



Full wwPDB X-ray Structure Validation Report ⓘ

May 25, 2020 – 07:00 am BST

PDB ID : 5MSV
Title : Structure of the phosphopantetheine modified PCP-R didomain of carboxylic acid reductase (CAR) in complex with NADP
Authors : Gahloth, D.; Leys, D.
Deposited on : 2017-01-05
Resolution : 2.34 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 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 : 2.11
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) : 2.11

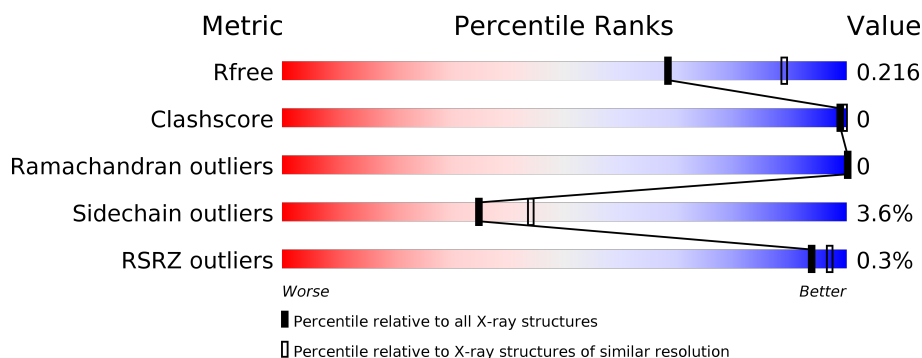
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.34 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	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 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	1188	<div> <div style="width: 41%; background-color: green;"></div> <div style="width: 57%; background-color: grey;"></div> </div>
1	B	1188	<div> <div style="width: 41%; background-color: green;"></div> <div style="width: 57%; background-color: grey;"></div> </div>
1	C	1188	<div> <div style="width: 41%; background-color: green;"></div> <div style="width: 57%; background-color: grey;"></div> </div>
1	D	1188	<div> <div style="width: 42%; background-color: green;"></div> <div style="width: 57%; background-color: grey;"></div> </div>

2 Entry composition [i](#)

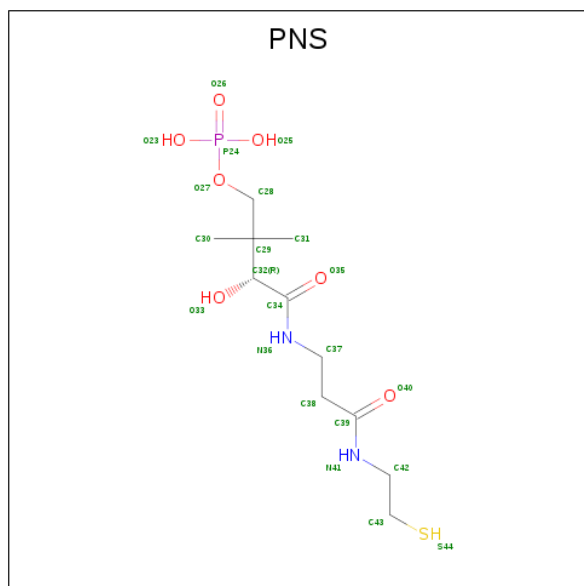
There are 4 unique types of molecules in this entry. The entry contains 16781 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Thioester reductase domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	514	Total	C	N	O	S	0	0	0
			3932	2498	684	749	1			
1	B	514	Total	C	N	O	S	0	0	0
			3938	2500	687	750	1			
1	C	514	Total	C	N	O	S	0	0	0
			3938	2501	687	749	1			
1	D	515	Total	C	N	O	S	0	0	0
			3939	2500	688	750	1			

- Molecule 2 is 4'-PHOSPHOPANTETHEINE (three-letter code: PNS) (formula: C₁₁H₂₃N₂O₇PS).



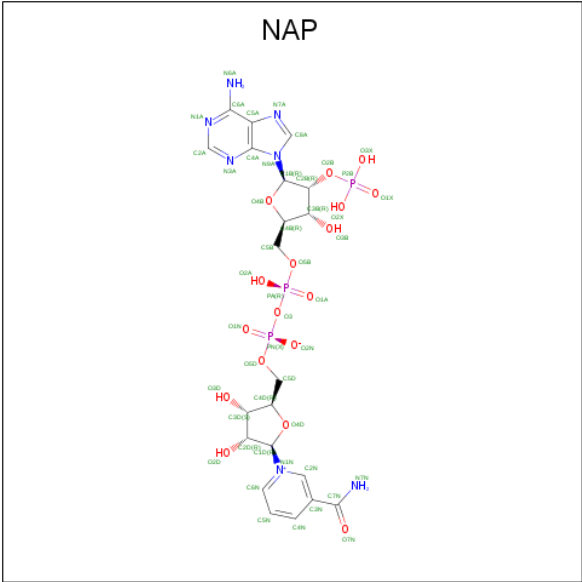
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P S	0	0
			21	11	2	6	1 1		
2	B	1	Total	C	N	O	P S	0	0
			21	11	2	6	1 1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	C	1	Total	C	N	O	P	S	0	0
			21	11	2	6	1	1		
2	D	1	Total	C	N	O	P	S	0	0
			21	11	2	6	1	1		

- Molecule 3 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: C₂₁H₂₈N₇O₁₇P₃).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			40	15	6	16	3		
3	B	1	Total	C	N	O	P	0	0
			40	15	6	16	3		
3	C	1	Total	C	N	O	P	0	0
			40	15	6	16	3		
3	D	1	Total	C	N	O	P	0	0
			40	15	6	16	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	223	Total	O	0	0
			223	223		
4	B	195	Total	O	0	0
			195	195		
4	C	164	Total	O	0	0
			164	164		

Continued on next page...

