WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 22, 2024 – 10:24 pm BST

PDB ID	:	8QX3
Title	:	Arabidopsis thaliana Phosphoenolpyruvate carboxylase PPC1 R886G mutant
Authors	:	Loris, R.; Haesaerts, S.; Larsen, P.B.
Deposited on	:	2023-10-21
Resolution	:	2.95 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

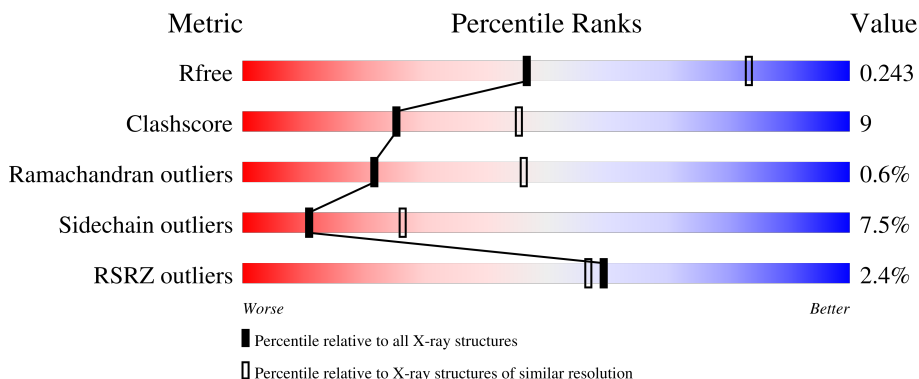|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.95 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	Percentile Ranks	Value
Rfree		0.243
Clashscore		9
Ramachandran outliers		0.6%
Sidechain outliers		7.5%
RSRZ outliers		2.4%

Worse *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1067 (2.96-2.92)
Clashscore	180529	1122 (2.96-2.92)
Ramachandran outliers	177936	1075 (2.96-2.92)
Sidechain outliers	177891	1075 (2.96-2.92)
RSRZ outliers	164620	1067 (2.96-2.92)

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	974	2% 67% 22% • 8%
1	B	974	.% 67% 23% • 8%
1	C	974	3% 67% 22% • 9%
1	D	974	2% 66% 22% • 10%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 28637 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 Phosphoenolpyruvate carboxylase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	897	Total	C	N	O	S	0	0	0
			7221	4581	1263	1348	29			
1	B	899	Total	C	N	O	S	0	0	0
			7233	4588	1263	1353	29			
1	C	883	Total	C	N	O	S	0	0	0
			7103	4507	1243	1324	29			
1	D	877	Total	C	N	O	S	0	0	0
			7052	4475	1228	1320	29			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	initiating methionine	UNP Q9MAH0
A	-5	HIS	-	expression tag	UNP Q9MAH0
A	-4	HIS	-	expression tag	UNP Q9MAH0
A	-3	HIS	-	expression tag	UNP Q9MAH0
A	-2	HIS	-	expression tag	UNP Q9MAH0
A	-1	HIS	-	expression tag	UNP Q9MAH0
A	0	HIS	-	expression tag	UNP Q9MAH0
A	886	GLY	ARG	engineered mutation	UNP Q9MAH0
B	-6	MET	-	initiating methionine	UNP Q9MAH0
B	-5	HIS	-	expression tag	UNP Q9MAH0
B	-4	HIS	-	expression tag	UNP Q9MAH0
B	-3	HIS	-	expression tag	UNP Q9MAH0
B	-2	HIS	-	expression tag	UNP Q9MAH0
B	-1	HIS	-	expression tag	UNP Q9MAH0
B	0	HIS	-	expression tag	UNP Q9MAH0
B	886	GLY	ARG	engineered mutation	UNP Q9MAH0
C	-6	MET	-	initiating methionine	UNP Q9MAH0
C	-5	HIS	-	expression tag	UNP Q9MAH0
C	-4	HIS	-	expression tag	UNP Q9MAH0
C	-3	HIS	-	expression tag	UNP Q9MAH0
C	-2	HIS	-	expression tag	UNP Q9MAH0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	HIS	-	expression tag	UNP Q9MAH0
C	0	HIS	-	expression tag	UNP Q9MAH0
C	886	GLY	ARG	engineered mutation	UNP Q9MAH0
D	-6	MET	-	initiating methionine	UNP Q9MAH0
D	-5	HIS	-	expression tag	UNP Q9MAH0
D	-4	HIS	-	expression tag	UNP Q9MAH0
D	-3	HIS	-	expression tag	UNP Q9MAH0
D	-2	HIS	-	expression tag	UNP Q9MAH0
D	-1	HIS	-	expression tag	UNP Q9MAH0
D	0	HIS	-	expression tag	UNP Q9MAH0
D	886	GLY	ARG	engineered mutation	UNP Q9MAH0

