



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 24, 2023 – 09:02 AM EDT

PDB ID : 3AI3  
Title : The crystal structure of L-Sorbose reductase from *Gluconobacter frateurii* complexed with NADPH and L-sorbose  
Authors : Kubota, K.; Nagata, K.; Okai, M.; Miyazono, K.; Tanokura, M.  
Deposited on : 2010-05-07  
Resolution : 1.80 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.36  
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) : 2.36

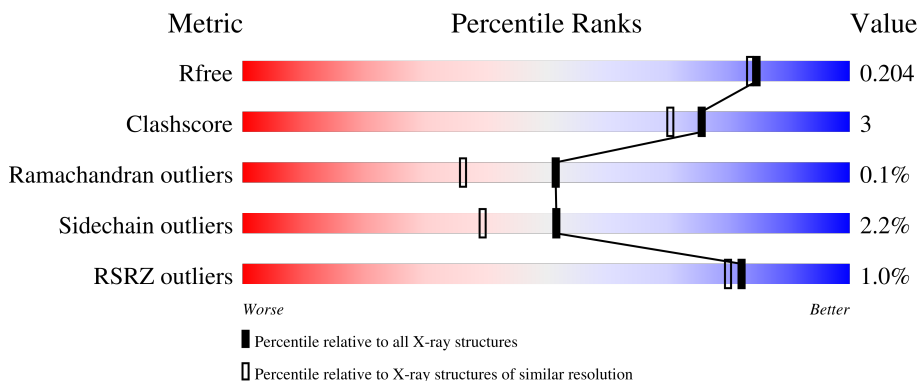
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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  $\geq 3$ , 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	263	 94% 5%
1	C	263	 94% 5%
1	E	263	 92% 7%
1	G	263	 93% 6%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 8624 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.

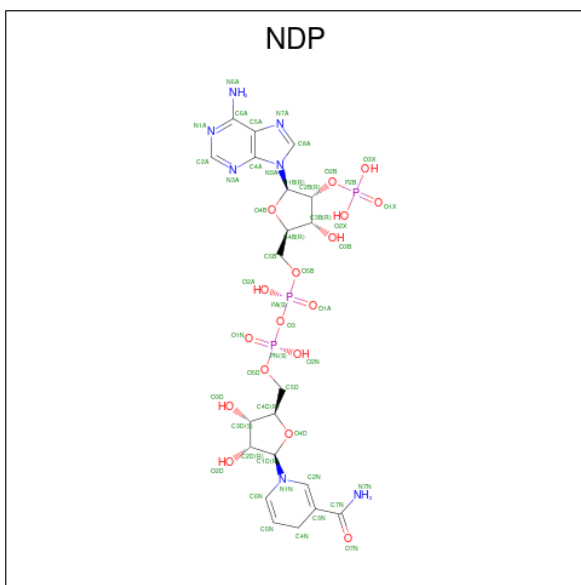
- Molecule 1 is a protein called NADPH-sorbose reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	263	1989	1266	340	372	11	0	0	0
1	C	263	1989	1266	340	372	11	0	0	0
1	E	263	1989	1266	340	372	11	0	0	0
1	G	263	1989	1266	340	372	11	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

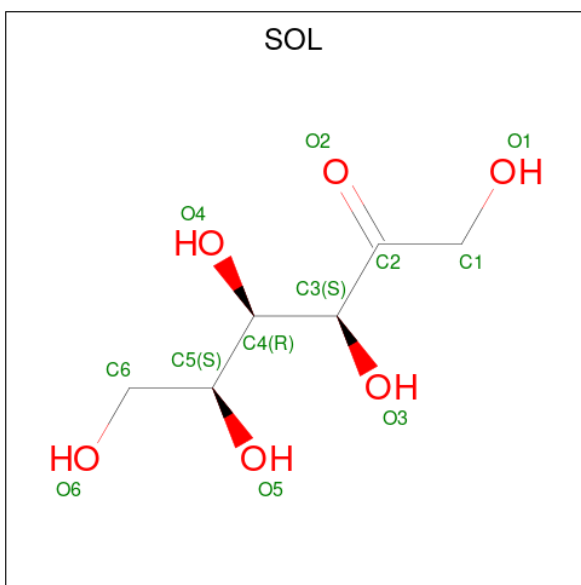
Chain	Residue	Modelled	Actual	Comment	Reference
A	116	LEU	HIS	engineered mutation	UNP A4PB64
C	116	LEU	HIS	engineered mutation	UNP A4PB64
E	116	LEU	HIS	engineered mutation	UNP A4PB64
G	116	LEU	HIS	engineered mutation	UNP A4PB64

- Molecule 2 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: C<sub>21</sub>H<sub>30</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>).



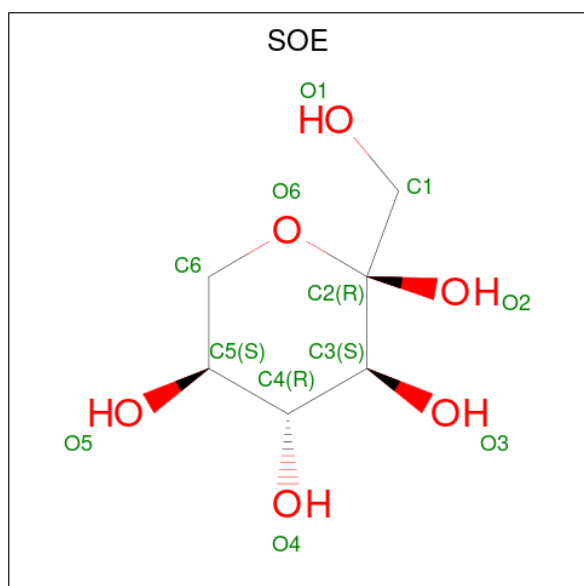
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	Total 48	C 21	N 7	O 17	P 3	0	0
2	C	1	Total 48	C 21	N 7	O 17	P 3	0	0
2	E	1	Total 48	C 21	N 7	O 17	P 3	0	0
2	G	1	Total 48	C 21	N 7	O 17	P 3	0	0

- Molecule 3 is L-sorbose (three-letter code: SOL) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 12 6 6	0	0
3	C	1	Total C O 12 6 6	0	0
3	E	1	Total C O 12 6 6	0	0
3	E	1	Total C O 12 6 6	0	0
3	G	1	Total C O 12 6 6	0	0

- Molecule 4 is alpha-L-sorbopyranose (three-letter code: SOE) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	E	1	Total C O 12 6 6	0	0
4	G	1	Total C O 12 6 6	0	0
4	G	1	Total C O 12 6 6	0	0
4	G	1	Total C O 12 6 6	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	89	Total O 89 89	0	0

