

wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 12:28 AM EDT

PDB ID	:	50XK
Title	:	PepTSt in complex with dipeptide Ala-Gln
Authors	:	Martinez Molledo, M.; Quistgaard, E.M.; Loew, C.
Deposited on		
Resolution	:	2.38 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 (1)) were used in the production of this report:

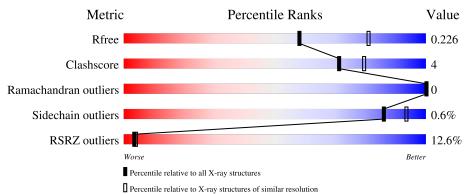
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 2.38 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R_{free}	130704	5509(2.40-2.36)
Clashscore	141614	6082(2.40-2.36)
Ramachandran outliers	138981	5973(2.40-2.36)
Sidechain outliers	138945	5975(2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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		
			11%		
1	А	490	81%	8%	11%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	GLN	А	519	-	-	-	Х



2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 3759 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.

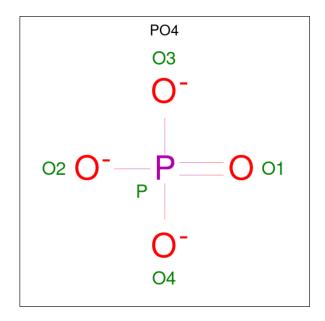
• Molecule 1 is a protein called Di-or tripeptide:H+ symporter.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	А	437	Total 3380	C 2284	N 515	O 564	S 17	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	484	ALA	-	expression tag	UNP Q5M4H8
А	485	GLU	-	expression tag	UNP Q5M4H8
A	486	ASN	-	expression tag	UNP Q5M4H8
А	487	LEU	-	expression tag	UNP Q5M4H8
А	488	TYR	-	expression tag	UNP Q5M4H8
А	489	PHE	-	expression tag	UNP Q5M4H8
А	490	GLN	-	expression tag	UNP Q5M4H8

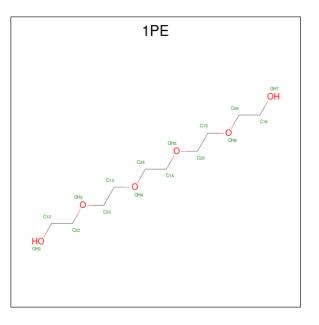
• Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{P} \\ 5 & 4 & 1 \end{array}$	0	0
2	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{P} \\ 5 & 4 & 1 \end{array}$	0	0
2	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{P} \\ 5 & 4 & 1 \end{array}$	0	0

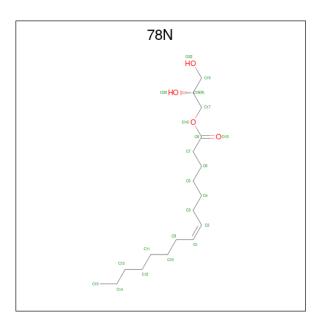
• Molecule 3 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	А	1	Total C 16 10	O 6	0	0

• Molecule 4 is (2R)-2,3-DIHYDROXYPROPYL(7Z)-PENTADEC-7-ENOATE (three-letter code: 78N) (formula: $C_{18}H_{34}O_4$).

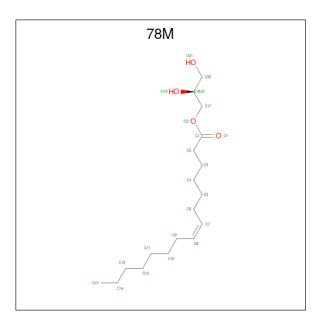




Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	1	Total C O 22 18 4	0	0
4	А	1	Total C O 22 18 4	0	0
4	А	1	Total C O 22 18 4	0	0
4	А	1	Total C O 22 18 4	0	0
4	А	1	Total C O 22 18 4	0	0
4	А	1	Total C O 22 18 4	0	0
4	А	1	Total C O 22 18 4	0	0
4	А	1	Total C O 22 18 4	0	0
4	А	1	Total C O 22 18 4	0	0
4	А	1	Total C O 22 18 4	0	0
4	А	1	Total C O 22 18 4	0	0
4	А	1	Total C O 22 18 4	0	0

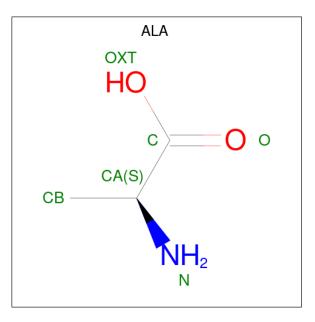
• Molecule 5 is (2S)-2,3-DIHYDROXYPROPYL(7Z)-PENTADEC-7-ENOATE (three-letter code: 78M) (formula: C₁₈H₃₄O₄).





