wwPDB X-ray Structure Validation Summary Report (i)

Nov 22, 2023 – 07:56 PM JST

PDB ID	:	7XHP
Title	:	Structure of a Glucose 6-Phosphate Dehydrogenase from Zymomonas mobilis
Authors	:	Meng, D.D.; Liu, M.X.; Liu, W.D.; You, C.
Deposited on	:	2022-04-09
Resolution	:	2.78 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 (i)) were used in the production of this report:

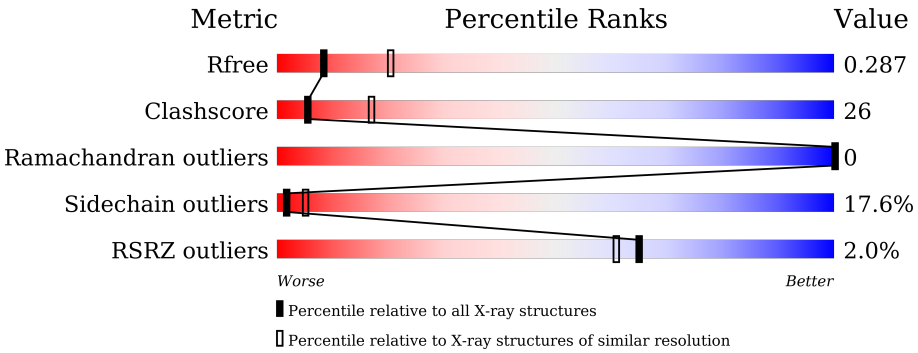
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.78 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.



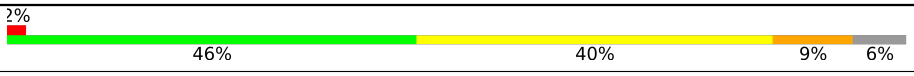
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4107 (2.80-2.76)
Clashscore	141614	4575 (2.80-2.76)
Ramachandran outliers	138981	4487 (2.80-2.76)
Sidechain outliers	138945	4489 (2.80-2.76)
RSRZ outliers	127900	4027 (2.80-2.76)

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	493	2% 49% 36% 6% 9%
1	B	493	2% 42% 42% 9% 6%
1	C	493	2% 37% 29% • 30%
1	D	493	% 40% 29% • 27%
1	E	493	2% 45% 39% 6% 10%
1	F	493	2% 49% 41% 7% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	493	
1	H	493	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 26350 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 Dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	451	Total	C	N	O	S	0	0	0
			3319	2106	565	638	10			
1	B	461	Total	C	N	O	S	0	0	0
			3513	2235	601	666	11			
1	C	345	Total	C	N	O	S	0	0	0
			2684	1716	468	493	7			
1	D	361	Total	C	N	O	S	0	0	0
			2750	1749	486	508	7			
1	E	446	Total	C	N	O	S	0	0	0
			3324	2116	567	631	10			
1	F	480	Total	C	N	O	S	0	0	0
			3628	2305	620	692	11			
1	G	465	Total	C	N	O	S	0	0	0
			3521	2243	598	669	11			
1	H	474	Total	C	N	O	S	0	0	0
			3499	2224	591	674	10			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	17	Total	O	0	0
			17	17		
2	B	12	Total	O	0	0
			12	12		
2	C	11	Total	O	0	0
			11	11		
2	D	10	Total	O	0	0
			10	10		
2	E	8	Total	O	0	0
			8	8		
2	F	22	Total	O	0	0
			22	22		

Continued on next page...

Continued from previous page...

