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

wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 06:13 PM GMT

PDB ID	:	4KWV
Title	:	Crystal Structure of human apo-QPRT
Authors	:	Malik, S.S.; Patterson, D.N.; Ncube, Z.; Toth, E.A.
Deposited on	:	2013-05-24
Resolution	:	2.80 Å(reported)

The following versions of software and data (see references (1)) were used in the production of this report:

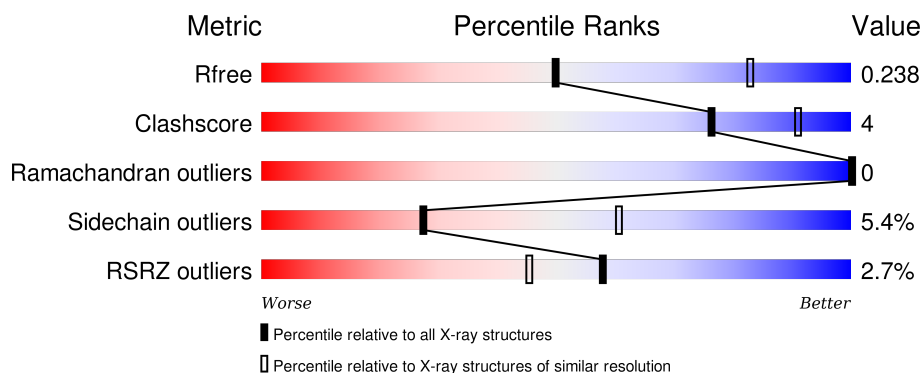|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.80 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		0.238
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		5.4%
RSRZ outliers		2.7%

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$_{free}$	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	301	2% — 84% ... 9% · 5%
1	B	301	4% — 86% ... 8% · 5%
1	C	301	3% — 85% ... 9% · 5%
1	D	301	85% ... 9% · 5%
1	E	301	6% — 84% ... 10% · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	301	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12001 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 Nicotinate-nucleotide pyrophosphorylase [carboxylating].

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	285	Total	C	N	O	S	0	0	0
			1995	1281	337	368	9			
1	B	286	Total	C	N	O	S	0	0	0
			1976	1268	330	369	9			
1	C	285	Total	C	N	O	S	0	0	0
			1951	1243	331	368	9			
1	D	285	Total	C	N	O	S	0	0	0
			2005	1285	341	370	9			
1	E	288	Total	C	N	O	S	0	0	0
			2023	1295	345	374	9			
1	F	291	Total	C	N	O	S	0	0	0
			2036	1301	345	381	9			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	EXPRESSION TAG	UNP Q15274
A	-2	PRO	-	EXPRESSION TAG	UNP Q15274
A	-1	GLY	-	EXPRESSION TAG	UNP Q15274
A	0	SER	-	EXPRESSION TAG	UNP Q15274
B	-3	GLY	-	EXPRESSION TAG	UNP Q15274
B	-2	PRO	-	EXPRESSION TAG	UNP Q15274
B	-1	GLY	-	EXPRESSION TAG	UNP Q15274
B	0	SER	-	EXPRESSION TAG	UNP Q15274
C	-3	GLY	-	EXPRESSION TAG	UNP Q15274
C	-2	PRO	-	EXPRESSION TAG	UNP Q15274
C	-1	GLY	-	EXPRESSION TAG	UNP Q15274
C	0	SER	-	EXPRESSION TAG	UNP Q15274
D	-3	GLY	-	EXPRESSION TAG	UNP Q15274
D	-2	PRO	-	EXPRESSION TAG	UNP Q15274
D	-1	GLY	-	EXPRESSION TAG	UNP Q15274
D	0	SER	-	EXPRESSION TAG	UNP Q15274
E	-3	GLY	-	EXPRESSION TAG	UNP Q15274

Continued on next page...