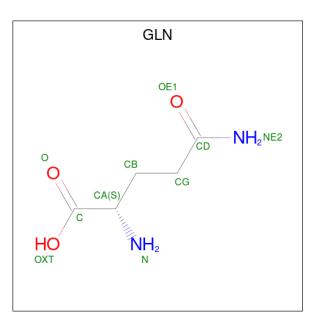
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	
5	А	1	Total	С	0	0	0
		-	22	18	4		5

• Molecule 6 is ALANINE (three-letter code: ALA) (formula: $C_3H_7NO_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	А	1	Total C N O 5 3 1 1	0	0





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	А	1	Total 10	C 5	~	O 3	0	0

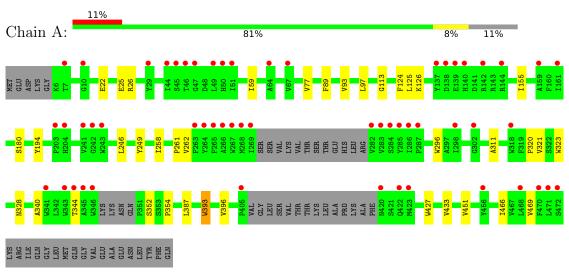
• Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	
8	А	47	$\begin{array}{cc} \text{Total} & \text{O} \\ 47 & 47 \end{array}$	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 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: Di-or tripeptide:H+ symporter



4 Data and refinement statistics (i)

Property	Value	Source	
Space group	C 2 2 21	Depositor	
Cell constants	100.70Å 110.20Å 105.20Å	Depositor	
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor	
Resolution (Å)	48.81 - 2.38	Depositor	
Resolution (A)	48.81 - 2.38	EDS	
% Data completeness	99.8 (48.81-2.38)	Depositor	
(in resolution range)	99.8 (48.81-2.38)	EDS	
R _{merge}	0.11	Depositor	
R _{sym}	(Not available)	Depositor	
$< I/\sigma(I) > 1$	$1.71 (at 2.37 \text{\AA})$	Xtriage	
Refinement program	PHENIX 1.9_1692	Depositor	
D D.	0.208 , 0.222	Depositor	
R, R_{free}	0.210 , 0.226	DCC	
R_{free} test set	1188 reflections (5.00%)	wwPDB-VP	
Wilson B-factor $(Å^2)$	47.1	Xtriage	
Anisotropy	0.440	Xtriage	
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.34 , 60.5	EDS	
L-test for twinning ²	$ \langle L \rangle = 0.49, \langle L^2 \rangle = 0.33$	Xtriage	
Estimated twinning fraction	No twinning to report.	Xtriage	
F_o, F_c correlation	0.94	EDS	
Total number of atoms	3759	wwPDB-VP	
Average B, all atoms $(Å^2)$	61.0	wwPDB-VP	

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.91% of the height of the origin peak. No significant pseudotranslation is detected.

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.



¹Intensities estimated from amplitudes.

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: $78\mathrm{N},$ 1PE, $78\mathrm{M},$ PO4

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).

Mal	Chain	Bond	lengths	Bond	angles
	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5
1	А	0.26	0/3484	0.41	0/4754

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 (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	А	3380	0	3446	24	0
2	А	15	0	0	0	0
3	А	16	0	22	0	0
4	А	264	0	408	10	0
5	А	22	0	34	1	0
6	А	5	0	4	1	0
7	А	10	0	8	0	0
8	А	47	0	0	0	0
All	All	3759	0	3922	27	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 27 close contacts within the same asymmetric unit are listed below, sorted by their



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:352:SER:HB2	1:A:354:PRO:HD2	1.84	0.60
1:A:59:ILE:HD11	1:A:246:LEU:HD21	1.84	0.59
1:A:262:VAL:HG11	4:A:514:78N:H112	1.85	0.58
1:A:451:VAL:HG11	4:A:505:78N:H192	1.89	0.55
1:A:249:TYR:CD1	4:A:512:78N:H72	2.43	0.54

clash magnitude.

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	А	429/490~(88%)	423~(99%)	6 (1%)	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.