*Continued on next page...*

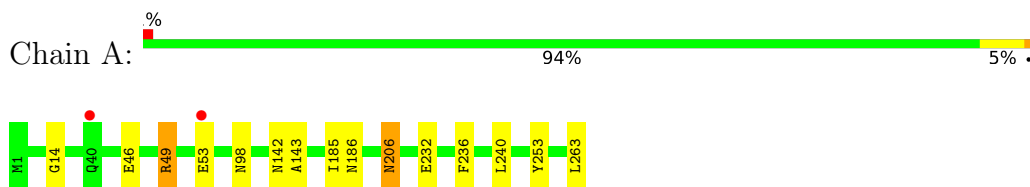
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
5	C	79	Total O 79 79	0	0
5	E	103	Total O 103 103	0	0
5	G	97	Total O 97 97	0	0

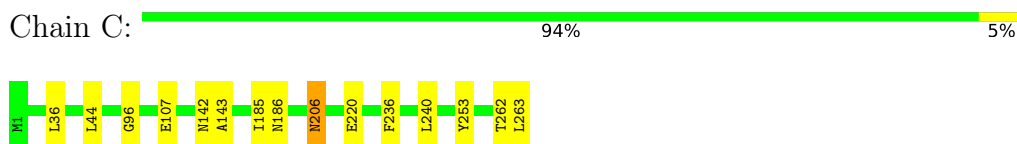
### 3 Residue-property plots [i](#)

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.

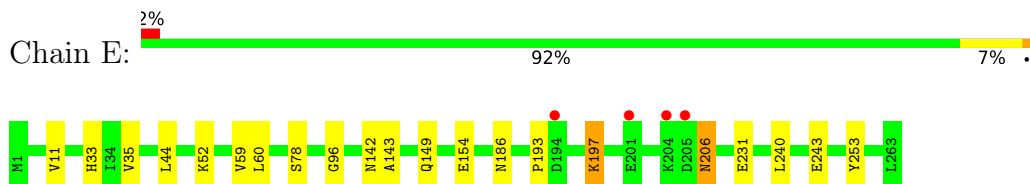
- Molecule 1: NADPH-sorbose reductase



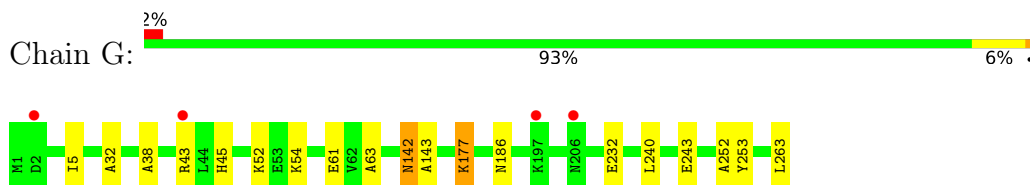
- Molecule 1: NADPH-sorbose reductase



- Molecule 1: NADPH-sorbose reductase



- Molecule 1: NADPH-sorbose reductase



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.04Å 124.41Å 124.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.69 – 1.80 19.69 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (19.69-1.80) 99.9 (19.69-1.80)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.46 (at 1.80Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.212 , 0.249 0.174 , 0.204	Depositor DCC
$R_{free}$ test set	4432 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.2	Xtrriage
Anisotropy	0.037	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.44 , 50.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.000 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8624	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 60.74 % 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 1.4551e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: NDP, SOE, SOL

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	0.66	0/2026	0.66	0/2743
1	C	0.63	0/2026	0.65	0/2743
1	E	0.61	0/2026	0.65	0/2743
1	G	0.63	0/2026	0.67	0/2743
All	All	0.63	0/8104	0.66	0/10972

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1989	0	2004	12	0
1	C	1989	0	2004	10	0
1	E	1989	0	2004	17	0
1	G	1989	0	2004	15	0
2	A	48	0	26	1	0
2	C	48	0	26	0	0
2	E	48	0	26	0	0
2	G	48	0	26	0	0
3	A	12	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	12	0	0	1	0
3	E	24	0	0	2	0
3	G	12	0	0	0	0
4	E	12	0	12	2	0
4	G	36	0	36	6	0
5	A	89	0	0	0	0
5	C	79	0	0	0	0
5	E	103	0	0	1	0
5	G	97	0	0	1	0
All	All	8624	0	8168	50	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 3.

