wwPDB X-ray Structure Validation Summary Report (i)

Nov 8, 2022 – 04:47 pm GMT

PDB ID	:	7QY5
Title	:	Crystal structure of the S.pombe Ars2-Red1 complex.
Authors	:	Foucher, A.E.; Kadlec, J.
Deposited on	:	2022-01-27
Resolution	:	2.77 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:

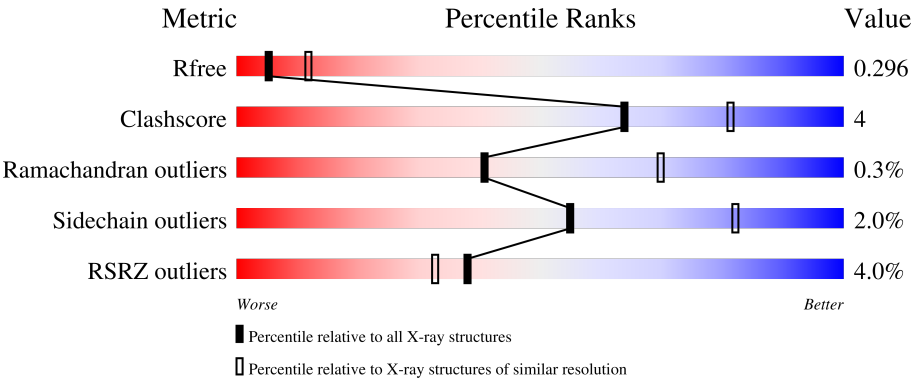
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 2.77 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.



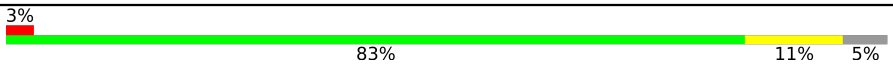
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4107 (2.80-2.76)
Clashscore	141614	4575 (2.80-2.76)
Ramachandran outliers	138981	4487 (2.80-2.76)
Sidechain outliers	138945	4489 (2.80-2.76)
RSRZ outliers	127900	4027 (2.80-2.76)

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	120	4% 88% 7% 5%
1	C	120	3% 78% 14% 8%
2	F	21	10% 29% 71%
2	G	21	14% 19% 14% 67%
3	B	338	3% 82% 12% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	338	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7174 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 NURS complex subunit pir2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	114	Total	C	N	O	S	0	0	0
			969	611	161	195	2			
1	C	111	Total	C	N	O	S	0	0	0
			947	596	158	191	2			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	64	GLY	-	expression tag	UNP O94326
A	65	ALA	-	expression tag	UNP O94326
A	66	MET	-	expression tag	UNP O94326
A	67	GLU	-	expression tag	UNP O94326
C	64	GLY	-	expression tag	UNP O94326
C	65	ALA	-	expression tag	UNP O94326
C	66	MET	-	expression tag	UNP O94326
C	67	GLU	-	expression tag	UNP O94326

- Molecule 2 is a protein called RNA elimination defective protein Red1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	F	6	Total	C	N	O	0	0	0
			44	25	6	13			
2	G	7	Total	C	N	O	0	0	0
			53	31	8	14			

- Molecule 3 is a protein called NURS complex subunit pir2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	320	Total	C	N	O	S	0	0	0
			2578	1634	444	489	11			
3	D	320	Total	C	N	O	S	0	0	0
			2581	1634	443	493	11			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	203	MET	-	initiating methionine	UNP O94326
B	204	GLU	-	expression tag	UNP O94326
B	205	MET	-	expression tag	UNP O94326
B	531	SER	-	expression tag	UNP O94326
B	532	ALA	-	expression tag	UNP O94326
B	533	TRP	-	expression tag	UNP O94326
B	534	SER	-	expression tag	UNP O94326
B	535	HIS	-	expression tag	UNP O94326
B	536	PRO	-	expression tag	UNP O94326
B	537	GLN	-	expression tag	UNP O94326
B	538	PHE	-	expression tag	UNP O94326
B	539	GLU	-	expression tag	UNP O94326
B	540	LYS	-	expression tag	UNP O94326
D	203	MET	-	initiating methionine	UNP O94326
D	204	GLU	-	expression tag	UNP O94326
D	205	MET	-	expression tag	UNP O94326
D	531	SER	-	expression tag	UNP O94326
D	532	ALA	-	expression tag	UNP O94326
D	533	TRP	-	expression tag	UNP O94326
D	534	SER	-	expression tag	UNP O94326
D	535	HIS	-	expression tag	UNP O94326
D	536	PRO	-	expression tag	UNP O94326
D	537	GLN	-	expression tag	UNP O94326
D	538	PHE	-	expression tag	UNP O94326
D	539	GLU	-	expression tag	UNP O94326
D	540	LYS	-	expression tag	UNP O94326

