Full wwPDB X-ray Structure Validation Report (i)

Apr 29, 2025 – 07:28 AM EDT

PDB ID	:	1QYR / pdb_00001qyr
Title	:	2.1 Angstrom Crystal structure of KsgA: A Universally Conserved Adenosine Dimethyltransferase
Authors	:	O'Farrell, H.C.; Scarsdale, J.N.; Wright, H.T.; Rife, J.P.
Deposited on	:	2003-09-11
Resolution	:	2.10 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 (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:

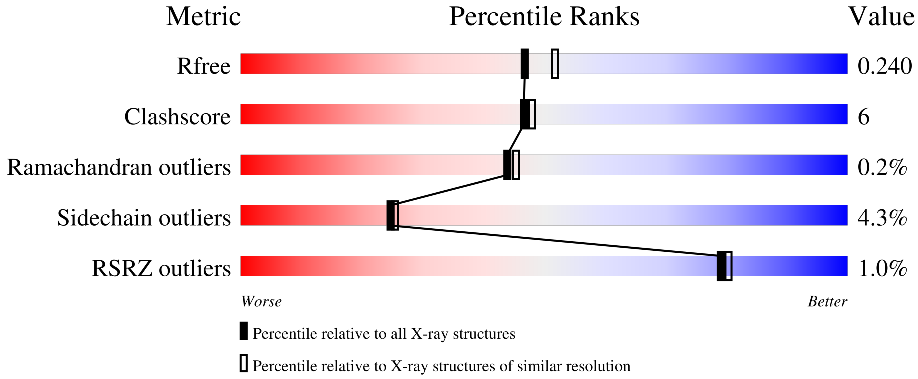
MolProbity	:	4-5-2 with Phenix2.0rc1
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

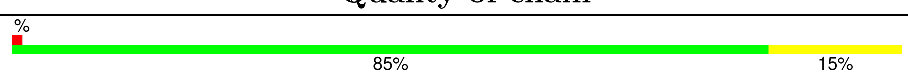
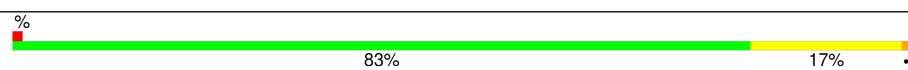
The reported resolution of this entry is 2.10 Å.

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}	164625	6234 (2.10-2.10)
Clashscore	180529	6893 (2.10-2.10)
Ramachandran outliers	177936	6839 (2.10-2.10)
Sidechain outliers	177891	6840 (2.10-2.10)
RSRZ outliers	164620	6234 (2.10-2.10)

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	252	 % 85% 15%
1	B	252	 % 83% 17% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4112 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 High level Kasugamycin resistance protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	252	1949	1241	336	359	13	0	0	0
1	B	252	1927	1231	328	355	13	0	0	0

- Molecule 2 is water.

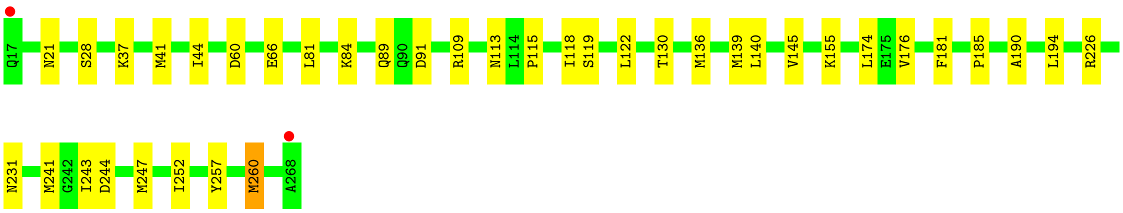
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	139	139	139	0	0
2	B	97	97	97	0	0

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.

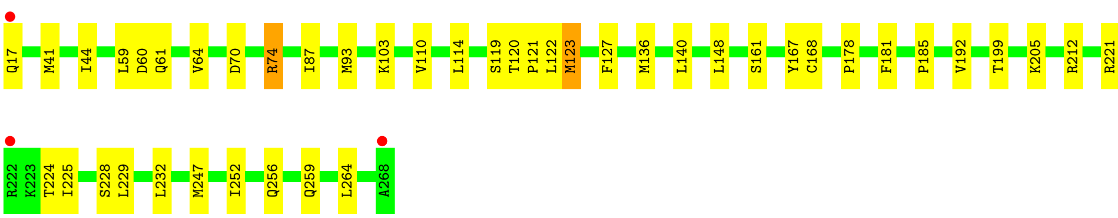
- Molecule 1: High level Kasugamycin resistance protein

Chain A:



- Molecule 1: High level Kasugamycin resistance protein

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	173.91Å 38.43Å 82.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.10 40.00 – 2.10	Depositor EDS
% Data completeness (in resolution range)	96.5 (40.00-2.10) 96.5 (40.00-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	3.16 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.1.24, CNS 1.0	Depositor
R, R_{free}	0.189 , 0.238 0.189 , 0.240	Depositor DCC
R_{free} test set	3159 reflections (10.02%)	wwPDB-VP
Wilson B-factor (Å^2)	32.4	Xtriage
Anisotropy	0.420	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 33.1	EDS
L-test for twinning[2]	$< \|L\| >$ = 0.49, $< L^2 >$ = 0.32	Xtriage
Estimated twinning fraction	0.020 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4112	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.63% 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 (i)

5.1 Standard geometry (i)

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	0.77	1/1991 (0.1%)	0.96	0/2710
1	B	0.87	1/1969 (0.1%)	0.95	3/2683 (0.1%)
All	All	0.82	2/3960 (0.1%)	0.96	3/5393 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	123	MET	SD-CE	-12.58	1.48	1.79
1	A	260	MET	SD-CE	-6.01	1.64	1.79

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	114	LEU	N-CA-C	5.63	116.76	109.72
1	B	60	ASP	N-CA-C	-5.58	106.83	113.97
1	B	168	CYS	N-CA-C	5.12	117.16	109.23

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	1949	0	1953	20	0
1	B	1927	0	1919	23	0
2	A	139	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	97	0	0	1	0
All	All	4112	0	3872	43	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 6.