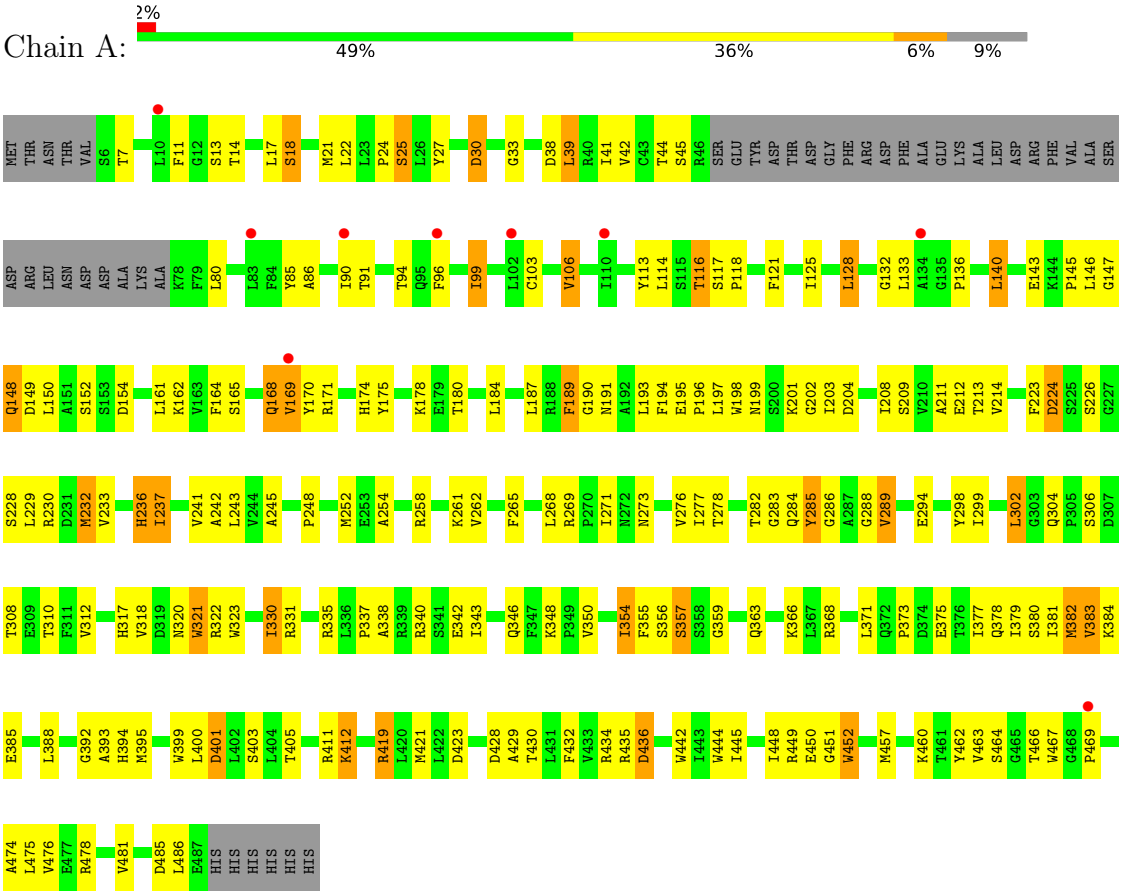
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	19	Total O 19 19	0	0
2	H	13	Total O 13 13	0	0

3 Residue-property plots (i)

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.

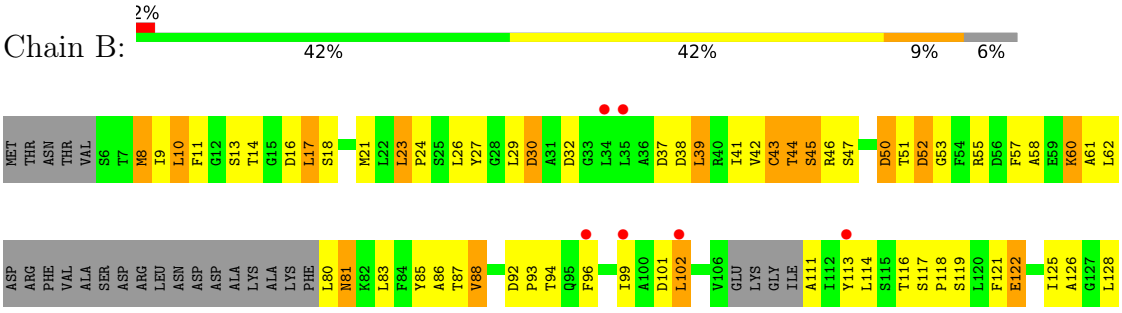
• Molecule 1: Glucose 6-Phosphate Dehydrogenase

