wwPDB X-ray Structure Validation Summary Report (i)

Jun 24, 2024 – 03:08 PM EDT

PDB ID	:	6GHB
Title	:	Crystal structure of Spx in complex with YjbH (oxidized)
Authors	:	Awad, W.; Logan, D.T.; von Wachenfeldt, C.
Deposited on	:	2018-05-06
Resolution	:	3.10 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:

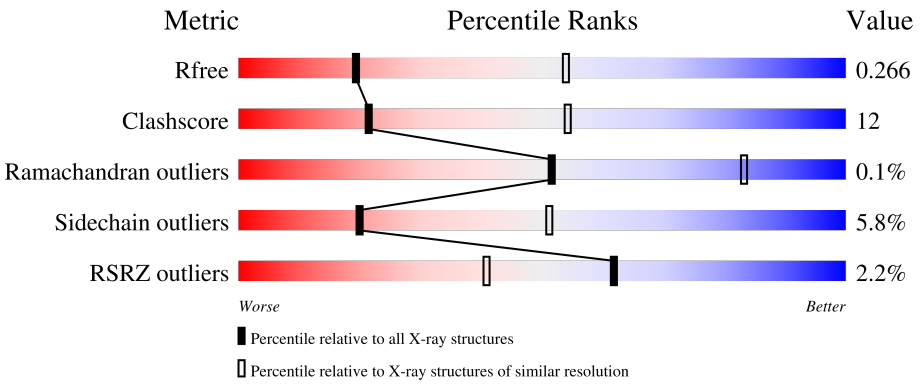
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	132	2%, 53%, 32%, •, 12%
1	C	132	%, 60%, 31%, •, 7%
2	B	298	%, 63%, 24%, •, 12%
2	D	298	3%, 69%, 19%, •, 10%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6155 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 Regulatory protein Spx.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	116	947	601	165	176	5	0	0	0
1	C	123	993	629	172	187	5	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP O31602
C	0	SER	-	expression tag	UNP O31602

- Molecule 2 is a protein called UPF0413 protein GK0824.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	263	2094	1357	339	385	13	0	0	0
2	D	267	2106	1362	341	390	13	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	SER	-	expression tag	UNP Q5L1S1
D	0	SER	-	expression tag	UNP Q5L1S1

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
3	B	1	1	1	0	0
3	D	1	1	1	0	0

- Molecule 4 is water.

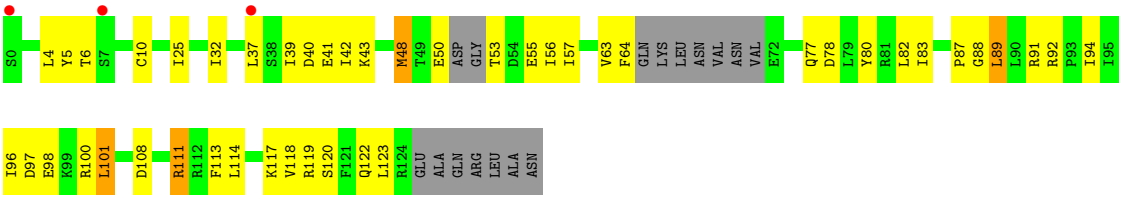
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	3	Total O 3 3	0	0
4	B	4	Total O 4 4	0	0
4	C	1	Total O 1 1	0	0
4	D	5	Total O 5 5	0	0

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.

