Full wwPDB X-ray Structure Validation Report

Sep 20, 2023 – 07:13 pm BST

PDB ID	:	8QKB
Title	:	Crystal structure of human cathepsin L in complex with the vinyl sulfone inhibitor K777
Authors	:	Falke, S.; Lieske, J.; Guenther, S.; Reinke, P.Y.A.; Ewert, W.; Loboda, J.; Karnicar, K.; Usenik, A.; Lindic, N.; Sekirnik, A.; Chapman, H.N.; Hinrichs, W.; Turk, D.; Meents, A.
Deposited on	:	2023-09-14
Resolution	:	1.60 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

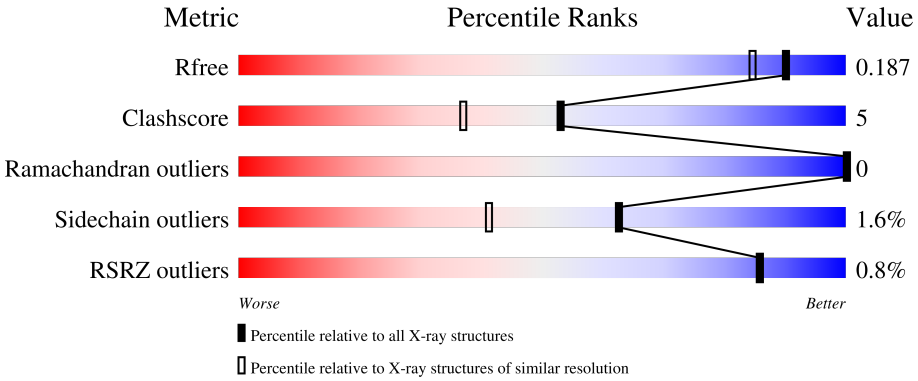
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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	220	90% 10%
1	B	220	% 88% 12%
1	C	220	% 87% 11% .
1	D	220	93% 7%

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
4	EDO	D	608	-	-	-	X

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 8065 atoms, of which 200 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.

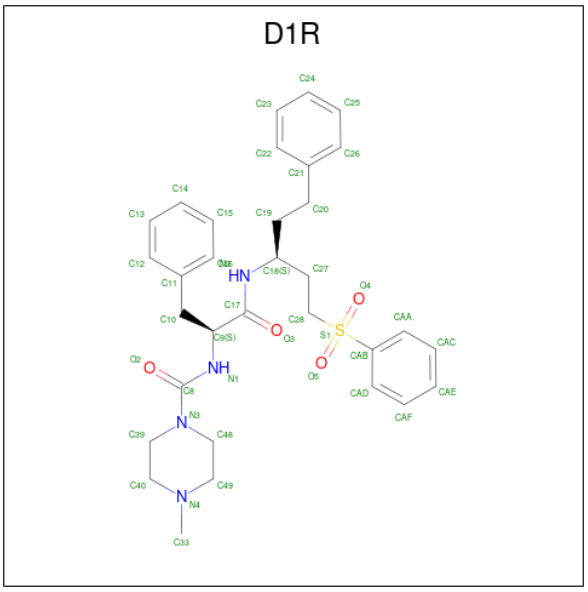
- Molecule 1 is a protein called Cathepsin L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	220	Total	C	N	O	S	0	2	0
			1709	1068	283	344	14			
1	B	220	Total	C	N	O	S	0	4	0
			1726	1079	285	347	15			
1	C	216	Total	C	N	O	S	0	2	0
			1676	1051	278	332	15			
1	D	220	Total	C	N	O	S	0	8	0
			1754	1094	290	354	16			

There are 4 discrepancies between the modelled and reference sequences:

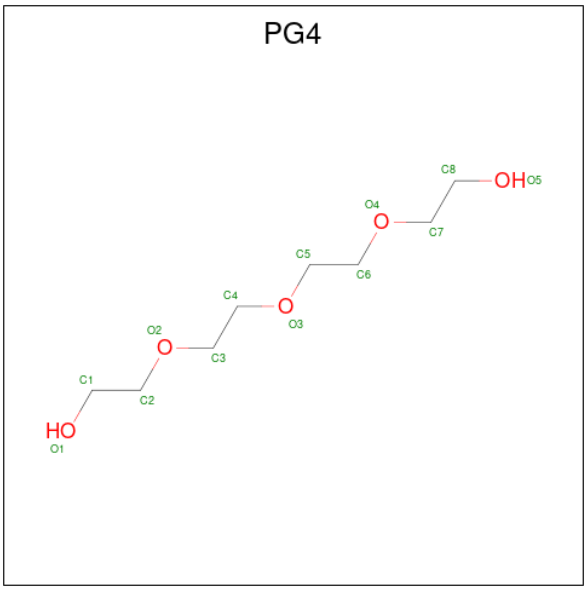
Chain	Residue	Modelled	Actual	Comment	Reference
A	110	ALA	THR	engineered mutation	UNP P07711
B	110	ALA	THR	engineered mutation	UNP P07711
C	110	ALA	THR	engineered mutation	UNP P07711
D	110	ALA	THR	engineered mutation	UNP P07711

- Molecule 2 is NALPHA-[(4-METHYLPIPERAZIN-1-YL)CARBONYL]-N-{(1S)-3-PHENYL-1-[2-(PHENYLSULFONYL)ETHYL]PROPYL}-L-PHENYLALANINAMIDE (three-letter code: D1R) (formula: $C_{32}H_{40}N_4O_4S$) (labeled as "Ligand of Interest" by depositor).



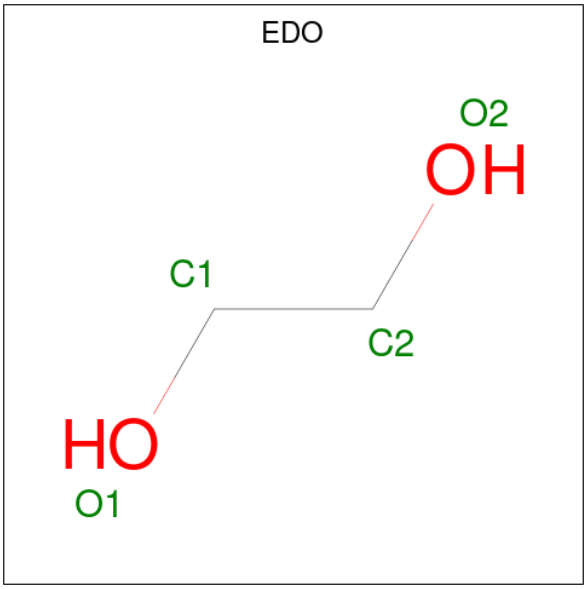
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			41	32	4	4	1		
2	B	1	Total	C	N	O	S	0	0
			41	32	4	4	1		
2	C	1	Total	C	N	O	S	0	0
			41	32	4	4	1		
2	D	1	Total	C	N	O	S	0	0
			41	32	4	4	1		

- Molecule 3 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 13 8 5	0	0

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



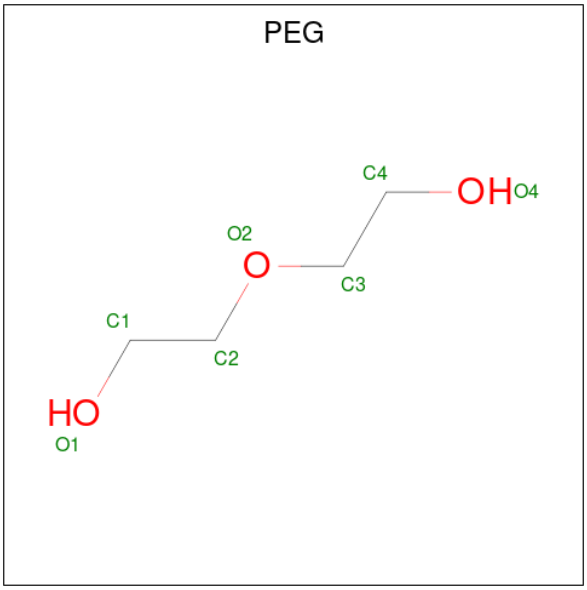
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C H O 10 2 6 2	0	0
4	A	1	Total C H O 10 2 6 2	0	0
4	A	1	Total C H O 10 2 6 2	0	0
4	A	1	Total C H O 10 2 6 2	0	0
4	B	1	Total C H O 10 2 6 2	0	0
4	B	1	Total C H O 10 2 6 2	0	0
4	B	1	Total C H O 10 2 6 2	0	0
4	B	1	Total C H O 10 2 6 2	0	0
4	B	1	Total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

Continued on next page...

Continued from previous page...

