

# wwPDB X-ray Structure Validation Summary Report (i)

Sep 23, 2024 – 04:17 pm BST

PDB ID	:	8QMT
Title	:	Succinic semialdehyde dehydrogenase from E. coli with Q262R substitution
		and bound NAD+, succinic semialdehyde
Authors	:	He, H.; Zarzycki, J.; Erb, T.J.
Deposited on	:	2023-09-25
Resolution	:	1.80  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 (1)) were used in the production of this report:

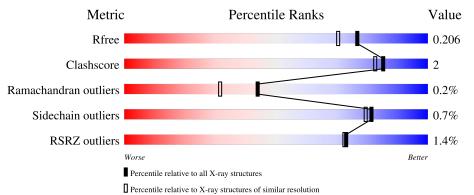
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as $541$ be (2020)
Xtriage (Phenix)	:	1.13
$\mathrm{EDS}$	:	3.0
buster-report	:	1.1.7(2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.002 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.2

## 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
$R_{free}$	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (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 >=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	А	462	% 92%	8%
1	В	462	3% 94%	6%
1	С	462	93%	6% •
1	D	462	% 95%	•



## 2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 15409 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Δ	460	Total	С	Ν	0	$\mathbf{S}$	0	0	0
	А	400	3478	2190	617	650	21	0	0	0
1	В	460	Total	С	Ν	0	S	0	0	0
	D	400	3478	2190	617	650	21	0	0	0
1	С	460	Total	С	Ν	0	S	0	0	0
	U	400	3478	2190	617	650	21	0		0
1	Л	460	Total	С	Ν	0	S	0	0	0
		460	3478	2190	617	650	21	0	0	U

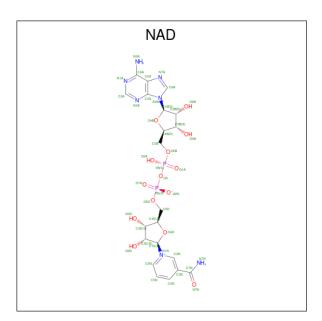
• Molecule 1 is a protein called Succinate semialdehyde dehydrogenase [NAD(P)+] Sad.

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	262	ARG	GLN	engineered mutation	UNP P76149
В	262	ARG	GLN	engineered mutation	UNP P76149
С	262	ARG	GLN	engineered mutation	UNP P76149
D	262	ARG	GLN	engineered mutation	UNP P76149

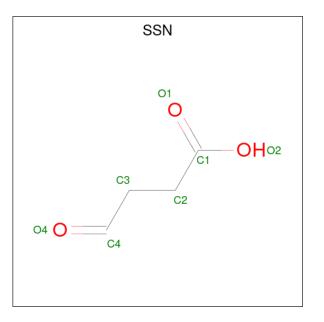
• Molecule 2 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: C<sub>21</sub>H<sub>27</sub>N<sub>7</sub>O<sub>14</sub>P<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
2	٨	1	Total	С	Ν	Ο	Р	0	0
	A	1	44	21	$\overline{7}$	14	2	0	0
2	р	1	Total	С	Ν	Ο	Р	0	0
	D	1	44	21	$\overline{7}$	14	2	0	0
2	С	1	Total	С	Ν	Ο	Р	0	0
	U	1	44	21	7	14	2	0	0
2	Л	1	Total	С	Ν	Ο	Р	0	0
2	D	1	44	21	7	14	2	0	0

• Molecule 3 is 4-oxobutanoic acid (three-letter code: SSN) (formula:  $C_4H_6O_3$ ) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 7 & 4 & 3 \end{array}$	0	0
3	В	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 7 & 4 & 3 \end{array}$	0	0
3	С	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 7 & 4 & 3 \end{array}$	0	0
3	D	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 7 & 4 & 3 \end{array}$	0	0

• Molecule 4 is water.