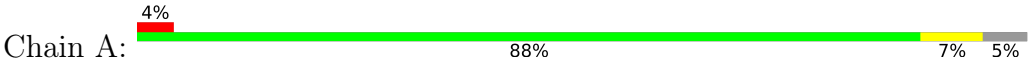
- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Zn 1 1	0	0
4	D	1	Total Zn 1 1	0	0

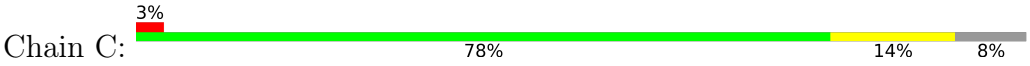
3 Residue-property plots (i)

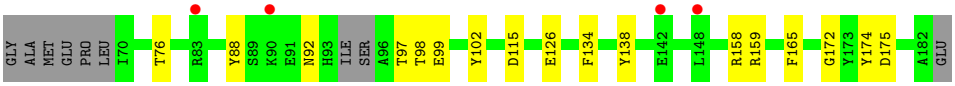
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: NURS complex subunit pir2

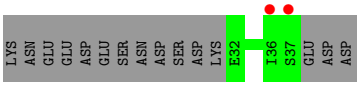


• Molecule 1: NURS complex subunit pir2

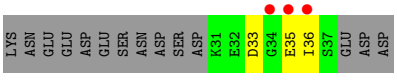




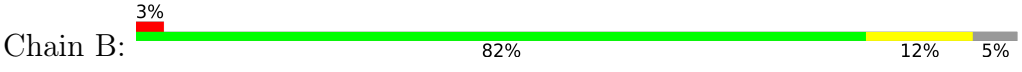
• Molecule 2: RNA elimination defective protein Red1

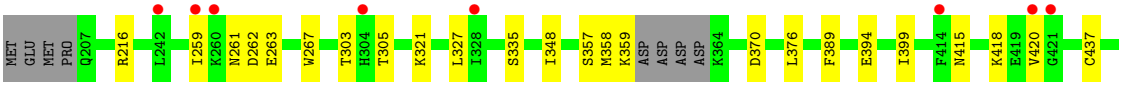


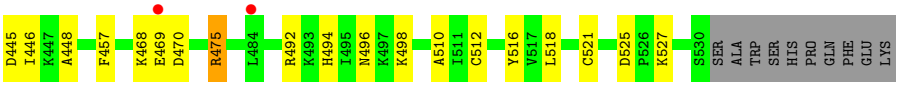
• Molecule 2: RNA elimination defective protein Red1



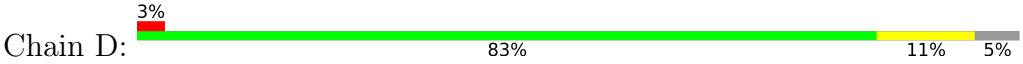
• Molecule 3: NURS complex subunit pir2

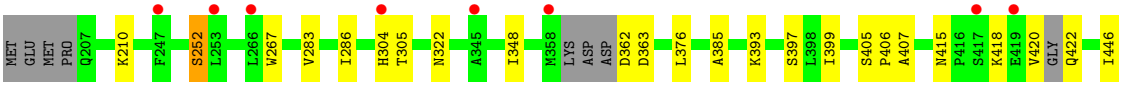


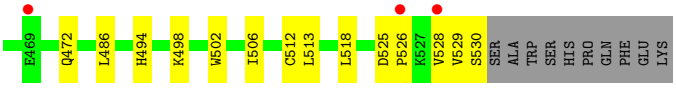




• Molecule 3: NURS complex subunit pir2







4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	64.67Å 128.28Å 89.26Å 90.00° 108.08° 90.00°	Depositor
Resolution (Å)	84.85 – 2.77 84.85 – 2.77	Depositor EDS
% Data completeness (in resolution range)	76.0 (84.85-2.77) 76.0 (84.85-2.77)	Depositor EDS
R_{merge}	0.29	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.00 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.228 , 0.298 0.234 , 0.296	Depositor DCC
R_{free} test set	1313 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å^2)	50.1	Xtriage
Anisotropy	0.053	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	7174	wwPDB-VP
Average B, all atoms (Å^2)	56.0	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 48.54 % 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 8.4798e-05. 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

