wwPDB X-ray Structure Validation Summary Report (i)

Oct 9, 2023 – 10:13 AM EDT

PDB ID	:	7K9O
Title	:	Co-crystal structure of alpha glucosidase with compound 3
Authors	:	Karade, S.S.; Mariuzza, R.A.
Deposited on	:	2020-09-29
Resolution	:	2.30 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 (i) 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 (i)) were used in the production of this report:

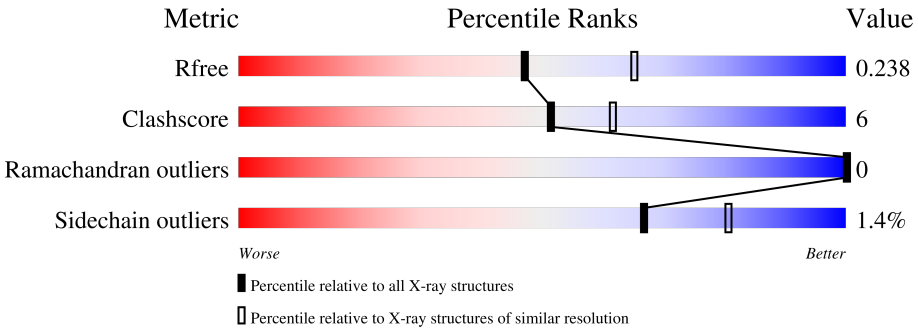
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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 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%

Mol	Chain	Length	Quality of chain
1	A	977	75% 13% 13%
1	C	977	75% 12% 12%
2	B	547	13% · 85%
2	D	547	14% · 84%

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 15936 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 Alpha glucosidase 2 alpha neutral subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	854	Total	C	N	O	S	0	10	0
			6925	4441	1194	1260	30			
1	C	856	Total	C	N	O	S	0	11	0
			6939	4450	1194	1265	30			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	MET	-	initiating methionine	UNP A1A4T2
A	3	GLY	-	expression tag	UNP A1A4T2
A	4	ILE	-	expression tag	UNP A1A4T2
A	5	LEU	-	expression tag	UNP A1A4T2
A	6	PRO	-	expression tag	UNP A1A4T2
A	7	SER	-	expression tag	UNP A1A4T2
A	8	PRO	-	expression tag	UNP A1A4T2
A	9	GLY	-	expression tag	UNP A1A4T2
A	10	MET	-	expression tag	UNP A1A4T2
A	11	PRO	-	expression tag	UNP A1A4T2
A	12	ALA	-	expression tag	UNP A1A4T2
A	13	LEU	-	expression tag	UNP A1A4T2
A	14	LEU	-	expression tag	UNP A1A4T2
A	15	SER	-	expression tag	UNP A1A4T2
A	16	LEU	-	expression tag	UNP A1A4T2
A	17	VAL	-	expression tag	UNP A1A4T2
A	18	SER	-	expression tag	UNP A1A4T2
A	19	LEU	-	expression tag	UNP A1A4T2
A	20	LEU	-	expression tag	UNP A1A4T2
A	21	SER	-	expression tag	UNP A1A4T2
A	22	VAL	-	expression tag	UNP A1A4T2
A	23	LEU	-	expression tag	UNP A1A4T2
A	24	LEU	-	expression tag	UNP A1A4T2
A	25	MET	-	expression tag	UNP A1A4T2
A	26	GLY	-	expression tag	UNP A1A4T2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	27	CYS	-	expression tag	UNP A1A4T2
A	28	VAL	-	expression tag	UNP A1A4T2
A	29	ALA	-	expression tag	UNP A1A4T2
A	30	GLU	-	expression tag	UNP A1A4T2
A	31	THR	-	expression tag	UNP A1A4T2
A	32	GLY	-	expression tag	UNP A1A4T2
A	97	ASP	ASN	engineered mutation	UNP A1A4T2
A	967	SER	-	expression tag	UNP A1A4T2
A	968	ALA	-	expression tag	UNP A1A4T2
A	969	TRP	-	expression tag	UNP A1A4T2
A	970	SER	-	expression tag	UNP A1A4T2
A	971	HIS	-	expression tag	UNP A1A4T2
A	972	PRO	-	expression tag	UNP A1A4T2
A	973	GLN	-	expression tag	UNP A1A4T2
A	974	PHE	-	expression tag	UNP A1A4T2
A	975	GLU	-	expression tag	UNP A1A4T2
A	976	LYS	-	expression tag	UNP A1A4T2
A	977	LEU	-	expression tag	UNP A1A4T2
A	978	GLU	-	expression tag	UNP A1A4T2
C	2	MET	-	initiating methionine	UNP A1A4T2
C	3	GLY	-	expression tag	UNP A1A4T2
C	4	ILE	-	expression tag	UNP A1A4T2
C	5	LEU	-	expression tag	UNP A1A4T2
C	6	PRO	-	expression tag	UNP A1A4T2
C	7	SER	-	expression tag	UNP A1A4T2
C	8	PRO	-	expression tag	UNP A1A4T2
C	9	GLY	-	expression tag	UNP A1A4T2
C	10	MET	-	expression tag	UNP A1A4T2
C	11	PRO	-	expression tag	UNP A1A4T2
C	12	ALA	-	expression tag	UNP A1A4T2
C	13	LEU	-	expression tag	UNP A1A4T2
C	14	LEU	-	expression tag	UNP A1A4T2
C	15	SER	-	expression tag	UNP A1A4T2
C	16	LEU	-	expression tag	UNP A1A4T2
C	17	VAL	-	expression tag	UNP A1A4T2
C	18	SER	-	expression tag	UNP A1A4T2
C	19	LEU	-	expression tag	UNP A1A4T2
C	20	LEU	-	expression tag	UNP A1A4T2
C	21	SER	-	expression tag	UNP A1A4T2
C	22	VAL	-	expression tag	UNP A1A4T2
C	23	LEU	-	expression tag	UNP A1A4T2
C	24	LEU	-	expression tag	UNP A1A4T2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	25	MET	-	expression tag	UNP A1A4T2
C	26	GLY	-	expression tag	UNP A1A4T2
C	27	CYS	-	expression tag	UNP A1A4T2
C	28	VAL	-	expression tag	UNP A1A4T2
C	29	ALA	-	expression tag	UNP A1A4T2
C	30	GLU	-	expression tag	UNP A1A4T2
C	31	THR	-	expression tag	UNP A1A4T2
C	32	GLY	-	expression tag	UNP A1A4T2
C	97	ASP	ASN	engineered mutation	UNP A1A4T2
C	967	SER	-	expression tag	UNP A1A4T2
C	968	ALA	-	expression tag	UNP A1A4T2
C	969	TRP	-	expression tag	UNP A1A4T2
C	970	SER	-	expression tag	UNP A1A4T2
C	971	HIS	-	expression tag	UNP A1A4T2
C	972	PRO	-	expression tag	UNP A1A4T2
C	973	GLN	-	expression tag	UNP A1A4T2
C	974	PHE	-	expression tag	UNP A1A4T2
C	975	GLU	-	expression tag	UNP A1A4T2
C	976	LYS	-	expression tag	UNP A1A4T2
C	977	LEU	-	expression tag	UNP A1A4T2
C	978	GLU	-	expression tag	UNP A1A4T2

- Molecule 2 is a protein called Glucosidase 2 subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	83	Total	C	N	O	S	0	0	0
			603	359	98	136	10			
2	D	85	Total	C	N	O	S	0	0	0
			607	362	97	138	10			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	MET	-	initiating methionine	UNP O08795
B	-15	GLY	-	expression tag	UNP O08795
B	-14	ILE	-	expression tag	UNP O08795
B	-13	LEU	-	expression tag	UNP O08795
B	-12	PRO	-	expression tag	UNP O08795
B	-11	SER	-	expression tag	UNP O08795
B	-10	PRO	-	expression tag	UNP O08795
B	-9	GLY	-	expression tag	UNP O08795
B	-8	MET	-	expression tag	UNP O08795

Continued on next page...