All (50) 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:G:263:LEU:O	4:G:266:SOE:H1	1.11	1.25
1:G:263:LEU:O	4:G:266:SOE:C1	2.07	1.01
1:A:49:ARG:O	1:A:53:GLU:HG3	1.92	0.69
1:C:253:TYR:HD1	1:G:253:TYR:HD1	1.42	0.68
1:E:154:GLU:OE2	3:E:0:SOL:O3	2.11	0.67
1:A:49:ARG:HH11	1:A:49:ARG:HB3	1.60	0.67
1:E:243:GLU:HG2	4:E:264:SOE:O4	1.95	0.66
1:E:33:HIS:ND1	5:E:304:HOH:O	2.17	0.66
1:G:5:ILE:HG22	4:G:265:SOE:O5	1.99	0.62
1:G:45:HIS:HD2	1:G:61:GLU:OE1	1.88	0.56
1:A:253:TYR:HD1	1:E:253:TYR:HD1	1.54	0.55
1:G:240:LEU:HD11	1:G:253:TYR:CD2	2.42	0.55
1:A:46:GLU:OE2	1:A:49:ARG:NH1	2.40	0.54
1:A:206:ASN:C	1:A:206:ASN:HD22	2.10	0.53
1:A:46:GLU:CD	1:A:49:ARG:NH1	2.63	0.51
1:E:60:LEU:HD21	1:E:78:SER:OG	2.11	0.51
1:E:243:GLU:CG	4:E:264:SOE:O4	2.60	0.50
1:A:253:TYR:HD1	1:E:253:TYR:CD1	2.30	0.50
1:C:263:LEU:O	4:G:266:SOE:O5	2.27	0.49
1:E:206:ASN:C	1:E:206:ASN:HD22	2.15	0.49
1:E:240:LEU:HD11	1:E:253:TYR:HD2	1.78	0.48
1:C:206:ASN:C	1:C:206:ASN:HD22	2.17	0.48
1:E:240:LEU:HD11	1:E:253:TYR:CD2	2.48	0.48
1:G:143:ALA:HB3	1:G:186:ASN:OD1	2.13	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:263:LEU:HD23	1:C:263:LEU:HA	1.60	0.47
1:G:5:ILE:CD1	1:G:32:ALA:HB2	2.47	0.45
1:G:54:LYS:HB3	1:G:54:LYS:HE3	1.74	0.45
1:E:52:LYS:NZ	1:E:59:VAL:HB	2.33	0.44
1:C:96:GLY:CA	3:C:0:SOL:O4	2.65	0.44
1:A:263:LEU:HA	1:A:263:LEU:HD23	1.60	0.44
1:C:185:ILE:HD11	1:C:240:LEU:HD12	1.99	0.44
1:C:36:LEU:HB3	1:C:44:LEU:HD11	2.00	0.44
1:E:231:GLU:CD	1:E:231:GLU:H	2.21	0.44
1:G:252:ALA:O	4:G:266:SOE:O4	2.31	0.43
1:A:240:LEU:HD11	1:A:253:TYR:CD2	2.54	0.42
1:A:143:ALA:HB3	1:A:186:ASN:OD1	2.19	0.42
1:G:240:LEU:HD11	1:G:253:TYR:HD2	1.81	0.42
1:A:14:GLY:HA2	2:A:270:NDP:H1B	2.02	0.42
1:A:185:ILE:HD11	1:A:240:LEU:HD12	2.02	0.42
1:C:143:ALA:HB3	1:C:186:ASN:OD1	2.18	0.42
1:E:96:GLY:HA3	3:E:0:SOL:O4	2.19	0.42
1:E:193:PRO:O	1:E:197:LYS:HD3	2.19	0.42
1:G:243:GLU:HG2	4:G:265:SOE:O4	2.19	0.42
1:G:177:LYS:HB2	5:G:486:HOH:O	2.20	0.41
1:E:143:ALA:HB3	1:E:186:ASN:OD1	2.20	0.41
1:C:107:GLU:CD	1:C:107:GLU:H	2.24	0.40
1:C:262:THR:HG22	1:E:149:GLN:HB2	2.03	0.40
1:E:11:VAL:HA	1:E:35:VAL:HB	2.03	0.40
1:G:142:ASN:C	1:G:142:ASN:HD22	2.23	0.40
1:G:38:ALA:O	1:G:63:ALA:HA	2.21	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	261/263 (99%)	254 (97%)	6 (2%)	1 (0%)	34	21
1	C	261/263 (99%)	253 (97%)	8 (3%)	0	100	100
1	E	261/263 (99%)	252 (97%)	9 (3%)	0	100	100
1	G	261/263 (99%)	251 (96%)	10 (4%)	0	100	100
All	All	1044/1052 (99%)	1010 (97%)	33 (3%)	1 (0%)	51	36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	98	ASN

### 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	206/206 (100%)	201 (98%)	5 (2%)	49	36
1	C	206/206 (100%)	202 (98%)	4 (2%)	57	46
1	E	206/206 (100%)	202 (98%)	4 (2%)	57	46
1	G	206/206 (100%)	201 (98%)	5 (2%)	49	36
All	All	824/824 (100%)	806 (98%)	18 (2%)	52	39

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	49	ARG
1	A	142	ASN
1	A	206	ASN
1	A	232	GLU
1	A	236	PHE
1	C	142	ASN
1	C	206	ASN
1	C	220	GLU
1	C	236	PHE
1	E	44	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	142	ASN
1	E	197	LYS
1	E	206	ASN
1	G	43	ARG
1	G	52	LYS
1	G	142	ASN
1	G	177	LYS
1	G	232	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	92	ASN
1	A	142	ASN
1	A	206	ASN
1	C	92	ASN
1	C	142	ASN
1	C	206	ASN
1	E	40	GLN
1	E	92	ASN
1	E	142	ASN
1	E	206	ASN
1	E	215	GLN
1	G	45	HIS
1	G	92	ASN
1	G	142	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

13 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
4	SOE	G	264	-	12,12,12	0.89	1 (8%)	18,18,18	0.97	1 (5%)
3	SOL	C	0	-	10,11,11	0.78	0	9,14,14	1.60	1 (11%)
2	NDP	A	270	-	45,52,52	1.61	6 (13%)	53,80,80	1.45	5 (9%)
4	SOE	G	266	-	12,12,12	1.31	2 (16%)	18,18,18	3.36	12 (66%)
2	NDP	E	270	-	45,52,52	1.78	6 (13%)	53,80,80	1.37	5 (9%)
2	NDP	C	270	-	45,52,52	1.72	6 (13%)	53,80,80	1.40	6 (11%)
4	SOE	G	265	-	12,12,12	1.06	1 (8%)	18,18,18	1.38	1 (5%)
2	NDP	G	270	-	45,52,52	1.67	6 (13%)	53,80,80	1.33	6 (11%)
3	SOL	G	0	-	10,11,11	0.42	0	9,14,14	0.79	0
3	SOL	A	0	-	10,11,11	0.62	0	9,14,14	1.07	1 (11%)
3	SOL	E	265	-	10,11,11	0.56	0	9,14,14	1.25	1 (11%)
4	SOE	E	264	-	12,12,12	1.03	1 (8%)	18,18,18	1.16	2 (11%)
3	SOL	E	0	-	10,11,11	0.32	0	9,14,14	1.39	2 (22%)

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	SOE	G	264	-	-	0/3/23/23	0/1/1/1
3	SOL	C	0	-	-	10/16/16/16	-
2	NDP	A	270	-	-	3/30/77/77	0/5/5/5
4	SOE	G	266	-	-	1/3/23/23	0/1/1/1
2	NDP	E	270	-	-	7/30/77/77	0/5/5/5
2	NDP	C	270	-	-	6/30/77/77	0/5/5/5
4	SOE	G	265	-	-	0/3/23/23	0/1/1/1
2	NDP	G	270	-	-	5/30/77/77	0/5/5/5

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SOL	G	0	-	-	8/16/16/16	-
3	SOL	A	0	-	-	8/16/16/16	-
3	SOL	E	265	-	-	8/16/16/16	-
4	SOE	E	264	-	-	0/3/23/23	0/1/1/1
3	SOL	E	0	-	-	6/16/16/16	-

