Full wwPDB X-ray Structure Validation Report ⓘ

Jan 5, 2024 – 02:36 am GMT

PDB ID	:	5FPM
Title	:	Structure of heat shock-related 70kDA protein 2 with small-molecule ligand 5 -phenyl-1,3,4-oxadiazole-2-thiol (AT809) in an alternate binding site.
Authors	:	Jhoti, H.; Ludlow, R.F.; Patel, S.; Saini, H.K.; Tickle, I.J.; Verdonk, M.
Deposited on	:	2015-12-02
Resolution	:	1.96 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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 ⓘ) were used in the production of this report:

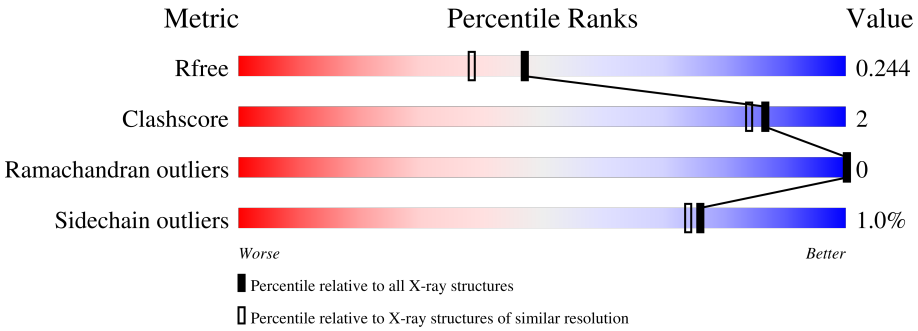
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 1.96 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)

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%

Mol	Chain	Length	Quality of chain
1	A	387	93% 5% •
1	B	387	92% 6% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6905 atoms, of which 12 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.

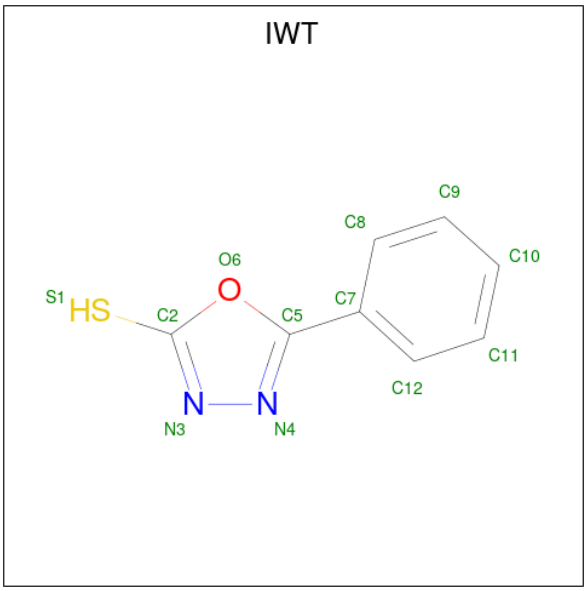
- Molecule 1 is a protein called HEAT SHOCK-RELATED 70KDA PROTEIN 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	381	Total	C	N	O	S	0	0	2
			2889	1822	501	559	7			
1	B	379	Total	C	N	O	S	0	0	0
			2887	1820	496	563	8			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP P54652
A	-2	SER	-	expression tag	UNP P54652
A	-1	HIS	-	expression tag	UNP P54652
A	?	-	SER	deletion	UNP P54652
A	?	-	ALA	deletion	UNP P54652
B	-3	GLY	-	expression tag	UNP P54652
B	-2	SER	-	expression tag	UNP P54652
B	-1	HIS	-	expression tag	UNP P54652
B	?	-	SER	deletion	UNP P54652
B	?	-	ALA	deletion	UNP P54652

- Molecule 2 is 5-PHENYL-1,3,4-OXADIAZOLE-2-THIOL (three-letter code: IWT) (formula: $C_8H_6N_2OS$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	S	0	0
			18	8	6	2	1	1		
2	B	1	Total	C	H	N	O	S	0	0
			18	8	6	2	1	1		

- Molecule 3 is water.

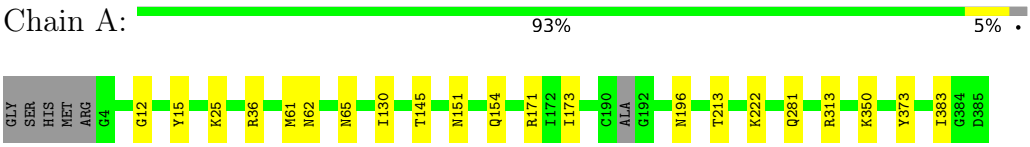
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	559	Total	O	0	0
			559	559		
3	B	534	Total	O	0	0
			534	534		

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. 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. 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.

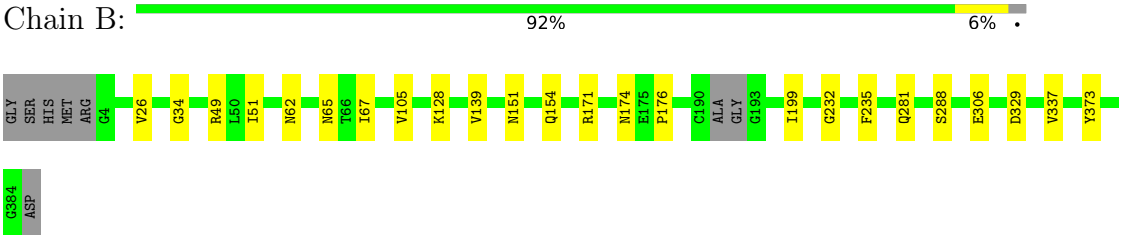
• Molecule 1: HEAT SHOCK-RELATED 70KDA PROTEIN 2

