WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 07:51 PM GMT

PDB ID	:	4WMJ
Title	:	Colias eurytheme Phosphoglucose isomerase. Homodimer from 4-5(18) geno-type.
Authors	:	Hill, J.A.; Watt, W.B.
Deposited on	:	2014-10-09
Resolution	:	1.50 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
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)	:	rb-20026982

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 1.50 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.
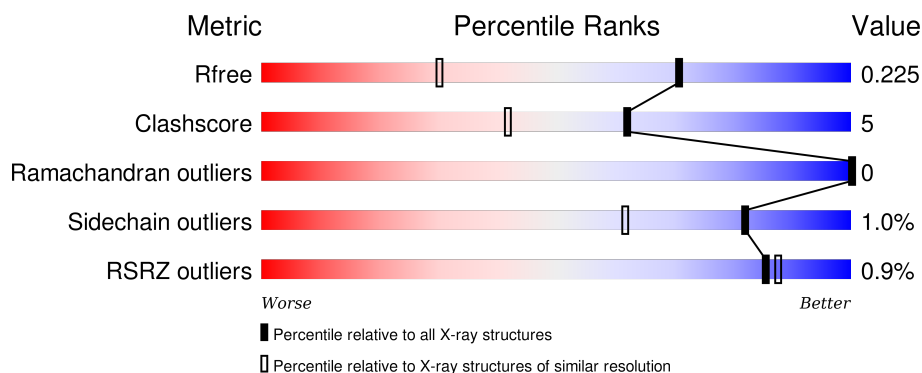
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2072 (1.50-1.50)
Clashscore	102246	2274 (1.50-1.50)
Ramachandran outliers	100387	2218 (1.50-1.50)
Sidechain outliers	100360	2216 (1.50-1.50)
RSRZ outliers	91569	2075 (1.50-1.50)

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	556	88% 11% •
1	B	556	86% 14%
1	C	556	87% 13%
1	D	556	87% 12% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 18508 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 Glucose-6-phosphate isomerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	553	Total 4348	C 2791	N 740	O 805	S 12	0	0	0
1	B	555	Total 4364	C 2801	N 742	O 809	S 12	0	0	0
1	C	554	Total 4355	C 2796	N 741	O 806	S 12	0	0	0
1	D	553	Total 4348	C 2791	N 740	O 805	S 12	0	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	264	Total 264	O 264	0	0
2	B	251	Total 251	O 251	0	0
2	C	289	Total 289	O 289	0	0
2	D	289	Total 289	O 289	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: Glucose-6-phosphate isomerase
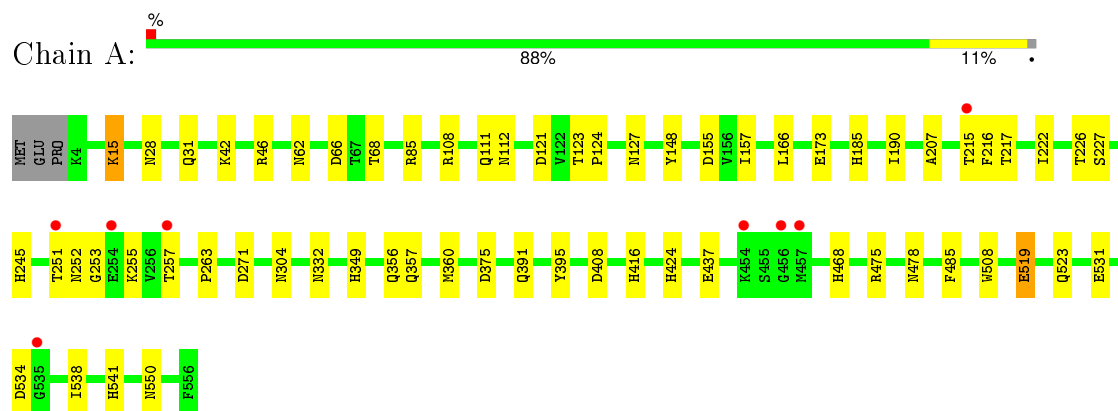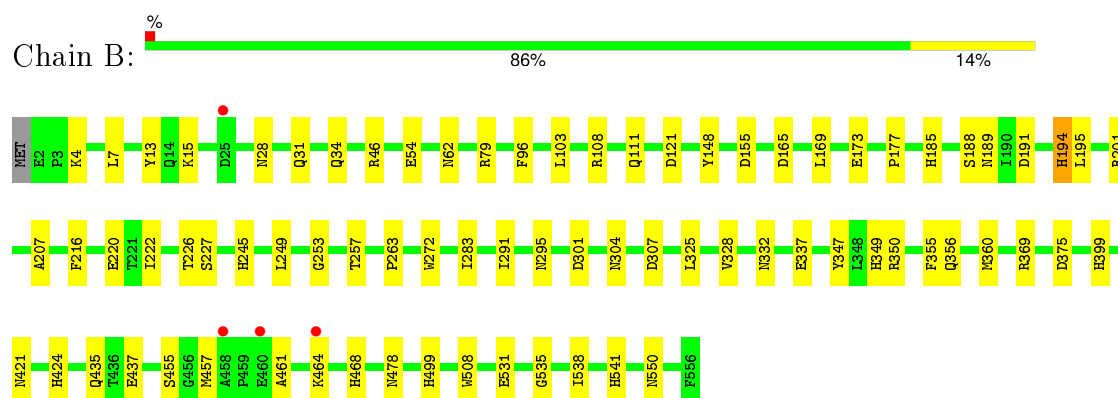
- Molecule 1: Glucose-6-phosphate isomerase

