



wwPDB X-ray Structure Validation Summary Report ⓘ

May 12, 2025 – 12:09 PM JST

PDB ID : 9KNL / pdb_00009knl
Title : Crystal structure of triethylene glycol-bound full-length PHA synthase (PhaC) from *Aeromonas caviae*
Authors : Chek, M.F.; Kim, S.Y.; Mori, T.; Hakoshima, T.
Deposited on : 2024-11-19
Resolution : 3.20 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 : 4-5-2 with Phenix2.0rc1
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 2.0rc1
EDS : 3.0
buster-report : 1.1.7 (2018)
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) : 2.43.1

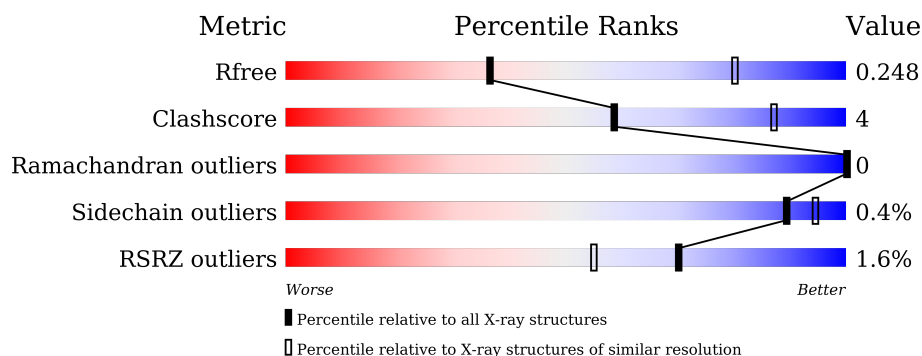
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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	1370 (3.20-3.20)
Clashscore	180529	1497 (3.20-3.20)
Ramachandran outliers	177936	1479 (3.20-3.20)
Sidechain outliers	177891	1478 (3.20-3.20)
RSRZ outliers	164620	1371 (3.20-3.20)

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	596	<div> <div>%</div> <div> <div></div> <div>79%</div> <div>12%</div> <div>9%</div> </div> </div>
1	B	596	<div> <div>%</div> <div> <div></div> <div>80%</div> <div>8%</div> <div>12%</div> </div> </div>
1	C	596	<div> <div>3%</div> <div> <div></div> <div>72%</div> <div>11%</div> <div>17%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 12399 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.

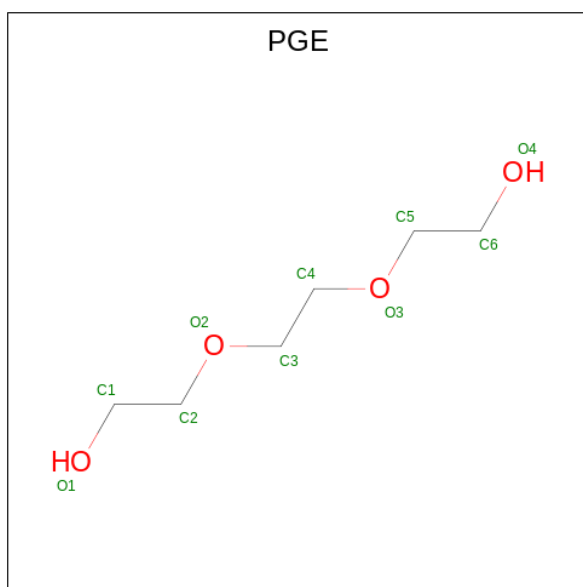
- Molecule 1 is a protein called PHA synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	545	Total	C	N	O	S	0	0	0
			4301	2746	750	791	14			
1	B	525	Total	C	N	O	S	0	0	0
			4154	2655	724	762	13			
1	C	494	Total	C	N	O	S	0	0	0
			3914	2510	677	714	13			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP O32471
A	0	PRO	-	expression tag	UNP O32471
B	-1	GLY	-	expression tag	UNP O32471
B	0	PRO	-	expression tag	UNP O32471
C	-1	GLY	-	expression tag	UNP O32471
C	0	PRO	-	expression tag	UNP O32471