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).
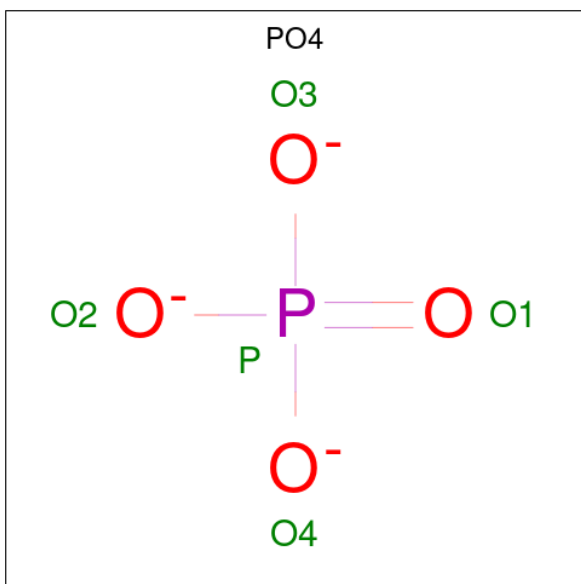
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 5	O 4	P 1	0	0
2	B	1	Total 5	O 4	P 1	0	0
2	C	1	Total 5	O 4	P 1	0	0
2	D	1	Total 5	O 4	P 1	0	0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	5	Total 5	O 5	0	0
3	B	3	Total 3	O 3	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.

