Full wwPDB X-ray Structure Validation Report (i)

Jan 16, 2024 – 12:42 am GMT

PDB ID	:	8C75
Title	:	PBP AccA from A. tumefaciens Bo542 in apoform 3
Authors	:	Morera, S.; Vigouroux, A.; Legrand, P.
Deposited on	:	2023-01-12
Resolution	:	1.67 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 (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 (1)) were used in the production of this report:

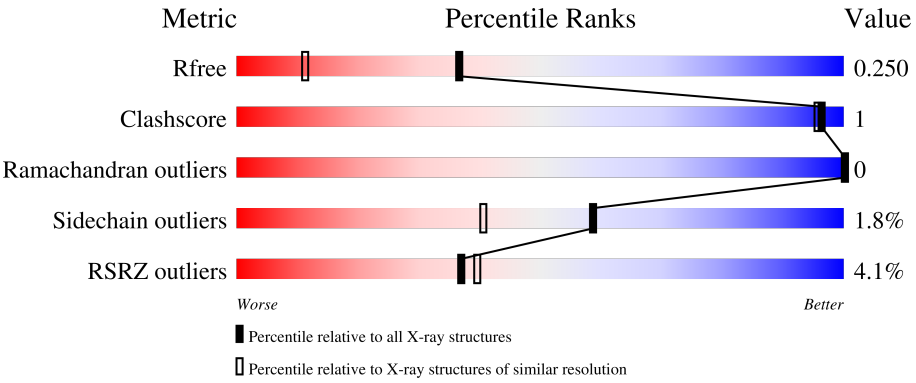
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.67 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	6780 (1.70-1.66)
Clashscore	141614	7310 (1.70-1.66)
Ramachandran outliers	138981	7173 (1.70-1.66)
Sidechain outliers	138945	7172 (1.70-1.66)
RSRZ outliers	127900	6661 (1.70-1.66)

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	498	4% 94% 5%
1	B	498	4% 95%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 16305 atoms, of which 7732 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 Agrocinopine utilization periplasmic binding protein AccA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	491	Total	C	H	N	O	S	0	1	0
			7758	2501	3849	669	725	14			
1	B	491	Total	C	H	N	O	S	0	2	0
			7768	2504	3853	670	727	14			

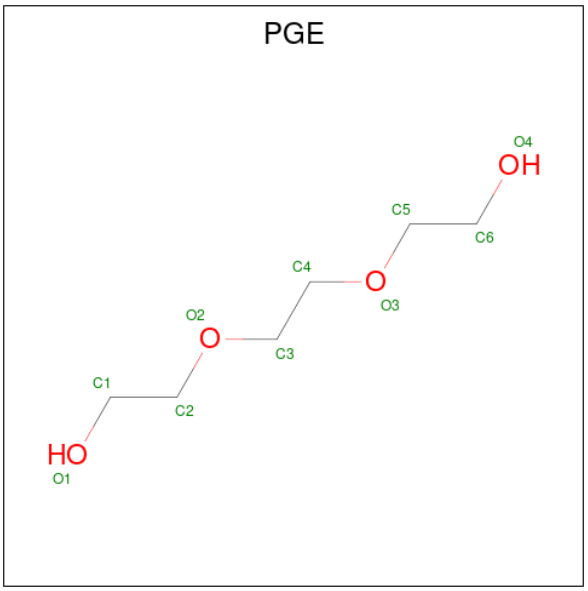
There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	30	MET	-	initiating methionine	UNP A0A2P0QK24
A	522	HIS	-	expression tag	UNP A0A2P0QK24
A	523	HIS	-	expression tag	UNP A0A2P0QK24
A	524	HIS	-	expression tag	UNP A0A2P0QK24
A	525	HIS	-	expression tag	UNP A0A2P0QK24
A	526	HIS	-	expression tag	UNP A0A2P0QK24
A	527	HIS	-	expression tag	UNP A0A2P0QK24
B	30	MET	-	initiating methionine	UNP A0A2P0QK24
B	522	HIS	-	expression tag	UNP A0A2P0QK24
B	523	HIS	-	expression tag	UNP A0A2P0QK24
B	524	HIS	-	expression tag	UNP A0A2P0QK24
B	525	HIS	-	expression tag	UNP A0A2P0QK24
B	526	HIS	-	expression tag	UNP A0A2P0QK24
B	527	HIS	-	expression tag	UNP A0A2P0QK24