- Molecule 2 is TRIETHYLENE GLYCOL (CCD ID: PGE) (formula: C₆H₁₄O₄) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			10	6	4		
2	A	1	Total	C	O	0	0
			10	6	4		

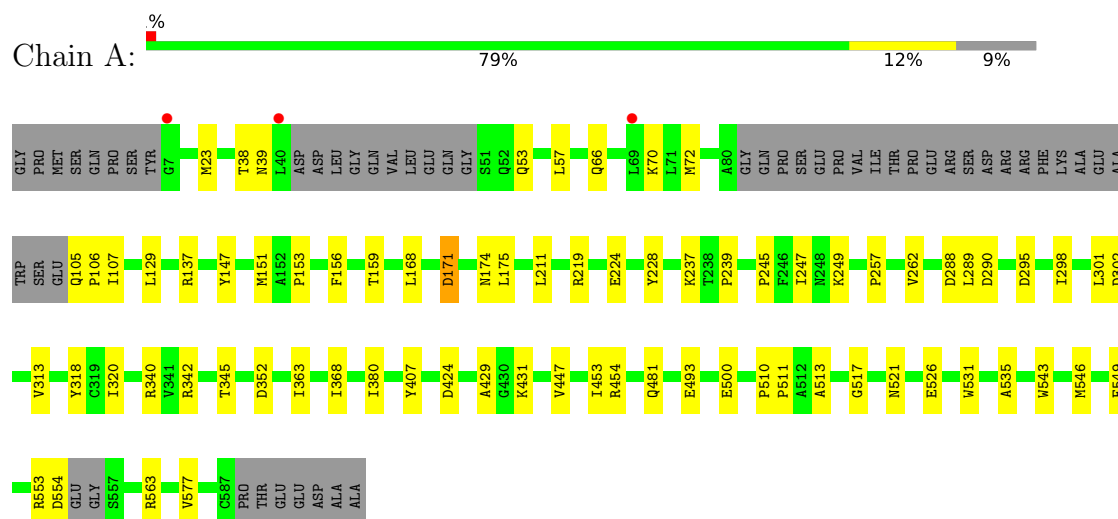
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	8	Total	O	0	0
			8	8		
3	B	2	Total	O	0	0
			2	2		

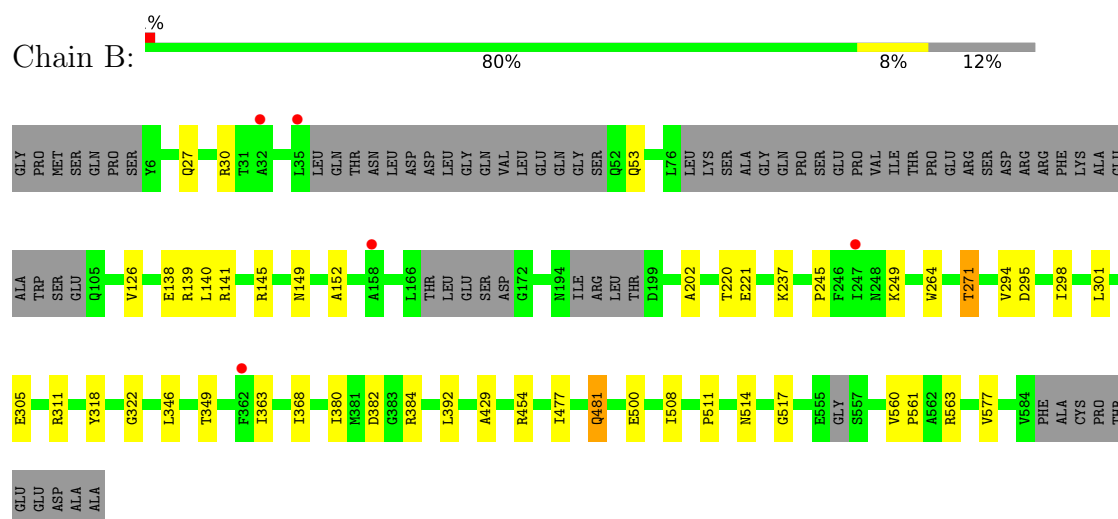
3 Residue-property plots [i](#)

