Full wwPDB X-ray Structure Validation Report ⓘ

Apr 14, 2025 – 07:34 PM EDT

PDB ID	:	4K5L / pdb_00004k5l
Title	:	Phosphonic Arginine Mimetics as Inhibitors of the M1 Aminopeptidases from Plasmodium falciparum
Authors	:	McGowan, S.
Deposited on	:	2013-04-14
Resolution	:	1.91 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:

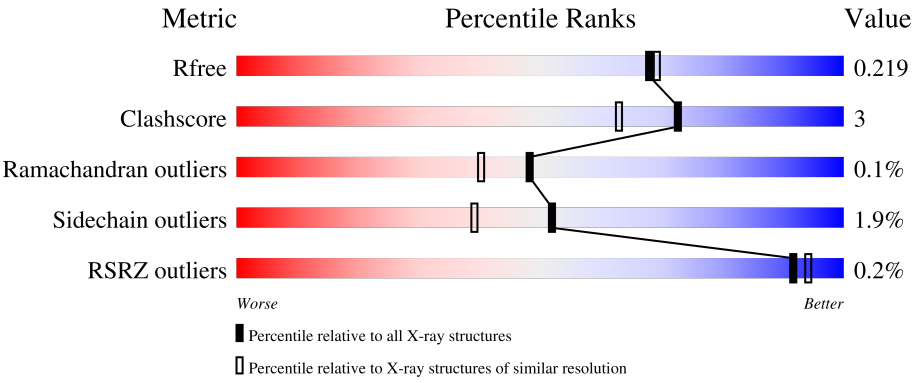
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.91 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	1028 (1.92-1.92)
Clashscore	180529	1100 (1.92-1.92)
Ramachandran outliers	177936	1087 (1.92-1.92)
Sidechain outliers	177891	1087 (1.92-1.92)
RSRZ outliers	164620	1028 (1.92-1.92)

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	A	898	91% 8%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7872 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 M1 family aminopeptidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	889	Total	C	N	O	S	0	18	0
			7312	4714	1185	1384	29			

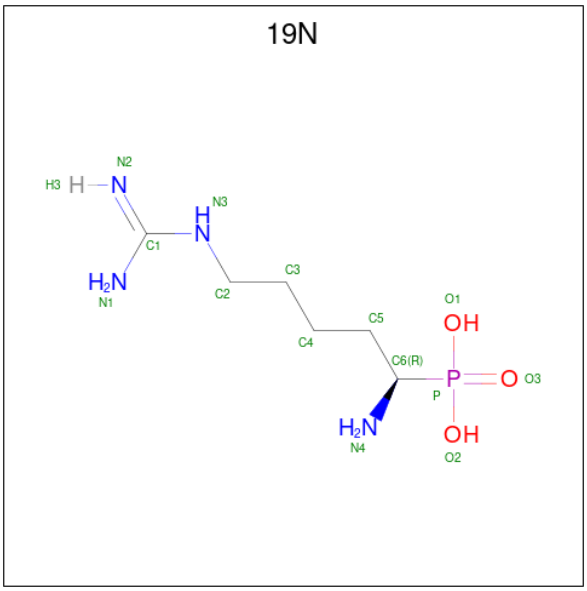
There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	194	MET	-	expression tag	UNP O96935
A	213	GLN	ASN	engineered mutation	UNP O96935
A	223	GLN	ASN	engineered mutation	UNP O96935
A	378	PRO	HIS	engineered mutation	UNP O96935
A	501	GLN	ASN	engineered mutation	UNP O96935
A	745	GLN	ASN	engineered mutation	UNP O96935
A	795	GLN	ASN	engineered mutation	UNP O96935
A	1069	GLN	ASN	engineered mutation	UNP O96935
A	1086	HIS	-	expression tag	UNP O96935
A	1087	HIS	-	expression tag	UNP O96935
A	1088	HIS	-	expression tag	UNP O96935
A	1089	HIS	-	expression tag	UNP O96935
A	1090	HIS	-	expression tag	UNP O96935
A	1091	HIS	-	expression tag	UNP O96935

- 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 [(1R)-1-amino-5-carbamimidamidopentyl]phosphonic acid (CCD ID: 19N) (formula: $C_6H_{17}N_4O_3P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			14	6	4	3	1		

- Molecule 4 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Mg	0	0
			1	1		

- Molecule 5 is water.

