



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 21, 2020 – 10:35 PM BST

PDB ID : 6XW7  
Title : Crystal structure of murine norovirus P domain in complex with Nanobody NB-5829 and glycochenodeoxycholate (GCDCA)  
Authors : Kilic, T.; Sabin, C.; Hansman, G.  
Deposited on : 2020-01-23  
Resolution : 2.15 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](mailto: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 : 2.13.1  
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

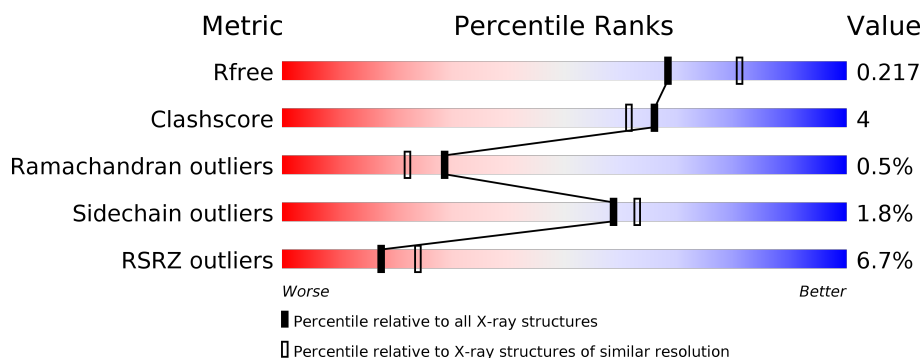
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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  $\geq 3$ , 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	309	<div> <div>2%</div> <div> <div></div> <div>94%</div> <div></div> </div> <div>• •</div> </div>
1	B	309	<div> <div>2%</div> <div> <div></div> <div>89%</div> <div></div> </div> <div>9%</div> <div>•</div> </div>
1	E	309	<div> <div>3%</div> <div> <div></div> <div>87%</div> <div></div> </div> <div>10%</div> <div>• •</div> </div>
1	F	309	<div> <div>4%</div> <div> <div></div> <div>88%</div> <div></div> </div> <div>9%</div> <div>•</div> </div>
2	C	123	<div> <div></div> <div> <div></div> <div>82%</div> <div></div> </div> <div>12%</div> <div>• 5%</div> </div>
2	D	123	<div> <div>%</div> <div> <div></div> <div>84%</div> <div></div> </div> <div>11%</div> <div>• 5%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	123	<div><div></div><div>3%</div><div>87%</div><div>8%</div><div>5%</div></div>
2	H	123	<div><div></div><div>61%</div><div>71%</div><div>20%</div><div>7%</div></div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 13940 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 Capsid protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	0	1	0
			2351	1516	384	444	7			
1	B	305	Total	C	N	O	S	0	0	0
			2344	1505	386	446	7			
1	E	303	Total	C	N	O	S	0	1	0
			2345	1509	385	444	7			
1	F	301	Total	C	N	O	S	0	0	0
			2317	1490	381	439	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	225	SER	-	expression tag	UNP Q80J94
B	225	SER	-	expression tag	UNP Q80J94
E	225	SER	-	expression tag	UNP Q80J94
F	225	SER	-	expression tag	UNP Q80J94

- Molecule 2 is a protein called Nanobody NB-5829.

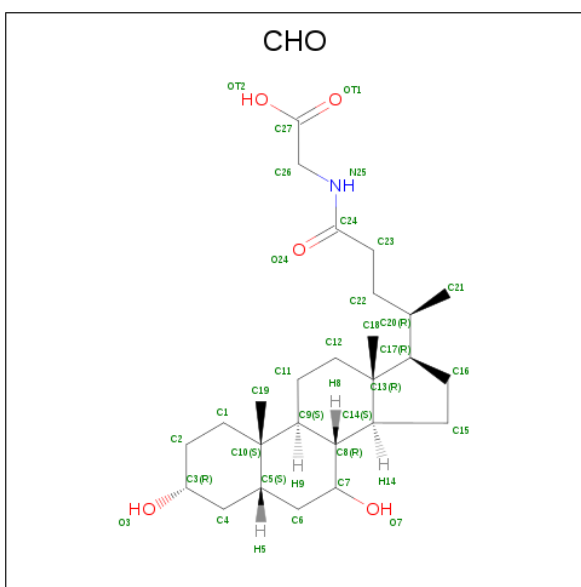
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	117	Total	C	N	O	S	0	0	0
			886	551	157	174	4			
2	D	117	Total	C	N	O	S	0	0	0
			886	551	157	174	4			
2	G	117	Total	C	N	O	S	0	0	0
			886	551	157	174	4			
2	H	114	Total	C	N	O	S	0	0	0
			865	540	153	168	4			

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	E	1	Total	C	O	0	0
			4	2	2		
3	E	1	Total	C	O	0	0
			4	2	2		
3	C	1	Total	C	O	0	0
			4	2	2		
3	D	1	Total	C	O	0	0
			4	2	2		
3	G	1	Total	C	O	0	0
			4	2	2		
3	H	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is GLYCOCHENODEOXYCHOLIC ACID (three-letter code: CHO) (formula:  $C_{26}H_{43}NO_5$ ) (labeled as "Ligand of Interest" by author).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 32	C 26	N 1	O 5	0	0
4	B	1	Total 32	C 26	N 1	O 5	0	0
4	E	1	Total 32	C 26	N 1	O 5	0	0
4	F	1	Total 32	C 26	N 1	O 5	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Mg 1 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	223	Total O 223 223	0	0
6	B	191	Total O 191 191	0	0
6	E	159	Total O 159 159	0	0
6	F	93	Total O 93 93	0	0

*Continued on next page...*

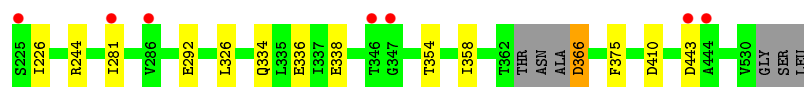
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	C	88	Total 88	O 88	0	0
6	D	81	Total 81	O 81	0	0
6	G	61	Total 61	O 61	0	0
6	H	3	Total 3	O 3	0	0

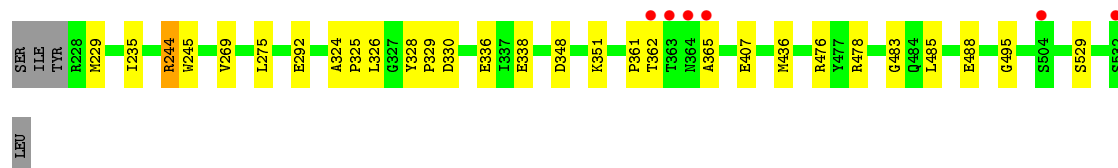
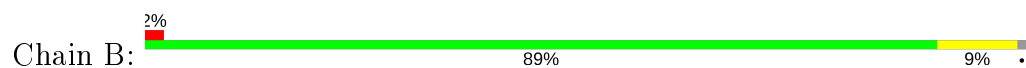
### 3 Residue-property plots [i](#)

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.