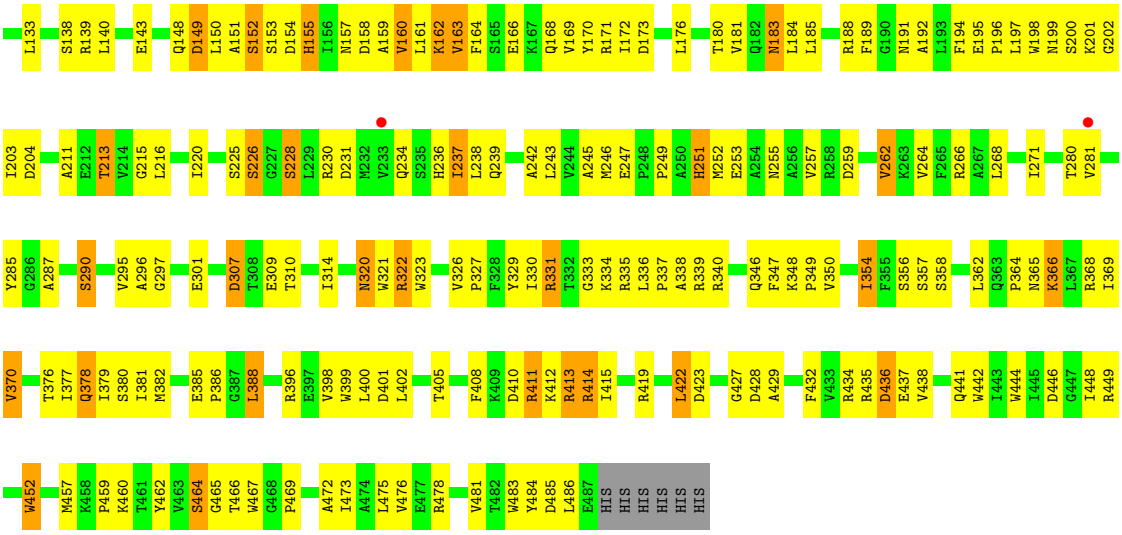
Chain A:



• Molecule 1: Glucose 6-Phosphate Dehydrogenase

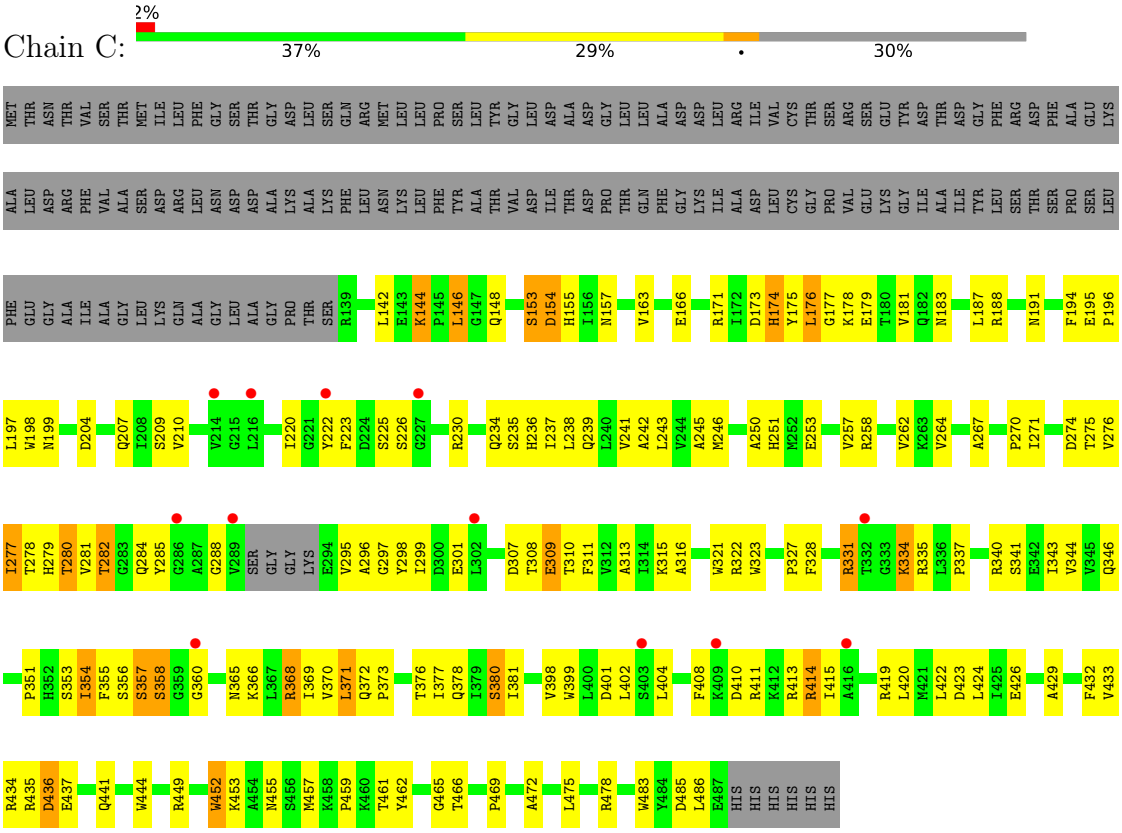
Chain B:





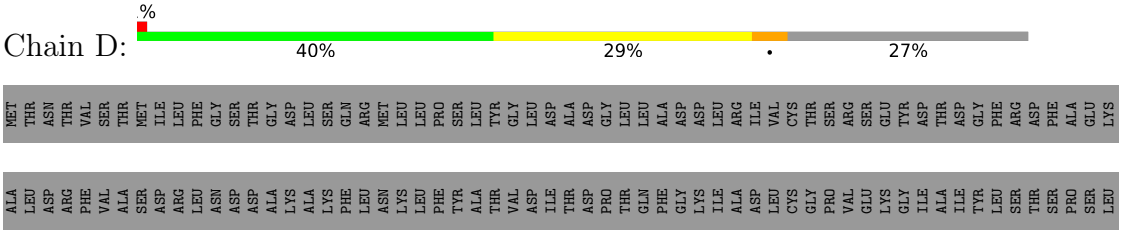
• Molecule 1: Glucose 6-Phosphate Dehydrogenase

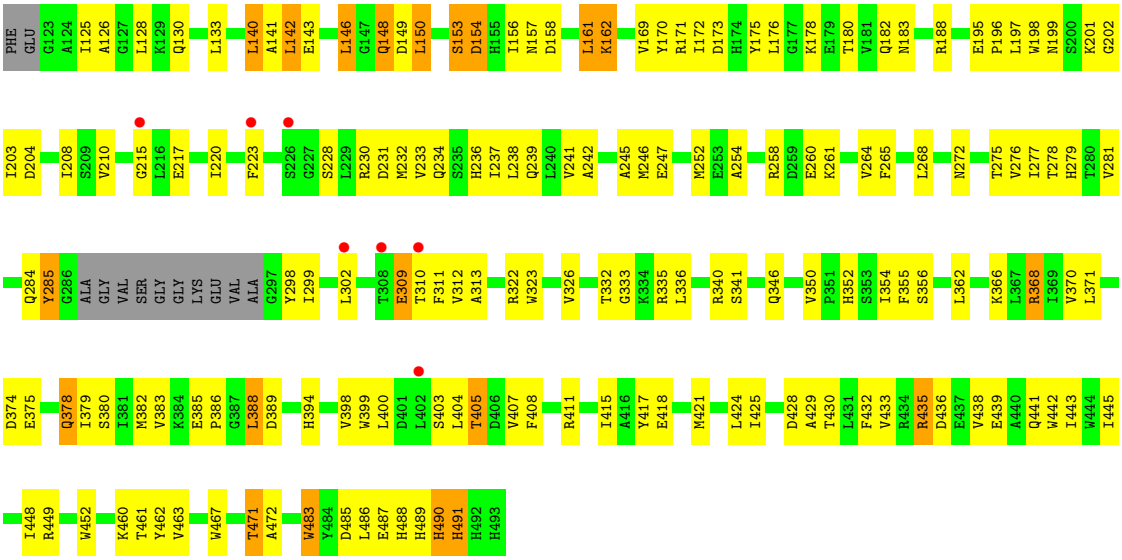
Chain C:



• Molecule 1: Glucose 6-Phosphate Dehydrogenase

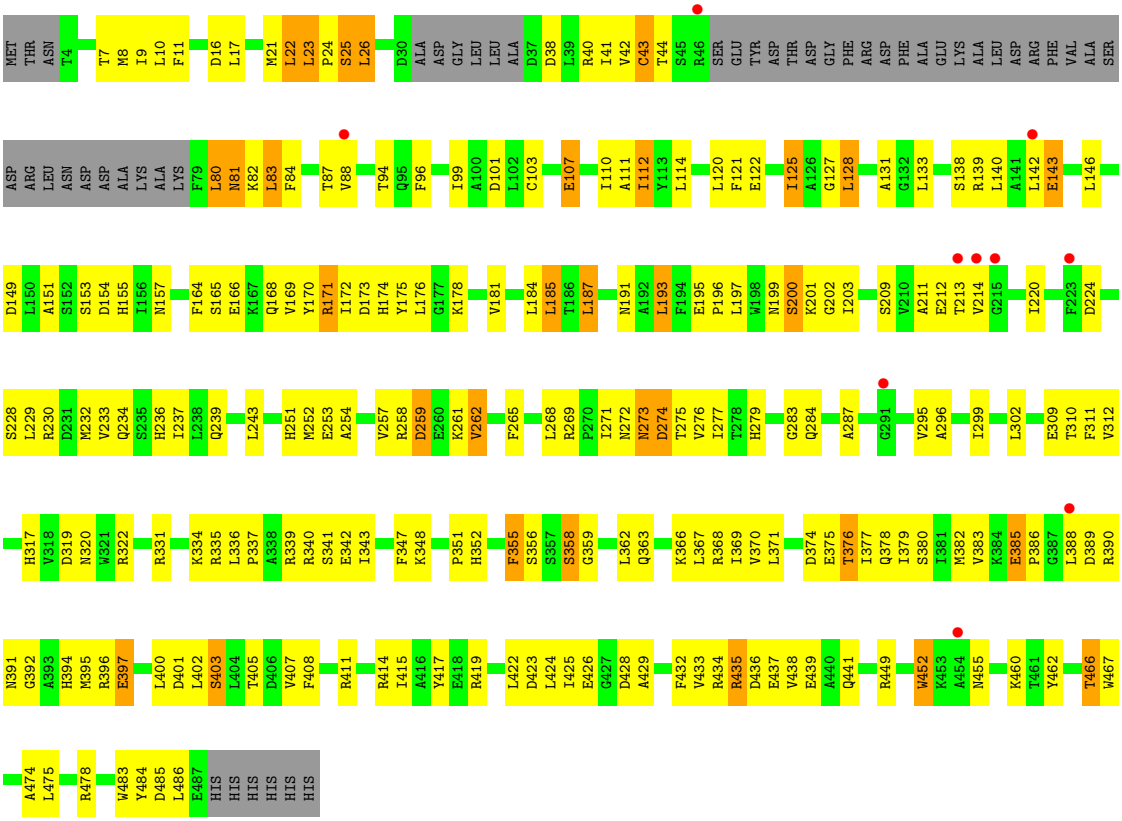
Chain D:





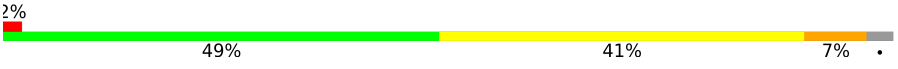
• Molecule 1: Glucose 6-Phosphate Dehydrogenase

