wwPDB X-ray Structure Validation Summary Report (i)

Jan 3, 2024 – 09:32 pm GMT

PDB ID	:	5FJR
Title	:	N-acyl amino acid racemase from Amycolatopsis sp. Ts-1-60: Q26A M50I G291D F323Y mutant in complex with N-acetyl napthylalanine
Authors	:	Sanchez Carron, G.; Campopiano, D.; Grogan, G.
Deposited on	:	2015-10-12
Resolution	:	2.44 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

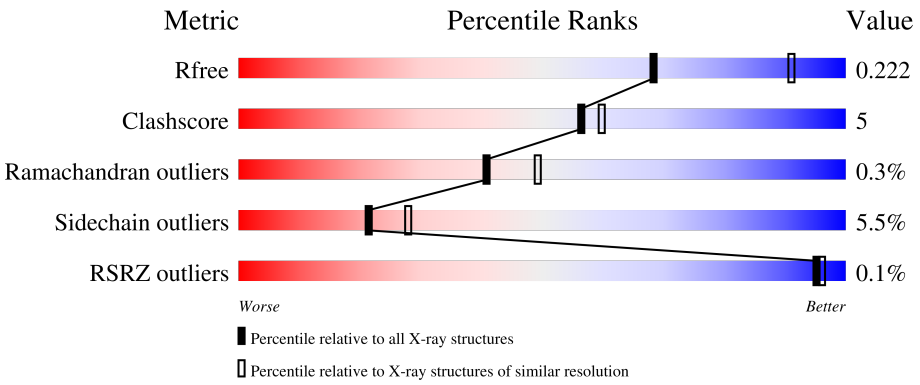
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.44 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	1564 (2.46-2.42)
Clashscore	141614	1631 (2.46-2.42)
Ramachandran outliers	138981	1617 (2.46-2.42)
Sidechain outliers	138945	1617 (2.46-2.42)
RSRZ outliers	127900	1547 (2.46-2.42)

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

Mol	Chain	Length	Quality of chain
1	A	368	87% 11% •
1	B	368	81% 16% •
1	C	368	88% 11% •
1	D	368	79% 18% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 11609 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.

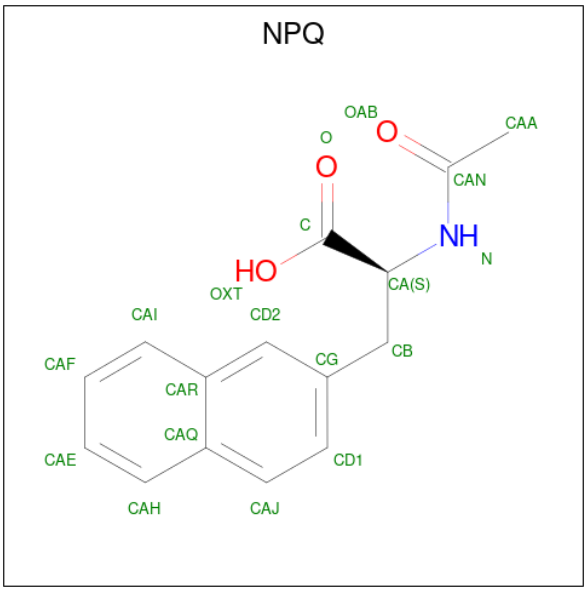
- Molecule 1 is a protein called O-SUCCINYLBENZOATE SYNTHASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	368	2764	1758	480	515	11	0	0	0
1	B	367	2726	1742	473	501	10	0	0	0
1	C	367	2720	1735	475	500	10	0	0	0
1	D	367	2702	1726	463	503	10	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	ALA	GLN	engineered mutation	UNP Q44244
A	50	ILE	MET	engineered mutation	UNP Q44244
A	291	ASP	GLY	engineered mutation	UNP Q44244
A	323	TYR	PHE	engineered mutation	UNP Q44244
B	26	ALA	GLN	engineered mutation	UNP Q44244
B	50	ILE	MET	engineered mutation	UNP Q44244
B	291	ASP	GLY	engineered mutation	UNP Q44244
B	323	TYR	PHE	engineered mutation	UNP Q44244
C	26	ALA	GLN	engineered mutation	UNP Q44244
C	50	ILE	MET	engineered mutation	UNP Q44244
C	291	ASP	GLY	engineered mutation	UNP Q44244
C	323	TYR	PHE	engineered mutation	UNP Q44244
D	26	ALA	GLN	engineered mutation	UNP Q44244
D	50	ILE	MET	engineered mutation	UNP Q44244
D	291	ASP	GLY	engineered mutation	UNP Q44244
D	323	TYR	PHE	engineered mutation	UNP Q44244

- Molecule 2 is N-acetyl naphthylalanine (three-letter code: NPQ) (formula: $C_{15}H_{15}NO_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	19	15	1	3	0	0
2	B	1	19	15	1	3	0	0
2	C	1	19	15	1	3	0	0
2	D	1	19	15	1	3	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
3	A	1	1	1	0	0
3	B	1	1	1	0	0
3	C	1	1	1	0	0
3	D	1	1	1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	246	246	246	0	0

Continued on next page...

