wwPDB X-ray Structure Validation Summary Report (i)

May 4, 2024 – 06:12 pm BST

PDB ID	:	6RU0
Title	:	Light-Regulation of Imidazole Glycerol Phosphate Synthase by Interference with its Allosteric Machinery through Photo-Sensitive Unnatural Amino Acids
Authors	:	Kneuttinger, A.; Rajendran, C.; Sterner, R.
Deposited on	:	2019-05-27
Resolution	:	2.65 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 (i) 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 (1)) were used in the production of this report:

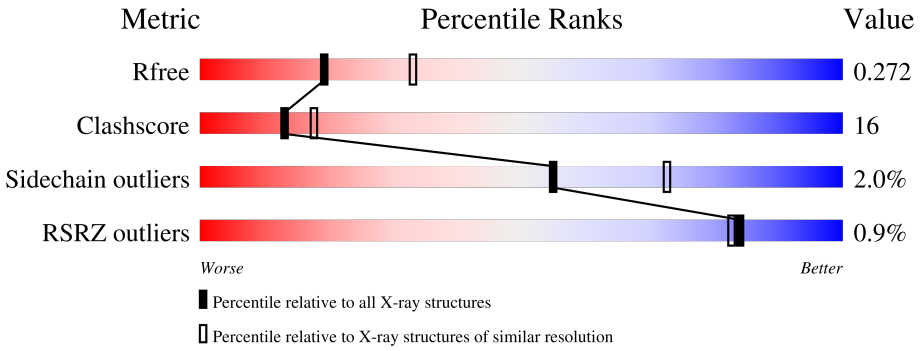
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.65 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}	130704	1426 (2.66-2.62)
Clashscore	141614	1472 (2.66-2.62)
Sidechain outliers	138945	1446 (2.66-2.62)
RSRZ outliers	127900	1408 (2.66-2.62)

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	253	74% 26%
1	C	253	73% 26%
1	E	253	66% 32%
2	B	201	58% 40%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	PO4	A	301	-	-	X	-
3	PO4	C	301	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 10752 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 Imidazole glycerol phosphate synthase subunit HisF.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	253	Total	C	N	O	S	0	1	0
			1958	1250	330	372	6			
1	C	253	Total	C	N	O	S	0	1	0
			1961	1251	333	372	5			
1	E	253	Total	C	N	O	S	0	0	0
			1943	1240	327	370	6			

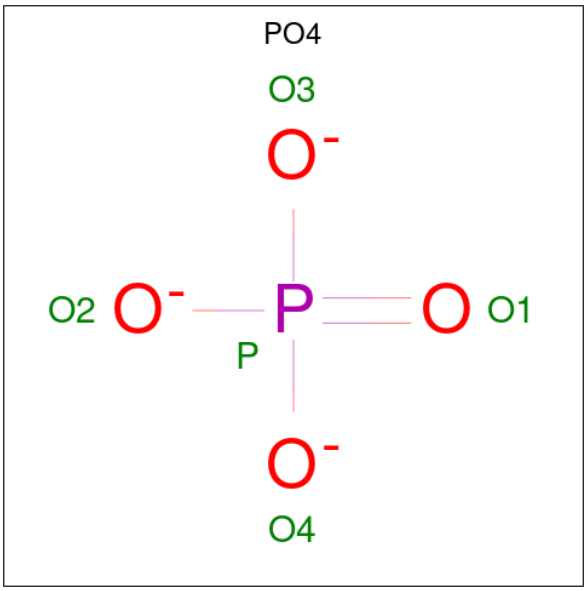
There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	55	PHE	SER	conflict	UNP Q9X0C6
C	55	PHE	SER	conflict	UNP Q9X0C6
E	55	PHE	SER	conflict	UNP Q9X0C6

- Molecule 2 is a protein called Imidazole glycerol phosphate synthase subunit HisH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	200	Total	C	N	O	S	0	0	0
			1595	1018	281	288	8			
2	D	200	Total	C	N	O	S	0	1	0
			1616	1027	288	293	8			
2	F	199	Total	C	N	O	S	0	0	0
			1604	1023	280	293	8			

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O P 5 4 1	0	0
3	C	1	Total O P 5 4 1	0	0
3	E	1	Total O P 5 4 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	15	Total O 15 15	0	0
4	B	8	Total O 8 8	0	0
4	C	23	Total O 23 23	0	0
4	D	5	Total O 5 5	0	0
4	E	5	Total O 5 5	0	0
4	F	4	Total O 4 4	0	0

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	95.42Å 95.42Å 165.93Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.96 – 2.65 45.96 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.6 (45.96-2.65) 99.6 (45.96-2.65)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.14 (at 2.65Å)	Xtriage
Refinement program	PHENIX (1.14_3260: ???)	Depositor
R, R_{free}	0.207 , 0.272 0.208 , 0.272	Depositor DCC
R_{free} test set	2446 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	74.6	Xtriage
Anisotropy	0.117	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 36.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.033 for -h,-k,l 0.118 for h,-h-k,-l 0.042 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10752	wwPDB-VP
Average B, all atoms (Å^2)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.08% 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.