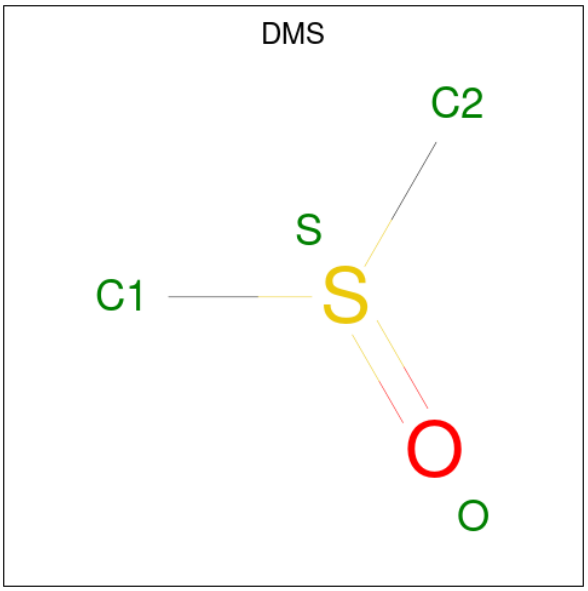
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	D	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



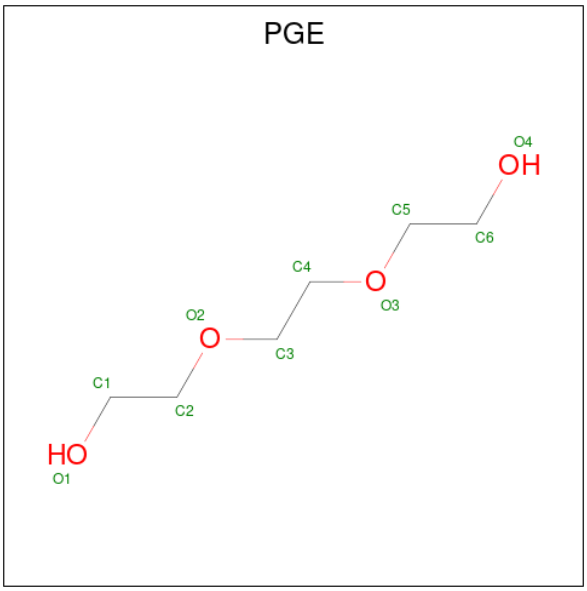
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	0
			17	4	10	3		
5	A	1	Total	C	H	O	0	0
			17	4	10	3		
5	C	1	Total	C	H	O	0	0
			17	4	10	3		
5	D	1	Total	C	H	O	0	0
			17	4	10	3		
5	D	1	Total	C	H	O	0	0
			17	4	10	3		

- Molecule 6 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).



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

- Molecule 7 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	1	Total C H O 24 6 14 4	0	0
7	D	1	Total C H O 24 6 14 4	0	0
7	D	1	Total C H O 24 6 14 4	0	0

- Molecule 8 is water.

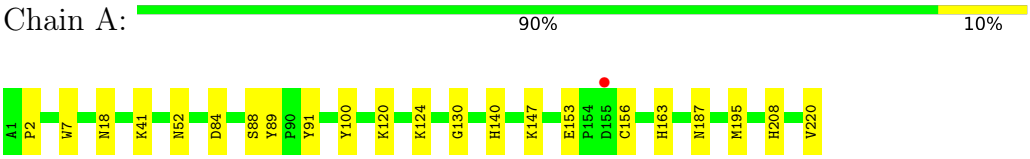
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	158	Total O 158 158	0	0
8	B	183	Total O 183 183	0	0
8	C	145	Total O 145 145	0	0
8	D	200	Total O 200 200	0	0

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.

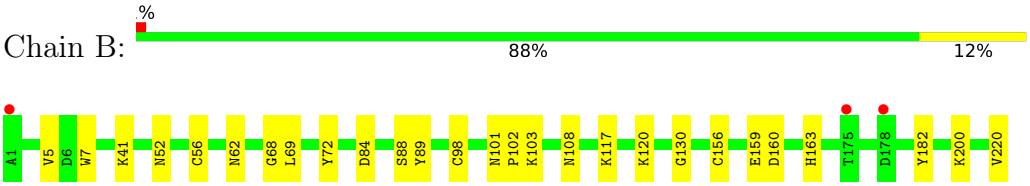
• Molecule 1: Cathepsin L

Chain A:



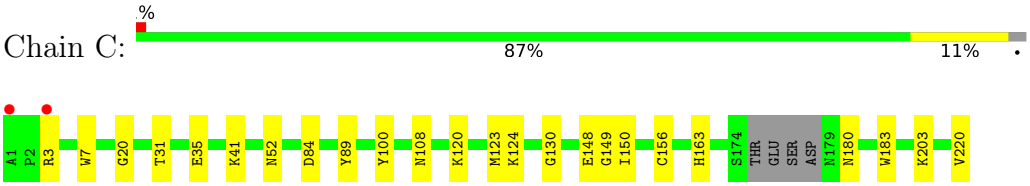
• Molecule 1: Cathepsin L

Chain B:



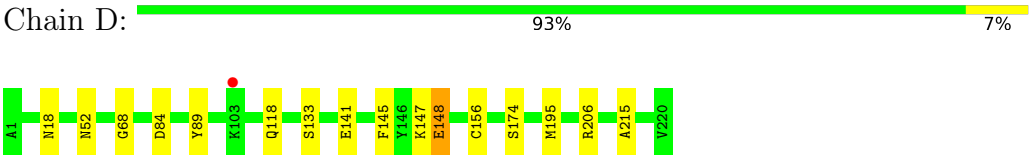
• Molecule 1: Cathepsin L

Chain C:



• Molecule 1: Cathepsin L

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	56.99Å 62.44Å 67.75Å 105.41° 93.65° 115.53°	Depositor
Resolution (Å)	44.19 – 1.60 44.19 – 1.60	Depositor EDS
% Data completeness (in resolution range)	91.8 (44.19-1.60) 91.8 (44.19-1.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.00 (at 1.60Å)	Xtriage
Refinement program	PHENIX 1.18_3861	Depositor
R, R_{free}	0.157 , 0.187 0.157 , 0.187	Depositor DCC
R_{free} test set	1366 reflections (1.42%)	wwPDB-VP
Wilson B-factor (Å^2)	13.4	Xtriage
Anisotropy	0.141	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 44.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	8065	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.78% 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: D1R, PEG, EDO, PGE, PG4, DMS

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	0.41	0/1751	0.57	0/2366
1	B	0.40	0/1768	0.59	0/2389
1	C	0.40	0/1720	0.57	0/2322
1	D	0.42	0/1796	0.59	0/2425
All	All	0.41	0/7035	0.58	0/9502

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	C	0	1
1	D	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	3	ARG	Sidechain
1	D	148[A]	GLU	Mainchain

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1709	0	1580	22	0
1	B	1726	0	1595	25	0
1	C	1676	0	1558	16	0
1	D	1754	0	1619	13	0
2	A	41	0	39	3	0
2	B	41	0	39	2	0
2	C	41	0	39	2	0
2	D	41	0	39	1	0
3	A	13	0	18	2	0
4	A	16	24	24	1	0
4	B	20	30	29	5	0
4	C	4	6	6	0	0
4	D	8	12	12	0	0
5	A	14	20	20	3	0
5	C	7	10	10	0	0
5	D	14	20	20	1	0
6	A	4	6	6	0	0
6	B	8	12	12	0	0
6	C	8	12	12	0	0
6	D	4	6	6	0	0
7	B	10	14	14	0	0
7	D	20	28	28	1	0
8	A	158	0	0	1	0
8	B	183	0	0	1	0
8	C	145	0	0	3	0
8	D	200	0	0	1	0
All	All	7865	200	6725	75	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 5.