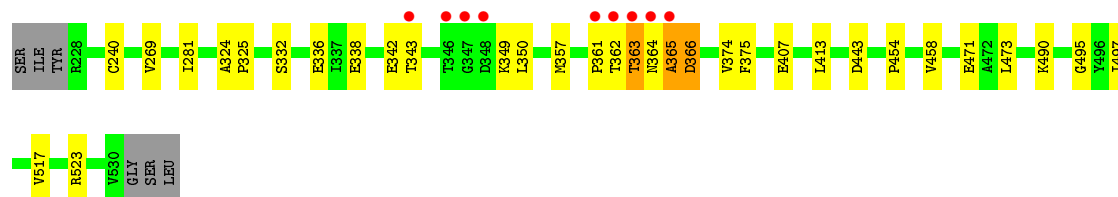
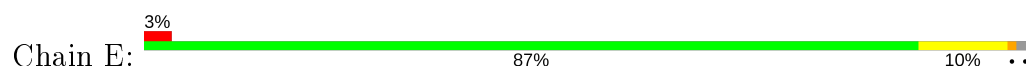
- Molecule 1: Capsid protein



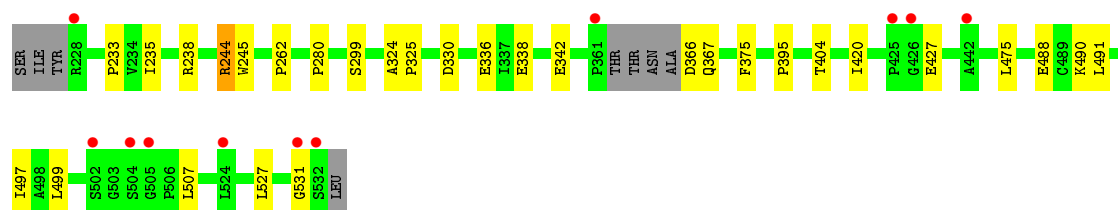
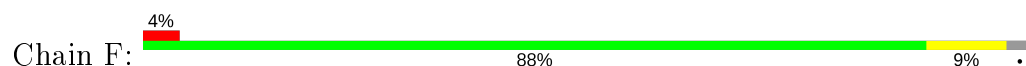
- Molecule 1: Capsid protein



- Molecule 1: Capsid protein




- Molecule 1: Capsid protein






- Molecule 2: Nanobody NB-5829

Chain C:  82% 12% • 5%




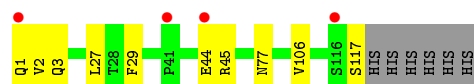
- Molecule 2: Nanobody NB-5829

Chain D:  84% 11% • 5%



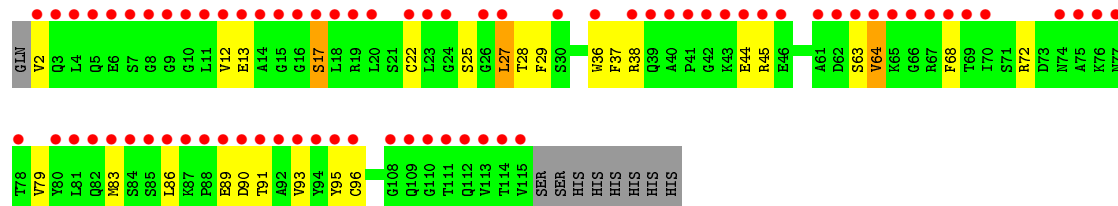
- Molecule 2: Nanobody NB-5829

Chain G:  87% 8% 5%



- Molecule 2: Nanobody NB-5829

Chain H:  61% 71% 20% • 7%



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	190.73Å 124.93Å 113.13Å 90.00° 125.30° 90.00°	Depositor
Resolution (Å)	49.21 – 2.15 49.16 – 2.15	Depositor EDS
% Data completeness (in resolution range)	97.1 (49.21-2.15) 97.2 (49.16-2.15)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.04 (at 2.14Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.171 , 0.215 0.178 , 0.217	Depositor DCC
$R_{free}$ test set	5759 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.8	Xtriage
Anisotropy	0.395	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 38.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.013 for -h-2*k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13940	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, EDO, CHO

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	0.71	0/2415	0.88	1/3305 (0.0%)
1	B	0.74	2/2407 (0.1%)	0.91	2/3296 (0.1%)
1	E	0.73	1/2409 (0.0%)	0.85	0/3299
1	F	0.69	0/2379	0.85	1/3255 (0.0%)
2	C	0.75	1/903 (0.1%)	0.90	0/1218
2	D	0.74	0/903	0.93	0/1218
2	G	0.71	0/903	0.91	0/1218
2	H	0.76	0/882	0.87	0/1190
All	All	0.72	4/13201 (0.0%)	0.88	4/17999 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	336	GLU	CD-OE1	6.68	1.32	1.25
1	E	336	GLU	CD-OE1	6.14	1.32	1.25
2	C	89	GLU	CD-OE2	5.23	1.31	1.25
1	B	488	GLU	CD-OE2	5.03	1.31	1.25

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	244	ARG	NE-CZ-NH1	9.97	125.28	120.30
1	B	244	ARG	NE-CZ-NH2	-8.95	115.83	120.30
1	F	244	ARG	NE-CZ-NH2	-5.41	117.59	120.30
1	A	244	ARG	NE-CZ-NH1	-5.26	117.67	120.30

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	2351	0	2300	14	0
1	B	2344	0	2294	18	0
1	E	2345	0	2294	22	0
1	F	2317	0	2268	19	0
2	C	886	0	856	11	0
2	D	886	0	856	9	0
2	G	886	0	856	5	0
2	H	865	0	837	20	0
3	A	8	0	12	0	0
3	C	4	0	6	0	0
3	D	4	0	6	0	0
3	E	8	0	12	0	0
3	G	4	0	6	0	0
3	H	4	0	6	0	0
4	A	32	0	42	0	0
4	B	32	0	42	0	0
4	E	32	0	42	0	0
4	F	32	0	42	0	0
5	B	1	0	0	0	0
6	A	223	0	0	3	0
6	B	191	0	0	1	0
6	C	88	0	0	1	0
6	D	81	0	0	1	0
6	E	159	0	0	1	0
6	F	93	0	0	2	0
6	G	61	0	0	0	0
6	H	3	0	0	0	0
All	All	13940	0	12777	104	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

