Full wwPDB X-ray Structure Validation Report ⓘ

Jun 16, 2024 – 11:01 PM EDT

PDB ID	:	5N5S
Title	:	Crystal structure of aldehyde dehydrogenase 21 (ALDH21) from Physcomitrella patens in complex with NADP+
Authors	:	Kopecny, D.; Vigouroux, A.; Briozzo, P.; Morera, S.
Deposited on	:	2017-02-14
Resolution	:	2.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The 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:

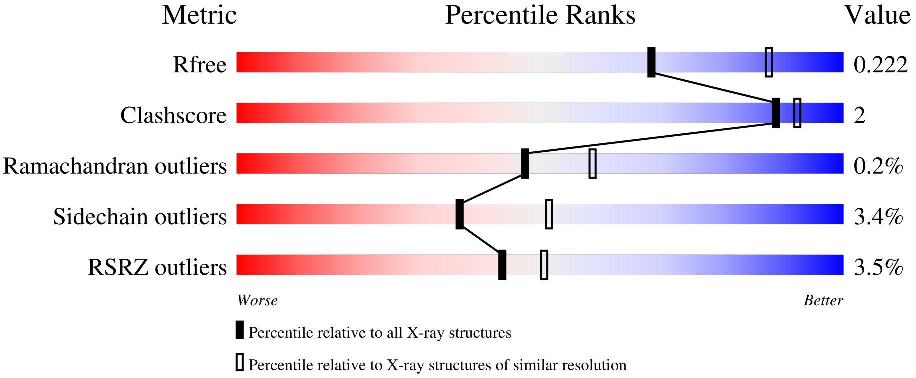
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance ⓘ

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

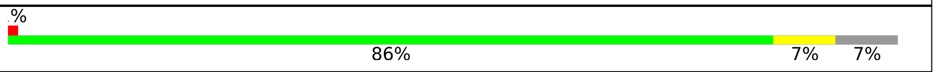
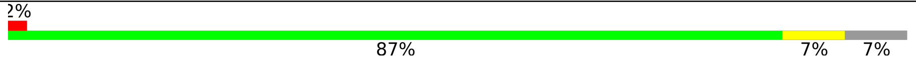
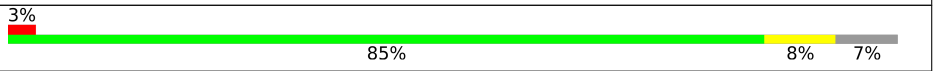
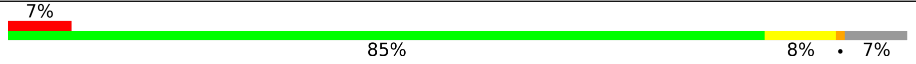
The reported resolution of this entry is 2.30 Å.

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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	515	
1	B	515	
1	C	515	
1	D	515	

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
3	EDO	D	501	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 15245 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 Aldehyde dehydrogenase 21 (ALDH21).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	480	Total 3713	C 2374	N 626	O 695	S 18	0	0	0
1	B	481	Total 3720	C 2379	N 627	O 696	S 18	0	0	0
1	C	481	Total 3720	C 2379	N 627	O 696	S 18	0	0	0
1	D	480	Total 3713	C 2374	N 626	O 695	S 18	0	0	0

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-17	MET	-	initiating methionine	UNP A9SS48
A	-16	GLY	-	expression tag	UNP A9SS48
A	-15	SER	-	expression tag	UNP A9SS48
A	-14	SER	-	expression tag	UNP A9SS48
A	-13	HIS	-	expression tag	UNP A9SS48
A	-12	HIS	-	expression tag	UNP A9SS48
A	-11	HIS	-	expression tag	UNP A9SS48
A	-10	HIS	-	expression tag	UNP A9SS48
A	-9	HIS	-	expression tag	UNP A9SS48
A	-8	HIS	-	expression tag	UNP A9SS48
A	-7	SER	-	expression tag	UNP A9SS48
A	-6	GLN	-	expression tag	UNP A9SS48
A	-5	ASP	-	expression tag	UNP A9SS48
A	-4	PRO	-	expression tag	UNP A9SS48
A	-3	ASN	-	expression tag	UNP A9SS48
A	-2	SER	-	expression tag	UNP A9SS48
A	-1	SER	-	expression tag	UNP A9SS48
A	0	SER	-	expression tag	UNP A9SS48
B	-17	MET	-	initiating methionine	UNP A9SS48
B	-16	GLY	-	expression tag	UNP A9SS48
B	-15	SER	-	expression tag	UNP A9SS48

Continued on next page...

Continued from previous page...

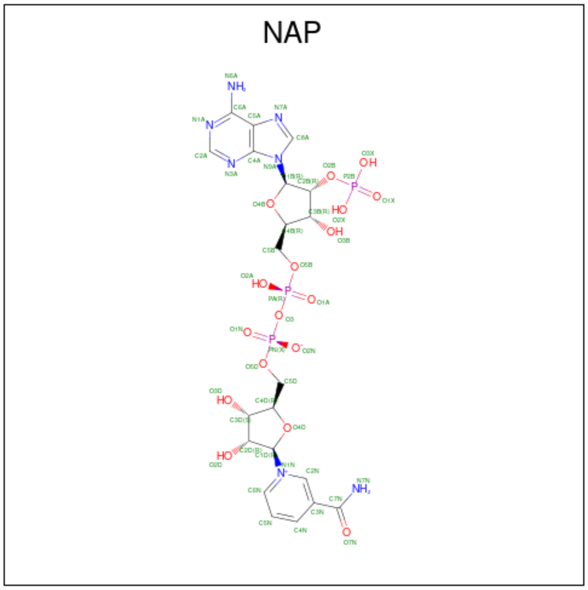
Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	SER	-	expression tag	UNP A9SS48
B	-13	HIS	-	expression tag	UNP A9SS48
B	-12	HIS	-	expression tag	UNP A9SS48
B	-11	HIS	-	expression tag	UNP A9SS48
B	-10	HIS	-	expression tag	UNP A9SS48
B	-9	HIS	-	expression tag	UNP A9SS48
B	-8	HIS	-	expression tag	UNP A9SS48
B	-7	SER	-	expression tag	UNP A9SS48
B	-6	GLN	-	expression tag	UNP A9SS48
B	-5	ASP	-	expression tag	UNP A9SS48
B	-4	PRO	-	expression tag	UNP A9SS48
B	-3	ASN	-	expression tag	UNP A9SS48
B	-2	SER	-	expression tag	UNP A9SS48
B	-1	SER	-	expression tag	UNP A9SS48
B	0	SER	-	expression tag	UNP A9SS48
C	-17	MET	-	initiating methionine	UNP A9SS48
C	-16	GLY	-	expression tag	UNP A9SS48
C	-15	SER	-	expression tag	UNP A9SS48
C	-14	SER	-	expression tag	UNP A9SS48
C	-13	HIS	-	expression tag	UNP A9SS48
C	-12	HIS	-	expression tag	UNP A9SS48
C	-11	HIS	-	expression tag	UNP A9SS48
C	-10	HIS	-	expression tag	UNP A9SS48
C	-9	HIS	-	expression tag	UNP A9SS48
C	-8	HIS	-	expression tag	UNP A9SS48
C	-7	SER	-	expression tag	UNP A9SS48
C	-6	GLN	-	expression tag	UNP A9SS48
C	-5	ASP	-	expression tag	UNP A9SS48
C	-4	PRO	-	expression tag	UNP A9SS48
C	-3	ASN	-	expression tag	UNP A9SS48
C	-2	SER	-	expression tag	UNP A9SS48
C	-1	SER	-	expression tag	UNP A9SS48
C	0	SER	-	expression tag	UNP A9SS48
D	-17	MET	-	initiating methionine	UNP A9SS48
D	-16	GLY	-	expression tag	UNP A9SS48
D	-15	SER	-	expression tag	UNP A9SS48
D	-14	SER	-	expression tag	UNP A9SS48
D	-13	HIS	-	expression tag	UNP A9SS48
D	-12	HIS	-	expression tag	UNP A9SS48
D	-11	HIS	-	expression tag	UNP A9SS48
D	-10	HIS	-	expression tag	UNP A9SS48
D	-9	HIS	-	expression tag	UNP A9SS48