Mo	Chain	Analysed	Rotameric	Outliers	Percentiles		
1	А	351/397~(88%)	349~(99%)	2(1%)	86 93		

All (2) residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	А	22	GLU
1	А	393	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	А	420	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)

19 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	Turne	Chain	Res	Link	Bo	ond leng	ths	В	ond ang	les
10101	Type	Unam	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z > 2
4	78N	А	513	-	21,21,21	0.94	1 (4%)	22,22,22	0.95	1 (4%)
7	GLN	А	519	-	8,9,9	0.84	1 (12%)	10,11,11	1.30	2 (20%)
6	ALA	А	518	-	$3,\!4,\!5$	0.61	0	2,4,6	0.79	0
2	PO4	А	501	-	4,4,4	0.88	0	$6,\!6,\!6$	0.64	0
4	78N	А	516	-	21,21,21	0.97	1 (4%)	22,22,22	0.95	1 (4%)
4	78N	А	506	-	21,21,21	0.96	1 (4%)	22,22,22	1.06	1 (4%)



Mol	Turne	Chain	Res	Link	Bo	ond leng	ths	B	ond ang	jles
NIOI	Type	Unam	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z > 2
5	78M	А	507	-	21,21,21	1.14	1 (4%)	22,22,22	1.03	1 (4%)
4	78N	А	515	-	21,21,21	0.94	1 (4%)	22,22,22	0.91	1 (4%)
4	78N	А	510	-	21,21,21	0.93	1 (4%)	22,22,22	1.05	1 (4%)
3	1PE	А	504	-	$15,\!15,\!15$	0.64	0	14,14,14	0.46	0
4	78N	А	505	-	21,21,21	0.94	1 (4%)	22,22,22	1.00	1 (4%)
4	78N	А	517	-	21,21,21	0.96	1 (4%)	22,22,22	1.00	1 (4%)
4	78N	А	509	-	21,21,21	0.95	1 (4%)	22,22,22	0.99	1 (4%)
4	78N	А	508	-	21,21,21	0.96	1 (4%)	22,22,22	0.94	1 (4%)
2	PO4	А	502	-	4,4,4	0.81	0	6,6,6	0.59	0
2	PO4	А	503	-	4,4,4	0.72	0	6,6,6	0.50	0
4	78N	А	514	-	21,21,21	0.98	1 (4%)	22,22,22	0.94	1 (4%)
4	78N	А	512	-	21,21,21	0.93	1 (4%)	22,22,22	0.98	1 (4%)
4	78N	А	511	-	21,21,21	0.94	1 (4%)	22,22,22	1.00	1 (4%)

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	78N	А	513	-	-	8/21/21/21	-
4	78N	А	511	-	-	9/21/21/21	-
4	78N	А	516	-	-	8/21/21/21	-
4	78N	А	506	-	-	12/21/21/21	-
4	78N	А	517	-	-	6/21/21/21	-
5	78M	А	507	-	-	9/21/21/21	-
4	78N	А	515	-	-	10/21/21/21	-
4	78N	А	510	-	-	15/21/21/21	-
7	GLN	А	519	-	-	3/9/9/9	-
3	1PE	А	504	-	-	12/13/13/13	-
4	78N	А	514	-	-	15/21/21/21	-
4	78N	А	512	-	-	11/21/21/21	-
6	ALA	А	518	-	-	0/0/2/4	-
4	78N	А	509	-	_	8/21/21/21	-
4	78N	А	508	-	-	10/21/21/21	_
4	78N	А	505	-	_	7/21/21/21	-



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	А	507	78M	O2-C1	3.34	1.43	1.33
4	А	506	78N	O16-C8	3.08	1.42	1.33
4	А	516	78N	O16-C8	2.99	1.42	1.33
4	А	514	78N	O16-C8	2.99	1.42	1.33
4	А	517	78N	O16-C8	2.97	1.42	1.33

The worst 5 of 14 bond length outliers are listed below:

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
7	А	519	GLN	OXT-C-O	-3.17	116.88	124.09
4	А	506	78N	O16-C8-C7	3.16	121.82	111.91
5	А	507	78M	O2-C1-C2	2.98	121.25	111.91
4	А	517	78N	O16-C8-C7	2.87	120.91	111.91
4	А	514	78N	O16-C8-C7	2.82	120.76	111.91

There are no chirality outliers.

5 of 143 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	А	506	78N	C17-C18-C19-O22
4	А	506	78N	O16-C17-C18-C19
4	А	506	78N	O16-C17-C18-O20
4	А	508	78N	C17-C18-C19-O22
4	А	509	78N	O16-C17-C18-O20

There are no ring outliers.

