Full wwPDB X-ray Structure Validation Report (i)

Jan 7, 2025 – 06:08 PM JST

PDB ID	:	8XQZ
Title	:	Crystal structure of AtHPPD-NHPA complex
Authors	:	Dong, J.; Lin, H.-Y.; Yang, G.-F.
Deposited on	:	2024-01-05
Resolution	:	1.78 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.78 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.
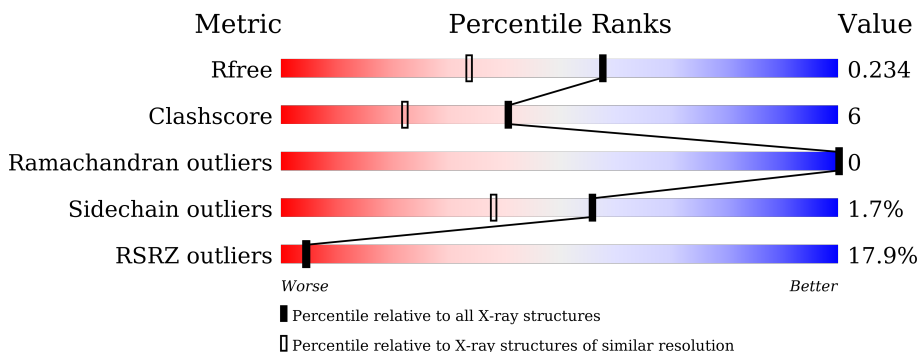
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1191 (1.78-1.78)
Clashscore	180529	1282 (1.78-1.78)
Ramachandran outliers	177936	1270 (1.78-1.78)
Sidechain outliers	177891	1270 (1.78-1.78)
RSRZ outliers	164620	1191 (1.78-1.78)

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	417	16% / 77% / 14% / 9%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3063 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 4-hydroxyphenylpyruvate dioxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	379	Total 2880	C 1833	N 490	O 543	S 14	0	3	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	29	GLY	-	expression tag	UNP P93836
A	30	SER	-	expression tag	UNP P93836
A	31	HIS	-	expression tag	UNP P93836
A	32	MET	-	expression tag	UNP P93836

- Molecule 2 is COBALT (II) ION (three-letter code: CO) (formula: Co).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Co 1	0	0

- Molecule 3 is 3-(4-aminophenyl)-2-oxidanylidene-propanoic acid (three-letter code: A1LWQ) (formula: $C_9H_9NO_3$).

A1LWQ

C05

N07

NH₂

C04

C06

C08 C03 C01

C02

O12 O C09

C10 O O11

OH

O13
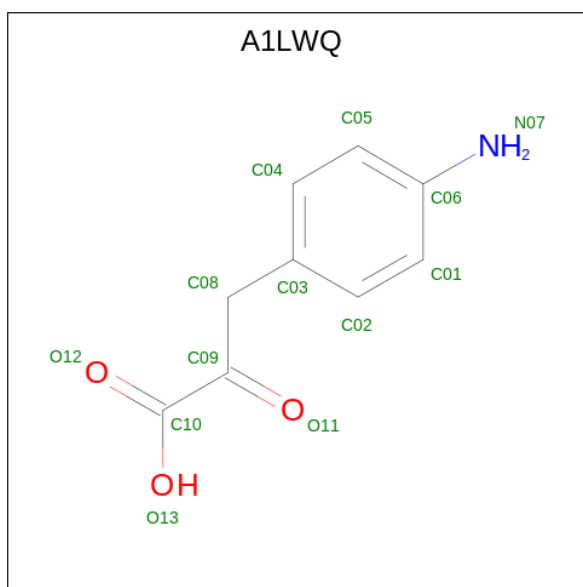
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total C N O 13 9 1 3				0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	169	Total O 169 169		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: 4-hydroxyphenylpyruvate dioxygenase
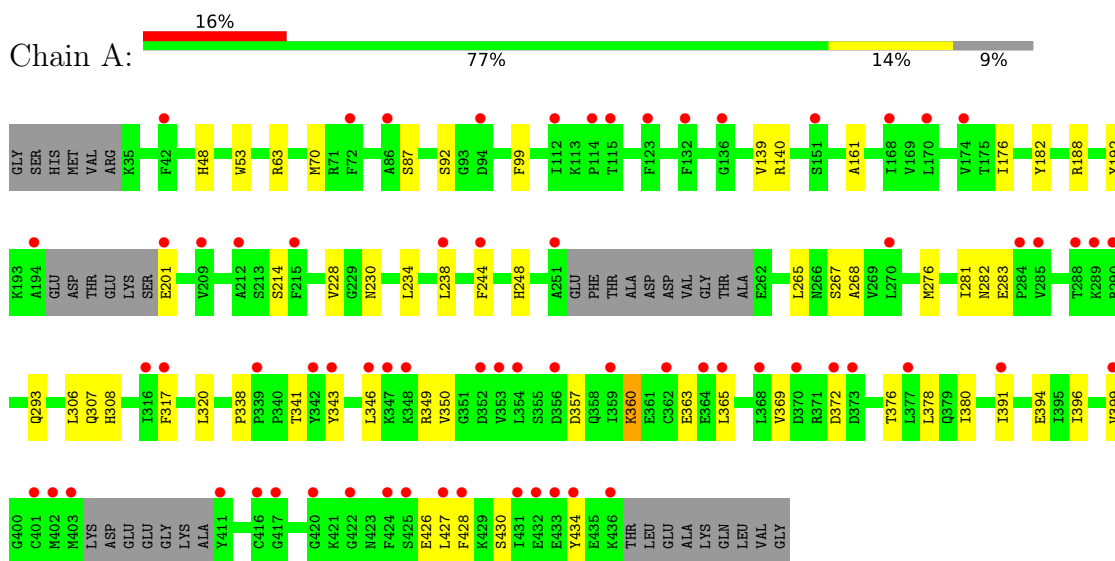
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	76.74Å　83.77Å　61.54Å 90.00°　100.77°　90.00°	Depositor
Resolution (Å)	28.02　–　1.78 28.02　–　1.78	Depositor EDS
% Data completeness (in resolution range)	98.3 (28.02-1.78) 98.4 (28.02-1.78)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.76 (at 1.78Å)	Xtriage
Refinement program	PHENIX 1.14-3260	Depositor
R, R_{free}	0.210　,　0.233 0.211　,　0.234	Depositor DCC
R_{free} test set	1795 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	26.8	Xtriage
Anisotropy	0.713	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 42.9	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3063	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.31% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CO, A1LWQ

