wwPDB X-ray Structure Validation Summary Report (i)

Aug 7, 2020 – 04:50 AM BST

PDB ID	:	5OHS
Title	:	A GH31 family sulfoquinovosidase mutant D455N in complex with pNPSQ
Authors	:	Jin, Y.; Williams, S.J.; Goddard-Borger, E.; Davies, G.J.
Deposited on	:	2017-07-18
Resolution	:	1.97 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 following versions of software and data (see references (i)) were used in the production of this report:

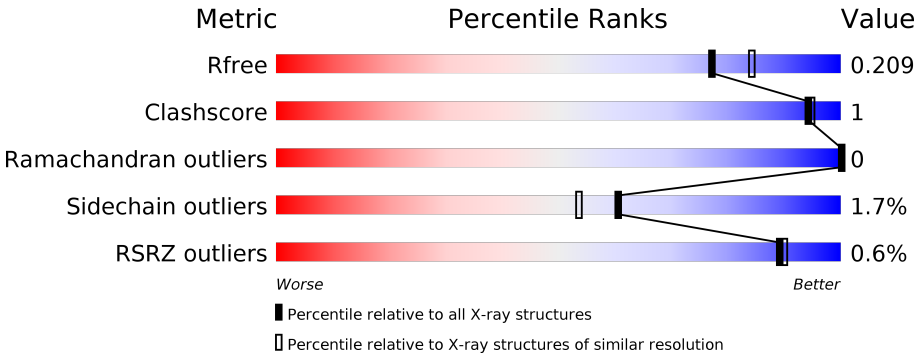
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 1.97 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 on 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	672	91% 7%
1	B	672	% 91% 7%
1	C	672	3% 90% 8%
1	D	672	90% 7%
1	E	672	90% 7%
1	F	672	90% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	672	92% 6%
1	H	672	90% 8%

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
5	EDO	D	707	-	-	X	-

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 44146 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Alpha-glucosidase yihQ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	661	Total	C	N	O	S	0	1	0
			5207	3324	897	964	22			
1	B	663	Total	C	N	O	S	0	1	0
			5215	3330	897	966	22			
1	C	661	Total	C	N	O	S	0	1	0
			5183	3309	890	962	22			
1	D	662	Total	C	N	O	S	0	0	0
			5187	3314	890	962	21			
1	E	661	Total	C	N	O	S	0	1	0
			5206	3325	897	963	21			
1	F	664	Total	C	N	O	S	0	0	0
			5219	3331	899	968	21			
1	G	661	Total	C	N	O	S	0	0	0
			5186	3310	894	961	21			
1	H	661	Total	C	N	O	S	0	0	0
			5197	3318	894	964	21			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	370	ALA	GLU	engineered mutation	UNP A0A083ZKV2
A	371	ALA	GLU	engineered mutation	UNP A0A083ZKV2
A	455	ASN	ASP	engineered mutation	UNP A0A083ZKV2
A	665	LEU	-	expression tag	UNP A0A083ZKV2
A	666	GLU	-	expression tag	UNP A0A083ZKV2
A	667	HIS	-	expression tag	UNP A0A083ZKV2
A	668	HIS	-	expression tag	UNP A0A083ZKV2
A	669	HIS	-	expression tag	UNP A0A083ZKV2
A	670	HIS	-	expression tag	UNP A0A083ZKV2
A	671	HIS	-	expression tag	UNP A0A083ZKV2
A	672	HIS	-	expression tag	UNP A0A083ZKV2
B	370	ALA	GLU	engineered mutation	UNP A0A083ZKV2
B	371	ALA	GLU	engineered mutation	UNP A0A083ZKV2

Continued on next page...

Continued from previous page...

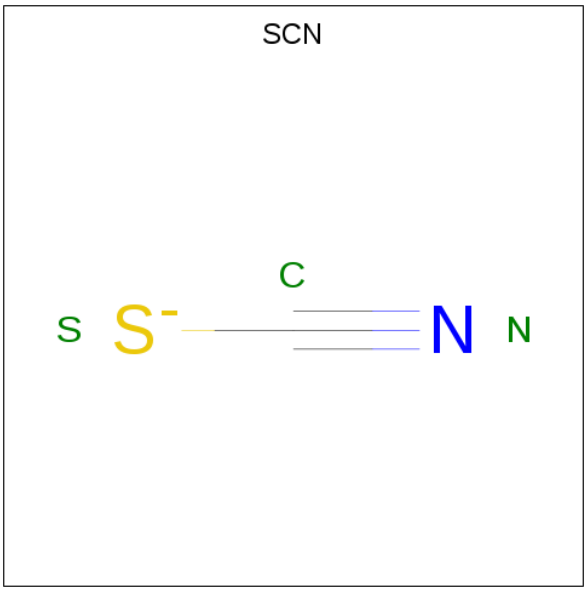
Chain	Residue	Modelled	Actual	Comment	Reference
B	455	ASN	ASP	engineered mutation	UNP A0A083ZKV2
B	665	LEU	-	expression tag	UNP A0A083ZKV2
B	666	GLU	-	expression tag	UNP A0A083ZKV2
B	667	HIS	-	expression tag	UNP A0A083ZKV2
B	668	HIS	-	expression tag	UNP A0A083ZKV2
B	669	HIS	-	expression tag	UNP A0A083ZKV2
B	670	HIS	-	expression tag	UNP A0A083ZKV2
B	671	HIS	-	expression tag	UNP A0A083ZKV2
B	672	HIS	-	expression tag	UNP A0A083ZKV2
C	370	ALA	GLU	engineered mutation	UNP A0A083ZKV2
C	371	ALA	GLU	engineered mutation	UNP A0A083ZKV2
C	455	ASN	ASP	engineered mutation	UNP A0A083ZKV2
C	665	LEU	-	expression tag	UNP A0A083ZKV2
C	666	GLU	-	expression tag	UNP A0A083ZKV2
C	667	HIS	-	expression tag	UNP A0A083ZKV2
C	668	HIS	-	expression tag	UNP A0A083ZKV2
C	669	HIS	-	expression tag	UNP A0A083ZKV2
C	670	HIS	-	expression tag	UNP A0A083ZKV2
C	671	HIS	-	expression tag	UNP A0A083ZKV2
C	672	HIS	-	expression tag	UNP A0A083ZKV2
D	370	ALA	GLU	engineered mutation	UNP A0A083ZKV2
D	371	ALA	GLU	engineered mutation	UNP A0A083ZKV2
D	455	ASN	ASP	engineered mutation	UNP A0A083ZKV2
D	665	LEU	-	expression tag	UNP A0A083ZKV2
D	666	GLU	-	expression tag	UNP A0A083ZKV2
D	667	HIS	-	expression tag	UNP A0A083ZKV2
D	668	HIS	-	expression tag	UNP A0A083ZKV2
D	669	HIS	-	expression tag	UNP A0A083ZKV2
D	670	HIS	-	expression tag	UNP A0A083ZKV2
D	671	HIS	-	expression tag	UNP A0A083ZKV2
D	672	HIS	-	expression tag	UNP A0A083ZKV2
E	370	ALA	GLU	engineered mutation	UNP A0A083ZKV2
E	371	ALA	GLU	engineered mutation	UNP A0A083ZKV2
E	455	ASN	ASP	engineered mutation	UNP A0A083ZKV2
E	665	LEU	-	expression tag	UNP A0A083ZKV2
E	666	GLU	-	expression tag	UNP A0A083ZKV2
E	667	HIS	-	expression tag	UNP A0A083ZKV2
E	668	HIS	-	expression tag	UNP A0A083ZKV2
E	669	HIS	-	expression tag	UNP A0A083ZKV2
E	670	HIS	-	expression tag	UNP A0A083ZKV2
E	671	HIS	-	expression tag	UNP A0A083ZKV2
E	672	HIS	-	expression tag	UNP A0A083ZKV2

Continued on next page...

Continued from previous page...

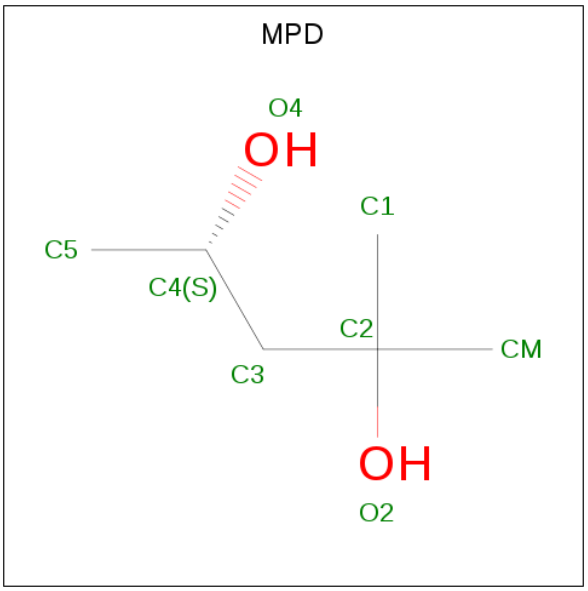
Chain	Residue	Modelled	Actual	Comment	Reference
F	370	ALA	GLU	engineered mutation	UNP A0A083ZKV2
F	371	ALA	GLU	engineered mutation	UNP A0A083ZKV2
F	455	ASN	ASP	engineered mutation	UNP A0A083ZKV2
F	665	LEU	-	expression tag	UNP A0A083ZKV2
F	666	GLU	-	expression tag	UNP A0A083ZKV2
F	667	HIS	-	expression tag	UNP A0A083ZKV2
F	668	HIS	-	expression tag	UNP A0A083ZKV2
F	669	HIS	-	expression tag	UNP A0A083ZKV2
F	670	HIS	-	expression tag	UNP A0A083ZKV2
F	671	HIS	-	expression tag	UNP A0A083ZKV2
F	672	HIS	-	expression tag	UNP A0A083ZKV2
G	370	ALA	GLU	engineered mutation	UNP A0A083ZKV2
G	371	ALA	GLU	engineered mutation	UNP A0A083ZKV2
G	455	ASN	ASP	engineered mutation	UNP A0A083ZKV2
G	665	LEU	-	expression tag	UNP A0A083ZKV2
G	666	GLU	-	expression tag	UNP A0A083ZKV2
G	667	HIS	-	expression tag	UNP A0A083ZKV2
G	668	HIS	-	expression tag	UNP A0A083ZKV2
G	669	HIS	-	expression tag	UNP A0A083ZKV2
G	670	HIS	-	expression tag	UNP A0A083ZKV2
G	671	HIS	-	expression tag	UNP A0A083ZKV2
G	672	HIS	-	expression tag	UNP A0A083ZKV2
H	370	ALA	GLU	engineered mutation	UNP A0A083ZKV2
H	371	ALA	GLU	engineered mutation	UNP A0A083ZKV2
H	455	ASN	ASP	engineered mutation	UNP A0A083ZKV2
H	665	LEU	-	expression tag	UNP A0A083ZKV2
H	666	GLU	-	expression tag	UNP A0A083ZKV2
H	667	HIS	-	expression tag	UNP A0A083ZKV2
H	668	HIS	-	expression tag	UNP A0A083ZKV2
H	669	HIS	-	expression tag	UNP A0A083ZKV2
H	670	HIS	-	expression tag	UNP A0A083ZKV2
H	671	HIS	-	expression tag	UNP A0A083ZKV2
H	672	HIS	-	expression tag	UNP A0A083ZKV2

- Molecule 2 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C N S 3 1 1 1	0	0
2	A	1	Total C N S 3 1 1 1	0	0
2	D	1	Total C N S 3 1 1 1	0	0
2	E	1	Total C N S 3 1 1 1	0	0
2	F	1	Total C N S 3 1 1 1	0	0
2	F	1	Total C N S 3 1 1 1	0	0
2	G	1	Total C N S 3 1 1 1	0	0
2	H	1	Total C N S 3 1 1 1	0	0