All (75) 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:41:LYS:HD3	1:C:220:VAL:HG12	1.39	1.00

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:108:ASN:HD22	1:D:206:ARG:HH22	1.22	0.86
1:C:41:LYS:HD3	1:C:220:VAL:CG1	2.05	0.84
1:C:123:MET:HE2	8:C:423:HOH:O	1.84	0.78
1:A:153:GLU:H	1:A:208:HIS:CE1	2.01	0.78
1:A:195:MET:HE2	8:C:539:HOH:O	1.82	0.78
1:A:41:LYS:HD2	1:A:220:VAL:HG12	1.68	0.76
1:B:108:ASN:ND2	1:D:206:ARG:HH22	1.83	0.75
1:B:120:LYS:HD3	4:B:308:EDO:H21	1.69	0.75
1:B:5:VAL:HG13	4:B:306:EDO:H11	1.69	0.74
1:B:41:LYS:HD2	1:B:220:VAL:HG12	1.70	0.71
1:A:147:LYS:HD2	1:D:148[A]:GLU:HG3	1.72	0.71
1:B:68:GLY:HA2	2:B:301:D1R:H332	1.72	0.69
1:C:150:ILE:HD12	1:C:180:ASN:HB3	1.74	0.69
1:A:163:HIS:CD2	2:A:301:D1R:H281	2.29	0.68
1:A:153:GLU:H	1:A:208:HIS:HE1	1.41	0.68
1:A:163:HIS:HD2	2:A:301:D1R:H281	1.60	0.67
1:D:147:LYS:C	1:D:148[A]:GLU:HG2	2.17	0.64
1:C:150:ILE:CD1	1:C:180:ASN:HB3	2.27	0.64
1:D:147:LYS:O	1:D:148[A]:GLU:HG2	1.99	0.63
1:B:52:ASN:ND2	1:B:84:ASP:H	1.97	0.62
1:B:163:HIS:ND1	2:B:301:D1R:H281	2.16	0.61
1:D:68:GLY:HA2	2:D:602:D1R:H332	1.85	0.59
1:D:52:ASN:ND2	1:D:84:ASP:H	2.02	0.58
1:A:2:PRO:HB3	1:B:159:GLU:HB3	1.86	0.57
1:B:41:LYS:HD2	1:B:220:VAL:CG1	2.35	0.56
1:C:52:ASN:HD22	1:C:84:ASP:H	1.54	0.56
1:B:62:ASN:ND2	1:B:69:LEU:H	2.04	0.56
1:C:52:ASN:ND2	1:C:84:ASP:H	2.03	0.56
1:B:56:CYS:SG	1:B:98[B]:CYS:SG	3.04	0.55
1:B:117:LYS:O	1:B:117:LYS:HG2	2.07	0.55
1:C:183:TRP:CE2	1:C:203:LYS:HG3	2.42	0.55
1:C:163:HIS:ND1	2:C:301:D1R:H281	2.22	0.55
1:B:102:PRO:HG2	1:B:103:LYS:HE2	1.89	0.55
1:D:141:GLU:HG3	1:D:145:PHE:CE2	2.42	0.54
1:A:220:VAL:OXT	4:A:305:EDO:H11	2.08	0.54
1:A:52:ASN:ND2	1:A:84:ASP:H	2.06	0.53
1:B:62:ASN:HD21	1:B:69:LEU:H	1.54	0.53
1:D:52:ASN:HD22	1:D:84:ASP:H	1.55	0.52
1:D:195[B]:MET:HE1	8:D:740:HOH:O	2.07	0.52
1:B:52:ASN:HD22	1:B:84:ASP:H	1.58	0.52
1:A:7:TRP:CE2	1:A:130:GLY:HA2	2.44	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:163:HIS:HE1	1:A:187:ASN:OD1	1.93	0.52
1:D:18:ASN:HB3	7:D:605:PGE:H1	1.92	0.52
1:A:18:ASN:H	5:A:306:PEG:H41	1.76	0.51
1:C:148:GLU:HG2	1:C:149:GLY:N	2.26	0.51
1:D:147:LYS:O	1:D:148[A]:GLU:CG	2.60	0.49
1:A:52:ASN:HD22	1:A:84:ASP:H	1.60	0.48
1:C:7:TRP:CE2	1:C:130:GLY:HA2	2.48	0.48
1:A:88:SER:O	3:A:302:PG4:H82	2.14	0.47
1:A:2:PRO:HA	1:B:159:GLU:O	2.14	0.47
1:C:31:THR:O	1:C:35:GLU:HG3	2.15	0.46
1:B:69:LEU:HD12	1:B:72:TYR:CZ	2.50	0.46
3:A:302:PG4:H82	3:A:302:PG4:H61	1.63	0.45
4:B:308:EDO:H22	8:B:541:HOH:O	2.16	0.45
1:B:103:LYS:HE2	1:B:103:LYS:N	2.32	0.44
1:C:120:LYS:O	1:C:124:LYS:HG3	2.18	0.44
1:C:108:ASN:HB3	8:C:506:HOH:O	2.17	0.44
1:A:140:HIS:HD2	8:A:475:HOH:O	2.01	0.43
2:C:301:D1R:HAA	5:D:601:PEG:H31	2.01	0.42
1:B:7:TRP:CE2	1:B:130:GLY:HA2	2.54	0.42
1:A:120:LYS:HG2	1:A:124:LYS:HE3	2.01	0.42
1:B:69:LEU:HD12	1:B:72:TYR:CE2	2.55	0.42
1:C:52:ASN:HD21	1:C:100:TYR:HD1	1.67	0.41
1:A:52:ASN:HD22	1:A:84:ASP:HB2	1.85	0.41
1:A:147:LYS:O	1:A:195:MET:HG2	2.20	0.41
1:B:101:ASN:OD1	1:B:103:LYS:HG2	2.21	0.41
1:B:102:PRO:HG2	1:B:103:LYS:CE	2.50	0.41
1:B:88:SER:O	4:B:304:EDO:H21	2.21	0.41
2:A:301:D1R:HAA	5:A:308:PEG:H31	2.03	0.40
1:A:52:ASN:HD21	1:A:100:TYR:HD1	1.68	0.40
1:D:133:SER:O	1:D:215:ALA:HA	2.21	0.40
1:B:182:TYR:CZ	1:B:200:LYS:HE2	2.57	0.40
1:A:91:TYR:HB3	5:A:306:PEG:H32	2.02	0.40
4:B:308:EDO:H12	1:C:20:GLY:HA3	2.02	0.40

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	220/220 (100%)	215 (98%)	5 (2%)	0	100	100
1	B	222/220 (101%)	216 (97%)	6 (3%)	0	100	100
1	C	214/220 (97%)	208 (97%)	6 (3%)	0	100	100
1	D	226/220 (103%)	219 (97%)	7 (3%)	0	100	100
All	All	882/880 (100%)	858 (97%)	24 (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	180/178 (101%)	178 (99%)	2 (1%)	73	57
1	B	182/178 (102%)	179 (98%)	3 (2%)	62	41
1	C	176/178 (99%)	174 (99%)	2 (1%)	73	57
1	D	186/178 (104%)	182 (98%)	4 (2%)	52	27
All	All	724/712 (102%)	713 (98%)	11 (2%)	62	44

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	89	TYR
1	A	156	CYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	89	TYR
1	B	156	CYS
1	B	160	ASP
1	C	89	TYR
1	C	156	CYS
1	D	89	TYR
1	D	118	GLN
1	D	156	CYS
1	D	174	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	52	ASN
1	A	75	GLN
1	A	80	ASN
1	A	163	HIS
1	A	180	ASN
1	A	208	HIS
1	B	52	ASN
1	B	62	ASN
1	B	108	ASN
1	C	52	ASN
1	C	108	ASN
1	D	52	ASN
1	D	179	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

31 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	D1R	A	301	1	44,44,44	2.33	8 (18%)	56,59,59	2.37	13 (23%)
4	EDO	A	305	-	3,3,3	0.46	0	2,2,2	0.55	0
4	EDO	B	308	-	3,3,3	0.83	0	2,2,2	0.20	0
6	DMS	A	309	-	3,3,3	0.66	0	3,3,3	0.45	0
4	EDO	C	304	-	3,3,3	0.48	0	2,2,2	0.34	0
5	PEG	D	603	-	6,6,6	0.17	0	5,5,5	0.38	0
4	EDO	D	606	-	3,3,3	0.43	0	2,2,2	0.60	0
2	D1R	C	301	1	44,44,44	1.98	10 (22%)	56,59,59	2.48	15 (26%)
6	DMS	C	305	-	3,3,3	1.01	0	3,3,3	0.73	0
4	EDO	D	608	-	3,3,3	0.52	0	2,2,2	0.19	0
7	PGE	D	605	-	9,9,9	0.32	0	8,8,8	0.48	0
4	EDO	A	303	-	3,3,3	0.50	0	2,2,2	0.29	0
5	PEG	D	601	-	6,6,6	0.74	0	5,5,5	1.02	0
4	EDO	A	307	-	3,3,3	0.48	0	2,2,2	0.28	0
4	EDO	B	303	-	3,3,3	0.48	0	2,2,2	0.43	0
3	PG4	A	302	-	12,12,12	0.31	0	11,11,11	1.40	1 (9%)
5	PEG	C	302	-	6,6,6	0.13	0	5,5,5	0.23	0
2	D1R	D	602	1	44,44,44	2.16	11 (25%)	56,59,59	2.23	16 (28%)
4	EDO	A	304	-	3,3,3	0.52	0	2,2,2	0.30	0
6	DMS	D	604	-	3,3,3	0.63	0	3,3,3	0.68	0
5	PEG	A	308	-	6,6,6	0.14	0	5,5,5	0.30	0
6	DMS	C	303	-	3,3,3	0.63	0	3,3,3	0.53	0
4	EDO	B	304	-	3,3,3	0.45	0	2,2,2	0.42	0
4	EDO	B	307	-	3,3,3	0.54	0	2,2,2	0.24	0
7	PGE	D	607	-	9,9,9	0.28	0	8,8,8	0.33	0
6	DMS	B	309	-	3,3,3	0.61	0	3,3,3	0.25	0
4	EDO	B	306	-	3,3,3	0.45	0	2,2,2	0.31	0
6	DMS	B	305	-	3,3,3	0.65	0	3,3,3	0.61	0
2	D1R	B	301	1	44,44,44	2.27	11 (25%)	56,59,59	2.86	16 (28%)
7	PGE	B	302	-	9,9,9	0.30	0	8,8,8	0.43	0
5	PEG	A	306	-	6,6,6	0.63	0	5,5,5	0.91	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	D1R	A	301	1	-	6/37/47/47	0/4/4/4
4	EDO	A	305	-	-	1/1/1/1	-
4	EDO	B	308	-	-	1/1/1/1	-
4	EDO	C	304	-	-	1/1/1/1	-
5	PEG	D	603	-	-	2/4/4/4	-
4	EDO	D	606	-	-	0/1/1/1	-
2	D1R	C	301	1	-	5/37/47/47	0/4/4/4
4	EDO	D	608	-	-	0/1/1/1	-
7	PGE	D	605	-	-	3/7/7/7	-
4	EDO	A	303	-	-	1/1/1/1	-
5	PEG	D	601	-	-	2/4/4/4	-
4	EDO	A	307	-	-	1/1/1/1	-
4	EDO	B	303	-	-	0/1/1/1	-
3	PG4	A	302	-	-	6/10/10/10	-
5	PEG	C	302	-	-	1/4/4/4	-
2	D1R	D	602	1	-	5/37/47/47	0/4/4/4
4	EDO	A	304	-	-	0/1/1/1	-
5	PEG	A	308	-	-	1/4/4/4	-
4	EDO	B	304	-	-	0/1/1/1	-
4	EDO	B	307	-	-	1/1/1/1	-
7	PGE	D	607	-	-	1/7/7/7	-
4	EDO	B	306	-	-	0/1/1/1	-
2	D1R	B	301	1	-	8/37/47/47	0/4/4/4
7	PGE	B	302	-	-	3/7/7/7	-
5	PEG	A	306	-	-	3/4/4/4	-

