



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 06:02 PM UTC

PDB ID : 4MKP / pdb_00004mkp
Title : Crystal structure of human cGAS apo form
Authors : Kato, K.; Ishii, R.; Ishitani, R.; Nureki, O.
Deposited on : 2013-09-05
Resolution : 1.95 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

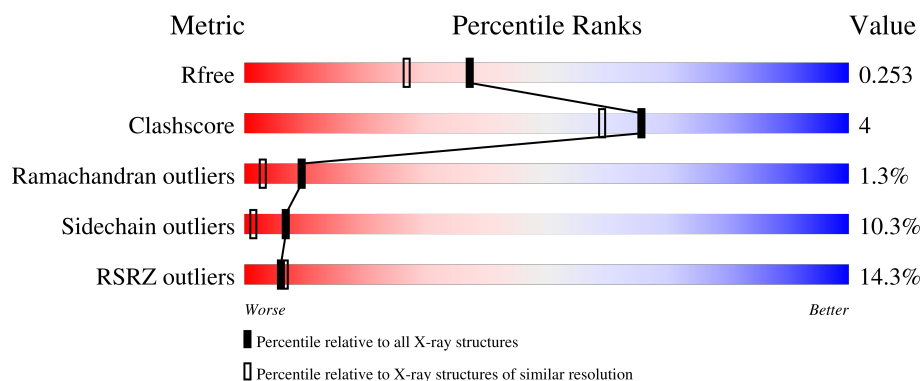
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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}	180053	3494 (1.96-1.96)
Clashscore	190562	3612 (1.96-1.96)
Ramachandran outliers	187476	3587 (1.96-1.96)
Sidechain outliers	187428	3587 (1.96-1.96)
RSRZ outliers	180081	3495 (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 $\leq 5\%$. 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	364	13% 74% 12% • 12%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2660 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 Cyclic GMP-AMP synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	322	Total	C	N	O	S	0	5	0
			2545	1642	426	462	15			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	159	GLY	-	expression tag	UNP Q8N884
A	160	SER	-	expression tag	UNP Q8N884

- Molecule 2 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		

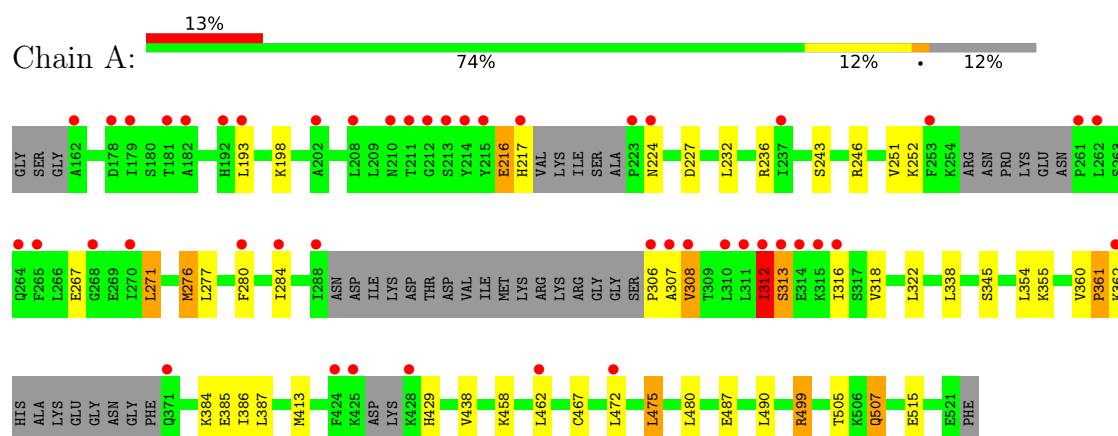
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	114	Total	O	0	0
			114	114		

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: Cyclic GMP-AMP synthase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	123.55Å 48.31Å 59.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.88 – 1.95 42.88 – 1.95	Depositor EDS
% Data completeness (in resolution range)	97.6 (42.88-1.95) 92.9 (42.88-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.32 (at 1.95Å)	Xtriage
Refinement program	PHENIX 1.7.2 _869	Depositor
R, R_{free}	0.204 , 0.252 0.207 , 0.253	Depositor DCC
R_{free} test set	1311 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	30.7	Xtriage
Anisotropy	0.310	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 63.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	2660	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.24% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.64	0/2609	0.86	3/3514 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	312	ILE	N-CA-C	6.95	118.13	108.12
1	A	216	GLU	N-CA-C	5.60	119.96	113.18
1	A	313	SER	N-CA-C	-5.50	101.89	110.42

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	312	ILE	Peptide

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	2545	0	2458	19	0
2	A	1	0	0	0	0
3	A	114	0	0	0	0
All	All	2660	0	2458	19	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.