All (29) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	270	NDP	O7N-C7N	6.84	1.40	1.24
2	G	270	NDP	O7N-C7N	6.27	1.39	1.24
2	C	270	NDP	O7N-C7N	5.91	1.38	1.24
2	A	270	NDP	O7N-C7N	5.41	1.37	1.24
2	E	270	NDP	C4N-C3N	-4.85	1.40	1.49
2	G	270	NDP	C4N-C3N	-4.60	1.40	1.49
2	C	270	NDP	C4N-C3N	-4.55	1.41	1.49
2	A	270	NDP	C2A-N3A	4.26	1.39	1.32
2	C	270	NDP	C6N-C5N	4.15	1.40	1.33
2	A	270	NDP	C4N-C3N	-4.14	1.41	1.49
2	C	270	NDP	C2A-N3A	3.93	1.38	1.32
2	G	270	NDP	C6N-C5N	3.58	1.39	1.33
2	E	270	NDP	C6N-C5N	3.54	1.39	1.33
2	E	270	NDP	C2A-N3A	3.53	1.37	1.32
2	G	270	NDP	C2A-N3A	3.38	1.37	1.32
2	A	270	NDP	C6N-C5N	3.38	1.39	1.33
2	E	270	NDP	C4N-C5N	-3.38	1.40	1.48
2	A	270	NDP	C4N-C5N	-3.33	1.40	1.48
2	G	270	NDP	C4N-C5N	-3.24	1.40	1.48
2	C	270	NDP	C4N-C5N	-2.96	1.41	1.48
4	G	265	SOE	O6-C2	2.60	1.45	1.42
4	E	264	SOE	O6-C2	2.55	1.45	1.42
2	E	270	NDP	C2A-N1A	2.44	1.38	1.33
4	G	264	SOE	O2-C2	2.30	1.44	1.40
4	G	266	SOE	C4-C3	-2.23	1.49	1.53
2	A	270	NDP	C2A-N1A	2.20	1.38	1.33
4	G	266	SOE	O6-C2	2.17	1.44	1.42
2	C	270	NDP	C2A-N1A	2.04	1.37	1.33
2	G	270	NDP	C2A-N1A	2.00	1.37	1.33

All (43) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	270	NDP	N3A-C2A-N1A	-6.20	118.99	128.68
2	C	270	NDP	N3A-C2A-N1A	-6.00	119.31	128.68
2	G	270	NDP	N3A-C2A-N1A	-5.94	119.39	128.68
2	E	270	NDP	N3A-C2A-N1A	-5.94	119.40	128.68
4	G	266	SOE	C1-C2-C3	-5.65	100.76	111.94
4	G	266	SOE	C6-C5-C4	5.63	116.58	109.67
4	G	266	SOE	O6-C2-C1	5.12	111.92	105.58
4	G	266	SOE	O5-C5-C4	-4.48	101.17	110.14
4	G	266	SOE	O4-C4-C3	-4.03	101.53	109.73
4	G	266	SOE	O2-C2-C3	3.86	114.03	107.79
4	G	266	SOE	O1-C1-C2	-3.68	104.04	111.86
2	A	270	NDP	C3N-C2N-N1N	-3.51	118.08	123.10
3	C	0	SOL	O3-C3-C4	-3.47	103.09	110.45
4	G	266	SOE	C2-C3-C4	3.37	115.23	110.70
4	G	266	SOE	C5-C4-C3	3.15	114.01	110.48
2	C	270	NDP	C3N-C2N-N1N	-3.00	118.81	123.10
2	C	270	NDP	O7N-C7N-C3N	-2.99	115.27	120.90
2	G	270	NDP	C3N-C2N-N1N	-2.98	118.84	123.10
2	A	270	NDP	O7N-C7N-C3N	-2.88	115.48	120.90
4	E	264	SOE	O2-C2-C1	-2.85	106.69	111.12
2	C	270	NDP	C3N-C7N-N7N	2.83	122.69	117.67
4	G	266	SOE	C6-O6-C2	-2.82	110.90	114.09
4	G	265	SOE	O3-C3-C2	2.82	114.42	109.95
4	G	264	SOE	C6-O6-C2	2.78	117.24	114.09
2	A	270	NDP	O2N-PN-O1N	2.74	125.77	112.24
4	E	264	SOE	O6-C2-C1	2.73	108.97	105.58
4	G	266	SOE	O3-C3-C4	-2.69	104.06	110.03
3	E	0	SOL	O4-C4-C3	-2.65	104.38	109.21
3	A	0	SOL	O3-C3-C4	-2.60	104.94	110.45
2	G	270	NDP	O7N-C7N-C3N	-2.58	116.03	120.90
2	E	270	NDP	O2N-PN-O1N	2.50	124.58	112.24
2	E	270	NDP	C3N-C2N-N1N	-2.49	119.54	123.10
4	G	266	SOE	O6-C6-C5	2.47	114.81	111.11
2	G	270	NDP	O3X-P2B-O2X	2.39	116.76	107.64
2	G	270	NDP	C3N-C7N-N7N	2.28	121.72	117.67
2	C	270	NDP	O4D-C1D-C2D	-2.25	101.74	106.64
2	A	270	NDP	C3N-C7N-N7N	2.21	121.59	117.67
3	E	265	SOL	O3-C3-C4	-2.13	105.94	110.45
2	G	270	NDP	C1B-N9A-C4A	-2.12	122.92	126.64
2	E	270	NDP	O4D-C1D-C2D	-2.12	102.03	106.64
2	C	270	NDP	O2N-PN-O1N	2.10	122.62	112.24
3	E	0	SOL	O1-C1-C2	-2.01	107.25	112.66
2	E	270	NDP	O7N-C7N-C3N	-2.01	117.11	120.90



There are no chirality outliers.

All (62) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	270	NDP	C5D-O5D-PN-O1N
2	C	270	NDP	C5D-O5D-PN-O2N
2	E	270	NDP	C5D-O5D-PN-O1N
2	E	270	NDP	C5D-O5D-PN-O2N
2	G	270	NDP	C5D-O5D-PN-O2N
3	A	0	SOL	C2-C3-C4-O4
3	A	0	SOL	C2-C3-C4-C5
3	A	0	SOL	O3-C3-C4-O4
3	A	0	SOL	O3-C3-C4-C5
3	A	0	SOL	C3-C4-C5-O5
3	A	0	SOL	O4-C4-C5-O5
3	C	0	SOL	C2-C3-C4-O4
3	C	0	SOL	C2-C3-C4-C5
3	C	0	SOL	O3-C3-C4-O4
3	C	0	SOL	O3-C3-C4-C5
3	C	0	SOL	C3-C4-C5-O5
3	E	0	SOL	O2-C2-C3-O3
3	E	0	SOL	C4-C5-C6-O6
3	E	0	SOL	O5-C5-C6-O6
3	E	265	SOL	O1-C1-C2-O2
3	E	265	SOL	O1-C1-C2-C3
3	E	265	SOL	C1-C2-C3-O3
3	E	265	SOL	C1-C2-C3-C4
3	E	265	SOL	O2-C2-C3-O3
3	E	265	SOL	O2-C2-C3-C4
3	G	0	SOL	O1-C1-C2-O2
3	G	0	SOL	O1-C1-C2-C3
3	G	0	SOL	C2-C3-C4-O4
3	G	0	SOL	C2-C3-C4-C5
3	G	0	SOL	O3-C3-C4-O4
3	G	0	SOL	O3-C3-C4-C5
4	G	266	SOE	O1-C1-C2-C3
3	C	0	SOL	C4-C5-C6-O6
3	A	0	SOL	O4-C4-C5-C6
3	A	0	SOL	C3-C4-C5-C6
3	C	0	SOL	O4-C4-C5-O5
3	E	265	SOL	C4-C5-C6-O6
3	C	0	SOL	O5-C5-C6-O6
2	A	270	NDP	O4D-C1D-N1N-C6N
2	E	270	NDP	O4D-C1D-N1N-C6N