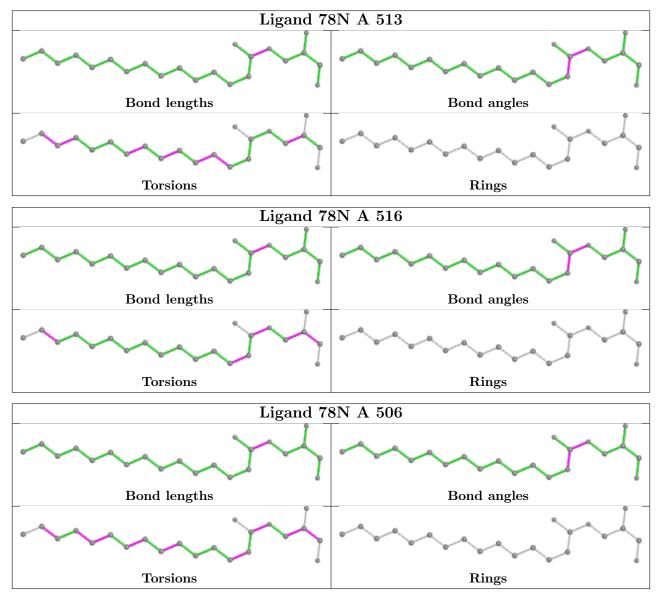
10 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	А	518	ALA	1	0
4	А	516	78N	2	0
5	А	507	78M	1	0
4	А	510	78N	2	0
4	А	505	78N	1	0
4	А	509	78N	1	0
4	А	508	78N	1	0
4	А	514	78N	1	0
4	А	512	78N	2	0
4	А	511	78N	1	0

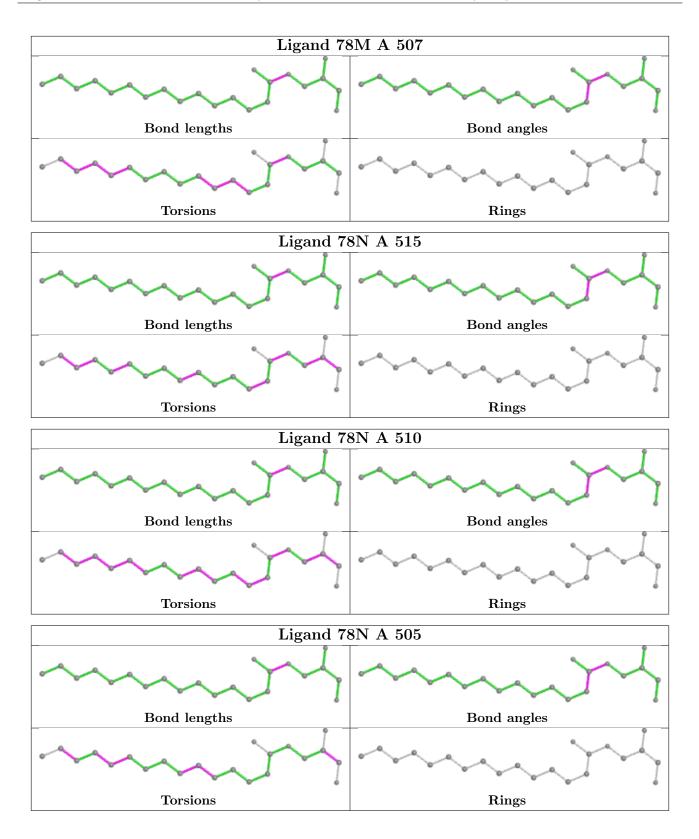
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,