Continued from previous page...

| Chain	Residue	Modelled	Actual	Comment	Reference
E	-2	PRO	-	EXPRESSION TAG	UNP Q15274
E	-1	GLY	-	EXPRESSION TAG	UNP Q15274
E	0	SER	-	EXPRESSION TAG	UNP Q15274
F	-3	GLY	-	EXPRESSION TAG	UNP Q15274
F	-2	PRO	-	EXPRESSION TAG	UNP Q15274
F	-1	GLY	-	EXPRESSION TAG	UNP Q15274
F	0	SER	-	EXPRESSION TAG	UNP Q15274

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	4	Total 4	O 4	0	0
2	B	1	Total 1	O 1	0	0
2	C	1	Total 1	O 1	0	0
2	D	3	Total 3	O 3	0	0
2	E	1	Total 1	O 1	0	0
2	F	5	Total 5	O 5	0	0

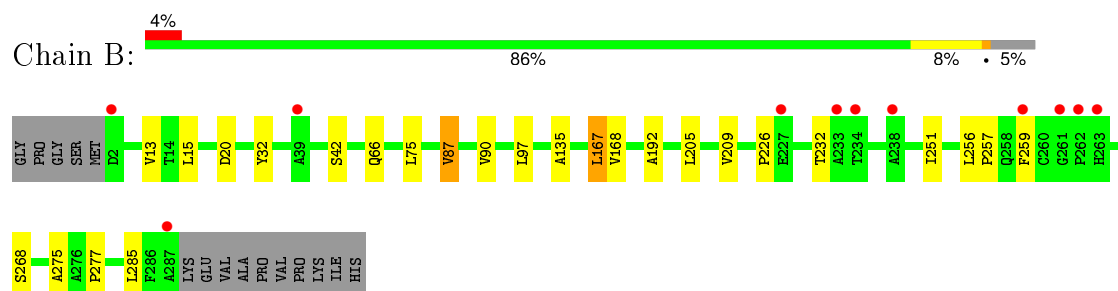# 3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.
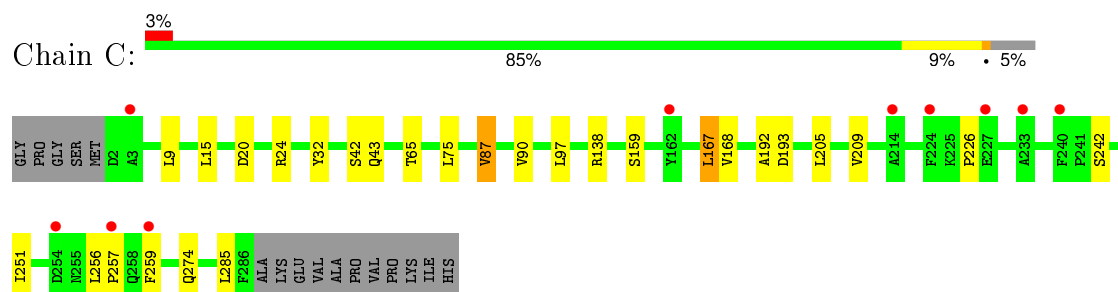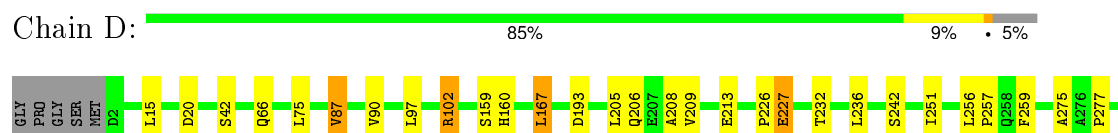
- Molecule 1: Nicotinate-nucleotide pyrophosphorylase [carboxylating]
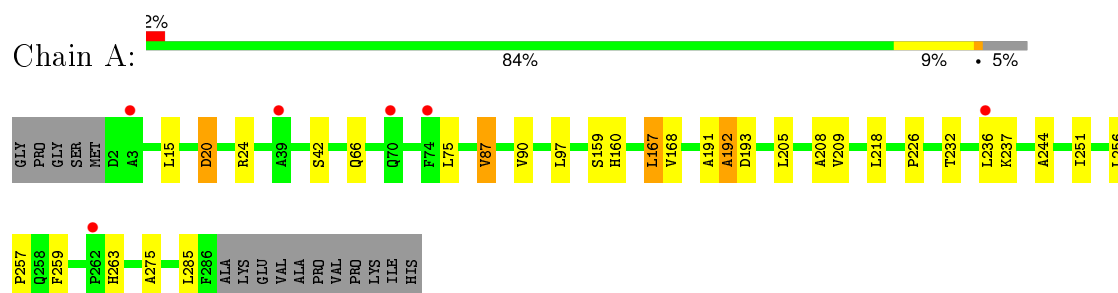
- Molecule 1: Nicotinate-nucleotide pyrophosphorylase [carboxylating]