4 Model quality (i)

4.1 Standard geometry (i)

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	0.47	0/1986	0.68	0/2677
1	C	0.50	0/1989	0.69	1/2681 (0.0%)
1	E	0.43	0/1971	0.64	0/2659
2	B	0.57	2/1626 (0.1%)	0.88	5/2191 (0.2%)
2	D	0.46	0/1650	0.65	0/2219
2	F	0.45	0/1638	0.60	0/2204
All	All	0.48	2/10860 (0.0%)	0.69	6/14631 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	D	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	62	ARG	CB-CG	-11.05	1.22	1.52
2	B	62	ARG	CZ-NH2	5.00	1.39	1.33

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	62	ARG	NE-CZ-NH2	-22.02	109.29	120.30
2	B	62	ARG	CG-CD-NE	-8.76	93.41	111.80
2	B	62	ARG	NH1-CZ-NH2	7.90	128.09	119.40
2	B	62	ARG	N-CA-CB	-7.81	96.54	110.60
2	B	62	ARG	CB-CG-CD	-5.67	96.87	111.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	57	GLY	Peptide

4.2 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	1958	0	1998	46	0
1	C	1961	0	2002	54	1
1	E	1943	0	1970	69	0
2	B	1595	0	1580	64	1
2	D	1616	0	1601	83	0
2	F	1604	0	1598	42	0
3	A	5	0	0	2	0
3	C	5	0	0	2	0
3	E	5	0	0	0	0
4	A	15	0	0	1	0
4	B	8	0	0	1	0
4	C	23	0	0	3	0
4	D	5	0	0	0	0
4	E	5	0	0	1	0
4	F	4	0	0	0	0
All	All	10752	0	10749	350	1

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 16.

The worst 5 of 350 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)
2:B:144:ARG:NH1	2:B:145:ALA:O	1.58	1.32
2:D:129:LYS:NZ	2:D:152:VAL:O	2.01	0.93
1:C:45:ASP:OD1	1:C:249:ARG:NH1	2.03	0.92
2:B:6:ILE:HG13	2:B:49:PRO:HD2	1.51	0.92
2:D:5:ILE:HG22	2:D:47:PHE:HB2	1.52	0.91

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:62:ARG:NH2	1:C:28:ASP:OD1[1_565]	1.72	0.48

4.3 Torsion angles (i)

4.3.1 Protein backbone (i)

There are no protein backbone outliers to report in this entry.

4.3.2 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	208/208 (100%)	207 (100%)	1 (0%)	88 94
1	C	208/208 (100%)	207 (100%)	1 (0%)	88 94
1	E	205/208 (99%)	198 (97%)	7 (3%)	37 53
2	B	171/177 (97%)	162 (95%)	9 (5%)	22 36
2	D	174/177 (98%)	172 (99%)	2 (1%)	73 85
2	F	175/177 (99%)	172 (98%)	3 (2%)	60 76
All	All	1141/1155 (99%)	1118 (98%)	23 (2%)	55 72

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	13	LYS
1	E	101	SER
1	E	55	PHE
1	E	154	ARG
2	B	65	ASP

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

Mol	Chain	Res	Type
2	B	151	HIS
2	D	151	HIS

4.3.3 RNA (i)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

3 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	$\#\|Z\| > 2$
3	PO4	A	301	-	4,4,4	1.04	0	6,6,6	0.49	0
3	PO4	E	301	-	4,4,4	0.95	0	6,6,6	0.27	0
3	PO4	C	301	-	4,4,4	0.79	0	6,6,6	0.94	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.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	301	PO4	2	0
3	C	301	PO4	2	0

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 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, 95^{th} 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	253/253 (100%)	-0.26	0 100 100	46, 68, 94, 110	0
1	C	253/253 (100%)	-0.32	1 (0%) 92 93	47, 68, 93, 118	0
1	E	253/253 (100%)	-0.13	2 (0%) 86 85	70, 96, 122, 133	0
2	B	200/201 (99%)	-0.29	4 (2%) 65 61	58, 86, 114, 123	0
2	D	200/201 (99%)	-0.09	3 (1%) 73 71	56, 93, 128, 140	0
2	F	199/201 (99%)	-0.24	2 (1%) 82 81	65, 83, 107, 131	0
All	All	1358/1362 (99%)	-0.22	12 (0%) 84 83	46, 82, 117, 140	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	118	LEU	3.9
2	F	199	SER	3.9
1	E	50	LEU	3.1
2	B	145	ALA	2.6
2	B	29	ASP	2.4

5.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.3 Carbohydrates

There are no monosaccharides in this entry.

5.4 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, 95^{th} 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	PO4	E	301	5/5	0.85	0.14	135,136,141,147	0
3	PO4	C	301	5/5	0.87	0.12	127,131,136,137	0
3	PO4	A	301	5/5	0.95	0.12	142,142,147,150	0

5.5 Other polymers

There are no such residues in this entry.