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	0.46	1/2964 (0.0%)	0.57	0/4017

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	281	ILE	C-N	10.48	1.58	1.34

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	2880	0	2743	35	0
2	A	1	0	0	0	0
3	A	13	0	0	1	0
4	A	169	0	0	3	0
All	All	3063	0	2743	36	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.

All (36) 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:248:HIS:ND1	4:A:601:HOH:O	2.23	0.72
1:A:228:VAL:HG21	1:A:308:HIS:CE1	2.33	0.64
1:A:376:THR:HB	1:A:399:VAL:HG13	1.83	0.60
1:A:265:LEU:HD23	1:A:293:GLN:HE21	1.67	0.59
1:A:376:THR:HB	1:A:399:VAL:CG1	2.36	0.55
1:A:320:LEU:HD11	1:A:380:ILE:HG21	1.88	0.55
3:A:502:A1LWQ:O13	3:A:502:A1LWQ:C04	2.55	0.54
1:A:238:LEU:HD11	1:A:268:ALA:HB3	1.90	0.54
1:A:176:ILE:HD11	1:A:188:ARG:HD3	1.90	0.53
1:A:48:HIS:CE1	1:A:192:TYR:HH	2.27	0.51
1:A:201:GLU:N	4:A:620:HOH:O	2.42	0.51
1:A:230:ASN:OD1	1:A:282:ASN:HB2	2.10	0.51
1:A:139:VAL:HG11	1:A:306:LEU:HD13	1.92	0.51
1:A:267:SER:OG	1:A:282:ASN:OD1	2.29	0.51
1:A:338:PRO:HG2	1:A:343:TYR:CE2	2.46	0.51
1:A:369:VAL:HG22	1:A:378:LEU:HD23	1.94	0.50
1:A:357:ASP:O	1:A:360:LYS:HG3	2.13	0.49
1:A:161:ALA:HB2	1:A:244:PHE:CG	2.48	0.49
1:A:341:THR:HG21	1:A:434:TYR:HB2	1.94	0.48
1:A:341:THR:HG21	1:A:430:SER:O	2.14	0.48
1:A:234:LEU:HB2	1:A:283:GLU:HB3	1.97	0.47
1:A:341:THR:CG2	1:A:434:TYR:HB2	2.45	0.46
1:A:346:LEU:HD11	1:A:369:VAL:HB	1.98	0.46
1:A:343:TYR:CG	1:A:363:GLU:HB2	2.52	0.45
1:A:307:GLN:O	1:A:391:ILE:HA	2.18	0.44
1:A:343:TYR:CZ	1:A:363:GLU:HA	2.53	0.43
1:A:276:MET:HG3	4:A:640:HOH:O	2.19	0.43
1:A:317:PHE:HD1	1:A:365:LEU:HD12	1.83	0.42
1:A:394:GLU:HG2	1:A:396:ILE:HG23	2.02	0.42
1:A:214:SER:O	1:A:214:SER:OG	2.36	0.41
1:A:349:ARG:NH1	1:A:426:GLU:HG2	2.35	0.41
1:A:338:PRO:HB3	1:A:427:LEU:HD21	2.02	0.41
1:A:53:TRP:CE2	1:A:140:ARG:HG2	2.56	0.41
1:A:70:MET:HG2	1:A:92:SER:HB2	2.02	0.40
1:A:87:SER:HA	1:A:99:PHE:O	2.20	0.40
1:A:350:VAL:HG21	1:A:369:VAL:HG11	2.02	0.40