Continued on next page...

Continued from previous page...

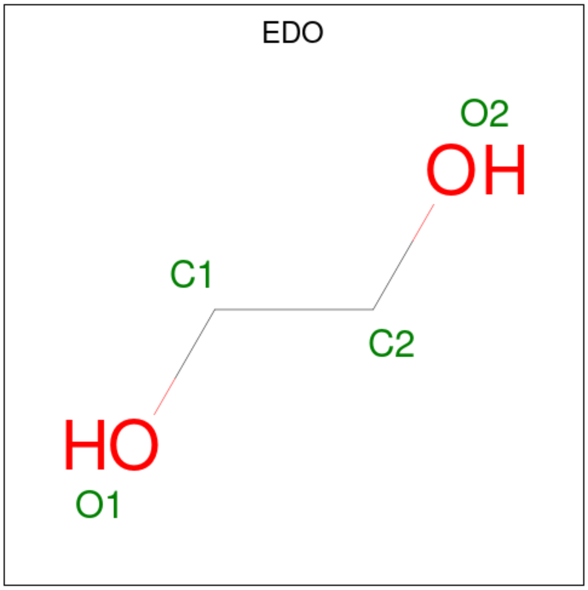
Chain	Residue	Modelled	Actual	Comment	Reference
D	-8	HIS	-	expression tag	UNP A9SS48
D	-7	SER	-	expression tag	UNP A9SS48
D	-6	GLN	-	expression tag	UNP A9SS48
D	-5	ASP	-	expression tag	UNP A9SS48
D	-4	PRO	-	expression tag	UNP A9SS48
D	-3	ASN	-	expression tag	UNP A9SS48
D	-2	SER	-	expression tag	UNP A9SS48
D	-1	SER	-	expression tag	UNP A9SS48
D	0	SER	-	expression tag	UNP A9SS48

- Molecule 2 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: $C_{21}H_{28}N_7O_{17}P_3$).



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

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 4 is water.

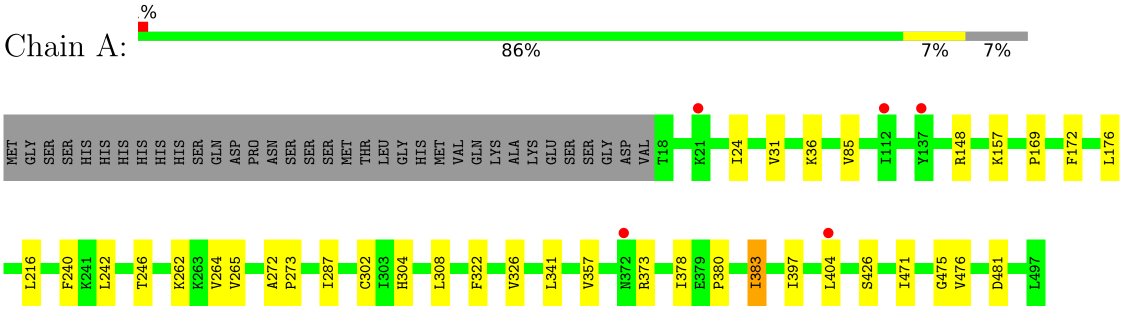
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	65	Total O 65 65	0	0
4	B	42	Total O 42 42	0	0
4	C	37	Total O 37 37	0	0
4	D	31	Total O 31 31	0	0

3 Residue-property plots

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.

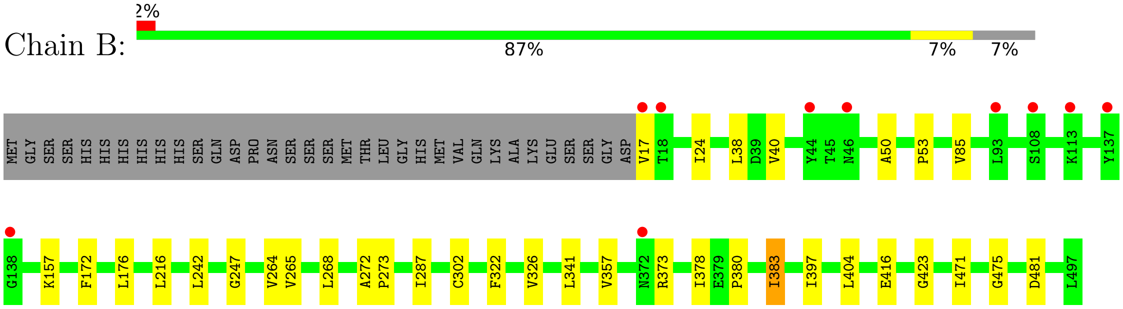
- Molecule 1: Aldehyde dehydrogenase 21 (ALDH21)

Chain A:



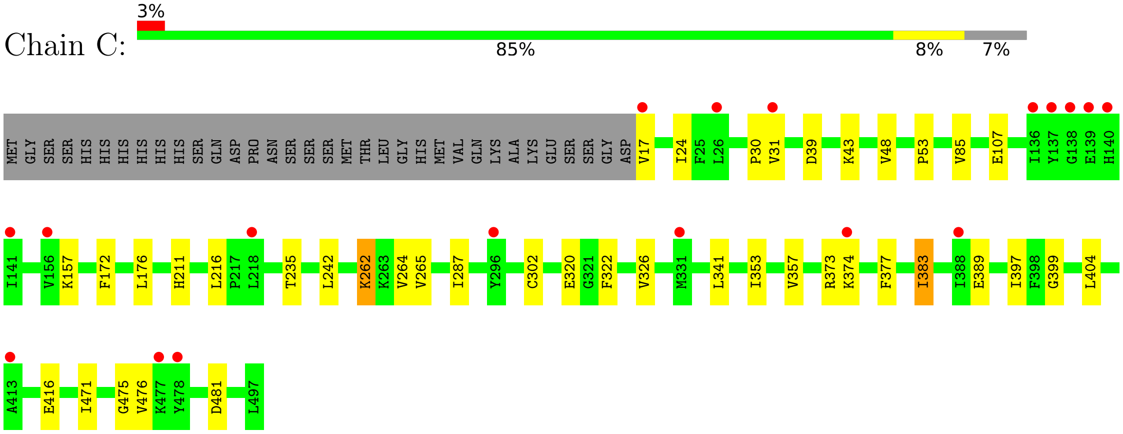
- Molecule 1: Aldehyde dehydrogenase 21 (ALDH21)