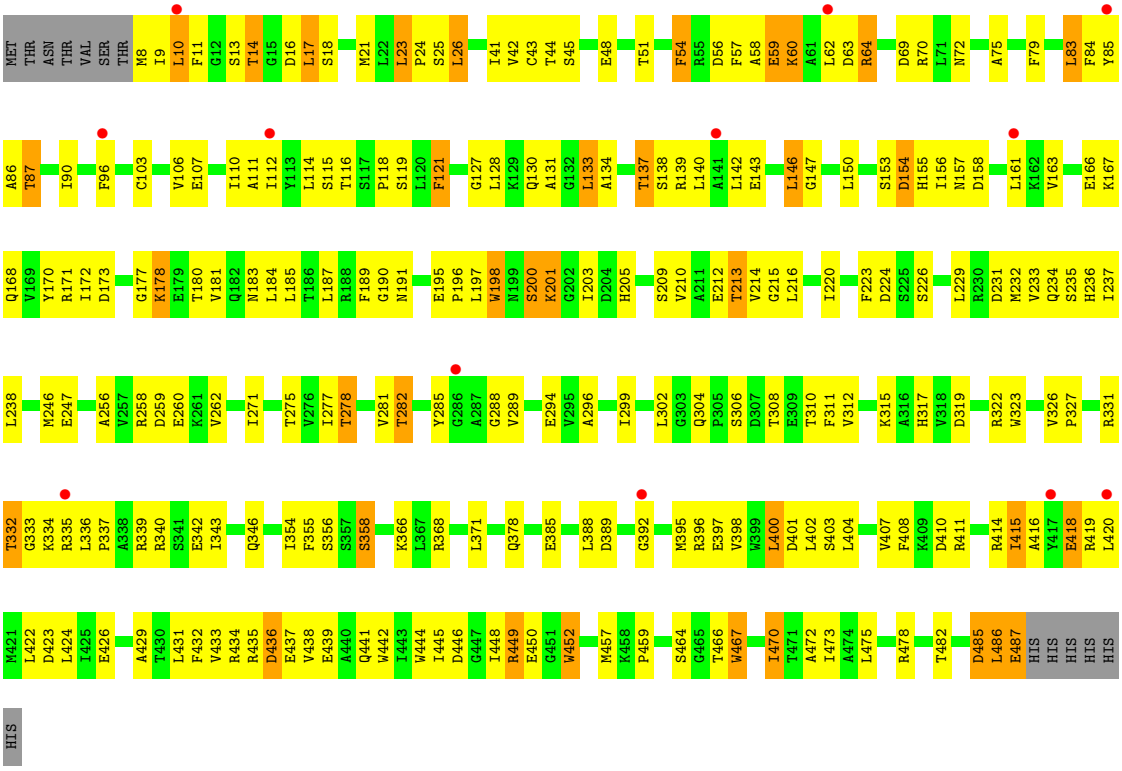
Chain E:



• Molecule 1: Glucose 6-Phosphate Dehydrogenase

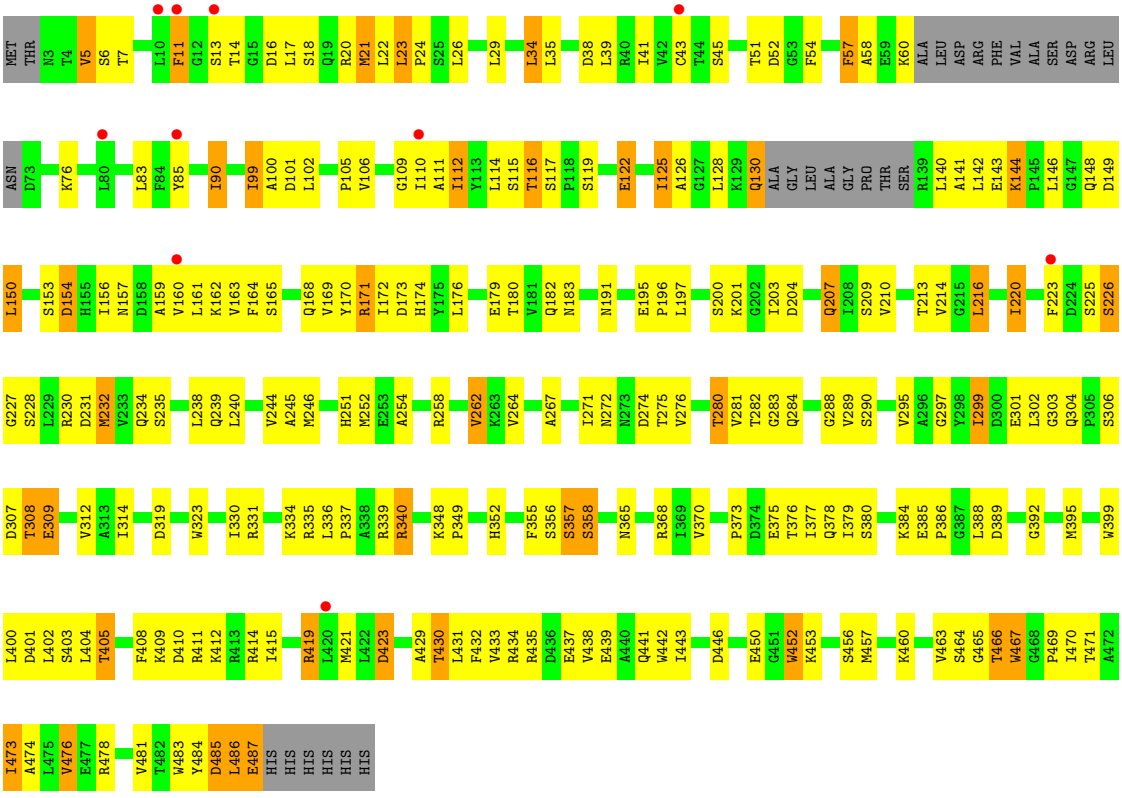
Chain F:





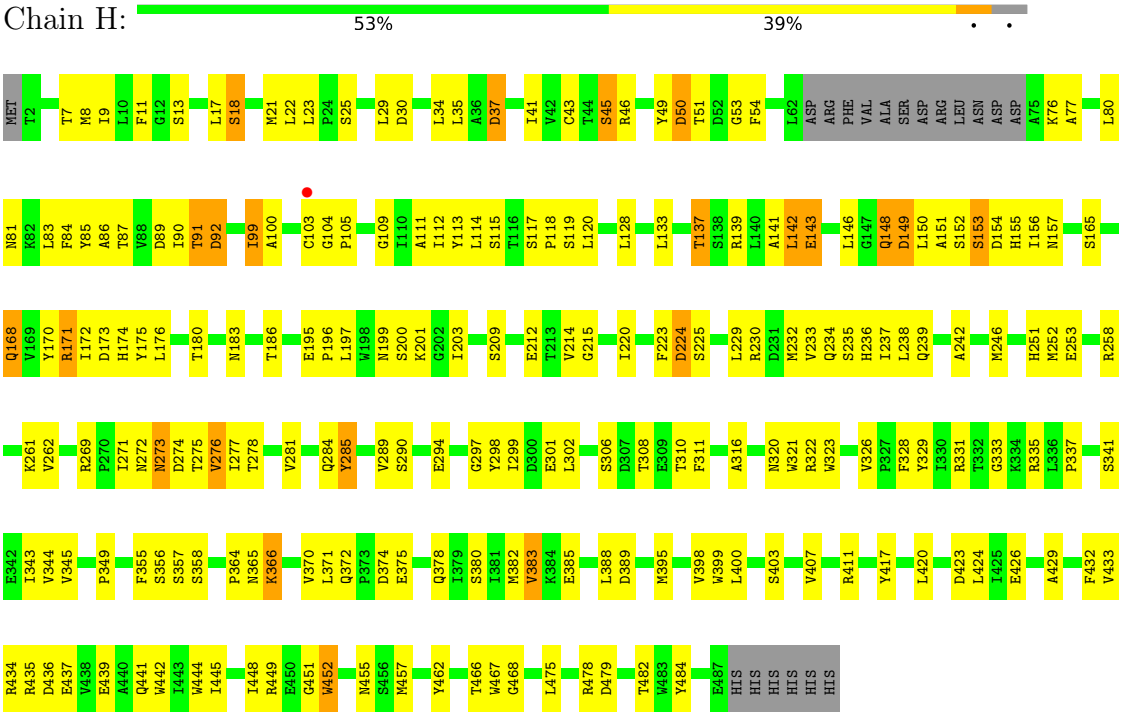
• Molecule 1: Glucose 6-Phosphate Dehydrogenase

Chain G:



• Molecule 1: Glucose 6-Phosphate Dehydrogenase

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	80.61Å 101.07Å 321.07Å 90.00° 90.02° 90.00°	Depositor
Resolution (Å)	78.18 – 2.78 85.53 – 2.78	Depositor EDS
% Data completeness (in resolution range)	90.8 (78.18-2.78) 89.4 (85.53-2.78)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.73 (at 2.77Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.233 , 0.289 0.247 , 0.287	Depositor DCC
R_{free} test set	5876 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	56.6	Xtriage
Anisotropy	0.181	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 79.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.35$, $< L^2 > = 0.18$	Xtriage
Estimated twinning fraction	0.439 for h,-k,-l	Xtriage
Reported twinning fraction	0.500 for h,-k,-l	Depositor
Outliers	0 of 118119 reflections	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	26350	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.80% 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