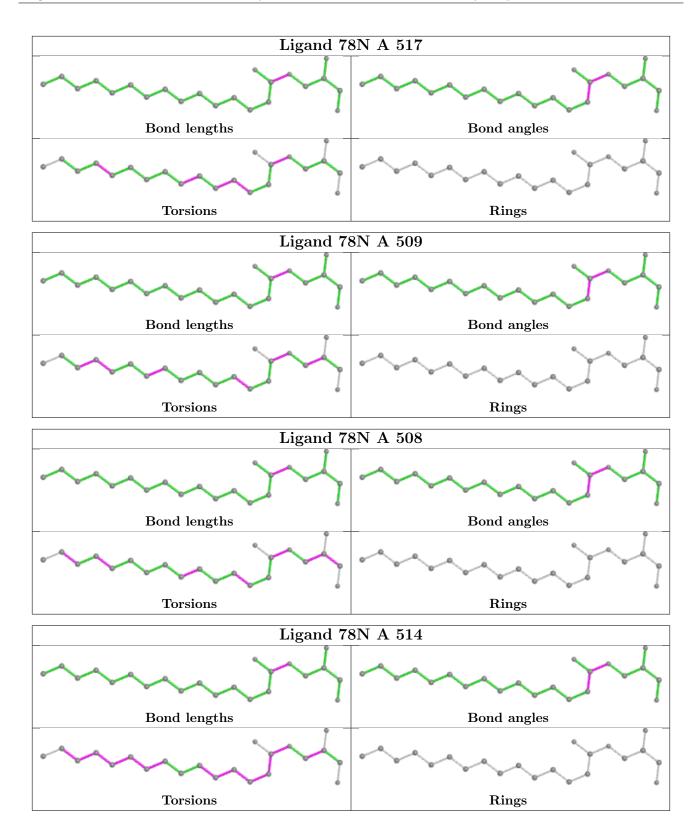
bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 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.



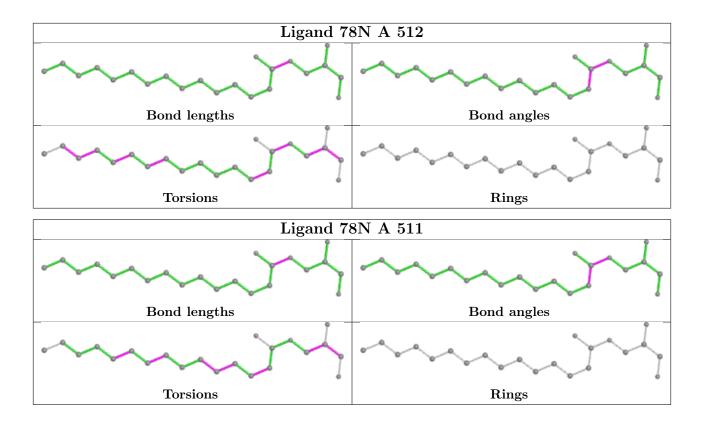












5.7 Other polymers (i)

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, 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	А	437/490~(89%)	0.75	55 (12%) 3 4	37, 51, 97, 133	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	267	TRP	9.5
1	А	269	ILE	7.7
1	А	471	LEU	7.6
1	А	142	ARG	7.5
1	А	283	VAL	7.5

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, 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	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q < 0.9
4	78N	А	514	22/22	0.57	0.32	85,92,96,101	0
4	78N	А	515	22/22	0.62	0.28	83,100,109,110	0

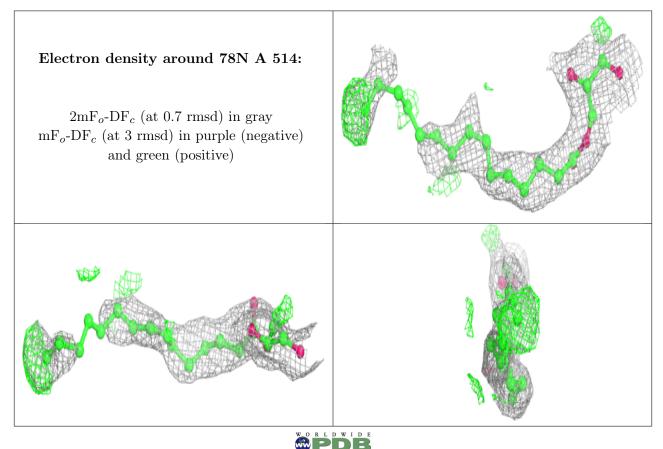
Continued on next page...