Chain B:



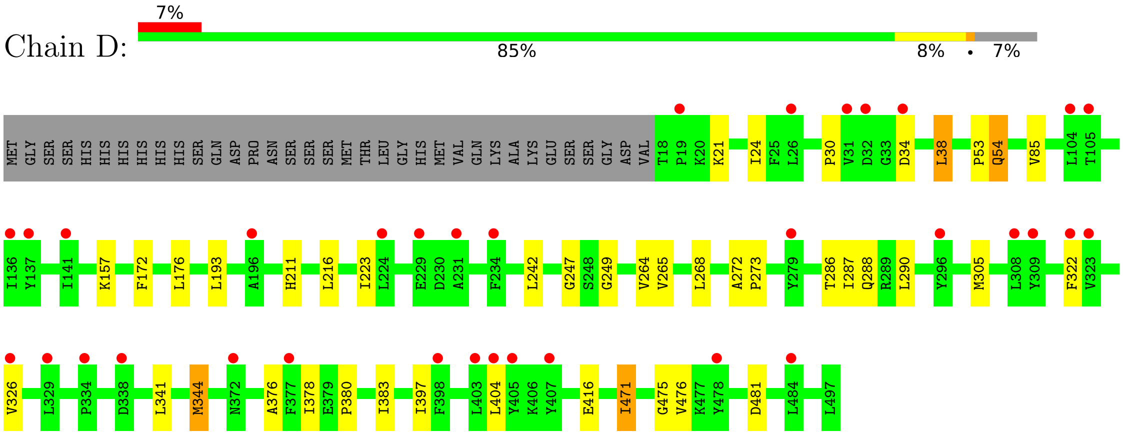
- Molecule 1: Aldehyde dehydrogenase 21 (ALDH21)

Chain C:



- Molecule 1: Aldehyde dehydrogenase 21 (ALDH21)

Chain D:


7%
85%
8%
•
7%
MET
GLY
SER
SER
HIS
HIS
HIS
HIS
HIS
HIS
SER
GLN
ASP
PRO
ASN
SER
SER
SER
MET
THR
LEU
GLY
HIS
MET
VAL
GLN
LYS
ALA
LYS
GLU
SER
SER
GLY
ASP
VAL
T18
P19
K20
K21
I24
F25
L26
P30
V31
D32
G33
D34
L38
P53
Q54
V85
L104
T105
I136
Y137
I141
K157
F172
L176
L193
A196
H211
L216
I223
L224
E229
D230
A231
F234
L242
G247
S248
G249
V264
V265
L268
A272
P273
Y279
T286
I287
Q288
R289
L290
Y296
M305
L308
Y309
F322
V323
V326
L329
P334
D338
L341
M344
N372
A376
F377
I378
E379
P380
I383
I397
F398
L403
L404
Y405
K406
Y407
E416
I471
G475
V476
K477
Y478
D481
L484
L497

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	89.38Å 165.28Å 138.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.13 – 2.30 46.13 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (46.13-2.30) 99.9 (46.13-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	1.33 (at 2.29Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.183 , 0.207 0.200 , 0.222	Depositor DCC
R_{free} test set	4579 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	61.2	Xtriage
Anisotropy	0.428	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 47.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	15245	wwPDB-VP
Average B, all atoms (Å^2)	83.0	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 32.17 % 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 9.8245e-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 (i)

5.1 Standard geometry (i)

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

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	0.44	0/3789	0.66	1/5123 (0.0%)
1	B	0.44	0/3796	0.66	1/5133 (0.0%)
1	C	0.43	0/3796	0.67	1/5133 (0.0%)
1	D	0.43	0/3789	0.66	1/5123 (0.0%)
All	All	0.44	0/15170	0.66	4/20512 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	475	GLY	N-CA-C	-5.74	98.75	113.10
1	D	475	GLY	N-CA-C	-5.74	98.76	113.10
1	C	475	GLY	N-CA-C	-5.47	99.42	113.10
1	A	475	GLY	N-CA-C	-5.31	99.83	113.10

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	3713	0	3756	13	0
1	B	3720	0	3765	12	0
1	C	3720	0	3765	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	3713	0	3756	16	0
2	A	48	0	25	1	0
2	B	48	0	25	1	0
2	C	48	0	25	1	0
2	D	48	0	25	0	0
3	B	4	0	6	0	0
3	D	8	0	12	0	0
4	A	65	0	0	0	0
4	B	42	0	0	0	0
4	C	37	0	0	0	0
4	D	31	0	0	0	0
All	All	15245	0	15160	57	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 2.