Continued from previous page...

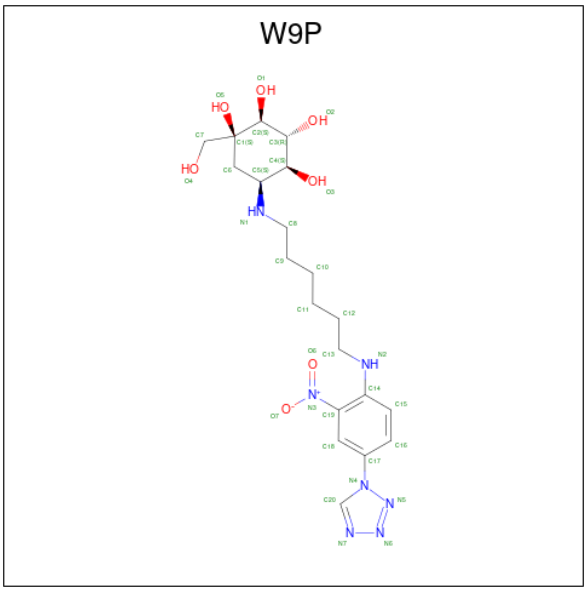
Chain	Residue	Modelled	Actual	Comment	Reference
B	-7	PRO	-	expression tag	UNP O08795
B	-6	ALA	-	expression tag	UNP O08795
B	-5	LEU	-	expression tag	UNP O08795
B	-4	LEU	-	expression tag	UNP O08795
B	-3	SER	-	expression tag	UNP O08795
B	-2	LEU	-	expression tag	UNP O08795
B	-1	VAL	-	expression tag	UNP O08795
B	0	SER	-	expression tag	UNP O08795
B	1	LEU	-	expression tag	UNP O08795
B	2	LEU	-	expression tag	UNP O08795
B	3	SER	-	expression tag	UNP O08795
B	4	VAL	-	expression tag	UNP O08795
B	5	LEU	-	expression tag	UNP O08795
B	6	LEU	-	expression tag	UNP O08795
B	7	MET	-	expression tag	UNP O08795
B	8	GLY	-	expression tag	UNP O08795
B	9	CYS	-	expression tag	UNP O08795
B	10	VAL	-	expression tag	UNP O08795
B	11	ALA	-	expression tag	UNP O08795
B	12	GLU	-	expression tag	UNP O08795
B	13	THR	-	expression tag	UNP O08795
B	14	GLY	-	expression tag	UNP O08795
B	518	SER	-	expression tag	UNP O08795
B	519	ALA	-	expression tag	UNP O08795
B	520	TRP	-	expression tag	UNP O08795
B	521	LEU	-	expression tag	UNP O08795
B	522	GLU	-	expression tag	UNP O08795
B	523	THR	-	expression tag	UNP O08795
B	524	LYS	-	expression tag	UNP O08795
B	525	HIS	-	expression tag	UNP O08795
B	526	HIS	-	expression tag	UNP O08795
B	527	HIS	-	expression tag	UNP O08795
B	528	HIS	-	expression tag	UNP O08795
B	529	HIS	-	expression tag	UNP O08795
B	530	HIS	-	expression tag	UNP O08795
D	-16	MET	-	initiating methionine	UNP O08795
D	-15	GLY	-	expression tag	UNP O08795
D	-14	ILE	-	expression tag	UNP O08795
D	-13	LEU	-	expression tag	UNP O08795
D	-12	PRO	-	expression tag	UNP O08795
D	-11	SER	-	expression tag	UNP O08795
D	-10	PRO	-	expression tag	UNP O08795

Continued on next page...

Continued from previous page...

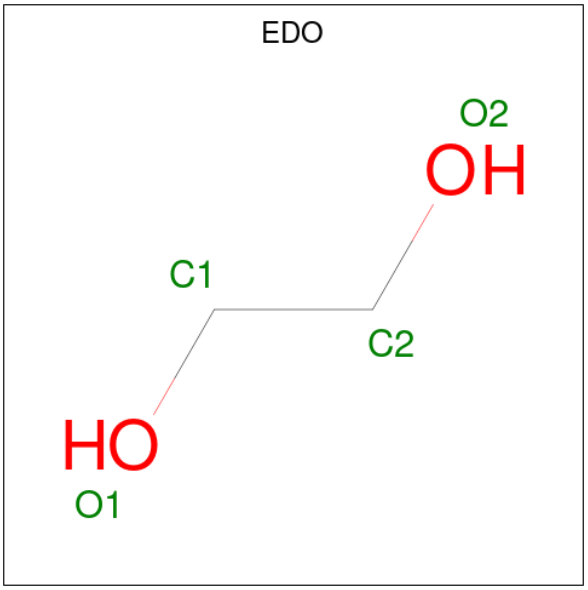
Chain	Residue	Modelled	Actual	Comment	Reference
D	-9	GLY	-	expression tag	UNP O08795
D	-8	MET	-	expression tag	UNP O08795
D	-7	PRO	-	expression tag	UNP O08795
D	-6	ALA	-	expression tag	UNP O08795
D	-5	LEU	-	expression tag	UNP O08795
D	-4	LEU	-	expression tag	UNP O08795
D	-3	SER	-	expression tag	UNP O08795
D	-2	LEU	-	expression tag	UNP O08795
D	-1	VAL	-	expression tag	UNP O08795
D	0	SER	-	expression tag	UNP O08795
D	1	LEU	-	expression tag	UNP O08795
D	2	LEU	-	expression tag	UNP O08795
D	3	SER	-	expression tag	UNP O08795
D	4	VAL	-	expression tag	UNP O08795
D	5	LEU	-	expression tag	UNP O08795
D	6	LEU	-	expression tag	UNP O08795
D	7	MET	-	expression tag	UNP O08795
D	8	GLY	-	expression tag	UNP O08795
D	9	CYS	-	expression tag	UNP O08795
D	10	VAL	-	expression tag	UNP O08795
D	11	ALA	-	expression tag	UNP O08795
D	12	GLU	-	expression tag	UNP O08795
D	13	THR	-	expression tag	UNP O08795
D	14	GLY	-	expression tag	UNP O08795
D	518	SER	-	expression tag	UNP O08795
D	519	ALA	-	expression tag	UNP O08795
D	520	TRP	-	expression tag	UNP O08795
D	521	LEU	-	expression tag	UNP O08795
D	522	GLU	-	expression tag	UNP O08795
D	523	THR	-	expression tag	UNP O08795
D	524	LYS	-	expression tag	UNP O08795
D	525	HIS	-	expression tag	UNP O08795
D	526	HIS	-	expression tag	UNP O08795
D	527	HIS	-	expression tag	UNP O08795
D	528	HIS	-	expression tag	UNP O08795
D	529	HIS	-	expression tag	UNP O08795
D	530	HIS	-	expression tag	UNP O08795