All (104) 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)
2:C:2:VAL:HG13	2:C:27:LEU:HD23	1.61	0.80
1:A:281:ILE:HG21	1:B:235:ILE:CD1	2.17	0.74
1:A:375[A]:PHE:CZ	6:A:797:HOH:O	2.40	0.74
1:A:338:GLU:OE2	1:B:338:GLU:OE2	2.16	0.63
2:G:2:VAL:HG12	2:G:106:VAL:HG21	1.80	0.63
1:A:281:ILE:HG21	1:B:235:ILE:HD11	1.79	0.62
1:B:361:PRO:HA	1:B:407:GLU:HB2	1.82	0.61
2:H:64:VAL:HG13	2:H:68:PHE:CD1	2.37	0.59
1:F:366:ASP:O	1:F:375:PHE:HE2	1.86	0.58
2:H:64:VAL:CG1	2:H:68:PHE:CG	2.87	0.58
1:B:348:ASP:HB2	1:B:351:LYS:HE2	1.86	0.58
2:H:63:SER:O	2:H:64:VAL:HG23	2.04	0.57
1:A:281:ILE:HG21	1:B:235:ILE:HD12	1.86	0.56
2:G:45:ARG:HG2	2:G:45:ARG:HH11	1.71	0.55
1:F:507:LEU:HD22	1:F:531:GLY:HA2	1.88	0.55
1:A:334:GLN:HG3	1:A:358:ILE:HD13	1.89	0.54
1:B:476:ARG:HD3	1:B:485:LEU:HD11	1.88	0.54
1:E:281:ILE:HG21	1:F:235:ILE:CD1	2.36	0.54
1:E:361:PRO:O	1:E:363:THR:N	2.41	0.54
2:H:64:VAL:HG13	2:H:68:PHE:CG	2.43	0.54
1:B:362:THR:HB	1:B:365:ALA:HB2	1.89	0.53
2:H:17:SER:HB3	2:H:83:MET:O	2.07	0.53
2:C:27:LEU:C	2:C:27:LEU:HD12	2.30	0.52
1:A:366:ASP:HB2	1:A:410:ASP:OD1	2.11	0.50
1:A:292:GLU:HG3	6:A:905:HOH:O	2.10	0.50
1:E:473:LEU:HD22	1:E:517:VAL:HG13	1.93	0.50
1:F:262:PRO:HG2	1:F:420:ILE:HD12	1.93	0.50
1:E:443:ASP:CG	1:F:342:GLU:OE2	2.50	0.50
2:H:64:VAL:O	2:H:64:VAL:HG12	2.10	0.50
1:B:244:ARG:HD2	1:B:330:ASP:OD2	2.12	0.49
2:D:83:MET:HB3	2:D:86:LEU:HD21	1.94	0.49
1:E:365:ALA:O	1:E:366:ASP:HB2	2.13	0.49
2:C:18:LEU:HB2	6:C:801:HOH:O	2.13	0.48
2:H:22:CYS:CB	2:H:96:CYS:SG	3.01	0.48
2:D:62:ASP:HA	2:D:65:LYS:HE2	1.94	0.48
1:E:324:ALA:HB1	1:E:325:PRO:HD2	1.95	0.48
1:F:366:ASP:O	1:F:375:PHE:CE2	2.67	0.47
1:A:366:ASP:CG	1:A:366:ASP:O	2.53	0.47
2:D:89:GLU:HB3	6:D:845:HOH:O	2.15	0.47
2:C:2:VAL:HG12	2:C:106:VAL:HG21	1.97	0.47
1:E:342:GLU:HA	1:E:349:LYS:O	2.15	0.47
1:A:336:GLU:HG2	1:A:354:THR:CG2	2.44	0.47

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:332:SER:HB3	1:E:361:PRO:HD3	1.97	0.46
1:E:363:THR:HB	1:E:407:GLU:CD	2.36	0.46
1:E:490:LYS:O	1:E:497:ILE:HA	2.15	0.46
1:E:363:THR:HG22	6:E:847:HOH:O	2.15	0.46
2:C:12:VAL:HG11	2:C:86:LEU:HD13	1.99	0.45
1:B:324:ALA:HB1	1:B:325:PRO:HD2	1.99	0.45
1:F:324:ALA:HB1	1:F:325:PRO:HD2	1.98	0.45
2:H:37:PHE:O	2:H:95:TYR:N	2.43	0.45
1:E:357:MET:HE3	1:E:374:VAL:HB	1.97	0.45
2:H:64:VAL:O	2:H:64:VAL:CG1	2.64	0.45
1:B:326:LEU:CD1	2:D:102:VAL:HB	2.47	0.45
1:B:292:GLU:HB3	6:B:705:HOH:O	2.16	0.45
1:E:365:ALA:HB1	1:E:375[B]:PHE:CE2	2.52	0.45
1:F:336:GLU:HG3	6:F:1059:HOH:O	2.17	0.45
1:F:427:GLU:HA	1:F:499:LEU:O	2.16	0.45
2:H:22:CYS:SG	2:H:96:CYS:SG	3.15	0.45
2:H:27:LEU:HD13	2:H:29:PHE:HA	1.99	0.45
1:B:275:LEU:CD1	2:D:102:VAL:HG12	2.47	0.45
2:H:68:PHE:CE2	2:H:83:MET:HG2	2.52	0.44
1:A:326:LEU:HD11	2:C:102:VAL:HB	1.99	0.44
1:F:238:ARG:HD3	6:F:1005:HOH:O	2.17	0.44
2:C:2:VAL:CG1	2:C:27:LEU:HD23	2.41	0.43
1:E:363:THR:HB	1:E:407:GLU:OE2	2.17	0.43
1:B:326:LEU:HD11	2:D:102:VAL:HB	1.98	0.43
1:F:490:LYS:HG3	1:F:527:LEU:HD21	2.00	0.43
1:E:363:THR:OG1	1:E:364:ASN:N	2.49	0.43
1:E:269:VAL:HG11	1:E:458:VAL:HA	1.99	0.43
2:H:36:TRP:HA	2:H:95:TYR:O	2.18	0.43
1:B:478:ARG:NH1	1:B:483:GLY:HA2	2.33	0.43
1:E:281:ILE:HG21	1:F:235:ILE:HD12	2.00	0.43
1:E:471:GLU:HG2	1:E:523:ARG:HB2	2.01	0.43
1:F:244:ARG:HD2	1:F:330:ASP:OD2	2.19	0.43
2:H:64:VAL:CG1	2:H:68:PHE:HB2	2.49	0.43
1:E:338:GLU:OE2	1:F:338:GLU:OE2	2.36	0.43
2:H:38:ARG:HA	2:H:93:VAL:O	2.19	0.42
2:D:2:VAL:HG13	2:D:27:LEU:HD23	2.00	0.42
2:D:30:SER:O	2:D:53:ARG:HB3	2.19	0.42
2:G:45:ARG:CG	2:G:45:ARG:HH11	2.31	0.42
1:F:475:LEU:HD11	1:F:491:LEU:HB2	2.01	0.42
2:H:89:GLU:C	2:H:91:THR:H	2.22	0.42
2:C:68:PHE:HA	2:C:82:GLN:O	2.20	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:12:VAL:HG22	2:H:13:GLU:H	1.85	0.42
2:D:68:PHE:HA	2:D:82:GLN:O	2.19	0.42
1:E:240:CYS:HB3	1:E:454:PRO:HB3	2.01	0.42
1:E:350:LEU:HD23	1:E:350:LEU:HA	1.90	0.42
2:H:72:ARG:HB3	2:H:79:VAL:HG23	2.01	0.42
1:E:269:VAL:O	1:E:495:GLY:HA3	2.20	0.42
1:A:336:GLU:HG2	1:A:354:THR:HG23	2.02	0.42
1:A:226:ILE:HG22	1:F:488:GLU:HG2	2.02	0.41
2:H:63:SER:O	2:H:64:VAL:CG2	2.68	0.41
1:A:336:GLU:HB2	6:A:764:HOH:O	2.21	0.41
1:B:245:TRP:CD2	1:B:436:MET:HG2	2.56	0.41
1:B:269:VAL:O	1:B:495:GLY:HA3	2.21	0.41
1:F:490:LYS:O	1:F:497:ILE:HA	2.20	0.41
1:F:245:TRP:CH2	1:F:395:PRO:HB2	2.56	0.41
2:C:27:LEU:CD1	2:C:29:PHE:HA	2.50	0.41
2:C:62:ASP:OD1	2:C:65:LYS:HE2	2.21	0.41
2:G:29:PHE:CD2	2:G:77:ASN:HA	2.56	0.41
2:H:2:VAL:HG13	2:H:27:LEU:HD23	2.02	0.41
2:C:1:GLN:NE2	2:G:3:GLN:NE2	2.69	0.40
1:F:233:PRO:HB2	1:F:235:ILE:HG22	2.04	0.40
1:B:328:TYR:CG	1:B:329:PRO:HD2	2.57	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	300/309 (97%)	288 (96%)	11 (4%)	1 (0%)	41	37
1	B	303/309 (98%)	293 (97%)	10 (3%)	0	100	100
1	E	302/309 (98%)	284 (94%)	14 (5%)	4 (1%)	12	6
1	F	297/309 (96%)	288 (97%)	9 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	115/123 (94%)	112 (97%)	3 (3%)	0	100	100
2	D	115/123 (94%)	112 (97%)	3 (3%)	0	100	100
2	G	115/123 (94%)	112 (97%)	3 (3%)	0	100	100
2	H	112/123 (91%)	94 (84%)	14 (12%)	4 (4%)	3	0
All	All	1659/1728 (96%)	1583 (95%)	67 (4%)	9 (0%)	29	22