- Molecule 3 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 8 6 2	0	0
3	A	1	Total C O 8 6 2	0	0
3	A	1	Total C O 8 6 2	0	0
3	B	1	Total C O 8 6 2	0	0
3	D	1	Total C O 8 6 2	0	0
3	D	1	Total C O 8 6 2	0	0
3	E	1	Total C O 8 6 2	0	0
3	F	1	Total C O 8 6 2	0	0
3	F	1	Total C O 8 6 2	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

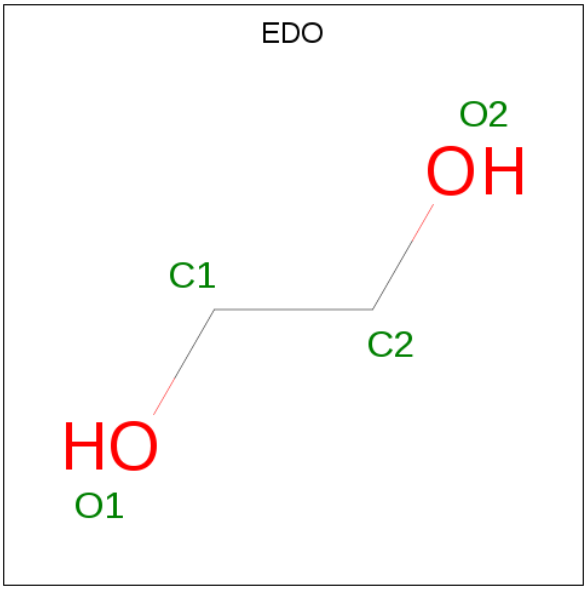
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	G	1	Total Cl 1 1	0	0
4	E	1	Total Cl 1 1	0	0
4	H	1	Total Cl 1 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Cl 1 1	0	0
4	A	1	Total Cl 1 1	0	0
4	F	1	Total Cl 1 1	0	0

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



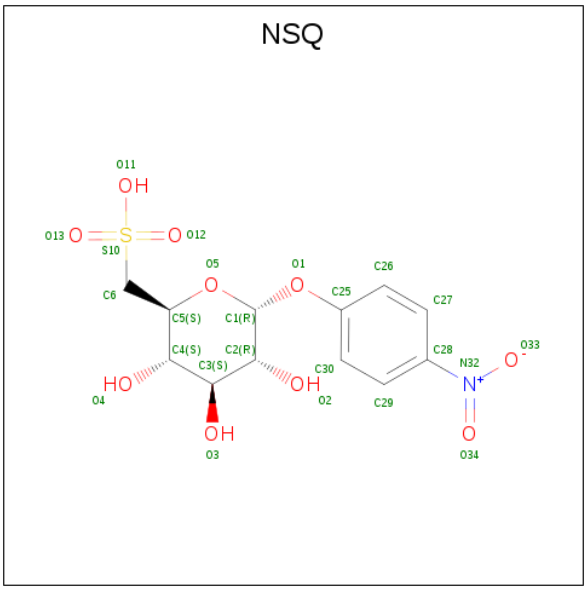
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 4 2 2	0	0
5	A	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0
5	D	1	Total C O 4 2 2	0	0
5	D	1	Total C O 4 2 2	0	0
5	D	1	Total C O 4 2 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	1	Total C O 4 2 2	0	0
5	F	1	Total C O 4 2 2	0	0
5	F	1	Total C O 4 2 2	0	0
5	F	1	Total C O 4 2 2	0	0
5	F	1	Total C O 4 2 2	0	0
5	F	1	Total C O 4 2 2	0	0
5	G	1	Total C O 4 2 2	0	0
5	G	1	Total C O 4 2 2	0	0
5	G	1	Total C O 4 2 2	0	0
5	H	1	Total C O 4 2 2	0	0

- Molecule 6 is 4-nitrophenyl alpha-D-6-sulfoquinovoside (three-letter code: NSQ) (formula: $C_{12}H_{15}NO_{10}S$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C N O S 24 12 1 10 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	N	O	S	0	0
			24	12	1	10	1		
6	C	1	Total	C	N	O	S	0	0
			24	12	1	10	1		
6	D	1	Total	C	N	O	S	0	0
			24	12	1	10	1		
6	E	1	Total	C	N	O	S	0	0
			24	12	1	10	1		
6	F	1	Total	C	N	O	S	0	0
			24	12	1	10	1		
6	G	1	Total	C	N	O	S	0	0
			24	12	1	10	1		
6	H	1	Total	C	N	O	S	0	0
			24	12	1	10	1		

- Molecule 7 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	D	1	Total	K	0	0
			1	1		
7	C	1	Total	K	0	0
			1	1		

- Molecule 8 is water.

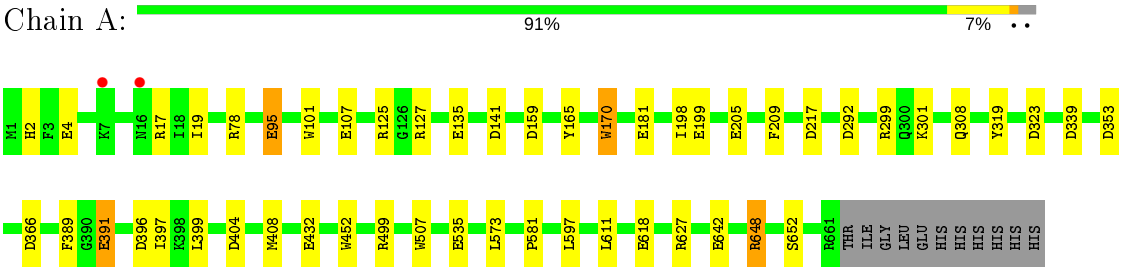
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	296	Total	O	0	2
			298	298		
8	B	265	Total	O	0	0
			265	265		
8	C	237	Total	O	0	1
			238	238		
8	D	305	Total	O	0	2
			307	307		
8	E	270	Total	O	0	2
			272	272		
8	F	292	Total	O	0	2
			294	294		
8	G	231	Total	O	0	0
			231	231		
8	H	272	Total	O	0	1
			273	273		

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.

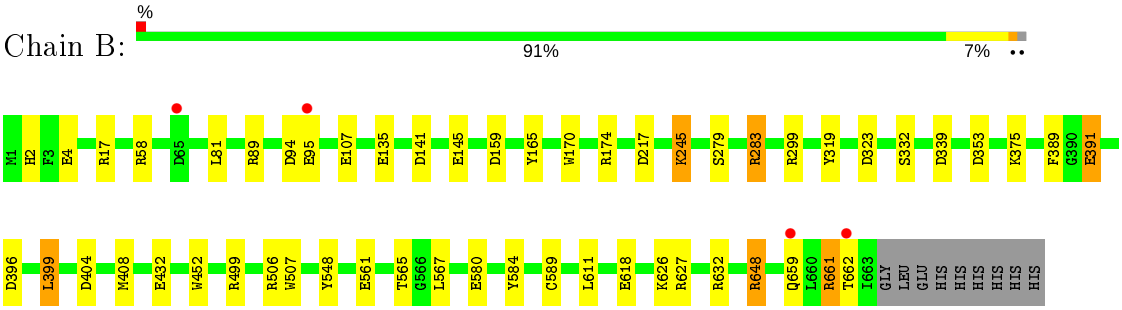
- Molecule 1: Alpha-glucosidase yihQ

Chain A:



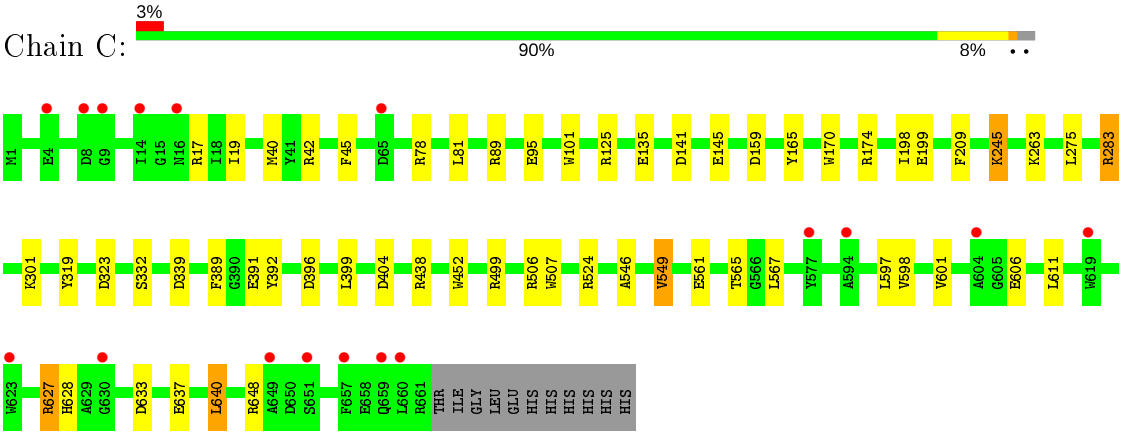
- Molecule 1: Alpha-glucosidase yihQ

Chain B:



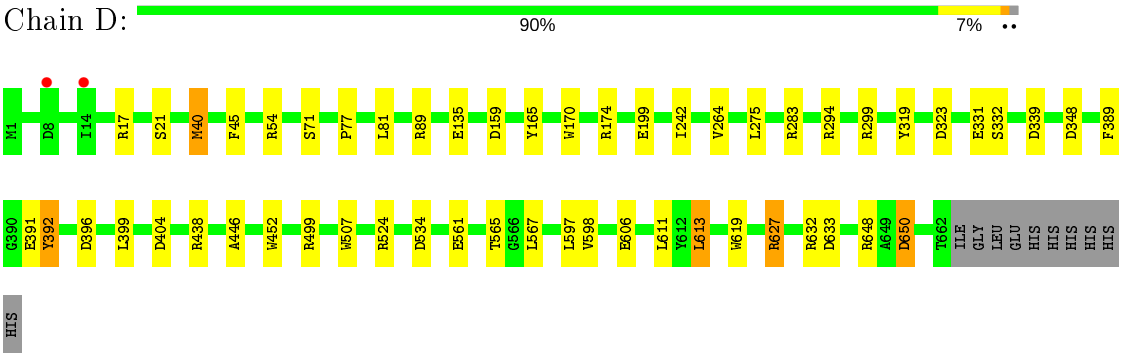
- Molecule 1: Alpha-glucosidase yihQ

Chain C:



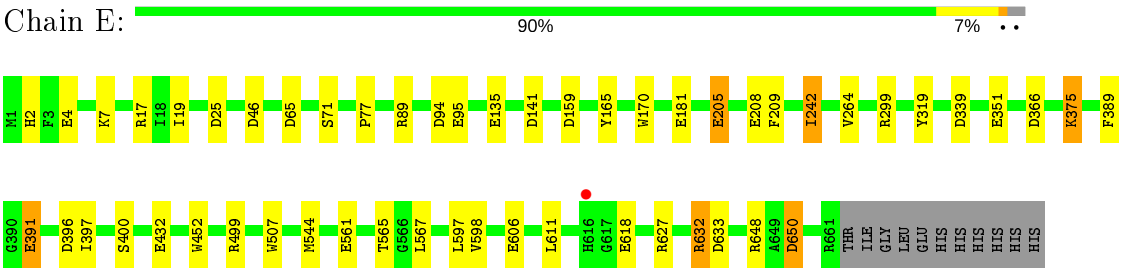
- Molecule 1: Alpha-glucosidase yihQ

Chain D:



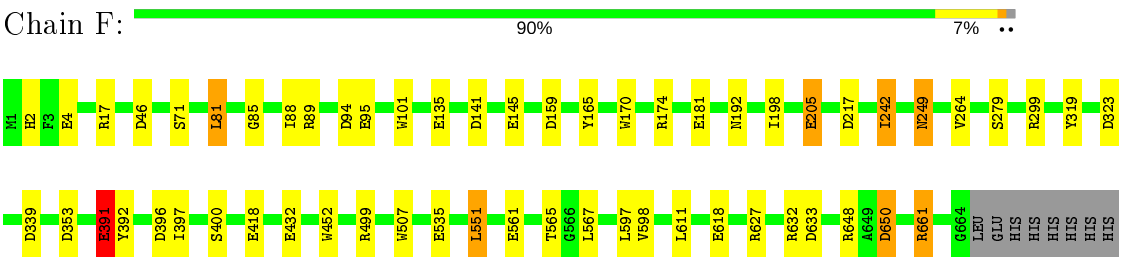
• Molecule 1: Alpha-glucosidase yihQ

Chain E:



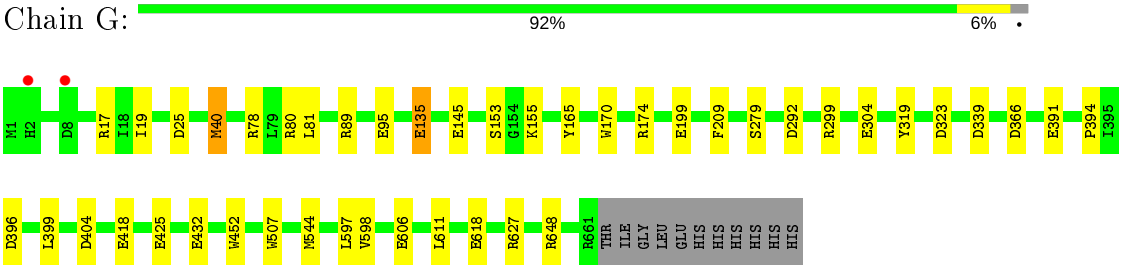
• Molecule 1: Alpha-glucosidase yihQ

Chain F:



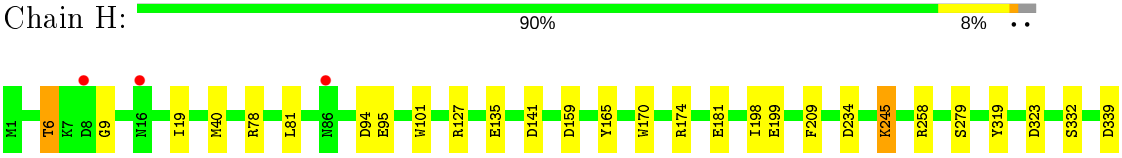
• Molecule 1: Alpha-glucosidase yihQ

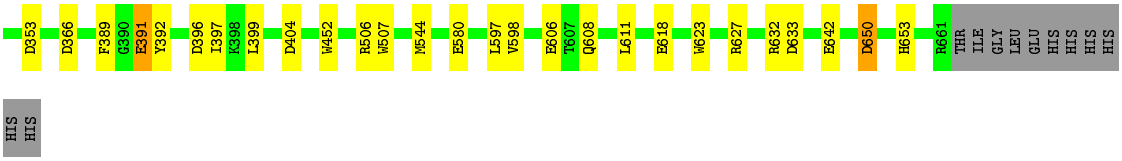
Chain G:



• Molecule 1: Alpha-glucosidase yihQ

Chain H:




D353
D366
F389
G390
E391
Y392
D396
I397
K398
L399
D404
W452
R506
W507
M544
E580
L597
V598
E606
T607
Q608
L611
E618
W623
R627
R632
D633
E642
D650
H653
R661
THR
ILE
GLY
LEU
GLU
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	104.20Å 169.19Å 169.69Å 90.00° 92.80° 90.00°	Depositor
Resolution (Å)	75.77 – 1.97 75.77 – 1.97	Depositor EDS
% Data completeness (in resolution range)	99.8 (75.77-1.97) 99.8 (75.77-1.97)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.34 (at 1.97Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.181 , 0.203 0.188 , 0.209	Depositor DCC
R_{free} test set	20006 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	25.8	Xtriage
Anisotropy	0.395	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 55.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k 0.000 for -h,-l,-k 0.020 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	44146	wwPDB-VP
Average B, all atoms (Å^2)	34.0	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 40.30 % 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 2.7939e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: MPD, SCN, CL, K, EDO, NSQ

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	1.00	12/5359 (0.2%)	1.01	25/7286 (0.3%)
1	B	0.98	13/5367 (0.2%)	1.08	31/7299 (0.4%)
1	C	0.97	9/5332 (0.2%)	1.07	43/7251 (0.6%)
1	D	0.99	10/5334 (0.2%)	1.07	40/7254 (0.6%)
1	E	0.96	13/5358 (0.2%)	1.02	28/7284 (0.4%)
1	F	1.00	19/5368 (0.4%)	1.03	29/7300 (0.4%)
1	G	0.93	13/5335 (0.2%)	0.97	26/7257 (0.4%)
1	H	0.95	13/5345 (0.2%)	1.01	29/7268 (0.4%)
All	All	0.97	102/42798 (0.2%)	1.03	251/58199 (0.4%)

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	A	0	1
1	F	0	1
All	All	0	2

The worst 5 of 102 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	391	GLU	CD-OE1	-12.33	1.12	1.25
1	A	135	GLU	CD-OE1	11.11	1.37	1.25
1	G	391	GLU	CD-OE1	-10.92	1.13	1.25
1	C	245	LYS	CB-CG	-10.82	1.23	1.52
1	H	391	GLU	CD-OE1	-10.77	1.13	1.25

The worst 5 of 251 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	283	ARG	NE-CZ-NH2	23.38	131.99	120.30
1	B	283	ARG	NE-CZ-NH1	-22.20	109.20	120.30
1	F	17	ARG	NE-CZ-NH1	14.73	127.67	120.30
1	B	661	ARG	NE-CZ-NH1	13.58	127.09	120.30
1	D	319	TYR	CB-CG-CD1	-13.00	113.20	121.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	308	GLN	Sidechain
1	F	391	GLU	Sidechain

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	5207	0	4962	19	0
1	B	5215	0	4965	11	0
1	C	5183	0	4933	18	0
1	D	5187	0	4931	14	0
1	E	5206	0	4964	9	0
1	F	5219	0	4967	17	0
1	G	5186	0	4918	6	0
1	H	5197	0	4945	18	0
2	A	6	0	0	0	0
2	D	3	0	0	0	0
2	E	3	0	0	0	0
2	F	6	0	0	0	0
2	G	3	0	0	0	0
2	H	3	0	0	0	0
3	A	24	0	42	2	0
3	B	8	0	14	0	0
3	D	16	0	28	0	0
3	E	8	0	14	0	0
3	F	16	0	28	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	1	0	0	0	0
4	F	1	0	0	0	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
5	A	8	0	12	0	0
5	B	8	0	12	0	0
5	C	4	0	6	0	0
5	D	12	0	18	6	0
5	E	4	0	6	0	0
5	F	20	0	30	0	0
5	G	12	0	18	0	0
5	H	4	0	6	0	0
6	A	24	0	5	0	0
6	B	24	0	5	2	0
6	C	24	0	5	0	0
6	D	24	0	4	0	0
6	E	24	0	4	0	0
6	F	24	0	5	0	0
6	G	24	0	5	0	0
6	H	24	0	5	0	0
7	C	1	0	0	0	0
7	D	1	0	0	0	0
8	A	298	0	0	1	0
8	B	265	0	0	1	0
8	C	238	0	0	3	0
8	D	307	0	0	1	0
8	E	272	0	0	1	0
8	F	294	0	0	2	0
8	G	231	0	0	3	0
8	H	273	0	0	3	0
All	All	44146	0	39857	102	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

The worst 5 of 102 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:F:81:LEU:HD11	1:F:88:ILE:CG2	2.04	0.87
1:F:81:LEU:CD1	1:F:88:ILE:CG2	2.57	0.82
1:H:623:TRP:CD1	1:H:642:GLU:HG3	2.17	0.79

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:561:GLU:O	1:C:565:THR:HG22	1.83	0.79
1:E:561:GLU:O	1:E:565:THR:HG22	1.84	0.78

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	660/672 (98%)	640 (97%)	20 (3%)	0	100	100
1	B	662/672 (98%)	641 (97%)	21 (3%)	0	100	100
1	C	660/672 (98%)	642 (97%)	18 (3%)	0	100	100
1	D	660/672 (98%)	637 (96%)	23 (4%)	0	100	100
1	E	660/672 (98%)	639 (97%)	21 (3%)	0	100	100
1	F	662/672 (98%)	642 (97%)	20 (3%)	0	100	100
1	G	659/672 (98%)	639 (97%)	20 (3%)	0	100	100
1	H	659/672 (98%)	638 (97%)	21 (3%)	0	100	100
All	All	5282/5376 (98%)	5118 (97%)	164 (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	533/545 (98%)	527 (99%)	6 (1%)	73	70
1	B	533/545 (98%)	524 (98%)	9 (2%)	60	53
1	C	528/545 (97%)	519 (98%)	9 (2%)	60	53
1	D	527/545 (97%)	520 (99%)	7 (1%)	69	64
1	E	532/545 (98%)	521 (98%)	11 (2%)	53	47
1	F	533/545 (98%)	521 (98%)	12 (2%)	50	44
1	G	527/545 (97%)	520 (99%)	7 (1%)	69	64
1	H	530/545 (97%)	521 (98%)	9 (2%)	60	53
All	All	4243/4360 (97%)	4173 (98%)	70 (2%)	60	56

5 of 70 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	7	LYS
1	E	611	LEU
1	H	170	TRP
1	E	77	PRO
1	E	170	TRP

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

Mol	Chain	Res	Type
1	D	300	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 51 ligands modelled in this entry, 8 are monoatomic - leaving 43 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	MPD	A	705	-	7,7,7	0.68	0	9,10,10	1.00	1 (11%)
5	EDO	F	708	-	3,3,3	0.62	0	2,2,2	0.17	0
2	SCN	G	701	-	1,2,2	0.09	0	0,1,1	0.00	-
2	SCN	E	701	-	1,2,2	0.43	0	0,1,1	0.00	-
5	EDO	A	708	-	3,3,3	0.89	0	2,2,2	0.32	0
6	NSQ	D	708	-	23,25,25	1.46	5 (21%)	32,37,37	1.82	8 (25%)
2	SCN	A	702	-	1,2,2	0.49	0	0,1,1	0.00	-
5	EDO	G	704	-	3,3,3	0.52	0	2,2,2	0.57	0
5	EDO	E	704	-	3,3,3	0.47	0	2,2,2	0.43	0
3	MPD	E	702	-	7,7,7	0.29	0	9,10,10	0.71	0
5	EDO	F	709	-	3,3,3	1.01	0	2,2,2	0.64	0
6	NSQ	G	706	-	23,25,25	1.04	1 (4%)	32,37,37	1.46	5 (15%)
2	SCN	F	702	-	1,2,2	0.09	0	0,1,1	0.00	-
5	EDO	F	706	-	3,3,3	0.43	0	2,2,2	0.75	0
3	MPD	A	704	-	7,7,7	0.76	0	9,10,10	1.27	2 (22%)
2	SCN	F	701	-	1,2,2	0.01	0	0,1,1	0.00	-
5	EDO	G	703	-	3,3,3	0.51	0	2,2,2	0.20	0
2	SCN	A	701	-	1,2,2	0.23	0	0,1,1	0.00	-
6	NSQ	C	703	-	23,25,25	1.49	4 (17%)	32,37,37	2.18	6 (18%)
2	SCN	D	702	-	1,2,2	0.25	0	0,1,1	0.00	-
5	EDO	B	704	-	3,3,3	1.05	0	2,2,2	0.31	0
6	NSQ	B	705	-	23,25,25	1.44	4 (17%)	32,37,37	1.45	5 (15%)
3	MPD	F	703	-	7,7,7	0.47	0	9,10,10	0.38	0
5	EDO	F	707	-	3,3,3	0.71	0	2,2,2	0.10	0
5	EDO	F	710	-	3,3,3	0.56	0	2,2,2	0.42	0
5	EDO	G	705	-	3,3,3	0.43	0	2,2,2	0.37	0
5	EDO	C	702	-	3,3,3	0.52	0	2,2,2	0.36	0
5	EDO	D	706	-	3,3,3	0.87	0	2,2,2	0.72	0
3	MPD	F	704	-	7,7,7	0.48	0	9,10,10	0.60	0
3	MPD	D	704	-	7,7,7	0.56	0	9,10,10	1.05	0
3	MPD	A	703	-	7,7,7	0.45	0	9,10,10	0.79	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	EDO	A	707	-	3,3,3	0.53	0	2,2,2	0.29	0
5	EDO	D	707	-	3,3,3	0.88	0	2,2,2	0.99	0
5	EDO	B	703	-	3,3,3	0.54	0	2,2,2	0.69	0
3	MPD	D	703	-	7,7,7	0.33	0	9,10,10	0.54	0
2	SCN	H	701	-	1,2,2	0.48	0	0,1,1	0.00	-
5	EDO	H	703	-	3,3,3	0.99	0	2,2,2	0.47	0
6	NSQ	E	705	-	23,25,25	1.16	2 (8%)	32,37,37	1.69	5 (15%)
5	EDO	D	705	-	3,3,3	0.30	0	2,2,2	0.37	0
6	NSQ	H	704	-	23,25,25	1.09	3 (13%)	32,37,37	1.35	4 (12%)
3	MPD	B	701	-	7,7,7	0.57	0	9,10,10	1.21	1 (11%)
6	NSQ	F	711	-	23,25,25	1.31	4 (17%)	32,37,37	1.26	2 (6%)
6	NSQ	A	709	-	23,25,25	1.51	6 (26%)	32,37,37	2.01	7 (21%)