All (57) 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:C:302:CYS:SG	2:C:501:NAP:C4N	2.85	0.65
1:A:302:CYS:SG	2:A:600:NAP:C4N	2.90	0.60
1:D:264:VAL:HG12	1:D:471:ILE:HD11	1.84	0.58
1:D:38:LEU:HB2	1:D:54:GLN:HG2	1.87	0.57
1:B:302:CYS:SG	2:B:502:NAP:C4N	2.93	0.57
1:C:172:PHE:HB2	1:C:176:LEU:HD12	1.92	0.51
1:C:235:THR:O	1:C:262:LYS:HE2	2.11	0.51
1:A:242:LEU:HD11	1:A:265:VAL:HG23	1.92	0.50
1:D:24:ILE:HG23	1:D:53:PRO:HD2	1.94	0.50
1:B:172:PHE:HB2	1:B:176:LEU:HD12	1.94	0.49
1:B:242:LEU:HD11	1:B:265:VAL:HG23	1.93	0.49
1:C:242:LEU:HD11	1:C:265:VAL:HG23	1.94	0.49
1:D:172:PHE:HB2	1:D:176:LEU:HD12	1.94	0.48
1:D:193:LEU:HD23	1:D:223:ILE:HG12	1.94	0.48
1:D:242:LEU:HD11	1:D:265:VAL:HG23	1.94	0.48
1:A:172:PHE:HB2	1:A:176:LEU:HD12	1.94	0.47
1:D:286:THR:O	1:D:290:LEU:HG	2.15	0.46
1:C:264:VAL:HG12	1:C:471:ILE:HD11	1.98	0.46
1:B:264:VAL:HG12	1:B:471:ILE:HD11	1.98	0.46
1:D:85:VAL:HG11	1:D:216:LEU:HD13	1.98	0.45
1:A:264:VAL:HG12	1:A:471:ILE:HD11	1.98	0.45
1:C:24:ILE:HG23	1:C:53:PRO:HD2	1.98	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:247:GLY:O	1:B:268:LEU:HA	2.16	0.44
1:C:30:PRO:HD3	1:C:211:HIS:CD2	2.51	0.44
1:C:85:VAL:HG11	1:C:216:LEU:HD13	1.99	0.44
1:C:322:PHE:O	1:C:326:VAL:HG23	2.18	0.44
1:D:290:LEU:CD2	1:D:305:MET:HE1	2.49	0.43
1:D:322:PHE:O	1:D:326:VAL:HG23	2.18	0.43
1:A:85:VAL:HG11	1:A:216:LEU:HD13	1.99	0.43
1:A:322:PHE:O	1:A:326:VAL:HG23	2.18	0.43
1:C:43:LYS:HD3	1:C:107:GLU:O	2.18	0.43
1:B:322:PHE:O	1:B:326:VAL:HG23	2.19	0.43
1:A:326:VAL:HG13	1:A:380:PRO:HB2	2.01	0.43
1:A:357:VAL:HG22	1:A:383:ILE:HD11	2.02	0.42
1:B:272:ALA:HA	1:B:273:PRO:HD3	1.94	0.42
1:B:24:ILE:HG23	1:B:53:PRO:HD2	2.01	0.42
1:B:85:VAL:HG11	1:B:216:LEU:HD13	2.01	0.42
1:C:24:ILE:HG13	1:C:31:VAL:HG23	2.02	0.42
1:D:344:MET:HB2	1:D:376:ALA:O	2.20	0.42
1:D:30:PRO:HD3	1:D:211:HIS:CE1	2.55	0.41
1:A:272:ALA:HA	1:A:273:PRO:HD3	1.96	0.41
1:D:326:VAL:HG13	1:D:380:PRO:HB2	2.03	0.41
1:D:247:GLY:O	1:D:268:LEU:HA	2.20	0.41
1:B:326:VAL:HG13	1:B:380:PRO:HB2	2.02	0.41
1:B:357:VAL:HG22	1:B:383:ILE:HD11	2.03	0.41
1:A:24:ILE:HG13	1:A:31:VAL:HG23	2.01	0.41
1:C:353:ILE:HG12	1:C:399:GLY:HA3	2.03	0.41
1:D:249:GLY:HA2	1:D:268:LEU:HD13	2.03	0.41
1:D:272:ALA:HA	1:D:273:PRO:HD3	2.00	0.41
1:A:169:PRO:HG3	1:A:246:THR:HG22	2.03	0.40
1:C:357:VAL:HG22	1:C:383:ILE:HD11	2.03	0.40
1:C:374:LYS:HG3	1:C:377:PHE:HB2	2.02	0.40
1:C:39:ASP:HB3	1:C:48:VAL:HG12	2.04	0.40
1:C:172:PHE:HB2	1:C:176:LEU:CD1	2.51	0.40
1:A:240:PHE:HB2	1:A:262:LYS:HE3	2.03	0.40
1:A:308:LEU:HD23	1:A:404:LEU:CD1	2.50	0.40
1:B:40:VAL:HB	1:B:50:ALA:HB3	2.03	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	478/515 (93%)	458 (96%)	19 (4%)	1 (0%)	47	58
1	B	479/515 (93%)	458 (96%)	20 (4%)	1 (0%)	47	58
1	C	479/515 (93%)	460 (96%)	18 (4%)	1 (0%)	47	58
1	D	478/515 (93%)	459 (96%)	18 (4%)	1 (0%)	47	58
All	All	1914/2060 (93%)	1835 (96%)	75 (4%)	4 (0%)	47	58

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	476	VAL
1	A	476	VAL
1	D	476	VAL
1	B	423	GLY

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	394/425 (93%)	382 (97%)	12 (3%)	41	57
1	B	395/425 (93%)	383 (97%)	12 (3%)	41	57
1	C	395/425 (93%)	382 (97%)	13 (3%)	38	53
1	D	394/425 (93%)	378 (96%)	16 (4%)	30	43
All	All	1578/1700 (93%)	1525 (97%)	53 (3%)	37	51

All (53) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	36	LYS
1	A	148	ARG
1	A	157	LYS
1	A	287	ILE
1	A	304	HIS
1	A	341	LEU
1	A	373	ARG
1	A	378	ILE
1	A	383	ILE
1	A	397	ILE
1	A	426	SER
1	A	481	ASP
1	B	17	VAL
1	B	38	LEU
1	B	157	LYS
1	B	287	ILE
1	B	341	LEU
1	B	373	ARG
1	B	378	ILE
1	B	383	ILE
1	B	397	ILE
1	B	404	LEU
1	B	416	GLU
1	B	481	ASP
1	C	17	VAL
1	C	157	LYS
1	C	262	LYS
1	C	287	ILE
1	C	320	GLU
1	C	341	LEU
1	C	373	ARG
1	C	383	ILE
1	C	389	GLU
1	C	397	ILE
1	C	404	LEU
1	C	416	GLU
1	C	481	ASP
1	D	21	LYS
1	D	34	ASP
1	D	38	LEU
1	D	54	GLN
1	D	157	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	287	ILE
1	D	288	GLN
1	D	341	LEU
1	D	344	MET
1	D	378	ILE
1	D	383	ILE
1	D	397	ILE
1	D	404	LEU
1	D	416	GLU
1	D	471	ILE
1	D	481	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	421	HIS
1	C	211	HIS
1	C	421	HIS

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 ⓘ

7 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	$\#\|Z\| > 2$
3	EDO	D	502	-	3,3,3	0.71	0	2,2,2	0.19	0
3	EDO	B	501	-	3,3,3	0.75	0	2,2,2	0.06	0
2	NAP	A	600	-	45,52,52	0.99	1 (2%)	56,80,80	0.95	2 (3%)
2	NAP	B	502	-	45,52,52	1.01	1 (2%)	56,80,80	1.11	2 (3%)
3	EDO	D	501	-	3,3,3	0.79	0	2,2,2	0.18	0
2	NAP	D	503	-	45,52,52	1.06	1 (2%)	56,80,80	0.98	2 (3%)
2	NAP	C	501	-	45,52,52	1.01	1 (2%)	56,80,80	1.00	2 (3%)

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	EDO	D	502	-	-	0/1/1/1	-
3	EDO	B	501	-	-	0/1/1/1	-
2	NAP	A	600	-	-	2/31/67/67	0/5/5/5
2	NAP	B	502	-	-	3/31/67/67	0/5/5/5
3	EDO	D	501	-	-	0/1/1/1	-
2	NAP	D	503	-	-	7/31/67/67	0/5/5/5
2	NAP	C	501	-	-	4/31/67/67	0/5/5/5

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	503	NAP	C2N-N1N	5.92	1.42	1.35
2	C	501	NAP	C2N-N1N	5.77	1.42	1.35
2	B	502	NAP	C2N-N1N	5.46	1.41	1.35
2	A	600	NAP	C2N-N1N	5.43	1.41	1.35

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	502	NAP	O4D-C1D-C2D	-5.43	98.99	106.93
2	D	503	NAP	O4D-C1D-C2D	-4.43	100.46	106.93
2	C	501	NAP	O4D-C1D-C2D	-4.05	101.01	106.93

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	600	NAP	O4D-C1D-C2D	-3.70	101.51	106.93
2	A	600	NAP	C5A-C6A-N6A	2.29	123.83	120.35
2	D	503	NAP	C5A-C6A-N6A	2.29	123.83	120.35
2	B	502	NAP	C5A-C6A-N6A	2.20	123.69	120.35
2	C	501	NAP	C5A-C6A-N6A	2.16	123.63	120.35

There are no chirality outliers.