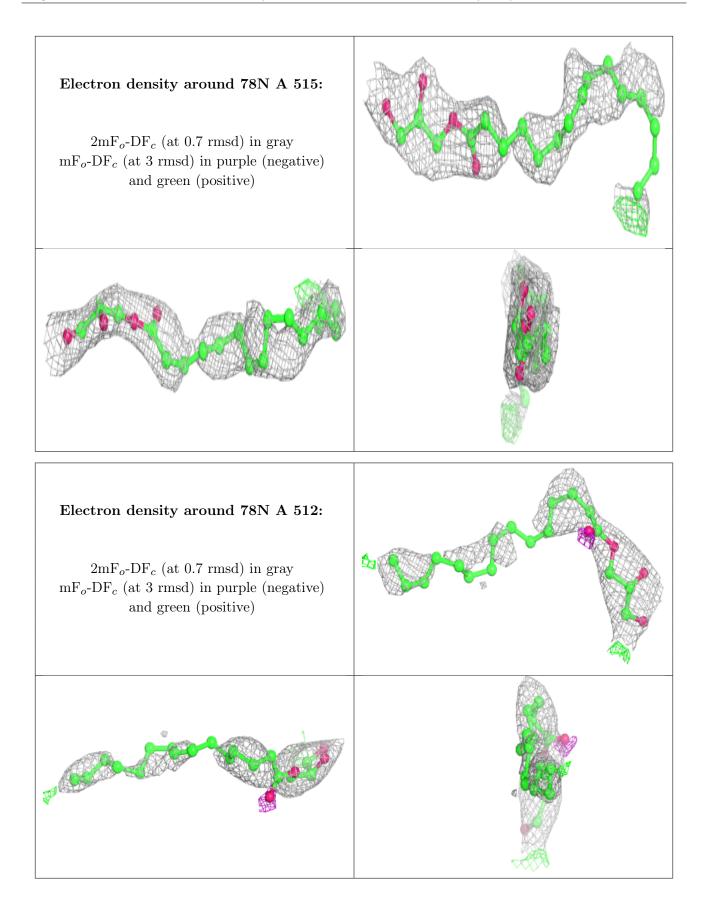


Mol	Type	m previoi Chain	Res	Atoms	RSCC	RSR	B -factors($Å^2$)	Q<0.9
3	1PE	А	504	16/16	0.66	0.25	82,87,92,94	0
4	78N	А	512	22/22	0.68	0.29	82,87,123,128	0
4	78N	А	511	22/22	0.74	0.34	73,90,128,132	0
4	78N	А	510	22/22	0.75	0.24	79,90,95,96	0
4	78N	А	513	22/22	0.77	0.29	68, 86, 100, 102	0
4	78N	А	517	22/22	0.78	0.36	71,86,108,109	0
7	GLN	А	519	10/10	0.78	0.45	78,90,114,115	0
5	78M	А	507	22/22	0.79	0.33	69,81,94,94	0
4	78N	А	506	22/22	0.79	0.36	72,92,105,108	0
4	78N	А	509	22/22	0.81	0.28	80,87,91,95	0
4	78N	А	505	22/22	0.83	0.26	$61,\!75,\!87,\!88$	0
4	78N	А	508	22/22	0.83	0.22	$65,\!89,\!99,\!101$	0
4	78N	А	516	22/22	0.85	0.31	70,84,88,91	0
6	ALA	А	518	5/6	0.88	0.27	71,71,77,78	0
2	PO4	А	502	5/5	0.88	0.34	96,97,112,114	0
2	PO4	А	503	5/5	0.92	0.26	94,102,105,109	0
2	PO4	А	501	5/5	0.96	0.13	76,80,88,91	0

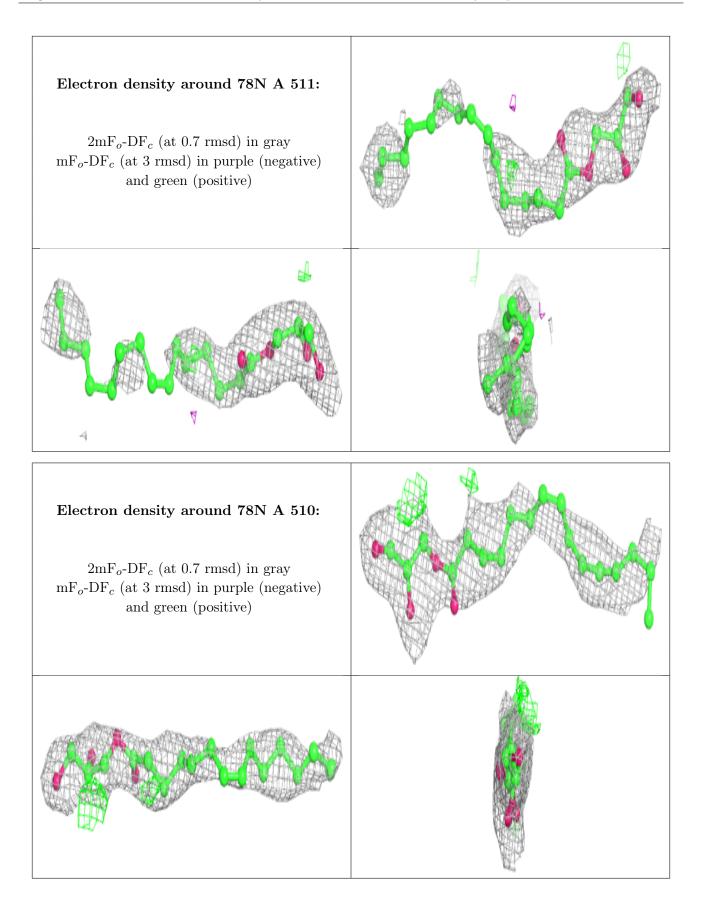
Continued from previous page...