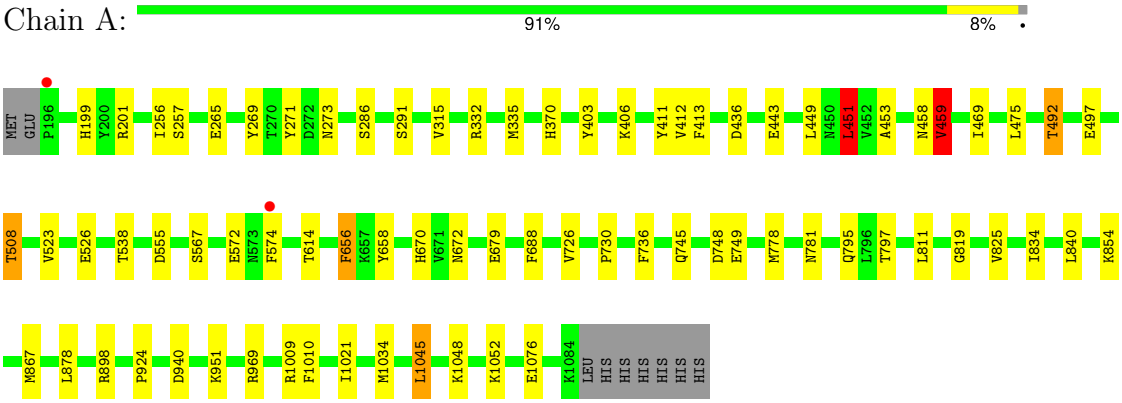
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	544	Total	O	0	0
			544	544		

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: M1 family aminopeptidase

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	76.09Å 109.34Å 118.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	80.42 – 1.91 80.42 – 1.91	Depositor EDS
% Data completeness (in resolution range)	99.6 (80.42-1.91) 99.7 (80.42-1.91)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.64 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.172 , 0.220 0.172 , 0.219	Depositor DCC
R_{free} test set	3918 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	21.2	Xtriage
Anisotropy	0.526	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 43.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7872	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.95% 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: MG, ZN, 19N

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	1.08	5/7532 (0.1%)	0.87	8/10195 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1076	GLU	CB-CG	6.18	1.64	1.52
1	A	656	PHE	CG-CD2	5.65	1.47	1.38
1	A	459	VAL	CA-CB	5.32	1.66	1.54
1	A	736	PHE	CE2-CZ	5.01	1.46	1.37
1	A	825	VAL	CB-CG2	5.01	1.63	1.52

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	451	LEU	CB-CG-CD1	7.29	123.39	111.00
1	A	748	ASP	CB-CG-OD1	6.79	124.41	118.30
1	A	898	ARG	NE-CZ-NH2	-6.48	117.06	120.30
1	A	458	ASN	N-CA-C	5.71	126.42	111.00
1	A	898	ARG	NE-CZ-NH1	5.66	123.13	120.30
1	A	508	THR	N-CA-CB	-5.61	99.64	110.30
1	A	574	PHE	N-CA-CB	-5.55	100.61	110.60
1	A	748	ASP	CB-CG-OD2	-5.26	113.56	118.30

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	7312	0	7214	45	1
2	A	1	0	0	0	0
3	A	14	0	14	1	0
4	A	1	0	0	0	0
5	A	544	0	0	17	1
All	All	7872	0	7228	45	1

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 3.