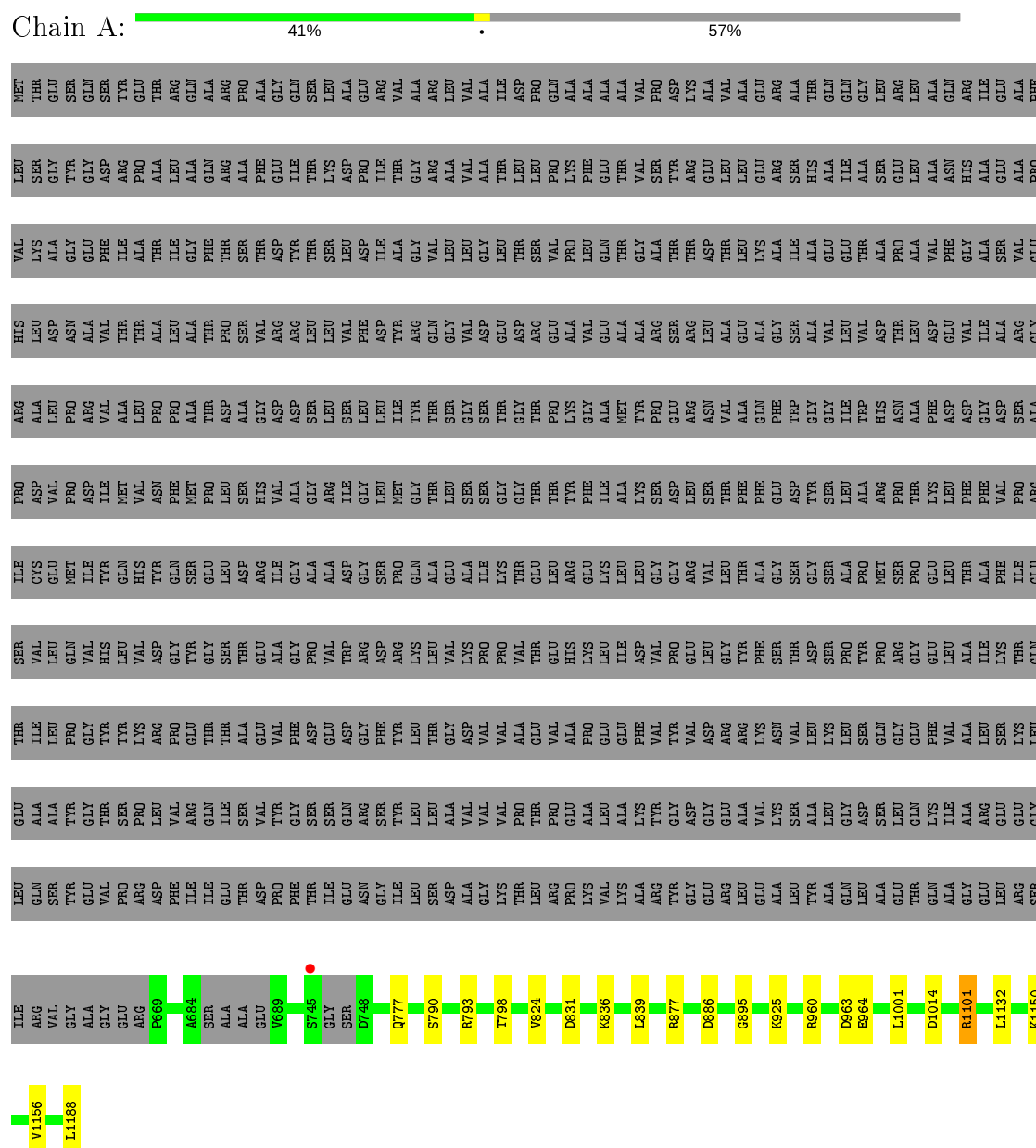
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	208	Total 208	O 208	0	0

