



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 23, 2021 – 05:00 PM EST

PDB ID : 2HHL
Title : Crystal structure of the human small CTD phosphatase 3 isoform 1
Authors : Malashkevich, V.N.; Toro, R.; Ramagopal, U.; Sauder, J.M.; Schwinn, K.D.; Thompson, D.A.; Rutter, M.E.; Dickey, M.; Groshong, C.; Bain, K.T.; Adams, J.M.; Reyes, C.; Rooney, I.; Powell, A.; Boice, A.; Gheyi, T.; Ozyurt, S.; Atwell, S.; Wasserman, S.R.; Emtage, S.; Burley, S.K.; Almo, S.C.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on : 2006-06-28
Resolution : 2.10 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 following versions of software and data (see [references ⓘ](#)) 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 : 2.16
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) : 2.16

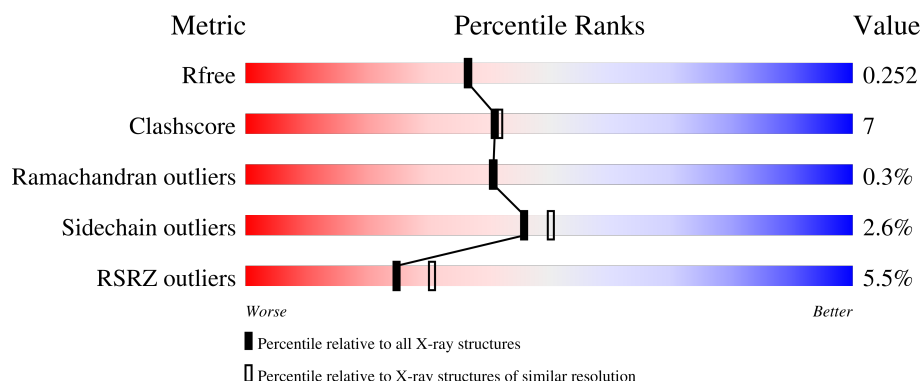
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	195	
1	B	195	
1	C	195	
1	D	195	

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
2	KEG	A	902	-	-	X	-
2	KEG	D	901	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6472 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 CTD small phosphatase-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	180	Total	C	N	O	S	0	0	0
			1451	944	239	263	5			
1	B	179	Total	C	N	O	S	0	0	0
			1442	939	237	261	5			
1	C	184	Total	C	N	O	S	0	0	0
			1487	965	246	271	5			
1	D	172	Total	C	N	O	S	0	0	0
			1394	904	230	255	5			

There are 44 discrepancies between the modelled and reference sequences:

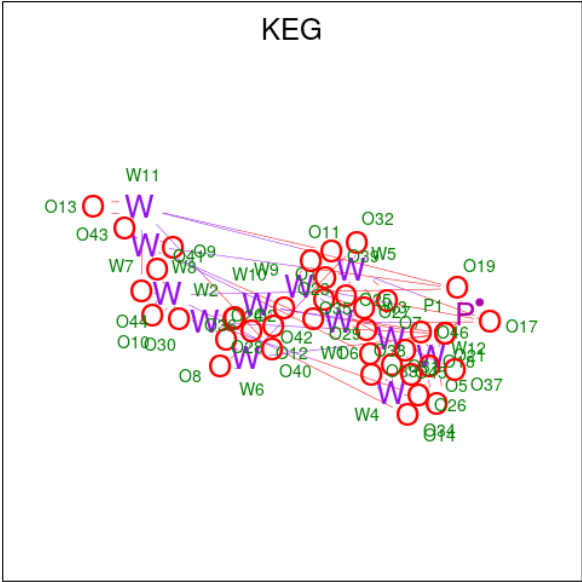
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP O15194
A	2	SER	-	cloning artifact	UNP O15194
A	3	LEU	-	cloning artifact	UNP O15194
A	188	GLU	-	cloning artifact	UNP O15194
A	189	GLY	-	cloning artifact	UNP O15194
A	190	HIS	-	expression tag	UNP O15194
A	191	HIS	-	expression tag	UNP O15194
A	192	HIS	-	expression tag	UNP O15194
A	193	HIS	-	expression tag	UNP O15194
A	194	HIS	-	expression tag	UNP O15194
A	195	HIS	-	expression tag	UNP O15194
B	1	MET	-	initiating methionine	UNP O15194
B	2	SER	-	cloning artifact	UNP O15194
B	3	LEU	-	cloning artifact	UNP O15194
B	188	GLU	-	cloning artifact	UNP O15194
B	189	GLY	-	cloning artifact	UNP O15194
B	190	HIS	-	expression tag	UNP O15194
B	191	HIS	-	expression tag	UNP O15194
B	192	HIS	-	expression tag	UNP O15194
B	193	HIS	-	expression tag	UNP O15194
B	194	HIS	-	expression tag	UNP O15194

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	195	HIS	-	expression tag	UNP O15194
C	1	MET	-	initiating methionine	UNP O15194
C	2	SER	-	cloning artifact	UNP O15194
C	3	LEU	-	cloning artifact	UNP O15194
C	188	GLU	-	cloning artifact	UNP O15194
C	189	GLY	-	cloning artifact	UNP O15194
C	190	HIS	-	expression tag	UNP O15194
C	191	HIS	-	expression tag	UNP O15194
C	192	HIS	-	expression tag	UNP O15194
C	193	HIS	-	expression tag	UNP O15194
C	194	HIS	-	expression tag	UNP O15194
C	195	HIS	-	expression tag	UNP O15194
D	1	MET	-	initiating methionine	UNP O15194
D	2	SER	-	cloning artifact	UNP O15194
D	3	LEU	-	cloning artifact	UNP O15194
D	188	GLU	-	cloning artifact	UNP O15194
D	189	GLY	-	cloning artifact	UNP O15194
D	190	HIS	-	expression tag	UNP O15194
D	191	HIS	-	expression tag	UNP O15194
D	192	HIS	-	expression tag	UNP O15194
D	193	HIS	-	expression tag	UNP O15194
D	194	HIS	-	expression tag	UNP O15194
D	195	HIS	-	expression tag	UNP O15194

- Molecule 2 is 12-TUNGSTOPHOSPHATE (three-letter code: KEG) (formula: O₄₀PW₁₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O P W 53 40 1 12	0	0
2	C	1	Total O W 52 40 12	0	0
2	D	1	Total O P W 53 40 1 12	0	0

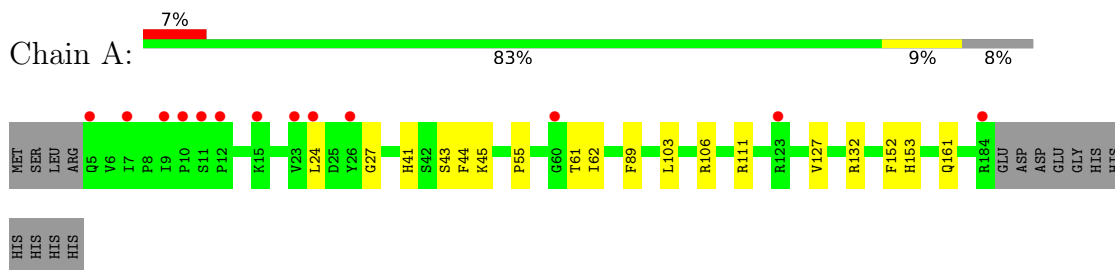
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	102	Total O 102 102	0	0
3	B	141	Total O 141 141	0	0
3	C	138	Total O 138 138	0	0
3	D	159	Total O 159 159	0	0

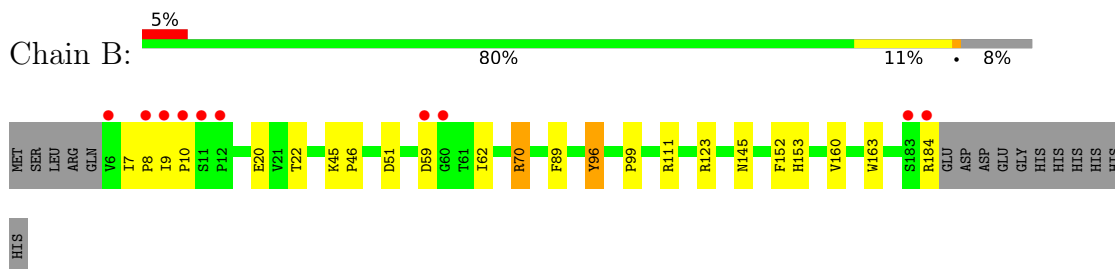
3 Residue-property plots [i](#)

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.