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	362	THR
1	E	365	ALA
1	E	366	ASP
1	E	363	THR
2	H	64	VAL
2	H	90	ASP
2	H	45	ARG
1	A	443	ASP
2	H	86	LEU

### 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	252/255 (99%)	251 (100%)	1 (0%)	91	93
1	B	251/255 (98%)	249 (99%)	2 (1%)	81	86
1	E	251/255 (98%)	249 (99%)	2 (1%)	81	86
1	F	248/255 (97%)	244 (98%)	4 (2%)	62	67
2	C	93/99 (94%)	90 (97%)	3 (3%)	39	38
2	D	93/99 (94%)	90 (97%)	3 (3%)	39	38
2	G	93/99 (94%)	89 (96%)	4 (4%)	29	27
2	H	90/99 (91%)	85 (94%)	5 (6%)	21	16

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1371/1416 (97%)	1347 (98%)	24 (2%)	59 63

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	366	ASP
1	B	229	MET
1	B	529	SER
1	E	343	THR
1	E	413	LEU
1	F	280	PRO
1	F	299	SER
1	F	367	GLN
1	F	404	THR
2	C	7	SER
2	C	28	THR
2	C	62	ASP
2	D	27	LEU
2	D	28	THR
2	D	44	GLU
2	G	1	GLN
2	G	27	LEU
2	G	44	GLU
2	G	117	SER
2	H	17	SER
2	H	25	SER
2	H	27	LEU
2	H	28	THR
2	H	44	GLU

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

Mol	Chain	Res	Type
1	E	364	ASN
1	E	455	GLN
2	C	112	GLN
2	D	39	GLN
2	D	112	GLN
2	G	3	GLN
2	G	112	GLN
2	H	3	GLN

### 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 ⓘ

Of 13 ligands modelled in this entry, 1 is monoatomic - leaving 12 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	$\# Z  > 2$
4	CHO	F	900	-	32,35,35	0.55	0	51,54,54	0.83	1 (1%)
3	EDO	C	700	-	3,3,3	0.56	0	2,2,2	0.44	0
3	EDO	G	700	-	3,3,3	0.31	0	2,2,2	0.38	0
3	EDO	D	700	-	3,3,3	0.43	0	2,2,2	0.42	0
4	CHO	A	603	-	32,35,35	0.47	0	51,54,54	0.92	2 (3%)
4	CHO	B	601	-	32,35,35	0.48	0	51,54,54	0.84	1 (1%)
3	EDO	A	601	-	3,3,3	0.31	0	2,2,2	0.36	0
3	EDO	H	700	-	3,3,3	0.11	0	2,2,2	0.09	0
3	EDO	A	602	-	3,3,3	0.16	0	2,2,2	0.42	0
3	EDO	E	602	-	3,3,3	0.29	0	2,2,2	0.30	0
4	CHO	E	603	-	32,35,35	0.42	0	51,54,54	1.00	1 (1%)
3	EDO	E	601	-	3,3,3	0.26	0	2,2,2	0.35	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	CHO	F	900	-	-	5/12/75/75	0/4/4/4
3	EDO	C	700	-	-	1/1/1/1	-
3	EDO	G	700	-	-	1/1/1/1	-
3	EDO	D	700	-	-	0/1/1/1	-
4	CHO	A	603	-	-	7/12/75/75	0/4/4/4
4	CHO	B	601	-	-	5/12/75/75	0/4/4/4
3	EDO	A	601	-	-	0/1/1/1	-
3	EDO	H	700	-	-	0/1/1/1	-
3	EDO	A	602	-	-	0/1/1/1	-
3	EDO	E	602	-	-	0/1/1/1	-
4	CHO	E	603	-	-	4/12/75/75	0/4/4/4
3	EDO	E	601	-	-	0/1/1/1	-

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	603	CHO	C16-C17-C20	3.74	117.93	112.15
4	A	603	CHO	C16-C17-C20	3.26	117.19	112.15
4	B	601	CHO	C16-C17-C20	2.87	116.58	112.15
4	A	603	CHO	C27-C26-N25	2.31	114.89	110.43
4	F	900	CHO	C4-C5-C10	2.28	115.08	112.66

There are no chirality outliers.