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
3	MPD	A	705	-	-	3/5/5/5	-
5	EDO	F	708	-	-	0/1/1/1	-
5	EDO	A	708	-	-	1/1/1/1	-
6	NSQ	D	708	-	-	4/11/33/33	0/2/2/2
5	EDO	G	704	-	-	0/1/1/1	-
5	EDO	E	704	-	-	1/1/1/1	-
3	MPD	E	702	-	-	1/5/5/5	-
5	EDO	F	709	-	-	1/1/1/1	-
6	NSQ	G	706	-	-	2/11/33/33	0/2/2/2
5	EDO	F	706	-	-	0/1/1/1	-
3	MPD	A	704	-	-	1/5/5/5	-
5	EDO	G	703	-	-	0/1/1/1	-
6	NSQ	C	703	-	-	4/11/33/33	0/2/2/2
5	EDO	B	704	-	-	1/1/1/1	-
6	NSQ	B	705	-	-	2/11/33/33	0/2/2/2
3	MPD	F	703	-	-	0/5/5/5	-
5	EDO	F	707	-	-	0/1/1/1	-
5	EDO	F	710	-	-	1/1/1/1	-
5	EDO	G	705	-	-	1/1/1/1	-
5	EDO	C	702	-	-	0/1/1/1	-
5	EDO	D	706	-	-	1/1/1/1	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	MPD	F	704	-	-	1/5/5/5	-
3	MPD	D	704	-	-	3/5/5/5	-
3	MPD	A	703	-	-	0/5/5/5	-
5	EDO	A	707	-	-	1/1/1/1	-
5	EDO	D	707	-	-	0/1/1/1	-
5	EDO	B	703	-	-	0/1/1/1	-
3	MPD	D	703	-	-	2/5/5/5	-
5	EDO	H	703	-	-	1/1/1/1	-
6	NSQ	E	705	-	-	2/11/33/33	0/2/2/2
5	EDO	D	705	-	-	1/1/1/1	-
6	NSQ	H	704	-	-	2/11/33/33	0/2/2/2
3	MPD	B	701	-	-	1/5/5/5	-
6	NSQ	F	711	-	-	2/11/33/33	0/2/2/2
6	NSQ	A	709	-	-	2/11/33/33	0/2/2/2

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	709	NSQ	O1-C1	4.30	1.47	1.41
6	C	703	NSQ	O1-C1	4.10	1.47	1.41
6	B	705	NSQ	O1-C1	3.68	1.46	1.41
6	D	708	NSQ	O34-N32	3.67	1.29	1.22
6	C	703	NSQ	O34-N32	3.24	1.28	1.22

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	709	NSQ	O12-S10-C6	-6.56	99.14	106.94
6	C	703	NSQ	C29-C28-N32	6.47	124.25	119.38
6	C	703	NSQ	O12-S10-C6	5.98	114.05	106.94
6	E	705	NSQ	O11-S10-C6	5.06	113.81	105.74
6	D	708	NSQ	C27-C28-N32	5.04	123.17	119.38

There are no chirality outliers.

5 of 42 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	D	708	NSQ	C27-C28-N32-O34
6	D	708	NSQ	C29-C28-N32-O34
6	C	703	NSQ	C27-C28-N32-O34

Continued on next page...

Continued from previous page...

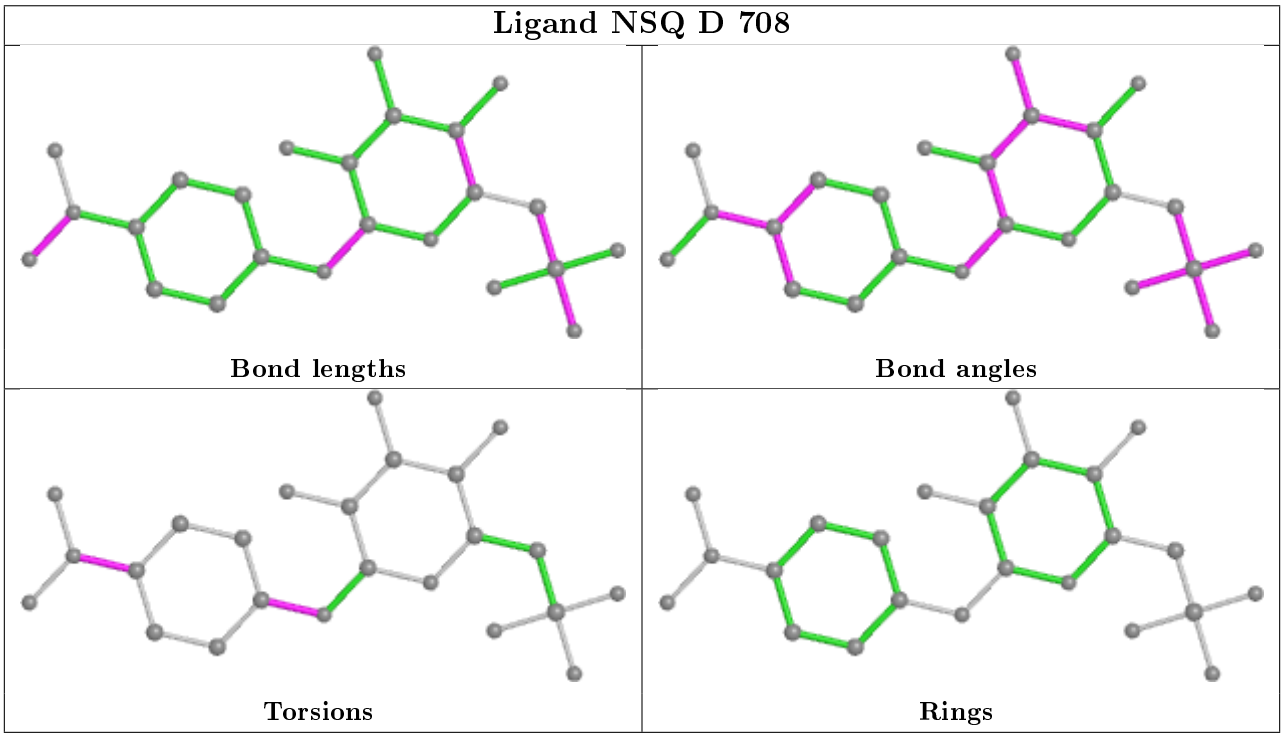
Mol	Chain	Res	Type	Atoms
6	C	703	NSQ	C29-C28-N32-O34
3	D	704	MPD	C2-C3-C4-O4

There are no ring outliers.

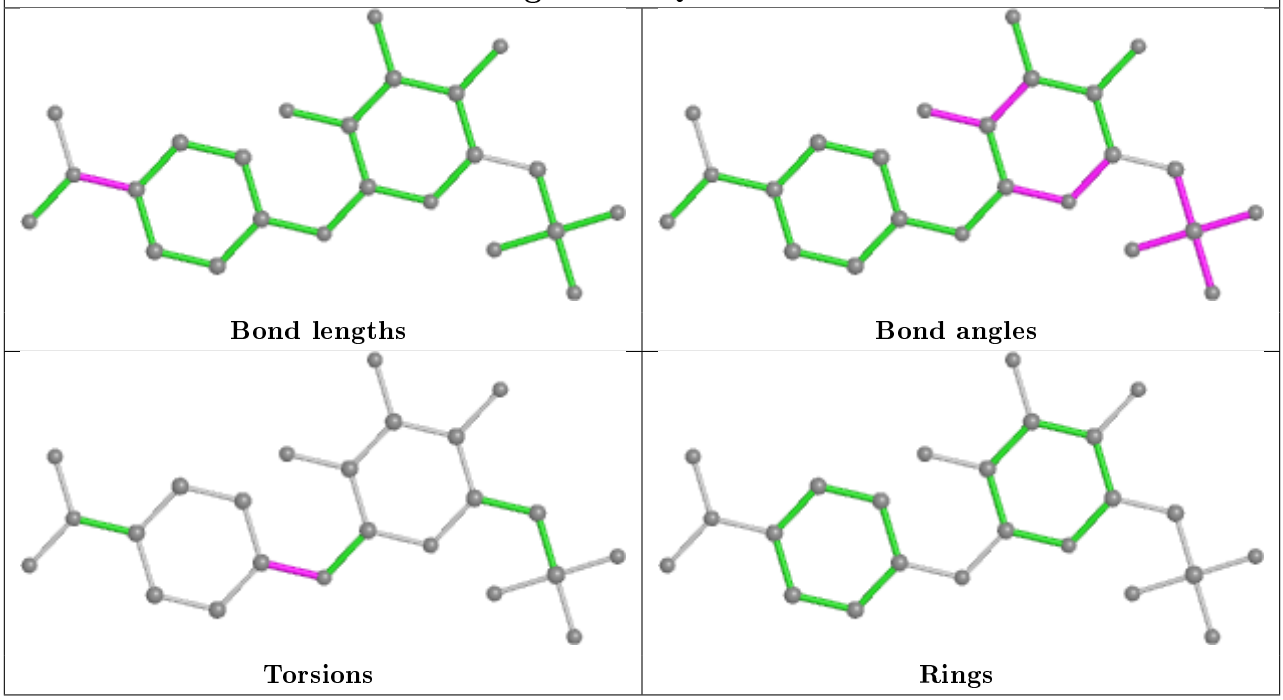
4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	704	MPD	2	0
6	B	705	NSQ	2	0
5	D	706	EDO	1	0
5	D	707	EDO	5	0

