



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 23, 2024 – 04:45 AM EDT

PDB ID : 4QFL
Title : Crystal structure of dipeptide binding protein from pseudoalteromonas sp. SM9913 in complex with Ala-Phe
Authors : Li, C.Y.; Zhang, Y.Z.
Deposited on : 2014-05-21
Resolution : 1.75 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

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 : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.37.1
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.37.1

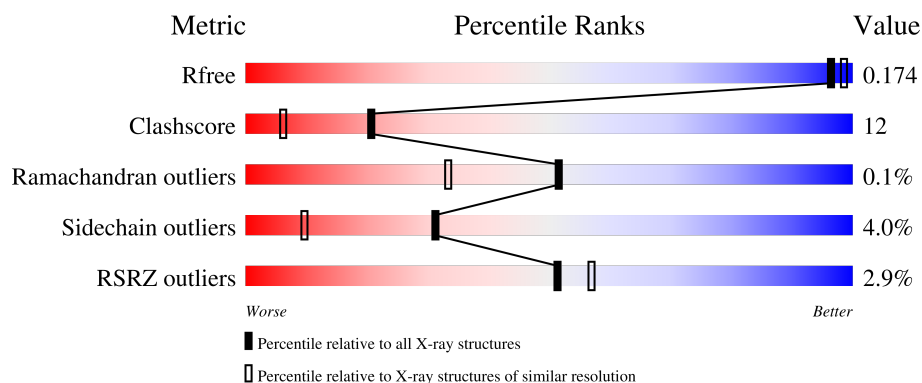
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.75 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}	130704	3764 (1.76-1.72)
Clashscore	141614	3923 (1.76-1.72)
Ramachandran outliers	138981	3878 (1.76-1.72)
Sidechain outliers	138945	3878 (1.76-1.72)
RSRZ outliers	127900	3705 (1.76-1.72)

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	541	<div> <div>3%</div> <div>71%</div> <div>21%</div> <div>6%</div> </div>
1	B	541	<div> <div>3%</div> <div>70%</div> <div>22%</div> <div>6%</div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 8932 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.

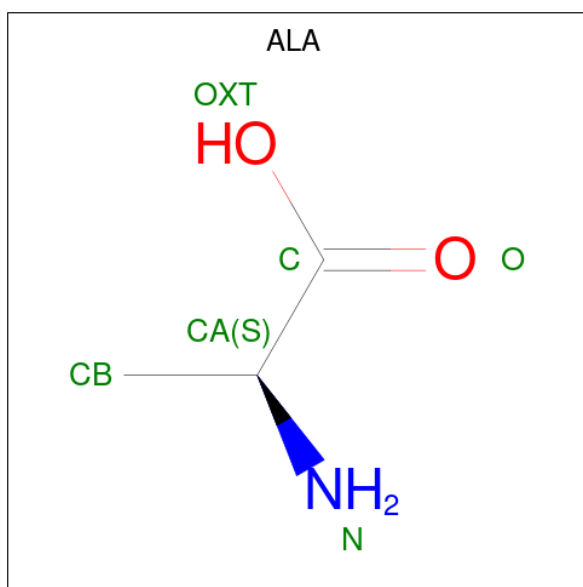
- Molecule 1 is a protein called ABC transporter periplasmic peptide-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	507	Total	C	N	O	S	0	2	0
			4098	2628	691	765	14			
1	B	507	Total	C	N	O	S	0	3	0
			4102	2629	692	767	14			

There are 12 discrepancies between the modelled and reference sequences:

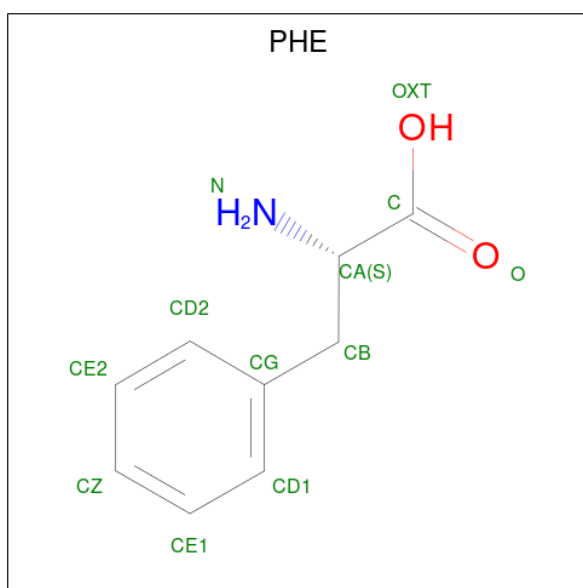
Chain	Residue	Modelled	Actual	Comment	Reference
A	536	HIS	-	expression tag	UNP A7Y7W1
A	537	HIS	-	expression tag	UNP A7Y7W1
A	538	HIS	-	expression tag	UNP A7Y7W1
A	539	HIS	-	expression tag	UNP A7Y7W1
A	540	HIS	-	expression tag	UNP A7Y7W1
A	541	HIS	-	expression tag	UNP A7Y7W1
B	536	HIS	-	expression tag	UNP A7Y7W1
B	537	HIS	-	expression tag	UNP A7Y7W1
B	538	HIS	-	expression tag	UNP A7Y7W1
B	539	HIS	-	expression tag	UNP A7Y7W1
B	540	HIS	-	expression tag	UNP A7Y7W1
B	541	HIS	-	expression tag	UNP A7Y7W1

- Molecule 2 is ALANINE (three-letter code: ALA) (formula: C₃H₇NO₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			5	3	1	1		
2	B	1	Total	C	N	O	0	0
			5	3	1	1		

- Molecule 3 is PHENYLALANINE (three-letter code: PHE) (formula: C₉H₁₁NO₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			12	9	1	2		
3	B	1	Total	C	N	O	0	0
			12	9	1	2		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	349	Total	O	0	0
			349	349		
5	B	337	Total	O	0	0
			337	337		

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.

- Chain A:

Sequence logo for Chain A, showing conservation across 20 positions. The y-axis represents information content in bits (0.00 to 0.25). The x-axis shows positions 1 to 20. The logo indicates high conservation at positions 1, 2, 3, 4, 11, 12, 13, 14, and 15, with lower conservation at positions 5, 6, 7, 8, 9, 10, 16, 17, 18, 19, and 20.

- Chain B:
-
- 70% 3% 22% 6%
- Met His Lys Leu Leu Ala Leu Leu Ser Ser Leu Val Gly Cys Ile Asp Ser Lys Glu Glu Ile Leu Glu Glu Lys Asn Q29 C34 N43 V46 N58 Q59 K83 I84 S85 K86 K97 G98 G99 Y99 R111 G136 P137 A138 N139 V158 H161 Q162 M169 L175 K188 M192 Q199 L202 P212 E217 Y218 R219 R220 D221 H222 L223 V224 Y227 K228 Y232 W233 K234 V237 A238 L239 L242 L243 Y244 N249 G250 T251 T252 R253 L254 A255 A256 L257 L258 L259 K260 E261 S269 A270 A271 Q272 L273 S274 L275 L276 A277 Q278 E285 R286 E287 N291 Y294 N298 T299 E300 R301 P302 D305 R306 L307 R310 Q311 V314 R315 A316 N330 G331 R332 R333 A334 W343 E346 K349 P352 Q357 L358 K361 L362 L363 T364 E365 E369 D373 M379 P380 V381 S382 P383 K391 L395 M396 L400 I403 G404 V405 N406 V407 Y412 E413 W414 L430 L431 G432 W433 D436 N442 L447 T453 K457 N458 P459 A460 N461 T477 T478 D479 L480 N481 L482 R483 K484 Q485 Y486 E496 Q497 Y501 H505 G506 M507 R508 F509 Q510

E517	G518	I519	T520	L521	G522	S528	L529	K536	HIS	HIS	HIS	HIS	HIS	HIS

4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	106.14Å 106.14Å 100.98Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	33.99 – 1.75 33.99 – 1.75	Depositor EDS
% Data completeness (in resolution range)	98.4 (33.99-1.75) 98.4 (33.99-1.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.32 (at 1.75Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, R_{free}	0.160 , 0.184 0.157 , 0.174	Depositor DCC
R_{free} test set	6348 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	23.1	Xtriage
Anisotropy	0.119	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 35.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	0.078 for -h,-k,l 0.320 for h,-h-k,-l 0.079 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	8932	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.84% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL

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	0.38	1/4204 (0.0%)	0.55	1/5718 (0.0%)
1	B	0.36	0/4207	0.53	0/5722
All	All	0.37	1/8411 (0.0%)	0.54	1/11440 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	387	PRO	N-CD	5.55	1.55	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	386	ASN	C-N-CD	5.23	139.37	128.40

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	4098	0	4001	103	0
1	B	4102	0	4006	96	0
2	A	5	0	4	0	0
2	B	5	0	4	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	12	0	9	0	0
3	B	12	0	9	0	0
4	A	6	0	8	1	0
4	B	6	0	8	1	0
5	A	349	0	0	13	0
5	B	337	0	0	20	0
All	All	8932	0	8049	200	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 12.