Continued from previous page...

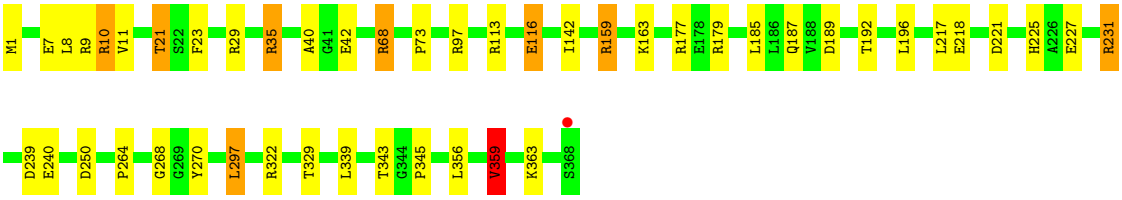
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	B	145	145	145	0	0
			Total	O		
4	C	118	118	118	0	0
			Total	O		
4	D	108	108	108	0	0

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: O-SUCCINYLBENZOATE SYNTHASE

Chain A: 87% 11% •

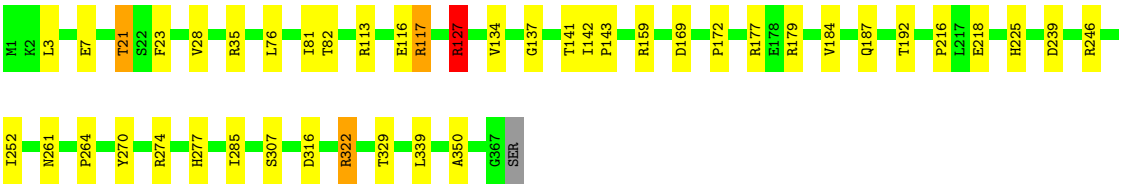


• Molecule 1: O-SUCCINYLBENZOATE SYNTHASE

Chain B: 81% 16% •

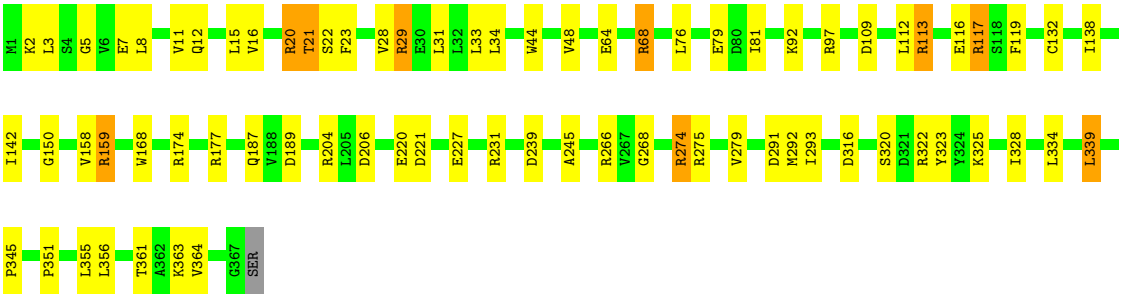
• Molecule 1: O-SUCCINYLBENZOATE SYNTHASE

Chain C: 88% 11% •



• Molecule 1: O-SUCCINYLBENZOATE SYNTHASE

Chain D: 79% 18% •


M1
K2
L3
S4
G5
V6
E7
L8
V11
Q12
L15
V16
R20
T21
S22
F23
V28
R29
E30
L31
L32
L33
L34
W44
V48
E64
R68
L76
E79
D80
I81
K92
R97
D109
L112
R113
E116
R117
S118
F119
C132
I138
I142
G150
V158
R159
W168
R174
R177
Q187
V188
D189
R204
L205
D206
E220
D221
E227
R231
D239
A245
R266
V267
G268
R274
R275
V279
D291
M292
I293
D316
S320
D321
R322
Y323
Y324
K325
I328
L334
L339
P345
P351
L355
L356
T361
A362
K363
V364
G367
SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	217.67Å 217.67Å 263.15Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	107.89 – 2.44 107.89 – 2.44	Depositor EDS
% Data completeness (in resolution range)	100.0 (107.89-2.44) 100.0 (107.89-2.44)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.26 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, R_{free}	0.172 , 0.221 0.173 , 0.222	Depositor DCC
R_{free} test set	4220 reflections (4.76%)	wwPDB-VP
Wilson B-factor (Å^2)	47.7	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 34.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.022 for -2/3*h-1/3*k+2/3*l,-1/3*h-2/3*k-2/3*l,2/3*h-2/3*k+1/3*l 0.016 for -h,1/3*h-1/3*k+2/3*l,2/3*h+4/3*k+1/3*l 0.012 for -1/3*h+1/3*k-2/3*l,-k,-4/3*h-2/3*k+1/3*l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	11609	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.18% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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