- Molecule 1: Glucose-6-phosphate isomerase

R475 N478 V482 F485 T489 H499 W508 E531 G535 I538 H541 S544 N550 E554 N555 F556

● Molecule 1: Glucose-6-phosphate isomerase

Chain D:

87%　　　　12%

MET GLU PRO K4 K8 K15 Y19 E54 S60 K61 N62 D65 R85 K93 F96 R108 Q111 N118 D121 V122 T123 Y148 D155 E173 H185 A207 F216 E220 T221 I222 T226 S227 H245 F246 V247 A248 N252

G253 E254 K255 V256 T257 P263 K264 N265 M266 F267 D271 L279 N304 D307 A319 P320 N332 H349 V367 T368 R369 D375 Q391 Y395 D408 H416 H424 E437 D445 K454 S465 G456 M457 A458 P459 E460 H468 R475 N478

F485 H499 W508 F513 D514 Q515 V518 D534 G535 K536 K537 I538 N550 F556
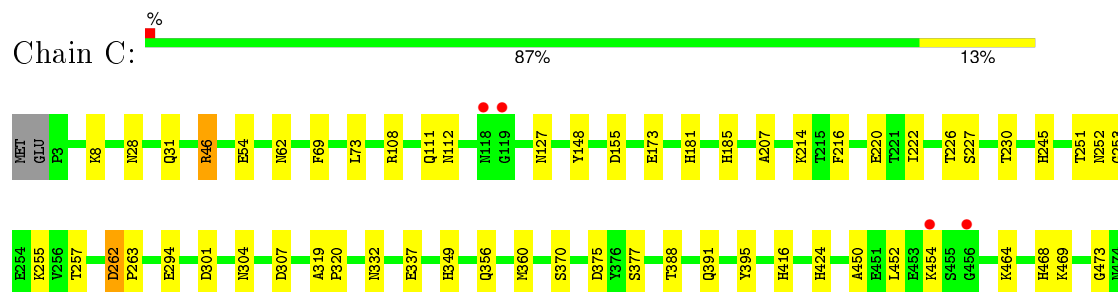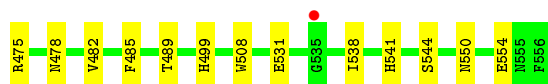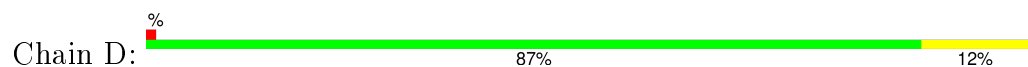
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	76.29Å 115.50Å 135.08Å 90.00° 101.70° 90.00°	Depositor
Resolution (Å)	41.19 – 1.50 41.19 – 1.50	Depositor EDS
% Data completeness (in resolution range)	77.7 (41.19-1.50) 77.7 (41.19-1.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.10 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.183 , 0.218 0.192 , 0.225	Depositor DCC
R_{free} test set	14356 reflections (5.32%)	DCC
Wilson B-factor (Å^2)	15.8	Xtriage
Anisotropy	0.135	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 34.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Outliers	1 of 284531 reflections (0.000%)	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	18508	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.75% 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	0.98	1/4455 (0.0%)	0.95	10/6045 (0.2%)
1	B	0.99	2/4472 (0.0%)	1.02	11/6069 (0.2%)
1	C	0.99	6/4463 (0.1%)	0.97	8/6056 (0.1%)
1	D	1.04	3/4455 (0.1%)	0.96	8/6045 (0.1%)
All	All	1.00	12/17845 (0.1%)	0.98	37/24215 (0.2%)

