	2.2
1	C	332	ASN	2.2
1	C	80	ASP	2.2
1	B	213	SER	2.2
1	A	222	GLY	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	80	ASP	2.1
1	B	173	ASP	2.1
1	B	21	GLY	2.1
1	A	322	THR	2.1
1	C	290	ALA	2.1
1	A	290	ALA	2.0
1	B	248	ASP	2.0
1	C	50	ASP	2.0
1	B	64	ASN	2.0

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	NA	B	402	1/1	0.89	0.12	45,45,45,45	1
2	CL	B	401	1/1	0.91	0.14	66,66,66,66	0
2	CL	C	401	1/1	0.95	0.10	54,54,54,54	0
3	NA	C	403	1/1	0.95	0.14	53,53,53,53	0
2	CL	C	402	1/1	0.97	0.10	38,38,38,38	0
3	NA	A	403	1/1	0.97	0.15	43,43,43,43	0
2	CL	A	401	1/1	0.98	0.07	41,41,41,41	0
2	CL	A	402	1/1	1.00	0.05	33,33,33,33	0

6.5 Other polymers [i](#)

There are no such residues in this entry.