All (40) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	301	D1R	CAB-S1	-9.89	1.60	1.76
2	B	301	D1R	O4-S1	-6.48	1.34	1.44
2	B	301	D1R	O5-S1	-6.28	1.34	1.44
2	D	602	D1R	O5-S1	-5.86	1.35	1.44
2	A	301	D1R	O5-S1	-5.65	1.35	1.44
2	A	301	D1R	C28-S1	-5.65	1.66	1.77
2	D	602	D1R	O3-C17	-4.92	1.13	1.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	301	D1R	O5-S1	-4.82	1.37	1.44
2	A	301	D1R	O4-S1	-4.71	1.37	1.44
2	D	602	D1R	C8-N3	4.69	1.45	1.36
2	B	301	D1R	C17-N2	4.69	1.44	1.34
2	D	602	D1R	O2-C8	-4.46	1.14	1.23
2	C	301	D1R	C17-N2	4.39	1.43	1.34
2	B	301	D1R	C8-N3	4.26	1.44	1.36
2	B	301	D1R	O3-C17	-4.25	1.14	1.23
2	C	301	D1R	O3-C17	-4.21	1.15	1.23
2	C	301	D1R	C8-N3	4.15	1.44	1.36
2	D	602	D1R	C17-N2	3.97	1.42	1.34
2	C	301	D1R	O2-C8	-3.93	1.15	1.23
2	D	602	D1R	O4-S1	-3.83	1.38	1.44
2	B	301	D1R	O2-C8	-3.79	1.16	1.23
2	B	301	D1R	C8-N1	3.51	1.43	1.35
2	C	301	D1R	C8-N1	3.42	1.42	1.35
2	C	301	D1R	C28-S1	3.30	1.84	1.77
2	D	602	D1R	C8-N1	3.29	1.42	1.35
2	D	602	D1R	CAD-CAB	-3.24	1.33	1.38
2	D	602	D1R	C48-N3	-3.21	1.41	1.47
2	A	301	D1R	CAD-CAB	-3.09	1.34	1.38
2	C	301	D1R	O4-S1	-3.01	1.39	1.44
2	B	301	D1R	C28-S1	3.01	1.83	1.77
2	A	301	D1R	O3-C17	-2.95	1.17	1.23
2	D	602	D1R	C28-S1	2.93	1.83	1.77
2	B	301	D1R	C49-N4	-2.86	1.40	1.46
2	C	301	D1R	C48-N3	-2.73	1.42	1.47
2	B	301	D1R	C18-N2	-2.60	1.41	1.46
2	C	301	D1R	CAD-CAB	-2.48	1.34	1.38
2	D	602	D1R	CAB-S1	2.31	1.80	1.76
2	A	301	D1R	O2-C8	-2.30	1.18	1.23
2	A	301	D1R	C48-N3	-2.23	1.43	1.47
2	B	301	D1R	C22-C21	-2.22	1.34	1.38

All (61) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	301	D1R	O5-S1-O4	-10.74	105.89	118.44
2	B	301	D1R	C48-C49-N4	9.73	121.80	110.80
2	C	301	D1R	O5-S1-O4	-9.56	107.26	118.44
2	B	301	D1R	O5-S1-O4	-8.49	108.51	118.44
2	C	301	D1R	C48-C49-N4	8.28	120.16	110.80

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	301	D1R	C33-N4-C40	8.01	122.65	110.66
2	B	301	D1R	O4-S1-CAB	7.25	115.73	108.36
2	D	602	D1R	C48-C49-N4	7.02	118.73	110.80
2	A	301	D1R	O5-S1-CAB	-6.42	101.84	108.36
2	D	602	D1R	O5-S1-O4	-5.78	111.68	118.44
2	D	602	D1R	C33-N4-C40	5.49	118.87	110.66
2	C	301	D1R	C33-N4-C49	-5.19	102.89	110.66
2	A	301	D1R	C28-S1-CAB	4.86	111.67	105.03
2	B	301	D1R	C39-C40-N4	4.83	116.26	110.80
2	D	602	D1R	O5-S1-CAB	-4.72	103.56	108.36
2	C	301	D1R	O4-S1-CAB	4.70	113.14	108.36
2	D	602	D1R	O4-S1-CAB	4.66	113.10	108.36
2	B	301	D1R	C49-C48-N3	4.57	120.23	110.44
2	C	301	D1R	C49-C48-N3	4.53	120.15	110.44
2	A	301	D1R	CAF-CAD-CAB	4.32	123.44	118.95
2	D	602	D1R	C49-C48-N3	4.25	119.55	110.44
2	C	301	D1R	C49-N4-C40	4.16	115.34	109.52
2	B	301	D1R	C28-S1-CAB	4.13	110.67	105.03
2	D	602	D1R	C28-S1-CAB	4.05	110.57	105.03
2	D	602	D1R	C27-C18-N2	3.74	115.80	110.54
2	C	301	D1R	O5-S1-CAB	-3.64	104.66	108.36
3	A	302	PG4	C3-O2-C2	3.62	128.99	113.29
2	B	301	D1R	O5-S1-CAB	-3.60	104.70	108.36
2	A	301	D1R	C27-C18-N2	3.36	115.27	110.54
2	B	301	D1R	C27-C28-S1	-3.33	104.87	114.29
2	A	301	D1R	C49-N4-C40	3.33	114.17	109.52
2	A	301	D1R	O5-S1-C28	3.29	114.10	108.20
2	A	301	D1R	C39-C40-N4	3.12	114.33	110.80
2	C	301	D1R	C39-C40-N4	3.11	114.31	110.80
2	A	301	D1R	N1-C8-N3	3.08	122.80	117.21
2	C	301	D1R	C28-S1-CAB	3.00	109.13	105.03
2	B	301	D1R	C26-C21-C22	2.87	122.68	118.17
2	B	301	D1R	C39-N3-C8	2.86	132.49	121.94
2	C	301	D1R	C33-N4-C40	2.86	114.94	110.66
2	A	301	D1R	C49-C48-N3	2.75	116.32	110.44
2	D	602	D1R	C27-C28-S1	-2.65	106.78	114.29
2	D	602	D1R	C39-C40-N4	2.65	113.80	110.80
2	B	301	D1R	C19-C18-N2	-2.60	106.88	110.54
2	D	602	D1R	CAC-CAA-CAB	-2.57	116.29	118.95
2	C	301	D1R	C27-C28-S1	-2.50	107.21	114.29
2	B	301	D1R	C27-C18-N2	2.42	113.95	110.54
2	A	301	D1R	C48-C49-N4	2.42	113.53	110.80

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	301	D1R	C49-N4-C40	2.39	112.87	109.52
2	D	602	D1R	C19-C18-N2	-2.36	107.22	110.54
2	D	602	D1R	CAF-CAD-CAB	2.34	121.39	118.95
2	A	301	D1R	O4-S1-C28	2.31	112.34	108.20
2	C	301	D1R	CAC-CAA-CAB	-2.27	116.60	118.95
2	B	301	D1R	C10-C9-C17	-2.26	104.38	110.25
2	C	301	D1R	O4-S1-C28	2.23	112.20	108.20
2	D	602	D1R	C48-N3-C39	-2.17	108.44	112.62
2	D	602	D1R	C17-C9-N1	-2.17	105.26	111.16
2	A	301	D1R	C19-C18-N2	-2.12	107.55	110.54
2	B	301	D1R	CAC-CAA-CAB	-2.11	116.77	118.95
2	D	602	D1R	C39-N3-C8	2.08	129.60	121.94
2	C	301	D1R	N1-C8-N3	2.07	120.96	117.21
2	C	301	D1R	C19-C18-N2	-2.01	107.72	110.54

There are no chirality outliers.