- Molecule 3 is (1 {S},2 {S},3 {R},4 {S},5 {S})-1-(hydroxymethyl)-5-[6-[[2-[oxidanyl(oxidanylidene)-$l^{4}-azanyl]-4-(1,2,3,4-tetrazol-1-yl)phenyl]amino]hexylamino]cyclohexane-1,2,3,4-tetrol (three-letter code: W9P) (formula: $C_{20}H_{31}N_7O_7$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 34 20 7 7	0	0
3	C	1	Total C N O 34 20 7 7	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

Continued on next page...

Continued from previous page...

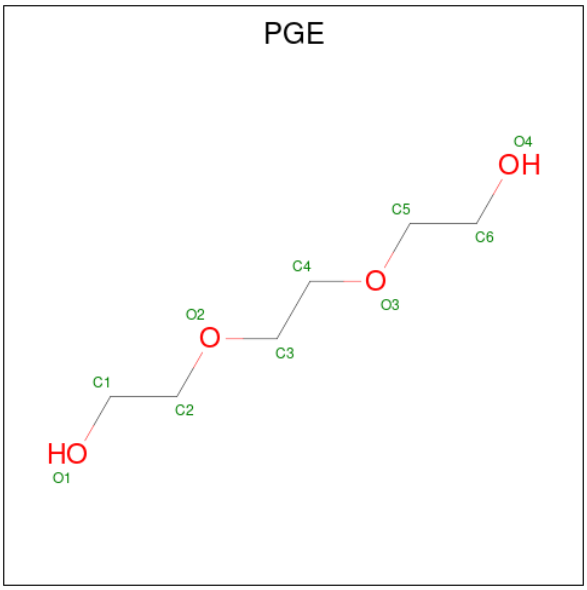
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total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

Continued on next page...

Continued from previous page...

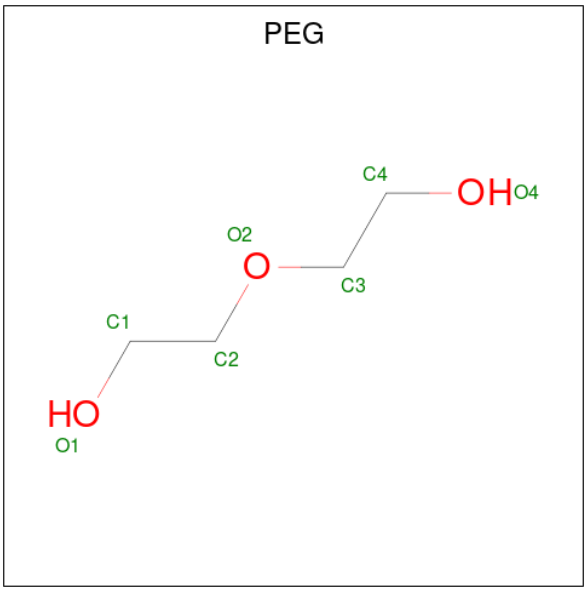
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



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

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

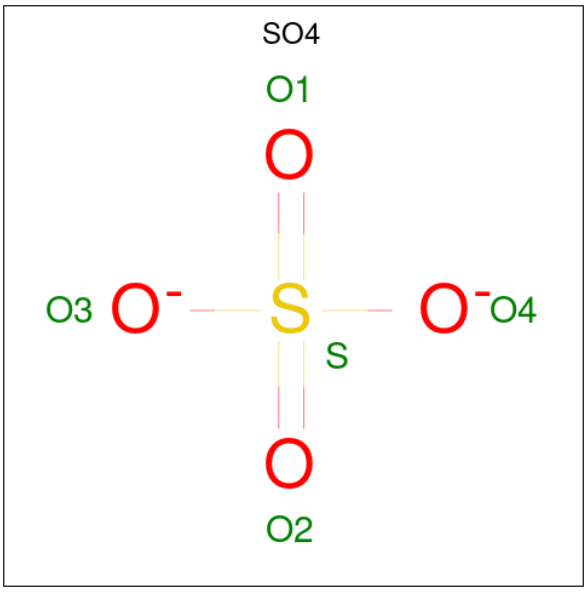


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 7 4 3	0	0
6	A	1	Total C O 7 4 3	0	0
6	A	1	Total C O 7 4 3	0	0
6	C	1	Total C O 7 4 3	0	0

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	2	Total Ca 2 2	0	0
7	D	2	Total Ca 2 2	0	0

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	C	1	Total O S 5 4 1	0	0

- Molecule 9 is water.

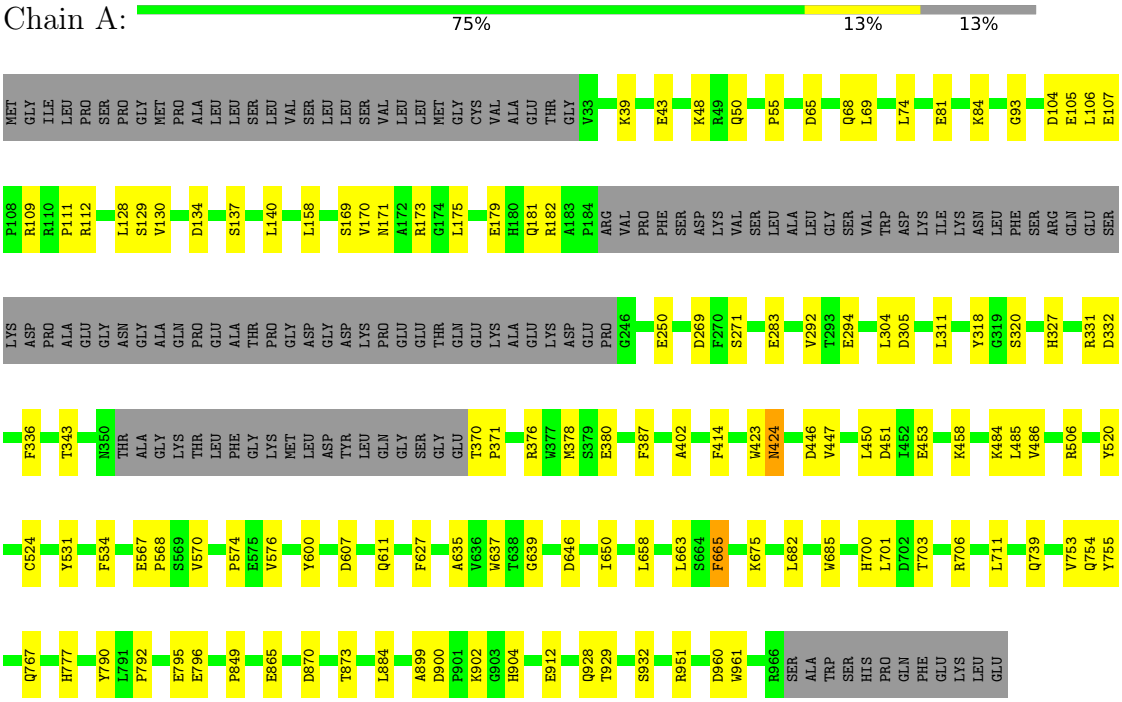
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	332	Total O 332 332	0	0
9	B	14	Total O 14 14	0	0
9	C	261	Total O 261 261	0	0
9	D	12	Total O 12 12	0	0

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. 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. 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.