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	E	0	SOL	O1-C1-C2-O2
2	C	270	NDP	O4D-C1D-N1N-C6N
3	C	0	SOL	C3-C4-C5-C6
3	G	0	SOL	O5-C5-C6-O6
3	E	265	SOL	O5-C5-C6-O6
2	G	270	NDP	O4D-C1D-N1N-C6N
2	C	270	NDP	C2B-O2B-P2B-O2X
2	G	270	NDP	C5D-O5D-PN-O3
3	G	0	SOL	C3-C4-C5-C6
2	G	270	NDP	C5D-O5D-PN-O1N
3	C	0	SOL	O4-C4-C5-C6
3	E	0	SOL	O1-C1-C2-C3
2	A	270	NDP	O4B-C4B-C5B-O5B
3	E	0	SOL	C2-C3-C4-C5
2	A	270	NDP	C5D-O5D-PN-O3
2	C	270	NDP	C5D-O5D-PN-O3
2	E	270	NDP	C5D-O5D-PN-O3
2	C	270	NDP	O4B-C4B-C5B-O5B
2	G	270	NDP	O4B-C4B-C5B-O5B
2	E	270	NDP	PA-O3-PN-O1N
2	E	270	NDP	PA-O3-PN-O2N
2	E	270	NDP	O4B-C4B-C5B-O5B

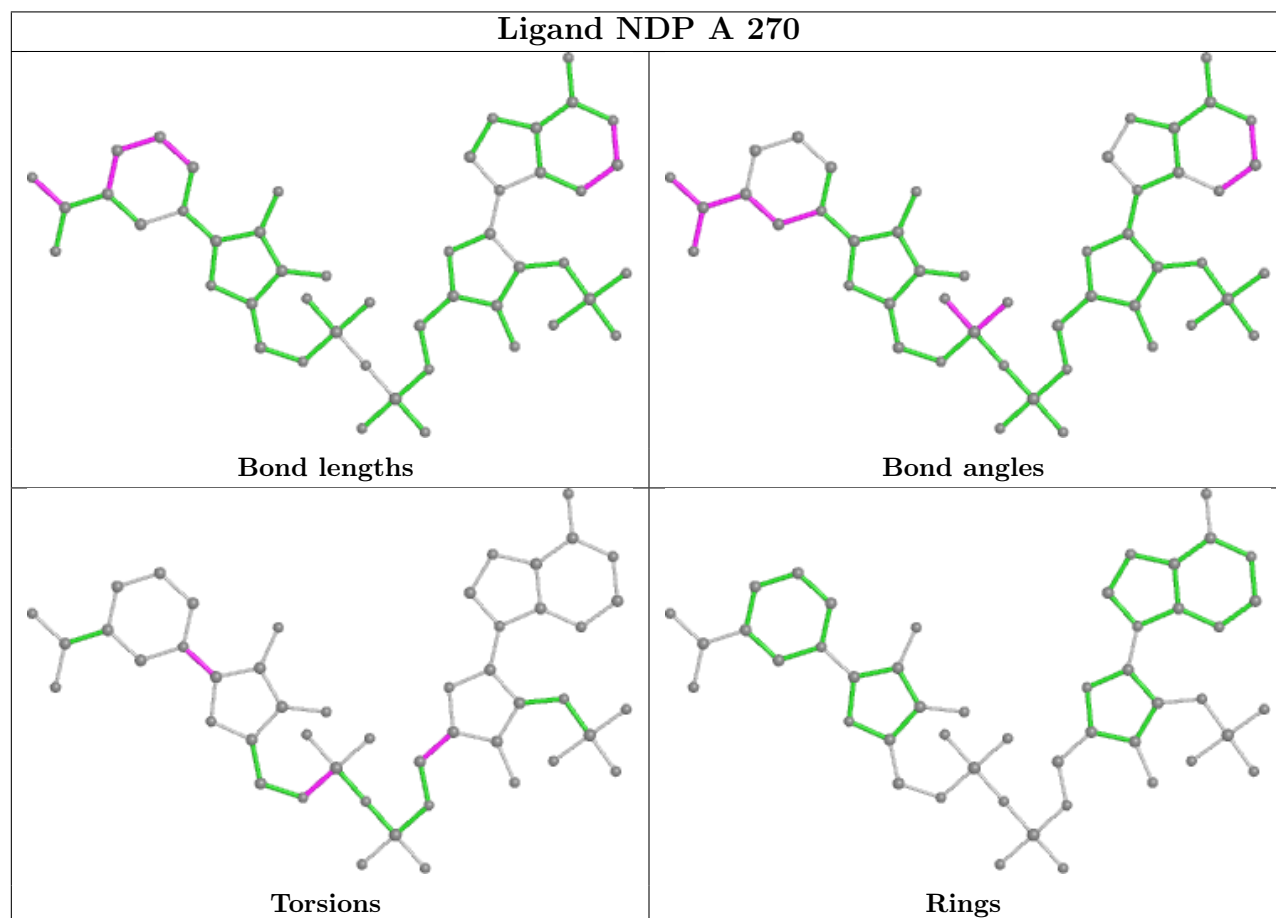
There are no ring outliers.