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	1.02	2/2820 (0.1%)	1.18	22/3844 (0.6%)
1	B	1.00	1/2782 (0.0%)	1.11	13/3796 (0.3%)
1	C	0.93	0/2776	1.07	7/3791 (0.2%)
1	D	1.01	3/2758 (0.1%)	1.18	21/3769 (0.6%)
All	All	0.99	6/11136 (0.1%)	1.14	63/15200 (0.4%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	220	GLU	CD-OE1	5.68	1.31	1.25
1	A	240	GLU	CD-OE2	5.36	1.31	1.25
1	D	239	ASP	CB-CG	5.31	1.62	1.51
1	D	361	THR	CB-CG2	5.23	1.69	1.52
1	A	250	ASP	CB-CG	5.21	1.62	1.51

The worst 5 of 63 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	113	ARG	NE-CZ-NH2	-17.84	111.38	120.30
1	A	113	ARG	NE-CZ-NH1	13.04	126.82	120.30
1	D	356	LEU	CB-CG-CD2	11.66	130.82	111.00
1	C	127	ARG	NE-CZ-NH1	10.79	125.70	120.30
1	D	159	ARG	NE-CZ-NH1	10.14	125.37	120.30

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	2764	0	2802	20	0
1	B	2726	0	2758	32	0
1	C	2720	0	2737	24	0
1	D	2702	0	2699	34	0
2	A	19	0	0	0	0
2	B	19	0	0	0	0
2	C	19	0	0	0	0
2	D	19	0	0	1	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	246	0	0	4	0
4	B	145	0	0	0	0
4	C	118	0	0	2	0
4	D	108	0	0	6	0
All	All	11609	0	10996	109	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 109 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:B:48:VAL:HB	1:B:292:MET:HG2	1.68	0.74
1:D:21:THR:HG23	1:D:23:PHE:H	1.55	0.72
1:C:127:ARG:HH11	1:C:127:ARG:HG3	1.53	0.71
1:A:11:VAL:HG13	1:A:359:VAL:HG13	1.74	0.70
1:B:10:ARG:HD3	1:B:64:GLU:OE1	1.94	0.68

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	366/368 (100%)	356 (97%)	10 (3%)	0	100	100
1	B	365/368 (99%)	347 (95%)	17 (5%)	1 (0%)	41	49
1	C	365/368 (99%)	346 (95%)	19 (5%)	0	100	100
1	D	365/368 (99%)	340 (93%)	22 (6%)	3 (1%)	19	22
All	All	1461/1472 (99%)	1389 (95%)	68 (5%)	4 (0%)	41	49

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	40	ALA
1	D	16	VAL
1	D	79	GLU
1	D	150	GLY

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	289/291 (99%)	273 (94%)	16 (6%)	21	28
1	B	279/291 (96%)	261 (94%)	18 (6%)	17	21
1	C	277/291 (95%)	265 (96%)	12 (4%)	29	38
1	D	274/291 (94%)	259 (94%)	15 (6%)	21	28
All	All	1119/1164 (96%)	1058 (94%)	61 (6%)	21	28

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	227	GLU
1	D	177	ARG
1	C	28	VAL
1	D	142	ILE
1	D	322	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	12	GLN
1	D	338	HIS

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 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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	$\#\|Z\| > 2$
2	NPQ	B	1368	3	20,20,20	1.87	5 (25%)	27,27,27	2.21	9 (33%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	NPQ	C	1368	3	20,20,20	1.70	4 (20%)	27,27,27	1.96	4 (14%)
2	NPQ	A	1369	3	20,20,20	1.63	4 (20%)	27,27,27	2.34	8 (29%)
2	NPQ	D	1368	3	20,20,20	2.07	3 (15%)	27,27,27	2.12	9 (33%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NPQ	B	1368	3	-	6/12/12/12	0/2/2/2
2	NPQ	C	1368	3	-	8/12/12/12	0/2/2/2
2	NPQ	A	1369	3	-	8/12/12/12	0/2/2/2
2	NPQ	D	1368	3	-	6/12/12/12	0/2/2/2

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1368	NPQ	CB-CG	-6.73	1.35	1.51
2	C	1368	NPQ	CB-CG	-4.69	1.40	1.51
2	B	1368	NPQ	CA-C	4.52	1.64	1.52
2	D	1368	NPQ	CA-C	4.40	1.64	1.52
2	B	1368	NPQ	CB-CG	-4.27	1.41	1.51

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1368	NPQ	CA-N-CAN	5.94	139.13	121.48
2	C	1368	NPQ	C-CA-N	-5.79	96.84	110.55
2	A	1369	NPQ	CA-N-CAN	5.18	136.88	121.48
2	C	1368	NPQ	CG-CB-CA	5.16	127.61	113.39
2	A	1369	NPQ	CB-CA-N	4.92	121.16	110.79

There are no chirality outliers.

