



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 03:35 PM GMT

PDB ID : 1CD3  
Title : PROCAPSID OF BACTERIOPHAGE PHIX174  
Authors : Rossmann, M.G.; Dokland, T.  
Deposited on : 1999-03-05  
Resolution : 3.50 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

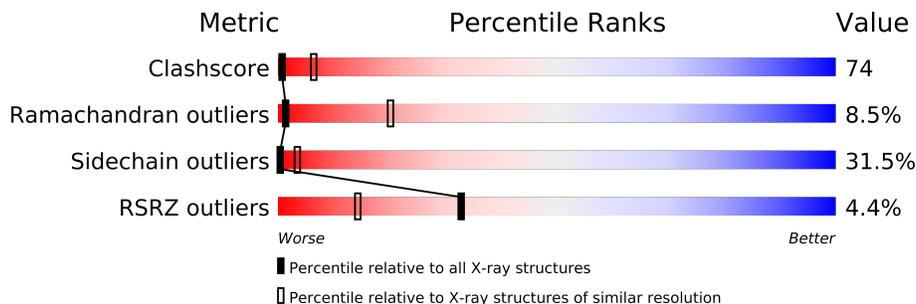
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.50 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))
Clashscore	79885	1039 (3.66-3.34)
Ramachandran outliers	78287	1000 (3.66-3.34)
Sidechain outliers	78261	1000 (3.66-3.34)
RSRZ outliers	66119	1243 (3.70-3.30)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	1	152	
1	2	152	
1	3	152	
1	4	152	
2	F	426	
3	G	175	
4	B	120	

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 9851 atoms, of which 0 are hydrogen and 0 are deuterium.

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 PROTEIN (SCAFFOLDING PROTEIN GPD).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	1	143	1125	716	194	211	4	0	0	0
1	2	135	1057	675	177	201	4	0	0	0
1	3	140	1099	699	187	209	4	0	0	0
1	4	146	1145	728	197	215	5	0	0	0

- Molecule 2 is a protein called PROTEIN (CAPSID PROTEIN GPF).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	F	426	3415	2173	590	638	14	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	216	ARG	HIS	CONFLICT	UNP P03641

- Molecule 3 is a protein called PROTEIN (SPIKE PROTEIN GPG).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	G	175	1340	856	221	255	8	0	0	0

- Molecule 4 is a protein called PROTEIN (SCAFFOLDING PROTEIN GPB).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	B	68	574	358	105	108	3	0	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	1	13	Total O 13 13	0	0
5	2	9	Total O 9 9	0	0
5	3	14	Total O 14 14	0	0
5	4	9	Total O 9 9	0	0
5	B	18	Total O 18 18	0	0
5	F	25	Total O 25 25	0	0
5	G	8	Total O 8 8	0	0

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

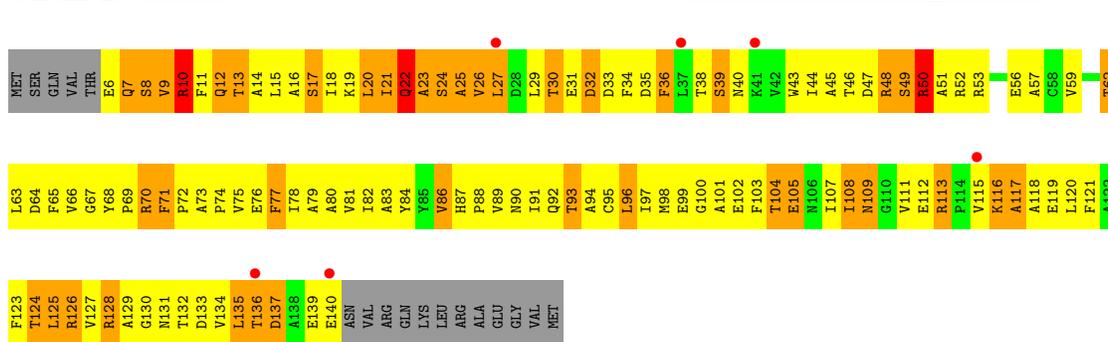
- Molecule 1: PROTEIN (SCAFFOLDING PROTEIN GPD)