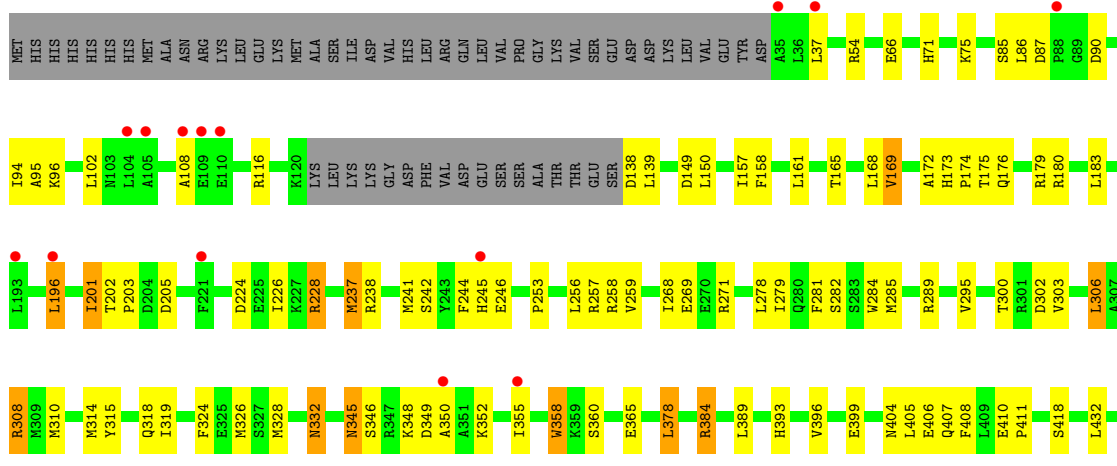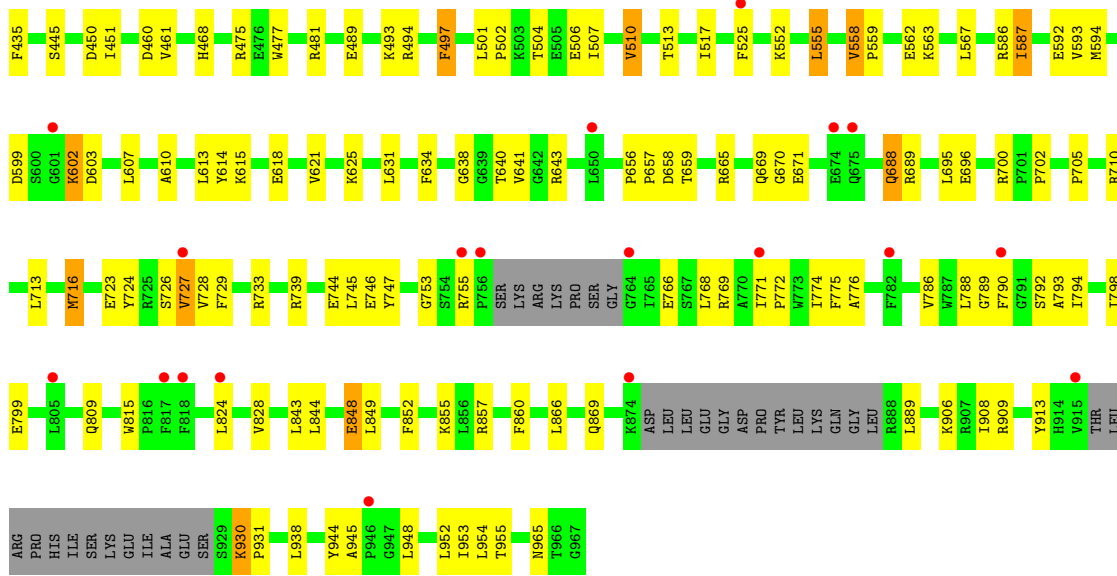• Molecule 1: Phosphoenolpyruvate carboxylase 1
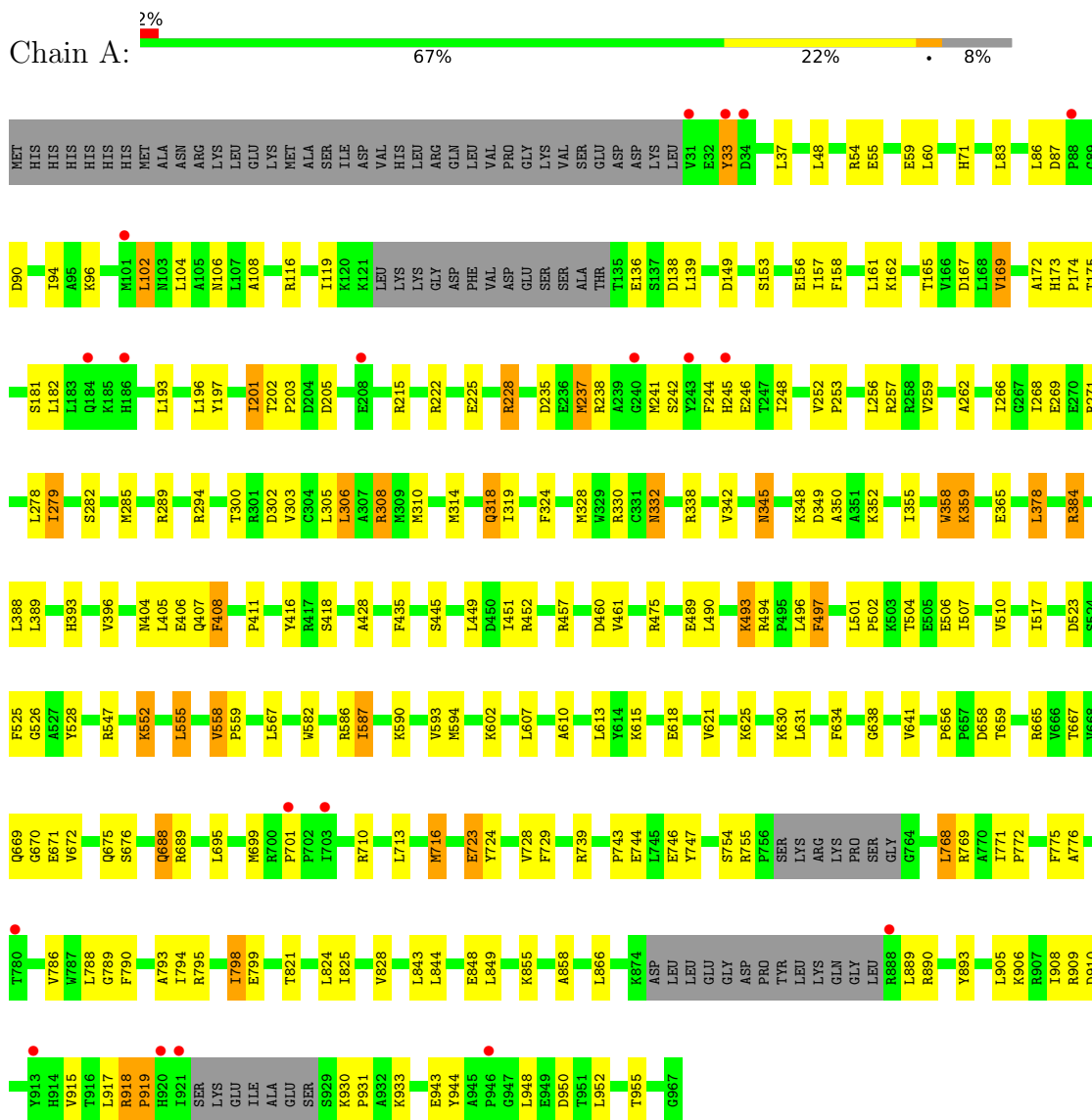
Chain A:

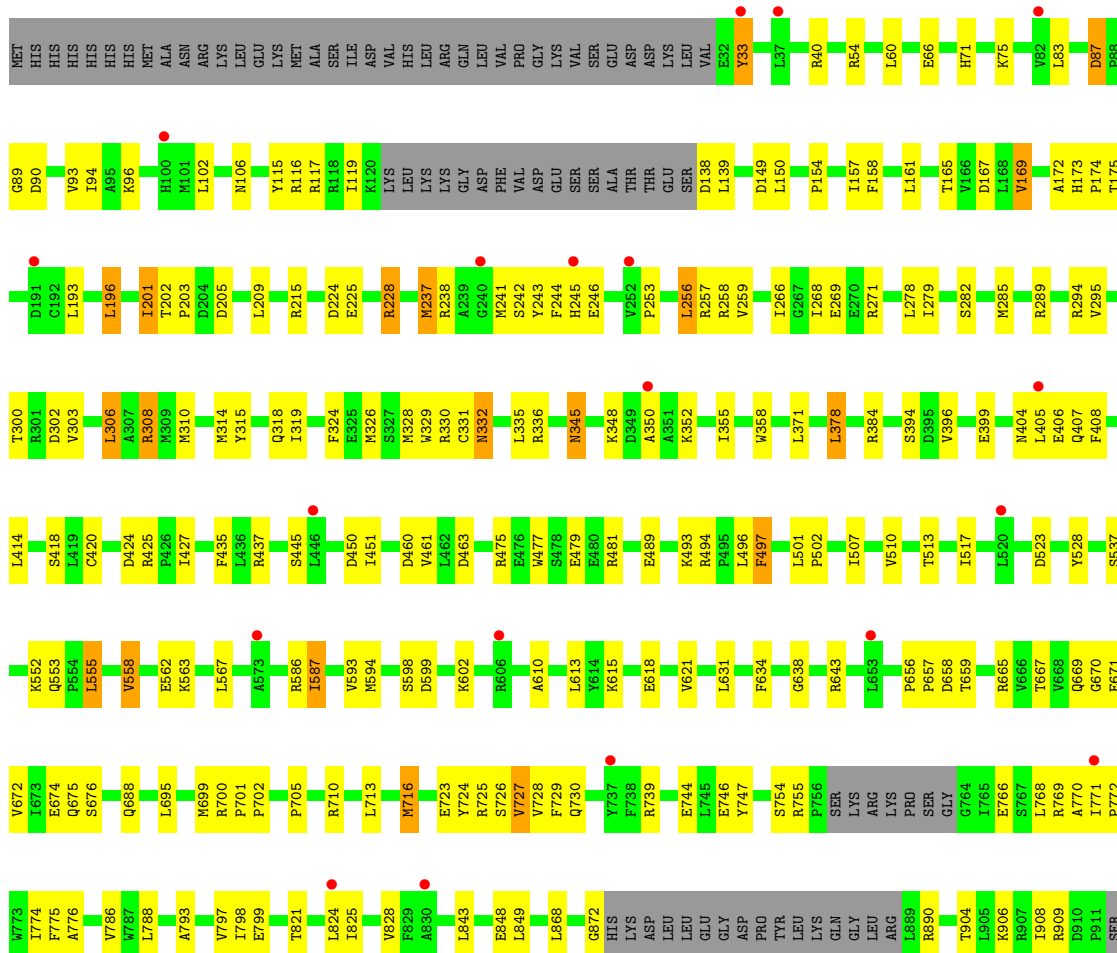• Molecule 1: Phosphoenolpyruvate carboxylase 1

Chain B:

• Molecule 1: Phosphoenolpyruvate carboxylase 1

Chain C:

3%

67% 22% • 9%

● Molecule 1: Phosphoenolpyruvate carboxylase 1

Chain D:
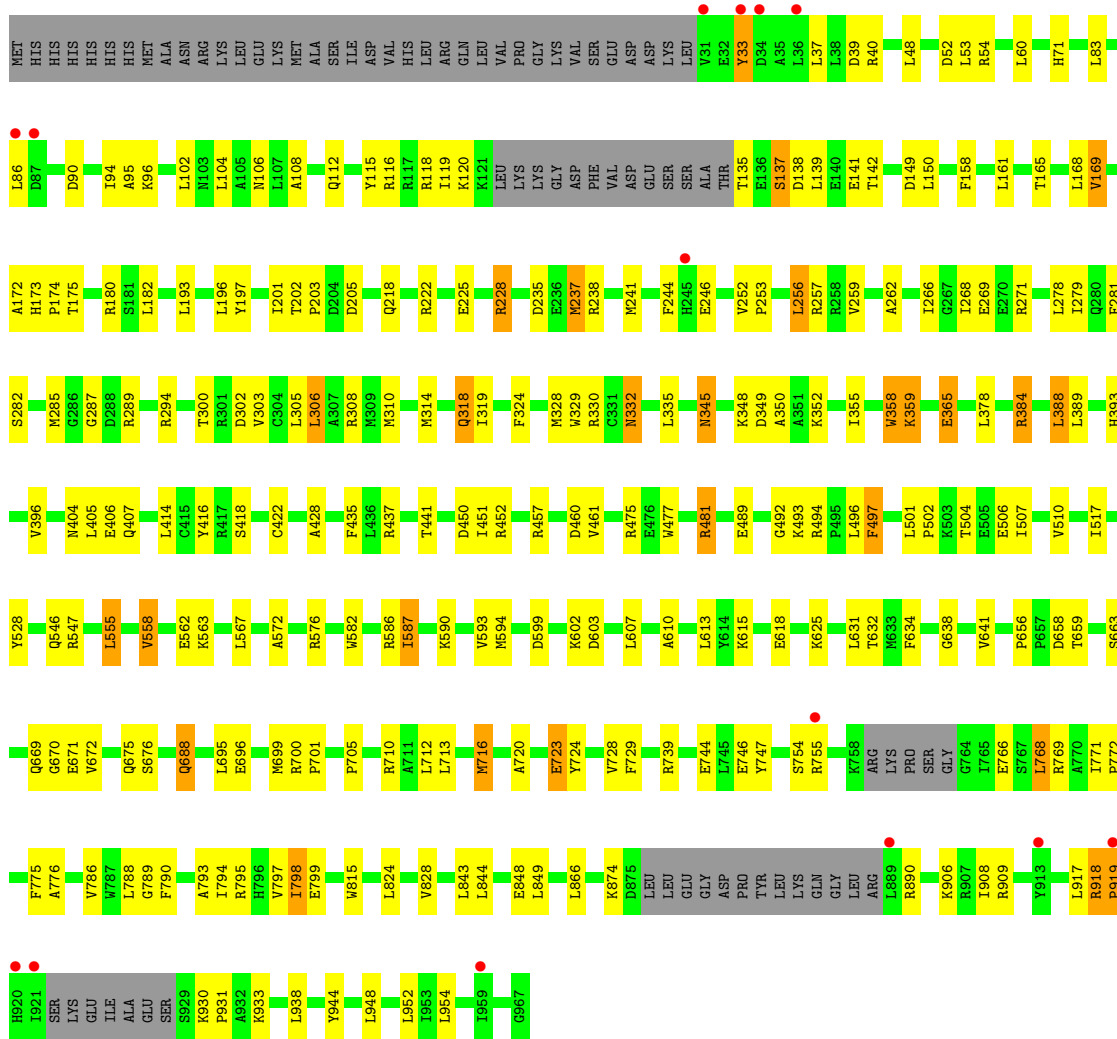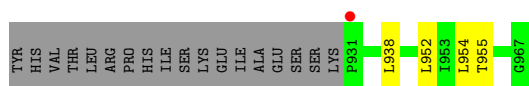
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	127.67Å 159.16Å 140.99Å 90.00° 92.03° 90.00°	Depositor
Resolution (Å)	48.99 – 2.95 48.99 – 2.95	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.99-2.95) 99.7 (48.99-2.95)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.08 (at 2.96Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.204 , 0.240 0.206 , 0.243	Depositor DCC
R_{free} test set	6105 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	91.6	Xtriage
Anisotropy	0.253	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 65.0	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.072 for h,-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	28637	wwPDB-VP
Average B, all atoms (Å^2)	114.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.63% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	0.29	0/7377	0.47	1/9985 (0.0%)
1	B	0.29	0/7389	0.48	1/10001 (0.0%)
1	C	0.27	0/7256	0.45	2/9820 (0.0%)
1	D	0.26	0/7203	0.42	0/9750
All	All	0.28	0/29225	0.45	4/39556 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	689	ARG	NE-CZ-NH2	-9.91	115.34	120.30
1	C	689	ARG	NE-CZ-NH1	7.38	123.99	120.30
1	B	768	LEU	CA-CB-CG	5.34	127.57	115.30
1	A	768	LEU	CA-CB-CG	5.29	127.46	115.30