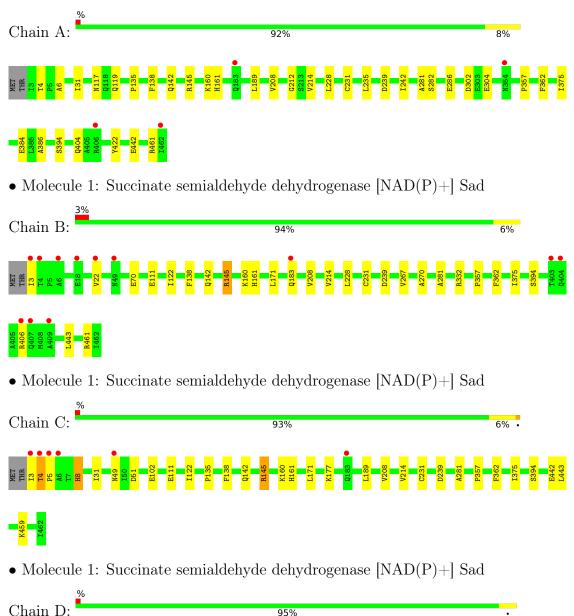
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	331	Total O 331 331	0	0
4	В	311	Total         O           311         311	0	0
4	С	315	Total O 315 315	0	0
4	D	336	Total O 336 336	0	0



#### 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: Succinate semialdehyde dehydrogenase [NAD(P)+] Sad



95%



#### 



## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	91.71Å 115.11Å 179.38Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	29.55 - 1.80	Depositor
Resolution (A)	29.55 - 1.80	EDS
% Data completeness	98.5 (29.55-1.80)	Depositor
(in resolution range)	98.7(29.55-1.80)	EDS
R <sub>merge</sub>	0.10	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.08 (at 1.79 \text{\AA})$	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
B B.	0.180 , $0.205$	Depositor
$R, R_{free}$	0.180 , $0.206$	DCC
$R_{free}$ test set	173729 reflections $(1.15%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	26.1	Xtriage
Anisotropy	0.786	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.40 , $50.6$	EDS
L-test for twinning <sup>2</sup>	$ < L >=0.49, < L^2>=0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	15409	wwPDB-VP
Average B, all atoms $(Å^2)$	32.0	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 37.57 % 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 4.2039e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: NAD, SSN

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		
	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.36	0/3544	0.58	0/4801	
1	В	0.35	0/3544	0.56	0/4801	
1	С	0.37	0/3544	0.58	1/4801~(0.0%)	
1	D	0.36	0/3544	0.58	0/4801	
All	All	0.36	0/14176	0.58	1/19204~(0.0%)	

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	А	0	2
1	В	0	2
1	С	0	1
1	D	0	1
All	All	0	6

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$Observed(^{o})$	$Ideal(^{o})$
1	С	8	HIS	CB-CA-C	-5.17	100.06	110.40

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	А	145	ARG	Sidechain
	•	a	1	

Continued on next page...



001000	Contentaca from precioas page									
Mol	Chain	$\mathbf{Res}$	Type	Group						
1	А	461	ARG	Sidechain						
1	В	145	ARG	Sidechain						
1	В	332	ARG	Sidechain						
1	С	145	ARG	Sidechain						

Continued from previous page...

#### 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	А	3478	0	3454	19	0
1	В	3478	0	3454	15	0
1	С	3478	0	3454	18	0
1	D	3478	0	3454	12	0
2	А	44	0	26	0	0
2	В	44	0	26	0	0
2	С	44	0	26	0	0
2	D	44	0	26	0	0
3	А	7	0	4	0	0
3	В	7	0	4	0	0
3	С	7	0	4	0	0
3	D	7	0	4	0	0
4	А	331	0	0	2	0
4	В	311	0	0	2	0
4	С	315	0	0	3	0
4	D	336	0	0	3	0
All	All	15409	0	13936	58	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.

The worst 5 of 58 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:A:4:THR:HG23	1:A:6:ALA:H	1.58	0.68
1:D:138:PHE:HB2	1:D:142:GLN:HG2	1.77	0.65
1:D:145:ARG:HG3	1:D:443:LEU:HD21	1.80	0.64

Continued on next page...