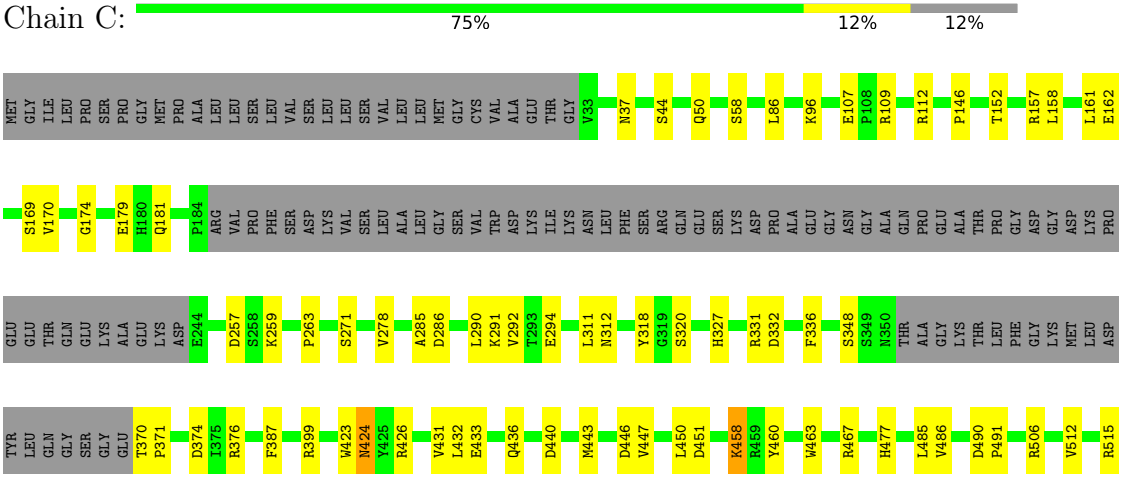
- Molecule 1: Alpha glucosidase 2 alpha neutral subunit

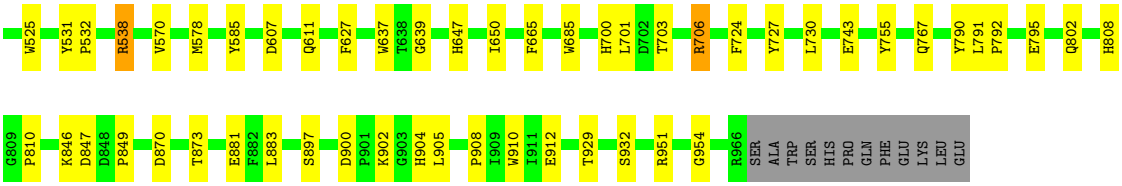
Chain A:



- Molecule 1: Alpha glucosidase 2 alpha neutral subunit

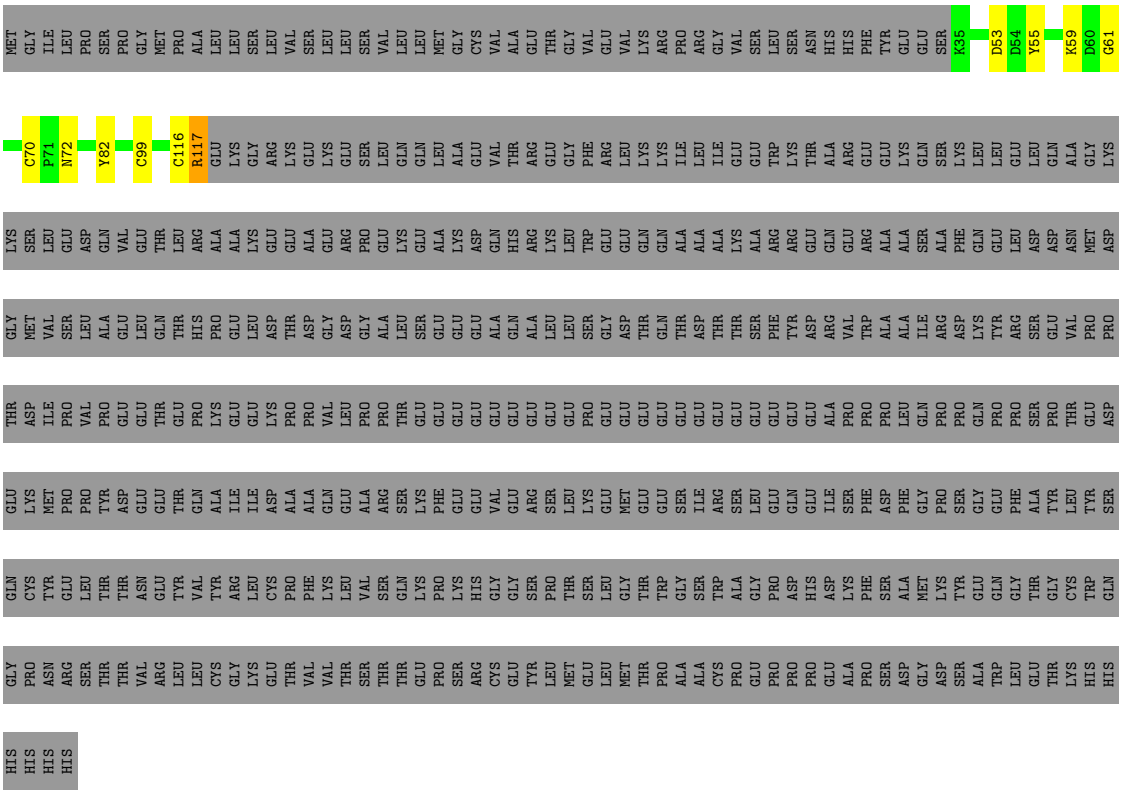
Chain C:

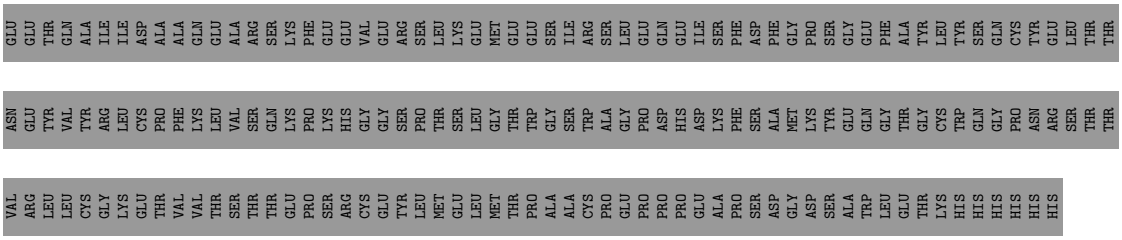




• Molecule 2: Glucosidase 2 subunit beta

Chain B:


MET GLY ILE LEU PRO SER PRO GLY MET PRO ALA LEU LEU SER LEU VAL SER LEU LEU SER VAL LEU LEU MET GLY CYS VAL ALA GLU THR GLY VAL GLU VAL LYS ARG PRO ARG GLY VAL SER LEU SER ASN HIS HIS PHE TYR GLU GLU SER K35 D53 D54 Y55 K59 D60 G61
C70 P71 N72 Y82 C99 C116 R117 GLU LYS GLY ARG LYS GLU LYS GLU ALA ALA GLU ARG PRO GLU LYS GLU ALA LYS ASP GLN HIS ARG LYS LEU TRP GLU GLN GLN LEU ALA ALA LYS GLU VAL THR ARG GLU GLY PHE ARG LEU LYS LYS ILE LEU ILE GLU GLU TRP LYS THR ALA ARG GLU GLU LYS GLN SER LYS LEU LEU GLU LEU GLN ALA GLY LYS LYS SER LEU GLU ASP ASP ASN MET ASP GLY MET VAL SER LEU ALA GLU
LYS SER LEU GLU ASP GLN VAL GLU GLU ALA LYS GLU GLU ALA GLU ARG PRO GLU LYS GLU ALA LYS ASP GLN HIS ARG LYS LEU TRP GLU GLN GLN ALA ALA LYS GLU VAL THR ARG GLU GLY PHE ARG LEU LYS LYS ILE LEU ILE GLU GLU TRP LYS THR ALA ARG GLU GLU LYS GLN SER LYS LEU LEU GLU LEU GLN ALA GLY LYS LYS SER LEU GLU ASP ASP ASN MET ASP
GLY MET VAL SER LEU ALA GLU LEU GLN THR HIS PRO GLU LEU ASP THR ASP GLY ASP GLY ALA LEU SER GLU GLU GLU ALA GLN THR ASP THR THR SER PHE TYR ASP ARG VAL TRP ALA ALA ILE ARG ASP LYS TYR ARG SER GLU VAL PRO PRO
THR ASP ILE PRO VAL PRO GLU GLU THR GLU PRO PRO LYS GLU GLU LYS PRO PRO VAL LEU PRO PRO THR GLU GLU GLU


GLU GLU THR GLN ALA ILE ILE ASP ALA ALA GLN GLU ALA ARG SER LYS PHE GLU GLU VAL GLU ARG SER LEU LYS GLU MET GLU GLU SER ILE ARG SER LEU GLU GLN GLU ILE SER PHE ASP PHE GLY PRO SER GLY GLU PHE ALA TYR LEU TYR SER GLN CYS TYR GLU LEU THR THR
ASN GLU TYR VAL TYR ARG LEU CYS PRO PHE LYS LEU VAL SER GLN LYS PRO LYS HIS GLY GLY SER PRO THR SER LEU GLY THR TRP GLY SER TRP ALA GLY PRO ASP HIS ASP LYS PHE SER ALA MET LYS TYR GLU GLN GLY THR GLY CYS TRP GLN GLY PRO ASN ARG SER THR THR
VAL ARG LEU LEU CYS GLY LYS GLU THR VAL THR SER THR THR GLU PRO SER ARG CYS GLU TYR LEU MET GLU LEU MET THR PRO ALA ALA CYS PRO GLU PRO PRO PRO PRO GLU ALA PRO SER ASP GLY ASP SER ALA TRP LEU GLU THR LYS HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	102.72Å 102.72Å 239.62Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.94 – 2.30 43.20 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (39.94-2.30) 97.0 (43.20-2.30)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.59 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.14_3228	Depositor
R, R_{free}	0.210 , 0.238 0.210 , 0.238	Depositor DCC
R_{free} test set	2012 reflections (1.61%)	wwPDB-VP
Wilson B-factor (Å^2)	40.3	Xtriage
Anisotropy	0.044	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 30.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.21$	Xtriage
Estimated twinning fraction	0.085 for -h,-k,l 0.247 for h,-h-k,-l 0.085 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	15936	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.48% 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: CA, SO4, EDO, PGE, W9P, PEG

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	0.28	0/7165	0.49	0/9751
1	C	0.27	0/7183	0.48	0/9777
2	B	0.28	0/615	0.59	1/842 (0.1%)
2	D	0.26	0/619	0.52	0/849
All	All	0.27	0/15582	0.49	1/21219 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	117	ARG	NE-CZ-NH1	6.56	123.58	120.30

There are no chirality outliers.

There are no planarity outliers.

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	6925	0	6689	84	1
1	C	6939	0	6692	74	1
2	B	603	0	500	5	0
2	D	607	0	493	5	0
3	A	34	0	0	0	0
3	C	34	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	56	0	84	3	0
4	B	4	0	6	0	0
4	C	48	0	72	5	0
5	A	20	0	28	5	0
5	C	10	0	14	1	0
6	A	21	0	29	1	0
6	C	7	0	10	0	0
7	B	2	0	0	0	0
7	D	2	0	0	0	0
8	C	5	0	0	0	0
9	A	332	0	0	24	0
9	B	14	0	0	1	0
9	C	261	0	0	21	0
9	D	12	0	0	1	0
All	All	15936	0	14617	166	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 6.

The worst 5 of 166 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:463:TRP:O	9:C:1102:HOH:O	1.81	0.99
1:C:611:GLN:OE1	9:C:1101:HOH:O	1.80	0.97
1:C:802:GLN:NE2	9:C:1105:HOH:O	1.99	0.94
1:A:574:PRO:O	9:A:1101:HOH:O	1.90	0.89
1:C:263:PRO:O	9:C:1103:HOH:O	1.93	0.85

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:182:ARG:NH2	1:C:795:GLU:OE1[3_665]	2.14	0.06

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	858/977 (88%)	825 (96%)	33 (4%)	0	100	100
1	C	861/977 (88%)	831 (96%)	30 (4%)	0	100	100
2	B	81/547 (15%)	80 (99%)	1 (1%)	0	100	100
2	D	83/547 (15%)	81 (98%)	2 (2%)	0	100	100
All	All	1883/3048 (62%)	1817 (96%)	66 (4%)	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	751/846 (89%)	743 (99%)	8 (1%)	73	86
1	C	752/846 (89%)	740 (98%)	12 (2%)	62	78
2	B	69/478 (14%)	67 (97%)	2 (3%)	42	58
2	D	68/478 (14%)	68 (100%)	0	100	100
All	All	1640/2648 (62%)	1618 (99%)	22 (1%)	67	82

