



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 05:58 PM UTC

PDB ID : 8XR2 / pdb_00008xr2
Title : Crystal structure of AKRtyl-apo1
Authors : Lin, S.; Dai, S.; Xiao, Z.
Deposited on : 2024-01-06
Resolution : 2.39 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

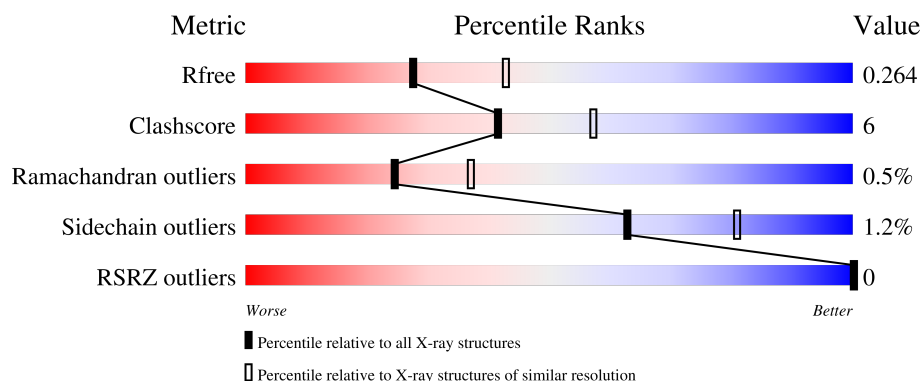
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.39 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	4912 (2.40-2.40)
Clashscore	190562	5391 (2.40-2.40)
Ramachandran outliers	187476	5320 (2.40-2.40)
Sidechain outliers	187428	5321 (2.40-2.40)
RSRZ outliers	180081	4916 (2.40-2.40)

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	331	 78% 10% • 11%
1	B	331	 73% 22% • •
1	C	331	 71% 18% 11%
1	D	331	 81% 12% • 5%
1	E	331	 75% 14% 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	331	 74%21%5%
1	G	331	 73%17%10%
1	H	331	 83%13%. .
1	I	331	 73%17%. 9%
1	J	331	 81%14%. .
1	K	331	 73%15%. 11%
1	L	331	 83%11%. 5%
1	M	331	 72%17%11%
1	N	331	 82%13%5%
1	O	331	 74%18%8%
1	P	331	 81%14%. .

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 39692 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	294	Total	C	N	O	S	0	0	0
			2290	1441	411	432	6			
1	B	318	Total	C	N	O	S	0	0	0
			2468	1554	440	466	8			
1	C	295	Total	C	N	O	S	0	0	0
			2294	1445	411	432	6			
1	D	313	Total	C	N	O	S	0	0	0
			2429	1531	434	456	8			
1	E	295	Total	C	N	O	S	0	0	0
			2297	1447	413	430	7			
1	F	315	Total	C	N	O	S	0	0	0
			2446	1542	436	460	8			
1	G	297	Total	C	N	O	S	0	0	0
			2311	1452	415	437	7			
1	H	322	Total	C	N	O	S	0	0	0
			2493	1568	447	470	8			
1	I	300	Total	C	N	O	S	0	0	0
			2334	1470	418	440	6			
1	J	317	Total	C	N	O	S	0	0	0
			2460	1550	439	463	8			
1	K	294	Total	C	N	O	S	0	0	0
			2293	1445	410	431	7			
1	L	314	Total	C	N	O	S	0	0	0
			2437	1537	435	457	8			
1	M	295	Total	C	N	O	S	0	0	0
			2297	1447	413	430	7			
1	N	315	Total	C	N	O	S	0	0	0
			2446	1542	436	460	8			
1	O	303	Total	C	N	O	S	0	0	0
			2355	1479	422	447	7			
1	P	319	Total	C	N	O	S	0	0	0
			2469	1555	440	466	8			

- Molecule 2 is water.