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.

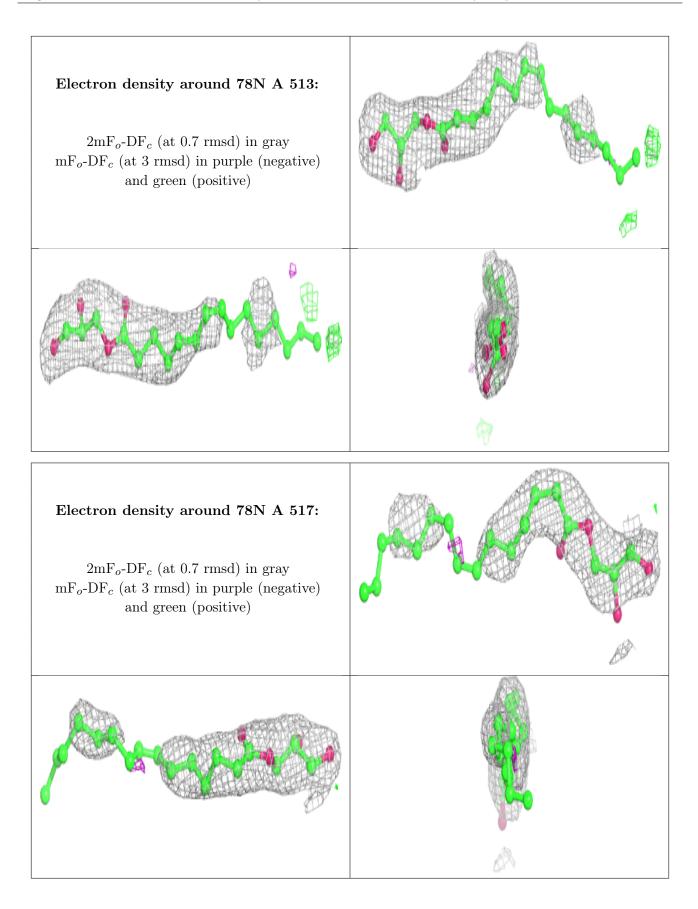




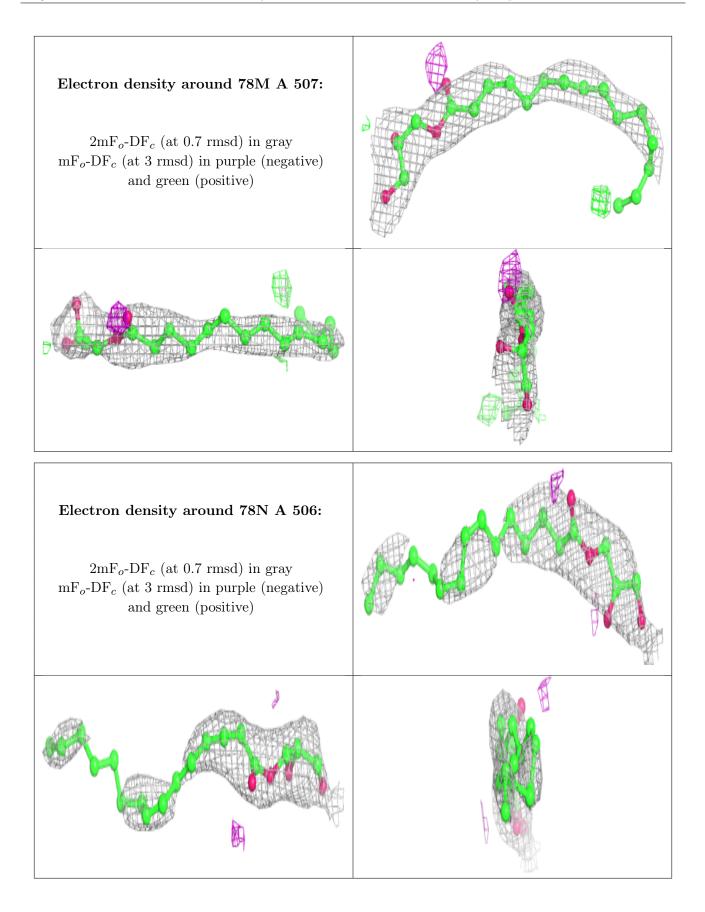