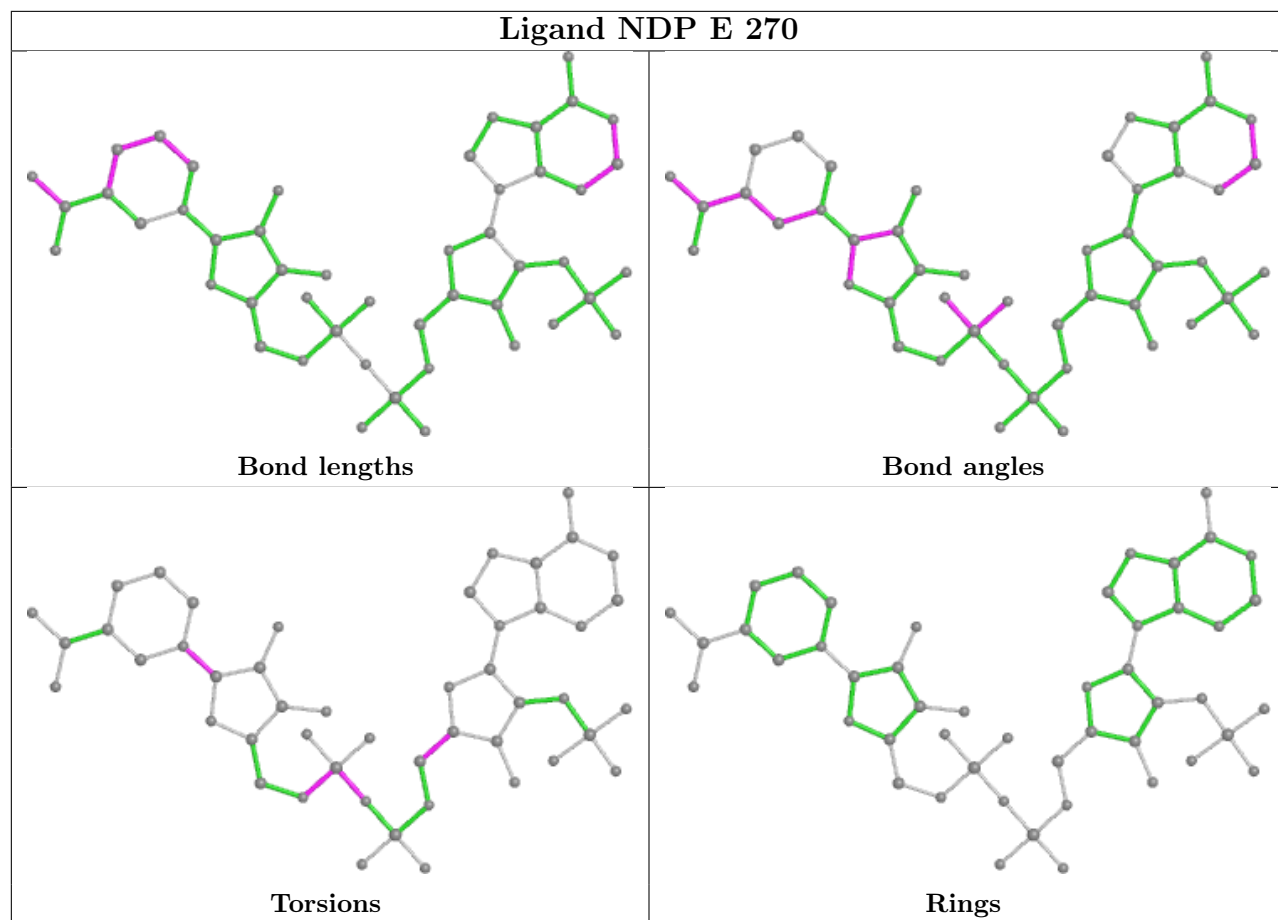
6 monomers are involved in 12 short contacts:

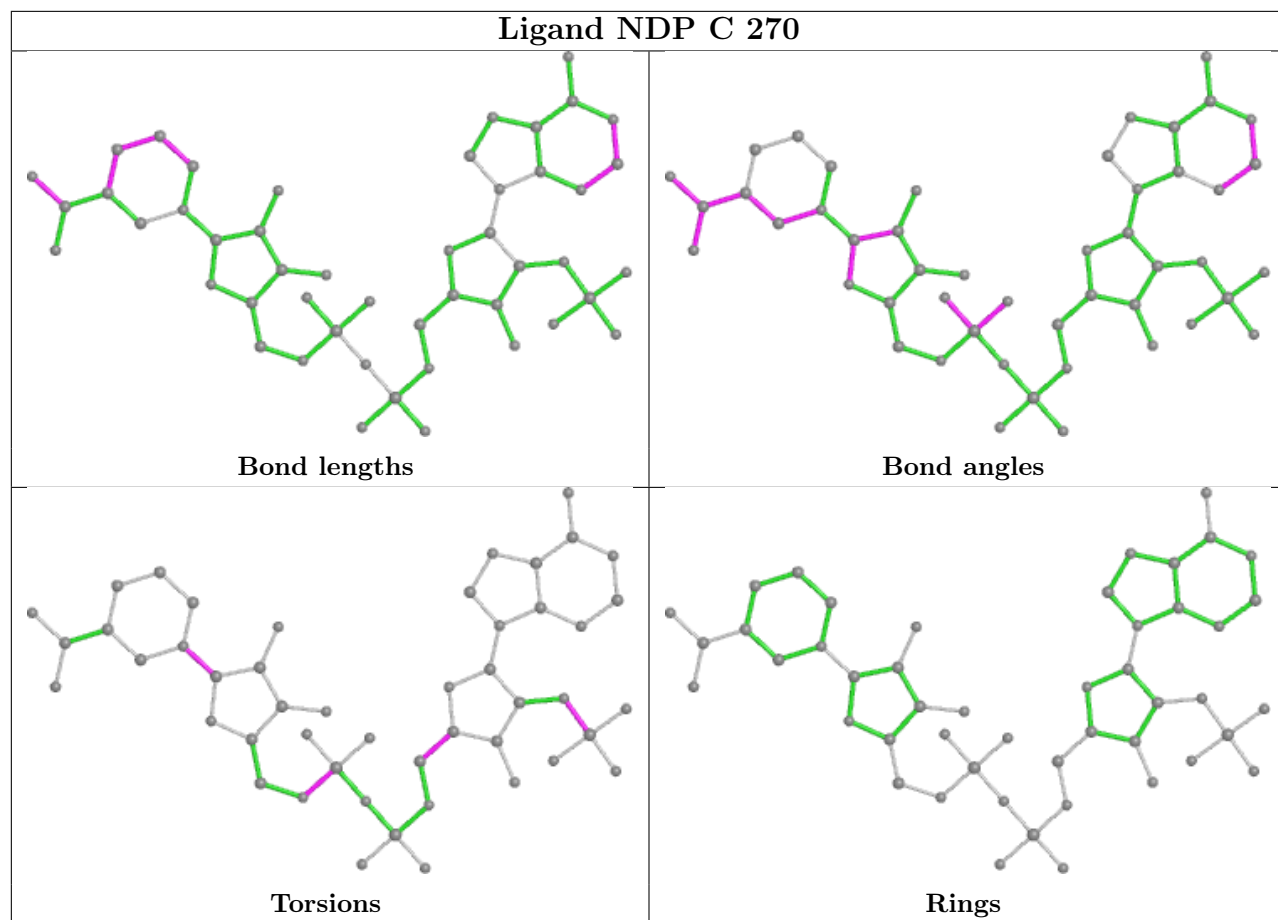
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	0	SOL	1	0
2	A	270	NDP	1	0
4	G	266	SOE	4	0
4	G	265	SOE	2	0
4	E	264	SOE	2	0
3	E	0	SOL	2	0