All (43) 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:140:LEU:HD12	1:B:192:VAL:HG11	1.65	0.78
1:A:247:MET:HE3	1:A:252:ILE:HD13	1.70	0.73
1:B:212:ARG:HH12	1:B:232:LEU:HD21	1.63	0.63
1:A:44:ILE:HD13	1:A:122:LEU:HD21	1.83	0.61
1:A:181:PHE:HB2	1:A:185:PRO:HD3	1.83	0.60
1:A:247:MET:HE3	1:A:252:ILE:CD1	2.32	0.59
1:A:115:PRO:HD2	1:A:118:ILE:HD11	1.84	0.57
1:B:181:PHE:HB2	1:B:185:PRO:HD3	1.85	0.57
1:B:140:LEU:CD1	1:B:192:VAL:HG11	2.35	0.56
1:B:41:MET:HG3	1:B:59:LEU:HD11	1.88	0.55
1:A:226:ARG:HD2	2:A:341:HOH:O	2.06	0.54
1:A:241:MET:HE1	1:A:260:MET:HG2	1.90	0.54
1:B:225:ILE:HD12	1:B:229:LEU:HD12	1.89	0.54
1:B:123:MET:HE1	1:B:148:LEU:HD11	1.89	0.53
1:A:241:MET:HE3	1:A:243:ILE:HD12	1.92	0.52
1:A:41:MET:HE3	1:A:109:ARG:HB2	1.92	0.51
1:A:28:SER:OG	1:A:174:LEU:HD11	2.11	0.50
1:A:41:MET:HE3	1:A:109:ARG:CB	2.41	0.50
1:B:178:PRO:HA	1:B:185:PRO:HD2	1.94	0.50
1:B:123:MET:HE3	1:B:148:LEU:HD21	1.95	0.48
1:A:244:ASP:HB3	1:A:247:MET:HG3	1.94	0.48
1:B:225:ILE:HG23	1:B:229:LEU:HD12	1.96	0.48
1:A:60:ASP:O	1:A:84:LYS:HG2	2.14	0.48
1:B:64:VAL:HG22	1:B:87:ILE:HG12	1.96	0.47
1:A:113:ASN:HA	1:A:139:MET:HB3	1.96	0.47
1:A:119:SER:HB2	1:A:140:LEU:HD22	1.97	0.47
1:B:44:ILE:HD13	1:B:122:LEU:HD21	1.97	0.47
1:A:136:MET:HB2	1:A:194:LEU:HB2	1.96	0.46
1:B:225:ILE:HA	1:B:228:SER:HB3	1.99	0.45
1:A:139:MET:HE1	1:A:176:VAL:HG21	1.99	0.45
1:B:70:ASP:O	1:B:74:ARG:HG3	2.17	0.44
1:B:212:ARG:HH12	1:B:232:LEU:CD2	2.27	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:145:VAL:HG21	1:A:190:ALA:HB3	2.00	0.43
1:B:120:THR:HB	1:B:121:PRO:HD3	1.99	0.43
1:B:256:GLN:HA	1:B:259:GLN:HG2	2.00	0.43
1:B:17:GLN:HA	2:B:362:HOH:O	2.18	0.42
1:A:66:GLU:O	1:A:89:GLN:HA	2.19	0.42
1:B:119:SER:O	1:B:123:MET:HG2	2.20	0.42
1:B:123:MET:CE	1:B:148:LEU:HD11	2.48	0.42
1:B:127:PHE:HB3	1:B:167:TYR:CD1	2.55	0.41
1:B:110:VAL:O	1:B:136:MET:HA	2.20	0.41
1:B:140:LEU:HD12	1:B:192:VAL:CG1	2.44	0.41
1:A:257:TYR:HA	1:A:260:MET:HE3	2.03	0.40

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	250/252 (99%)	247 (99%)	2 (1%)	1 (0%)	30	29
1	B	250/252 (99%)	246 (98%)	4 (2%)	0	100	100
All	All	500/504 (99%)	493 (99%)	6 (1%)	1 (0%)	44	45

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	130	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	213/216 (99%)	207 (97%)	6 (3%)	38	43
1	B	208/216 (96%)	196 (94%)	12 (6%)	17	15
All	All	421/432 (98%)	403 (96%)	18 (4%)	25	25

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	21	ASN
1	A	37	LYS
1	A	81	LEU
1	A	91	ASP
1	A	155	LYS
1	A	231	ASN
1	B	61	GLN
1	B	74	ARG
1	B	93	MET
1	B	103	LYS
1	B	161	SER
1	B	199	THR
1	B	205	LYS
1	B	221	ARG
1	B	224	THR
1	B	247	MET
1	B	252	ILE
1	B	264	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	36	GLN
1	A	89	GLN
1	A	165	GLN
1	B	227	ASN

5.3.3 RNA ⓘ

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 oligosaccharides 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, 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	252/252 (100%)	-0.32	2 (0%)	82	83	18, 25, 37, 49	0
1	B	252/252 (100%)	-0.15	3 (1%)	76	77	16, 26, 41, 50	0
All	All	504/504 (100%)	-0.23	5 (0%)	79	80	16, 25, 39, 50	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	268	ALA	3.4
1	A	268	ALA	2.8
1	B	222	ARG	2.5
1	A	17	GLN	2.1
1	B	17	GLN	2.0

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.