Chain A:



• Molecule 1: HEAT SHOCK-RELATED 70KDA PROTEIN 2

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	48.86Å 94.96Å 81.54Å 90.00° 89.98° 90.00°	Depositor
Resolution (Å)	38.37 – 1.96 38.34 – 1.95	Depositor EDS
% Data completeness (in resolution range)	97.6 (38.37-1.96) 96.2 (38.34-1.95)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.08 (at 1.95Å)	Xtriage
Refinement program	BUSTER 2.11.6	Depositor
R, R_{free}	0.170 , 0.218 0.234 , 0.244	Depositor DCC
R_{free} test set	2660 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	21.1	Xtriage
Anisotropy	0.168	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 53.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.36$	Xtriage
Estimated twinning fraction	0.360 for h,-k,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6905	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.73% 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

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

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	0.62	0/2934	0.62	0/3964
1	B	0.63	0/2932	0.62	0/3962
All	All	0.62	0/5866	0.62	0/7926

There are no bond length outliers.

There are no bond angle outliers.

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	2889	0	2852	13	0
1	B	2887	0	2843	12	2
2	A	12	6	6	1	0
2	B	12	6	6	1	0
3	A	559	0	0	4	3
3	B	534	0	0	0	3
All	All	6893	12	5707	24	4

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 2.

All (24) 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:49:ARG:HD3	1:B:51:ILE:HD11	1.62	0.81
1:A:130:ILE:HG12	3:A:2084:HOH:O	1.93	0.68
1:B:151:ASN:H	1:B:154:GLN:HE21	1.41	0.67
1:A:151:ASN:H	1:A:154:GLN:HE21	1.42	0.65
1:B:62:ASN:HD21	1:B:65:ASN:HD22	1.54	0.55
1:A:151:ASN:H	1:A:154:GLN:NE2	2.07	0.53
1:A:350:LYS:HG2	3:A:2528:HOH:O	2.08	0.52
1:A:36:ARG:HG2	3:A:2073:HOH:O	2.08	0.52
1:B:232:GLY:HA2	1:B:235:PHE:CD2	2.48	0.49
1:B:199:ILE:HD13	1:B:337:VAL:HB	1.95	0.49
1:B:171:ARG:HA	2:B:1385:IWT:S1	2.54	0.47
1:B:174:ASN:HB3	1:B:176:PRO:HD2	1.97	0.47
1:A:61:MET:O	1:B:34:GLY:HA3	2.14	0.47
1:A:12:GLY:HA3	1:A:15:TYR:O	2.17	0.45
1:A:213:THR:HG23	1:A:222:LYS:HG3	1.99	0.45
1:A:62:ASN:HD21	1:A:65:ASN:HD22	1.64	0.44
1:A:25:LYS:HG3	3:A:2046:HOH:O	2.17	0.44
1:B:128:LYS:HE2	1:B:139:VAL:O	2.18	0.43
1:A:196:ASN:OD1	1:A:213:THR:HG22	2.18	0.43
1:B:26:VAL:HG11	1:B:373:TYR:HA	2.01	0.42
1:B:151:ASN:H	1:B:154:GLN:NE2	2.13	0.42
1:B:67:ILE:HD11	1:B:105:VAL:CG1	2.50	0.42
1:A:145:THR:HG22	1:A:173:ILE:HG13	2.01	0.41
1:A:171:ARG:HA	2:A:1385:IWT:S1	2.61	0.40

All (4) 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)
1:B:306:GLU:OE2	3:B:2270:HOH:O[2_746]	1.90	0.30
1:B:329:ASP:OD2	3:A:2464:HOH:O[2_646]	2.08	0.12
3:A:2304:HOH:O	3:B:2312:HOH:O[1_554]	2.09	0.11
3:A:2411:HOH:O	3:B:2511:HOH:O[2_656]	2.09	0.11

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	377/387 (97%)	369 (98%)	8 (2%)	0	100	100
1	B	375/387 (97%)	371 (99%)	4 (1%)	0	100	100
All	All	752/774 (97%)	740 (98%)	12 (2%)	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	296/316 (94%)	292 (99%)	4 (1%)	67	62
1	B	298/316 (94%)	296 (99%)	2 (1%)	84	82
All	All	594/632 (94%)	588 (99%)	6 (1%)	76	74

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	281	GLN
1	A	313	ARG
1	A	373	TYR
1	A	383	ILE
1	B	281	GLN
1	B	288	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13)

such sidechains are listed below:

Mol	Chain	Res	Type
1	A	33	GLN
1	A	62	ASN
1	A	154	GLN
1	A	156	GLN
1	A	281	GLN
1	A	357	ASN
1	B	23	HIS
1	B	62	ASN
1	B	154	GLN
1	B	156	GLN
1	B	168	ASN
1	B	281	GLN
1	B	357	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)

2 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	$\#\|Z\|>2$
2	IWT	A	1385	-	7,13,13	0.32	0	10,17,17	0.94	1 (10%)
2	IWT	B	1385	-	7,13,13	0.24	0	10,17,17	0.87	1 (10%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	IWT	A	1385	-	-	0/2/4/4	0/2/2/2
2	IWT	B	1385	-	-	0/2/4/4	0/2/2/2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1385	IWT	C7-C5-N4	2.91	129.05	124.12
2	B	1385	IWT	C7-C5-N4	2.63	128.57	124.12

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1385	IWT	1	0
2	B	1385	IWT	1	0

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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.