3 Residue-property plots

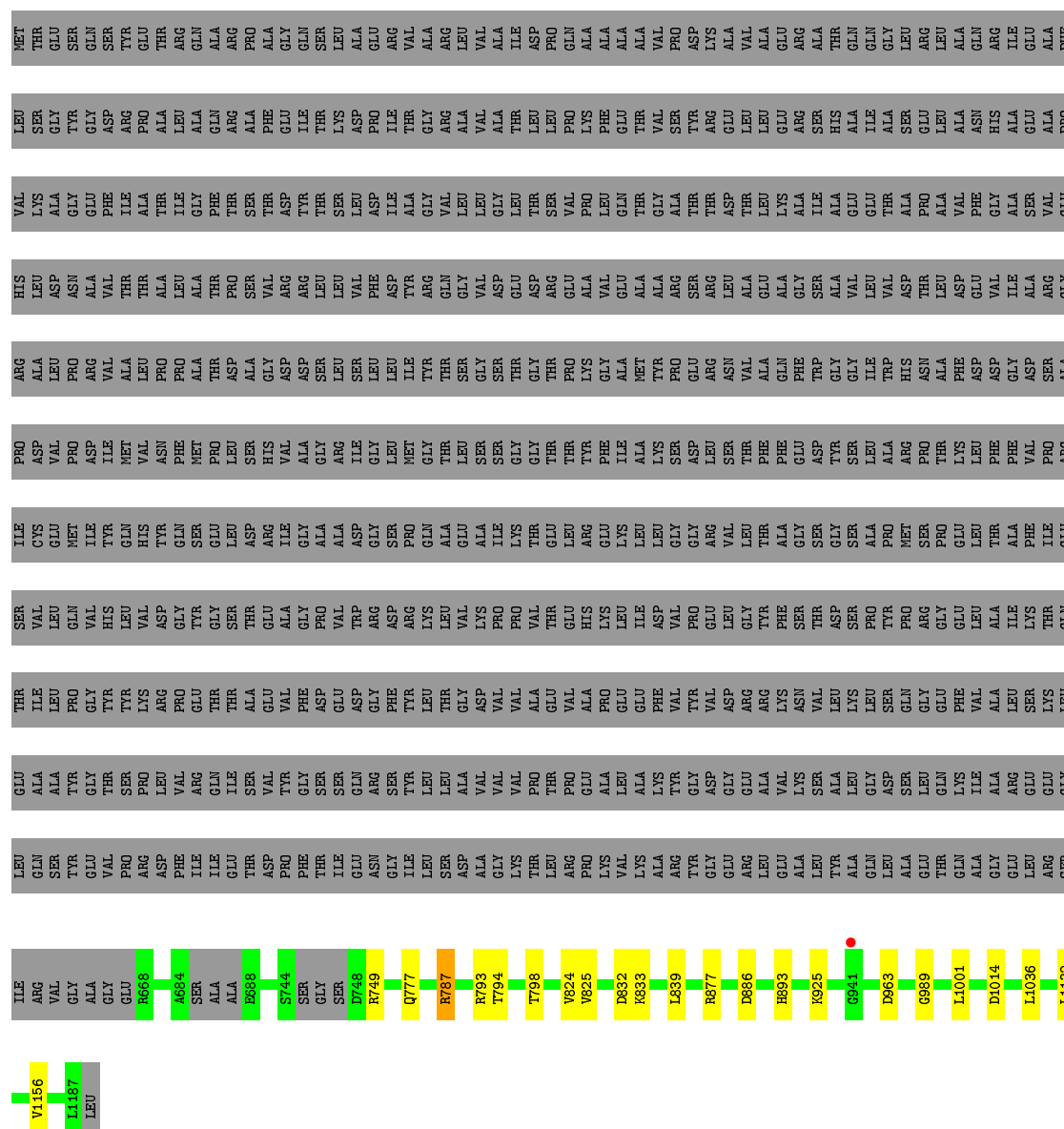
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Thioester reductase domain-containing protein



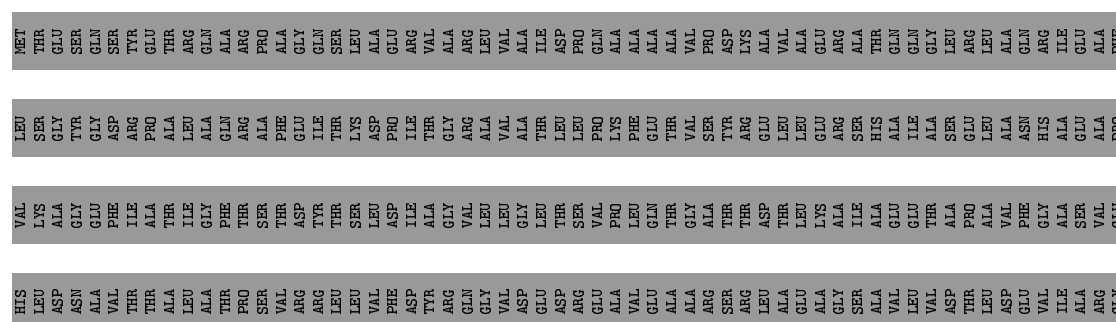
- Molecule 1: Thioester reductase domain-containing protein

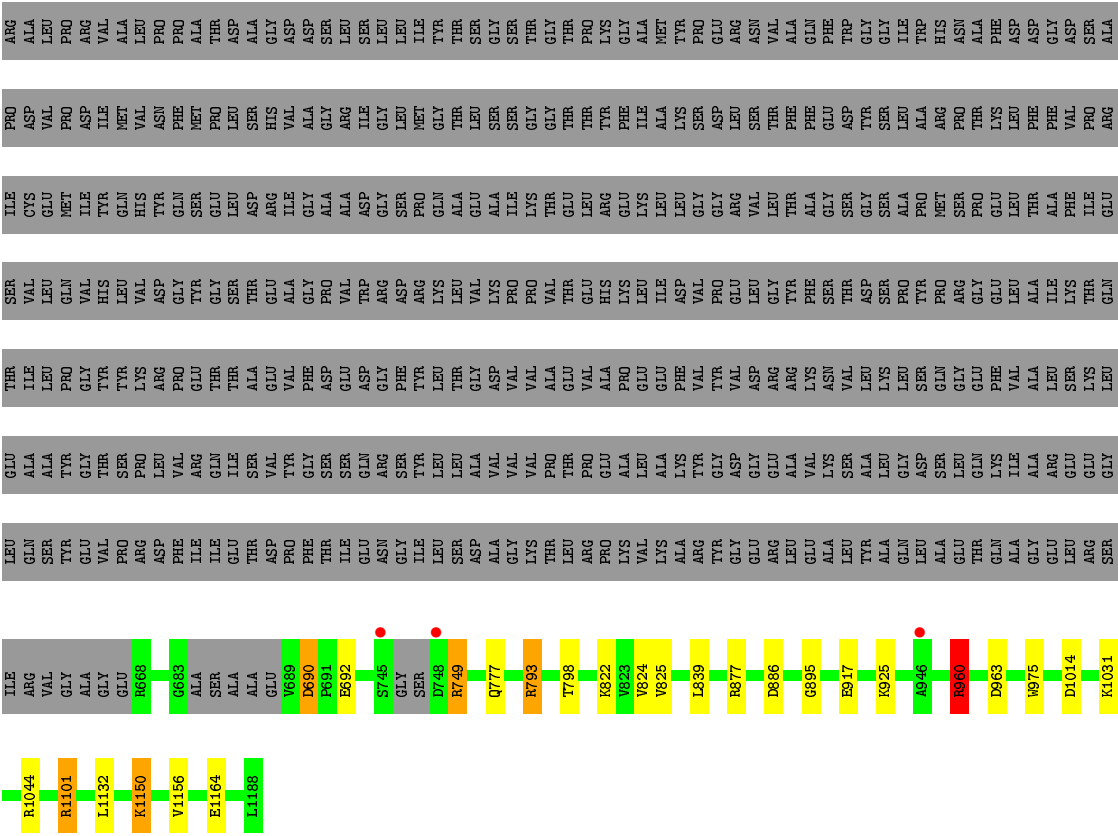
Chain B:



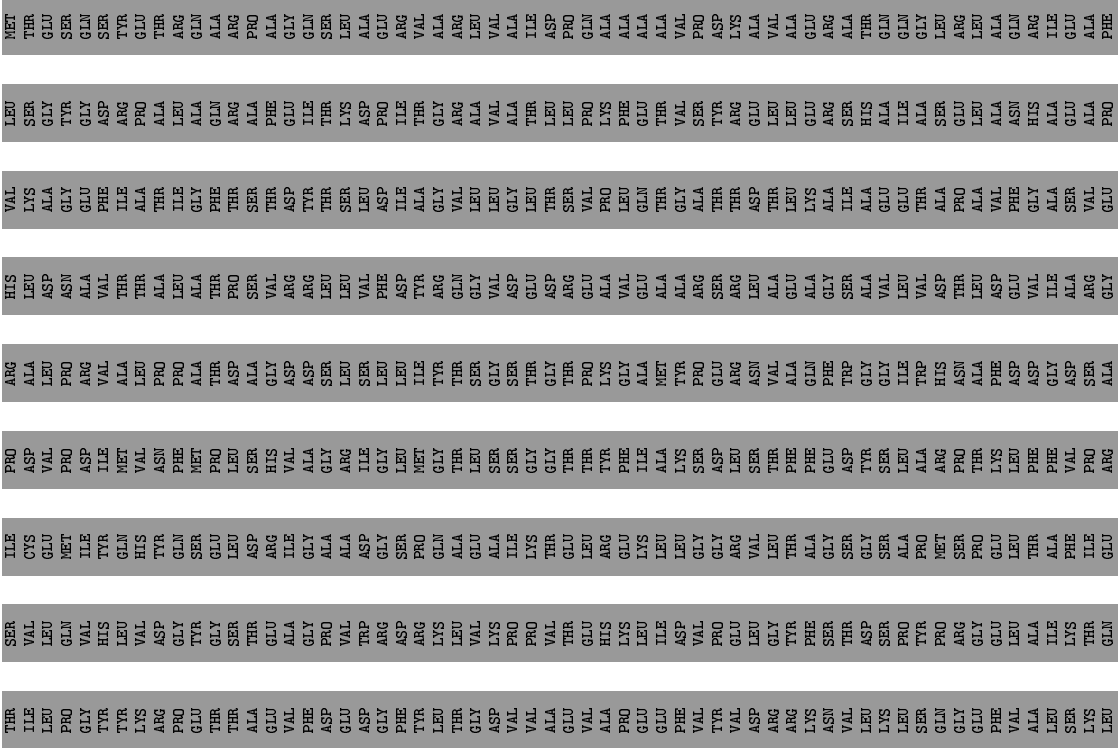
- Molecule 1: Thioester reductase domain-containing protein

Chain C:





● Molecule 1: Thioester reductase domain-containing protein



ILE	ARG	VAL	GLY	GLY	GLY	R668	A684	SER	ALA	ALA	GLU	V689	L732	S744	SER	G746	S747	D748	E771	Q777	R793	T798	V824	V825	L839	D886	G895	K925	D963	D1014	L1036	R1042	L1132	V1156	L1187	LEU																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
LEU	GLN	SER	TYR	GLY	VAL	PRO	ARG	ASP	PHE	ILE	ILE	THR	PRO	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	TYR	ALA	GLN	LEU	GLY	LEU	ASP	GLY	GLY	GLU	ALA	VAL	LYS	SER	ALA	LEU	LYS	ALA	TYR	THR	GLU	LEU	ASP	GLY	ASP	GLY	GLY	GLU	VAL	LYS	ALA	TYR	THR	GLN	LEU	SER	ALA	LEU	VAL	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	SER	GLY	TYR	THR	GLY	GLY	TYR	ALA	GLU	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA	ALA	GLY	LYS	THR	LEU	ARG	PRO	LYS	VAL	LYS	ALA	LEU	GLY	ASP	GLY	GLY	GLU	ALA	VAL	VAL	VAL	ALA	ALA	GLN	SER	GLN	SER	GLY	TYR	THR	GLY	VAL	PRO	ARG	ASP	PHE	THR	ILE	ASN	GLY	ILE	LEU	SER	ASP	ALA

4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α , β , γ	92.11Å 92.11Å 363.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	200.00 – 2.34 90.88 – 2.34	Depositor EDS
% Data completeness (in resolution range)	95.2 (200.00-2.34) 95.3 (90.88-2.34)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.03 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.191 , 0.207 0.202 , 0.216	Depositor DCC
R_{free} test set	5923 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	23.4	Xtriage
Anisotropy	0.030	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 27.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.068 for h,-k,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	16781	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.28% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: PNS, NAP

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	0.45	0/4018	0.73	8/5467 (0.1%)
1	B	0.44	0/4024	0.70	4/5475 (0.1%)
1	C	0.44	1/4024 (0.0%)	0.70	9/5475 (0.2%)
1	D	0.43	0/4025	0.66	1/5476 (0.0%)
All	All	0.44	1/16091 (0.0%)	0.70	22/21893 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	1164	GLU	CG-CD	5.56	1.60	1.51