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:228:LEU:HD11	1:B:214:VAL:HG13	1.81	0.63
1:A:214:VAL:HG13	1:B:228:LEU:HD11	1.81	0.62

Continued from previous page...

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	Perce	entiles
1	А	458/462~(99%)	449~(98%)	8 (2%)	1 (0%)	44	31
1	В	458/462~(99%)	448 (98%)	10 (2%)	0	100	100
1	С	458/462~(99%)	448 (98%)	9(2%)	1 (0%)	44	31
1	D	458/462~(99%)	450 (98%)	7 (2%)	1 (0%)	44	31
All	All	1832/1848~(99%)	1795 (98%)	34 (2%)	3 (0%)	44	31

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	135	PRO
1	С	135	PRO
1	D	135	PRO

#### 5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain 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 side chain conformation was analysed, and the total number of residues.



Mol	Chain	Analysed	Rotameric	Outliers	F	Perce	ntiles
1	А	345/347~(99%)	343~(99%)	2(1%)		84	82
1	В	345/347~(99%)	343 (99%)	2 (1%)		84	82
1	С	345/347~(99%)	342 (99%)	3 (1%)		75	72
1	D	345/347~(99%)	343~(99%)	2 (1%)		84	82
All	All	1380/1388~(99%)	1371 (99%)	9 (1%)		81	79

5 of 9 residues with a non-rotameric side chain are listed below:

Mol	Chain	$\mathbf{Res}$	Type
1	D	160	LYS
1	D	161	HIS
1	В	161	HIS
1	С	4	THR
1	С	160	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

#### 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 oligosaccharides in this entry.

#### 5.6 Ligand geometry (i)

8 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



Mal	Mol Type Chain		Res Link		Bo	ond leng	ths	Bond angles		
NIOI	туре	Chain	nes		Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2
2	NAD	D	501	-	42,48,48	1.13	5 (11%)	50,73,73	1.29	6 (12%)
2	NAD	В	501	-	42,48,48	1.16	6 (14%)	50,73,73	1.28	6 (12%)
3	SSN	А	502	1	6,6,6	0.85	0	6,6,6	1.62	1 (16%)
3	SSN	С	502	1	6,6,6	0.94	0	6,6,6	0.96	0
3	SSN	В	502	1	6,6,6	1.02	0	6,6,6	1.38	1 (16%)
3	SSN	D	502	1	6,6,6	0.88	0	6,6,6	1.56	1 (16%)
2	NAD	С	501	-	42,48,48	1.17	5 (11%)	50,73,73	1.28	5 (10%)
2	NAD	А	501	-	42,48,48	1.15	5 (11%)	50,73,73	1.32	6 (12%)

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).

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	NAD	D	501	-	-	3/26/62/62	0/5/5/5
2	NAD	В	501	-	-	5/26/62/62	0/5/5/5
3	SSN	А	502	1	-	1/3/4/4	-
3	SSN	С	502	1	-	0/3/4/4	-
3	SSN	В	502	1	-	0/3/4/4	-
3	SSN	D	502	1	-	0/3/4/4	-
2	NAD	С	501	-	-	5/26/62/62	0/5/5/5
2	NAD	А	501	-	-	5/26/62/62	0/5/5/5

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$\operatorname{Observed}(\operatorname{\AA})$	$\mathrm{Ideal}(\mathrm{\AA})$
2	С	501	NAD	O4D-C1D	2.89	1.45	1.41
2	С	501	NAD	O4B-C1B	2.86	1.45	1.41
2	А	501	NAD	O4D-C1D	2.83	1.45	1.41
2	D	501	NAD	O4D-C1D	2.79	1.45	1.41
2	В	501	NAD	O4B-C1B	2.75	1.44	1.41

The worst 5 of 26 bond angle outliers are listed below:

Mo	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	В	501	NAD	PN-O3-PA	-4.45	117.56	132.83

Continued on next page...



Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	А	501	NAD	PN-O3-PA	-4.44	117.60	132.83
2	D	501	NAD	PN-O3-PA	-4.34	117.95	132.83
2	С	501	NAD	PN-O3-PA	-4.15	118.60	132.83
2	С	501	NAD	N3A-C2A-N1A	-4.04	122.37	128.68

Continued from previous page...

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

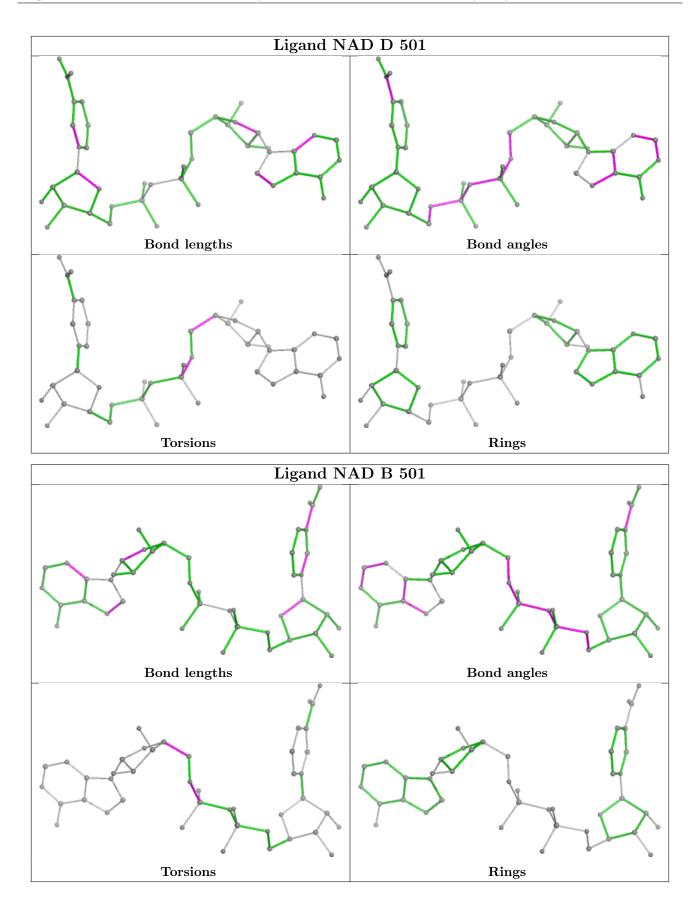
Mol	Chain	Res	Type	Atoms
2	А	501	NAD	С5В-О5В-РА-О1А
2	В	501	NAD	C5B-O5B-PA-O1A
2	С	501	NAD	C5B-O5B-PA-O1A
2	А	501	NAD	O4B-C4B-C5B-O5B
2	С	501	NAD	O4B-C4B-C5B-O5B

There are no ring outliers.

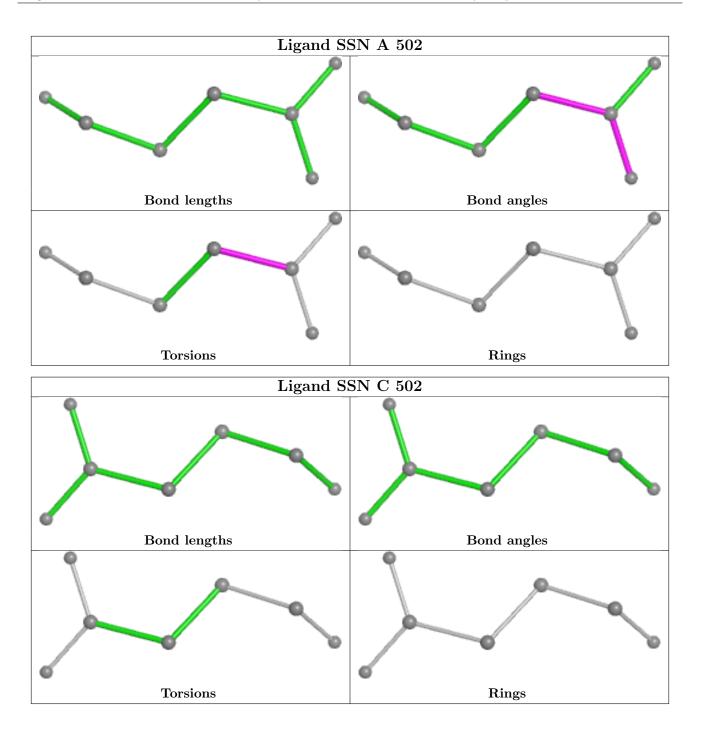
No monomer is involved in short contacts.