All (52) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	301	D1R	CAD-CAB-S1-C28
2	B	301	D1R	CAA-CAB-S1-O5
2	B	301	D1R	CAD-CAB-S1-O5
2	C	301	D1R	CAA-CAB-S1-O5
2	C	301	D1R	CAD-CAB-S1-O5
2	D	602	D1R	CAA-CAB-S1-O5
2	D	602	D1R	CAD-CAB-S1-O5
2	A	301	D1R	CAD-CAB-S1-O5
2	A	301	D1R	CAA-CAB-S1-C28
2	A	301	D1R	CAD-CAB-S1-C28
2	B	301	D1R	CAA-CAB-S1-C28
2	C	301	D1R	CAA-CAB-S1-C28
2	C	301	D1R	CAD-CAB-S1-C28
2	D	602	D1R	CAA-CAB-S1-C28
2	D	602	D1R	CAD-CAB-S1-C28
2	A	301	D1R	CAA-CAB-S1-O5
2	A	301	D1R	C18-C27-C28-S1
2	B	301	D1R	C18-C27-C28-S1
2	C	301	D1R	C18-C27-C28-S1
2	D	602	D1R	C18-C27-C28-S1
3	A	302	PG4	O2-C3-C4-O3
3	A	302	PG4	O3-C5-C6-O4
7	B	302	PGE	O3-C5-C6-O4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	D	603	PEG	O1-C1-C2-O2
5	A	306	PEG	O2-C3-C4-O4
3	A	302	PG4	C6-C5-O3-C4
4	B	307	EDO	O1-C1-C2-O2
5	A	308	PEG	O2-C3-C4-O4
5	A	306	PEG	O1-C1-C2-O2
4	A	303	EDO	O1-C1-C2-O2
4	B	308	EDO	O1-C1-C2-O2
4	C	304	EDO	O1-C1-C2-O2
3	A	302	PG4	O1-C1-C2-O2
5	D	603	PEG	C4-C3-O2-C2
7	D	605	PGE	C1-C2-O2-C3
7	B	302	PGE	C6-C5-O3-C4
5	A	306	PEG	C1-C2-O2-C3
7	B	302	PGE	C1-C2-O2-C3
2	A	301	D1R	C27-C18-C19-C20
5	D	601	PEG	C4-C3-O2-C2
3	A	302	PG4	O4-C7-C8-O5
5	D	601	PEG	O1-C1-C2-O2
4	A	307	EDO	O1-C1-C2-O2
2	B	301	D1R	C27-C18-C19-C20
7	D	607	PGE	C4-C3-O2-C2
2	B	301	D1R	C9-C10-C11-C16
2	B	301	D1R	C9-C10-C11-C12
4	A	305	EDO	O1-C1-C2-O2
7	D	605	PGE	C3-C4-O3-C5
5	C	302	PEG	O2-C3-C4-O4
3	A	302	PG4	C8-C7-O4-C6
7	D	605	PGE	C6-C5-O3-C4

There are no ring outliers.

13 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	D1R	3	0
4	A	305	EDO	1	0
4	B	308	EDO	3	0
2	C	301	D1R	2	0
7	D	605	PGE	1	0
5	D	601	PEG	1	0
3	A	302	PG4	2	0
2	D	602	D1R	1	0

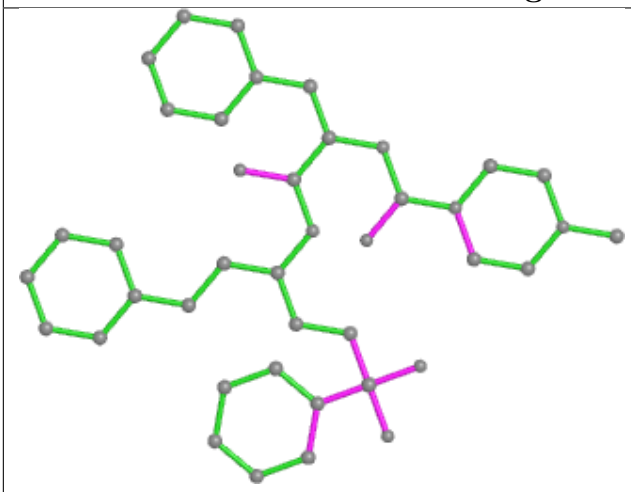
Continued on next page...

Continued from previous page...

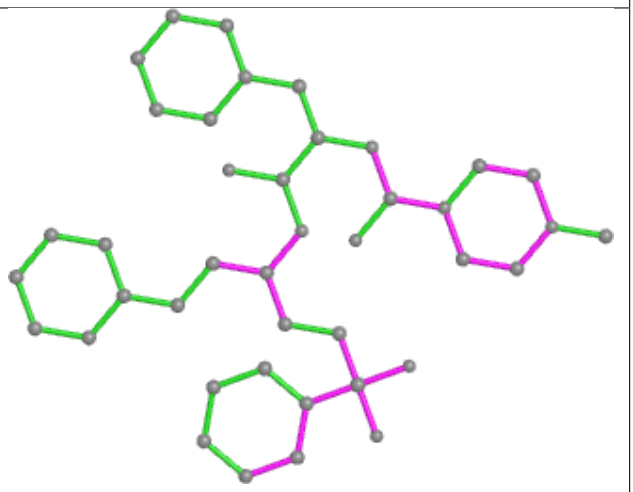
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	308	PEG	1	0
4	B	304	EDO	1	0
4	B	306	EDO	1	0
2	B	301	D1R	2	0
5	A	306	PEG	2	0

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 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.