5 of 28 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1369	NPQ	CAA-CAN-N-CA
2	A	1369	NPQ	OAB-CAN-N-CA
2	B	1368	NPQ	CAA-CAN-N-CA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	B	1368	NPQ	OAB-CAN-N-CA
2	C	1368	NPQ	CAA-CAN-N-CA

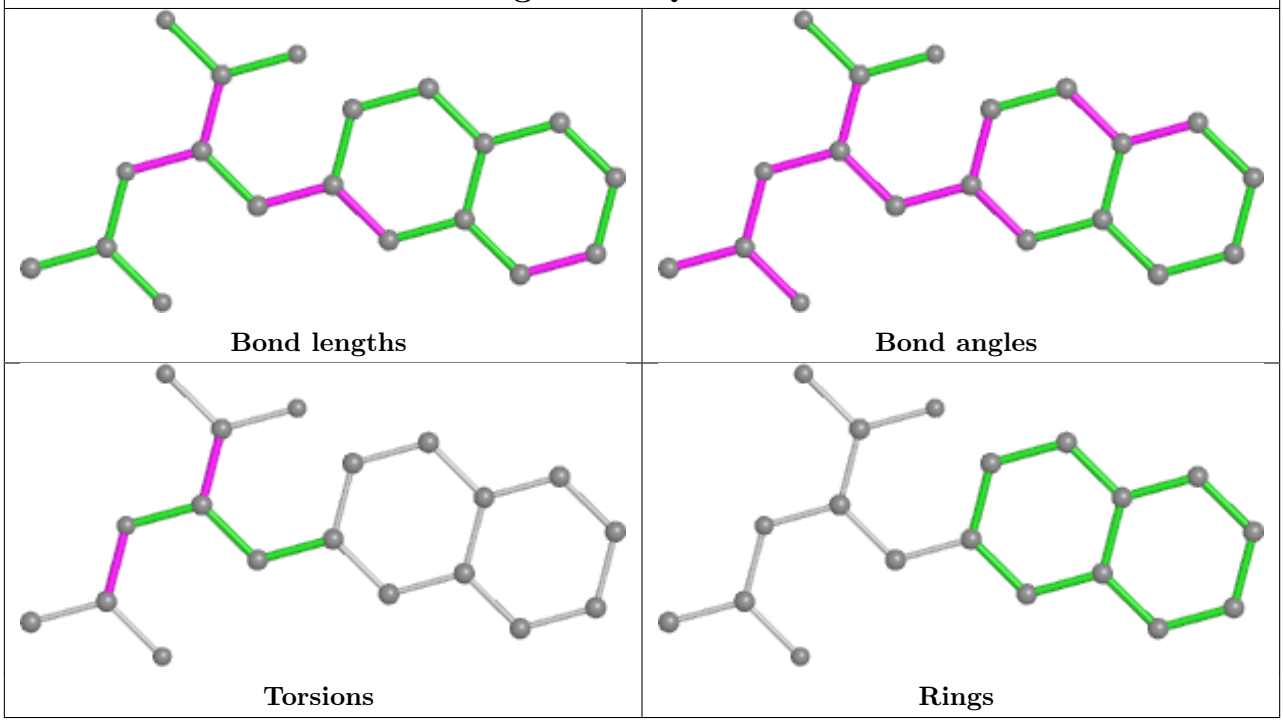
There are no ring outliers.

1 monomer is involved in 1 short contact:

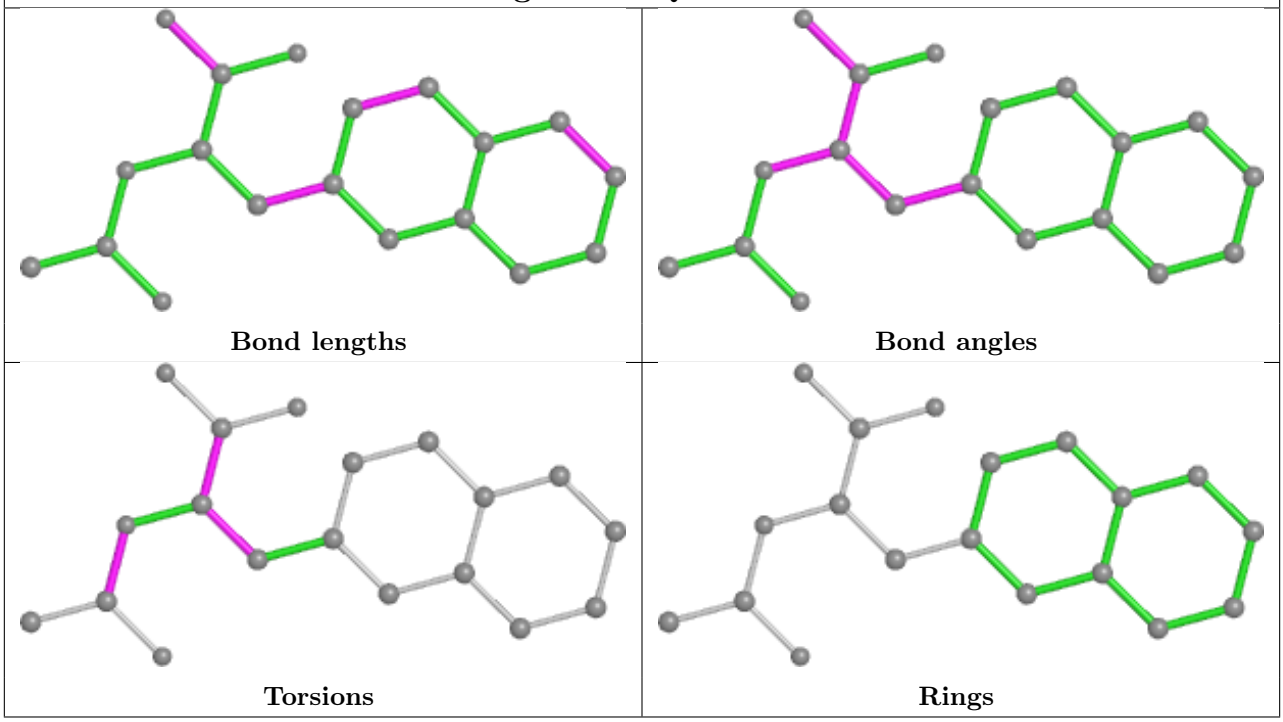
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1368	NPQ	1	0

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.