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	7221	0	7178	143	0
1	B	7233	0	7187	145	0
1	C	7103	0	7063	138	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	7052	0	7002	134	0
2	A	5	0	0	1	0
2	B	5	0	0	1	0
2	C	5	0	0	1	0
2	D	5	0	0	0	0
3	A	5	0	0	0	0
3	B	3	0	0	0	0
All	All	28637	0	28430	539	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 9.

The worst 5 of 539 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:404:ASN:HB2	1:A:407:GLN:HB2	1.63	0.81
1:D:460:ASP:OD1	1:D:475:ARG:NH1	2.14	0.79
1:A:173:HIS:ND1	1:A:669:GLN:OE1	2.15	0.79
1:A:638:GLY:HA2	1:A:669:GLN:HG3	1.64	0.77
1:B:638:GLY:HA2	1:B:669:GLN:HG3	1.67	0.76

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	887/974 (91%)	843 (95%)	37 (4%)	7 (1%)	16	38
1	B	889/974 (91%)	842 (95%)	40 (4%)	7 (1%)	16	38
1	C	873/974 (90%)	833 (95%)	36 (4%)	4 (0%)	25	50
1	D	867/974 (89%)	826 (95%)	38 (4%)	3 (0%)	37	60

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	3516/3896 (90%)	3344 (95%)	151 (4%)	21 (1%)	22	46

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	670	GLY
1	A	919	PRO
1	B	670	GLY
1	B	919	PRO
1	C	670	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	780/848 (92%)	719 (92%)	61 (8%)	10	25
1	B	782/848 (92%)	721 (92%)	61 (8%)	10	25
1	C	766/848 (90%)	710 (93%)	56 (7%)	11	28
1	D	760/848 (90%)	707 (93%)	53 (7%)	12	29
All	All	3088/3392 (91%)	2857 (92%)	231 (8%)	11	26

5 of 231 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	768	LEU
1	D	688	GLN
1	C	328	MET
1	D	602	LYS
1	D	308	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	332	ASN
1	C	332	ASN
1	D	332	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 oligosaccharides 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	$\#|Z| > 2$
2	PO4	C	1001	-	4,4,4	0.91	0	6,6,6	0.46	0
2	PO4	B	1001	-	4,4,4	0.91	0	6,6,6	0.45	0
2	PO4	A	1001	-	4,4,4	0.86	0	6,6,6	0.46	0
2	PO4	D	1001	-	4,4,4	0.91	0	6,6,6	0.42	0

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.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1001	PO4	1	0
2	B	1001	PO4	1	0
2	A	1001	PO4	1	0

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 (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	897/974 (92%)	-0.25	19 (2%)	63	61	49, 80, 131, 214	0
1	B	899/974 (92%)	-0.28	14 (1%)	70	68	51, 82, 135, 216	0
1	C	883/974 (90%)	0.06	33 (3%)	45	43	76, 140, 217, 268	0
1	D	877/974 (90%)	-0.03	20 (2%)	61	58	78, 133, 196, 277	0
All	All	3556/3896 (91%)	-0.13	86 (2%)	59	57	49, 105, 193, 277	0

The worst 5 of 86 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	196	LEU	6.0
1	B	245	HIS	5.2
1	A	245	HIS	4.5
1	A	31	VAL	4.5
1	C	105	ALA	3.8

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 monosaccharides 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
2	PO4	D	1001	5/5	0.90	0.09	99,124,131,141	0
2	PO4	C	1001	5/5	0.91	0.08	119,125,133,135	0
2	PO4	A	1001	5/5	0.95	0.09	57,71,112,127	0
2	PO4	B	1001	5/5	0.98	0.06	57,89,98,105	0

6.5 Other polymers (i)

There are no such residues in this entry.