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	374/417 (90%)	360 (96%)	14 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	304/350 (87%)	299 (98%)	5 (2%)	58	41

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	63	ARG
1	A	182	TYR
1	A	360	LYS
1	A	372	ASP
1	A	428	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	293	GLN
1	A	374	GLN

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 ⓘ

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

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	$\#|Z| > 2$
3	A1LWQ	A	502	2	13,13,13	1.43	4 (30%)	17,17,17	2.33	5 (29%)

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
3	A1LWQ	A	502	2	-	4/8/8/8	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	502	A1LWQ	C08-C09	2.92	1.54	1.51
3	A	502	A1LWQ	C06-N07	2.14	1.45	1.38
3	A	502	A1LWQ	O11-C09	-2.13	1.18	1.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	502	A1LWQ	O12-C10	2.04	1.28	1.22

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	502	A1LWQ	C08-C09-C10	6.05	128.07	117.85
3	A	502	A1LWQ	C03-C08-C09	4.62	119.64	113.64
3	A	502	A1LWQ	O13-C10-O12	-3.01	116.72	123.61
3	A	502	A1LWQ	O11-C09-C08	-3.00	114.56	120.84
3	A	502	A1LWQ	O13-C10-C09	2.70	121.37	113.97

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	502	A1LWQ	C03-C08-C09-C10
3	A	502	A1LWQ	C03-C08-C09-O11
3	A	502	A1LWQ	O11-C09-C10-O13
3	A	502	A1LWQ	C02-C03-C08-C09

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	A1LWQ	1	0

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	379/417 (90%)	1.26	68 (17%)	4	4	17, 33, 67, 85	3 (0%)

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	428	PHE	4.6
1	A	354	LEU	4.4
1	A	72	PHE	4.3
1	A	411	TYR	4.1
1	A	362	CYS	3.9
1	A	373	ASP	3.8
1	A	212	ALA	3.6
1	A	424	PHE	3.3
1	A	123	PHE	3.2
1	A	352	ASP	3.1
1	A	42	PHE	3.1
1	A	132	PHE	3.0
1	A	115	THR	2.9
1	A	209	VAL	2.9
1	A	432	GLU	2.9
1	A	431	ILE	2.9
1	A	151	SER	2.9
1	A	343	TYR	2.9
1	A	359	ILE	2.8
1	A	370	ASP	2.7
1	A	422	GLY	2.7
1	A	174	VAL	2.7
1	A	427	LEU	2.6
1	A	285	VAL	2.6
1	A	353	VAL	2.6
1	A	399	VAL	2.6
1	A	342	TYR	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	251	ALA	2.5
1	A	136	GLY	2.5
1	A	194	ALA	2.5
1	A	372	ASP	2.5
1	A	365	LEU	2.4
1	A	368	LEU	2.4
1	A	403	MET	2.4
1	A	434	TYR	2.4
1	A	377	LEU	2.4
1	A	284	PRO	2.3
1	A	215	PHE	2.3
1	A	347	LYS	2.3
1	A	288	THR	2.3
1	A	238	LEU	2.3
1	A	244	PHE	2.3
1	A	356	ASP	2.3
1	A	402	MET	2.3
1	A	364	GLU	2.3
1	A	416	CYS	2.2
1	A	112	ILE	2.2
1	A	316	ILE	2.2
1	A	317	PHE	2.2
1	A	391	ILE	2.2
1	A	436	LYS	2.2
1	A	289	LYS	2.2
1	A	270	LEU	2.2
1	A	290	ARG	2.1
1	A	433	GLU	2.1
1	A	339	PRO	2.1
1	A	346	LEU	2.1
1	A	94	ASP	2.1
1	A	420	GLY	2.1
1	A	114	PRO	2.1
1	A	170	LEU	2.1
1	A	401	CYS	2.1
1	A	86	ALA	2.0
1	A	201	GLU	2.0
1	A	348	LYS	2.0
1	A	417	GLY	2.0
1	A	425	SER	2.0
1	A	168	ILE	2.0

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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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	A1LWQ	A	502	13/13	0.80	0.15	34,48,58,63	0
2	CO	A	501	1/1	0.98	0.05	33,33,33,33	0

6.5 Other polymers ⓘ

There are no such residues in this entry.