



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 02:19 pm GMT

PDB ID : 1P1F  
Title : Crystal structure of apo 1L-myo-inositol 1-phosphate synthase  
Authors : Jin, X.; Geiger, J.H.  
Deposited on : 2003-04-12  
Resolution : 2.60 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

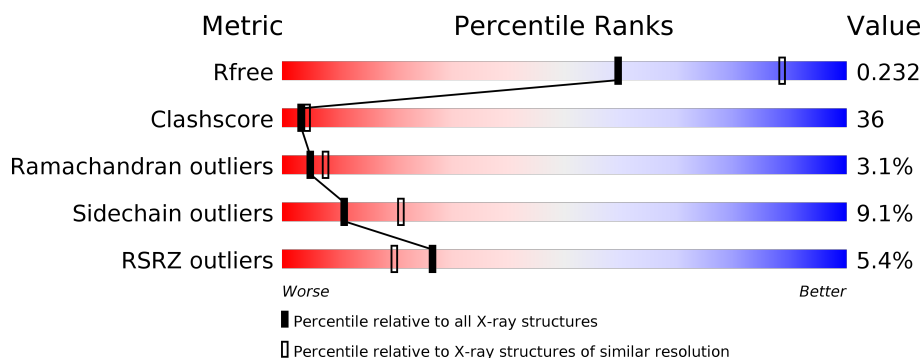
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	100719	2542 (2.60-2.60)
Clashscore	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)
RSRZ outliers	101464	2550 (2.60-2.60)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	533	<div> <div>5%</div> <div> <div></div> <div>43%</div> <div>44%</div> <div>7%</div> <div>6%</div> </div> </div>
1	B	533	<div> <div>6%</div> <div> <div></div> <div>40%</div> <div>45%</div> <div>8%</div> <div>7%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8005 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 Inositol-3-phosphate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	501	Total	C	N	O	S	0	0	0
			3953	2519	661	757	16			
1	B	496	Total	C	N	O	S	0	0	0
			3919	2500	655	748	16			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ARG	SEE REMARK 999	UNP P11986
A	14	VAL	LEU	SEE REMARK 999	UNP P11986
A	?	-	PHE	SEE REMARK 999	UNP P11986
A	60	LEU	GLU	SEE REMARK 999	UNP P11986
A	?	-	ALA	SEE REMARK 999	UNP P11986
A	98	GLU	LYS	SEE REMARK 999	UNP P11986
A	140	ASN	LYS	SEE REMARK 999	UNP P11986
A	141	ASP	HIS	SEE REMARK 999	UNP P11986
A	201	ASN	GLN	SEE REMARK 999	UNP P11986
A	444	PRO	ALA	SEE REMARK 999	UNP P11986
B	?	-	ARG	SEE REMARK 999	UNP P11986
B	14	VAL	LEU	SEE REMARK 999	UNP P11986
B	?	-	PHE	SEE REMARK 999	UNP P11986
B	60	LEU	GLU	SEE REMARK 999	UNP P11986
B	?	-	ALA	SEE REMARK 999	UNP P11986
B	98	GLU	LYS	SEE REMARK 999	UNP P11986
B	140	ASN	LYS	SEE REMARK 999	UNP P11986
B	141	ASP	HIS	SEE REMARK 999	UNP P11986
B	201	ASN	GLN	SEE REMARK 999	UNP P11986
B	444	PRO	ALA	SEE REMARK 999	UNP P11986