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	458	LYS
1	C	665	PHE
1	C	637	TRP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	685	TRP
1	A	685	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry

Of 41 ligands modelled in this entry, 4 are monoatomic - leaving 37 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	$\#\|Z\| > 2$
4	EDO	C	1011	-	3,3,3	0.41	0	2,2,2	0.49	0
4	EDO	A	1008	-	3,3,3	0.47	0	2,2,2	0.36	0
4	EDO	A	1015	-	3,3,3	0.44	0	2,2,2	0.42	0
5	PGE	C	1008	-	9,9,9	0.31	0	8,8,8	0.44	0
4	EDO	A	1004	-	3,3,3	0.48	0	2,2,2	0.30	0
6	PEG	A	1009	-	6,6,6	0.48	0	5,5,5	0.31	0
4	EDO	A	1005	-	3,3,3	0.44	0	2,2,2	0.43	0
3	W9P	A	1001	-	33,36,36	2.12	2 (6%)	39,50,50	1.21	4 (10%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	EDO	C	1013	-	3,3,3	0.49	0	2,2,2	0.37	0
4	EDO	A	1010	-	3,3,3	0.44	0	2,2,2	0.33	0
6	PEG	A	1014	-	6,6,6	0.48	0	5,5,5	0.34	0
4	EDO	C	1016	-	3,3,3	0.48	0	2,2,2	0.30	0
4	EDO	C	1015	-	3,3,3	0.42	0	2,2,2	0.56	0
4	EDO	C	1007	-	3,3,3	0.46	0	2,2,2	0.36	0
4	EDO	A	1011	-	3,3,3	0.46	0	2,2,2	0.43	0
4	EDO	B	603	-	3,3,3	0.47	0	2,2,2	0.33	0
6	PEG	C	1003	-	6,6,6	0.48	0	5,5,5	0.36	0
4	EDO	A	1018	-	3,3,3	0.42	0	2,2,2	0.48	0
4	EDO	A	1019	-	3,3,3	0.41	0	2,2,2	0.46	0
6	PEG	A	1020	-	6,6,6	0.48	0	5,5,5	0.25	0
8	SO4	C	1002	-	4,4,4	0.13	0	6,6,6	0.11	0
4	EDO	A	1012	-	3,3,3	0.47	0	2,2,2	0.35	0
3	W9P	C	1001	-	33,36,36	2.10	2 (6%)	39,50,50	1.45	3 (7%)
4	EDO	A	1013	-	3,3,3	0.48	0	2,2,2	0.32	0
4	EDO	C	1004	-	3,3,3	0.44	0	2,2,2	0.41	0
4	EDO	A	1003	-	3,3,3	0.45	0	2,2,2	0.59	0
4	EDO	C	1009	-	3,3,3	0.45	0	2,2,2	0.32	0
4	EDO	A	1017	-	3,3,3	0.44	0	2,2,2	0.43	0
4	EDO	A	1002	-	3,3,3	0.42	0	2,2,2	0.51	0
4	EDO	C	1010	-	3,3,3	0.46	0	2,2,2	0.44	0
5	PGE	A	1007	-	9,9,9	0.31	0	8,8,8	0.34	0
5	PGE	A	1006	-	9,9,9	0.28	0	8,8,8	0.40	0
4	EDO	A	1016	-	3,3,3	0.42	0	2,2,2	0.49	0
4	EDO	C	1006	-	3,3,3	0.45	0	2,2,2	0.41	0
4	EDO	C	1014	-	3,3,3	0.44	0	2,2,2	0.46	0
4	EDO	C	1005	-	3,3,3	0.44	0	2,2,2	0.32	0
4	EDO	C	1012	-	3,3,3	0.44	0	2,2,2	0.43	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
4	EDO	C	1011	-	-	1/1/1/1	-
4	EDO	A	1008	-	-	1/1/1/1	-
4	EDO	A	1015	-	-	0/1/1/1	-
5	PGE	C	1008	-	-	4/7/7/7	-
4	EDO	A	1004	-	-	1/1/1/1	-
6	PEG	A	1009	-	-	1/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	EDO	A	1005	-	-	0/1/1/1	-
3	W9P	A	1001	-	-	9/20/45/45	0/3/3/3
4	EDO	C	1013	-	-	0/1/1/1	-
4	EDO	A	1010	-	-	0/1/1/1	-
6	PEG	A	1014	-	-	3/4/4/4	-
4	EDO	C	1016	-	-	1/1/1/1	-
4	EDO	C	1015	-	-	1/1/1/1	-
4	EDO	C	1007	-	-	1/1/1/1	-
4	EDO	A	1011	-	-	1/1/1/1	-
4	EDO	B	603	-	-	1/1/1/1	-
6	PEG	C	1003	-	-	2/4/4/4	-
4	EDO	A	1018	-	-	1/1/1/1	-
4	EDO	A	1019	-	-	1/1/1/1	-
6	PEG	A	1020	-	-	1/4/4/4	-
4	EDO	A	1012	-	-	0/1/1/1	-
3	W9P	C	1001	-	-	8/20/45/45	0/3/3/3
4	EDO	A	1013	-	-	0/1/1/1	-
4	EDO	C	1004	-	-	1/1/1/1	-
4	EDO	A	1003	-	-	1/1/1/1	-
4	EDO	C	1009	-	-	1/1/1/1	-
4	EDO	A	1017	-	-	1/1/1/1	-
4	EDO	A	1002	-	-	0/1/1/1	-
4	EDO	C	1010	-	-	0/1/1/1	-
5	PGE	A	1007	-	-	5/7/7/7	-
5	PGE	A	1006	-	-	3/7/7/7	-
4	EDO	A	1016	-	-	1/1/1/1	-
4	EDO	C	1006	-	-	1/1/1/1	-
4	EDO	C	1014	-	-	0/1/1/1	-
4	EDO	C	1005	-	-	0/1/1/1	-
4	EDO	C	1012	-	-	0/1/1/1	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1001	W9P	O6-N3	10.56	1.40	1.22
3	C	1001	W9P	O6-N3	10.47	1.40	1.22
3	A	1001	W9P	C14-N2	4.21	1.48	1.37
3	C	1001	W9P	C14-N2	4.04	1.48	1.37

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	1001	W9P	C8-N1-C5	-5.84	105.94	114.20
3	A	1001	W9P	C17-N4-N5	4.39	126.10	117.19
3	C	1001	W9P	C17-N4-N5	4.10	125.52	117.19
3	A	1001	W9P	C18-C19-C14	-2.47	119.29	121.53
3	A	1001	W9P	C8-N1-C5	-2.15	111.15	114.20

There are no chirality outliers.

5 of 52 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1001	W9P	C19-C14-N2-C13
3	C	1001	W9P	C19-C14-N2-C13
5	C	1008	PGE	C1-C2-O2-C3
5	C	1008	PGE	O3-C5-C6-O4
3	C	1001	W9P	C9-C10-C11-C12

There are no ring outliers.