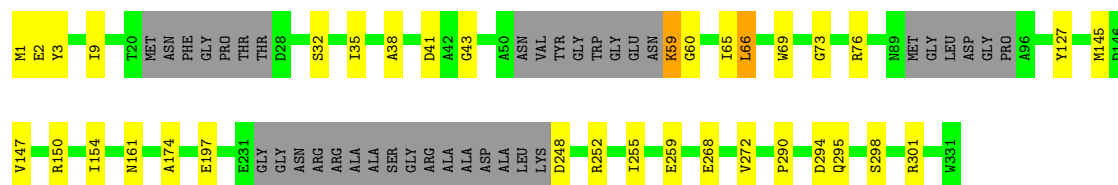
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	79	Total 79	O 79	0	0
2	B	105	Total 105	O 105	0	0
2	C	86	Total 86	O 86	0	0
2	D	111	Total 111	O 111	0	0
2	E	84	Total 84	O 84	0	0
2	F	85	Total 85	O 85	0	0
2	G	71	Total 71	O 71	0	0
2	H	111	Total 111	O 111	0	0
2	I	81	Total 81	O 81	0	0
2	J	110	Total 110	O 110	0	0
2	K	68	Total 68	O 68	0	0
2	L	122	Total 122	O 122	0	0
2	M	108	Total 108	O 108	0	0
2	N	129	Total 129	O 129	0	0
2	O	87	Total 87	O 87	0	0
2	P	136	Total 136	O 136	0	0

3 Residue-property plots

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.

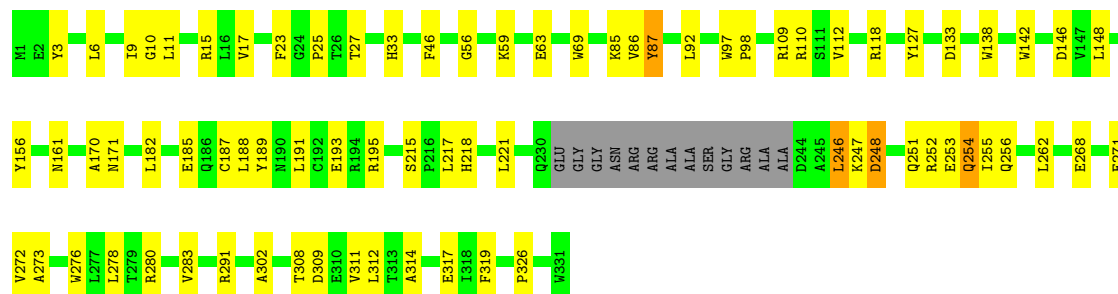
- Molecule 1: Aldo/keto reductase

Chain A: 



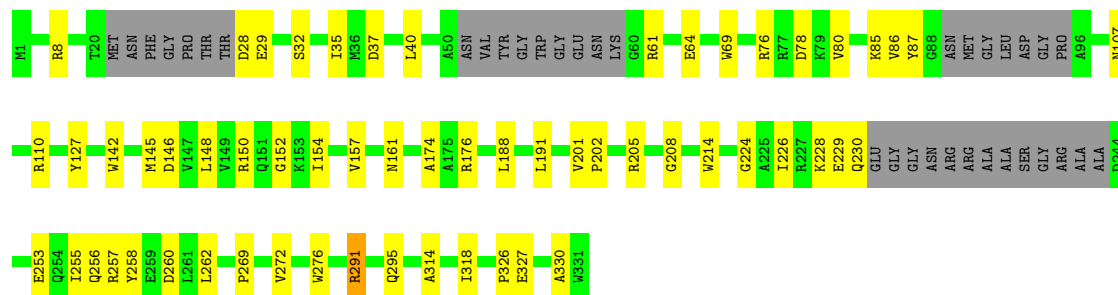
- Molecule 1: Aldo/keto reductase

Chain B: 