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: PHA synthase



• Molecule 1: PHA synthase



• Molecule 1: PHA synthase



ASP	GLU	GLY	SER	E558	R563	V577	P583	VAL	PHE	ALA	CYS	PRO	THR	GLU	GLU	ASP	ALA	ALA																																			
GLY	Y407	A415	F416	D417	L418	T426	N427	V428	A429	GLY	LYS	THR	H433	L436	Q445	LEU	VAL	I451	K452	I453	R454	R457	V467	L468	L479	L489	F490	Q494	L497	I504	A505	P511	K515	E524	ALA	GLU	S527	P544	N552	ARG													
PRO	K249	Y250	R256	M274	I275	S276	W277	V294	D295	G296	V297	I298	L301	V313	H314	G315	I316	G317	Y318	C319	I320	A334	Q337	LYS	GLN	R340	T345	L346	F347	L351	D352	Q355	N375	T380	M381	D382	G383	V388	L393	ARG	GLU	ASN	SER	L398	N401								
GLY	Q105	R120	L129	R141	L157	N160	PRO	GLU	LEU	LEU	LYS	THR	LEU	LEU	GLU	SER	ASP	GLY	GLN	ASN	VAL	ARG	GLY	L179	L186	I195	ARG	LEU	THR	ASP	GLU	SER	ALA	F203	Q218	R219	E224	K237	T238	P239	V240	L241	P244	P245	F246	I247	N248						
PRO	MET	SER	GLN	PRO	SER	TYR	G7	L43	GLY	GLN	VAL	LEU	GLU	GLN	SER	GLY	GLN	Q53	P54	L57	Q66	K70	L71	M72	T75	LEU	LEU	LYS	SER	ALA	GLY	GLN	PRO	SER	GLU	PRO	VAL	ILE	THR	PRO	GLU	ARG	SER	ASP	ARG	ARG	PHE	LYS	ALA	GLU	ALA	TRP	SER

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	156.71Å 95.53Å 141.17Å 90.00° 90.70° 90.00°	Depositor
Resolution (Å)	47.77 – 3.20 47.77 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.0 (47.77-3.20) 92.9 (47.77-3.20)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.98 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.19.1_4122	Depositor
R, R_{free}	0.221 , 0.247 0.221 , 0.248	Depositor DCC
R_{free} test set	32927 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å ²)	70.5	Xtriage
Anisotropy	0.461	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 60.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.013 for -1/2*h+3/2*k,1/2*h+1/2*k,-l 0.014 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.024 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.022 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l 0.022 for -h,-k,l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	12399	wwPDB-VP
Average B, all atoms (Å ²)	93.0	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

5.1 Standard geometry

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

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	0.22	0/4399	0.38	1/5980 (0.0%)
1	B	0.18	0/4250	0.35	1/5775 (0.0%)
1	C	0.11	0/4003	0.28	0/5439
All	All	0.18	0/12652	0.34	2/17194 (0.0%)

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	A	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	174	ASN	N-CA-C	-5.32	105.65	111.82
1	B	508	ILE	CA-C-O	-5.10	118.19	122.63

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	137	ARG	Sidechain

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	A	4301	0	4277	43	0
1	B	4154	0	4114	27	0
1	C	3914	0	3864	39	0
2	A	20	0	28	2	0
3	A	8	0	0	0	0
3	B	2	0	0	0	0
All	All	12399	0	12283	107	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 4.