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 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.

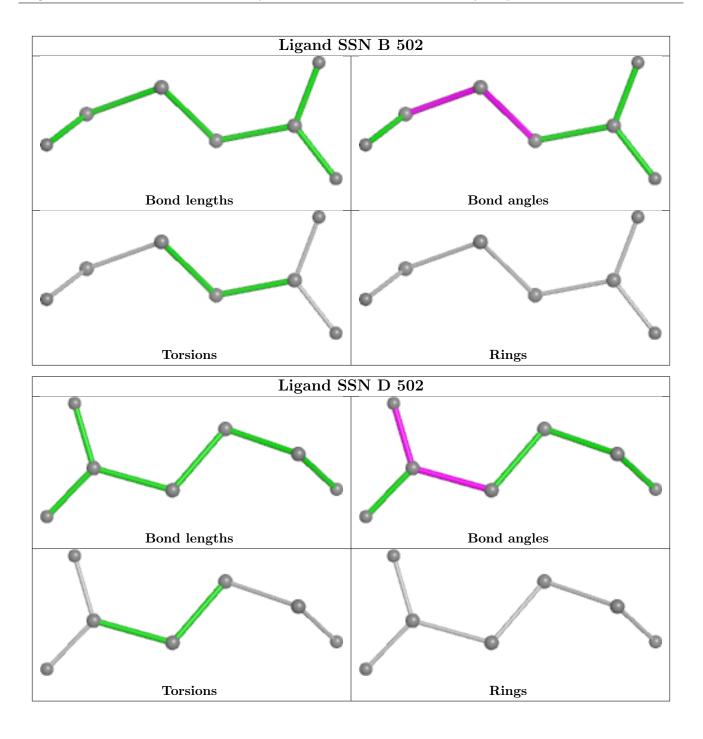




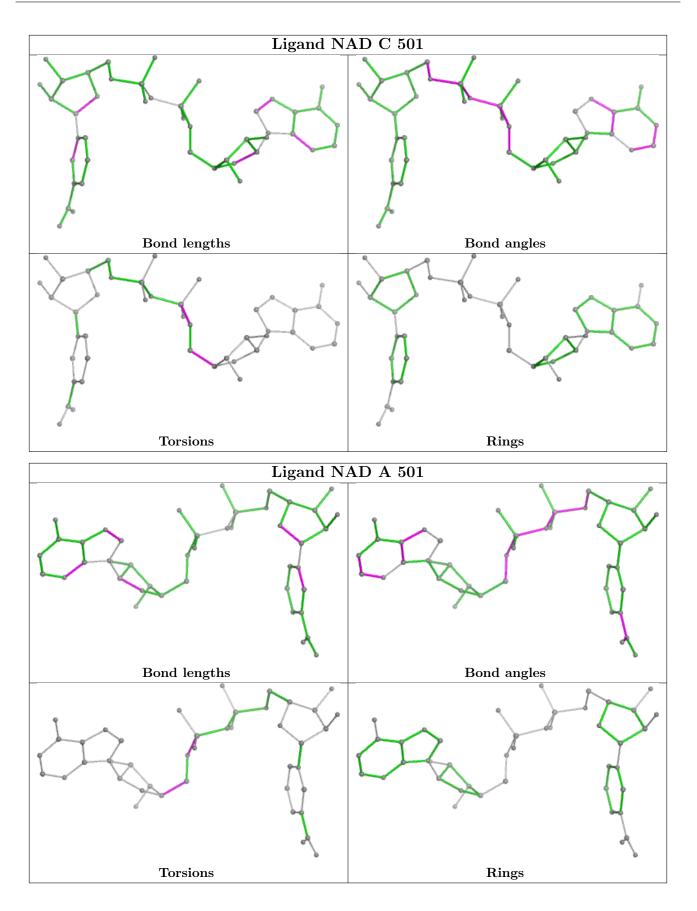














### 5.7 Other polymers (i)

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

Mol	Chain	Analysed	<RSRZ $>$	#RSRZ>2	$OWAB(Å^2)$	Q<0.9
1	А	460/462~(99%)	-0.08	4 (0%) 81 80	22, 29, 42, 76	0
1	В	460/462~(99%)	0.08	12 (2%) 57 55	24, 31, 47, 88	0
1	С	460/462~(99%)	-0.01	6 (1%) 74 74	23, 29, 44, 75	0
1	D	460/462~(99%)	-0.08	4 (0%) 81 80	23, 29, 40, 73	0
All	All	1840/1848~(99%)	-0.02	26 (1%) 73 72	22, 30, 43, 88	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	$\mathbf{Res}$	Type	RSRZ
1	С	3	ILE	3.0
1	В	406	ARG	2.9
1	В	407	GLN	2.9
