



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 10, 2025 – 07:13 pm BST

PDB ID : 9F82 / pdb\_00009f82  
Title : Arbitrium receptor from ATCC13952 phage in complex with GVVRGA peptide  
Authors : Gallego del Sol, F.; Marina, A.  
Deposited on : 2024-05-06  
Resolution : 3.10 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](mailto: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 Phenix2.0rc1
Xtriage (Phenix)	:	2.0rc1
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)	:	2.43.1

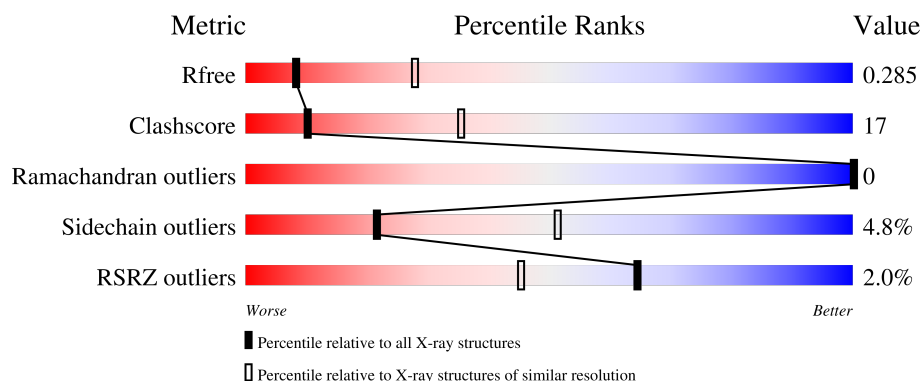
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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	1351 (3.10-3.10)
Clashscore	180529	1454 (3.10-3.10)
Ramachandran outliers	177936	1391 (3.10-3.10)
Sidechain outliers	177891	1391 (3.10-3.10)
RSRZ outliers	164620	1351 (3.10-3.10)

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  $\geq 3$ , 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	386	<div> <div>2%</div> <div>64% 20% 12%</div> </div>
1	C	386	<div> <div>%</div> <div>56% 26% 5% 12%</div> </div>
1	E	386	<div> <div>%</div> <div>64% 20% 12%</div> </div>
1	G	386	<div> <div>3%</div> <div>55% 29% 12%</div> </div>
2	B	6	<div> <div>50% 33% 17%</div> </div>

Continued on next page...

Mol	Chain	Length	Quality of chain
2	D	6	<p>17% 67% 17% 17%</p>
2	F	6	<p>50% 17% 33%</p>
2	K	6	<p>17% 50% 50%</p>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11488 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 Arbitrium receptor from ATCC13952 phage.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	341	Total	C	N	O	S	0	0	0
			2827	1822	457	539	9			
1	C	341	Total	C	N	O	S	0	0	0
			2827	1822	457	539	9			
1	E	341	Total	C	N	O	S	0	0	0
			2827	1822	457	539	9			
1	G	341	Total	C	N	O	S	0	0	0
			2827	1822	457	539	9			

- Molecule 2 is a protein called GLY-VAL-VAL-ARG-GLY-ALA.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	6	Total	C	N	O	0	0	0
			39	23	9	7			
2	D	6	Total	C	N	O	0	0	0
			39	23	9	7			
2	F	6	Total	C	N	O	0	0	0
			39	23	9	7			
2	K	6	Total	C	N	O	0	0	0
			39	23	9	7			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	5	Total	O	0	0
			5	5		
3	C	3	Total	O	0	0
			3	3		
3	E	6	Total	O	0	0
			6	6		
3	G	10	Total	O	0	0
			10	10		

i

- Molecule 1: Arbitrium receptor from ATCC13952 phage





- Molecule 2: GLY-VAL-VAL-ARG-GLY-ALA

Chain F:  50% 17% 33%



- Molecule 2: GLY-VAL-VAL-ARG-GLY-ALA

Chain K:  17% 50% 50%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.34Å 82.96Å 146.86Å 90.00° 90.88° 90.00°	Depositor
Resolution (Å)	146.84 – 3.10 146.84 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (146.84-3.10) 99.6 (146.84-3.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.22 (at 3.07Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, $R_{free}$	0.228 , 0.277 0.237 , 0.285	Depositor DCC
$R_{free}$ test set	1523 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	75.0	Xtriage
Anisotropy	0.604	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 32.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.027 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	11488	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	80.0	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 31.67 % 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 1.0654e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [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	0.90	0/2881	1.50	23/3872 (0.6%)
1	C	0.87	0/2881	1.59	54/3872 (1.4%)
1	E	0.77	0/2881	1.54	32/3872 (0.8%)
1	G	0.77	0/2881	1.57	41/3872 (1.1%)
2	B	1.21	0/38	1.75	1/48 (2.1%)
2	D	1.05	0/38	2.38	1/48 (2.1%)
2	F	1.15	0/38	1.80	1/48 (2.1%)
2	K	1.11	0/38	1.80	0/48
All	All	0.83	0/11676	1.56	153/15680 (1.0%)

There are no bond length outliers.