All (16) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	503	NAP	C5B-O5B-PA-O1A
2	D	503	NAP	C5D-O5D-PN-O1N
2	D	503	NAP	C4D-C5D-O5D-PN
2	D	503	NAP	C3B-C2B-O2B-P2B
2	A	600	NAP	C4D-C5D-O5D-PN
2	C	501	NAP	C4D-C5D-O5D-PN
2	B	502	NAP	C4D-C5D-O5D-PN
2	C	501	NAP	C3B-C2B-O2B-P2B
2	D	503	NAP	C1B-C2B-O2B-P2B
2	B	502	NAP	C2B-O2B-P2B-O1X
2	C	501	NAP	C2B-O2B-P2B-O1X
2	D	503	NAP	C2B-O2B-P2B-O1X
2	A	600	NAP	C2B-O2B-P2B-O3X
2	D	503	NAP	C5D-O5D-PN-O3
2	C	501	NAP	C1B-C2B-O2B-P2B
2	B	502	NAP	C5B-O5B-PA-O1A

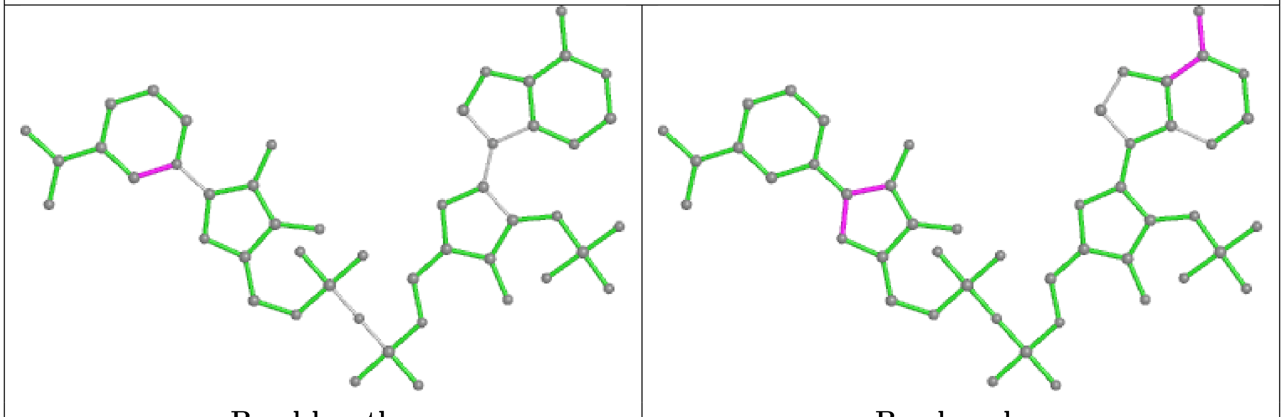
There are no ring outliers.

3 monomers are involved in 3 short contacts:

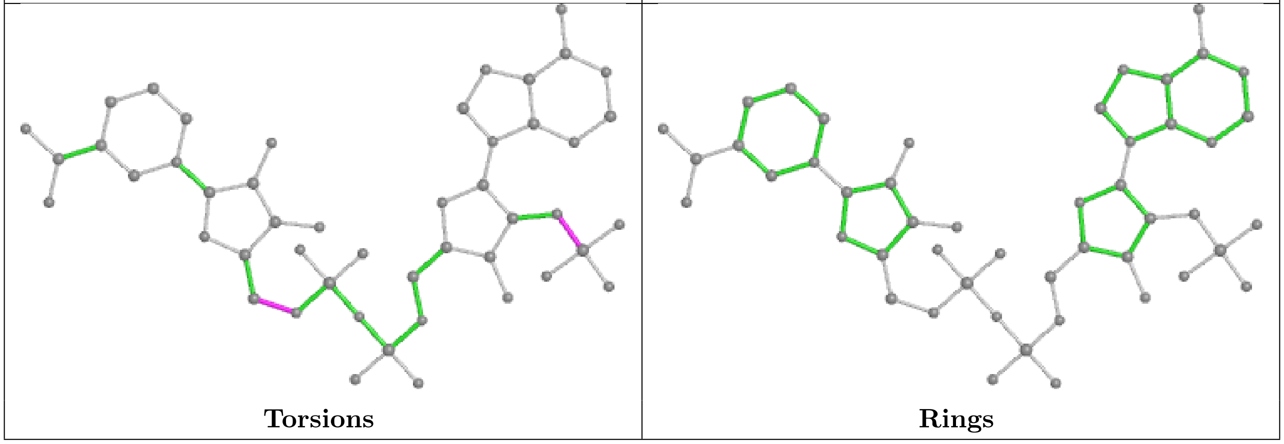
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	600	NAP	1	0
2	B	502	NAP	1	0
2	C	501	NAP	1	0