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

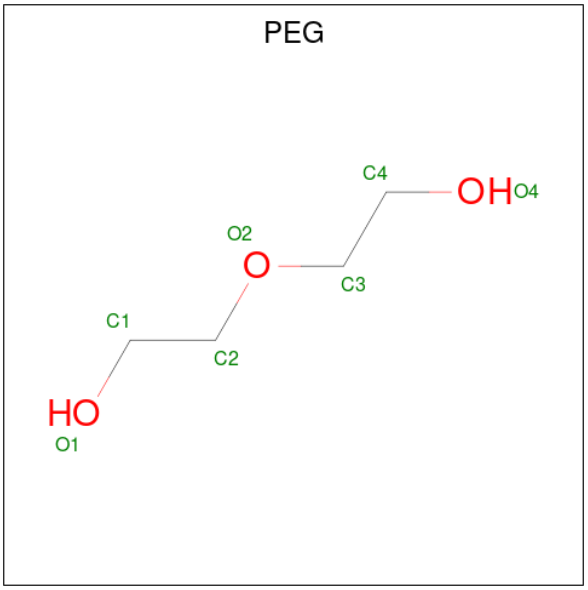
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Na	0	0
			1	1		
2	B	2	Total	Na	0	0
			2	2		

- Molecule 3 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



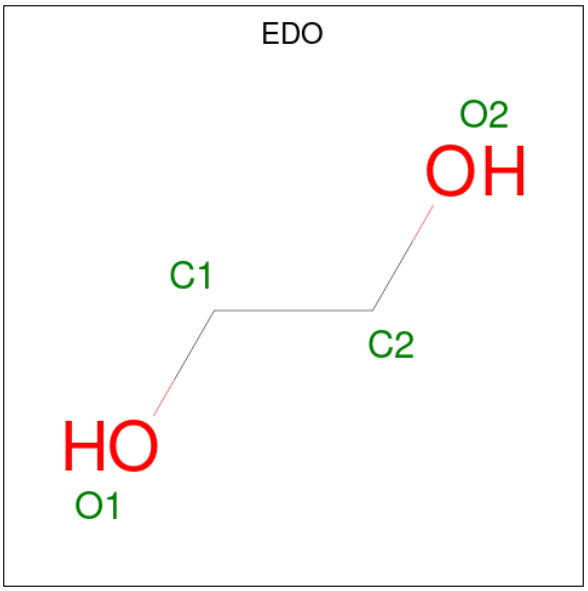
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			24	6	14	4		

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	H	O	0	0
			17	4	10	3		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	H	O	0	0
			10	2	6	2		

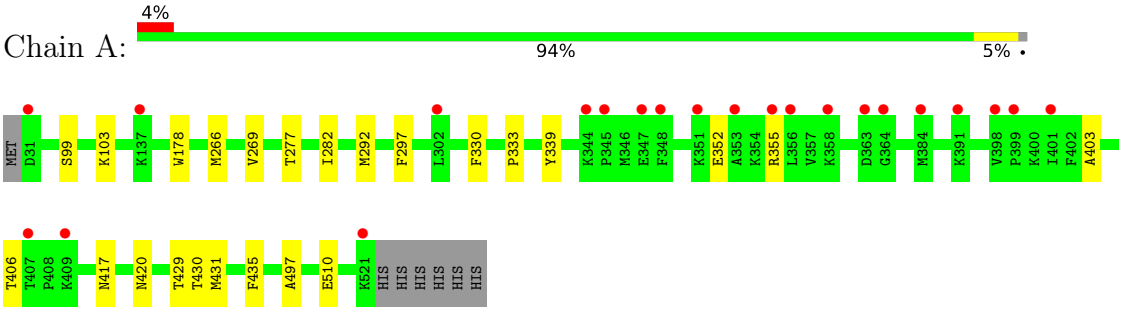
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	341	Total	O	0	0
			341	341		
6	B	384	Total	O	0	0
			384	384		