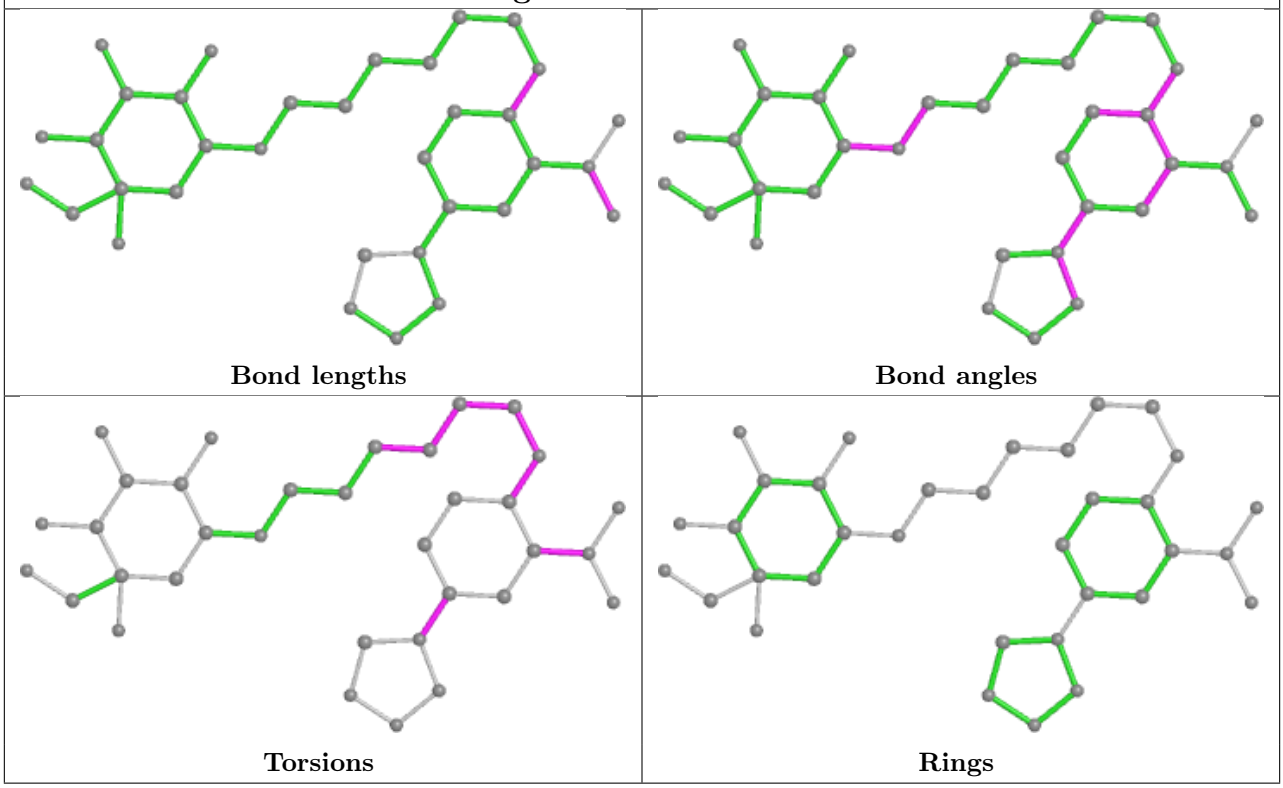
11 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	1011	EDO	1	0
5	C	1008	PGE	1	0
4	A	1004	EDO	1	0
6	A	1020	PEG	1	0
4	A	1003	EDO	1	0
4	C	1009	EDO	1	0
4	A	1002	EDO	1	0
5	A	1007	PGE	3	0
5	A	1006	PGE	2	0
4	C	1006	EDO	2	0
4	C	1014	EDO	1	0

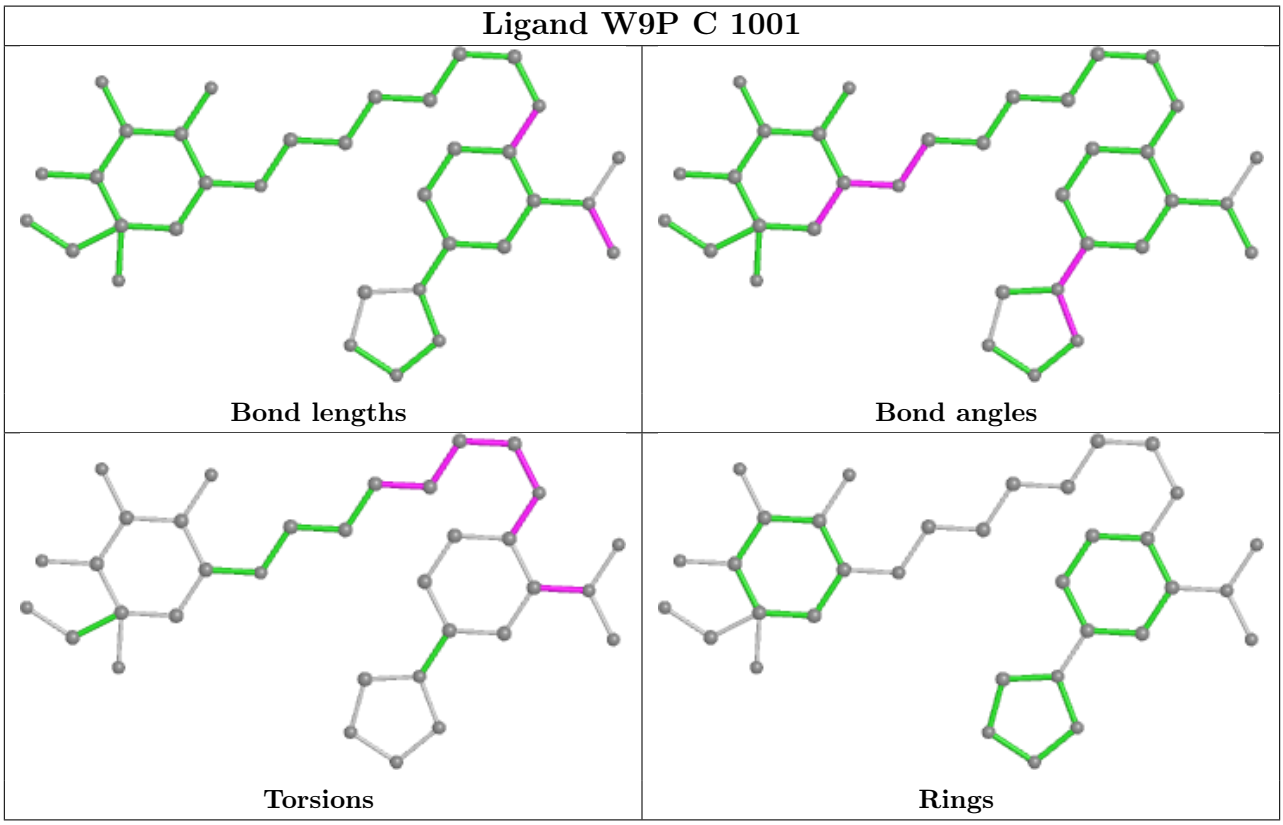
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 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.