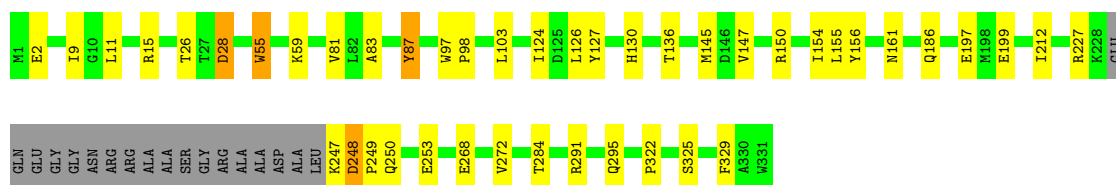
- Molecule 1: Aldo/keto reductase

Chain C: 



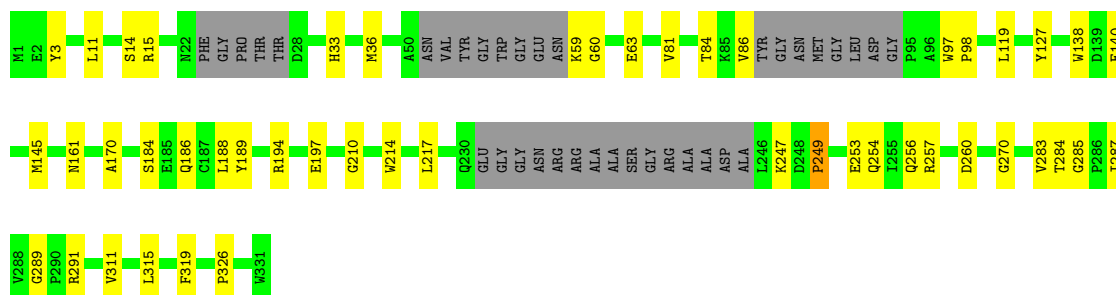
- Molecule 1: Aldo/keto reductase

Chain D:  81% 12% 5%



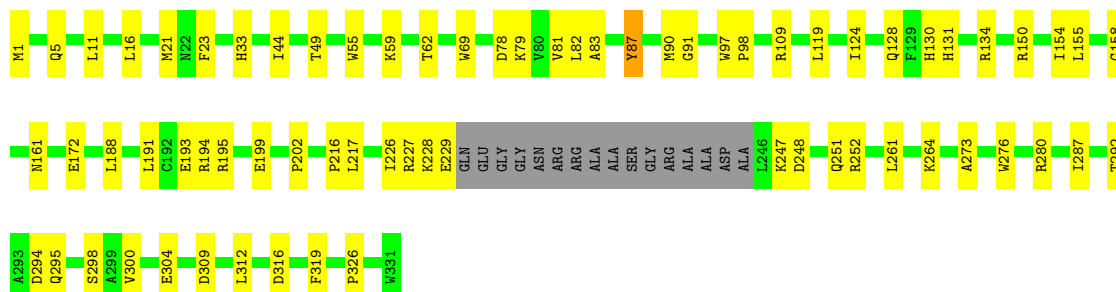
- Molecule 1: Aldo/keto reductase

Chain E:  75% 14% 11%



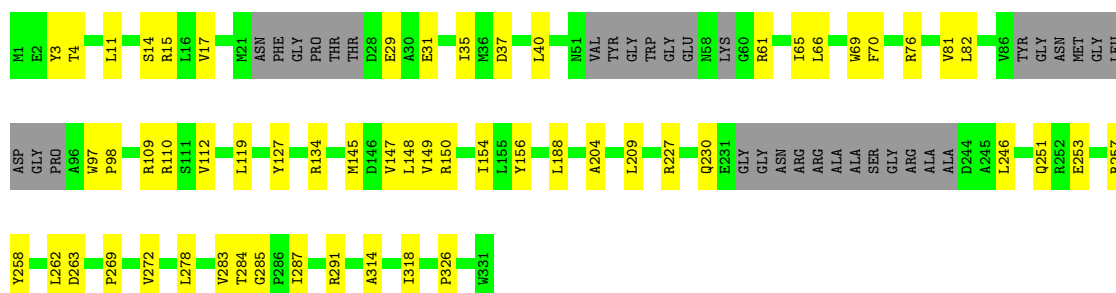
- Molecule 1: Aldo/keto reductase

Chain F:  74% 21% 5%



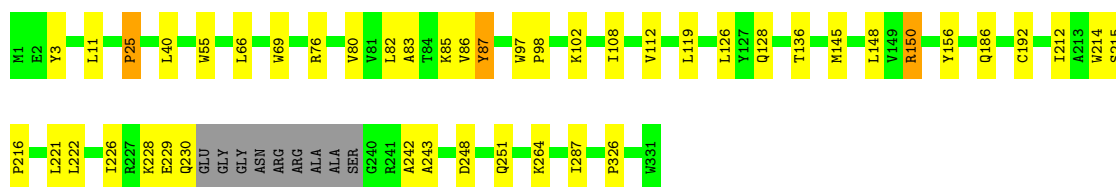
- Molecule 1: Aldo/keto reductase

Chain G:  73% 17% 10%



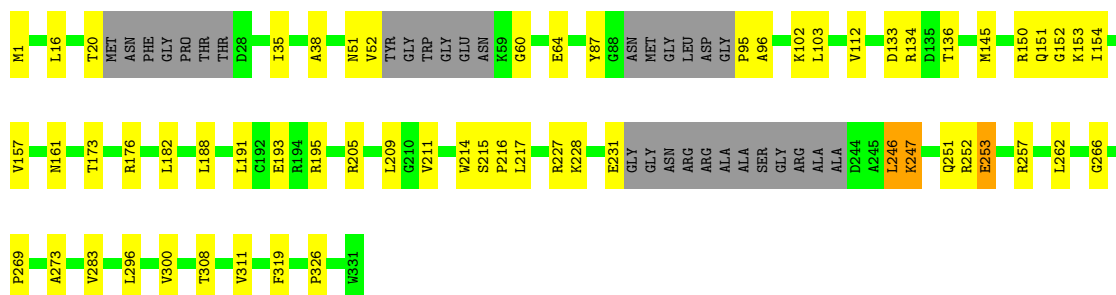
- Molecule 1: Aldo/keto reductase

Chain H:  83% 13% . .