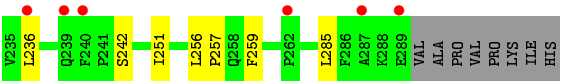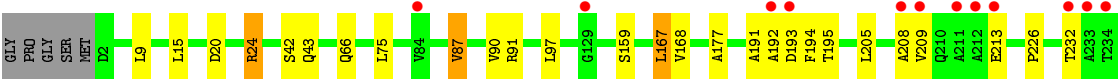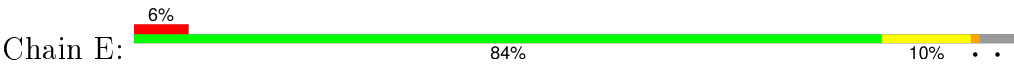- Molecule 1: Nicotinate-nucleotide pyrophosphorylase [carboxylating]

- Molecule 1: Nicotinate-nucleotide pyrophosphorylase [carboxylating]

L285 F286 ALA LYS GLU VAL ALA PRO VAL PRO LYS ILE HIS

● Molecule 1: Nicotinate-nucleotide pyrophosphorylase [carboxylating]

Chain E:

GLY PRO GLY SER MET D2 L9 L15 D20 R24 S42 Q43 Q66 L75 V84 V87 V90 R91 L97 G129 S159 L167 V168 A177 A191 A192 D193 F194 T195 L205 A208 V209 Q210 A211 A212 E213 P226 T232 A233 T234

V235 L236 Q239 F240 P241 S242 I251 L256 P257 Q258 F259 P262 L285 F286 A287 K288 E289 VAL ALA PRO VAL PRO LYS ILE HIS

● Molecule 1: Nicotinate-nucleotide pyrophosphorylase [carboxylating]

Chain F:

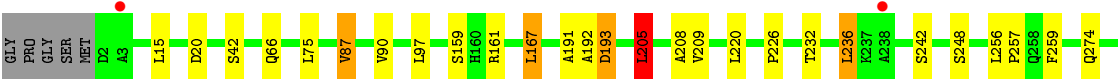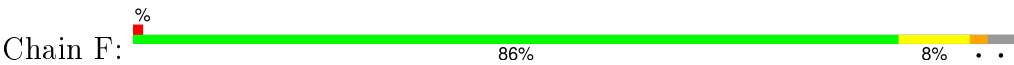GLY PRO GLY SER MET D2 A3 L15 D20 S42 Q66 L75 V87 V90 L97 S159 H160 R161 L167 A191 A192 D193 L205 A208 V209 L220 P226 T232 L236 K237 A238 S242 S248 L256 P257 Q258 F259 Q274

L285 K288 E289 V290 A291 P292 VAL PRO LYS ILE HIS
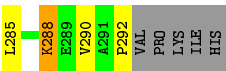
4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	178.56Å 178.56Å 121.60Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.93 – 2.80 42.89 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (42.93-2.80) 99.9 (42.89-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	2.06 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.212 , 0.237 0.212 , 0.238	Depositor DCC
R_{free} test set	2812 reflections (5.35%)	DCC
Wilson B-factor (Å^2)	54.4	Xtriage
Anisotropy	0.058	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 40.4	EDS
Estimated twinning fraction	0.021 for -h,-k,l	Xtriage
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Outliers	0 of 55377 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	12001	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.71% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.34	0/2038	0.54	2/2796 (0.1%)
1	B	0.38	0/2018	0.53	0/2773
1	C	0.37	0/1990	0.54	1/2731 (0.0%)
1	D	0.38	0/2048	0.54	0/2806
1	E	0.34	0/2065	0.55	0/2829
1	F	0.51	0/2078	0.60	4/2849 (0.1%)
All	All	0.39	0/12237	0.55	7/16784 (0.0%)