Ligand W9P A 1001



Ligand W9P C 1001



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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

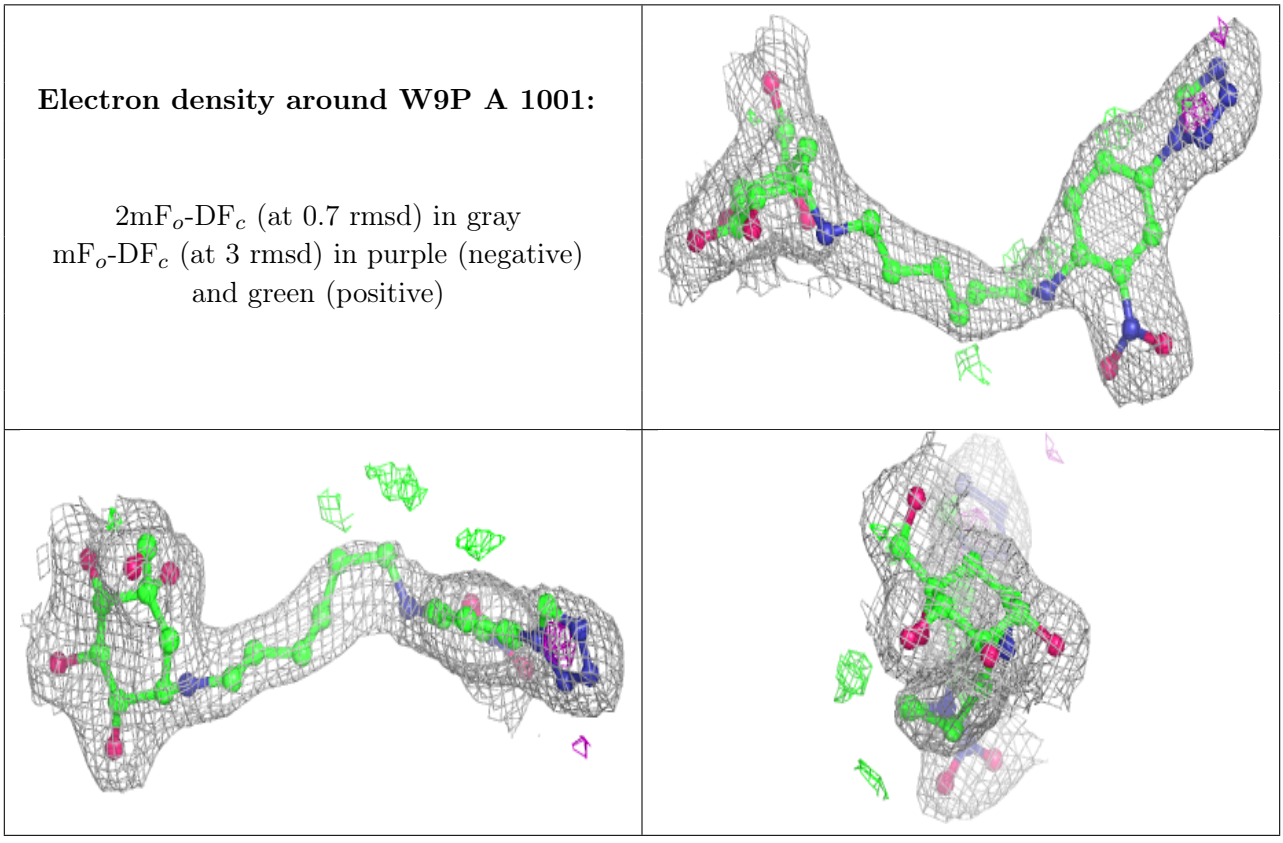
6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

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.

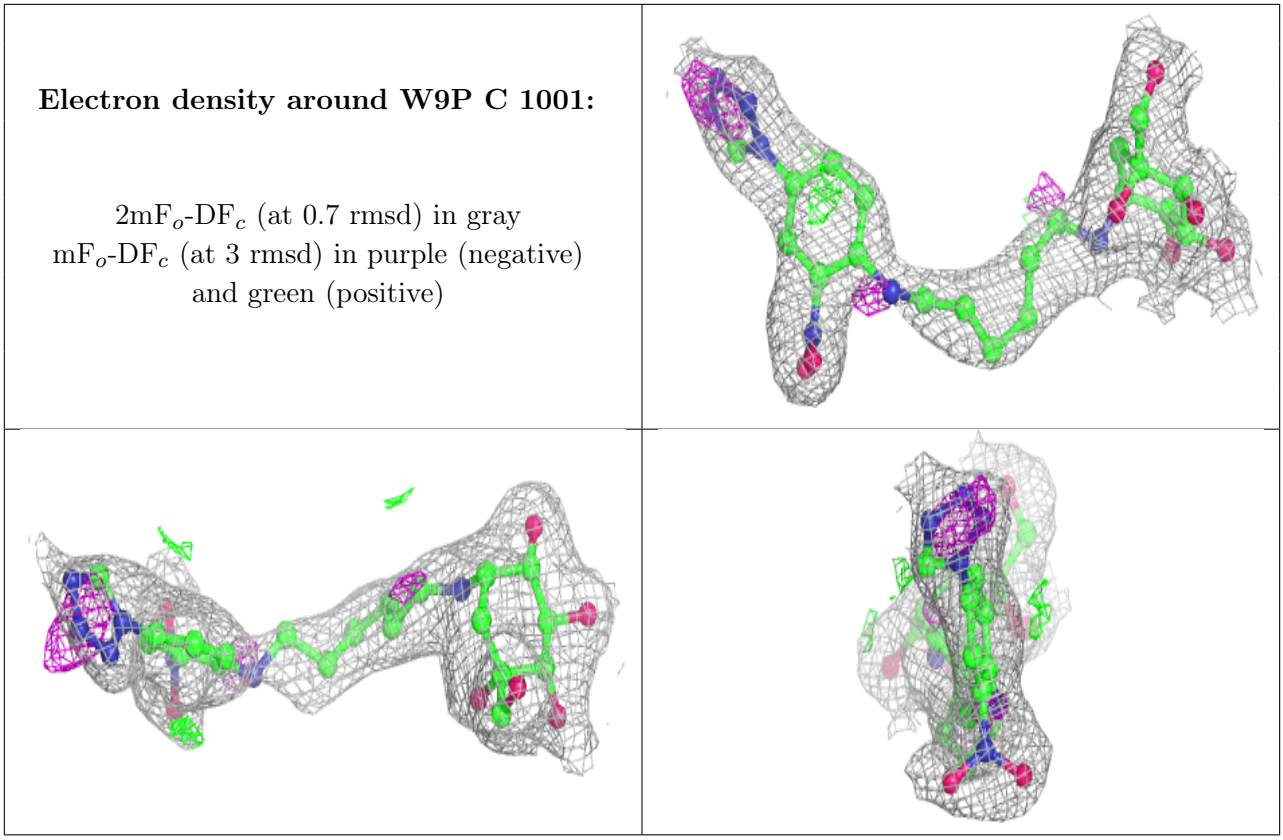
Electron density around W9P A 1001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



Electron density around W9P C 1001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.