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	963	ASP	CB-CG-OD2	12.93	129.94	118.30
1	A	963	ASP	CB-CG-OD1	-11.32	108.11	118.30
1	A	1101	ARG	CG-CD-NE	10.99	134.87	111.80
1	C	749	ARG	CG-CD-NE	8.96	130.61	111.80
1	B	749	ARG	NE-CZ-NH2	-6.90	116.85	120.30
1	C	793	ARG	NE-CZ-NH1	6.59	123.59	120.30
1	A	1101	ARG	NE-CZ-NH2	-6.17	117.21	120.30
1	C	963	ASP	CB-CG-OD1	5.91	123.62	118.30
1	A	1101	ARG	CD-NE-CZ	5.89	131.84	123.60
1	A	877	ARG	NE-CZ-NH1	5.88	123.24	120.30
1	B	963	ASP	CB-CG-OD1	5.88	123.59	118.30
1	C	877	ARG	NE-CZ-NH1	5.83	123.21	120.30
1	D	963	ASP	CB-CG-OD1	5.75	123.48	118.30
1	A	960	ARG	NE-CZ-NH2	5.62	123.11	120.30
1	A	877	ARG	NE-CZ-NH2	-5.58	117.51	120.30
1	C	1150	LYS	CD-CE-NZ	5.54	124.43	111.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	877	ARG	NE-CZ-NH2	-5.48	117.56	120.30
1	B	877	ARG	NE-CZ-NH1	5.40	123.00	120.30
1	C	1044	ARG	NE-CZ-NH1	5.22	122.91	120.30
1	C	1101	ARG	NE-CZ-NH2	-5.22	117.69	120.30
1	C	960	ARG	NE-CZ-NH2	-5.15	117.73	120.30
1	C	877	ARG	NE-CZ-NH2	-5.04	117.78	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3932	0	3879	3	0
1	B	3938	0	3881	4	0
1	C	3938	0	3886	6	0
1	D	3939	0	3883	2	0
2	A	21	0	21	0	0
2	B	21	0	21	0	0
2	C	21	0	21	0	0
2	D	21	0	21	0	0
3	A	40	0	19	0	0
3	B	40	0	19	0	0
3	C	40	0	19	0	0
3	D	40	0	19	0	0
4	A	223	0	0	0	0
4	B	195	0	0	0	0
4	C	164	0	0	0	0
4	D	208	0	0	0	0
All	All	16781	0	15689	13	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 0.

All (13) 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:960:ARG:HD2	1:C:975:TRP:CD2	2.34	0.62
1:C:1031:LYS:HA	1:C:1101:ARG:HD3	1.84	0.58
1:A:836:LYS:HD2	1:B:833:LYS:HE2	1.87	0.56
1:B:787:ARG:NH2	1:B:989:GLY:O	2.40	0.54
1:B:798:THR:OG1	1:B:893:HIS:HA	2.11	0.51
1:C:749:ARG:NH2	1:C:917:GLU:OE2	2.45	0.49
1:D:798:THR:HG22	1:D:895:GLY:HA3	1.95	0.48
1:C:960:ARG:HD2	1:C:975:TRP:CE2	2.49	0.48
1:A:798:THR:HG22	1:A:895:GLY:HA3	1.96	0.47
1:C:690:ASP:OD1	1:C:692:GLU:N	2.47	0.47
1:A:831:ASP:HB2	1:B:832:ASP:OD2	2.15	0.47
1:C:798:THR:HG22	1:C:895:GLY:HA3	1.97	0.45
1:D:689:VAL:HG12	1:D:732:LEU:CD2	2.49	0.42

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	508/1188 (43%)	496 (98%)	12 (2%)	0	100	100
1	B	508/1188 (43%)	496 (98%)	12 (2%)	0	100	100
1	C	508/1188 (43%)	496 (98%)	12 (2%)	0	100	100
1	D	509/1188 (43%)	496 (97%)	13 (3%)	0	100	100
All	All	2033/4752 (43%)	1984 (98%)	49 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	411/947 (43%)	396 (96%)	15 (4%)	35	44
1	B	411/947 (43%)	397 (97%)	14 (3%)	37	46
1	C	412/947 (44%)	398 (97%)	14 (3%)	37	46
1	D	411/947 (43%)	395 (96%)	16 (4%)	32	41
All	All	1645/3788 (43%)	1586 (96%)	59 (4%)	35	44

All (59) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	777	GLN
1	A	790	SER
1	A	793	ARG
1	A	824	VAL
1	A	839	LEU
1	A	886	ASP
1	A	925	LYS
1	A	964	GLU
1	A	1001	LEU
1	A	1014	ASP
1	A	1101	ARG
1	A	1132	LEU
1	A	1150	LYS
1	A	1156	VAL
1	A	1188	LEU
1	B	777	GLN
1	B	787	ARG
1	B	793	ARG
1	B	794	THR
1	B	824	VAL
1	B	825	VAL
1	B	839	LEU
1	B	886	ASP
1	B	925	LYS
1	B	1001	LEU
1	B	1014	ASP
1	B	1036	LEU
1	B	1132	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	1156	VAL
1	C	690	ASP
1	C	777	GLN
1	C	793	ARG
1	C	822	LYS
1	C	824	VAL
1	C	825	VAL
1	C	839	LEU
1	C	886	ASP
1	C	925	LYS
1	C	960	ARG
1	C	1014	ASP
1	C	1132	LEU
1	C	1150	LYS
1	C	1156	VAL
1	D	689	VAL
1	D	732	LEU
1	D	748	ASP
1	D	771	GLU
1	D	777	GLN
1	D	793	ARG
1	D	824	VAL
1	D	825	VAL
1	D	839	LEU
1	D	886	ASP
1	D	925	LYS
1	D	1014	ASP
1	D	1036	LEU
1	D	1042	ARG
1	D	1132	LEU
1	D	1156	VAL