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	192	ALA	CB-CA-C	5.88	118.91	110.10
1	F	193	ASP	CB-CA-C	-5.28	99.85	110.40
1	F	292	PRO	N-CA-CB	5.20	109.54	103.30
1	A	275	ALA	CB-CA-C	5.17	117.86	110.10
1	F	205	LEU	CB-CG-CD2	-5.15	102.24	111.00

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	1995	0	1941	15	0
1	B	1976	0	1899	15	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1951	0	1861	12	0
1	D	2005	0	1962	16	0
1	E	2023	0	1982	25	0
1	F	2036	0	1986	12	0
2	A	4	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	3	0	0	0	0
2	E	1	0	0	0	0
2	F	5	0	0	0	0
All	All	12001	0	11631	89	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 4.

The worst 5 of 89 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:E:168:VAL:HG21	1:E:192:ALA:HB2	1.31	1.13
1:D:275:ALA:O	1:D:277:PRO:HD3	1.78	0.83
1:B:275:ALA:O	1:B:277:PRO:HD3	1.79	0.82
1:E:168:VAL:CG2	1:E:192:ALA:HB2	2.11	0.80
1:E:168:VAL:HG21	1:E:192:ALA:CB	2.10	0.79

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	283/301 (94%)	282 (100%)	1 (0%)	0	100	100
1	B	284/301 (94%)	282 (99%)	2 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	283/301 (94%)	281 (99%)	2 (1%)	0	100	100
1	D	283/301 (94%)	282 (100%)	1 (0%)	0	100	100
1	E	286/301 (95%)	284 (99%)	2 (1%)	0	100	100
1	F	289/301 (96%)	286 (99%)	3 (1%)	0	100	100
All	All	1708/1806 (95%)	1697 (99%)	11 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	188/222 (85%)	179 (95%)	9 (5%)	31	66
1	B	183/222 (82%)	178 (97%)	5 (3%)	52	85
1	C	179/222 (81%)	167 (93%)	12 (7%)	20	50
1	D	191/222 (86%)	178 (93%)	13 (7%)	20	49
1	E	192/222 (86%)	183 (95%)	9 (5%)	32	67
1	F	195/222 (88%)	182 (93%)	13 (7%)	20	50
All	All	1128/1332 (85%)	1067 (95%)	61 (5%)	27	60

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	42	SER
1	D	206	GLN
1	F	236	LEU
1	D	87	VAL
1	D	159	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	66	GLN
1	D	66	GLN
1	E	66	GLN
1	C	43	GLN
1	D	206	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 carbohydrates 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(Å²)	Q<0.9
1	A	285/301 (94%)	0.07	6 (2%)	67	56	39, 63, 83, 116	0
1	B	286/301 (95%)	0.16	11 (3%)	44	32	40, 63, 89, 107	0
1	C	285/301 (94%)	0.23	10 (3%)	48	35	33, 63, 104, 120	0
1	D	285/301 (94%)	0.06	0	100	100	36, 52, 76, 100	0
1	E	288/301 (95%)	0.20	18 (6%)	23	14	30, 51, 83, 106	0
1	F	291/301 (96%)	-0.08	2 (0%)	89	84	27, 44, 70, 90	0
All	All	1720/1806 (95%)	0.11	47 (2%)	58	45	27, 56, 90, 120	0

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	39	ALA	4.7
1	C	3	ALA	4.0
1	B	287	ALA	3.4
1	C	233	ALA	3.3
1	C	227	GLU	3.1

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 carbohydrates 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.