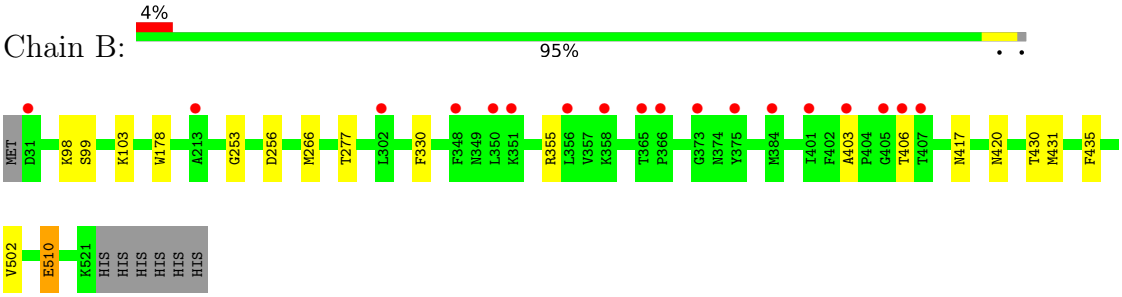
3 Residue-property plots (i)

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: Agrocinopine utilization periplasmic binding protein AccA



• Molecule 1: Agrocinopine utilization periplasmic binding protein AccA



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	46.99Å 182.07Å 58.61Å 90.00° 94.09° 90.00°	Depositor
Resolution (Å)	49.19 – 1.67 49.19 – 1.67	Depositor EDS
% Data completeness (in resolution range)	74.1 (49.19-1.67) 74.0 (49.19-1.67)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.65 (at 1.67Å)	Xtriage
Refinement program	BUSTER 2.10.4	Depositor
R, R_{free}	0.219 , 0.256 0.217 , 0.250	Depositor DCC
R_{free} test set	4208 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	22.2	Xtriage
Anisotropy	0.030	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 33.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	16305	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.19% 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: NA, PEG, EDO, 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	0.48	0/4011	0.68	0/5449
1	B	0.49	0/4017	0.68	0/5457
All	All	0.48	0/8028	0.68	0/10906

There are no bond length outliers.

There are no bond angle outliers.

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	3909	3849	3848	9	0
1	B	3915	3853	3851	7	0
2	A	1	0	0	0	0
2	B	2	0	0	0	0
3	A	10	14	14	0	0
4	B	7	10	10	0	0
5	B	4	6	6	0	0
6	A	341	0	0	0	0
6	B	384	0	0	2	0
All	All	8573	7732	7729	16	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 1.

All (16) 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:B:510:GLU:HG3	6:B:714:HOH:O	2.07	0.54
1:A:403:ALA:O	1:A:406:THR:OG1	2.28	0.51
1:B:256:ASP:HB3	1:B:502:VAL:HG21	1.93	0.51
1:B:403:ALA:O	1:B:406:THR:OG1	2.28	0.51
1:B:253:GLY:HA2	6:B:819:HOH:O	2.11	0.50
1:A:282:ILE:HD12	1:A:510:GLU:HG3	1.92	0.50
1:A:269:VAL:HG11	1:A:497:ALA:HB1	1.96	0.48
1:A:429:THR:HA	1:A:430:THR:HA	1.87	0.43
1:B:266:MET:HA	1:B:277:THR:HG21	2.01	0.42
1:B:420:ASN:HB3	1:B:430:THR:HG21	2.01	0.42
1:B:431:MET:O	1:B:435:PHE:HB2	2.19	0.42
1:A:292:MET:HG2	1:A:297:PHE:O	2.19	0.42
1:A:333:PRO:HA	1:A:339:TYR:CG	2.55	0.42
1:A:431:MET:O	1:A:435:PHE:HB2	2.20	0.42
1:A:266:MET:HA	1:A:277:THR:HG21	2.02	0.41
1:A:420:ASN:HB3	1:A:430:THR:HG21	2.02	0.41