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 NAP A 600

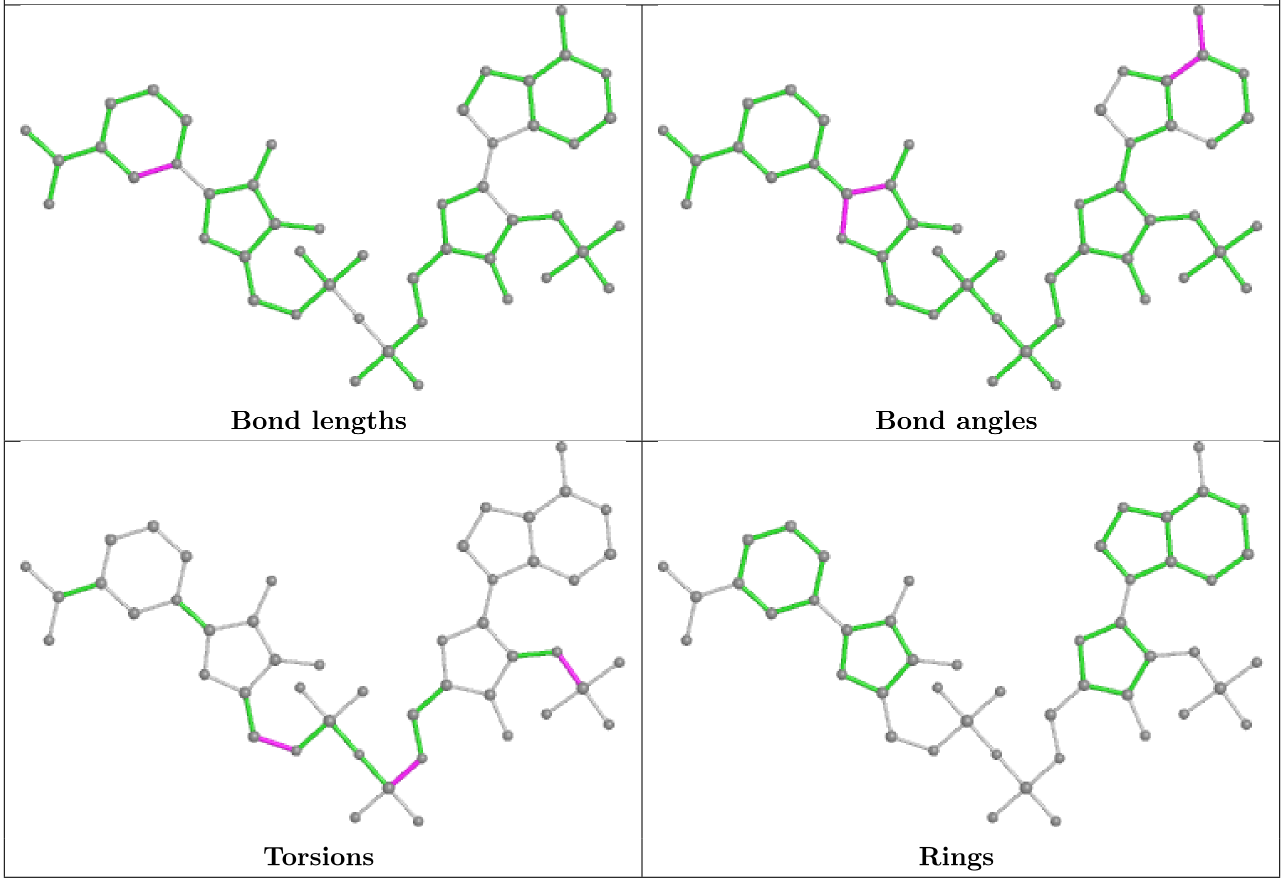


Bond lengths **Bond angles**



Torsions **Rings**

Ligand NAP B 502



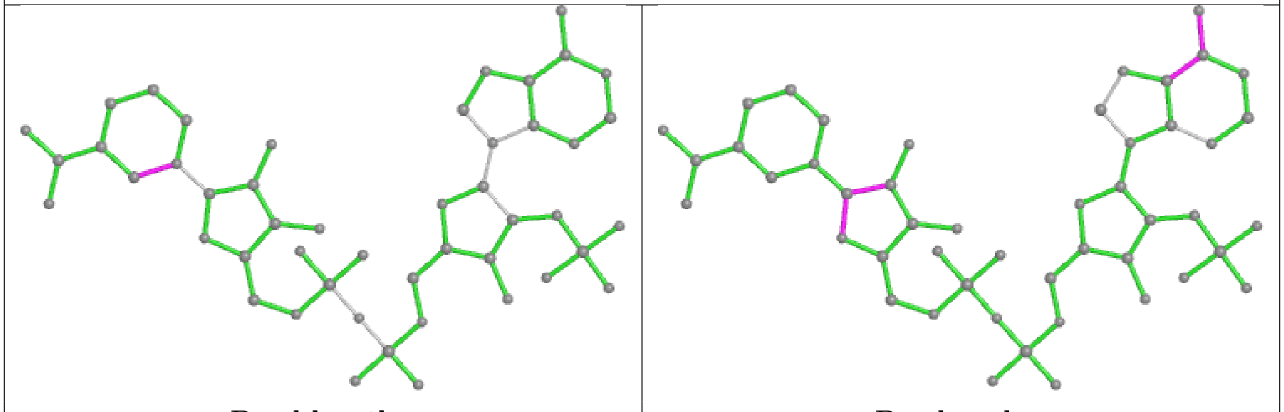
Bond lengths

Bond angles

Torsions

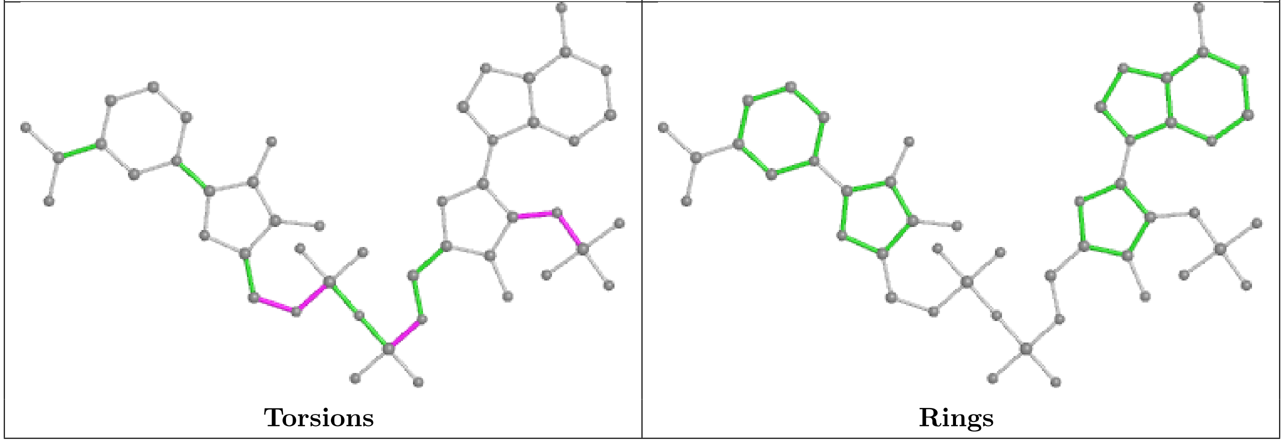
Rings

Ligand NAP D 503



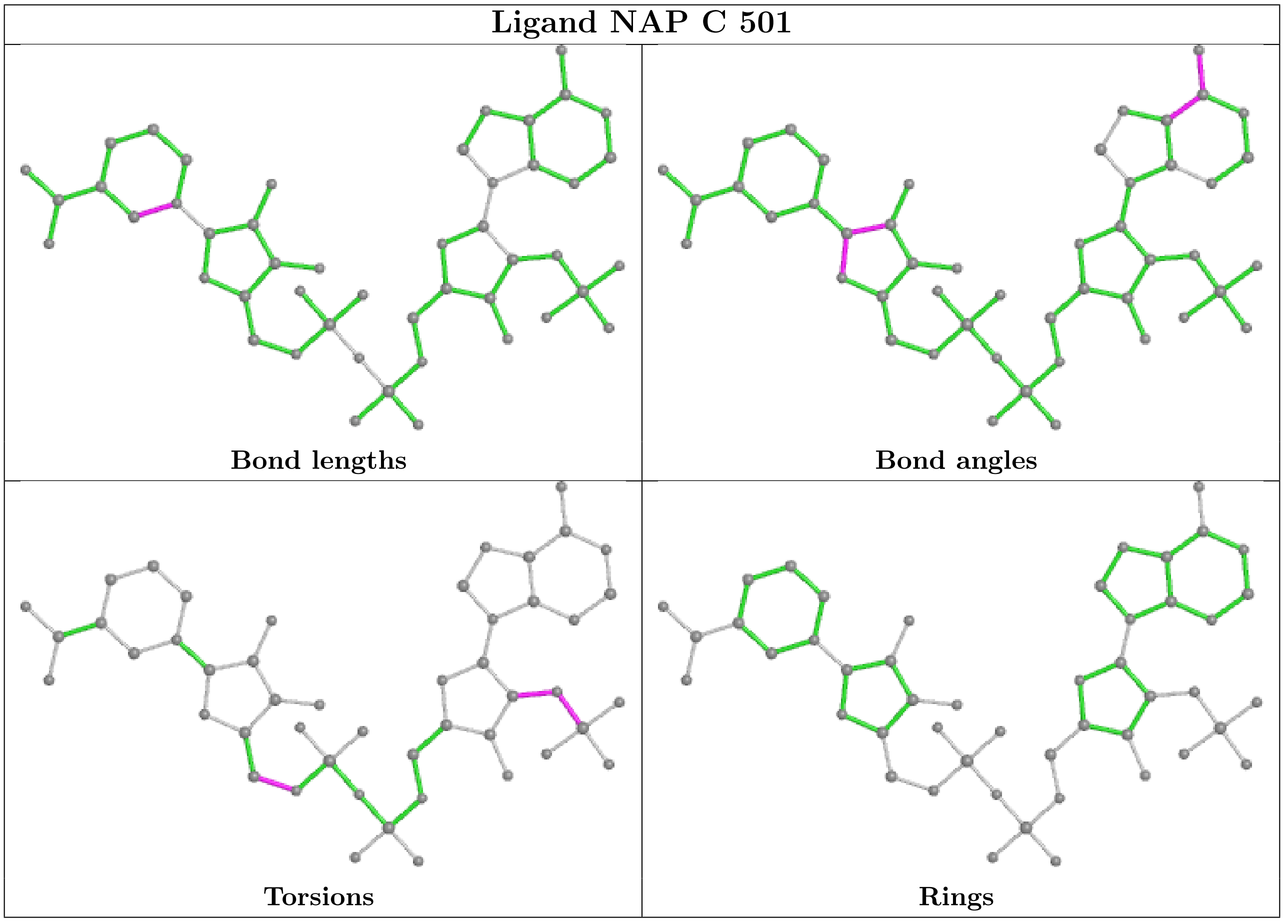
Bond lengths

Bond angles



Torsions

Rings



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, 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	480/515 (93%)	0.05	5 (1%)	82	86	51, 71, 103, 129	0
1	B	481/515 (93%)	0.09	10 (2%)	63	70	52, 76, 113, 144	0
1	C	481/515 (93%)	0.23	18 (3%)	41	48	50, 78, 115, 144	0
1	D	480/515 (93%)	0.49	34 (7%)	16	21	46, 93, 151, 173	0
All	All	1922/2060 (93%)	0.22	67 (3%)	44	51	46, 79, 127, 173	0

All (67) RSRZ outliers are listed below:

Mol	**Chain**	**Res**	**Type**	**RSRZ**
1	B	17	VAL	5.8
1	C	17	VAL	5.7
1	D	26	LEU	4.8
1	D	404	LEU	4.5
1	D	296	TYR	4.4
1	B	18	THR	4.4
1	D	326	VAL	3.9
1	D	32	ASP	3.9
1	A	372	ASN	3.8
1	A	21	LYS	3.6
1	C	137	TYR	3.6
1	D	279	TYR	3.5
1	B	372	ASN	3.5
1	B	44	TYR	3.4
1	D	334	PRO	3.2
1	D	323	VAL	3.1
1	D	229	GLU	3.1
1	D	104	LEU	3.1
1	D	105	THR	3.0
1	D	137	TYR	2.9
1	D	31	VAL	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	31	VAL	2.8
1	A	404	LEU	2.8