All (19) 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:277:LEU:HD22	1:A:308:VAL:HG13	1.83	0.60
1:A:280:PHE:HD2	1:A:308:VAL:HG11	1.67	0.59
1:A:413[A]:MET:SD	1:A:467:CYS:HB3	2.43	0.57
1:A:413[A]:MET:HB3	1:A:438:VAL:HG13	1.91	0.53
1:A:246:ARG:HG3	1:A:487:GLU:HB2	1.90	0.53
1:A:306:PRO:HG3	1:A:360:VAL:HB	1.93	0.51
1:A:271:LEU:HD21	1:A:276:MET:HE3	1.94	0.50
1:A:216:GLU:HG2	1:A:384:LYS:HA	1.95	0.49
1:A:306:PRO:HD2	1:A:361:PRO:HG2	1.95	0.48
1:A:355:LYS:HE3	1:A:385:GLU:OE2	2.13	0.48
1:A:475:LEU:HB3	1:A:505:THR:OG1	2.15	0.47
1:A:193:LEU:HD13	1:A:284:ILE:HG12	1.96	0.46
1:A:216:GLU:O	1:A:217:HIS:HB2	2.17	0.45
1:A:277:LEU:HD22	1:A:308:VAL:CG1	2.49	0.43
1:A:306:PRO:HA	1:A:307:ALA:HA	1.66	0.43
1:A:429:HIS:HB3	1:A:499:ARG:HH12	1.83	0.42
1:A:507:GLN:H	1:A:507:GLN:HG2	1.43	0.42
1:A:338:LEU:HD11	1:A:386[A]:ILE:HD11	2.02	0.41
1:A:413[A]:MET:HE2	1:A:438:VAL:HG13	2.04	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	315/364 (86%)	300 (95%)	11 (4%)	4 (1%)	9	3

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	345	SER
1	A	236	ARG
1	A	361	PRO
1	A	308	VAL

5.3.2 Protein sidechains ⓘ

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	266/335 (79%)	239 (90%)	27 (10%)	7	1

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	198	LYS
1	A	224	ASN
1	A	227	ASP
1	A	232	LEU
1	A	243	SER
1	A	251	VAL
1	A	252	LYS
1	A	267	GLU
1	A	271	LEU
1	A	276	MET
1	A	312	ILE
1	A	313	SER
1	A	316	ILE
1	A	318	VAL
1	A	322	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	354	LEU
1	A	362	LYS
1	A	387	LEU
1	A	458	LYS
1	A	462	LEU
1	A	472	LEU
1	A	475	LEU
1	A	480	LEU
1	A	490	LEU
1	A	499	ARG
1	A	507	GLN
1	A	515	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 1 ligands modelled in this entry, 1 is 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 ⓘ

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, 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	322/364 (88%)	0.77	46 (14%) 6 7	17, 49, 88, 107	5 (1%)

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	214	TYR	5.3
1	A	223	PRO	5.1
1	A	211	THR	5.0
1	A	261	PRO	4.3
1	A	308	VAL	4.1
1	A	217	HIS	4.0
1	A	313	SER	3.8
1	A	312	ILE	3.7
1	A	265	PHE	3.7
1	A	213	SER	3.7
1	A	210	ASN	3.3
1	A	288	ILE	3.3
1	A	253	PHE	3.2
1	A	316	ILE	3.2
1	A	310	LEU	3.1
1	A	424	PHE	3.1
1	A	428	LYS	3.0
1	A	425	LYS	2.9
1	A	179	ILE	2.8
1	A	212	GLY	2.8
1	A	193	LEU	2.7
1	A	280	PHE	2.6
1	A	462	LEU	2.6
1	A	371	GLN	2.5
1	A	262	LEU	2.5
1	A	268	GLY	2.5
1	A	178	ASP	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	472	LEU	2.4
1	A	314	GLU	2.4
1	A	311	LEU	2.4
1	A	215	TYR	2.4
1	A	264	GLN	2.3
1	A	270	ILE	2.3
1	A	284	ILE	2.3
1	A	192	HIS	2.3
1	A	237	ILE	2.3
1	A	162	ALA	2.2
1	A	315	LYS	2.2
1	A	224	ASN	2.2
1	A	182	ALA	2.2
1	A	307	ALA	2.2
1	A	362	LYS	2.2
1	A	208	LEU	2.1
1	A	202	ALA	2.1
1	A	181	THR	2.1
1	A	306	PRO	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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
2	ZN	A	1000	1/1	1.00	0.01	26,26,26,26	0

6.5 Other polymers [i](#)

There are no such residues in this entry.