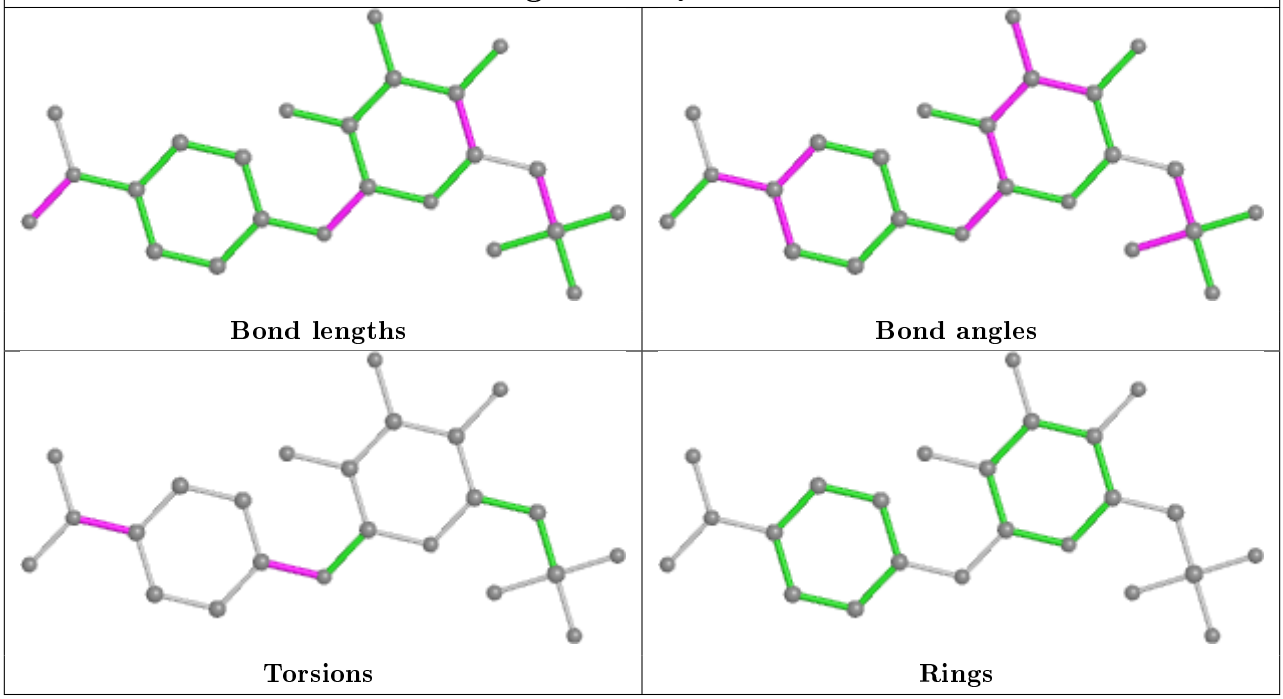
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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