All (23) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	E	603	CHO	C16-C17-C20-C21
4	E	603	CHO	C13-C17-C20-C21
4	E	603	CHO	C13-C17-C20-C22
4	F	900	CHO	C13-C17-C20-C22
4	B	601	CHO	C13-C17-C20-C22
4	E	603	CHO	C16-C17-C20-C22
4	A	603	CHO	C13-C17-C20-C22
4	F	900	CHO	C16-C17-C20-C21
4	B	601	CHO	C16-C17-C20-C21
4	A	603	CHO	C16-C17-C20-C21
4	F	900	CHO	C13-C17-C20-C21
4	B	601	CHO	C13-C17-C20-C21
4	A	603	CHO	C13-C17-C20-C21
3	G	700	EDO	O1-C1-C2-O2

*Continued on next page...*

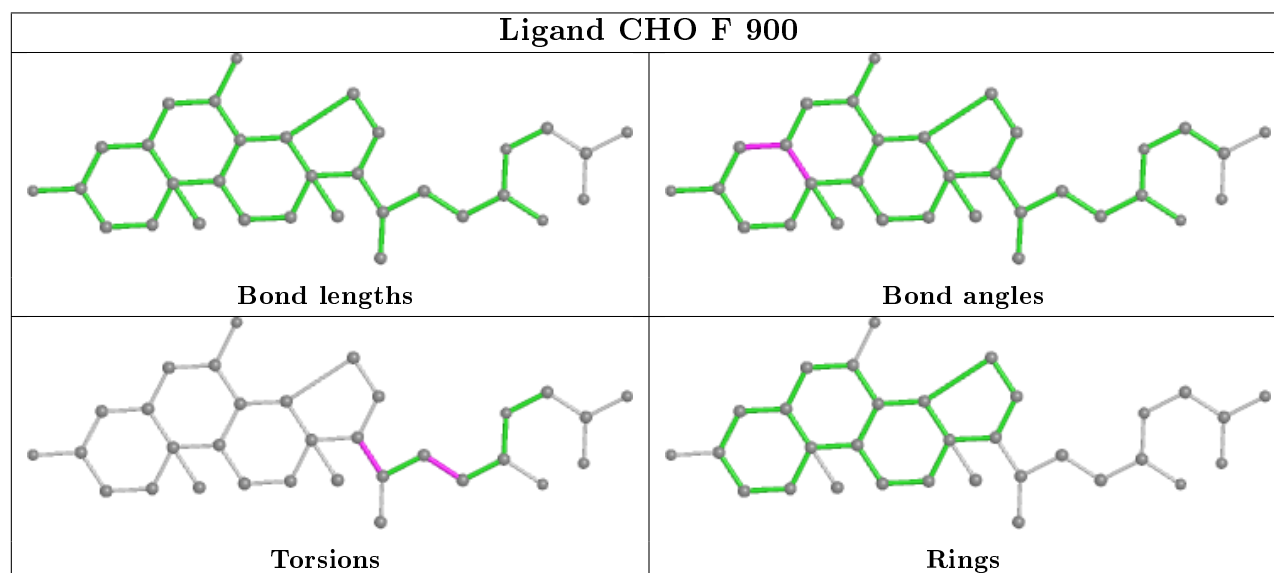
*Continued from previous page...*

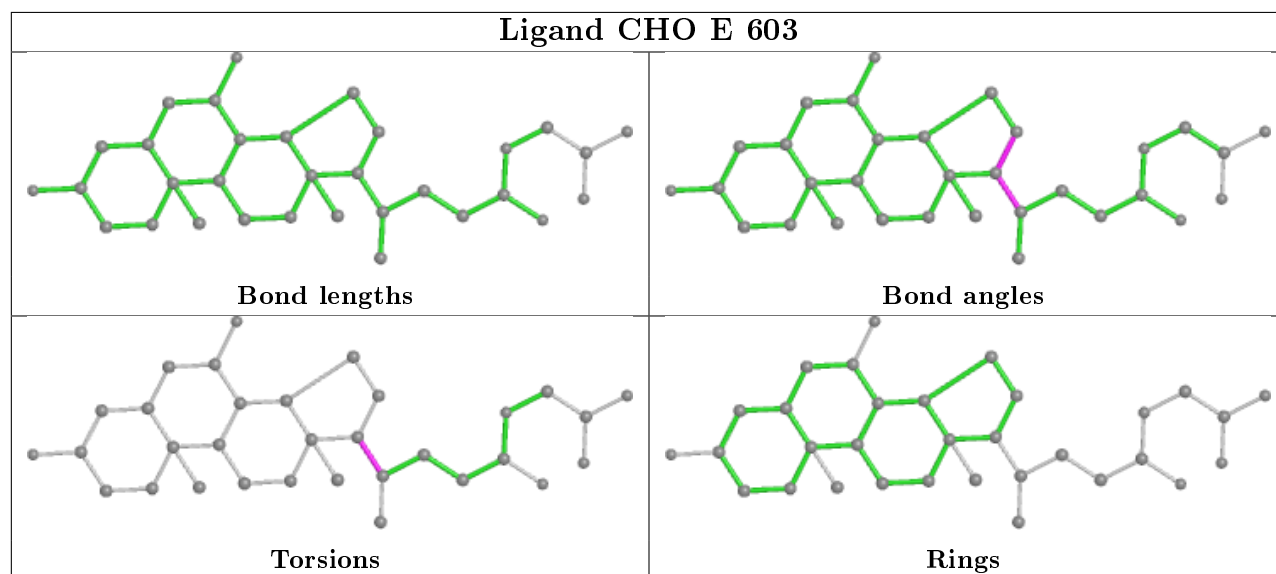
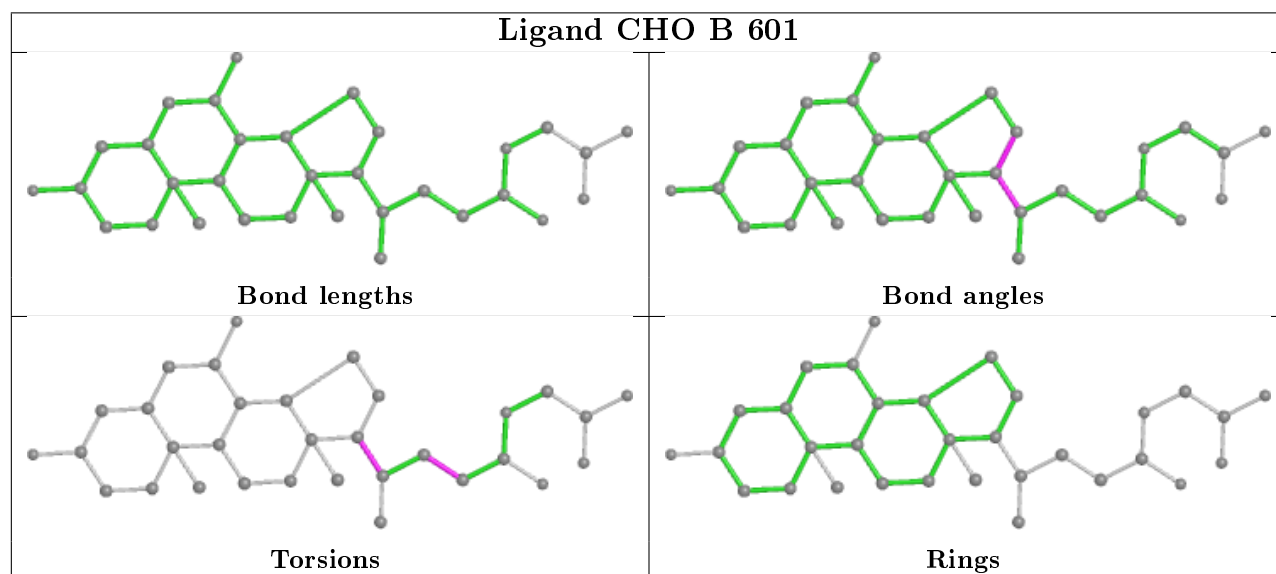
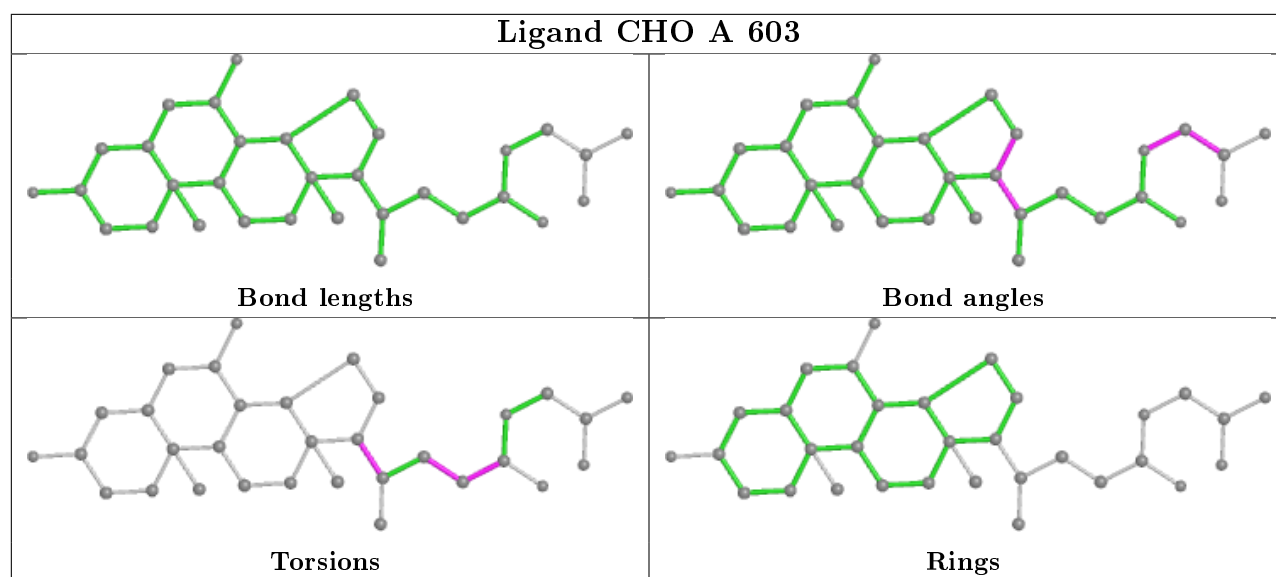
Mol	Chain	Res	Type	Atoms
4	F	900	CHO	C16-C17-C20-C22
4	A	603	CHO	C16-C17-C20-C22
4	B	601	CHO	C16-C17-C20-C22
4	A	603	CHO	C20-C22-C23-C24
3	C	700	EDO	O1-C1-C2-O2
4	F	900	CHO	C20-C22-C23-C24
4	B	601	CHO	C20-C22-C23-C24
4	A	603	CHO	C22-C23-C24-O24
4	A	603	CHO	C22-C23-C24-N25

There are no ring outliers.

No monomer is involved in short contacts.

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

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	303/309 (98%)	-0.12	7 (2%) 60 68	17, 25, 49, 75	0
1	B	305/309 (98%)	-0.13	6 (1%) 65 72	18, 26, 46, 78	0
1	E	303/309 (98%)	-0.22	9 (2%) 50 59	22, 32, 55, 108	0