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	220	GLU	CD-OE1	9.73	1.36	1.25
1	B	337	GLU	CD-OE2	-8.09	1.16	1.25
1	D	220	GLU	CD-OE1	7.02	1.33	1.25
1	A	519	GLU	CD-OE2	6.44	1.32	1.25
1	C	294	GLU	CG-CD	6.14	1.61	1.51

The worst 5 of 37 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	201	ARG	NE-CZ-NH1	15.57	128.08	120.30
1	B	201	ARG	NE-CZ-NH2	-8.44	116.08	120.30
1	C	8	LYS	CD-CE-NZ	7.77	129.57	111.70
1	A	46	ARG	NE-CZ-NH1	7.60	124.10	120.30
1	D	271	ASP	CB-CG-OD1	7.48	125.03	118.30

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	4348	0	4301	45	0
1	B	4364	0	4314	55	0
1	C	4355	0	4309	48	0
1	D	4348	0	4301	51	0
2	A	264	0	0	7	0
2	B	251	0	0	6	0
2	C	289	0	0	5	0
2	D	289	0	0	6	0
All	All	18508	0	17225	186	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 5.

The worst 5 of 186 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:7:LEU:HD21	1:B:13:TYR:CD2	1.63	1.33
1:C:468:HIS:CD2	1:D:367:VAL:HG11	1.81	1.16
1:B:7:LEU:HD21	1:B:13:TYR:CE2	1.83	1.11
1:A:111:GLN:HE22	1:A:304:ASN:HD21	1.15	0.94
1:B:7:LEU:CD2	1:B:13:TYR:CD2	2.51	0.94

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	551/556 (99%)	534 (97%)	17 (3%)	0	100	100
1	B	553/556 (100%)	532 (96%)	21 (4%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	552/556 (99%)	535 (97%)	17 (3%)	0	100	100
1	D	551/556 (99%)	533 (97%)	18 (3%)	0	100	100
All	All	2207/2224 (99%)	2134 (97%)	73 (3%)	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	459/462 (99%)	455 (99%)	4 (1%)	84	65
1	B	461/462 (100%)	456 (99%)	5 (1%)	80	58
1	C	460/462 (100%)	457 (99%)	3 (1%)	88	73
1	D	459/462 (99%)	453 (99%)	6 (1%)	76	50
All	All	1839/1848 (100%)	1821 (99%)	18 (1%)	82	62

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	421	ASN
1	C	108	ARG
1	D	216	PHE
1	B	194	HIS
1	B	216	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 67 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	533	GLN
1	C	185	HIS
1	D	424	HIS
1	B	541	HIS
1	C	28	ASN

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(Å^2)	Q<0.9
1	A	553/556 (99%)	-0.25	8 (1%)	78	81	10, 16, 32, 47	0
1	B	555/556 (99%)	-0.23	4 (0%)	89	91	11, 17, 31, 60	0
1	C	554/556 (99%)	-0.19	5 (0%)	85	87	10, 16, 30, 47	0
1	D	553/556 (99%)	-0.28	4 (0%)	89	91	10, 16, 28, 55	0
All	All	2215/2224 (99%)	-0.24	21 (0%)	85	87	10, 16, 31, 60	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	458	ALA	5.1
1	A	251	THR	4.0
1	B	458	ALA	3.7
1	C	535	GLY	3.3
1	C	456	GLY	3.3

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.