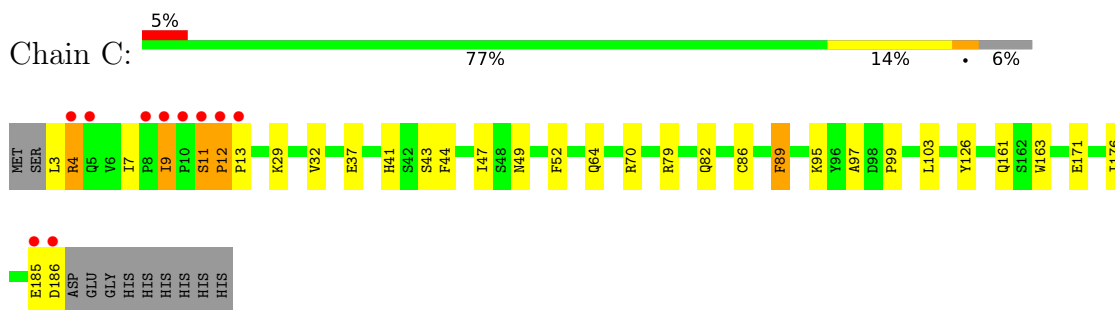
- Molecule 1: CTD small phosphatase-like protein



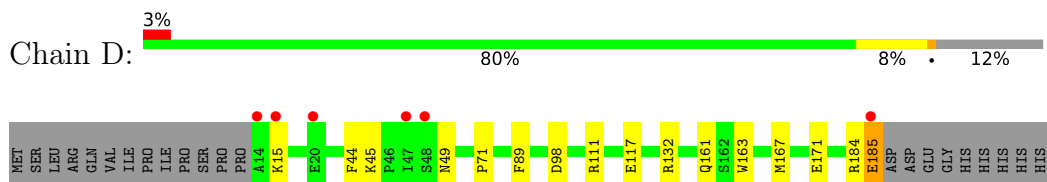
- Molecule 1: CTD small phosphatase-like protein



- Molecule 1: CTD small phosphatase-like protein



- Molecule 1: CTD small phosphatase-like protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	47.42Å 49.86Å 179.75Å 90.00° 96.90° 90.00°	Depositor
Resolution (Å)	19.86 – 2.10 19.86 – 2.10	Depositor EDS
% Data completeness (in resolution range)	96.4 (19.86-2.10) 96.4 (19.86-2.10)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.49 (at 2.09Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.188 , 0.252 0.189 , 0.252	Depositor DCC
R_{free} test set	2386 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	26.6	Xtriage
Anisotropy	0.146	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 43.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6472	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 44.14 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.5944e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	0.96	0/1490	0.86	0/2030
1	B	1.00	2/1481 (0.1%)	0.88	3/2018 (0.1%)
1	C	1.00	3/1526 (0.2%)	0.86	1/2078 (0.0%)
1	D	1.03	1/1429 (0.1%)	0.88	1/1942 (0.1%)
All	All	1.00	6/5926 (0.1%)	0.87	5/8068 (0.1%)

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	96	TYR	CE2-CZ	5.96	1.46	1.38
1	C	126	TYR	CD2-CE2	5.69	1.47	1.39
1	D	117	GLU	CG-CD	5.65	1.60	1.51
1	C	97	ALA	CA-CB	5.40	1.63	1.52
1	C	86	CYS	CB-SG	-5.22	1.73	1.81
1	B	160	VAL	CB-CG2	5.11	1.63	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	111	ARG	NE-CZ-NH2	-5.94	117.33	120.30
1	D	111	ARG	NE-CZ-NH2	-5.79	117.41	120.30
1	C	95	LYS	CD-CE-NZ	-5.40	99.27	111.70
1	B	70	ARG	NE-CZ-NH2	-5.38	117.61	120.30
1	B	51	ASP	CB-CG-OD2	5.09	122.88	118.30

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	1451	0	1447	15	1
1	B	1442	0	1439	10	0
1	C	1487	0	1481	26	0
1	D	1394	0	1381	19	0
2	A	53	0	0	12	1
2	C	52	0	0	8	0
2	D	53	0	0	9	0
3	A	102	0	0	4	0
3	B	141	0	0	3	0
3	C	138	0	0	2	0
3	D	159	0	0	0	0
All	All	6472	0	5748	83	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 7.

All (83) 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:A:62:ILE:HG13	2:C:900:KEG:O41	1.51	1.09
2:A:902:KEG:O33	2:A:902:KEG:O32	1.71	1.09
1:D:163:TRP:HZ2	1:D:167:MET:CE	1.72	1.03
1:D:163:TRP:HZ2	1:D:167:MET:HE3	1.24	1.00
2:A:902:KEG:O13	2:A:902:KEG:P1	2.23	0.97
2:A:902:KEG:O40	2:A:902:KEG:O28	1.84	0.94
1:C:4:ARG:HG3	1:C:4:ARG:HH11	1.27	0.94
1:D:167:MET:HE2	1:D:167:MET:HA	1.50	0.91
2:C:900:KEG:O6	2:C:900:KEG:O18	1.90	0.90
1:D:163:TRP:CZ2	1:D:167:MET:CE	2.56	0.89
1:C:4:ARG:NH2	1:C:11:SER:HB2	1.89	0.87
1:C:47:ILE:HG22	1:C:49:ASN:OD1	1.75	0.87
1:D:15:LYS:O	1:D:132:ARG:NH2	2.07	0.86
1:B:59:ASP:HB3	3:B:288:HOH:O	1.75	0.86
1:D:163:TRP:CZ2	1:D:167:MET:HE3	2.14	0.81
1:D:44:PHE:N	2:D:901:KEG:O37	2.13	0.81