1	F	301/309 (97%)	0.13	11 (3%) 41 49	26, 41, 64, 91	0
2	C	117/123 (95%)	-0.52	0 100 100	19, 25, 48, 66	0
2	D	117/123 (95%)	-0.36	1 (0%) 84 88	20, 27, 47, 70	0
2	G	117/123 (95%)	-0.21	4 (3%) 45 53	23, 34, 58, 79	0
2	H	114/123 (92%)	2.97	75 (65%) 0 0	43, 82, 123, 131	0
All	All	1677/1728 (97%)	0.06	113 (6%) 17 24	17, 31, 76, 131	0

All (113) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	91	THR	8.3
2	H	11	LEU	8.2
2	H	86	LEU	7.4
2	H	88	PRO	7.2
2	H	94	TYR	7.1
2	H	40	ALA	6.9
2	H	10	GLY	6.9
2	H	114	THR	6.8
2	H	111	THR	6.6
2	H	84	SER	6.4
1	B	363	THR	6.2
2	H	41	PRO	6.2
2	H	14	ALA	6.2
1	E	362	THR	6.1
2	H	115	VAL	5.8
2	H	82	GLN	5.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	H	12	VAL	5.8
2	H	8	GLY	5.7
2	H	20	LEU	5.7
2	H	68	PHE	5.7
2	H	17	SER	5.5
2	H	43	LYS	5.4
1	F	532	SER	5.4
2	H	19	ARG	5.3
2	H	113	VAL	5.3
2	H	89	GLU	5.2
2	H	42	GLY	5.1
2	H	15	GLY	5.1
2	H	16	GLY	5.1
2	H	93	VAL	5.1
2	H	92	ALA	5.0
1	E	346	THR	5.0
1	A	225	SER	4.8
2	H	13	GLU	4.8
2	H	18	LEU	4.7
1	E	363	THR	4.7
1	F	228	ARG	4.6
2	H	85	SER	4.5
2	H	83	MET	4.3
1	E	365	ALA	4.3
1	B	365	ALA	4.2
2	H	69	THR	4.1
1	A	346	THR	4.1
2	H	95	TYR	4.0
2	H	75	ALA	3.9
2	H	81	LEU	3.9
2	H	76	LYS	3.8
2	D	1	GLN	3.8
1	F	361	PRO	3.7
2	H	62	ASP	3.7
2	H	9	GLY	3.7
2	H	66	GLY	3.6
1	E	364	ASN	3.6
1	F	504	SER	3.6
1	B	364	ASN	3.5
1	B	362	THR	3.5
2	H	63	SER	3.5
2	H	23	LEU	3.5