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

Mol	Chain	Res	Type
1	B	777	GLN
1	C	777	GLN

5.3.3 RNA ⓘ

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 carbohydrates 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 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	NAP	A	1202	-	36,43,52	0.84	1 (2%)	44,67,80	1.19	5 (11%)
3	NAP	C	1202	-	36,43,52	0.81	1 (2%)	44,67,80	1.21	5 (11%)
2	PNS	C	1201	1	13,20,21	0.60	0	18,26,29	0.80	0
2	PNS	A	1201	1	13,20,21	0.53	0	18,26,29	0.94	0
2	PNS	B	1201	1	13,20,21	0.51	0	18,26,29	0.84	0
2	PNS	D	1201	1	13,20,21	0.52	0	18,26,29	0.83	0
3	NAP	D	1202	-	36,43,52	0.88	2 (5%)	44,67,80	1.29	6 (13%)
3	NAP	B	1202	-	36,43,52	0.82	1 (2%)	44,67,80	1.16	2 (4%)

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	NAP	A	1202	-	-	7/23/59/67	0/4/4/5
3	NAP	C	1202	-	-	5/23/59/67	0/4/4/5
2	PNS	C	1201	1	-	0/24/26/27	-
2	PNS	A	1201	1	-	0/24/26/27	-
2	PNS	B	1201	1	-	0/24/26/27	-
2	PNS	D	1201	1	-	0/24/26/27	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAP	D	1202	-	-	5/23/59/67	0/4/4/5
3	NAP	B	1202	-	-	7/23/59/67	0/4/4/5

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	1202	NAP	C5A-C4A	2.26	1.46	1.40
3	A	1202	NAP	C5A-C4A	2.25	1.46	1.40
3	C	1202	NAP	C5A-C4A	2.21	1.46	1.40
3	D	1202	NAP	O4B-C1B	2.09	1.44	1.41
3	B	1202	NAP	C5A-C4A	2.06	1.46	1.40

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1202	NAP	N3A-C2A-N1A	-4.13	122.22	128.68
3	D	1202	NAP	N3A-C2A-N1A	-3.87	122.63	128.68
3	A	1202	NAP	N3A-C2A-N1A	-3.74	122.83	128.68
3	B	1202	NAP	N3A-C2A-N1A	-3.58	123.08	128.68
3	D	1202	NAP	C3B-C2B-C1B	-2.93	97.38	102.89
3	D	1202	NAP	C4A-C5A-N7A	-2.87	106.41	109.40
3	B	1202	NAP	PN-O3-PA	-2.82	123.14	132.83
3	D	1202	NAP	C1B-N9A-C4A	-2.79	121.75	126.64
3	D	1202	NAP	PN-O3-PA	-2.52	124.19	132.83
3	A	1202	NAP	PN-O3-PA	-2.43	124.50	132.83
3	C	1202	NAP	PN-O3-PA	-2.41	124.55	132.83
3	C	1202	NAP	C4A-C5A-N7A	-2.35	106.95	109.40
3	C	1202	NAP	C1B-N9A-C4A	-2.23	122.72	126.64
3	C	1202	NAP	C2A-N1A-C6A	2.22	122.54	118.75
3	A	1202	NAP	C3B-C2B-C1B	-2.16	98.82	102.89
3	A	1202	NAP	C1B-N9A-C4A	-2.10	122.96	126.64
3	A	1202	NAP	C2A-N1A-C6A	2.04	122.24	118.75
3	D	1202	NAP	C2A-N1A-C6A	2.01	122.19	118.75

There are no chirality outliers.

All (24) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1202	NAP	C5D-O5D-PN-O2N
3	C	1202	NAP	C5D-O5D-PN-O2N
3	D	1202	NAP	C5D-O5D-PN-O1N

Continued on next page...