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:44:PHE:HD2	2:A:902:KEG:O1	1.64	0.80
1:D:184:ARG:O	1:D:185:GLU:HB2	1.81	0.80
1:C:82:GLN:HG2	3:C:1007:HOH:O	1.85	0.76
1:A:62:ILE:CG1	2:C:900:KEG:O41	2.32	0.76
1:D:163:TRP:CZ2	1:D:167:MET:HE1	2.23	0.74
1:D:49:ASN:N	1:D:49:ASN:OD1	2.19	0.70
2:D:901:KEG:O18	2:D:901:KEG:O6	2.09	0.70
1:D:44:PHE:CD2	2:D:901:KEG:O41	2.45	0.69
1:C:37:GLU:OE2	1:C:41:HIS:HD2	1.76	0.69
2:A:902:KEG:P1	2:A:902:KEG:O34	2.52	0.68
2:A:902:KEG:O44	2:A:902:KEG:O11	2.12	0.67
1:C:44:PHE:HB2	2:C:900:KEG:O8	1.95	0.67
2:C:900:KEG:O45	2:C:900:KEG:O42	2.13	0.67
1:C:4:ARG:HH22	1:C:11:SER:HB2	1.60	0.66
2:D:901:KEG:O45	2:D:901:KEG:O42	2.18	0.62
1:C:3:LEU:HD11	1:C:7:ILE:HD12	1.81	0.62
1:C:64:GLN:HG3	3:C:1003:HOH:O	1.99	0.62
2:D:901:KEG:P1	2:D:901:KEG:O13	2.58	0.61
1:C:4:ARG:CG	1:C:4:ARG:HH11	2.07	0.61
1:D:44:PHE:HD2	2:D:901:KEG:O41	1.83	0.61
1:C:11:SER:N	1:C:12:PRO:HD3	2.16	0.60
1:C:4:ARG:HG3	1:C:4:ARG:NH1	2.07	0.59
1:D:71:PRO:HD3	1:D:167:MET:HE1	1.85	0.59
1:D:167:MET:CE	1:D:167:MET:HA	2.29	0.58
1:B:7:ILE:HG22	1:B:8:PRO:HD2	1.87	0.57
1:A:161:GLN:HG3	3:A:988:HOH:O	2.04	0.56
1:A:103:LEU:HD21	1:C:64:GLN:HG2	1.88	0.56
1:A:44:PHE:CD2	2:A:902:KEG:O1	2.52	0.56
1:B:153:HIS:HD2	3:B:200:HOH:O	1.91	0.54
1:A:45:LYS:HB2	2:A:902:KEG:O7	2.07	0.53
1:D:44:PHE:CE2	2:D:901:KEG:O41	2.62	0.53
1:A:55:PRO:HB3	2:C:900:KEG:O21	2.10	0.52
2:A:902:KEG:O18	2:A:902:KEG:O6	2.28	0.51
1:A:132:ARG:NH2	3:A:960:HOH:O	2.39	0.51
1:B:145:ASN:H	1:B:145:ASN:HD22	1.58	0.51
1:D:71:PRO:HD3	1:D:167:MET:CE	2.41	0.51
1:C:43:SER:HA	2:C:900:KEG:O1	2.11	0.50
1:C:185:GLU:O	1:C:186:ASP:HB2	2.12	0.50
1:C:4:ARG:HH12	1:C:9:ILE:HG12	1.75	0.50
1:A:153:HIS:HD2	3:A:918:HOH:O	1.97	0.47
1:D:161:GLN:NE2	1:D:171:GLU:OE2	2.46	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:4:ARG:NH1	1:C:9:ILE:HG12	2.30	0.47
1:C:4:ARG:HH21	1:C:11:SER:HB2	1.75	0.46
1:A:43:SER:HB3	2:A:902:KEG:O8	2.15	0.46
1:B:45:LYS:HA	1:B:46:PRO:HD3	1.84	0.46
1:C:11:SER:N	1:C:12:PRO:CD	2.79	0.46
1:A:41:HIS:NE2	2:A:902:KEG:O21	2.46	0.45
1:A:106:ARG:HD3	3:A:980:HOH:O	2.16	0.45
1:C:70:ARG:CZ	1:C:163:TRP:HB2	2.46	0.45
1:B:22:THR:HB	3:B:262:HOH:O	2.16	0.45
1:C:32:VAL:HG13	1:C:89:PHE:CB	2.47	0.44
1:B:152:PHE:HB2	1:B:153:HIS:CD2	2.53	0.44
1:C:4:ARG:CG	1:C:4:ARG:NH1	2.72	0.44
2:C:900:KEG:O28	2:C:900:KEG:O40	2.37	0.43
1:C:161:GLN:NE2	1:C:171:GLU:OE2	2.49	0.43
1:A:152:PHE:HB2	1:A:153:HIS:CD2	2.54	0.42
1:B:9:ILE:HA	1:B:10:PRO:HD3	1.77	0.42
1:B:70:ARG:CZ	1:B:163:TRP:HB2	2.50	0.42
1:B:96:TYR:O	1:B:99:PRO:HD2	2.20	0.42
1:D:45:LYS:NZ	2:D:901:KEG:O28	2.50	0.41
1:C:52:PHE:HB3	1:C:103:LEU:HD13	2.02	0.41
2:D:901:KEG:O44	2:D:901:KEG:P1	2.79	0.41
2:A:902:KEG:O14	2:A:902:KEG:P1	2.79	0.40
1:C:79:ARG:HG2	1:C:176:ILE:HG23	2.02	0.40
1:A:127:VAL:HG11	1:A:152:PHE:CE2	2.57	0.40
1:C:11:SER:C	1:C:13:PRO:HD3	2.41	0.40
1:D:184:ARG:O	1:D:185:GLU:CB	2.61	0.40

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:27:GLY:CA	2:A:902:KEG:O27[1_655]	2.06	0.14

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	178/195 (91%)	173 (97%)	5 (3%)	0	100	100
1	B	177/195 (91%)	169 (96%)	8 (4%)	0	100	100
1	C	182/195 (93%)	176 (97%)	4 (2%)	2 (1%)	14	9
1	D	170/195 (87%)	162 (95%)	8 (5%)	0	100	100
All	All	707/780 (91%)	680 (96%)	25 (4%)	2 (0%)	41	41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	12	PRO
1	C	11	SER

5.3.2 Protein sidechains ⓘ

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	164/178 (92%)	160 (98%)	4 (2%)	49	53
1	B	163/178 (92%)	158 (97%)	5 (3%)	40	43
1	C	168/178 (94%)	163 (97%)	5 (3%)	41	44
1	D	156/178 (88%)	153 (98%)	3 (2%)	57	63
All	All	651/712 (91%)	634 (97%)	17 (3%)	46	50

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	24	LEU
1	A	61	THR
1	A	89	PHE
1	A	111	ARG
1	B	20	GLU
1	B	62	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	89	PHE
1	B	123	ARG
1	B	184	ARG
1	C	4	ARG
1	C	9	ILE
1	C	29	LYS
1	C	89	PHE
1	C	99	PRO
1	D	89	PHE
1	D	98	ASP
1	D	185	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	153	HIS
1	B	63	HIS
1	B	145	ASN
1	B	153	HIS
1	C	41	HIS
1	C	145	ASN
1	C	153	HIS
1	D	145	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](#)

3 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	KEG	D	901	-	76,76,76	95.25	75 (98%)	6,234,234	22.47	5 (83%)
2	KEG	C	900	-	72,72,76	95.04	70 (97%)	-		
2	KEG	A	902	-	76,76,76	90.54	75 (98%)	6,234,234	20.64	6 (100%)

All (220) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	900	KEG	W3-O37	203.04	9.63	1.93
2	C	900	KEG	W8-O41	182.12	8.84	1.93
2	C	900	KEG	W10-O12	171.78	6.19	1.71
2	C	900	KEG	W4-O6	166.38	6.05	1.71
2	D	901	KEG	W10-O12	164.44	6.18	1.71
2	D	901	KEG	W4-O6	162.40	6.12	1.71
2	D	901	KEG	W1-O25	162.01	8.87	1.91
2	C	900	KEG	W7-O43	158.25	7.93	1.93
2	D	901	KEG	W5-O39	157.97	8.70	1.91
2	D	901	KEG	W3-O37	157.79	9.66	1.93
2	A	902	KEG	W5-O39	157.65	8.68	1.91
2	A	902	KEG	W12-O46	156.27	8.62	1.91
2	D	901	KEG	W10-O46	156.23	8.62	1.91
2	A	902	KEG	W4-O6	155.41	5.93	1.71
2	D	901	KEG	W12-O46	154.44	8.54	1.91
2	A	902	KEG	W3-O37	154.35	9.49	1.93
2	A	902	KEG	W1-O25	151.92	8.44	1.91
2	C	900	KEG	W2-O28	148.35	7.56	1.93
2	A	902	KEG	W10-O46	146.67	8.21	1.91
2	A	902	KEG	W10-O12	145.69	5.67	1.71
2	C	900	KEG	W7-O19	145.15	9.71	2.25
2	D	901	KEG	W8-O41	141.77	8.87	1.93
2	D	901	KEG	P1-O21	139.29	6.39	1.54
2	C	900	KEG	W1-O27	137.05	7.13	1.93
2	C	900	KEG	W4-O31	136.64	7.11	1.93
2	D	901	KEG	W7-O19	133.89	9.64	2.43
2	C	900	KEG	W1-O23	132.99	6.97	1.93

