Full wwPDB X-ray Structure Validation Report (i)

Dec 16, 2024 – 03:30 pm GMT

PDB ID	:	9GVA
Title	:	Crystal structure of the gamma carbonic anhydrase from Porphyromonas gingivalis
Authors	:	Angeli, A.; Ferraroni, M.
Deposited on	:	2024-09-23
Resolution	:	2.40 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

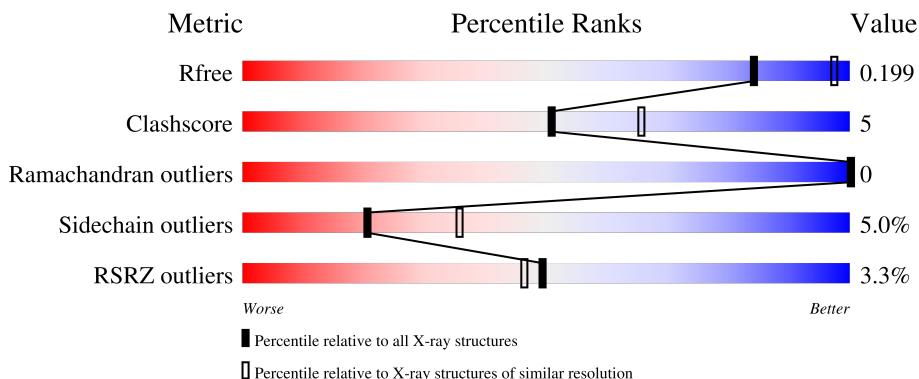|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 2.40 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	Percentile Ranks	Value
Rfree		0.199
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		5.0%
RSRZ outliers		3.3%

Worse ... *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	4642 (2.40-2.40)
Clashscore	180529	5218 (2.40-2.40)
Ramachandran outliers	177936	5158 (2.40-2.40)
Sidechain outliers	177891	5159 (2.40-2.40)
RSRZ outliers	164620	4642 (2.40-2.40)

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	AAA	192	5% / 80% / 8% ·· 10%
1	BBB	192	% / 81% / 7% · 11%
1	CCC	192	3% / 72% / 18% · 9%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4024 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 Hexapeptide transferase family protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	173	Total	C	N	O	S	0	0	0
			1312	831	228	246	7			
1	BBB	171	Total	C	N	O	S	0	0	0
			1295	822	226	240	7			
1	CCC	174	Total	C	N	O	S	0	0	0
			1318	834	229	248	7			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	1	Total	Zn	0	0
			1	1		
2	BBB	1	Total	Zn	0	0
			1	1		
2	CCC	1	Total	Zn	0	0
			1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	24	Total	O	0	0
			24	24		
3	BBB	44	Total	O	0	0
			44	44		
3	CCC	28	Total	O	0	0
			28	28		

3 Residue-property plots ⓘ

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.
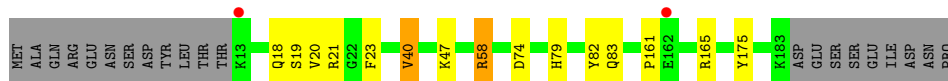
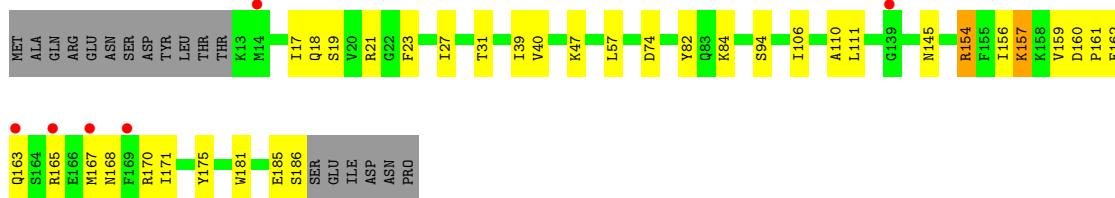- Molecule 1: Hexapeptide transferase family protein

Chain AAA:

- Molecule 1: Hexapeptide transferase family protein

Chain BBB:

- Molecule 1: Hexapeptide transferase family protein

Chain CCC:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	120.75Å 120.75Å 91.31Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.37 – 2.40 45.37 – 2.40	Depositor EDS
% Data completeness (in resolution range)	95.9 (45.37-2.40) 95.9 (45.37-2.40)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.83 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.165 , 0.197 0.175 , 0.199	Depositor DCC
R_{free} test set	1493 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	35.3	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 37.0	EDS
L-test for twinning[2]	$< |L| > = 0.52, < L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.027 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4024	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.78% 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: 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	AAA	0.75	0/1335	1.01	4/1810 (0.2%)
1	BBB	0.79	0/1318	0.93	2/1787 (0.1%)
1	CCC	0.75	0/1341	0.96	1/1818 (0.1%)
All	All	0.76	0/3994	0.97	7/5415 (0.1%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	AAA	58	ARG	NE-CZ-NH1	8.99	124.80	120.30
1	AAA	58	ARG	NE-CZ-NH2	-7.68	116.46	120.30
1	BBB	58	ARG	NE-CZ-NH2	-6.39	117.10	120.30
1	AAA	65	ARG	NE-CZ-NH1	5.82	123.21	120.30
1	BBB	58	ARG	NE-CZ-NH1	5.70	123.15	120.30
1	CCC	21	ARG	NE-CZ-NH1	5.20	122.90	120.30
1	AAA	21	ARG	NE-CZ-NH1	5.02	122.81	120.30

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	AAA	1312	0	1309	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	BBB	1295	0	1299	7	0
1	CCC	1318	0	1314	27	0
2	AAA	1	0	0	0	0
2	BBB	1	0	0	0	0
2	CCC	1	0	0	0	0
3	AAA	24	0	0	0	0
3	BBB	44	0	0	0	0
3	CCC	28	0	0	1	0
All	All	4024	0	3922	40	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 5.

All (40) 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:CCC:18:GLN:HB2	1:CCC:40:VAL:HG22	1.69	0.75
1:CCC:159:VAL:HB	1:CCC:163:GLN:HE21	1.63	0.63
1:CCC:163:GLN:O	1:CCC:167:MET:HB2	1.99	0.61
1:CCC:167:MET:SD	1:CCC:170:ARG:NH1	2.76	0.58
1:AAA:18:GLN:HB3	1:AAA:40:VAL:HG22	1.85	0.57
1:AAA:178:TYR:CD1	1:AAA:184:ASP:O	2.61	0.53
1:AAA:58:ARG:HG2	1:AAA:79:HIS:CD2	2.43	0.53
1:AAA:21:ARG:HH11	1:AAA:21:ARG:HB2	1.73	0.52
1:CCC:17:ILE:HA	1:CCC:39:ILE:O	2.09	0.52
1:CCC:94:SER:HB2	1:CCC:111:LEU:HD13	1.90	0.52
1:BBB:21:ARG:HD2	1:CCC:181:TRP:CE3	2.45	0.52
1:AAA:172:ALA:HA	1:AAA:177:MET:HE3	1.93	0.50
1:AAA:178:TYR:CE1	1:AAA:184:ASP:O	2.65	0.50
1:AAA:183:LYS:HE3	1:CCC:154:ARG:HD3	1.94	0.49
1:CCC:167:MET:HA	1:CCC:170:ARG:HH11	1.77	0.48
1:CCC:156:ILE:HG22	1:CCC:157:LYS:HD2	1.95	0.48
1:CCC:27:ILE:CG2	1:CCC:31:THR:HG21	2.43	0.48
1:BBB:58:ARG:HG2	1:BBB:79:HIS:CD2	2.48	0.48
1:CCC:163:GLN:O	1:CCC:167:MET:CB	2.61	0.48
1:BBB:21:ARG:HD2	1:CCC:181:TRP:CD2	2.49	0.47
1:CCC:145:ASN:O	1:CCC:159:VAL:HG22	2.15	0.47
1:CCC:159:VAL:HB	1:CCC:163:GLN:NE2	2.28	0.47
1:BBB:161:PRO:O	1:BBB:165:ARG:HG3	2.15	0.46
1:CCC:167:MET:O	1:CCC:171:ILE:HD12	2.15	0.46
1:CCC:19:SER:HA	1:CCC:23:PHE:O	2.16	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:BBB:82:TYR:O	1:BBB:83:GLN:HB2	2.16	0.46
1:BBB:19:SER:HA	1:BBB:23:PHE:O	2.16	0.46
1:CCC:168:ASN:HA	1:CCC:171:ILE:HD12	1.97	0.46
1:CCC:162:GLU:CG	3:CCC:316:HOH:O	2.64	0.45
1:CCC:186:SER:O	1:CCC:186:SER:OG	2.36	0.44
1:CCC:160:ASP:OD2	1:CCC:162:GLU:HB2	2.18	0.43
1:BBB:18:GLN:HB3	1:BBB:40:VAL:HG22	2.00	0.43
1:AAA:39:ILE:HD12	1:AAA:39:ILE:N	2.34	0.43
1:AAA:160:ASP:OD1	1:AAA:162:GLU:N	2.51	0.42
1:CCC:163:GLN:O	1:CCC:167:MET:CG	2.67	0.42
1:CCC:57:LEU:N	1:CCC:57:LEU:HD12	2.34	0.42
1:CCC:106:ILE:HG22	1:CCC:110:ALA:HB3	2.02	0.41
1:CCC:111:LEU:HD22	1:CCC:171:ILE:HD13	2.01	0.41
1:CCC:82:TYR:O	1:CCC:84:LYS:HD2	2.21	0.40
1:CCC:161:PRO:O	1:CCC:165:ARG:HG2	2.22	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	AAA	171/192 (89%)	163 (95%)	8 (5%)	0	100	100
1	BBB	169/192 (88%)	168 (99%)	1 (1%)	0	100	100
1	CCC	172/192 (90%)	164 (95%)	8 (5%)	0	100	100
All	All	512/576 (89%)	495 (97%)	17 (3%)	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	AAA	140/158 (89%)	130 (93%)	10 (7%)	12	20
1	BBB	138/158 (87%)	133 (96%)	5 (4%)	30	49
1	CCC	141/158 (89%)	135 (96%)	6 (4%)	25	42
All	All	419/474 (88%)	398 (95%)	21 (5%)	20	36