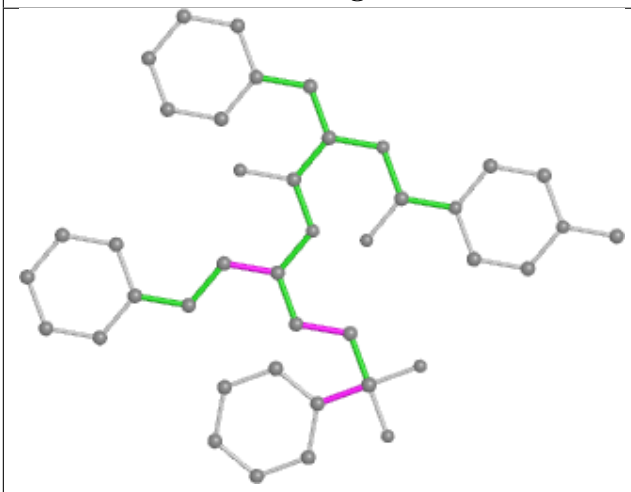
Ligand D1R A 301



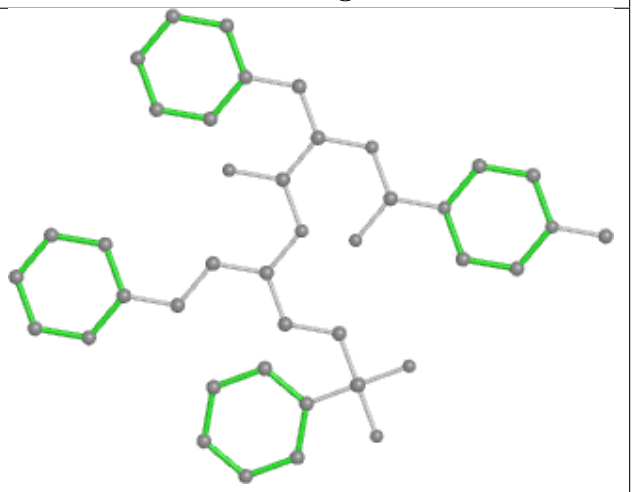
Bond lengths



Bond angles

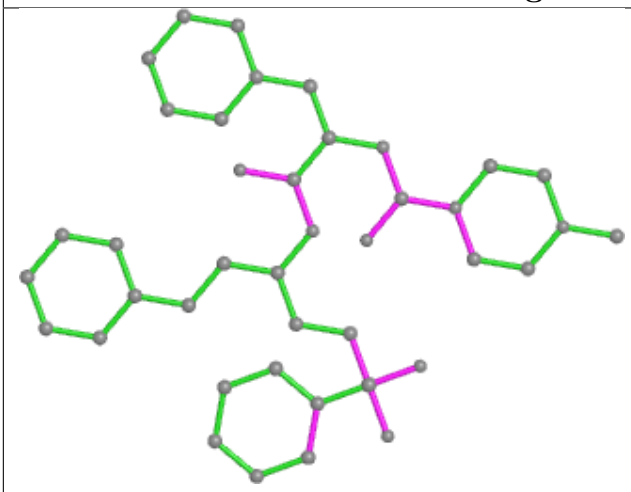


Torsions

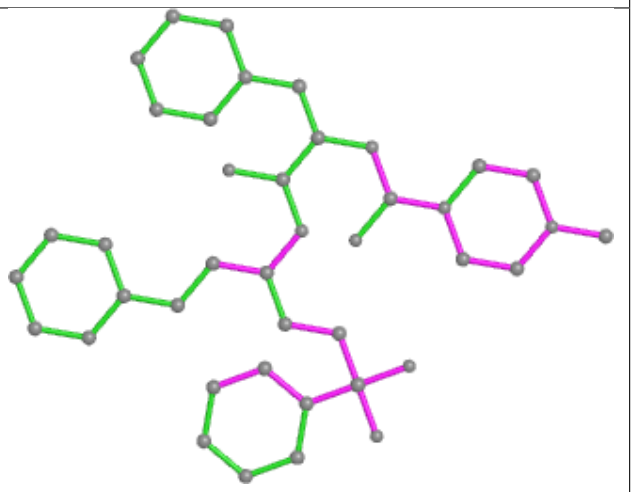


Rings

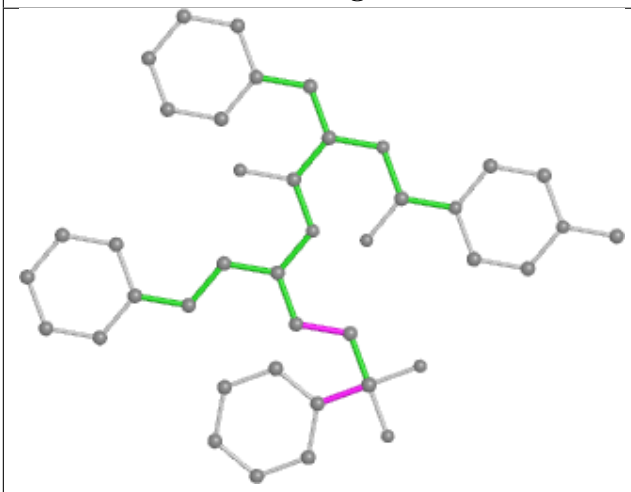
Ligand D1R C 301



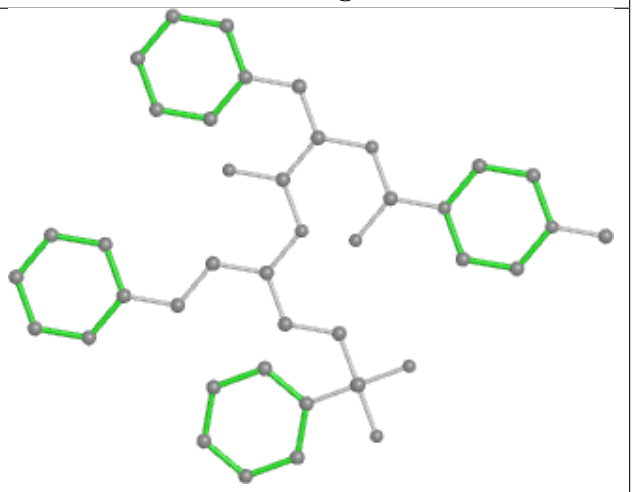
Bond lengths



Bond angles

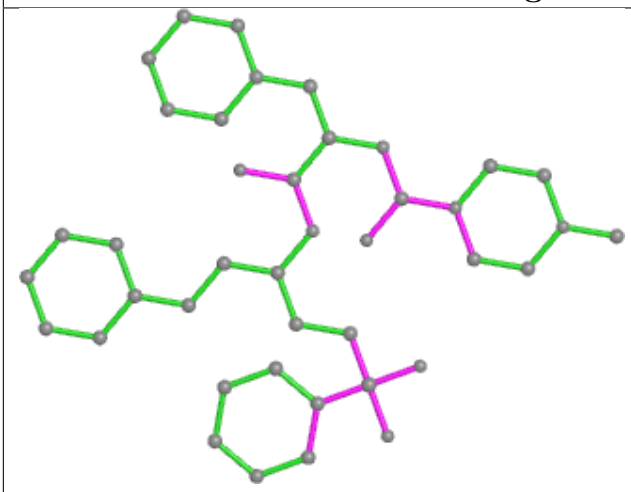


Torsions

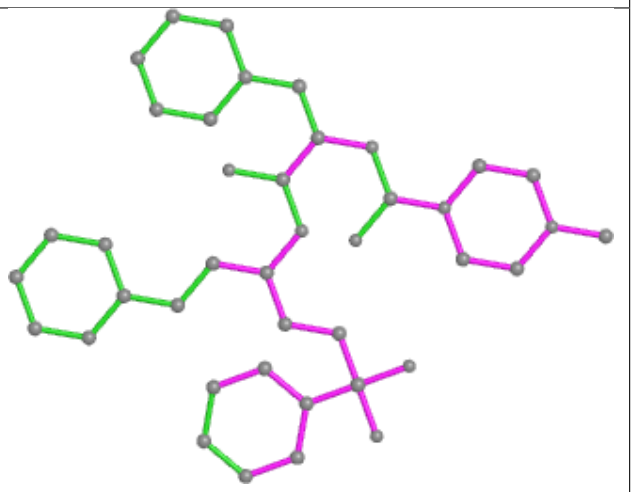


Rings

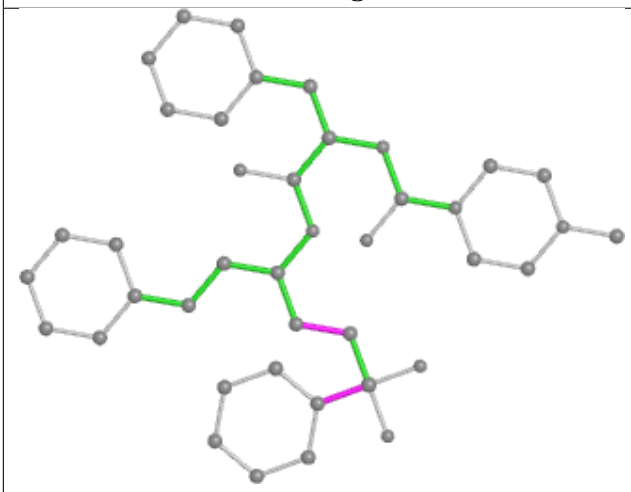
Ligand D1R D 602



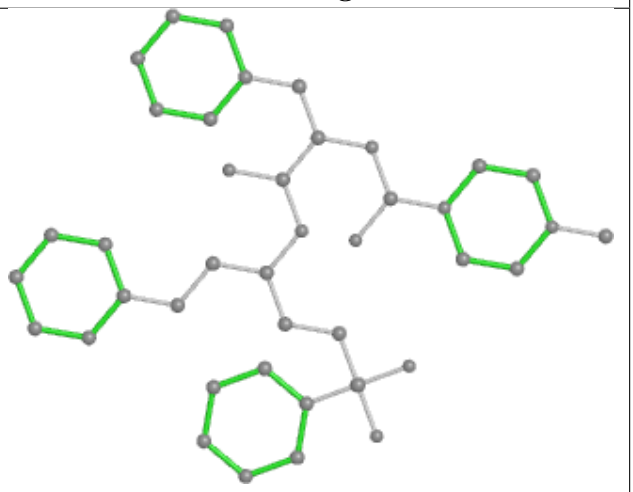
Bond lengths



Bond angles

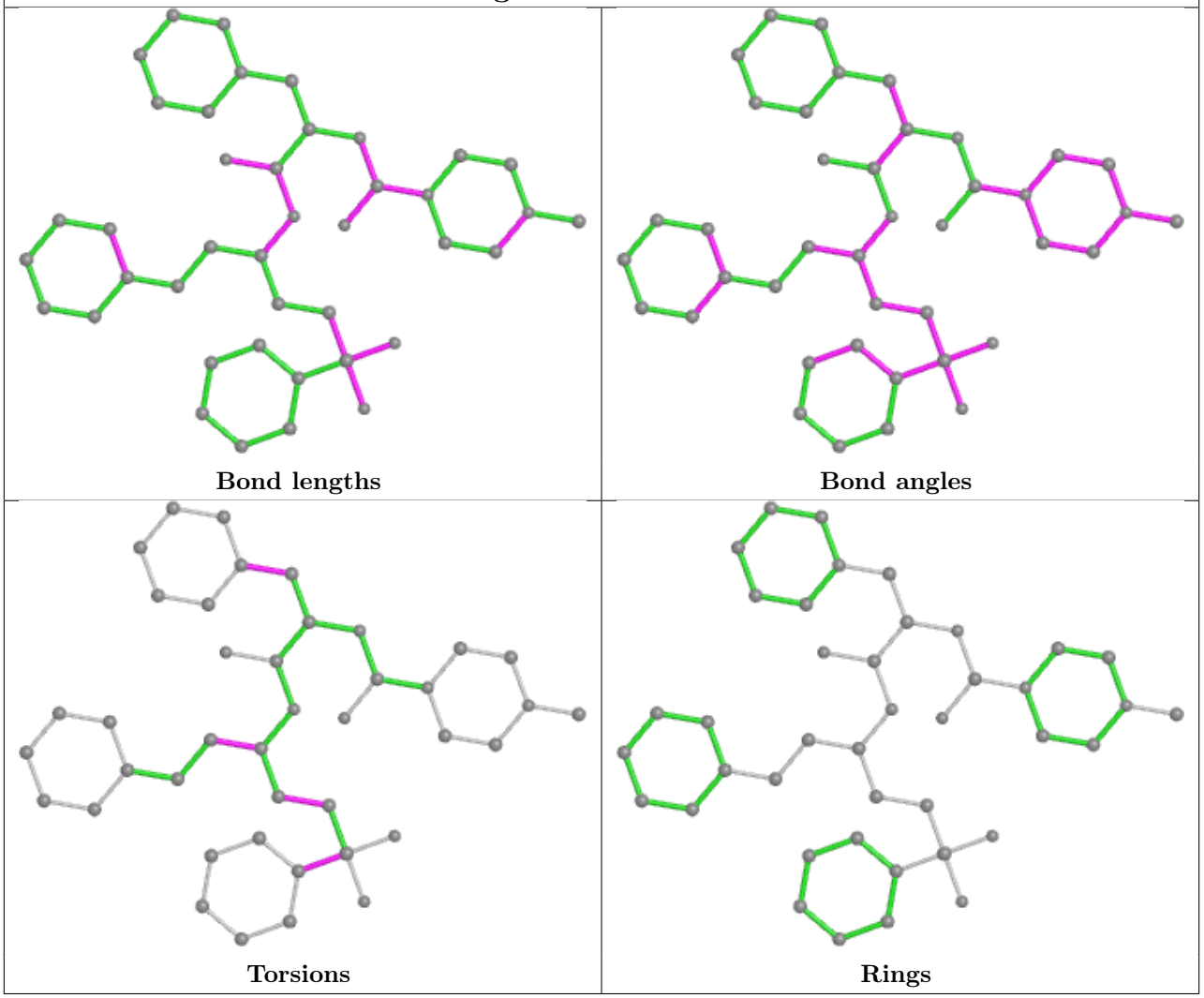


Torsions



Rings

Ligand D1R B 301



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	220/220 (100%)	-0.00	1 (0%)	91	90	10, 16, 29, 40	0
1	B	220/220 (100%)	-0.00	3 (1%)	75	75	7, 13, 27, 41	0
1	C	216/220 (98%)	-0.06	2 (0%)	84	84	10, 17, 31, 43	0
1	D	220/220 (100%)	-0.13	1 (0%)	91	90	8, 12, 25, 33	0
All	All	876/880 (99%)	-0.05	7 (0%)	86	86	7, 15, 28, 43	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1	ALA	5.0
1	C	3	ARG	2.6
1	B	1	ALA	2.4
1	B	175	THR	2.4
1	A	155	ASP	2.3
1	D	103[A]	LYS	2.3
1	B	178	ASP	2.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands (i)

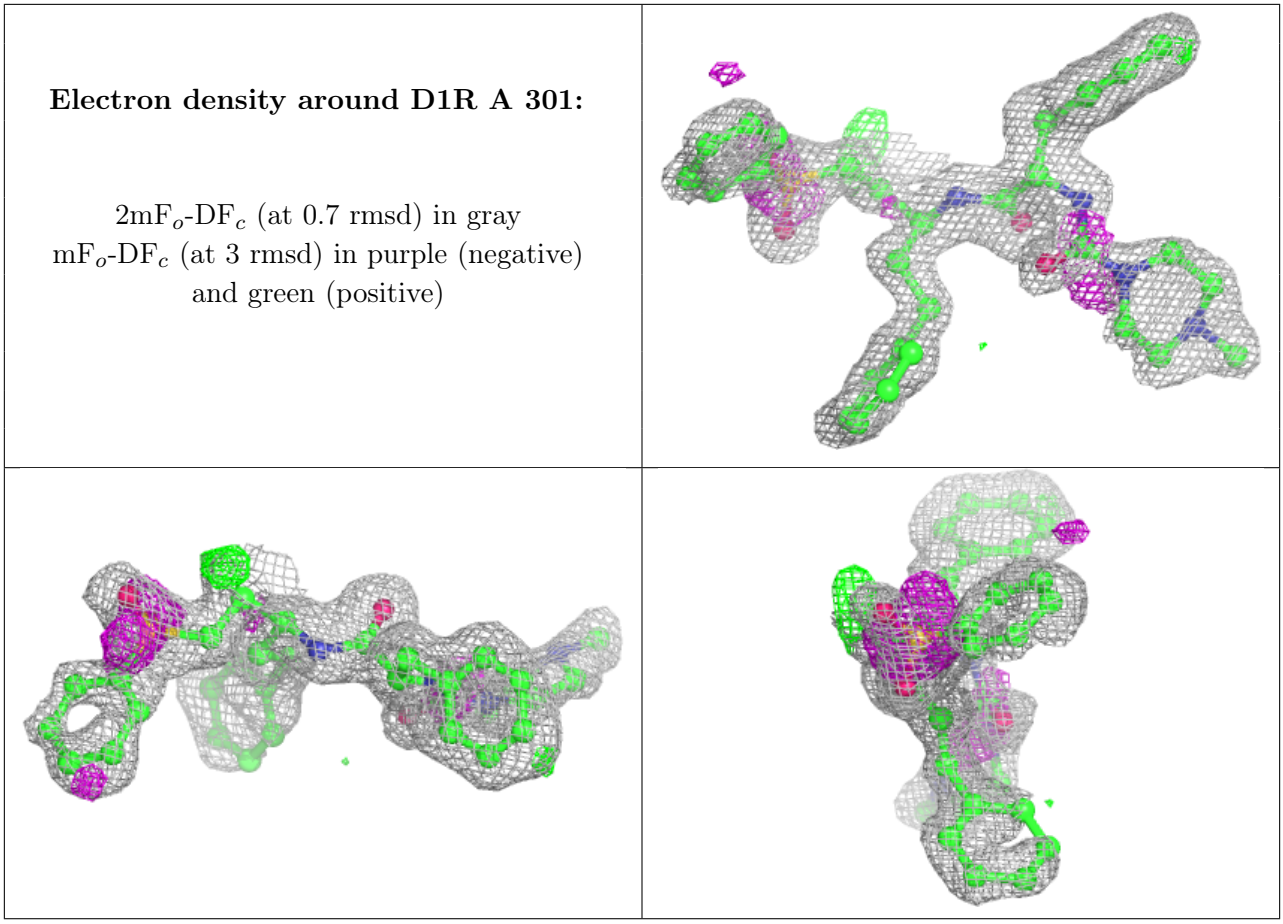
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	EDO	B	303	4/4	0.78	0.20	34,41,48,50	0