1	В	403	THR	2.8
1	В	3	ILE	2.8

#### 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.

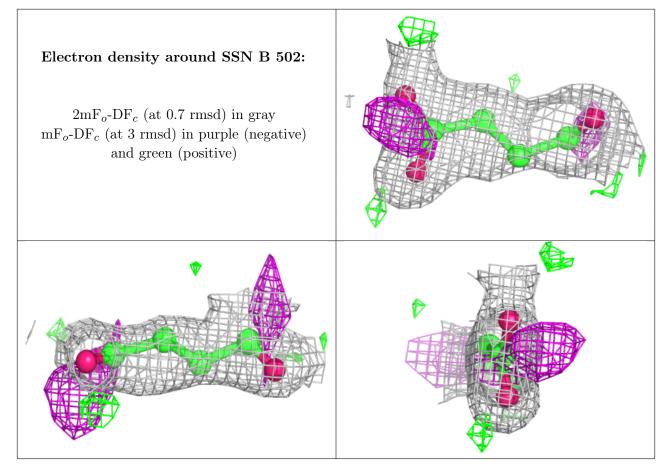
#### 6.4 Ligands (i)

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,  $95^{th}$  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.

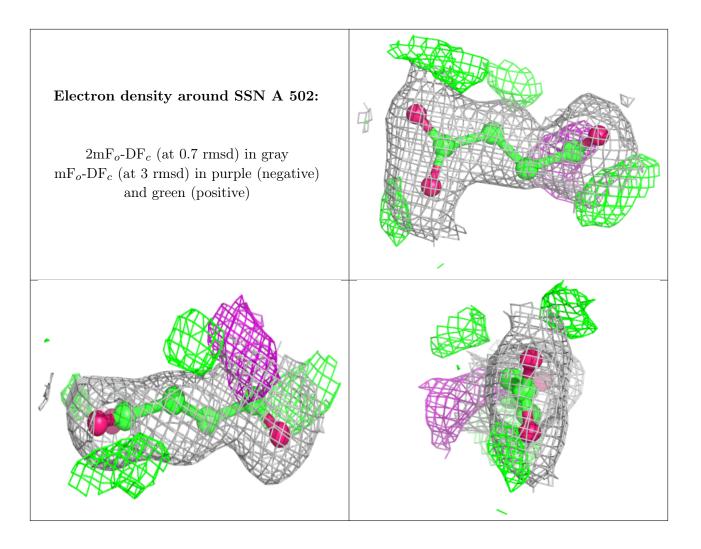


Mol	Type	Chain	Res	Atoms	RSCC	$\mathbf{RSR}$	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q < 0.9
3	SSN	В	502	7/7	0.77	0.13	37,39,43,44	0
3	SSN	А	502	7/7	0.82	0.11	34,36,38,39	0
3	SSN	С	502	7/7	0.83	0.11	32,34,37,40	0
3	SSN	D	502	7/7	0.83	0.10	32,36,39,39	0
2	NAD	В	501	44/44	0.92	0.08	27,33,38,39	0
2	NAD	D	501	44/44	0.92	0.08	27,31,34,35	0
2	NAD	С	501	44/44	0.93	0.07	27,30,34,35	0
2	NAD	А	501	44/44	0.93	0.08	23,30,34,36	0