The worst 5 of 200 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:153:ARG:HH21	1:A:154:LYS:HE2	1.18	1.08
1:A:364:THR:HG22	1:A:369:GLU:HA	1.35	1.07
1:A:366:ALA:H	1:A:367:GLY:HA2	0.92	1.04
1:A:366:ALA:N	1:A:367:GLY:HA2	1.68	1.02
1:A:224:VAL:HG13	1:A:244:TYR:HB2	1.44	0.98

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	507/541 (94%)	488 (96%)	18 (4%)	1 (0%)	47	29
1	B	508/541 (94%)	494 (97%)	14 (3%)	0	100	100
All	All	1015/1082 (94%)	982 (97%)	32 (3%)	1 (0%)	51	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	365	GLU

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	442/472 (94%)	430 (97%)	12 (3%)	44	21
1	B	443/472 (94%)	419 (95%)	24 (5%)	22	5
All	All	885/944 (94%)	849 (96%)	36 (4%)	31	9

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	363	LEU
1	B	517	GLU
1	B	431	LEU
1	B	482	LEU
1	B	34	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	58	ASN
1	B	531	ASN
1	B	249	ASN
1	B	485	GLN
1	B	139	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

6 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	ALA	A	601	3	3,4,5	0.64	0	2,4,6	0.85	0
2	ALA	B	601	3	3,4,5	0.78	0	2,4,6	0.84	0
4	GOL	A	603	-	5,5,5	0.39	0	5,5,5	0.35	0
3	PHE	B	602	2	11,12,12	0.71	1 (9%)	14,15,15	1.14	2 (14%)
3	PHE	A	602	2	11,12,12	0.80	1 (9%)	14,15,15	0.89	2 (14%)
4	GOL	B	603	-	5,5,5	0.38	0	5,5,5	0.59	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ALA	A	601	3	-	0/0/2/4	-
2	ALA	B	601	3	-	0/0/2/4	-
4	GOL	A	603	-	-	0/4/4/4	-
3	PHE	B	602	2	-	2/8/8/8	0/1/1/1
3	PHE	A	602	2	-	2/8/8/8	0/1/1/1
4	GOL	B	603	-	-	2/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	602	PHE	OXT-C	-2.11	1.23	1.30
3	B	602	PHE	OXT-C	-2.05	1.23	1.30

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	602	PHE	OXT-C-O	-3.03	117.21	124.09
3	B	602	PHE	OXT-C-CA	2.75	122.75	113.38
3	A	602	PHE	OXT-C-O	-2.32	118.82	124.09
3	A	602	PHE	OXT-C-CA	2.09	120.49	113.38

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	603	GOL	O1-C1-C2-C3
3	B	602	PHE	CA-CB-CG-CD1
3	B	602	PHE	CA-CB-CG-CD2
4	B	603	GOL	O1-C1-C2-O2
3	A	602	PHE	CA-CB-CG-CD1

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	603	GOL	1	0
4	B	603	GOL	1	0

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, 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	507/541 (93%)	0.37	14 (2%)	53	58	15, 22, 34, 46	0
1	B	507/541 (93%)	0.41	15 (2%)	50	56	17, 24, 34, 49	0
All	All	1014/1082 (93%)	0.39	29 (2%)	51	57	15, 23, 35, 49	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	366	ALA	5.8
1	A	368	TYR	4.8
1	B	271	ALA	3.8
1	A	227[A]	TYR	3.7
1	B	137	ASP	3.3

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](#)

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(\AA^2)	Q<0.9
4	GOL	A	603	6/6	0.86	0.15	19,21,26,26	0
2	ALA	A	601	5/6	0.89	0.12	15,15,17,20	0
3	PHE	B	602	12/12	0.92	0.11	18,21,23,24	0
4	GOL	B	603	6/6	0.92	0.18	19,24,28,33	0
2	ALA	B	601	5/6	0.95	0.10	18,18,20,21	0
3	PHE	A	602	12/12	0.96	0.09	14,17,19,19	0

6.5 Other polymers [i](#)

There are no such residues in this entry.