Chain 1:



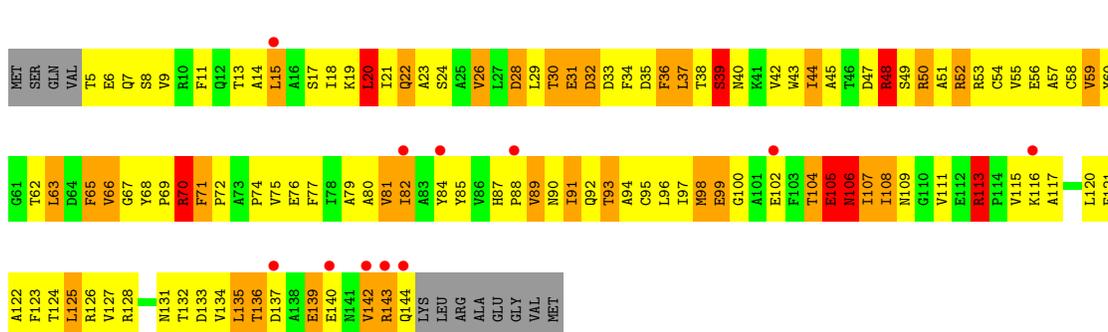
- Molecule 1: PROTEIN (SCAFFOLDING PROTEIN GPD)

Chain 2:



- Molecule 1: PROTEIN (SCAFFOLDING PROTEIN GPD)

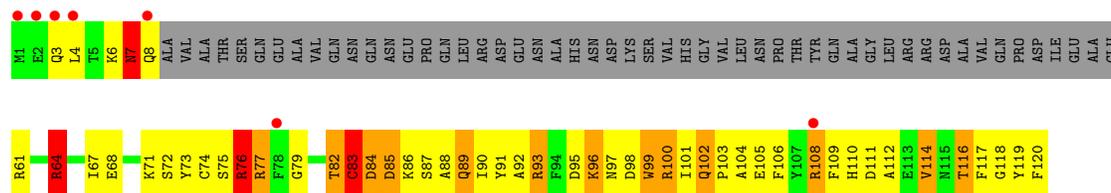
Chain 3:





- Molecule 4: PROTEIN (SCAFFOLDING PROTEIN GPB)

Chain B:



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	774.00Å 774.00Å 774.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 3.50 44.84 – 3.50	Depositor EDS
% Data completeness (in resolution range)	67.2 (8.00-3.50) 66.3 (44.84-3.50)	Depositor EDS
$R_{merge}$	0.22	Depositor
$R_{sym}$	0.22	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.93 (at 3.48Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.275 , (Not available) 0.437 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	78.2	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , -20.0	EDS
Estimated twinning fraction	0.023 for -l,-k,-h	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	6 of 631903 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.28	EDS
Total number of atoms	9851	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.37% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 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	1	0.79	0/1145	1.05	4/1557 (0.3%)
1	2	0.78	0/1077	1.05	4/1467 (0.3%)
1	3	0.87	1/1119 (0.1%)	1.10	6/1524 (0.4%)
1	4	0.79	0/1165	1.03	6/1582 (0.4%)
2	F	0.85	0/3511	1.05	11/4777 (0.2%)
3	G	0.75	0/1372	1.01	2/1872 (0.1%)
4	B	0.84	0/586	1.24	6/779 (0.8%)
All	All	0.82	1/9975 (0.0%)	1.06	39/13558 (0.3%)

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	3	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	3	50	ARG	CZ-NH1	6.50	1.41	1.33

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	64	ARG	NE-CZ-NH2	8.72	124.66	120.30
4	B	61	ARG	NE-CZ-NH2	8.37	124.48	120.30
1	1	50	ARG	NE-CZ-NH2	7.71	124.16	120.30
2	F	161	ARG	NE-CZ-NH2	7.63	124.12	120.30
1	3	70	ARG	NE-CZ-NH2	7.63	124.11	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	3	48	ARG	Sidechain
1	3	52	ARG	Sidechain