4	EDO	D	608	4/4	0.79	0.44	27,35,42,48	0
7	PGE	D	605	10/10	0.79	0.22	30,40,46,46	0
5	PEG	A	308	7/7	0.81	0.17	29,36,47,47	0
4	EDO	A	304	4/4	0.81	0.22	29,34,42,50	0
5	PEG	D	603	7/7	0.82	0.18	28,36,49,52	0
3	PG4	A	302	13/13	0.83	0.20	30,36,48,50	0
4	EDO	C	304	4/4	0.83	0.18	28,38,46,52	0
5	PEG	D	601	7/7	0.84	0.12	29,39,46,48	0
4	EDO	B	307	4/4	0.84	0.29	26,36,44,49	0
4	EDO	A	303	4/4	0.84	0.16	21,35,43,50	0
2	D1R	A	301	41/41	0.85	0.23	18,27,38,43	0
4	EDO	B	304	4/4	0.86	0.30	28,36,48,58	0
7	PGE	B	302	10/10	0.86	0.19	28,38,46,55	0
4	EDO	A	307	4/4	0.86	0.16	32,41,51,51	0
5	PEG	A	306	7/7	0.88	0.26	20,34,45,54	0
5	PEG	C	302	7/7	0.88	0.18	25,41,49,50	0
6	DMS	D	604	4/4	0.89	0.25	25,31,45,45	0
4	EDO	D	606	4/4	0.89	0.18	14,25,35,35	0
6	DMS	A	309	4/4	0.89	0.22	30,36,43,48	0
7	PGE	D	607	10/10	0.89	0.12	24,32,42,49	0
2	D1R	B	301	41/41	0.90	0.21	17,22,36,41	0
2	D1R	D	602	41/41	0.90	0.20	13,20,35,43	0
4	EDO	A	305	4/4	0.91	0.19	30,39,46,47	0
2	D1R	C	301	41/41	0.91	0.20	16,23,36,40	0
4	EDO	B	306	4/4	0.92	0.17	24,33,35,40	0
6	DMS	B	305	4/4	0.92	0.18	24,36,46,46	0
6	DMS	C	303	4/4	0.93	0.26	30,37,45,46	0
4	EDO	B	308	4/4	0.96	0.16	20,32,40,47	0
6	DMS	C	305	4/4	0.96	0.23	34,43,51,51	0
6	DMS	B	309	4/4	0.96	0.12	20,24,33,33	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.