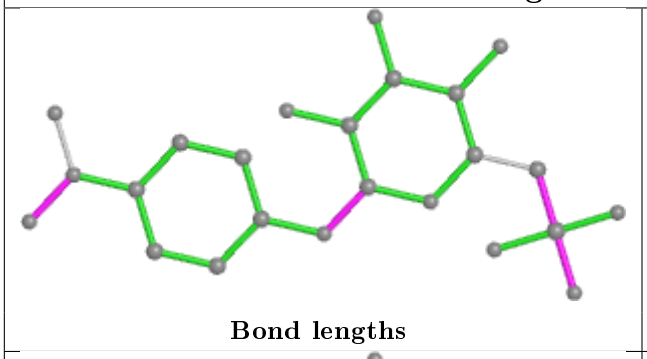
Ligand NSQ G 706



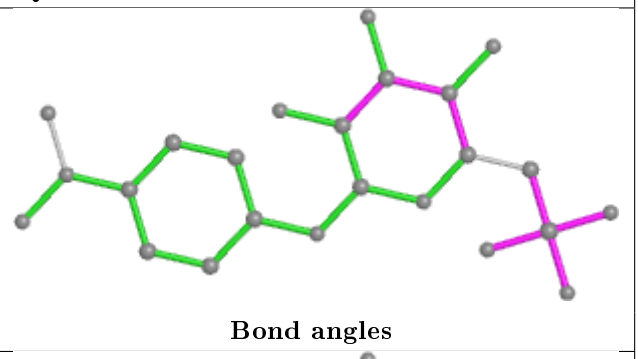
Ligand NSQ C 703



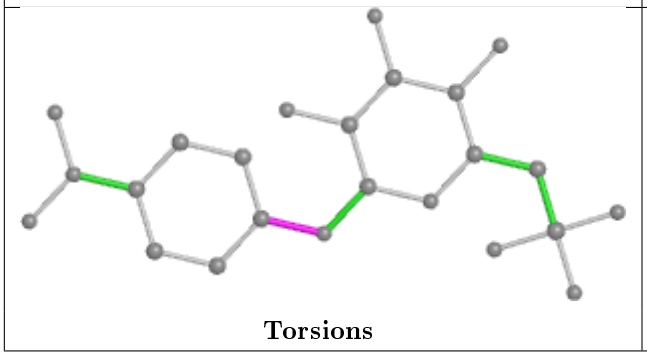
Ligand NSQ B 705



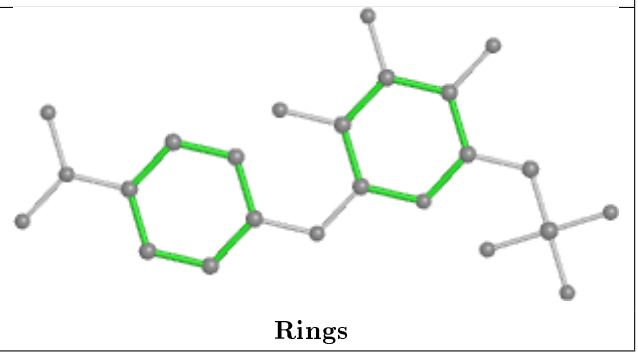
Bond lengths



Bond angles

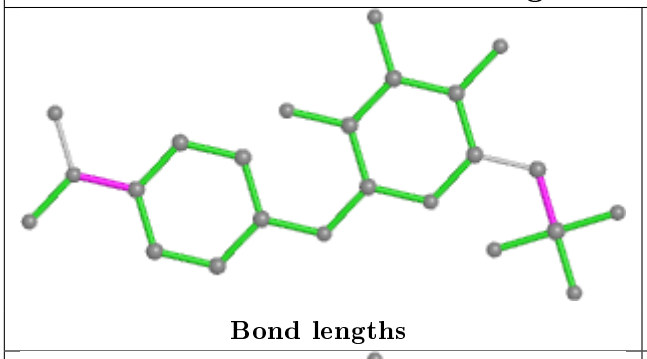


Torsions

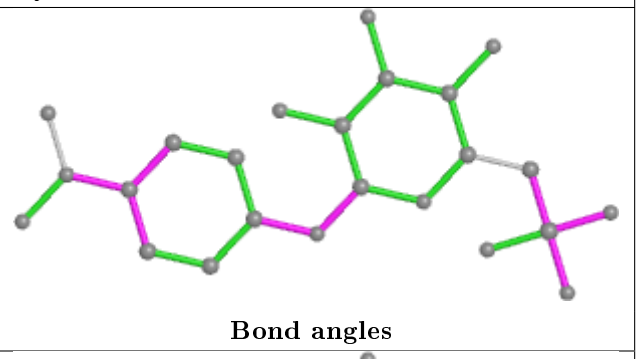


Rings

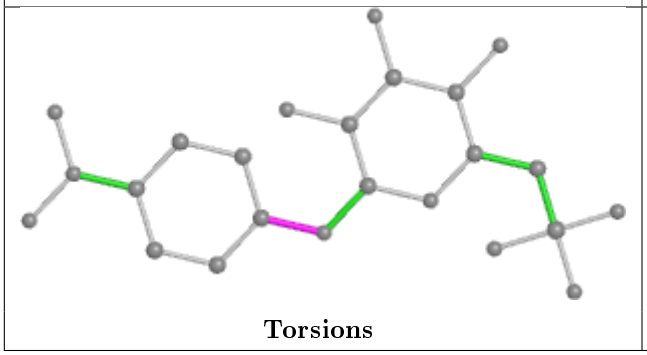
Ligand NSQ E 705



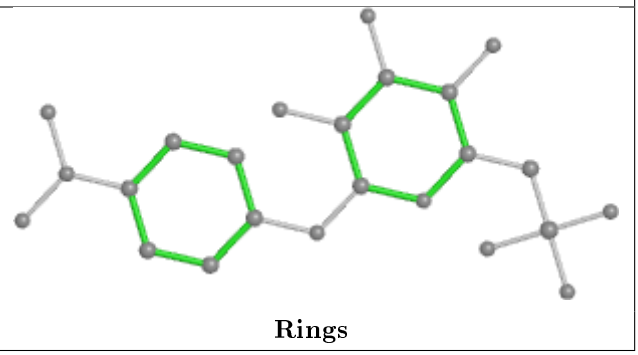
Bond lengths



Bond angles

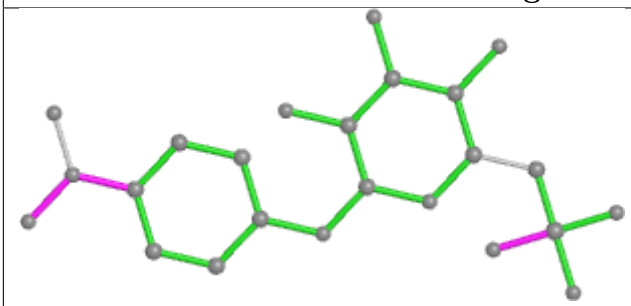


Torsions

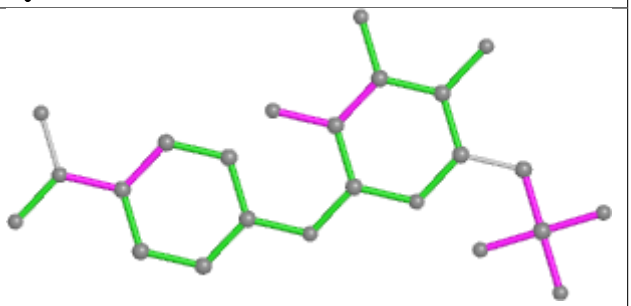


Rings

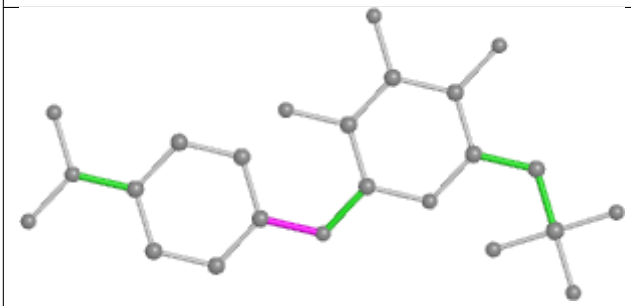
Ligand NSQ H 704



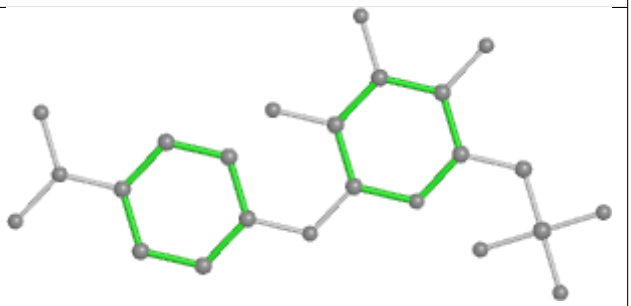
Bond lengths



Bond angles

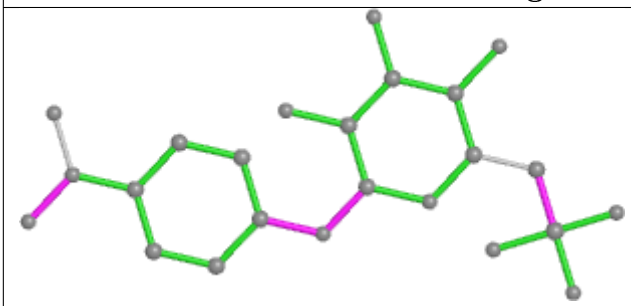


Torsions

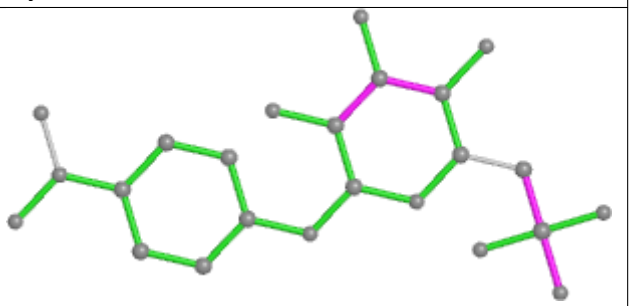


Rings

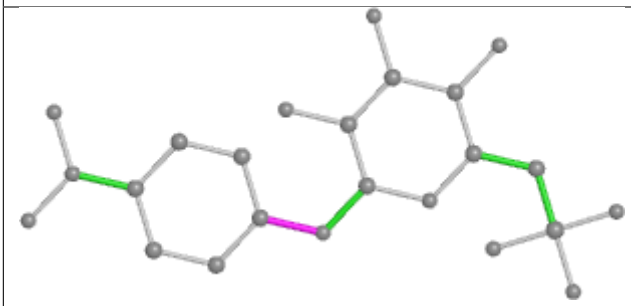
Ligand NSQ F 711



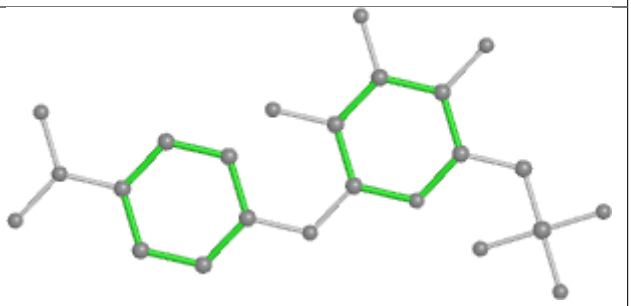
Bond lengths



Bond angles

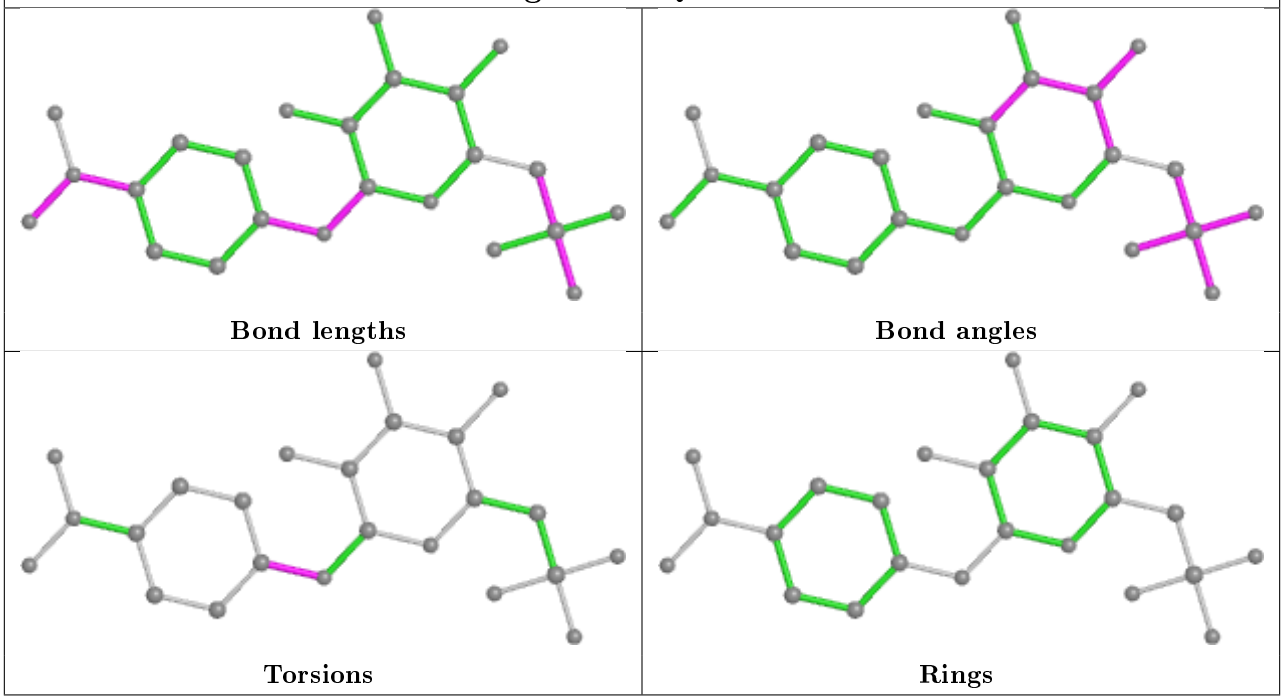


Torsions



Rings

Ligand NSQ A 709



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	661/672 (98%)	-0.27	2 (0%) 94 94	20, 30, 49, 77	0
1	B	663/672 (98%)	-0.23	4 (0%) 89 90	20, 32, 55, 78	0
1	C	661/672 (98%)	0.05	17 (2%) 56 58	22, 38, 69, 92	1 (0%)
1	D	662/672 (98%)	-0.24	2 (0%) 94 94	20, 31, 50, 75	0
1	E	661/672 (98%)	-0.21	1 (0%) 95 95	20, 32, 50, 70	0
1	F	664/672 (98%)	-0.17	0 100 100	19, 28, 47, 70	0
1	G	661/672 (98%)	-0.21	2 (0%) 94 94	21, 36, 59, 89	0
1	H	661/672 (98%)	-0.18	3 (0%) 91 91	20, 33, 58, 93	0
All	All	5294/5376 (98%)	-0.18	31 (0%) 89 90	19, 32, 57, 93	1 (0%)

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	8	ASP	3.3
1	D	8	ASP	3.1
1	C	660	LEU	3.1
1	C	630	GLY	2.9
1	C	604	ALA	2.8

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
5	EDO	D	706	4/4	0.69	0.24	49,51,53,54	0
5	EDO	D	707	4/4	0.76	0.20	39,40,42,43	0
5	EDO	F	707	4/4	0.77	0.20	32,39,41,43	4
3	MPD	D	704	8/8	0.80	0.21	49,56,58,62	0
5	EDO	H	703	4/4	0.80	0.15	40,43,44,46	0
5	EDO	F	708	4/4	0.81	0.17	46,46,49,50	0
3	MPD	A	705	8/8	0.81	0.23	38,43,44,47	8
5	EDO	F	709	4/4	0.83	0.16	41,42,43,48	0
5	EDO	B	703	4/4	0.85	0.14	46,48,49,50	0
5	EDO	A	708	4/4	0.85	0.13	38,39,40,44	0
3	MPD	B	701	8/8	0.86	0.19	42,48,52,54	0
5	EDO	D	705	4/4	0.87	0.21	40,41,42,44	4
5	EDO	E	704	4/4	0.87	0.19	31,32,32,35	4
3	MPD	A	703	8/8	0.89	0.16	39,43,45,58	0
5	EDO	G	704	4/4	0.89	0.12	47,49,50,51	0
4	CL	H	702	1/1	0.90	0.09	44,44,44,44	1
5	EDO	B	704	4/4	0.91	0.13	39,39,39,41	0
3	MPD	F	704	8/8	0.91	0.15	51,56,58,58	0
2	SCN	A	702	3/3	0.93	0.11	33,33,38,48	3
5	EDO	G	705	4/4	0.93	0.14	35,42,46,47	0
3	MPD	A	704	8/8	0.93	0.15	32,38,42,45	0
3	MPD	E	702	8/8	0.93	0.10	26,29,32,32	0
5	EDO	F	710	4/4	0.94	0.13	43,43,45,46	0
3	MPD	F	703	8/8	0.94	0.14	28,35,37,38	0
4	CL	G	702	1/1	0.94	0.11	39,39,39,39	1
5	EDO	A	707	4/4	0.95	0.12	41,44,45,47	0
5	EDO	G	703	4/4	0.95	0.09	40,42,44,48	0
5	EDO	C	702	4/4	0.95	0.10	40,46,47,54	0
3	MPD	D	703	8/8	0.95	0.21	41,47,49,52	0
4	CL	E	703	1/1	0.95	0.12	48,48,48,48	1
4	CL	B	702	1/1	0.96	0.09	46,46,46,46	0
4	CL	F	705	1/1	0.96	0.09	44,44,44,44	0
2	SCN	F	702	3/3	0.97	0.11	34,34,34,39	3
2	SCN	D	702	3/3	0.97	0.10	42,42,45,49	0
2	SCN	H	701	3/3	0.97	0.11	35,35,38,41	0
6	NSQ	D	708	24/24	0.97	0.10	20,26,60,67	0
4	CL	A	706	1/1	0.98	0.04	48,48,48,48	0

Continued on next page...

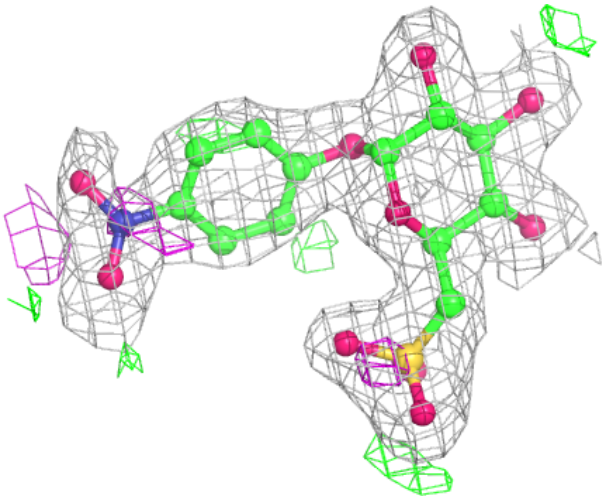
Continued from previous page...