1	D	19	PRO	2.7
1	D	309	TYR	2.7
1	D	224	LEU	2.7
1	D	329	LEU	2.7
1	D	322	PHE	2.7
1	A	137	TYR	2.7
1	D	196	ALA	2.7
1	C	139	GLU	2.6
1	C	478	TYR	2.6
1	D	234	PHE	2.6
1	D	308	LEU	2.6
1	A	112	ILE	2.6
1	D	338	ASP	2.5
1	C	136	ILE	2.5
1	C	26	LEU	2.5
1	C	331	MET	2.4
1	C	141	ILE	2.4
1	C	374	LYS	2.4
1	C	156	VAL	2.4
1	C	138	GLY	2.4
1	C	140	HIS	2.4
1	D	407	TYR	2.3
1	D	484	LEU	2.3
1	C	218	LEU	2.3
1	D	403	LEU	2.3
1	B	93	LEU	2.3
1	B	113	LYS	2.2
1	D	372	ASN	2.2
1	B	46	ASN	2.2
1	D	398	PHE	2.1
1	D	136	ILE	2.1
1	D	478	TYR	2.1
1	C	477	LYS	2.1
1	D	377	PHE	2.1
1	D	405	TYR	2.1
1	D	141	ILE	2.1
1	D	34	ASP	2.1
1	C	388	ILE	2.1
1	C	413	ALA	2.1
1	B	138	GLY	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	137	TYR	2.1
1	D	231	ALA	2.0
1	B	108	SER	2.0
1	C	296	TYR	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 ⓘ

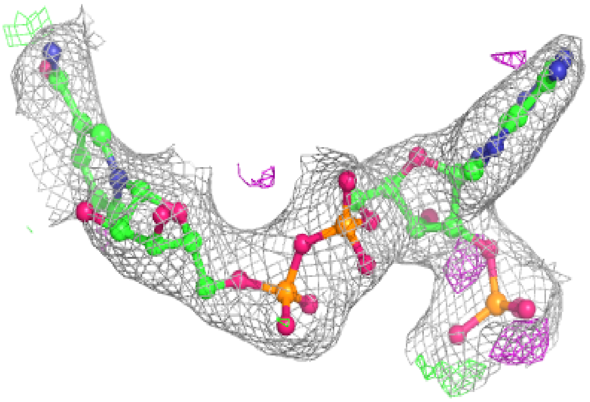
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

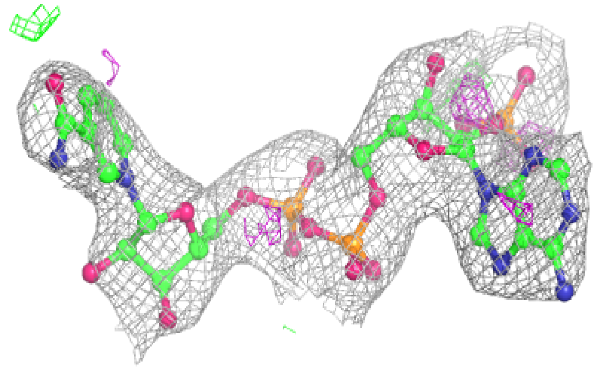
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	EDO	D	501	4/4	0.64	0.42	68,74,75,76	0
3	EDO	D	502	4/4	0.86	0.20	71,74,75,77	0
3	EDO	B	501	4/4	0.88	0.15	73,77,80,80	0
2	NAP	D	503	48/48	0.91	0.15	86,96,103,104	0
2	NAP	B	502	48/48	0.95	0.11	52,81,92,93	0
2	NAP	C	501	48/48	0.96	0.13	61,74,78,80	0
2	NAP	A	600	48/48	0.98	0.12	61,70,90,91	0