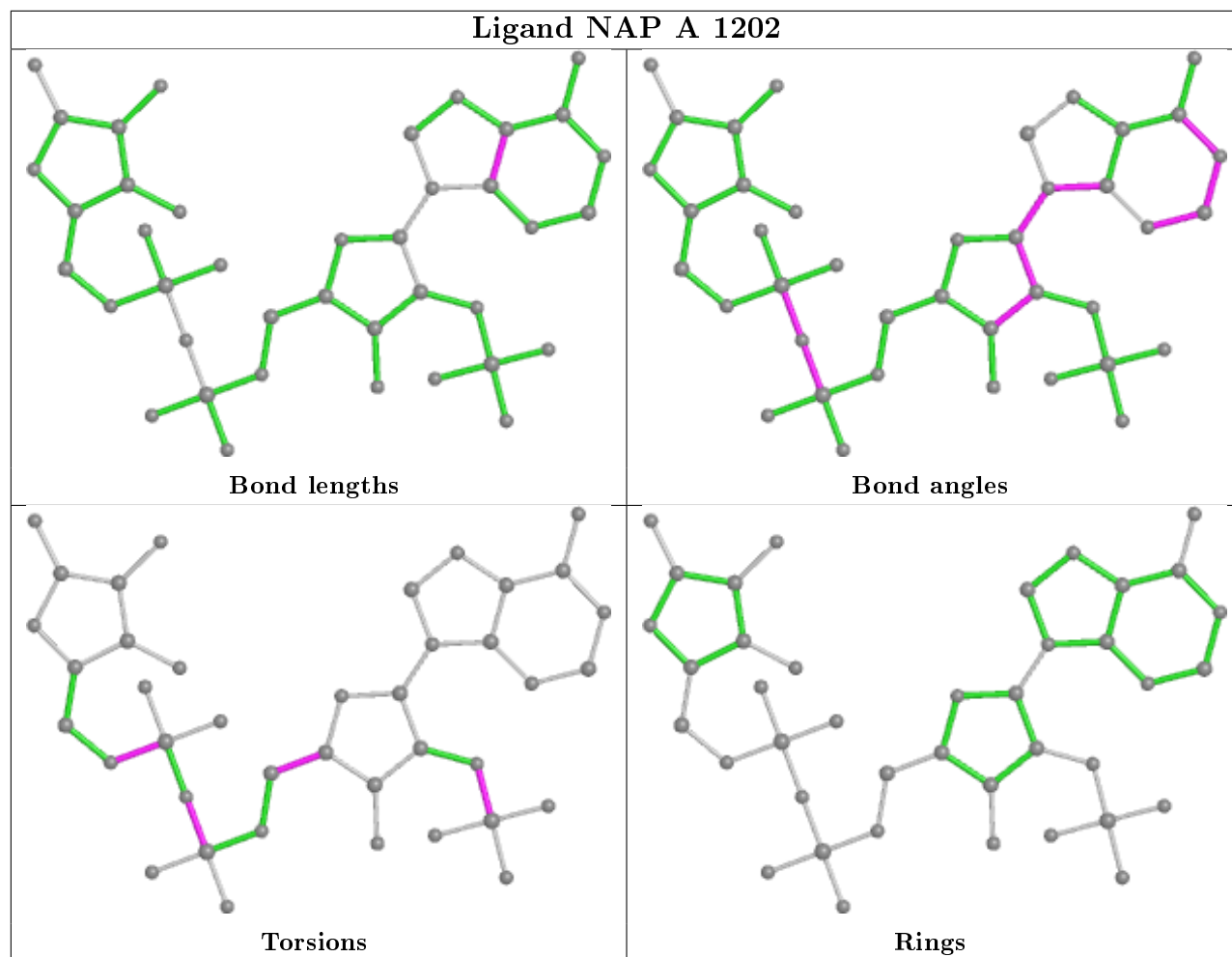
Continued from previous page...

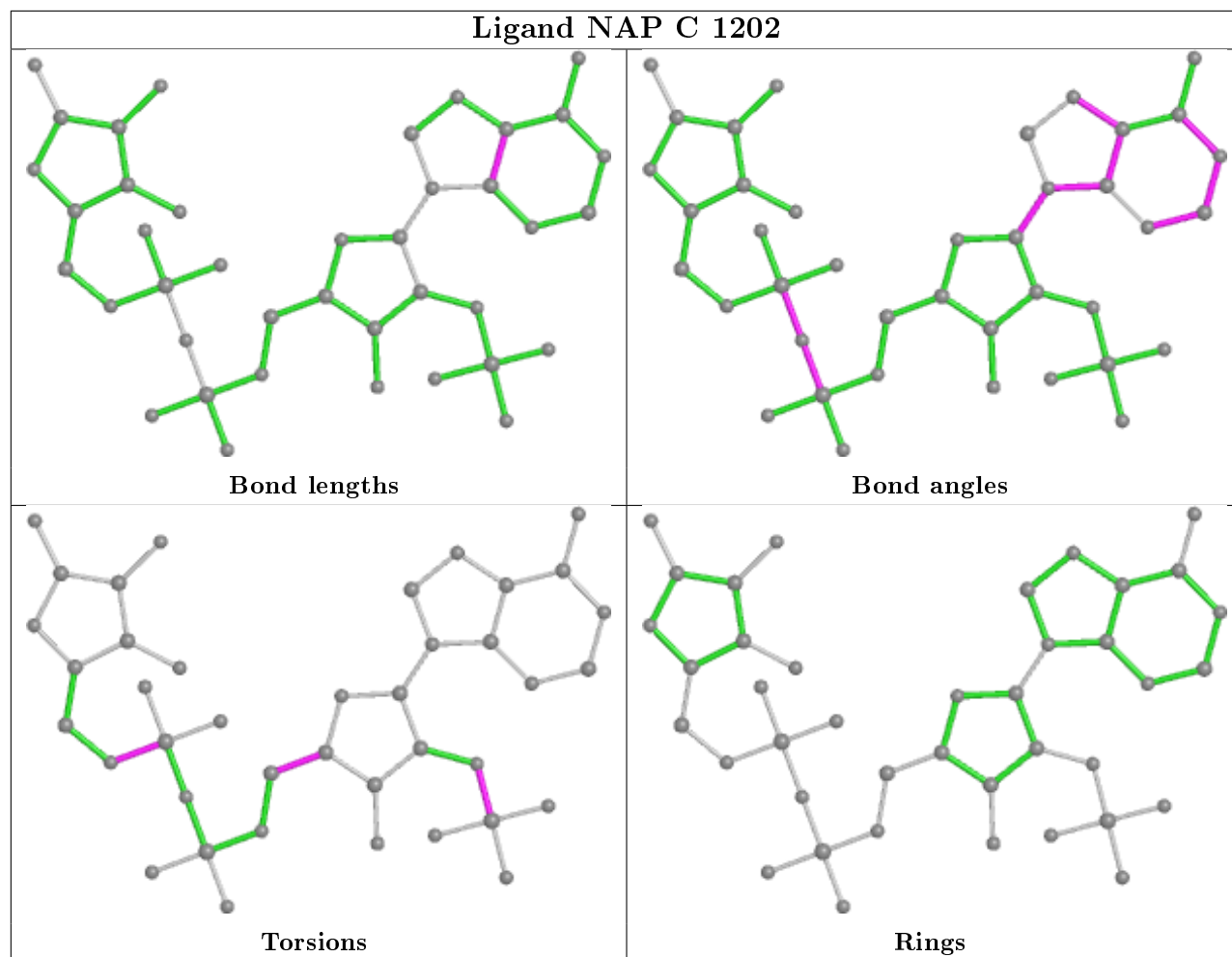
Mol	Chain	Res	Type	Atoms
3	D	1202	NAP	C5D-O5D-PN-O2N
3	B	1202	NAP	C5D-O5D-PN-O2N
3	A	1202	NAP	C2B-O2B-P2B-O3X
3	A	1202	NAP	C5D-O5D-PN-O3
3	C	1202	NAP	C5D-O5D-PN-O3
3	D	1202	NAP	C5D-O5D-PN-O3
3	B	1202	NAP	C5D-O5D-PN-O3
3	A	1202	NAP	C5D-O5D-PN-O1N
3	C	1202	NAP	C5D-O5D-PN-O1N
3	B	1202	NAP	C5D-O5D-PN-O1N
3	B	1202	NAP	PN-O3-PA-O1A
3	A	1202	NAP	PN-O3-PA-O1A
3	B	1202	NAP	PN-O3-PA-O2A
3	C	1202	NAP	C2B-O2B-P2B-O3X
3	D	1202	NAP	C2B-O2B-P2B-O3X
3	B	1202	NAP	C2B-O2B-P2B-O3X
3	C	1202	NAP	O4B-C4B-C5B-O5B
3	D	1202	NAP	O4B-C4B-C5B-O5B
3	A	1202	NAP	PN-O3-PA-O2A
3	A	1202	NAP	O4B-C4B-C5B-O5B
3	B	1202	NAP	O4B-C4B-C5B-O5B