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	490/498 (98%)	473 (96%)	17 (4%)	0	100	100
1	B	491/498 (99%)	477 (97%)	14 (3%)	0	100	100
All	All	981/996 (98%)	950 (97%)	31 (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	414/420 (99%)	407 (98%)	7 (2%)	60	43
1	B	415/420 (99%)	407 (98%)	8 (2%)	57	38
All	All	829/840 (99%)	814 (98%)	15 (2%)	59	40

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

Mol	Chain	Res	Type
1	A	99	SER
1	A	103	LYS
1	A	178	TRP
1	A	330	PHE
1	A	352	GLU
1	A	355	ARG
1	A	417	ASN
1	B	98	LYS
1	B	99	SER
1	B	103	LYS
1	B	178	TRP
1	B	330	PHE
1	B	355	ARG
1	B	417	ASN
1	B	510	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	284	ASN
1	B	519	ASN

5.3.3 RNA (i)

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)

Of 6 ligands modelled in this entry, 3 are monoatomic - leaving 3 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	$\#\|Z\| > 2$
5	EDO	B	604	-	3,3,3	0.44	0	2,2,2	0.14	0
3	PGE	A	602	-	9,9,9	0.23	0	8,8,8	0.42	0
4	PEG	B	601	-	6,6,6	0.28	0	5,5,5	0.35	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
5	EDO	B	604	-	-	0/1/1/1	-
3	PGE	A	602	-	-	3/7/7/7	-
4	PEG	B	601	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	602	PGE	O2-C3-C4-O3
3	A	602	PGE	C1-C2-O2-C3
4	B	601	PEG	O1-C1-C2-O2
4	B	601	PEG	C1-C2-O2-C3
3	A	602	PGE	C6-C5-O3-C4

There are no ring outliers.

No monomer is involved in short contacts.

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	491/498 (98%)	0.21	22 (4%)	33	35	13, 27, 67, 83	1 (0%)
1	B	491/498 (98%)	0.16	18 (3%)	41	44	13, 26, 68, 90	1 (0%)
All	All	982/996 (98%)	0.19	40 (4%)	37	39	13, 26, 68, 90	2 (0%)

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	405	GLY	5.0
1	A	384	MET	4.9
1	B	350	LEU	4.2
1	B	356	LEU	4.2
1	B	406	THR	4.1
1	A	407	THR	3.9
1	B	31	ASP	3.9
1	A	401	ILE	3.8
1	A	345	PRO	3.8
1	B	407	THR	3.7
1	B	384	MET	3.5
1	B	358	LYS	3.2
1	A	356	LEU	3.2
1	B	403	ALA	3.1
1	B	401	ILE	2.9
1	B	351	LYS	2.9
1	A	353	ALA	2.8
1	A	409	LYS	2.7
1	A	398	VAL	2.6
1	A	351	LYS	2.6
1	A	347	GLU	2.6
1	A	31	ASP	2.5
1	B	302	LEU	2.5
1	A	302	LEU	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	137	LYS	2.5
1	B	213	ALA	2.4
1	A	355	ARG	2.4
1	B	366	PRO	2.3
1	A	399	PRO	2.3
1	A	348	PHE	2.3
1	A	358	LYS	2.3
1	A	364	GLY	2.2
1	A	391	LYS	2.2
1	A	344	LYS	2.1
1	A	521	LYS	2.1
1	B	375	TYR	2.1
1	A	363	ASP	2.1
1	B	365	THR	2.1
1	B	348	PHE	2.0
1	B	373	GLY	2.0

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 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, 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	PEG	B	601	7/7	0.61	0.16	56,56,56,56	0
3	PGE	A	602	10/10	0.74	0.14	47,50,51,51	0
5	EDO	B	604	4/4	0.83	0.11	34,35,37,37	0
2	NA	A	601	1/1	0.96	0.05	33,33,33,33	0
2	NA	B	603	1/1	0.98	0.08	32,32,32,32	0
2	NA	B	602	1/1	0.99	0.06	25,25,25,25	0

6.5 Other polymers (i)

There are no such residues in this entry.