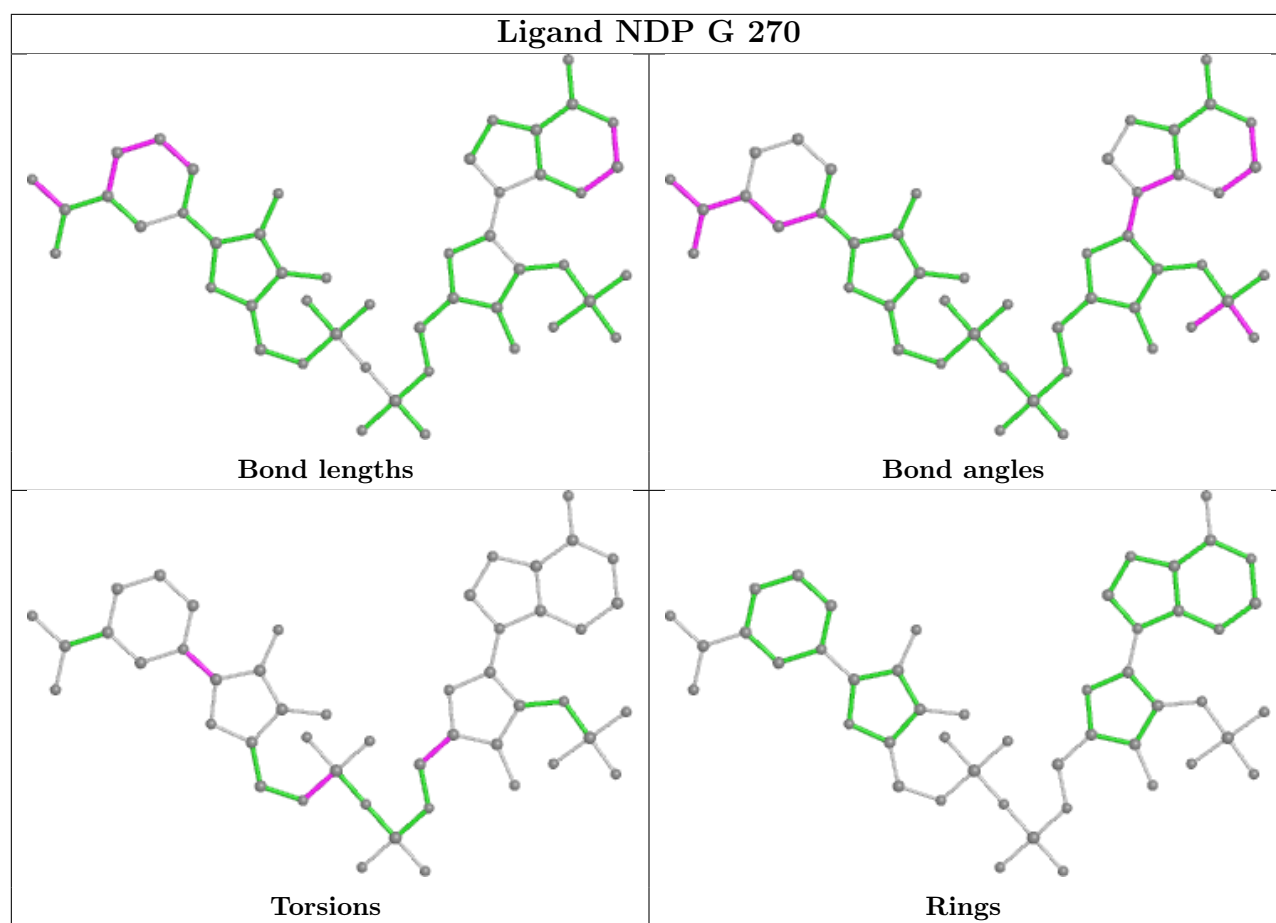
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<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	263/263 (100%)	-0.27	2 (0%) 86 84	5, 10, 20, 25	0
1	C	263/263 (100%)	-0.20	0 100 100	5, 11, 22, 31	0
1	E	263/263 (100%)	-0.18	4 (1%) 73 70	5, 11, 26, 33	0
1	G	263/263 (100%)	-0.22	4 (1%) 73 70	5, 10, 23, 32	0
All	All	1052/1052 (100%)	-0.22	10 (0%) 82 80	5, 11, 22, 33	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	2	ASP	3.0
1	E	204	LYS	2.9
1	E	205	ASP	2.8
1	E	201	GLU	2.8
1	E	194	ASP	2.3
1	G	197	LYS	2.2
1	G	206	ASN	2.1
1	G	43	ARG	2.0
1	A	53	GLU	2.0
1	A	40	GLN	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

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.

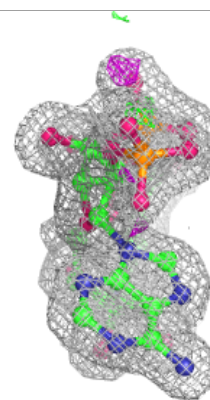
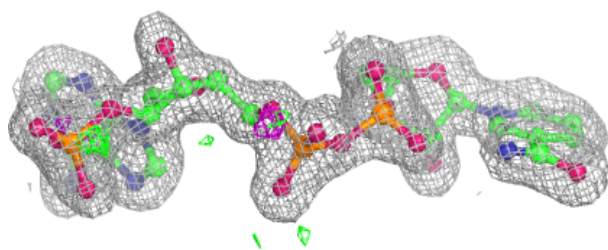
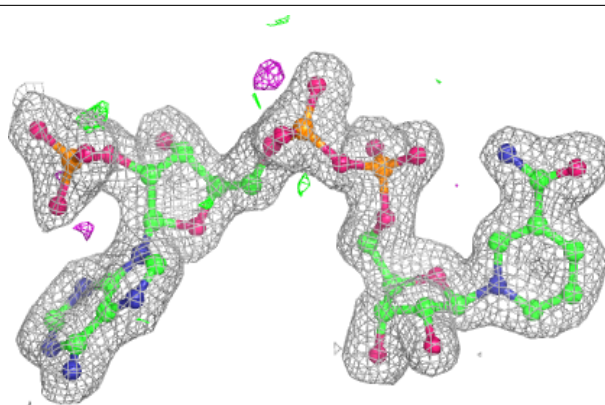
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	SOL	A	0	12/12	0.64	0.33	20,36,38,38	0
3	SOL	E	265	12/12	0.66	0.20	25,31,34,34	0
3	SOL	G	0	12/12	0.67	0.29	28,40,42,43	0
4	SOE	G	266	12/12	0.71	0.29	30,32,34,35	0
3	SOL	C	0	12/12	0.72	0.27	28,36,38,39	0
3	SOL	E	0	12/12	0.73	0.30	26,36,38,38	0
4	SOE	E	264	12/12	0.90	0.15	19,20,22,23	0
4	SOE	G	265	12/12	0.91	0.14	20,22,24,25	0
4	SOE	G	264	12/12	0.94	0.13	17,19,20,22	0
2	NDP	A	270	48/48	0.96	0.08	6,12,17,18	0
2	NDP	C	270	48/48	0.97	0.07	6,11,14,15	0
2	NDP	E	270	48/48	0.97	0.08	8,11,14,15	0
2	NDP	G	270	48/48	0.97	0.07	6,12,15,16	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.