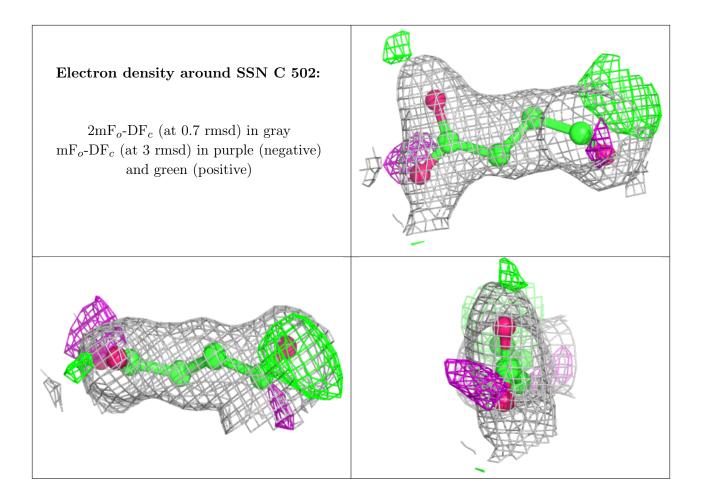
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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



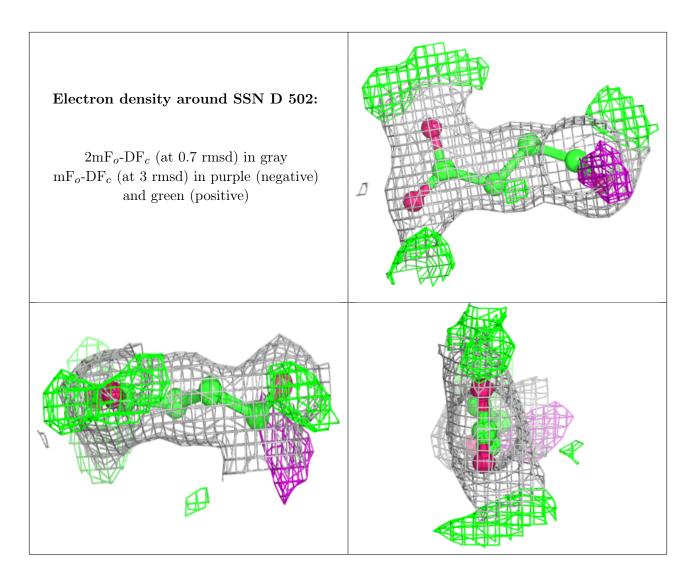




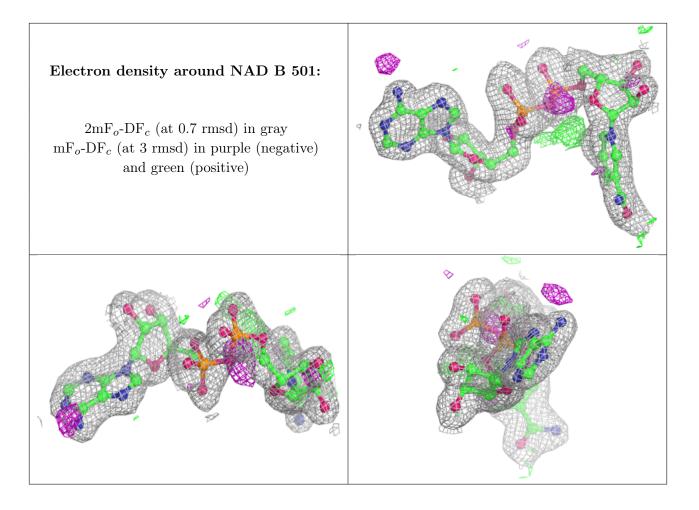