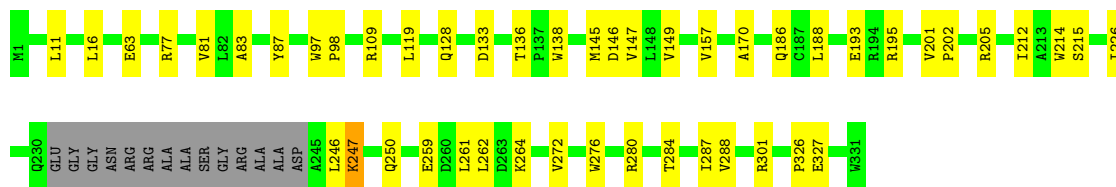
- Molecule 1: Aldo/keto reductase

Chain I:  73% 17% . 9%



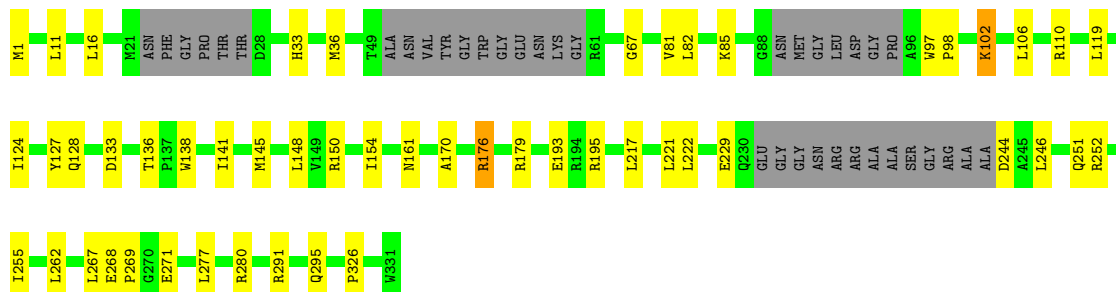
- Molecule 1: Aldo/keto reductase

Chain J:  81% 14% .