5.1 Standard geometry

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

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	0.67	0/996	0.77	0/1352
1	C	0.64	0/973	0.75	0/1319
2	F	0.72	0/43	0.77	0/56
2	G	0.81	0/52	0.86	0/67
3	B	0.66	0/2634	0.79	0/3563
3	D	0.65	0/2636	0.78	0/3566
All	All	0.66	0/7334	0.78	0/9923

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	969	0	873	4	0
1	C	947	0	845	9	0
2	F	44	0	34	0	0
2	G	53	0	47	3	0
3	B	2578	0	2545	25	0
3	D	2581	0	2536	30	0
4	B	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	D	1	0	0	0	0
All	All	7174	0	6880	59	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 4.

The worst 5 of 59 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)
3:D:502:TRP:CZ2	3:D:506:ILE:HD11	2.00	0.96
3:D:502:TRP:CE2	3:D:506:ILE:HD11	2.05	0.91
2:G:36:ILE:HD13	3:D:486:LEU:HD13	1.62	0.81
2:G:36:ILE:HG21	3:D:486:LEU:HD13	1.63	0.80
3:D:305:THR:HG21	3:D:406:PRO:HG2	1.70	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	112/120 (93%)	104 (93%)	7 (6%)	1 (1%)	17 44
1	C	107/120 (89%)	99 (92%)	7 (6%)	1 (1%)	17 44
2	F	4/21 (19%)	4 (100%)	0	0	100 100
2	G	5/21 (24%)	4 (80%)	1 (20%)	0	100 100
3	B	316/338 (94%)	293 (93%)	23 (7%)	0	100 100
3	D	314/338 (93%)	289 (92%)	24 (8%)	1 (0%)	41 70
All	All	858/958 (90%)	793 (92%)	62 (7%)	3 (0%)	41 70

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	141	ASP
3	D	304	HIS
1	C	98	THR

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	103/107 (96%)	101 (98%)	2 (2%)	57 83
1	C	100/107 (94%)	96 (96%)	4 (4%)	31 62
2	F	5/20 (25%)	5 (100%)	0	100 100
2	G	6/20 (30%)	4 (67%)	2 (33%)	0 0
3	B	289/306 (94%)	286 (99%)	3 (1%)	76 91
3	D	290/306 (95%)	285 (98%)	5 (2%)	60 85
All	All	793/866 (92%)	777 (98%)	16 (2%)	55 82

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	405	SER
3	D	322	ASN
3	B	335	SER
3	D	252	SER
2	G	35	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	499	HIS
3	D	425	ASN
3	D	355	HIS
3	B	496	ASN
3	D	415	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

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

There are no such residues in this entry.

5.8 Polymer linkage issues

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	114/120 (95%)	0.63	5 (4%)	34	28	33, 49, 85, 109	0
1	C	111/120 (92%)	0.58	4 (3%)	42	37	32, 50, 99, 116	0
2	F	6/21 (28%)	1.46	2 (33%)	0	0	70, 83, 90, 91	0
2	G	7/21 (33%)	2.11	3 (42%)	0	0	85, 98, 103, 113	0
3	B	320/338 (94%)	0.55	10 (3%)	49	44	28, 51, 97, 120	0
3	D	320/338 (94%)	0.60	11 (3%)	45	39	27, 49, 103, 127	0
All	All	878/958 (91%)	0.60	35 (3%)	38	33	27, 50, 101, 127	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	304	HIS	5.7
3	B	420	VAL	5.4
2	G	36	ILE	4.7
1	C	83	ARG	4.5
3	B	259	ILE	3.7

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

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, 95^{th} 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
4	ZN	B	601	1/1	0.99	0.17	70,70,70,70	0
4	ZN	D	601	1/1	0.99	0.19	64,64,64,64	0

6.5 Other polymers

There are no such residues in this entry.