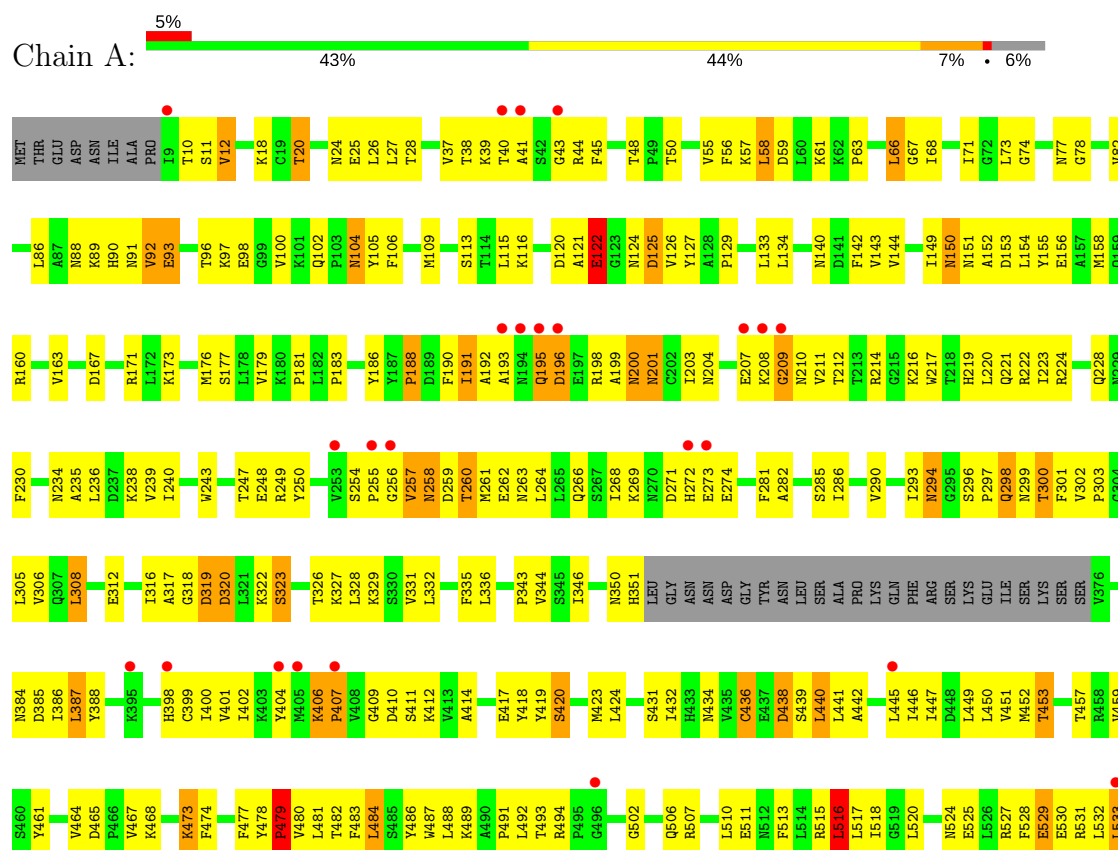
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	80	Total 80	O 80	0	0
2	B	53	Total 53	O 53	0	0

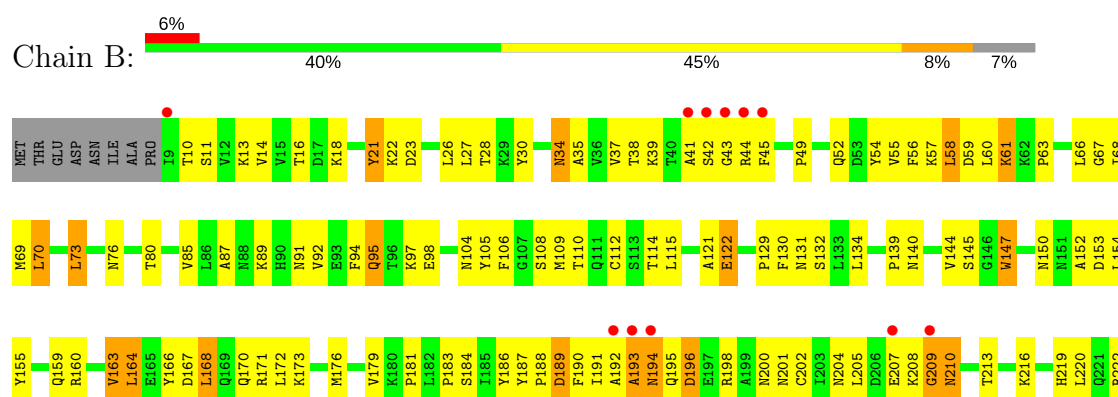
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Inositol-3-phosphate synthase