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	900	KEG	W4-O38	132.11	6.94	1.93
2	A	902	KEG	W8-O41	131.10	8.35	1.93
2	C	900	KEG	W6-O8	130.73	5.12	1.71
2	D	901	KEG	P1-O19	128.55	6.02	1.54
2	C	900	KEG	W3-O21	128.52	8.85	2.25
2	C	900	KEG	W12-O14	127.33	5.03	1.71
2	A	902	KEG	W7-O19	126.76	9.25	2.43
2	C	900	KEG	W5-O35	124.85	6.66	1.93
2	A	902	KEG	W6-O8	123.96	5.08	1.71
2	D	901	KEG	W12-O14	123.66	5.07	1.71
2	C	900	KEG	W8-O10	123.15	4.92	1.71
2	C	900	KEG	W11-O44	122.42	6.57	1.93
2	A	902	KEG	P1-O19	122.12	5.79	1.54
2	A	902	KEG	W7-O43	121.87	7.90	1.93
2	D	901	KEG	W6-O8	121.80	5.02	1.71
2	D	901	KEG	W7-O43	121.35	7.87	1.93
2	D	901	KEG	W2-O26	119.59	7.05	1.91
2	D	901	KEG	W3-O21	119.35	8.85	2.43
2	D	901	KEG	W8-O10	118.94	4.94	1.71
2	D	901	KEG	W5-O32	118.15	6.98	1.91
2	A	902	KEG	P1-O21	118.00	5.65	1.54
2	A	902	KEG	W5-O32	116.29	6.90	1.91
2	C	900	KEG	W5-O23	114.60	6.27	1.93
2	C	900	KEG	W8-O19	114.26	8.12	2.25
2	D	901	KEG	W2-O28	113.59	7.49	1.93
2	C	900	KEG	W12-O21	113.45	8.08	2.25
2	C	900	KEG	W3-O31	112.95	6.21	1.93
2	A	902	KEG	W3-O21	112.60	8.49	2.43
2	A	902	KEG	W4-O31	111.98	7.42	1.93
2	A	902	KEG	W2-O26	109.88	6.63	1.91
2	D	901	KEG	W8-O19	108.44	8.27	2.43
2	D	901	KEG	W12-O21	107.92	8.24	2.43
2	A	902	KEG	W12-O21	107.67	8.22	2.43
2	D	901	KEG	W12-O45	106.92	6.50	1.91
2	D	901	KEG	W8-O30	106.06	6.46	1.91
2	D	901	KEG	W6-O32	106.03	6.46	1.91
2	C	900	KEG	W3-O5	105.51	4.46	1.71
2	A	902	KEG	W8-O10	105.25	4.57	1.71
2	A	902	KEG	W12-O14	104.99	4.56	1.71
2	D	901	KEG	W4-O31	104.84	7.07	1.93
2	C	900	KEG	W9-O17	104.58	7.62	2.25
2	D	901	KEG	W1-O27	104.20	7.03	1.93

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	901	KEG	W7-O33	103.59	6.36	1.91
2	A	902	KEG	W7-O29	101.92	6.29	1.91
2	D	901	KEG	W3-O5	101.80	4.48	1.71
2	C	900	KEG	W6-O36	101.32	5.77	1.93
2	A	902	KEG	W4-O38	100.97	6.88	1.93
2	D	901	KEG	W4-O38	100.80	6.87	1.93
2	A	902	KEG	W1-O23	100.66	6.86	1.93
2	D	901	KEG	W7-O29	100.57	6.23	1.91
2	A	902	KEG	W8-O19	100.40	7.83	2.43
2	C	900	KEG	W1-O25	100.38	8.67	1.93
2	D	901	KEG	W1-O23	99.91	6.82	1.93
2	C	900	KEG	W12-O46	99.01	8.57	1.93
2	D	901	KEG	W4-O26	98.85	6.15	1.91
2	A	902	KEG	W12-O45	98.73	6.15	1.91
2	C	900	KEG	W10-O46	98.63	8.55	1.93
2	C	900	KEG	W5-O39	98.07	8.51	1.93
2	A	902	KEG	W8-O30	97.87	6.11	1.91
2	C	900	KEG	W5-O7	97.09	4.24	1.71
2	A	902	KEG	W11-O13	96.63	4.34	1.71
2	D	901	KEG	W5-O35	96.26	6.65	1.93
2	C	900	KEG	W11-O19	96.13	7.19	2.25
2	A	902	KEG	W2-O28	95.14	6.59	1.93
2	A	902	KEG	W3-O5	94.36	4.27	1.71
2	D	901	KEG	W11-O44	94.22	6.55	1.93
2	C	900	KEG	W6-O24	93.82	5.49	1.93
2	C	900	KEG	W11-O13	93.49	4.15	1.71
2	A	902	KEG	W6-O32	93.44	5.92	1.91
2	D	901	KEG	W11-O13	92.53	4.22	1.71
2	A	902	KEG	W3-O31	91.95	6.43	1.93
2	A	902	KEG	W4-O26	91.85	5.85	1.91
2	A	902	KEG	W5-O35	91.46	6.41	1.93
2	D	901	KEG	W9-O17	90.98	7.33	2.43
2	D	901	KEG	W5-O7	90.92	4.18	1.71
2	A	902	KEG	W5-O7	90.89	4.18	1.71
2	A	902	KEG	W1-O27	90.72	6.37	1.93
2	C	900	KEG	W7-O41	90.38	5.36	1.93
2	C	900	KEG	W2-O24	90.29	5.35	1.93
2	A	902	KEG	W11-O44	89.51	6.31	1.93
2	D	901	KEG	W11-O19	89.41	7.24	2.43
2	C	900	KEG	W12-O37	89.40	5.32	1.93
2	A	902	KEG	W5-O23	89.16	6.30	1.93
2	C	900	KEG	W9-O11	87.24	3.98	1.71