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	H	61	ALA	3.4
2	H	45	ARG	3.3
2	H	36	TRP	3.3
2	H	67	ARG	3.3
1	F	505	GLY	3.3
2	H	7	SER	3.3
2	H	108	GLY	3.2
2	H	112	GLN	3.2
2	H	110	GLY	3.2
1	A	443	ASP	3.1
1	B	532	SER	3.1
2	H	65	LYS	3.1
1	F	524	LEU	3.1
2	H	109	GLN	3.1
2	H	80	TYR	3.0
2	G	41	PRO	3.0
2	H	90	ASP	3.0
1	F	426	GLY	2.9
1	E	347	GLY	2.9
2	G	44	GLU	2.9
1	F	531	GLY	2.9
2	H	27	LEU	2.8
2	H	44	GLU	2.8
2	H	5	GLN	2.7
2	H	87	LYS	2.7
2	H	4	LEU	2.7
1	F	502	SER	2.7
2	H	96	CYS	2.6
1	F	442	ALA	2.6
2	H	64	VAL	2.5
2	H	3	GLN	2.5
1	B	504	SER	2.5
1	F	425	PRO	2.4
2	H	39	GLN	2.4
2	H	22	CYS	2.4
2	H	38	ARG	2.4
2	G	1	GLN	2.4
2	H	6	GLU	2.3
1	E	343	THR	2.3
1	E	348	ASP	2.3
1	A	281	ILE	2.3
2	H	78	THR	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	H	26	GLY	2.3
2	H	74	ASN	2.2
1	A	286	VAL	2.2
2	H	2	VAL	2.2
2	H	30	SER	2.2
1	A	444	ALA	2.1
2	H	46	GLU	2.1
1	A	347	GLY	2.1
2	H	24	GLY	2.1
2	G	116	SER	2.1
2	H	70	ILE	2.1
1	E	361	PRO	2.0
2	H	77	ASN	2.0

## 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 ⓘ

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<sup>th</sup> 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.

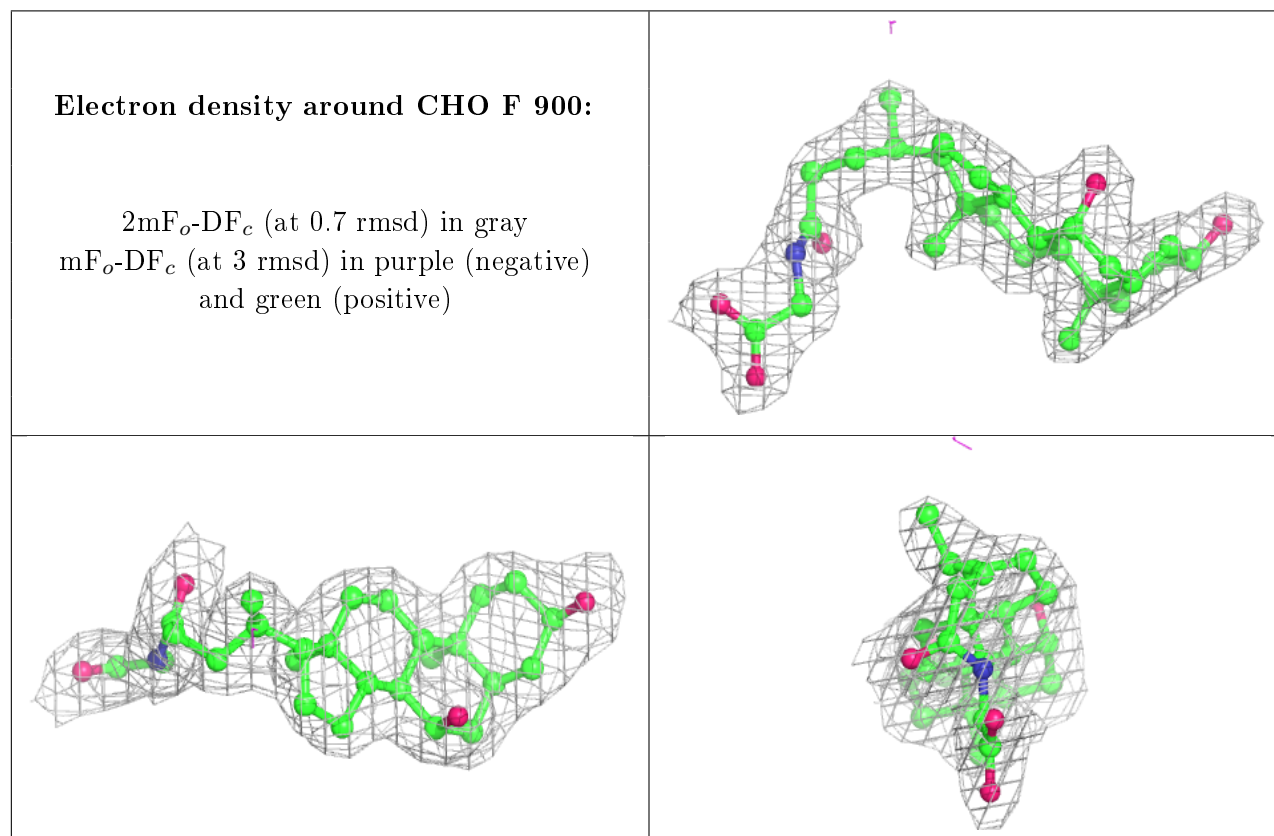
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	EDO	H	700	4/4	0.74	0.22	68,69,70,72	0
3	EDO	E	602	4/4	0.86	0.18	50,52,52,53	0
3	EDO	E	601	4/4	0.89	0.15	29,34,40,42	0
3	EDO	C	700	4/4	0.90	0.17	30,39,41,47	0
5	MG	B	602	1/1	0.91	0.07	41,41,41,41	0
3	EDO	G	700	4/4	0.93	0.13	32,36,36,41	0
3	EDO	A	601	4/4	0.94	0.11	26,28,29,31	0
4	CHO	F	900	32/32	0.95	0.09	27,36,37,40	0
3	EDO	A	602	4/4	0.95	0.14	38,39,40,41	0
3	EDO	D	700	4/4	0.95	0.14	28,30,31,34	0