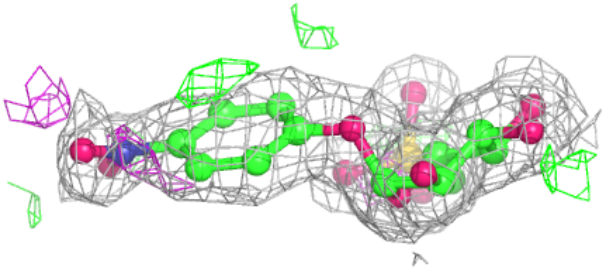
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	NSQ	F	711	24/24	0.98	0.10	19,22,29,34	0
6	NSQ	C	703	24/24	0.98	0.10	22,24,64,69	0
6	NSQ	E	705	24/24	0.98	0.10	22,26,35,42	0
5	EDO	F	706	4/4	0.98	0.06	30,32,33,35	0
6	NSQ	H	704	24/24	0.98	0.08	17,24,33,35	0
6	NSQ	B	705	24/24	0.98	0.09	19,24,37,38	0
2	SCN	A	701	3/3	0.98	0.10	33,33,36,40	0
2	SCN	E	701	3/3	0.98	0.12	41,41,53,56	0
6	NSQ	A	709	24/24	0.98	0.08	21,25,36,41	0
7	K	C	701	1/1	0.99	0.07	37,37,37,37	0
2	SCN	F	701	3/3	0.99	0.08	34,34,35,48	0
6	NSQ	G	706	24/24	0.99	0.07	20,25,29,35	0
7	K	D	701	1/1	0.99	0.06	33,33,33,33	0
2	SCN	G	701	3/3	0.99	0.10	39,39,43,47	0

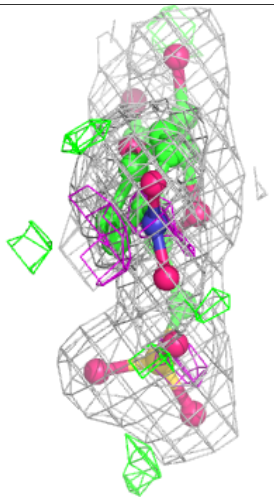
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around NSQ D 708:

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

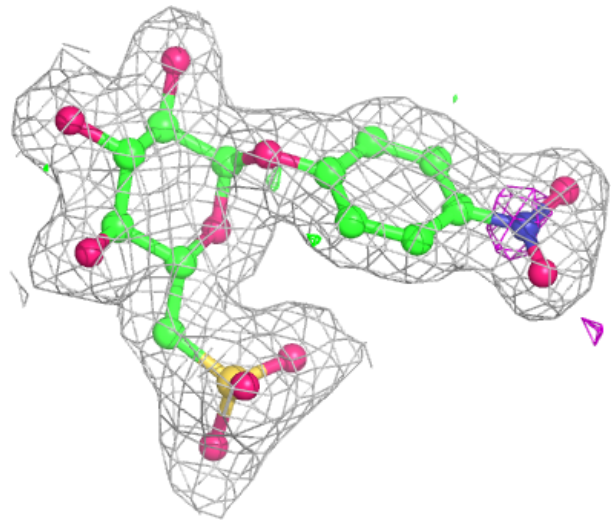


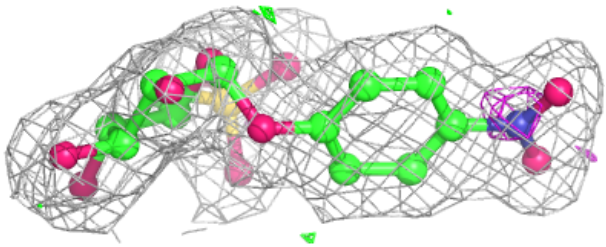


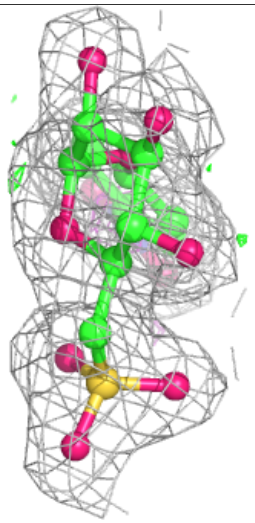


Electron density around NSQ F 711:

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

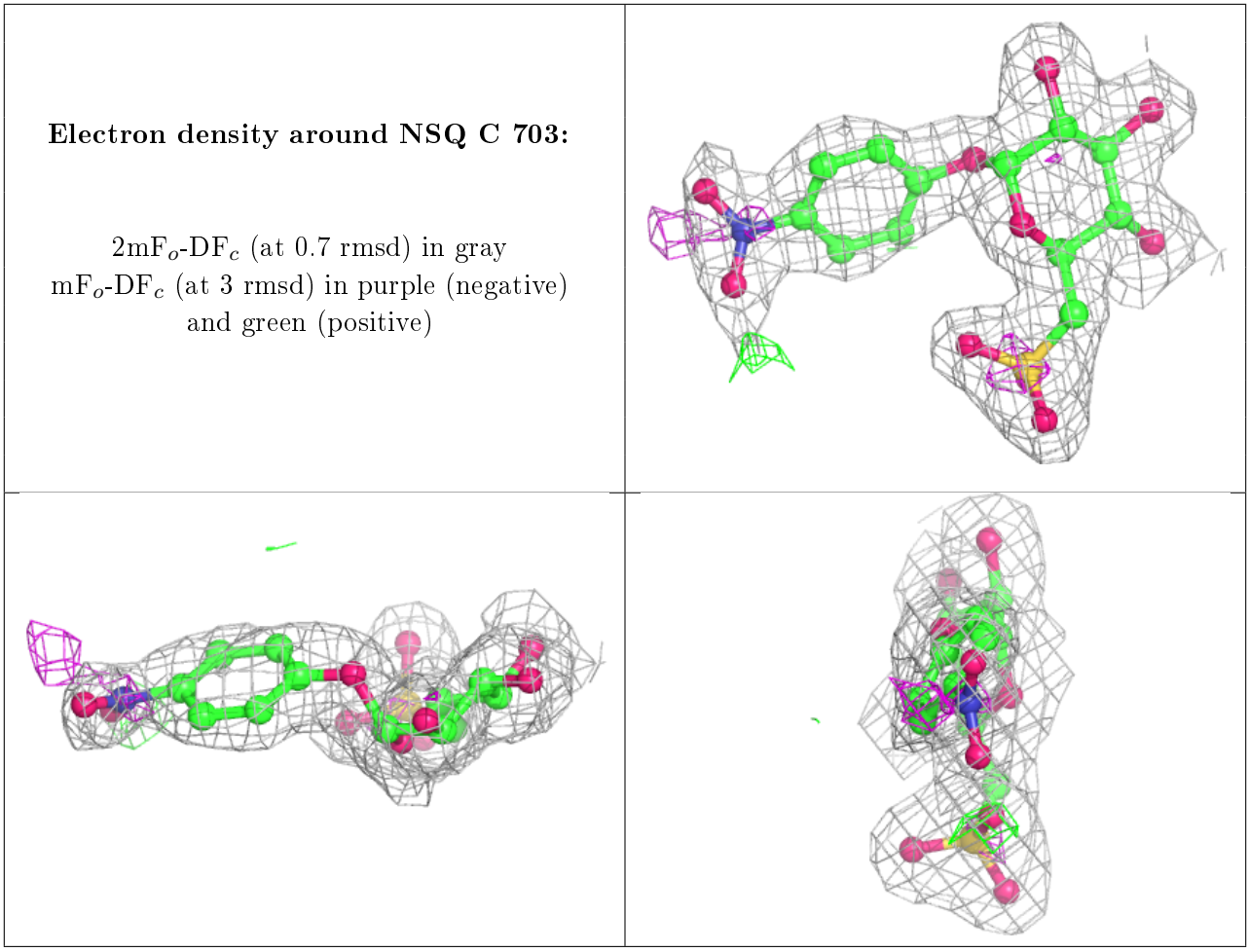






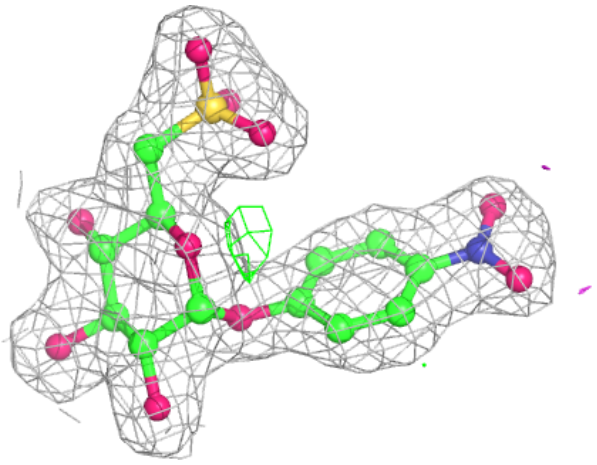
Electron density around NSQ C 703:

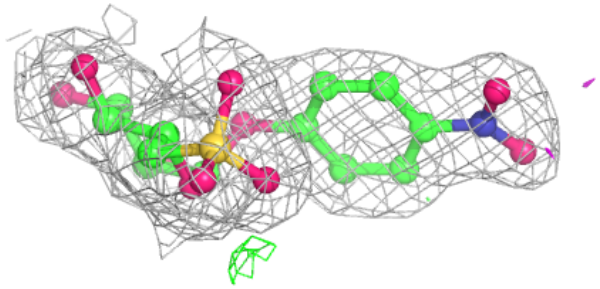
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

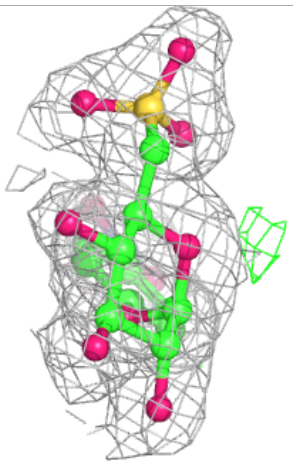


Electron density around NSQ E 705:

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

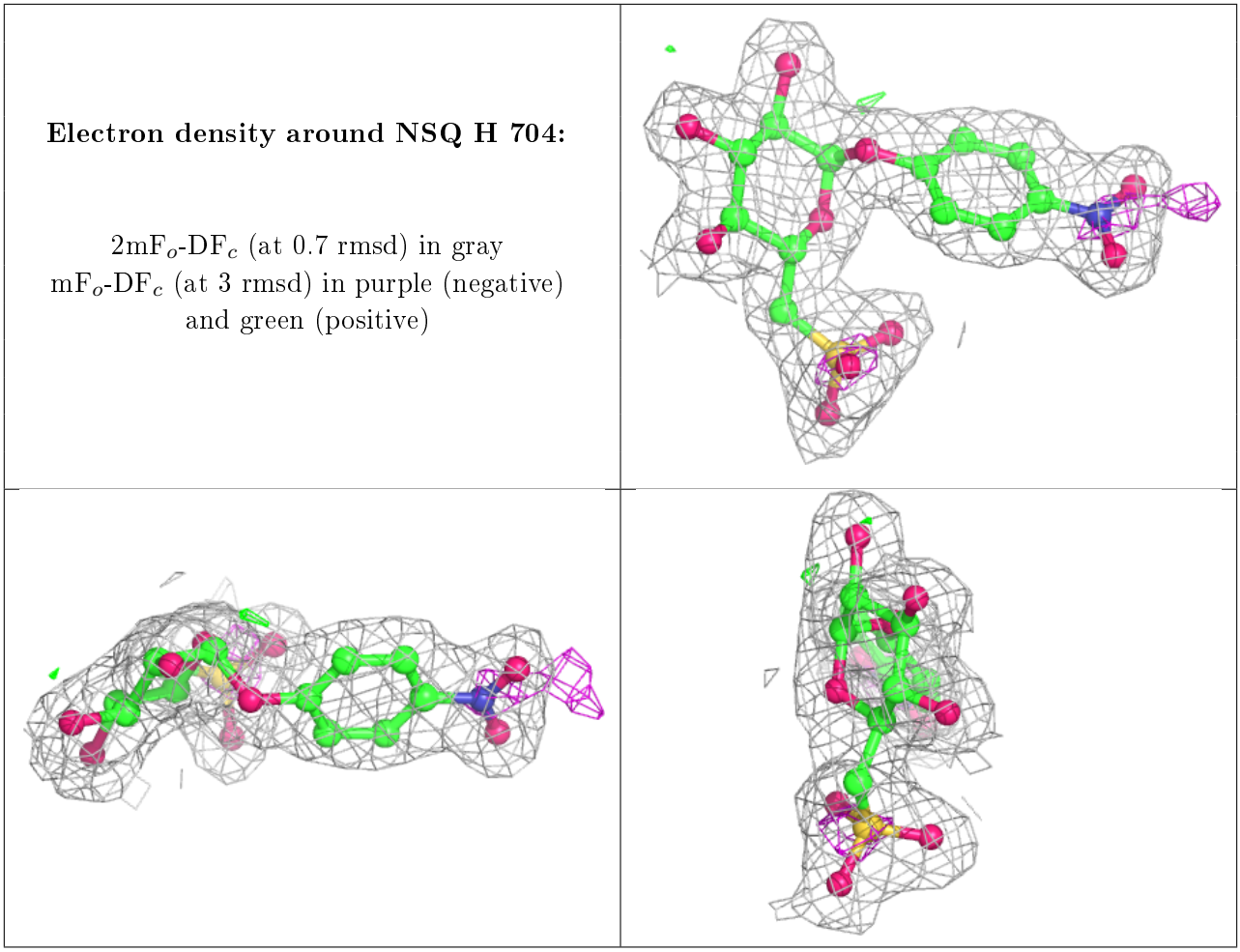






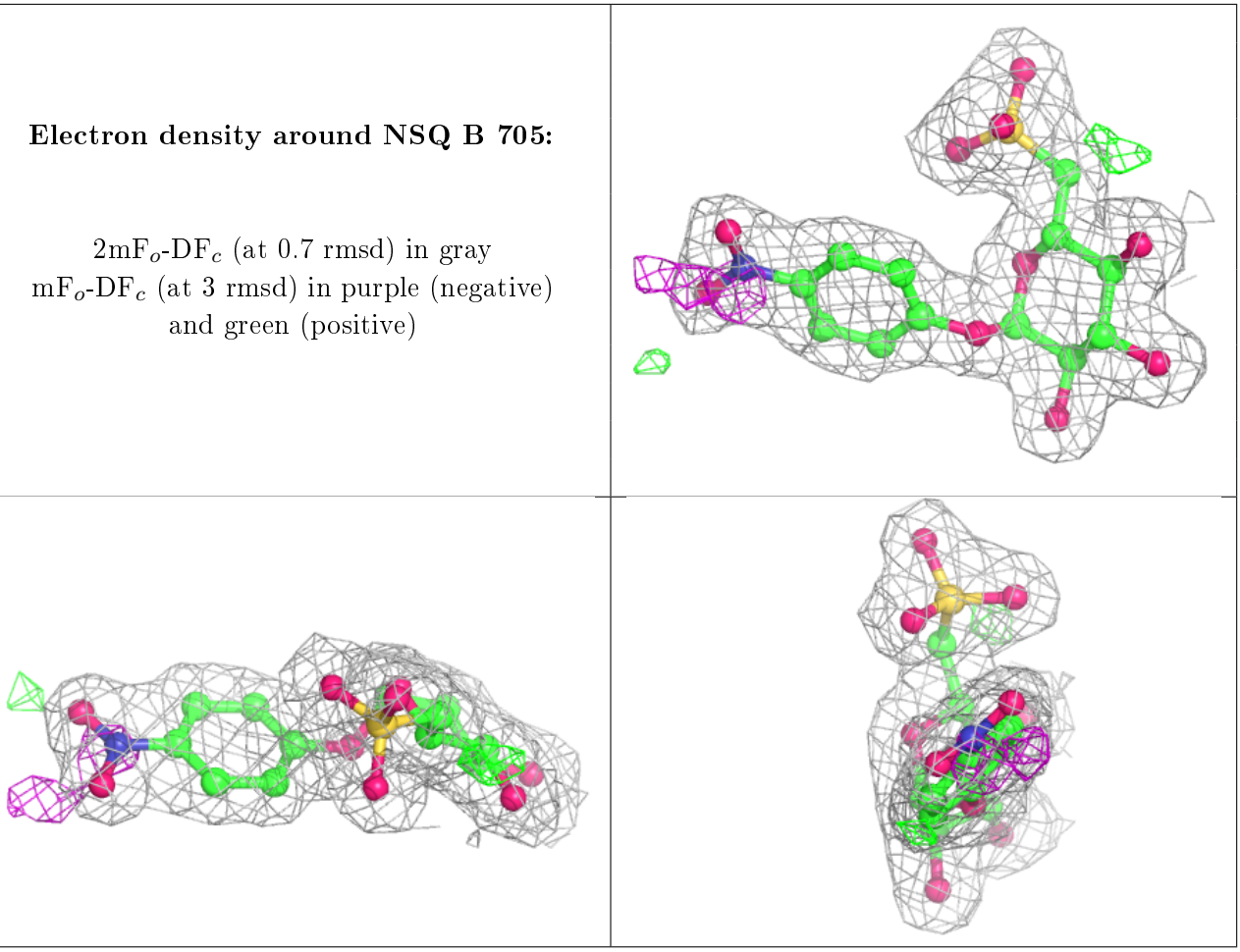
Electron density around NSQ H 704:

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



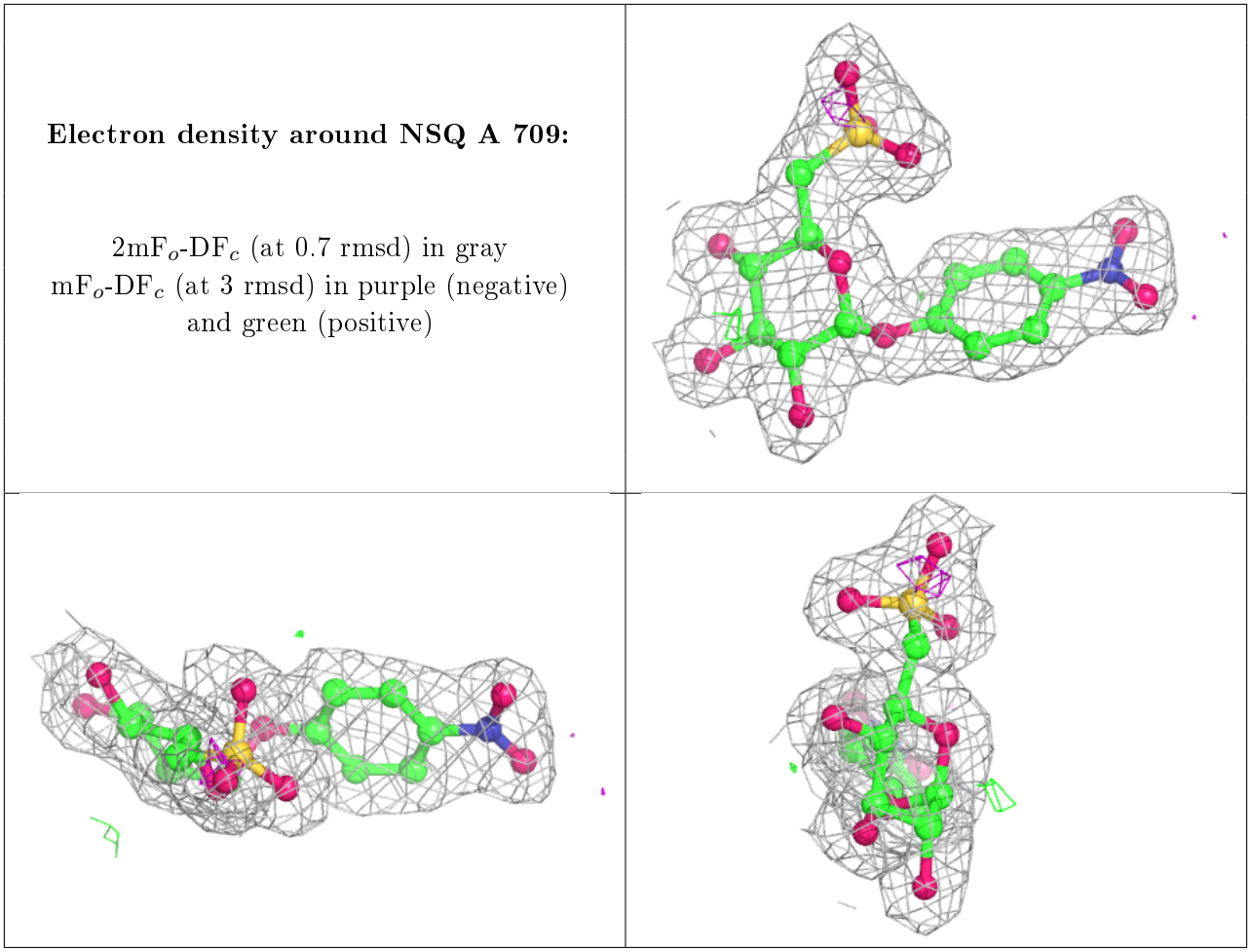
Electron density around NSQ B 705:

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



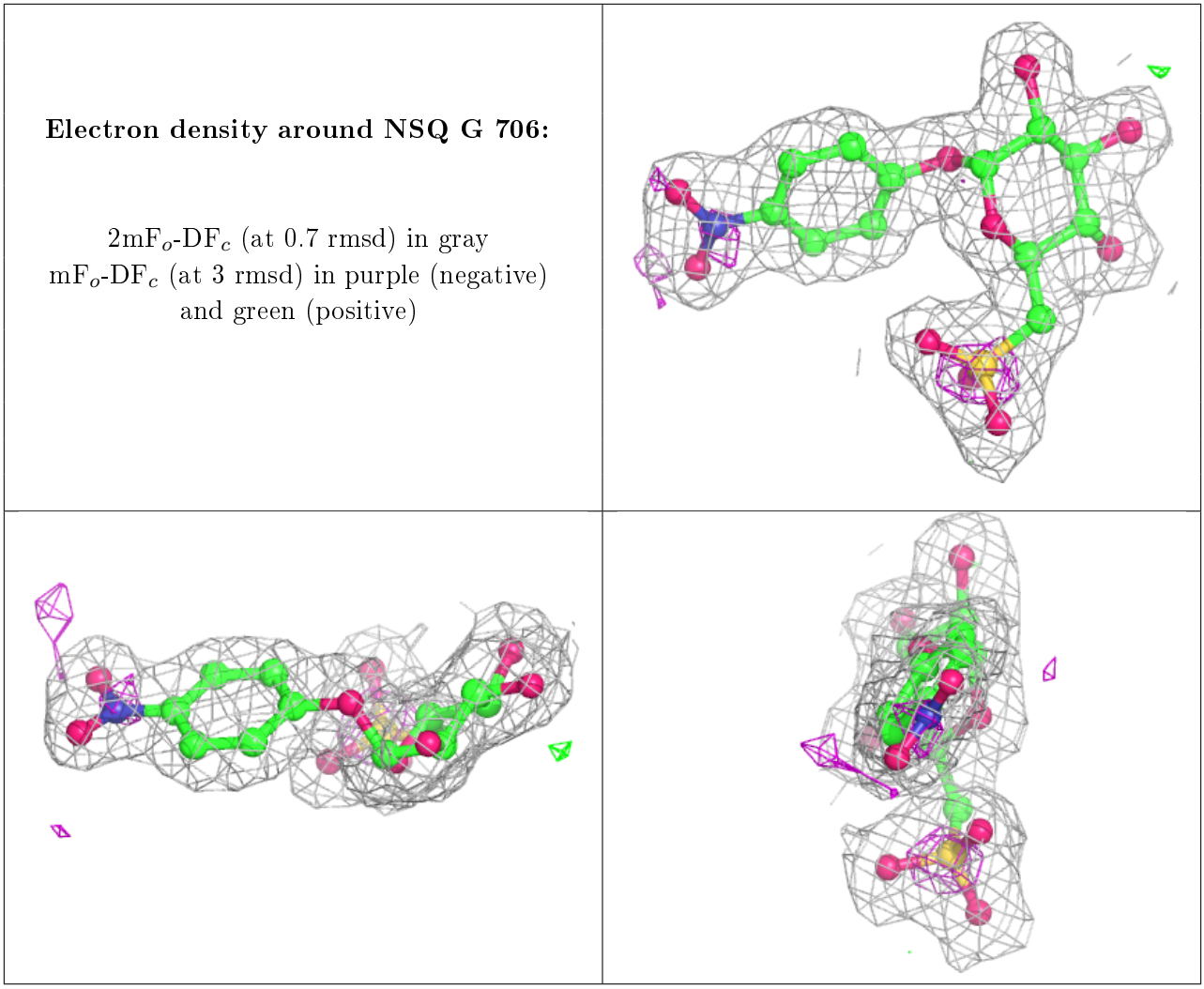
Electron density around NSQ A 709:

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



Electron density around NSQ G 706:

$2mF_o$-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.