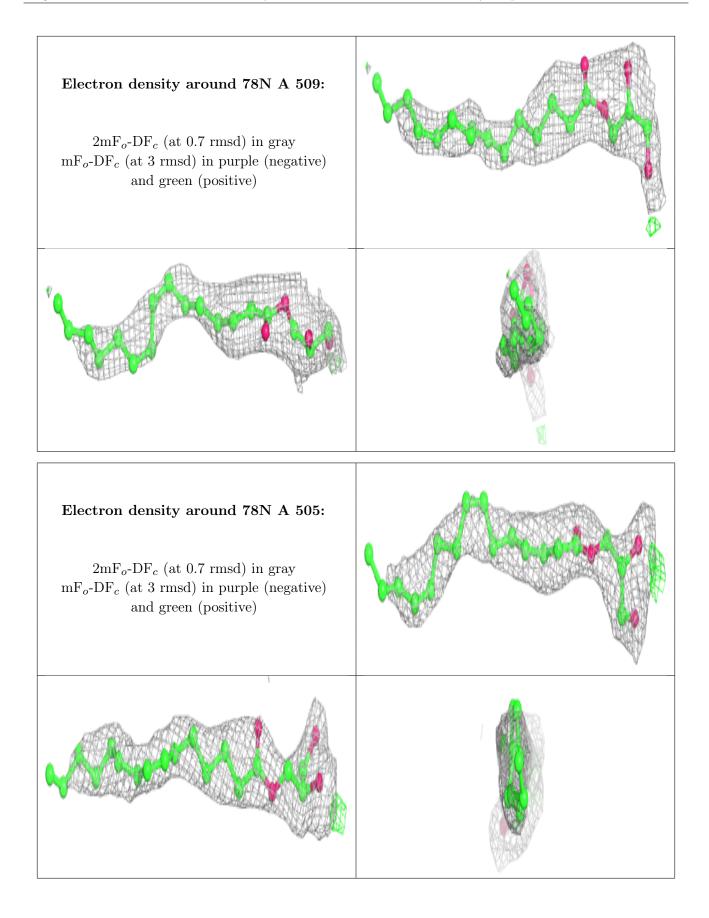




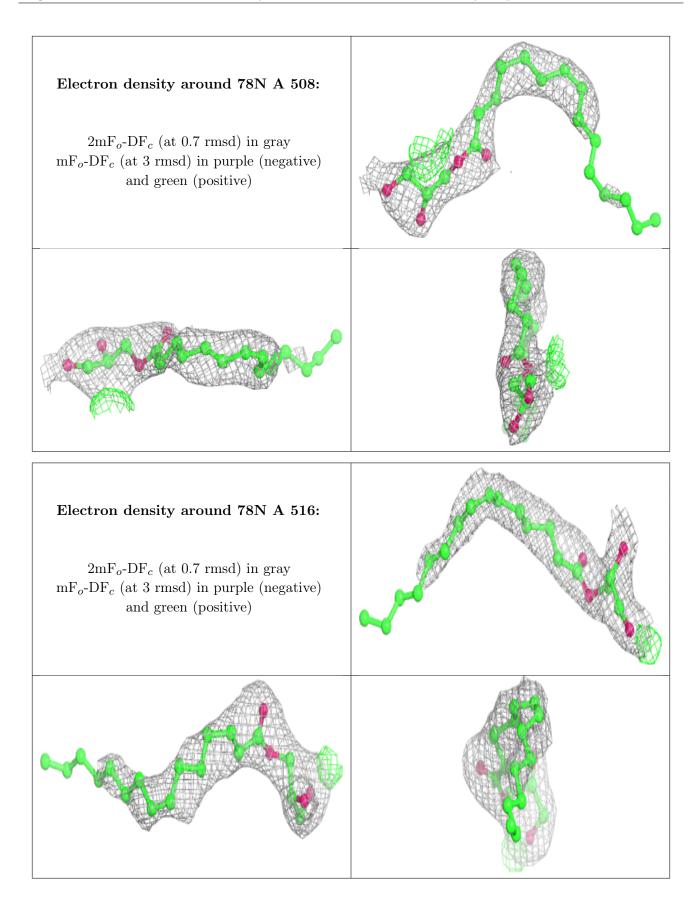














6.5 Other polymers (i)

There are no such residues in this entry.