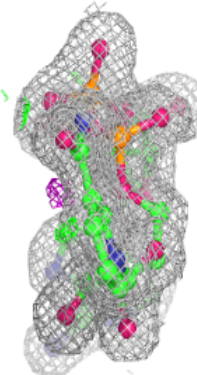
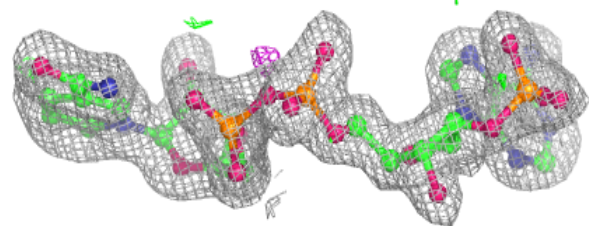
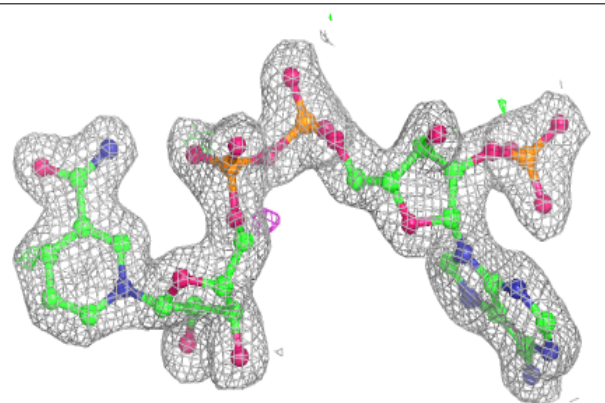


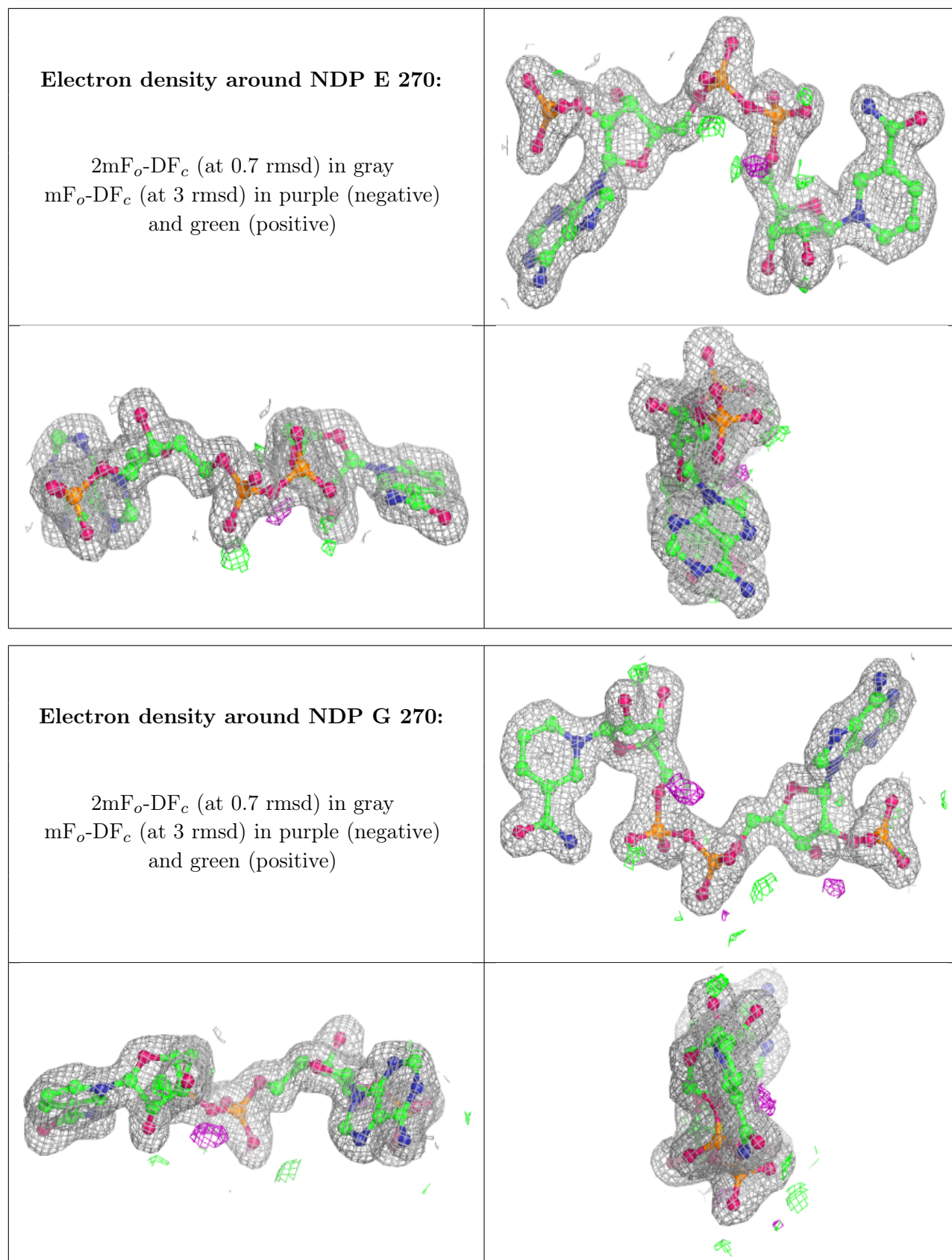
**Electron density around NDP A 270:**

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

**Electron density around NDP C 270:**

$2mF_o-DF_c$  (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.