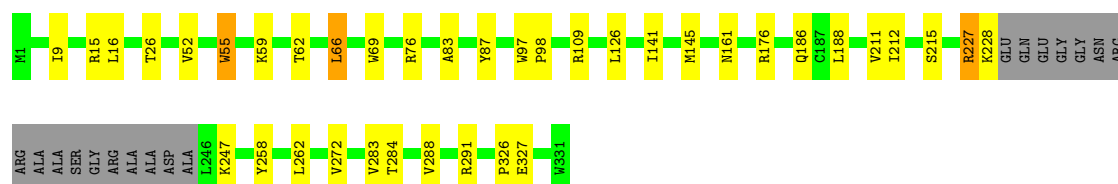
- Molecule 1: Aldo/keto reductase

Chain K:  73% 15% . 11%

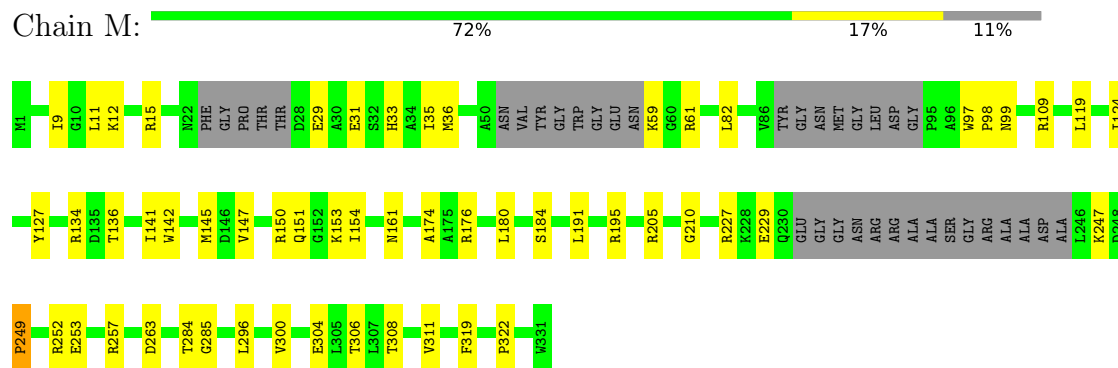


- Molecule 1: Aldo/keto reductase

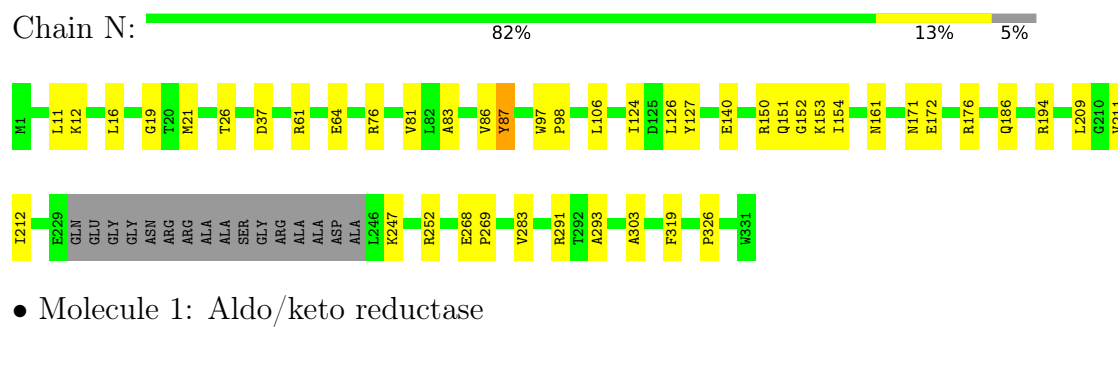
Chain L:  83% 11% . 5%



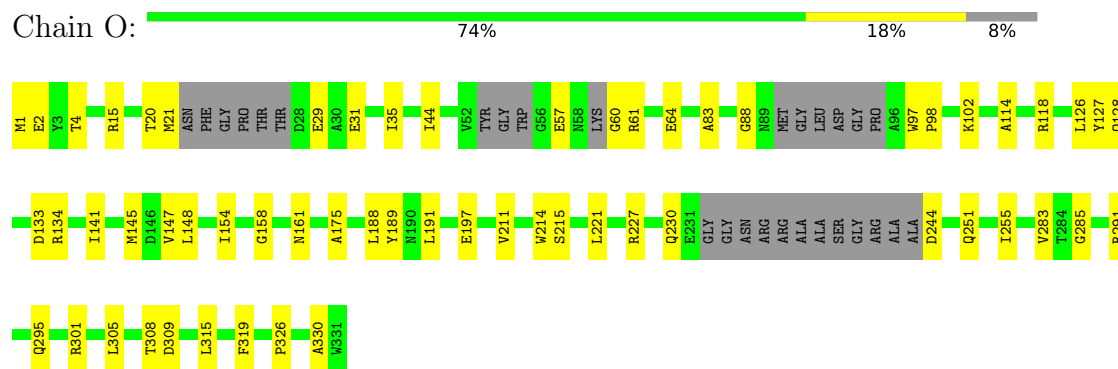
- Molecule 1: Aldo/keto reductase



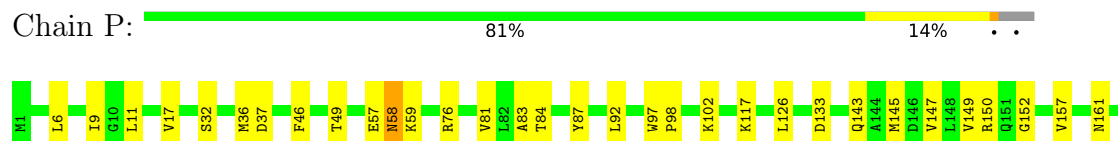
- Molecule 1: Aldo/keto reductase



- Molecule 1: Aldo/keto reductase



- Molecule 1: Aldo/keto reductase



Q186	E193	R194	R195	I212	E229	GLN	GLU	GLY	GLY	ASN	ARG	ARG	ALA	ALA	SER	GLY	ARG	A242	L246	K247	D248	P249	Q250	Q251	R252	L262	V272	I287	A314	I318	P326	E327	A330	W331
------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α , β , γ	110.78Å 110.78Å 561.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.67 – 2.39 49.67 – 2.39	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.67-2.39) 100.0 (49.67-2.39)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.85 (at 2.39Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.222 , 0.263 0.223 , 0.264	Depositor DCC
R_{free} test set	13209 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	27.7	Xtriage
Anisotropy	0.031	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 24.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	0.067 for h,-k,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	39692	wwPDB-VP
Average B, all atoms (Å ²)	31.0	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 22.77 % 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 5.3278e-03.*

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](#)