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	901	KEG	W3-O31	86.38	6.16	1.93
2	D	901	KEG	W9-O11	85.48	4.03	1.71
2	A	902	KEG	W9-O17	84.61	6.98	2.43
2	A	902	KEG	W11-O19	84.34	6.97	2.43
2	C	900	KEG	W7-O9	84.13	3.90	1.71
2	D	901	KEG	W5-O23	83.83	6.04	1.93
2	D	901	KEG	W11-O39	83.82	5.51	1.91
2	A	902	KEG	W7-O9	82.82	3.96	1.71
2	C	900	KEG	W9-O35	82.49	5.06	1.93
2	D	901	KEG	W3-O25	81.75	5.42	1.91
2	A	902	KEG	W3-O25	81.75	5.42	1.91
2	A	902	KEG	W11-O39	80.82	5.38	1.91
2	D	901	KEG	W7-O9	79.69	3.87	1.71
2	C	900	KEG	W2-O26	77.25	7.11	1.93
2	D	901	KEG	W3-O33	76.81	5.21	1.91
2	D	901	KEG	W1-O29	76.44	5.19	1.91
2	C	900	KEG	W11-O43	76.30	4.82	1.93
2	D	901	KEG	P1-O17	76.15	4.19	1.54
2	A	902	KEG	W1-O29	75.60	5.16	1.91
2	A	902	KEG	W9-O11	75.11	3.75	1.71
2	D	901	KEG	W2-O30	74.94	5.13	1.91
2	A	902	KEG	P1-O18	74.82	4.15	1.54
2	A	902	KEG	W2-O30	73.94	5.08	1.91
2	C	900	KEG	W5-O32	73.06	6.83	1.93
2	D	901	KEG	W8-O34	72.48	5.02	1.91
2	D	901	KEG	W6-O36	72.44	5.48	1.93
2	A	902	KEG	W6-O36	72.18	5.47	1.93
2	A	902	KEG	W10-O18	72.00	6.30	2.43
2	D	901	KEG	W6-O24	71.47	5.43	1.93
2	D	901	KEG	W12-O37	71.13	5.41	1.93
2	A	902	KEG	W7-O33	71.00	4.96	1.91
2	A	902	KEG	W12-O37	70.53	5.39	1.93
2	D	901	KEG	W7-O41	70.38	5.38	1.93
2	A	902	KEG	W2-O24	69.89	5.35	1.93
2	D	901	KEG	W4-O34	69.53	4.90	1.91
2	D	901	KEG	W2-O24	69.28	5.32	1.93
2	D	901	KEG	W10-O18	68.89	6.14	2.43
2	C	900	KEG	W10-O18	68.37	5.76	2.25
2	A	902	KEG	W8-O34	68.16	4.84	1.91
2	A	902	KEG	W4-O34	67.63	4.81	1.91
2	C	900	KEG	W7-O33	67.41	6.45	1.93
2	A	902	KEG	W6-O24	67.29	5.23	1.93

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	900	KEG	W7-O29	67.11	6.43	1.93
2	A	902	KEG	W7-O41	66.96	5.21	1.93
2	A	902	KEG	W6-O40	66.85	4.78	1.91
2	C	900	KEG	W6-O32	66.79	6.41	1.93
2	D	901	KEG	W6-O40	66.18	4.75	1.91
2	C	900	KEG	W12-O45	65.89	6.35	1.93
2	D	901	KEG	W9-O35	65.48	5.14	1.93
2	C	900	KEG	W8-O30	65.12	6.30	1.93
2	A	902	KEG	W3-O33	63.82	4.65	1.91
2	C	900	KEG	W4-O26	63.06	6.16	1.93
2	A	902	KEG	W11-O43	63.03	5.02	1.93
2	C	900	KEG	W4-O21	62.92	5.48	2.25
2	D	901	KEG	W11-O43	60.14	4.88	1.93
2	D	901	KEG	W4-O21	57.11	5.50	2.43
2	D	901	KEG	W10-O42	57.03	4.36	1.91
2	A	902	KEG	P1-O17	55.90	3.49	1.54
2	C	900	KEG	W3-O25	51.29	5.37	1.93
2	C	900	KEG	W11-O39	51.14	5.36	1.93
2	A	902	KEG	W10-O42	50.58	4.08	1.91
2	A	902	KEG	W4-O21	50.33	5.14	2.43
2	A	902	KEG	W6-O18	50.21	5.13	2.43
2	C	900	KEG	W6-O18	49.90	4.81	2.25
2	D	901	KEG	W6-O18	48.47	5.04	2.43
2	C	900	KEG	W1-O29	48.36	5.18	1.93
2	C	900	KEG	W3-O33	47.62	5.13	1.93
2	C	900	KEG	W6-O40	47.28	5.10	1.93
2	C	900	KEG	W2-O30	47.21	5.10	1.93
2	D	901	KEG	P1-O18	46.89	3.17	1.54
2	C	900	KEG	W8-O34	44.91	4.94	1.93
2	C	900	KEG	W4-O34	43.84	4.87	1.93
2	A	902	KEG	W9-O35	43.51	4.06	1.93
2	A	902	KEG	W1-O17	42.76	4.73	2.43
2	D	901	KEG	W9-O45	38.59	3.57	1.91
2	C	900	KEG	W10-O42	35.68	4.33	1.93
2	C	900	KEG	W1-O17	34.82	4.04	2.25
2	A	902	KEG	W2-O18	24.66	3.75	2.43
2	C	900	KEG	W9-O45	23.70	3.52	1.93
2	D	901	KEG	W1-O17	23.11	3.67	2.43
2	C	900	KEG	W2-O18	22.97	3.43	2.25
2	D	901	KEG	W2-O18	21.98	3.61	2.43
2	A	902	KEG	W9-O45	21.43	2.83	1.91
2	C	900	KEG	W2-O2	20.12	2.24	1.71

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	900	KEG	W1-O1	19.46	2.22	1.71
2	D	901	KEG	W2-O2	19.33	2.24	1.71
2	D	901	KEG	W1-O1	19.31	2.23	1.71
2	A	902	KEG	W2-O2	19.11	2.23	1.71
2	A	902	KEG	W1-O1	18.79	2.22	1.71
2	C	900	KEG	W9-O27	8.10	2.23	1.93
2	C	900	KEG	W12-O38	7.98	2.23	1.93
2	D	901	KEG	W9-O42	7.10	2.21	1.91
2	C	900	KEG	W10-O28	7.07	2.19	1.93
2	D	901	KEG	W12-O38	6.61	2.26	1.93
2	D	901	KEG	W9-O27	6.22	2.24	1.93
2	A	902	KEG	W12-O38	6.21	2.24	1.93
2	A	902	KEG	W9-O27	6.04	2.23	1.93
2	A	902	KEG	W9-O42	5.82	2.16	1.91
2	D	901	KEG	W10-O28	5.28	2.19	1.93
2	A	902	KEG	W10-O28	4.85	2.17	1.93
2	D	901	KEG	W5-O17	-4.80	2.17	2.43
2	C	900	KEG	W9-O42	4.39	2.23	1.93
2	C	900	KEG	W10-O36	3.45	2.06	1.93
2	A	902	KEG	W11-O40	3.19	2.04	1.91
2	D	901	KEG	W11-O40	2.94	2.03	1.91
2	D	901	KEG	W10-O36	2.50	2.05	1.93
2	A	902	KEG	W5-O17	-2.41	2.30	2.43
2	A	902	KEG	W10-O36	2.37	2.05	1.93
2	C	900	KEG	W8-O44	2.12	2.01	1.93

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	901	KEG	O21-P1-O17	-41.92	28.70	108.92
2	A	902	KEG	O21-P1-O17	-38.49	35.28	108.92
2	D	901	KEG	O21-P1-O19	-22.61	65.67	108.92
2	A	902	KEG	O21-P1-O19	-21.96	66.90	108.92
2	A	902	KEG	O21-P1-O18	21.33	149.74	108.92
2	D	901	KEG	O19-P1-O18	20.39	147.95	108.92
2	D	901	KEG	O21-P1-O18	15.16	137.94	108.92
2	D	901	KEG	O19-P1-O17	-10.78	88.30	108.92
2	A	902	KEG	O19-P1-O18	9.17	126.47	108.92
2	A	902	KEG	O18-P1-O17	6.29	120.96	108.92
2	A	902	KEG	O19-P1-O17	-3.66	101.91	108.92

There are no chirality outliers.