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	0.46	0/3390	0.69	0/4634
1	B	0.51	0/3586	0.74	0/4881
1	C	0.38	0/2744	0.65	0/3733
1	D	0.41	0/2817	0.67	0/3844
1	E	0.49	0/3396	0.71	0/4637
1	F	0.46	0/3706	0.69	0/5049
1	G	0.47	0/3594	0.73	1/4894 (0.0%)
1	H	0.40	0/3575	0.65	0/4886
All	All	0.45	0/26808	0.69	1/36558 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	469	PRO	N-CA-CB	5.88	110.36	103.30

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	3319	0	3109	155	0
1	B	3513	0	3417	209	0
1	C	2684	0	2635	128	0
1	D	2750	0	2611	155	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	3324	0	3150	176	0
1	F	3628	0	3471	191	0
1	G	3521	0	3377	212	0
1	H	3499	0	3273	181	0
2	A	17	0	0	2	0
2	B	12	0	0	0	0
2	C	11	0	0	4	0
2	D	10	0	0	0	0
2	E	8	0	0	1	0
2	F	22	0	0	5	0
2	G	19	0	0	1	0
2	H	13	0	0	1	0
All	All	26350	0	25043	1326	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 26.

The worst 5 of 1326 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:F:8:MET:HG2	1:F:111:ALA:HB3	1.43	1.00
1:H:100:ALA:HB2	1:H:133:LEU:HD11	1.45	0.96
1:B:215:GLY:HA3	1:B:335:ARG:HH11	1.32	0.95
1:G:125:ILE:HG21	1:G:160:VAL:HA	1.49	0.94
1:G:6:SER:HA	1:G:39:LEU:HD13	1.49	0.93

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	447/493 (91%)	415 (93%)	32 (7%)	0	100	100
1	B	455/493 (92%)	425 (93%)	30 (7%)	0	100	100
1	C	341/493 (69%)	310 (91%)	31 (9%)	0	100	100
1	D	357/493 (72%)	327 (92%)	30 (8%)	0	100	100
1	E	440/493 (89%)	396 (90%)	44 (10%)	0	100	100
1	F	478/493 (97%)	440 (92%)	38 (8%)	0	100	100
1	G	459/493 (93%)	421 (92%)	38 (8%)	0	100	100
1	H	470/493 (95%)	431 (92%)	39 (8%)	0	100	100
All	All	3447/3944 (87%)	3165 (92%)	282 (8%)	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	330/417 (79%)	263 (80%)	67 (20%)	1	3
1	B	366/417 (88%)	289 (79%)	77 (21%)	1	3
1	C	280/417 (67%)	238 (85%)	42 (15%)	3	8
1	D	278/417 (67%)	241 (87%)	37 (13%)	4	11
1	E	336/417 (81%)	264 (79%)	72 (21%)	1	2
1	F	370/417 (89%)	307 (83%)	63 (17%)	2	5
1	G	361/417 (87%)	288 (80%)	73 (20%)	1	3
1	H	348/417 (84%)	309 (89%)	39 (11%)	6	16
All	All	2669/3336 (80%)	2199 (82%)	470 (18%)	2	5

5 of 470 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	80	LEU
1	H	142	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	405	THR
1	H	91	THR
1	G	290	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	207	GLN
1	H	365	ASN
1	G	284	GLN
1	G	378	GLN
1	H	455	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 monosaccharides 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

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, 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	451/493 (91%)	0.10	9 (1%)	65	61	34, 51, 95, 109	0
1	B	461/493 (93%)	0.02	8 (1%)	70	67	31, 48, 88, 114	0
1	C	345/493 (69%)	0.15	12 (3%)	44	38	37, 58, 73, 85	0
1	D	361/493 (73%)	0.06	7 (1%)	66	63	40, 58, 72, 83	0
1	E	446/493 (90%)	0.09	10 (2%)	62	57	36, 61, 83, 94	0
1	F	480/493 (97%)	0.13	12 (2%)	57	52	33, 59, 82, 98	0
1	G	465/493 (94%)	0.07	10 (2%)	62	57	32, 48, 85, 105	0
1	H	474/493 (96%)	-0.08	1 (0%)	95	95	30, 47, 65, 79	0
All	All	3483/3944 (88%)	0.06	69 (1%)	65	61	30, 54, 83, 114	0

The worst 5 of 69 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	214	VAL	6.9
1	A	10	LEU	5.3
1	E	142	LEU	4.6
1	D	215	GLY	4.6
1	B	35	LEU	4.5

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.