The worst 5 of 107 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:C:319:CYS:SG	1:C:320:ILE:N	2.58	0.76
1:C:245:PRO:HD3	1:C:318:TYR:HB3	1.74	0.70
1:B:298:ILE:HA	1:B:301:LEU:HD12	1.75	0.69
1:A:510:PRO:HG2	1:A:513:ALA:HB2	1.74	0.67
1:B:294:VAL:HA	1:B:298:ILE:HD12	1.76	0.67

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	A	537/596 (90%)	517 (96%)	20 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	513/596 (86%)	498 (97%)	15 (3%)	0	100	100
1	C	472/596 (79%)	460 (98%)	12 (2%)	0	100	100
All	All	1522/1788 (85%)	1475 (97%)	47 (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	A	452/494 (92%)	449 (99%)	3 (1%)	81	92
1	B	434/494 (88%)	432 (100%)	2 (0%)	86	93
1	C	409/494 (83%)	409 (100%)	0	100	100
All	All	1295/1482 (87%)	1290 (100%)	5 (0%)	89	94

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	168	LEU
1	A	171	ASP
1	A	481	GLN
1	B	271	THR
1	B	481	GLN

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	C	149	ASN
1	C	551	GLN
1	B	433	HIS
1	B	434	ASN
1	B	481	GLN

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 ⓘ

2 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	PGE	A	601	-	9,9,9	0.51	0	8,8,8	0.24	0
2	PGE	A	602	-	9,9,9	0.51	0	8,8,8	0.24	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	PGE	A	601	-	-	5/7/7/7	-
2	PGE	A	602	-	-	2/7/7/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

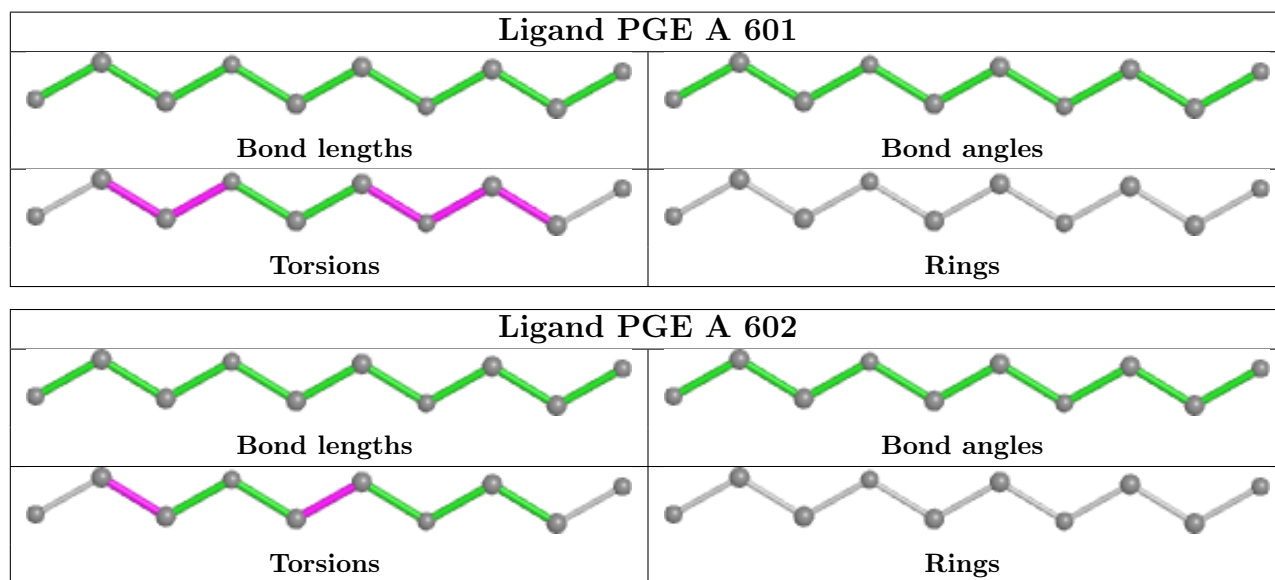
Mol	Chain	Res	Type	Atoms
2	A	601	PGE	O3-C5-C6-O4
2	A	602	PGE	O2-C3-C4-O3
2	A	602	PGE	O3-C5-C6-O4
2	A	601	PGE	C4-C3-O2-C2
2	A	601	PGE	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	PGE	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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, 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	545/596 (91%)	-0.23	3 (0%) 85 76	26, 62, 109, 143	0
1	B	525/596 (88%)	0.03	5 (0%) 79 66	49, 86, 140, 158	0
1	C	494/596 (82%)	0.48	17 (3%) 48 34	65, 126, 153, 168	0
All	All	1564/1788 (87%)	0.08	25 (1%) 70 55	26, 91, 145, 168	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	40	LEU	3.8
1	C	398	LEU	3.6
1	C	436	LEU	3.6
1	B	32	ALA	3.5
1	C	426	THR	3.5