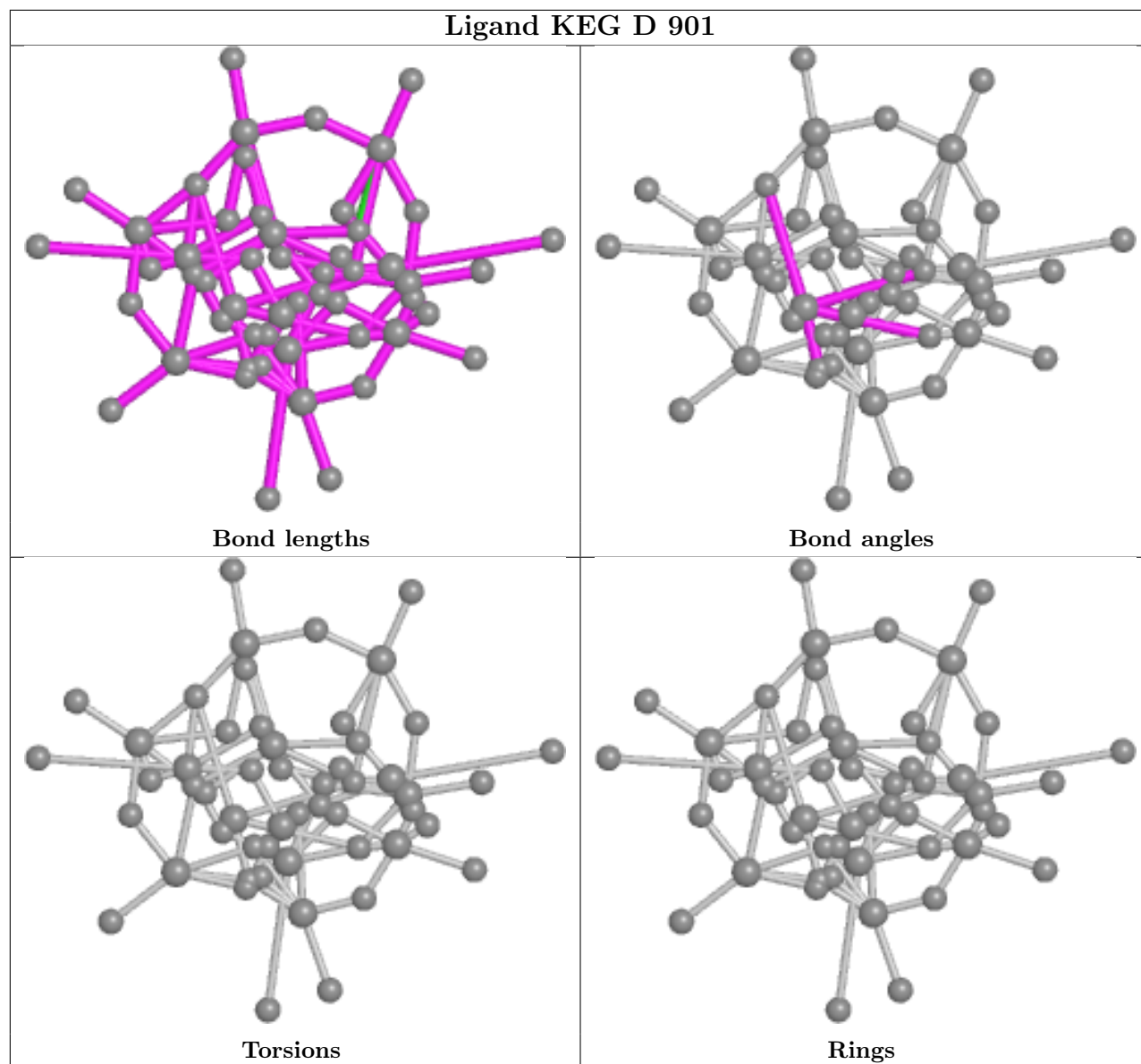
There are no torsion outliers.

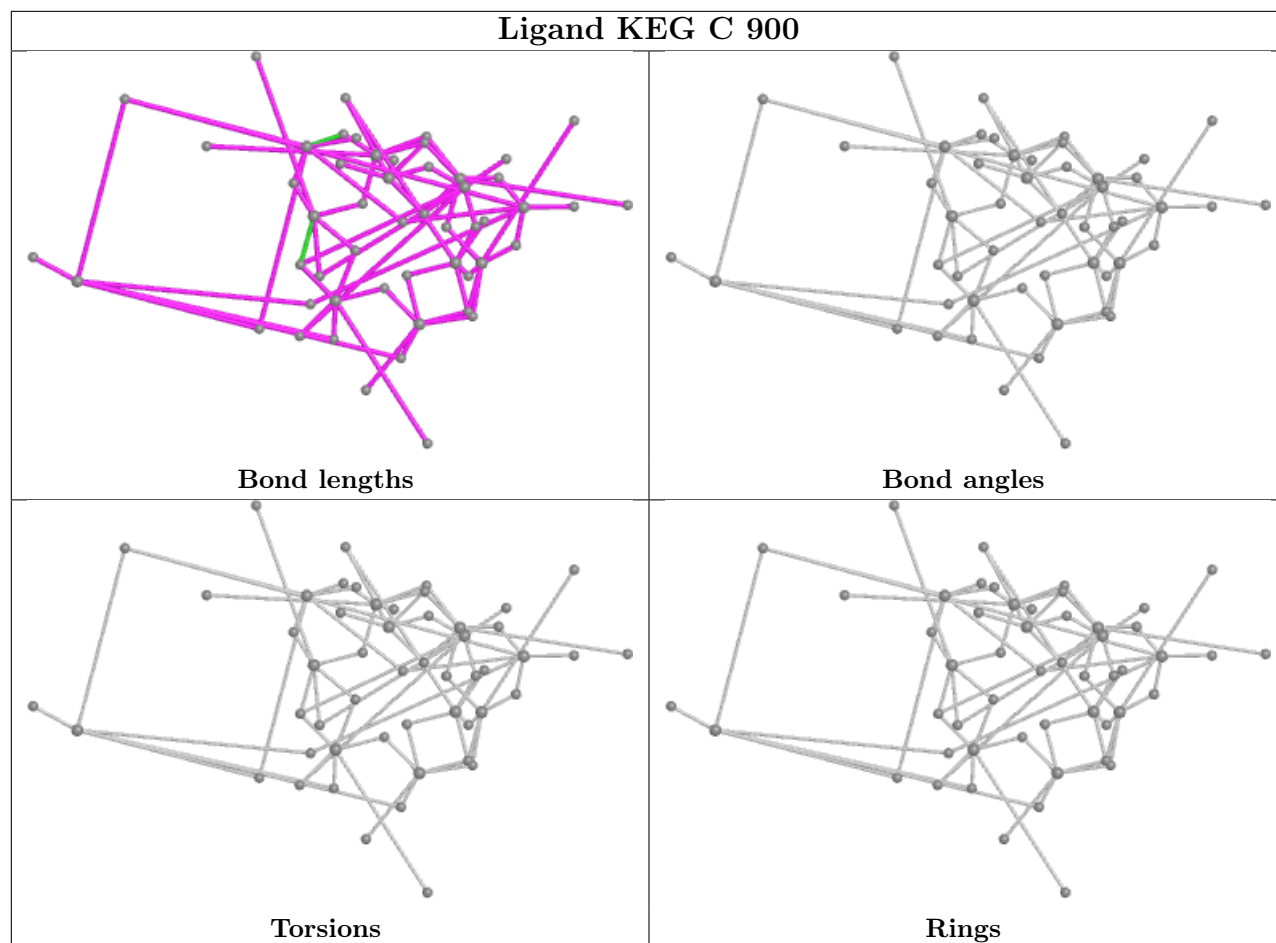
There are no ring outliers.