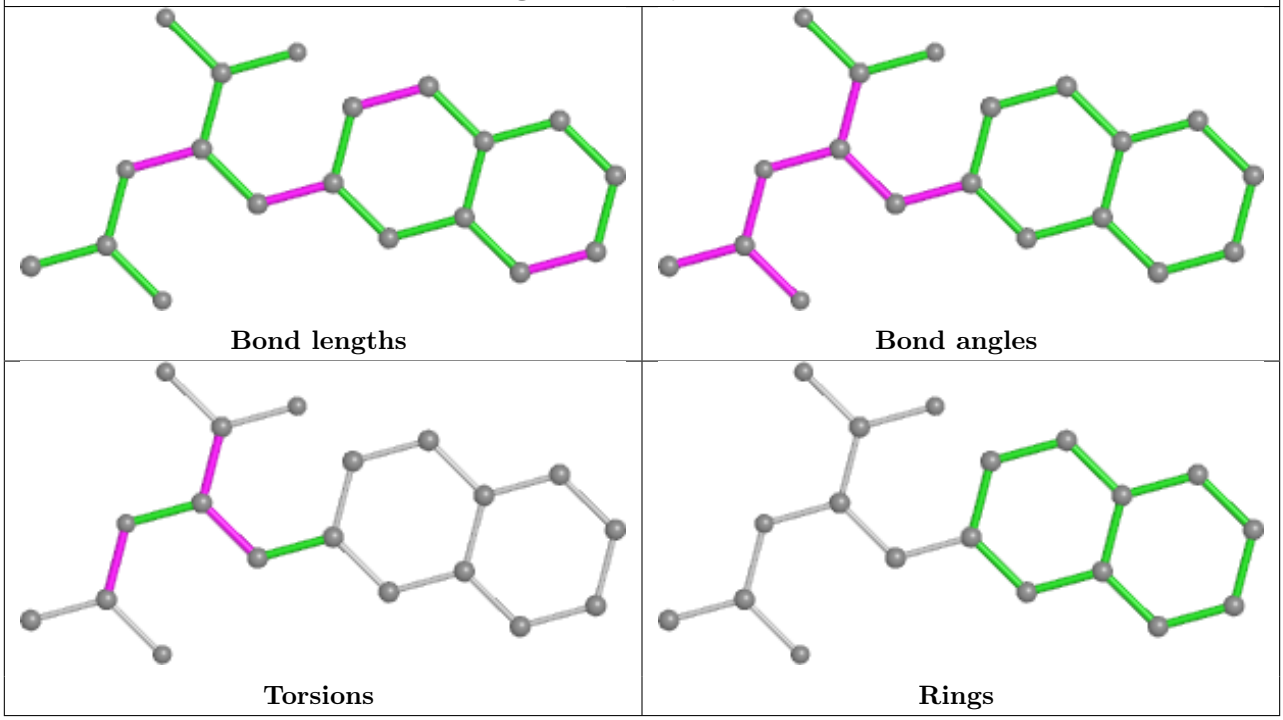
Ligand NPQ B 1368

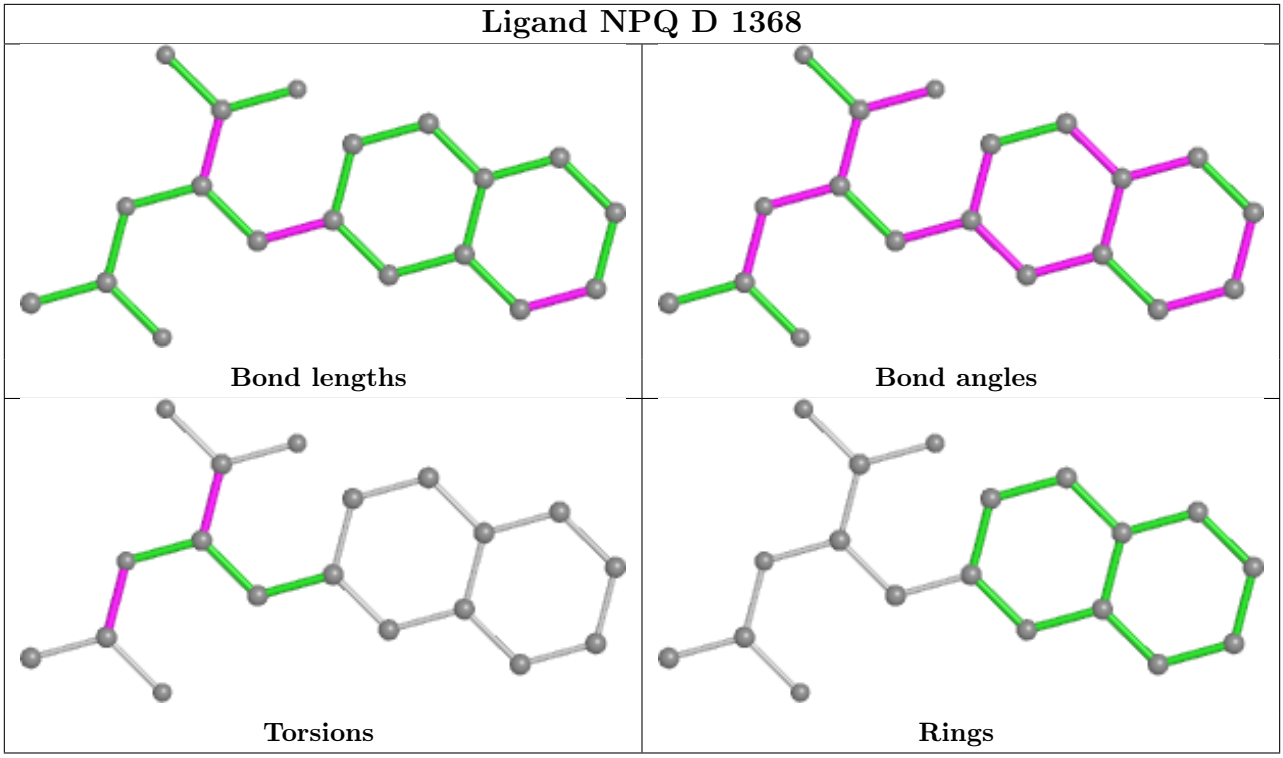


Ligand NPQ C 1368



Ligand NPQ A 1369





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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	368/368 (100%)	-0.42	1 (0%)	94	94	28, 38, 59, 83	0
1	B	367/368 (99%)	-0.32	1 (0%)	94	94	30, 49, 72, 99	0
1	C	367/368 (99%)	-0.38	0	100	100	29, 50, 72, 112	0
1	D	367/368 (99%)	-0.25	0	100	100	29, 52, 78, 94	0
All	All	1469/1472 (99%)	-0.34	2 (0%)	95	96	28, 47, 72, 112	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	368	SER	2.4
1	B	367	GLY	2.1