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	0.62	0/2336	0.70	2/3167 (0.1%)
1	B	0.52	0/2523	0.70	5/3427 (0.1%)
1	C	0.34	0/2340	0.56	0/3173
1	D	0.49	0/2484	0.62	2/3374 (0.1%)
1	E	0.42	0/2343	0.59	0/3175
1	F	0.43	0/2501	0.59	0/3397
1	G	0.42	0/2355	0.61	0/3191
1	H	0.50	0/2548	0.69	2/3460 (0.1%)
1	I	0.50	0/2381	0.68	3/3228 (0.1%)
1	J	0.42	0/2515	0.60	1/3416 (0.0%)
1	K	0.37	0/2339	0.58	0/3171
1	L	0.46	0/2492	0.63	0/3385
1	M	0.40	0/2343	0.59	0/3175
1	N	0.45	0/2501	0.60	0/3397
1	O	0.47	0/2400	0.65	1/3252 (0.0%)
1	P	0.47	0/2524	0.65	1/3429 (0.0%)
All	All	0.46	0/38925	0.63	17/52817 (0.0%)

There are no bond length outliers.

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	247	LYS	N-CA-C	-12.11	98.92	112.72
1	I	246	LEU	N-CA-C	-11.88	98.20	113.17
1	H	229	GLU	N-CA-C	-9.44	99.80	112.26
1	A	60	GLY	N-CA-C	7.36	124.96	115.32
1	I	87	TYR	N-CA-C	6.97	120.67	109.79

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	2290	0	2241	20	0
1	B	2468	0	2408	45	0
1	C	2294	0	2249	37	0
1	D	2429	0	2374	27	0
1	E	2297	0	2264	31	0
1	F	2446	0	2391	47	0
1	G	2311	0	2263	41	0
1	H	2493	0	2434	28	0
1	I	2334	0	2291	38	0
1	J	2460	0	2404	31	0
1	K	2293	0	2250	32	0
1	L	2437	0	2385	22	0
1	M	2297	0	2264	36	0
1	N	2446	0	2391	32	0
1	O	2355	0	2299	36	0
1	P	2469	0	2410	31	0
2	A	79	0	0	2	0
2	B	105	0	0	0	0
2	C	86	0	0	0	0
2	D	111	0	0	0	0
2	E	84	0	0	0	0
2	F	85	0	0	1	0
2	G	71	0	0	1	0
2	H	111	0	0	0	0
2	I	81	0	0	1	0
2	J	110	0	0	1	0
2	K	68	0	0	2	0
2	L	122	0	0	1	0
2	M	108	0	0	2	0
2	N	129	0	0	3	0
2	O	87	0	0	0	0
2	P	136	0	0	0	0
All	All	39692	0	37318	485	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 6.

The worst 5 of 485 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:G:127:TYR:HD2	1:G:145:MET:HE1	1.51	0.76
1:B:218:HIS:HD2	1:B:221:LEU:HD12	1.56	0.70
1:D:150:ARG:HH11	1:F:150:ARG:HD2	1.57	0.70
1:L:188:LEU:HD13	1:L:215:SER:HB2	1.77	0.67
1:H:108:ILE:HD13	1:H:145:MET:HE3	1.78	0.66

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	284/331 (86%)	271 (95%)	13 (5%)	0	100	100
1	B	314/331 (95%)	297 (95%)	16 (5%)	1 (0%)	36	50
1	C	285/331 (86%)	274 (96%)	11 (4%)	0	100	100
1	D	309/331 (93%)	292 (94%)	14 (4%)	3 (1%)	12	20
1	E	285/331 (86%)	269 (94%)	14 (5%)	2 (1%)	18	28
1	F	311/331 (94%)	297 (96%)	11 (4%)	3 (1%)	12	20
1	G	286/331 (86%)	277 (97%)	9 (3%)	0	100	100
1	H	318/331 (96%)	302 (95%)	13 (4%)	3 (1%)	14	22
1	I	290/331 (88%)	282 (97%)	8 (3%)	0	100	100
1	J	313/331 (95%)	298 (95%)	14 (4%)	1 (0%)	36	50
1	K	284/331 (86%)	273 (96%)	10 (4%)	1 (0%)	30	43
1	L	310/331 (94%)	294 (95%)	13 (4%)	3 (1%)	12	20
1	M	285/331 (86%)	276 (97%)	8 (3%)	1 (0%)	30	43
1	N	311/331 (94%)	297 (96%)	13 (4%)	1 (0%)	36	50
1	O	291/331 (88%)	278 (96%)	12 (4%)	1 (0%)	36	50