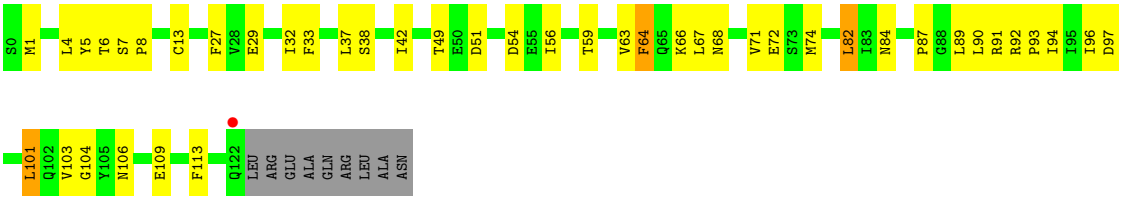
• Molecule 1: Regulatory protein Spx

Chain A:



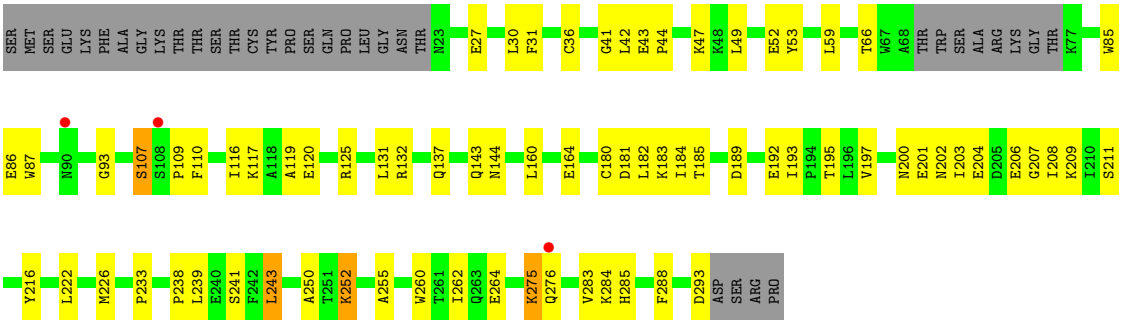
• Molecule 1: Regulatory protein Spx

Chain C:



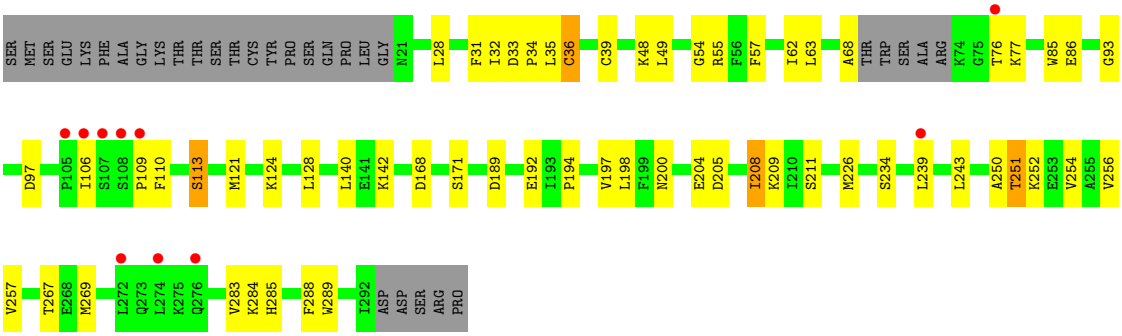
• Molecule 2: UPF0413 protein GK0824

Chain B:



• Molecule 2: UPF0413 protein GK0824

Chain D:


SER MET SER GLU LYS PHE ALA GLY LYS THR THR SER THR CYS TYR PRO SER GLN PRO LEU GLY N21 L28 F31 I32 D33 P34 L35 C36 C39 K48 L49 G54 R55 F56 F57 I62 L63 A68 THR TRP SER ALA ARG K74 G75 T76 K77 W85 E86 G93
D97 P105 I106 S107 S108 P109 F110 S113 M121 K124 L128 L140 E141 K142 D168 S171 D189 E192 I193 P194 V197 L198 F199 N200 E204 D205 I208 K209 I210 S211 M226 S234 L239 L243 A250 T251 K252 E253 V254 A255 V256
V257 T267 E268 M269 L272 Q273 L274 K275 Q276 V283 K284 H285 F288 W289 I292 ASP ASP SER ARG PRO