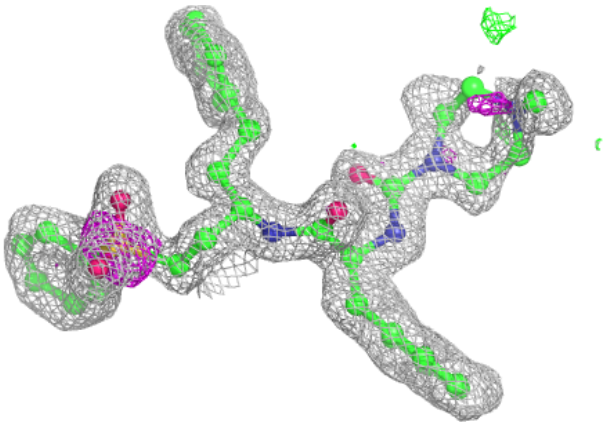
Electron density around D1R A 301:

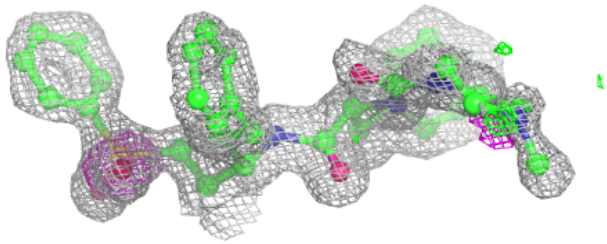
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

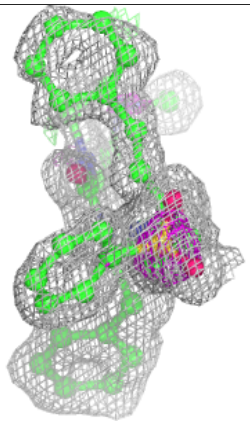


Electron density around D1R B 301:

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

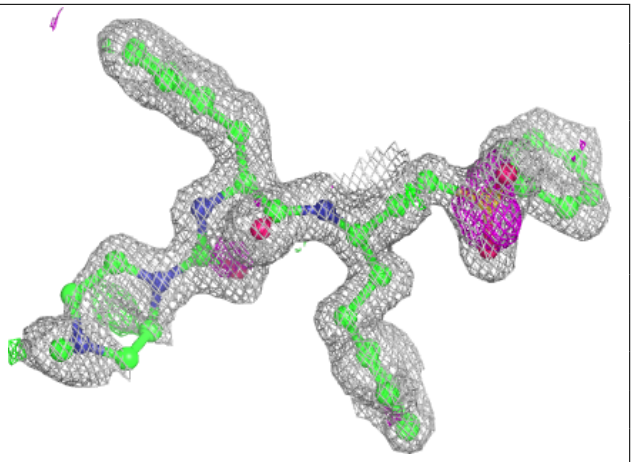


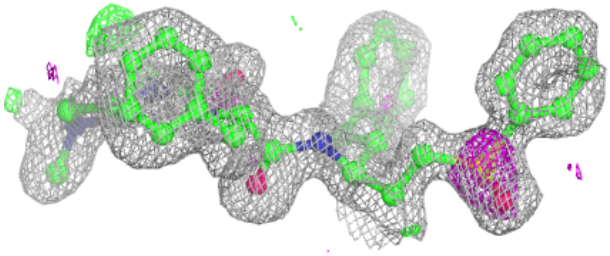


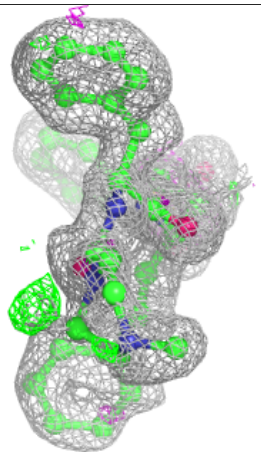


Electron density around D1R D 602:

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

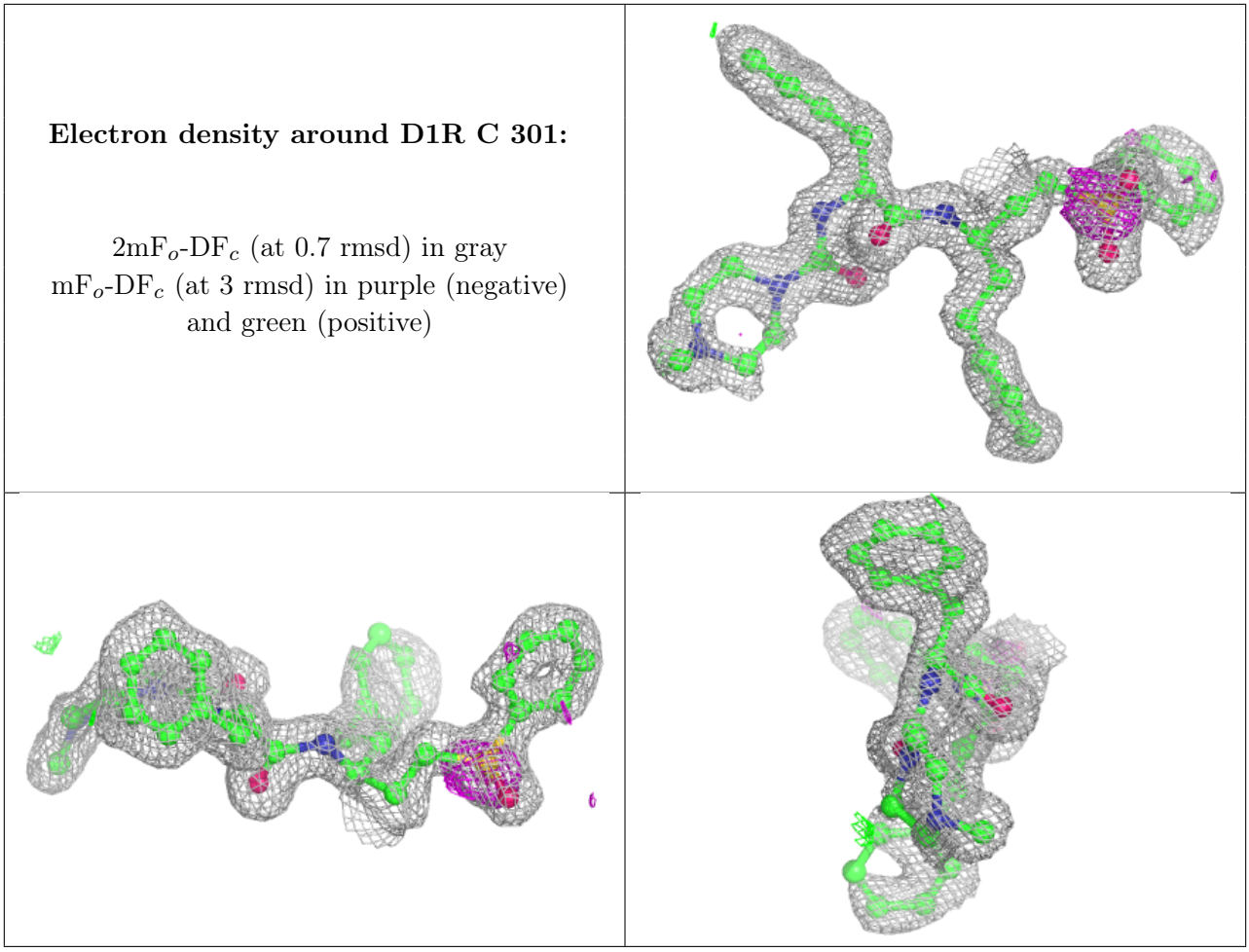






Electron density around D1R C 301:

$2mF_o$-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

There are no such residues in this entry.