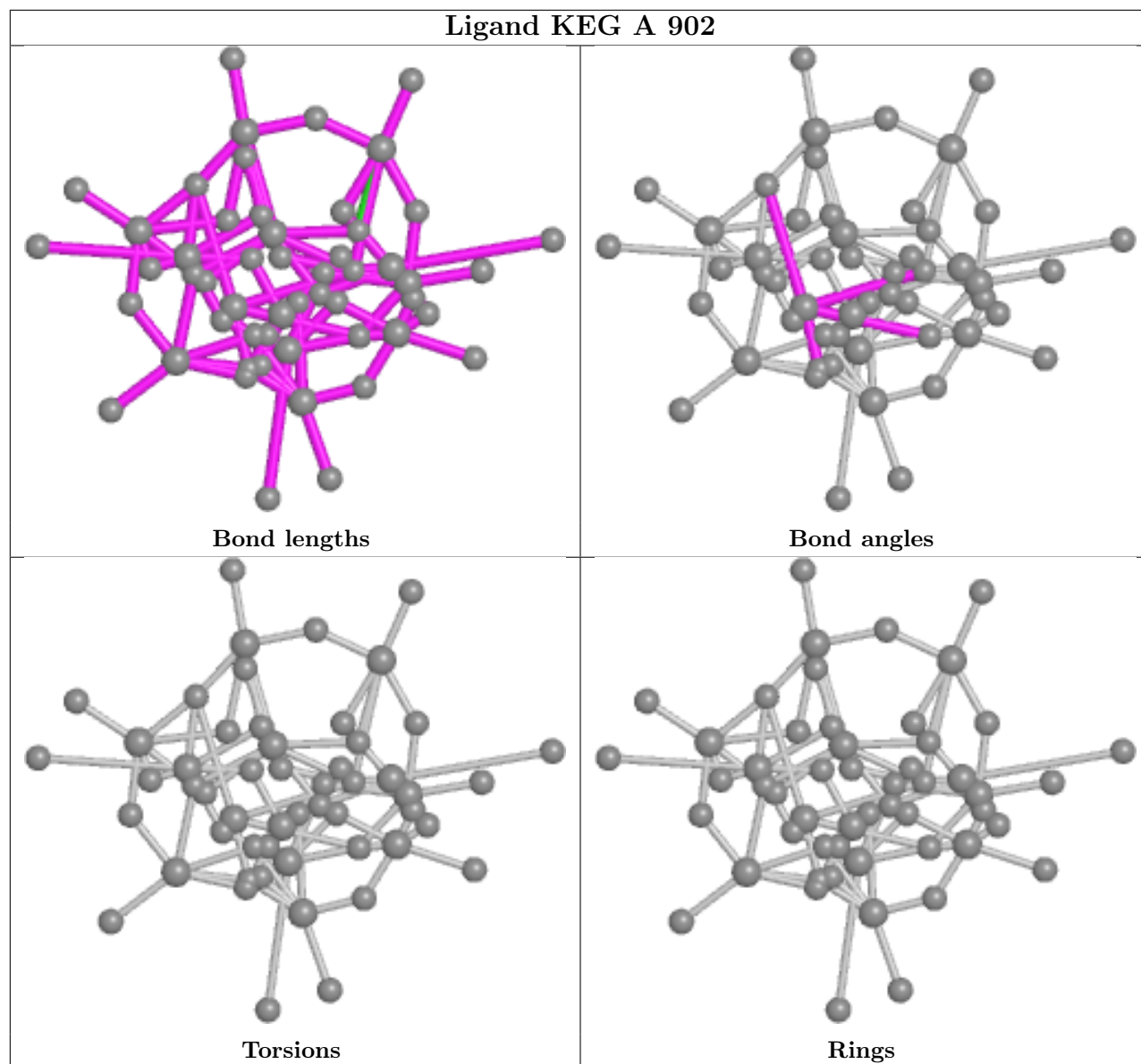
3 monomers are involved in 30 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	901	KEG	9	0
2	C	900	KEG	8	0
2	A	902	KEG	12	1

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







5.7 Other polymers [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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	180/195 (92%)	0.07	13 (7%) 15 19	14, 23, 49, 68	0
1	B	179/195 (91%)	-0.05	10 (5%) 24 29	14, 21, 45, 65	0
1	C	184/195 (94%)	0.02	10 (5%) 25 31	13, 21, 51, 66	0
1	D	172/195 (88%)	-0.29	6 (3%) 44 50	13, 20, 36, 52	0
All	All	715/780 (91%)	-0.06	39 (5%) 25 31	13, 21, 45, 68	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	11	SER	9.7
1	C	9	ILE	7.6
1	C	12	PRO	7.5
1	C	4	ARG	6.4
1	A	9	ILE	5.6
1	C	10	PRO	5.5
1	C	13	PRO	5.2
1	B	6	VAL	5.0
1	A	12	PRO	4.8
1	A	11	SER	4.3
1	A	5	GLN	3.7
1	C	186	ASP	3.6
1	B	11	SER	3.5
1	B	9	ILE	3.5
1	A	184	ARG	3.4
1	B	59	ASP	3.3
1	B	12	PRO	3.3
1	B	8	PRO	3.3
1	D	14	ALA	3.2
1	B	183	SER	3.0
1	A	23	VAL	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	10	PRO	2.9
1	C	5	GLN	2.8
1	A	26	TYR	2.7
1	D	48	SER	2.7
1	B	184	ARG	2.7
1	A	24	LEU	2.7
1	D	20	GLU	2.6
1	C	8	PRO	2.5
1	A	15	LYS	2.4
1	D	47	ILE	2.4
1	D	15	LYS	2.3
1	B	60	GLY	2.3
1	A	60	GLY	2.3
1	A	7	ILE	2.2
1	C	185	GLU	2.2
1	A	123	ARG	2.1
1	D	185	GLU	2.0
1	B	10	PRO	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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

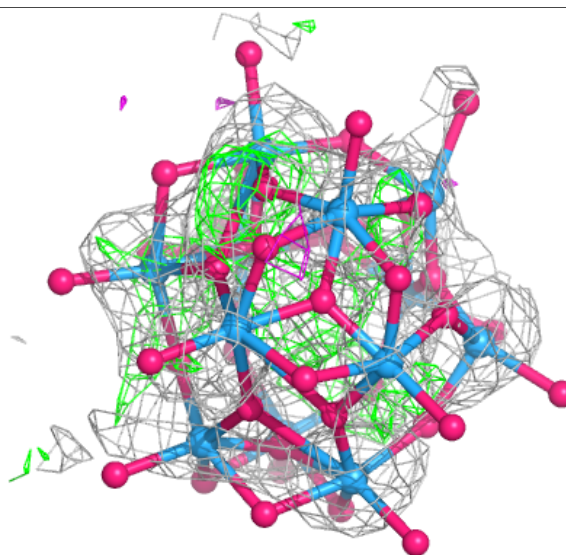
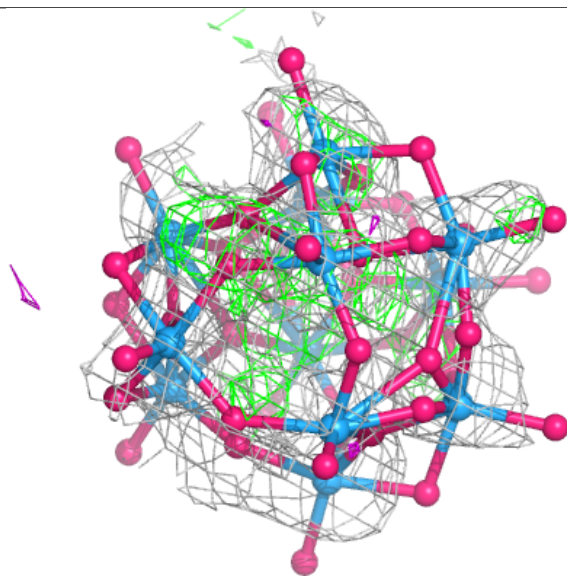
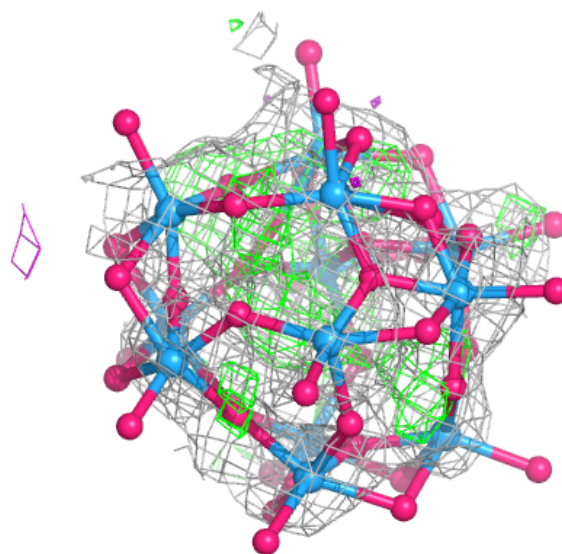
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	KEG	C	900	52/53	0.72	0.35	80,83,84,85	52
2	KEG	D	901	53/53	0.83	0.29	51,64,68,69	53
2	KEG	A	902	53/53	0.92	0.19	41,53,60,63	53