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

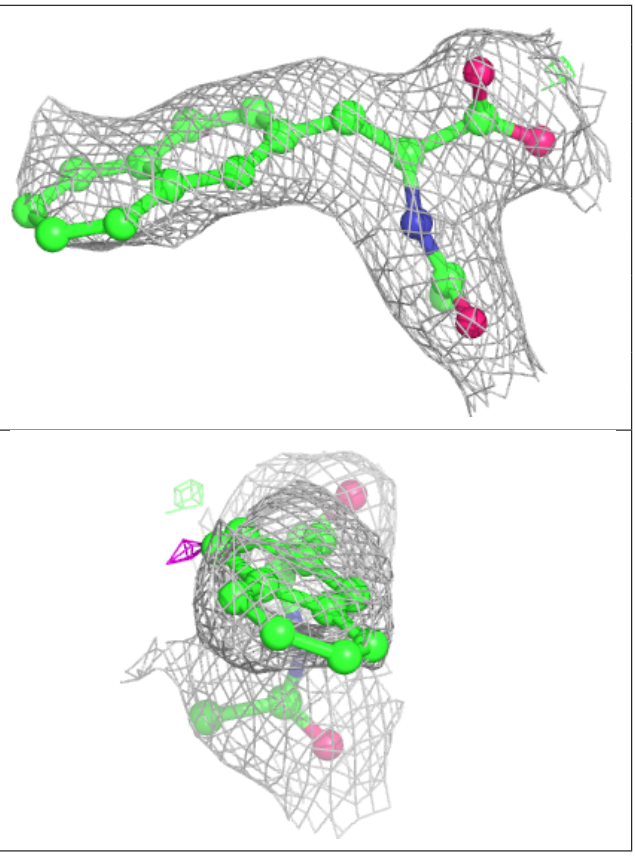
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	MG	B	1369	1/1	0.91	0.09	58,58,58,58	0
2	NPQ	D	1368	19/19	0.95	0.18	44,59,95,97	0
3	MG	A	1370	1/1	0.96	0.11	42,42,42,42	0
2	NPQ	B	1368	19/19	0.96	0.16	41,61,80,80	0
2	NPQ	C	1368	19/19	0.97	0.13	40,50,66,68	0
2	NPQ	A	1369	19/19	0.98	0.16	31,40,46,61	0
3	MG	C	1369	1/1	0.98	0.04	52,52,52,52	0
3	MG	D	1369	1/1	0.98	0.12	54,54,54,54	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.

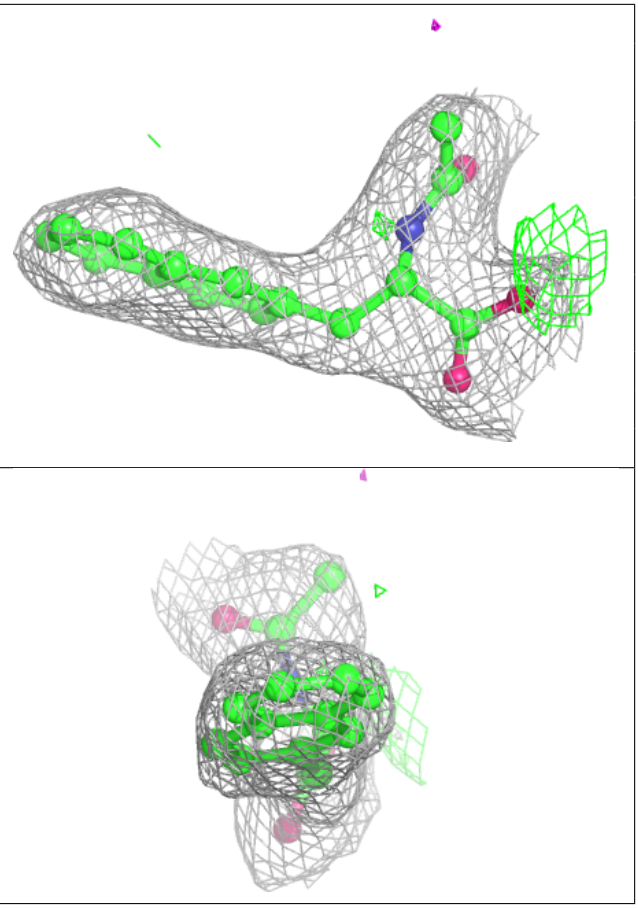
Electron density around NPQ D 1368:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)