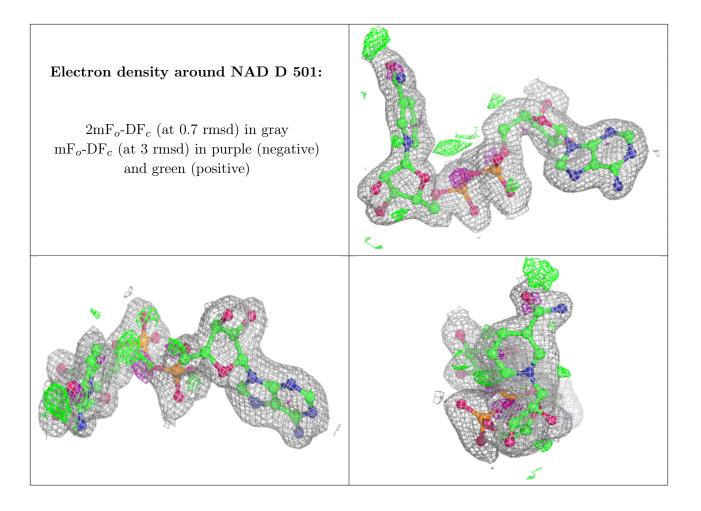




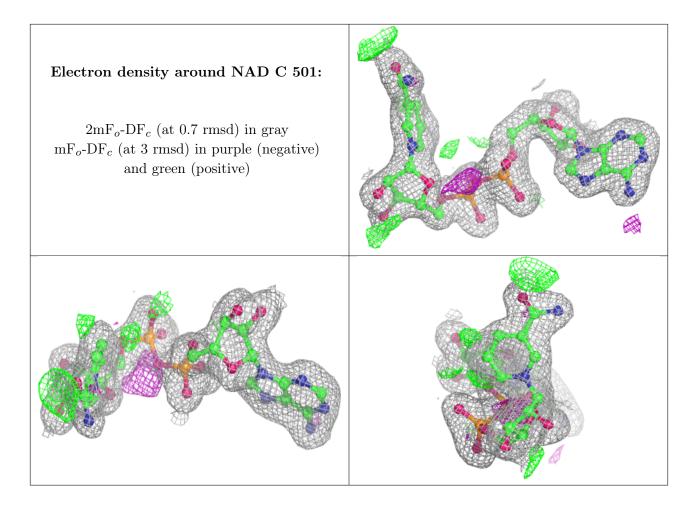




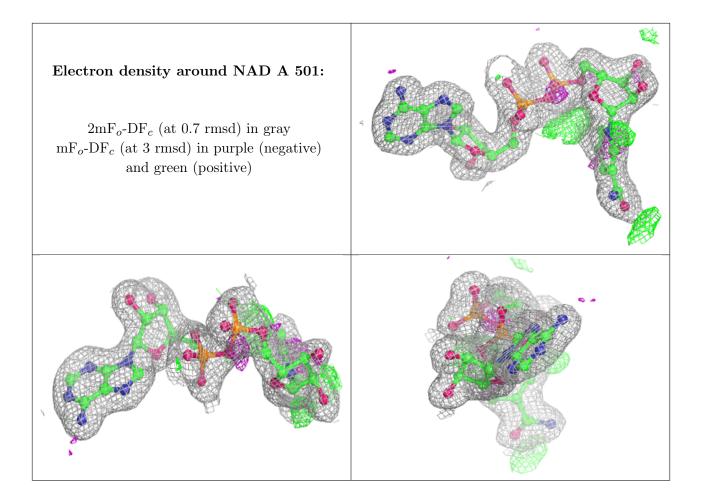












### 6.5 Other polymers (i)

There are no such residues in this entry.