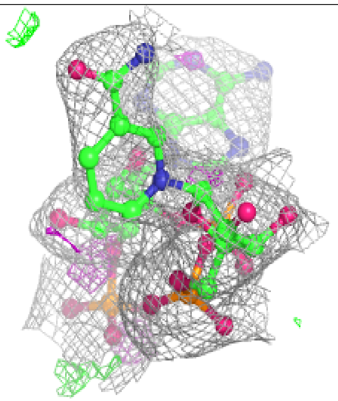
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around NAP D 503:

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

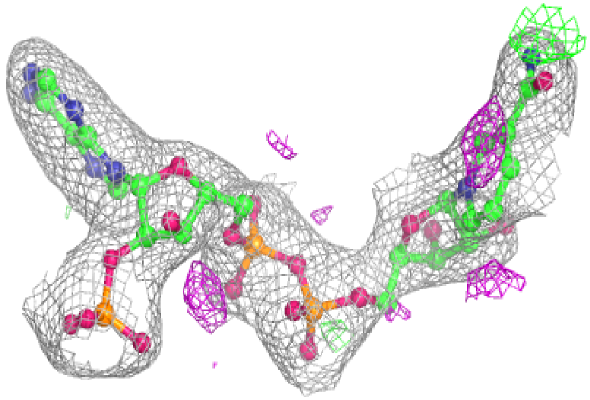


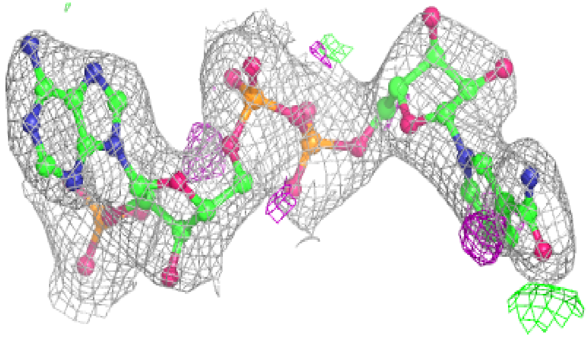


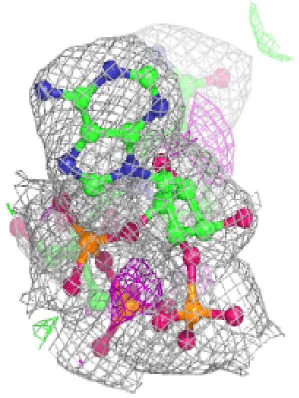


Electron density around NAP B 502:

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

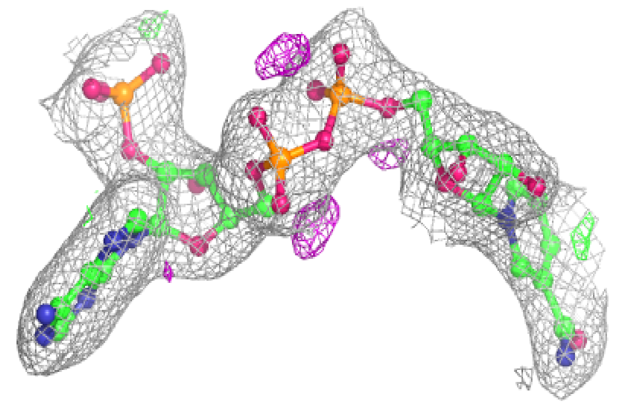


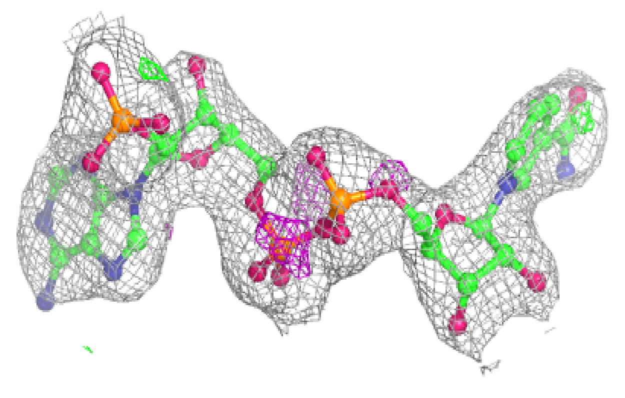


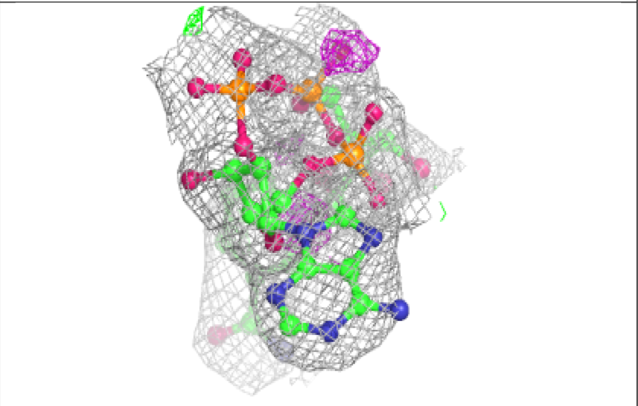


Electron density around NAP C 501:

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

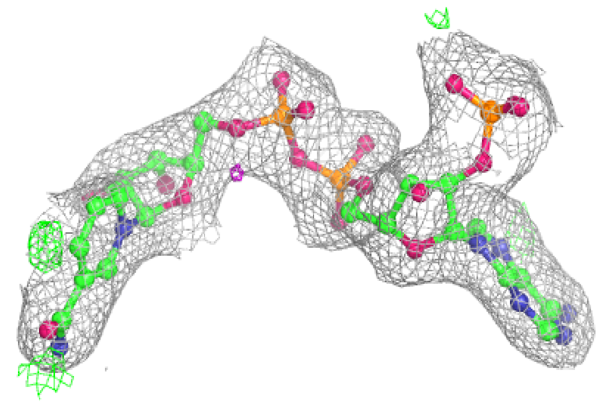


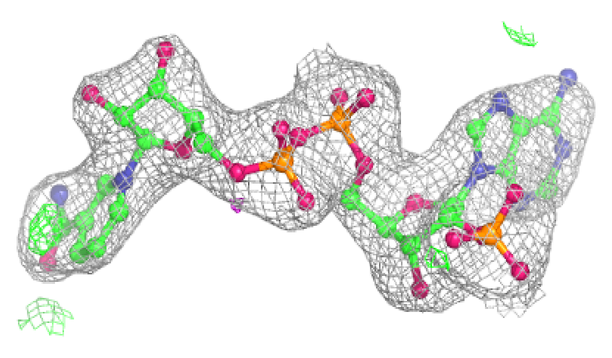


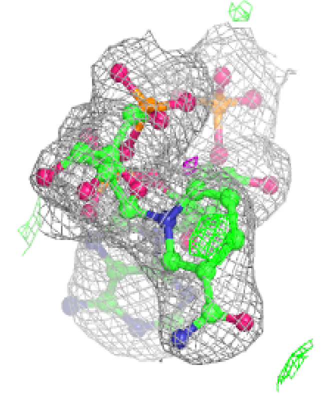


Electron density around NAP A 600:

$2mF_o$-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.