The worst 5 of 153 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	170	MET	N-CA-C	12.36	124.29	111.07
1	E	58	HIS	CA-CB-CG	-10.54	103.25	113.80
1	G	51	ILE	CB-CA-C	-10.45	98.59	111.97
1	G	121	ILE	N-CA-C	-10.28	100.76	110.42
1	G	121	ILE	N-CA-CB	9.76	121.68	110.65

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	2827	0	2846	77	3
1	C	2827	0	2843	118	0
1	E	2827	0	2845	84	5
1	G	2827	0	2846	135	2
2	B	39	0	44	9	0
2	D	39	0	44	4	0
2	F	39	0	44	3	0
2	K	39	0	44	7	0
3	A	5	0	0	0	0
3	C	3	0	0	1	0
3	E	6	0	0	1	0
3	G	10	0	0	3	0
All	All	11488	0	11556	390	5

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 17.

The worst 5 of 390 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:G:279:ASN:HB3	1:G:303:PHE:CD2	1.34	1.62
1:G:51:ILE:HD13	1:G:70:TYR:CE2	1.28	1.59
1:G:51:ILE:CD1	1:G:70:TYR:CZ	1.79	1.58
1:G:51:ILE:CD1	1:G:70:TYR:CE2	1.82	1.51
1:G:279:ASN:CB	1:G:303:PHE:HD2	1.46	1.28

All (5) 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:E:59:TYR:CZ	1:G:47:GLU:OE2[1_656]	1.61	0.59
1:E:59:TYR:OH	1:G:47:GLU:OE2[1_656]	1.69	0.51
1:A:350:LYS:NZ	1:E:322:GLU:O[2_646]	1.79	0.41
1:A:323:ARG:NH2	1:E:328:GLU:OE1[2_646]	2.10	0.10
1:A:323:ARG:CZ	1:E:328:GLU:OE1[2_646]	2.14	0.06

## 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	339/386 (88%)	330 (97%)	9 (3%)	0	100	100
1	C	339/386 (88%)	328 (97%)	11 (3%)	0	100	100
1	E	339/386 (88%)	330 (97%)	9 (3%)	0	100	100
1	G	339/386 (88%)	327 (96%)	12 (4%)	0	100	100
2	B	4/6 (67%)	4 (100%)	0	0	100	100
2	D	4/6 (67%)	4 (100%)	0	0	100	100
2	F	4/6 (67%)	4 (100%)	0	0	100	100
2	K	4/6 (67%)	3 (75%)	1 (25%)	0	100	100
All	All	1372/1568 (88%)	1330 (97%)	42 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	317/356 (89%)	306 (96%)	11 (4%)	31	61
1	C	317/356 (89%)	302 (95%)	15 (5%)	22	52
1	E	317/356 (89%)	304 (96%)	13 (4%)	26	57
1	G	317/356 (89%)	298 (94%)	19 (6%)	16	44
2	B	3/3 (100%)	2 (67%)	1 (33%)	0	0
2	D	3/3 (100%)	2 (67%)	1 (33%)	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	F	3/3 (100%)	2 (67%)	1 (33%)	0	0
2	K	3/3 (100%)	3 (100%)	0	100	100
All	All	1280/1436 (89%)	1219 (95%)	61 (5%)	21	51

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	51	ILE
1	G	196	ILE
1	E	211	GLU
1	G	182	ILE
1	G	328	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	289	ASN
1	G	329	ASN
1	C	326	GLN
1	C	377	ASN
1	E	58	HIS

### 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 ⓘ

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	341/386 (88%)	-0.00	6 (1%) 67 49	38, 72, 122, 167	0
1	C	341/386 (88%)	-0.04	5 (1%) 71 54	38, 72, 119, 154	0
1	E	341/386 (88%)	0.05	5 (1%) 71 54	36, 77, 117, 153	0
1	G	341/386 (88%)	0.25	10 (2%) 54 34	48, 86, 132, 157	0
2	B	6/6 (100%)	0.29	0 100 100	44, 62, 74, 101	0
2	D	6/6 (100%)	0.75	1 (16%) 5 3	42, 61, 81, 118	0
2	F	6/6 (100%)	0.62	0 100 100	47, 51, 71, 73	0
2	K	6/6 (100%)	0.93	1 (16%) 5 3	45, 59, 75, 156	0
All	All	1388/1568 (88%)	0.08	28 (2%) 64 45	36, 77, 123, 167	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	256	GLY	6.7
1	G	260	ALA	4.6
1	G	255	SER	4.3
1	E	258	GLU	4.0
1	A	89	ASP	3.7

### 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.