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](#)

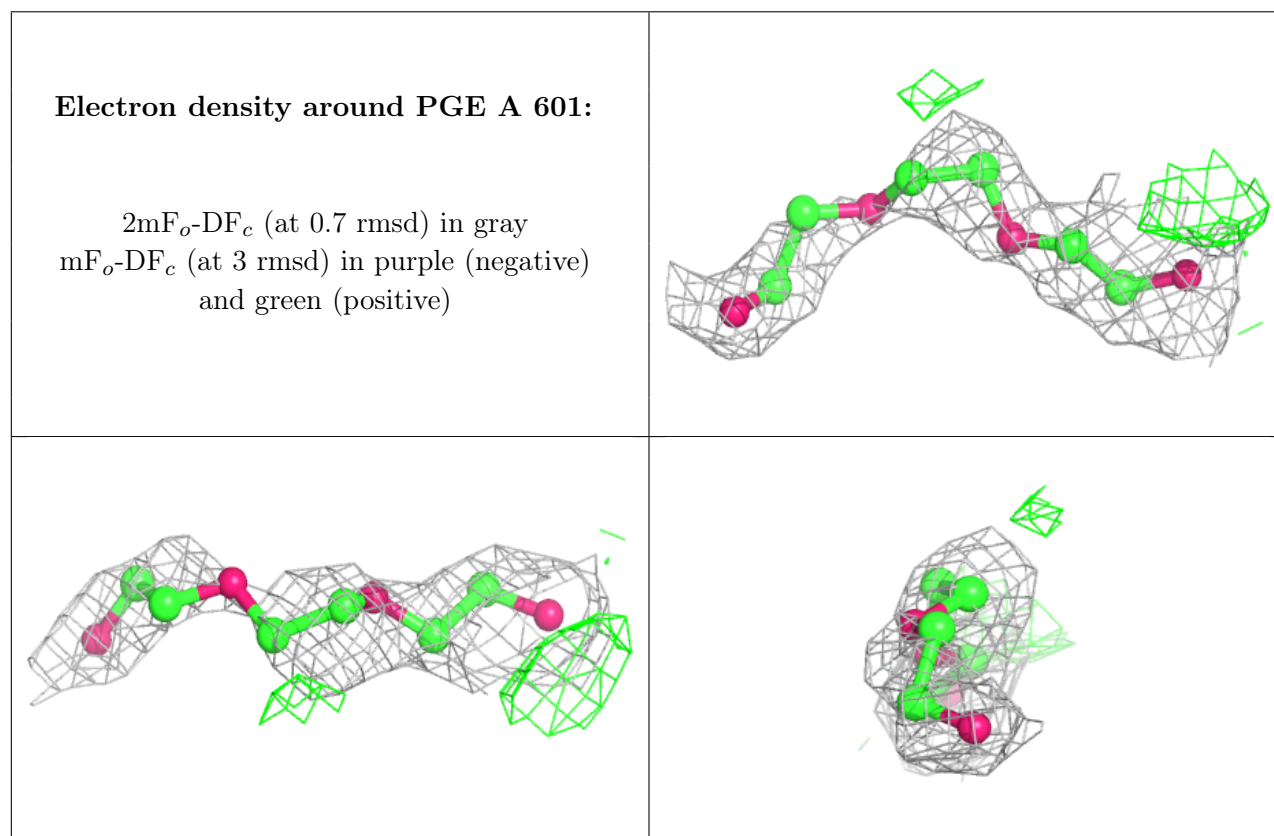
There are no monosaccharides in this entry.