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.60Å 88.70Å 117.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.43 – 3.10 48.43 – 3.11	Depositor EDS
% Data completeness (in resolution range)	99.0 (48.43-3.10) 98.9 (48.43-3.11)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.44 (at 3.12Å)	Xtriage
Refinement program	PHENIX (1.13_2998: 000)	Depositor
R, R_{free}	0.222 , 0.266 0.222 , 0.266	Depositor DCC
R_{free} test set	1642 reflections (9.94%)	wwPDB-VP
Wilson B-factor (Å^2)	65.0	Xtriage
Anisotropy	0.614	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 48.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.002 for k,h,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6155	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 24.84 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.5525e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MG

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	0.52	0/963	0.67	1/1299 (0.1%)
1	C	0.53	0/1011	0.71	0/1368
2	B	0.53	0/2149	0.62	0/2917
2	D	0.47	0/2159	0.57	0/2930
All	All	0.51	0/6282	0.63	1/8514 (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	A	89	LEU	CB-CG-CD1	-5.30	101.98	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	947	0	944	40	0
1	C	993	0	987	31	0
2	B	2094	0	2062	48	0
2	D	2106	0	2075	36	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	3	0	0	1	0
4	B	4	0	0	2	0
4	C	1	0	0	0	0
4	D	5	0	0	0	0
All	All	6155	0	6068	147	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 12.

The worst 5 of 147 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:A:63:VAL:HG23	1:A:64:PHE:HA	1.55	0.89
1:C:32:ILE:HG22	1:C:87:PRO:HB3	1.60	0.84
2:B:283:VAL:HG22	2:B:284:LYS:H	1.43	0.83
1:C:71:VAL:HG22	1:C:72:GLU:HA	1.68	0.75
1:C:96:ILE:HD11	1:C:101:LEU:HD12	1.68	0.73

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	110/132 (83%)	106 (96%)	4 (4%)	0	100 100
1	C	121/132 (92%)	116 (96%)	5 (4%)	0	100 100
2	B	259/298 (87%)	251 (97%)	7 (3%)	1 (0%)	34 69
2	D	263/298 (88%)	252 (96%)	11 (4%)	0	100 100
All	All	753/860 (88%)	725 (96%)	27 (4%)	1 (0%)	51 83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	109	PRO

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	107/125 (86%)	99 (92%)	8 (8%)	13	42
1	C	112/125 (90%)	104 (93%)	8 (7%)	14	44
2	B	225/258 (87%)	209 (93%)	16 (7%)	14	44
2	D	226/258 (88%)	219 (97%)	7 (3%)	40	70
All	All	670/766 (88%)	631 (94%)	39 (6%)	20	51

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	67	LEU
2	D	113	SER
1	C	68	ASN
2	D	36	CYS
2	D	251	THR

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

Mol	Chain	Res	Type
1	A	122	GLN
2	B	276	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

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	116/132 (87%)	0.16	3 (2%)	56	33	51, 69, 96, 106	0
1	C	123/132 (93%)	0.16	1 (0%)	86	72	54, 72, 94, 104	0
2	B	263/298 (88%)	-0.13	3 (1%)	80	64	37, 59, 82, 94	0
2	D	267/298 (89%)	0.09	10 (3%)	41	21	46, 66, 101, 123	0
All	All	769/860 (89%)	0.04	17 (2%)	62	41	37, 64, 94, 123	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	274	LEU	3.5
2	B	108	SER	3.2
2	B	276	GLN	3.2
1	A	7	SER	3.1
2	D	276	GLN	3.0

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 monosaccharides in this entry.

6.4 Ligands (i)

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, 95th 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	MG	D	301	1/1	0.69	0.27	66,66,66,66	0
3	MG	B	301	1/1	0.80	0.20	65,65,65,65	0

6.5 Other polymers (i)

There are no such residues in this entry.