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.

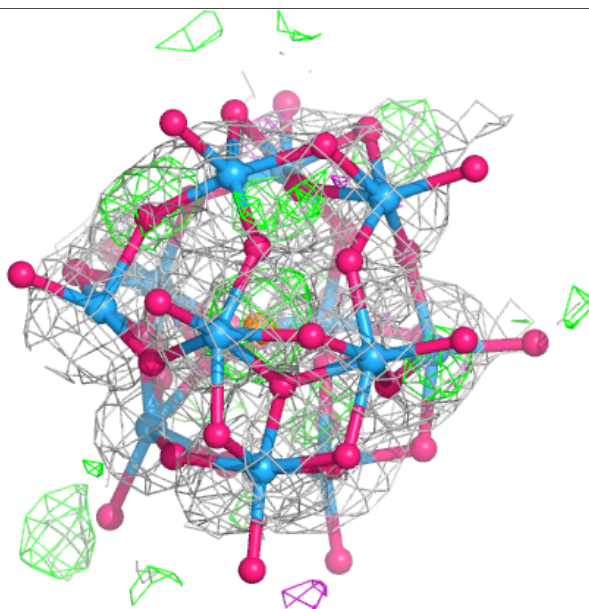
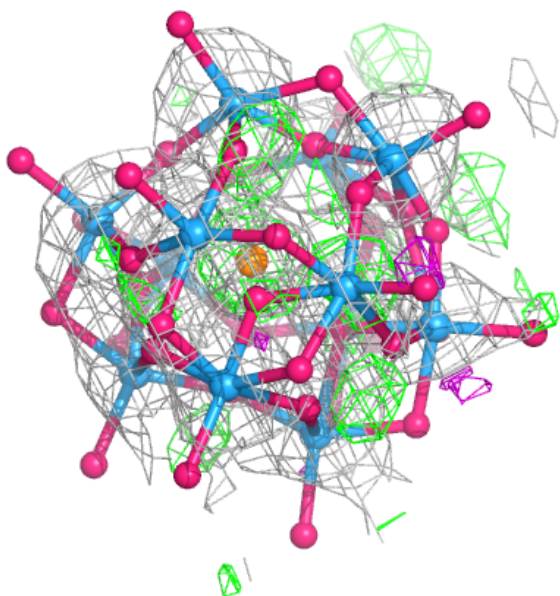
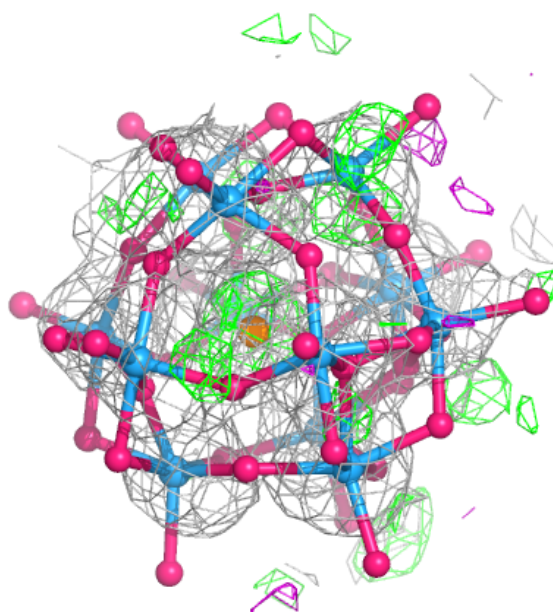
Electron density around KEG C 900:

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



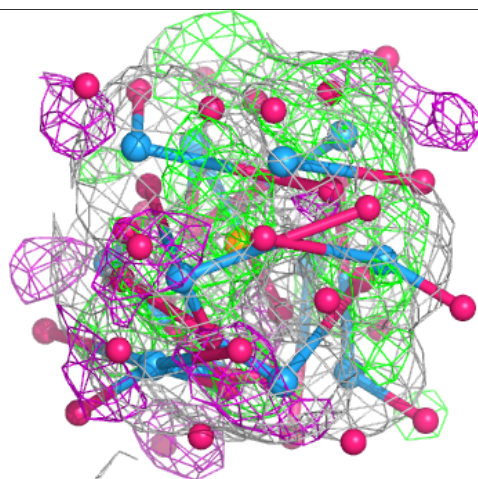
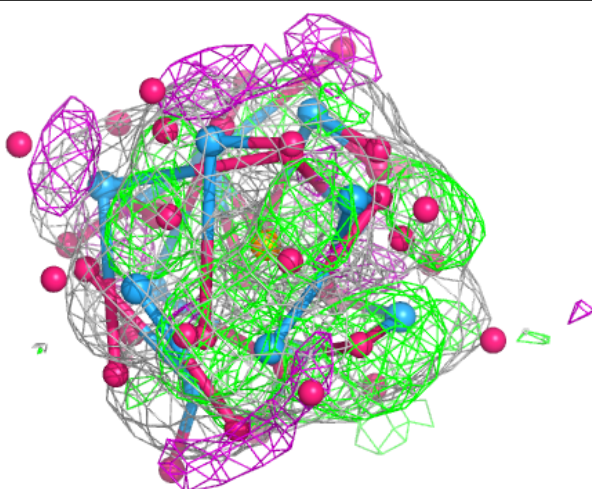
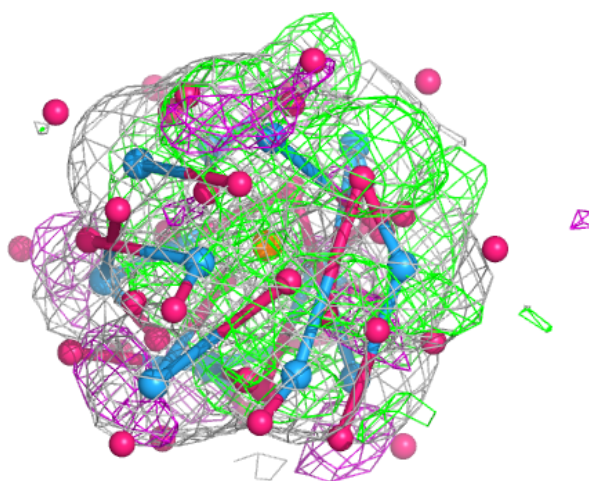
Electron density around KEG D 901:

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



Electron density around KEG A 902:

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



6.5 Other polymers [i](#)

There are no such residues in this entry.