*Continued on next page...*

*Continued from previous page...*

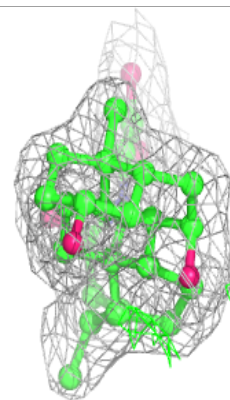
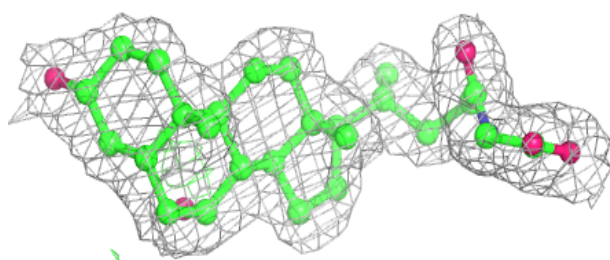
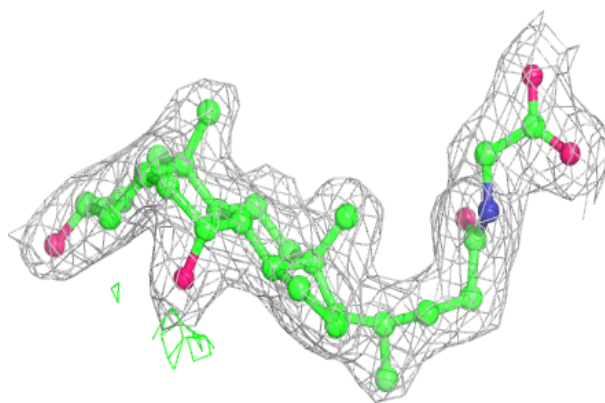
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	CHO	A	603	32/32	0.96	0.12	20,25,28,28	0
4	CHO	B	601	32/32	0.96	0.10	22,25,26,26	0
4	CHO	E	603	32/32	0.96	0.10	27,32,35,36	0

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.

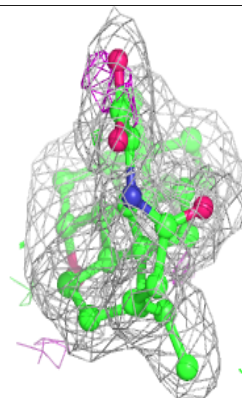
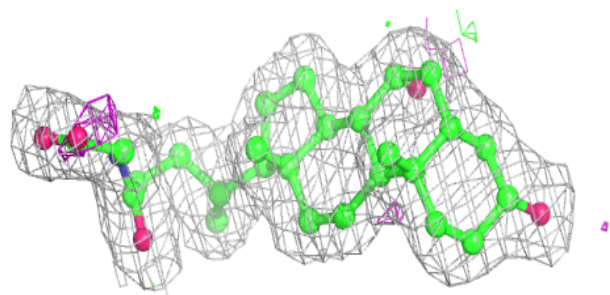
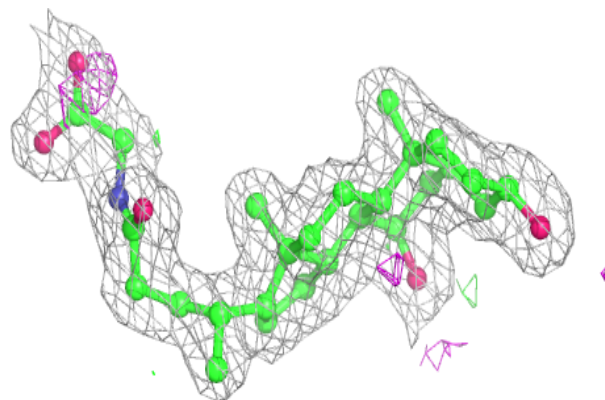


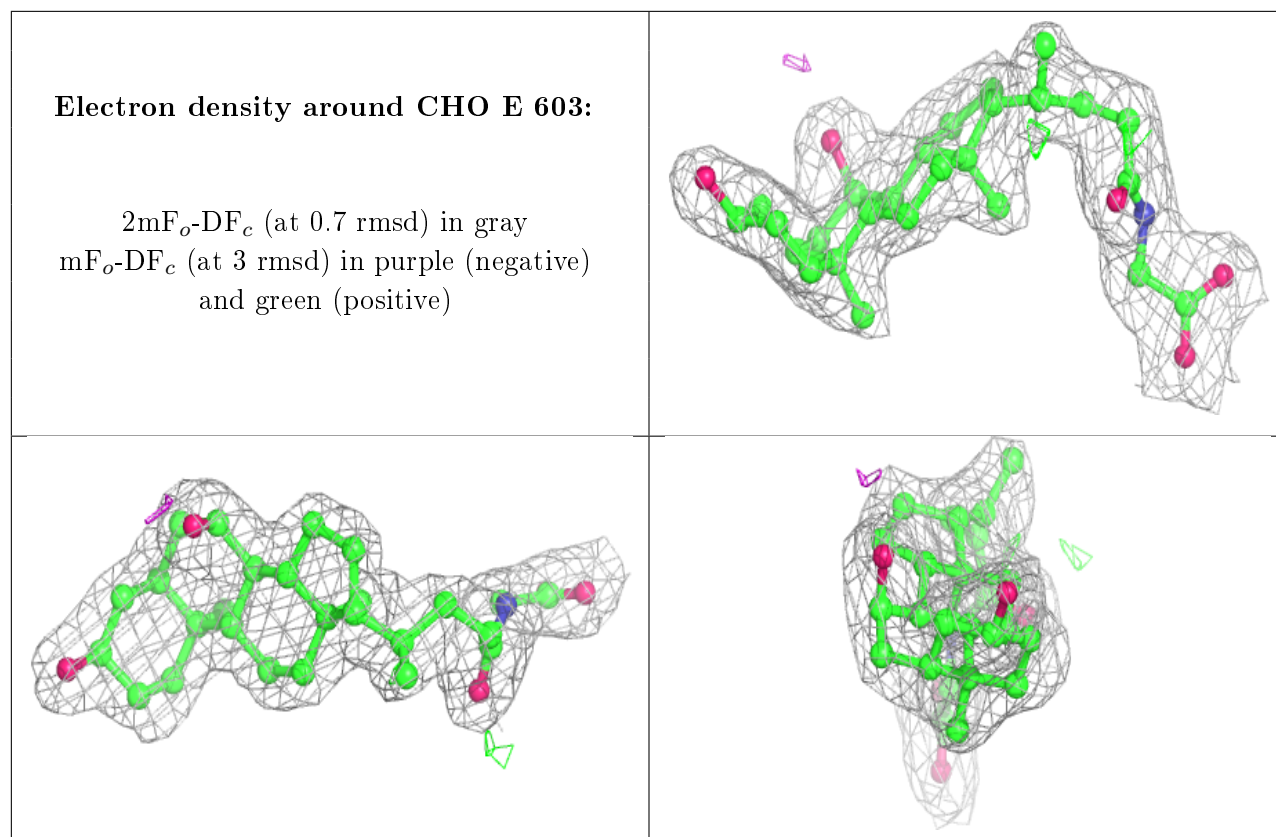
**Electron density around CHO A 603:**

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

**Electron density around CHO B 601:**

$2mF_o-DF_c$  (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.