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	1	1125	0	1121	182	0
1	2	1057	0	1043	212	0
1	3	1099	0	1086	187	0
1	4	1145	0	1142	201	0
2	F	3415	0	3305	422	0
3	G	1340	0	1323	194	0
4	B	574	0	538	93	0
5	1	13	0	0	0	0
5	2	9	0	0	3	0
5	3	14	0	0	0	0
5	4	9	0	0	0	0
5	B	18	0	0	1	0
5	F	25	0	0	0	0
5	G	8	0	0	0	0
All	All	9851	0	9558	1426	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 74.

The worst 5 of 1426 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
4:B:64:ARG:O	4:B:67:ILE:HG22	1.27	1.31
2:F:23:GLN:NE2	2:F:401:GLN:HG3	1.57	1.19
1:2:71:PHE:CD2	1:2:72:PRO:HD3	1.78	1.19
3:G:62:MET:HG3	3:G:131:VAL:HG12	1.29	1.11
4:B:64:ARG:O	4:B:67:ILE:CG2	2.00	1.09

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	1	141/152 (93%)	96 (68%)	33 (23%)	12 (8%)	1	20
1	2	133/152 (88%)	81 (61%)	39 (29%)	13 (10%)	1	16
1	3	138/152 (91%)	97 (70%)	28 (20%)	13 (9%)	1	17
1	4	144/152 (95%)	89 (62%)	44 (31%)	11 (8%)	2	23
2	F	424/426 (100%)	314 (74%)	71 (17%)	39 (9%)	1	18
3	G	173/175 (99%)	143 (83%)	19 (11%)	11 (6%)	2	29
4	B	64/120 (53%)	48 (75%)	12 (19%)	4 (6%)	2	29
All	All	1217/1329 (92%)	868 (71%)	246 (20%)	103 (8%)	1	20

5 of 103 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1	139	GLU
1	2	8	SER
1	2	9	VAL
1	2	12	GLN
1	2	23	ALA

### 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	1	119/127 (94%)	83 (70%)	36 (30%)	0	4
1	2	112/127 (88%)	77 (69%)	35 (31%)	0	3
1	3	117/127 (92%)	75 (64%)	42 (36%)	0	2
1	4	121/127 (95%)	85 (70%)	36 (30%)	0	4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	F	372/372 (100%)	252 (68%)	120 (32%)	0	3
3	G	153/153 (100%)	106 (69%)	47 (31%)	0	4
4	B	58/101 (57%)	43 (74%)	15 (26%)	1	5
All	All	1052/1134 (93%)	721 (68%)	331 (32%)	0	3

5 of 331 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	2	ASN
2	F	166	LYS
3	G	151	THR
2	F	15	SER
2	F	86	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	80	GLN
2	F	372	GLN
4	B	89	GLN
2	F	270	HIS
2	F	387	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

## 5.6 Ligand geometry

There are no ligands in this entry.

## 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 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	1	143/152 (94%)	0.99	8 (5%) 24 10	3, 23, 72, 90	0
1	2	135/152 (88%)	1.04	6 (4%) 33 14	5, 24, 86, 100	0
1	3	140/152 (92%)	1.04	11 (7%) 13 7	4, 24, 87, 96	0
1	4	146/152 (96%)	0.96	8 (5%) 24 10	4, 26, 82, 100	0
2	F	426/426 (100%)	0.91	11 (2%) 53 24	2, 8, 61, 92	0
3	G	175/175 (100%)	0.85	3 (1%) 67 34	2, 15, 32, 44	0
4	B	68/120 (56%)	1.03	7 (10%) 7 5	7, 44, 79, 98	0
All	All	1233/1329 (92%)	0.95	54 (4%) 33 14	2, 18, 76, 100	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	3	144	GLN	6.6
1	4	108	ILE	4.5
1	3	116	LYS	3.4
1	1	148	ALA	3.3
4	B	4	LEU	3.0

### 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 carbohydrates in this entry.

## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.