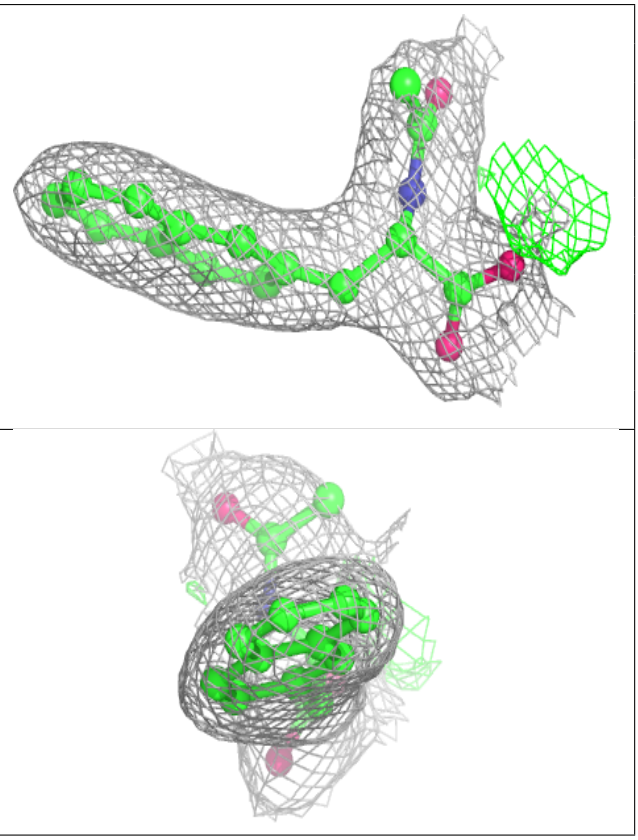
Electron density around NPQ B 1368:

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



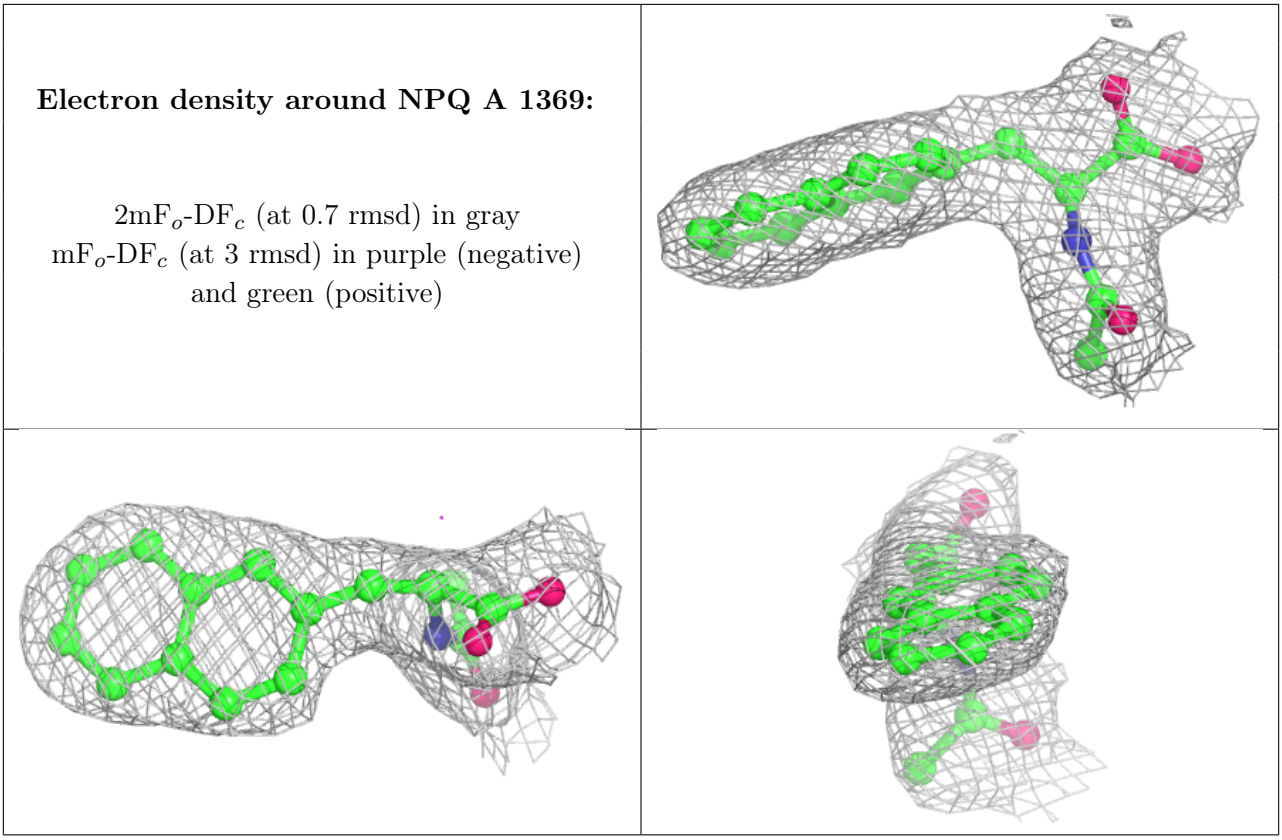
Electron density around NPQ C 1368:

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



Electron density around NPQ A 1369:

$2mF_o$-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.