#### • Molecule 1: Inositol-3-phosphate synthase



T508	A509	L510	E511	N512	F513	L514	R515	L516	L517	I518	G519	L520	P521	S522	Q523	L526	R527	F528	E529	E530	R531	L532	L533																																		
I446	L449	L450	V451	M452	T453	E454	F455	C456	T457	R458	V459	S460	Y461	K462	K463	V464	D465	F466	VAL	LYS	GLU	ASP	ALA	GLY																																	
L223	R224	R225	N229	F230	E233	N234	A235	L236	D237	K238	V239	L242	W243	T244	A245	N246	T247	V251	E252	V253	S254	P255	G256																																		
S256	F257	Q258	N259	T300	F301	V302	L305	V306	Q307	L308	A309	E310	H311	E312	G313	T314	F315	I316	A317	G318	D319	D320	L321	K322	S323	G324	Q325	T326	K327	L328	L332	F335	L336	V337	T341	K342	P343	V344	S345	I346	N350	F281	H351	L352	GLY	ASN	ASN	ASP	GLY	TYR	ASN	LEU	SER	ALA	PRO		
LYS	GLN	PHE	ARG	SER	LYS	GLU	ILE	SER	LYS	SER	SER	V376	I377	D378	D379	I380	I381	N384	L387	D390	K391	L392	K395	V396	D397	H398	I402	K403	Y404	M405	K406	P407	V408	G409	D410	S411	K412	M415	Y419	M423	G426	H427	M428	V500	N501	Q502	L503	Q506	R507								
L446	L449	L450	V451	M452	T453	E454	F455	C456	T457	R458	V459	S460	Y461	K462	K463	V464	D465	F466	VAL	LYS	GLU	ASP	ALA	GLY	K473	F474	E475	N476	F477	Y478	P479	Y480	L481	T482	F483	L484	S485	Y486	W487	L488	K489	L490	P491	L492	T493	R494	P495	G496	F497	H498	P499	V500	N501	Q502	L503	Q506	R507
L223	R224	R225	N229	F230	E233	N234	A235	L236	D237	K238	V239	L242	W243	T244	A245	N246	T247	V251	E252	V253	S254	P255	G256	D259	T260	M261	E262	L265	Q266	K269	E273	E274	I275	A276	T279	I280	F281	A282	A283	A284	S285	I286	L287	E288	G289	V290	I293	N294	G295								

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	153.54Å 97.06Å 122.06Å 90.00° 125.72° 90.00°	Depositor
Resolution (Å)	10.00 – 2.60 38.29 – 2.60	Depositor EDS
% Data completeness (in resolution range)	(Not available) (10.00-2.60) 91.0 (38.29-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.13 (at 2.61Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.190 , 0.278 0.228 , 0.232	Depositor DCC
$R_{free}$ test set	2228 reflections (5.06%)	DCC
Wilson B-factor (Å <sup>2</sup> )	38.9	Xtriage
Anisotropy	0.537	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 62.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	8005	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.15% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.38	0/4030	0.65	2/5465 (0.0%)
1	B	0.50	1/3995 (0.0%)	0.70	3/5417 (0.1%)
All	All	0.45	1/8025 (0.0%)	0.68	5/10882 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	463	LYS	C-N	20.08	1.80	1.34

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	463	LYS	C-N-CA	-11.40	93.20	121.70
1	B	463	LYS	O-C-N	8.52	136.34	122.70
1	B	463	LYS	CA-C-N	-6.60	102.67	117.20
1	A	516	LEU	CA-CB-CG	6.24	129.65	115.30
1	A	319	ASP	N-CA-C	5.38	125.53	111.00

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	3953	0	3964	282	0
1	B	3919	0	3933	308	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	80	0	0	8	0
2	B	53	0	0	7	0
All	All	8005	0	7897	565	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 36.

The worst 5 of 565 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:B:55:VAL:HG23	1:B:464:VAL:CG2	1.28	1.61
1:B:55:VAL:CG2	1:B:464:VAL:CG2	1.96	1.42
1:B:55:VAL:CG2	1:B:464:VAL:HG21	1.54	1.36
1:B:463:LYS:C	1:B:464:VAL:N	1.80	1.34
1:B:293:ILE:HA	1:B:317:ALA:HB2	1.33	1.10

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	497/533 (93%)	427 (86%)	55 (11%)	15 (3%)	5	8
1	B	490/533 (92%)	409 (84%)	65 (13%)	16 (3%)	4	7
All	All	987/1066 (93%)	836 (85%)	120 (12%)	31 (3%)	5	8

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	122	GLU
1	A	191	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	193	ALA
1	B	210	ASN
1	B	311	HIS

### 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	443/471 (94%)	404 (91%)	39 (9%)	12	22
1	B	440/471 (93%)	399 (91%)	41 (9%)	10	20
All	All	883/942 (94%)	803 (91%)	80 (9%)	11	21

5 of 80 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	529	GLU
1	B	73	LEU
1	B	412	LYS
1	A	530	GLU
1	B	34	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 58 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	294	ASN
1	A	512	ASN
1	B	428	ASN
1	A	298	GLN
1	A	311	HIS

### 5.3.3 RNA ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	501/533 (93%)	0.13	24 (4%) 31 24	13, 43, 77, 80	0
1	B	496/533 (93%)	0.21	30 (6%) 23 17	17, 44, 80, 80	0
All	All	997/1066 (93%)	0.17	54 (5%) 26 20	13, 43, 79, 80	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	404	TYR	6.7
1	B	407	PRO	5.6
1	B	464	VAL	5.1
1	A	194	ASN	4.4
1	A	195	GLN	4.1

### 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.