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	P	315/331 (95%)	304 (96%)	9 (3%)	2 (1%)	21	32
All	All	4791/5296 (90%)	4581 (96%)	188 (4%)	22 (0%)	24	37

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	87	TYR
1	D	227	ARG
1	E	249	PRO
1	J	87	TYR
1	M	249	PRO

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	234/258 (91%)	229 (98%)	5 (2%)	47	69
1	B	252/258 (98%)	245 (97%)	7 (3%)	38	60
1	C	234/258 (91%)	232 (99%)	2 (1%)	70	85
1	D	248/258 (96%)	243 (98%)	5 (2%)	48	70
1	E	236/258 (92%)	236 (100%)	0	100	100
1	F	250/258 (97%)	250 (100%)	0	100	100
1	G	237/258 (92%)	235 (99%)	2 (1%)	73	86
1	H	253/258 (98%)	248 (98%)	5 (2%)	48	70
1	I	239/258 (93%)	234 (98%)	5 (2%)	47	69
1	J	251/258 (97%)	249 (99%)	2 (1%)	73	86
1	K	235/258 (91%)	234 (100%)	1 (0%)	84	92
1	L	249/258 (96%)	244 (98%)	5 (2%)	48	70
1	M	236/258 (92%)	234 (99%)	2 (1%)	73	86
1	N	250/258 (97%)	250 (100%)	0	100	100
1	O	241/258 (93%)	237 (98%)	4 (2%)	53	74

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	P	251/258 (97%)	248 (99%)	3 (1%)	63	81
All	All	3896/4128 (94%)	3848 (99%)	48 (1%)	63	81

5 of 48 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	251	GLN
1	L	66	LEU
1	I	252	ARG
1	J	247	LYS
1	L	247	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 52 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	169	GLN
1	K	186	GLN
1	O	169	GLN
1	I	186	GLN
1	J	186	GLN

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 ⓘ

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	294/331 (88%)	-1.33	0 100 100	18, 32, 59, 71	0
1	B	318/331 (96%)	-1.37	0 100 100	17, 30, 61, 81	0
1	C	295/331 (89%)	-1.34	0 100 100	19, 32, 58, 73	0
1	D	313/331 (94%)	-1.41	0 100 100	15, 27, 53, 89	0
1	E	295/331 (89%)	-1.34	0 100 100	18, 32, 57, 82	0
1	F	315/331 (95%)	-1.36	0 100 100	17, 32, 53, 73	0
1	G	297/331 (89%)	-1.39	0 100 100	18, 30, 56, 72	0
1	H	322/331 (97%)	-1.38	0 100 100	17, 27, 60, 88	0
1	I	300/331 (90%)	-1.40	0 100 100	14, 27, 57, 77	0
1	J	317/331 (95%)	-1.45	0 100 100	15, 24, 48, 75	0
1	K	294/331 (88%)	-1.39	0 100 100	15, 30, 57, 74	0
1	L	314/331 (94%)	-1.49	0 100 100	16, 23, 41, 81	0
1	M	295/331 (89%)	-1.40	0 100 100	17, 29, 53, 80	0
1	N	315/331 (95%)	-1.48	0 100 100	15, 25, 43, 64	0
1	O	303/331 (91%)	-1.40	0 100 100	17, 29, 60, 75	0
1	P	319/331 (96%)	-1.48	0 100 100	13, 22, 46, 89	0
All	All	4906/5296 (92%)	-1.40	0 100 100	13, 28, 56, 89	0

There are no RSRZ outliers to report.

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 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.