All (45) 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:508:THR:HG23	5:A:1713:HOH:O	1.50	1.12
1:A:286:SER:HB2	5:A:1599:HOH:O	1.53	1.09
1:A:781[A]:ASN:ND2	5:A:1714:HOH:O	1.93	1.00
1:A:1034[B]:MET:HA	1:A:1034[B]:MET:HE2	1.55	0.88
1:A:265:GLU:OE2	5:A:1325:HOH:O	1.93	0.85
1:A:508:THR:CG2	5:A:1713:HOH:O	2.18	0.79
1:A:1009:ARG:HD3	5:A:1736:HOH:O	1.83	0.76
1:A:726:VAL:HG22	5:A:1493:HOH:O	1.89	0.72
1:A:940:ASP:CB	5:A:1372:HOH:O	2.42	0.68
1:A:1010:PHE:HE1	1:A:1021:ILE:HG13	1.59	0.67
1:A:256:ILE:HD11	1:A:269:TYR:CD1	2.32	0.64
1:A:749[A]:GLU:HG2	5:A:1495:HOH:O	2.01	0.61
1:A:656:PHE:O	1:A:745:GLN:HG3	2.03	0.59
1:A:1010:PHE:CE1	1:A:1021:ILE:HG13	2.38	0.59
1:A:670[B]:HIS:CD2	5:A:1470:HOH:O	2.56	0.57
1:A:403:TYR:N	1:A:443:GLU:OE2	2.35	0.55
1:A:969[B]:ARG:HD3	5:A:1512:HOH:O	2.06	0.55
1:A:1010:PHE:HE1	1:A:1021:ILE:CG1	2.22	0.52
1:A:1034[B]:MET:HE2	1:A:1034[B]:MET:CA	2.26	0.52
1:A:778[B]:MET:SD	1:A:797:THR:HG21	2.50	0.52
1:A:492:THR:HG23	1:A:526[A]:GLU:OE2	2.10	0.51
1:A:778[A]:MET:SD	5:A:1706:HOH:O	2.60	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:878:LEU:HD21	1:A:924:PRO:HG2	1.93	0.50
1:A:572:GLU:HA	3:A:1102:19N:H1	1.76	0.49
1:A:315:VAL:HG23	1:A:469:ILE:HD13	1.94	0.49
1:A:778[B]:MET:HE2	5:A:1702:HOH:O	2.12	0.49
1:A:811:LEU:HB3	1:A:867:MET:SD	2.52	0.49
1:A:508:THR:HG22	1:A:614:THR:HG22	1.94	0.48
1:A:201:ARG:HD2	5:A:1608:HOH:O	2.16	0.45
1:A:523:VAL:O	1:A:526[A]:GLU:HB3	2.17	0.45
1:A:795:GLN:HG3	5:A:1709:HOH:O	2.17	0.45
1:A:271:TYR:CZ	1:A:273:ASN:HA	2.53	0.44
1:A:412:VAL:HG12	1:A:451:LEU:HB2	1.99	0.44
1:A:256:ILE:HD11	1:A:269:TYR:CE1	2.53	0.43
1:A:411:TYR:HB3	1:A:413:PHE:CE2	2.53	0.43
1:A:658:TYR:CE1	1:A:730:PRO:HG3	2.54	0.43
1:A:257:SER:HB2	1:A:291:SER:OG	2.19	0.42
1:A:332:ARG:HG3	1:A:335:MET:HG2	2.01	0.41
1:A:834:ILE:HD13	1:A:1045:LEU:HD22	2.02	0.41
1:A:453:ALA:HB2	1:A:475:LEU:HD23	2.02	0.41
1:A:1048:LYS:HE2	5:A:1671:HOH:O	2.21	0.41
1:A:199:HIS:O	1:A:567:SER:HA	2.20	0.41
1:A:672:ASN:ND2	5:A:1427:HOH:O	2.53	0.41
1:A:538:THR:HG21	1:A:819:GLY:HA3	2.03	0.40

All (1) 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:A:679:GLU:OE2	5:A:1414:HOH:O[3_545]	1.96	0.24

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	905/898 (101%)	887 (98%)	17 (2%)	1 (0%)	48 40

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	459	VAL

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	A	814/830 (98%)	799 (98%)	15 (2%)	54 42

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	370	HIS
1	A	406	LYS
1	A	436	ASP
1	A	449	LEU
1	A	451	LEU
1	A	459	VAL
1	A	492	THR
1	A	497	GLU
1	A	555	ASP
1	A	688	PHE
1	A	840	LEU
1	A	854	LYS
1	A	951	LYS
1	A	1045	LEU
1	A	1052	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	212	ASN
1	A	527	ASN
1	A	745	GLN
1	A	912	ASN
1	A	915	ASN

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, 2 are monoatomic - leaving 1 for Mogul analysis.

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	$\#\|Z\| > 2$
3	19N	A	1102	2	12,13,13	1.36	1 (8%)	11,17,17	2.24	3 (27%)

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
3	19N	A	1102	2	-	5/13/14/14	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1102	19N	P-O3	3.76	1.56	1.49

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1102	19N	O2-P-O1	4.58	120.00	107.58
3	A	1102	19N	O1-P-O3	-4.18	102.96	113.45
3	A	1102	19N	C4-C5-C6	-3.49	103.75	113.41

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1102	19N	N1-C1-N3-C2
3	A	1102	19N	N2-C1-N3-C2
3	A	1102	19N	N3-C2-C3-C4
3	A	1102	19N	C5-C6-P-O3
3	A	1102	19N	C3-C2-N3-C1

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1102	19N	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

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, 95^{th} 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	889/898 (98%)	-0.52	2 (0%) 92 94	8, 21, 36, 53	18 (2%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	574	PHE	4.3
1	A	196	PRO	3.6

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

There are no monosaccharides in this entry.

6.4 Ligands

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, 95^{th} 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
4	MG	A	1103	1/1	0.96	0.10	43,43,43,43	0
3	19N	A	1102	14/14	0.97	0.08	12,17,42,45	0
2	ZN	A	1101	1/1	1.00	0.01	13,13,13,13	0

6.5 Other polymers (i)

There are no such residues in this entry.