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	14	MET
1	AAA	20	VAL
1	AAA	21	ARG
1	AAA	35	GLU
1	AAA	40	VAL
1	AAA	159	VAL
1	AAA	165	ARG
1	AAA	177	MET
1	AAA	182	PHE
1	AAA	183	LYS
1	BBB	20	VAL
1	BBB	40	VAL
1	BBB	47	LYS
1	BBB	74	ASP
1	BBB	175	TYR
1	CCC	47	LYS
1	CCC	74	ASP
1	CCC	154	ARG
1	CCC	157	LYS
1	CCC	175	TYR
1	CCC	185	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 3 ligands modelled in this entry, 3 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, 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	AAA	173/192 (90%)	-0.17	9 (5%)	34	31	24, 35, 87, 134	0
1	BBB	171/192 (89%)	-0.60	2 (1%)	76	73	24, 31, 57, 85	0
1	CCC	174/192 (90%)	-0.17	6 (3%)	48	45	26, 38, 81, 104	0
All	All	518/576 (89%)	-0.31	17 (3%)	49	46	24, 35, 74, 134	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	181	TRP	6.2
1	AAA	184	ASP	6.0
1	AAA	183	LYS	3.9
1	CCC	139	GLY	3.4
1	CCC	167	MET	3.2
1	CCC	169	PHE	3.2
1	AAA	179	ALA	3.1
1	AAA	13	LYS	2.9
1	AAA	175	TYR	2.7
1	AAA	182	PHE	2.7
1	CCC	14	MET	2.1
1	CCC	165	ARG	2.1
1	BBB	162	GLU	2.1
1	AAA	167	MET	2.1
1	AAA	185	GLU	2.1
1	CCC	163	GLN	2.1
1	BBB	13	LYS	2.0

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 (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
2	ZN	AAA	201	1/1	1.00	0.01	28,28,28,28	0
2	ZN	BBB	201	1/1	1.00	0.01	29,29,29,29	0
2	ZN	CCC	201	1/1	1.00	0.01	33,33,33,33	0

6.5 Other polymers (i)

There are no such residues in this entry.