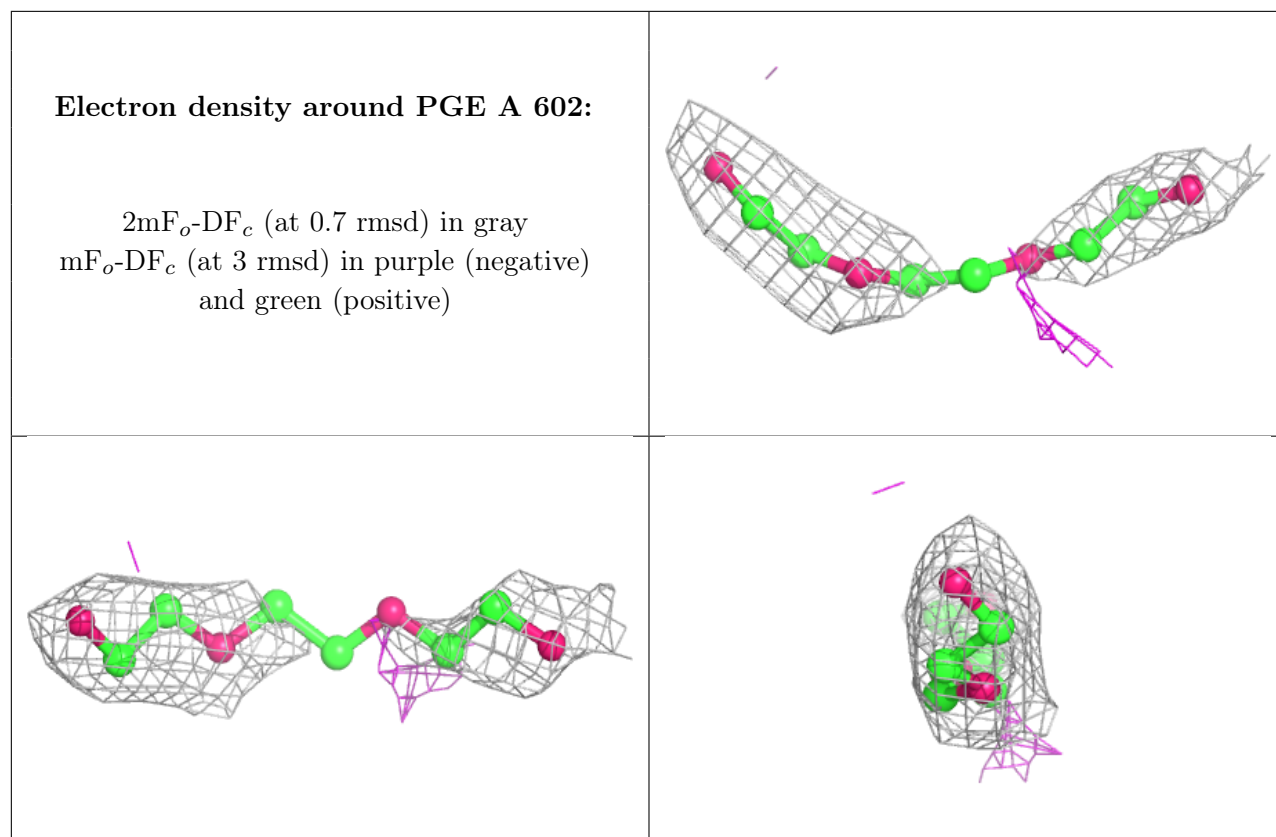
6.4 Ligands [i](#)

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
2	PGE	A	601	10/10	0.88	0.22	58,91,95,115	0
2	PGE	A	602	10/10	0.90	0.21	65,69,84,90	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.





6.5 Other polymers [i](#)

There are no such residues in this entry.