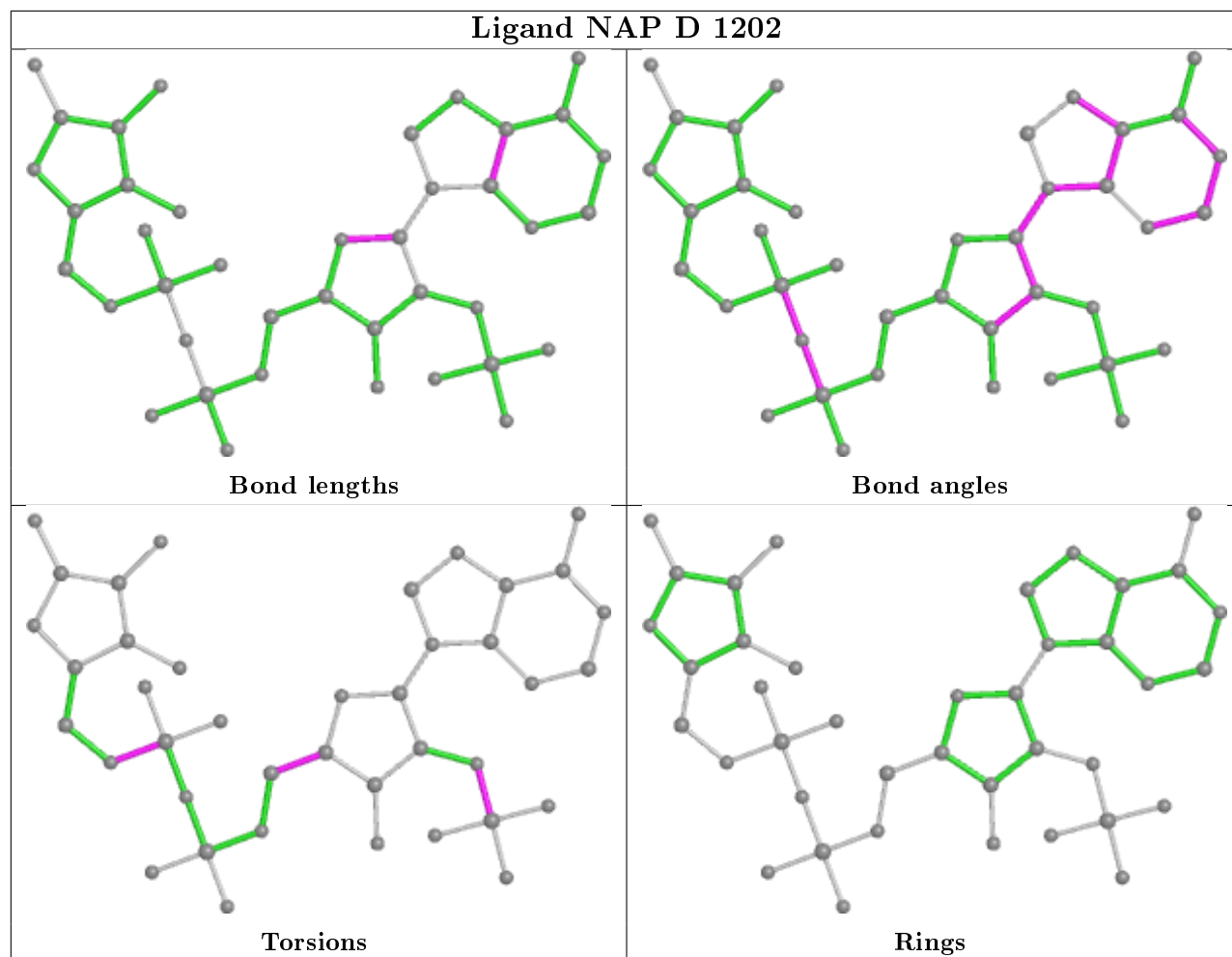
There are no ring outliers.

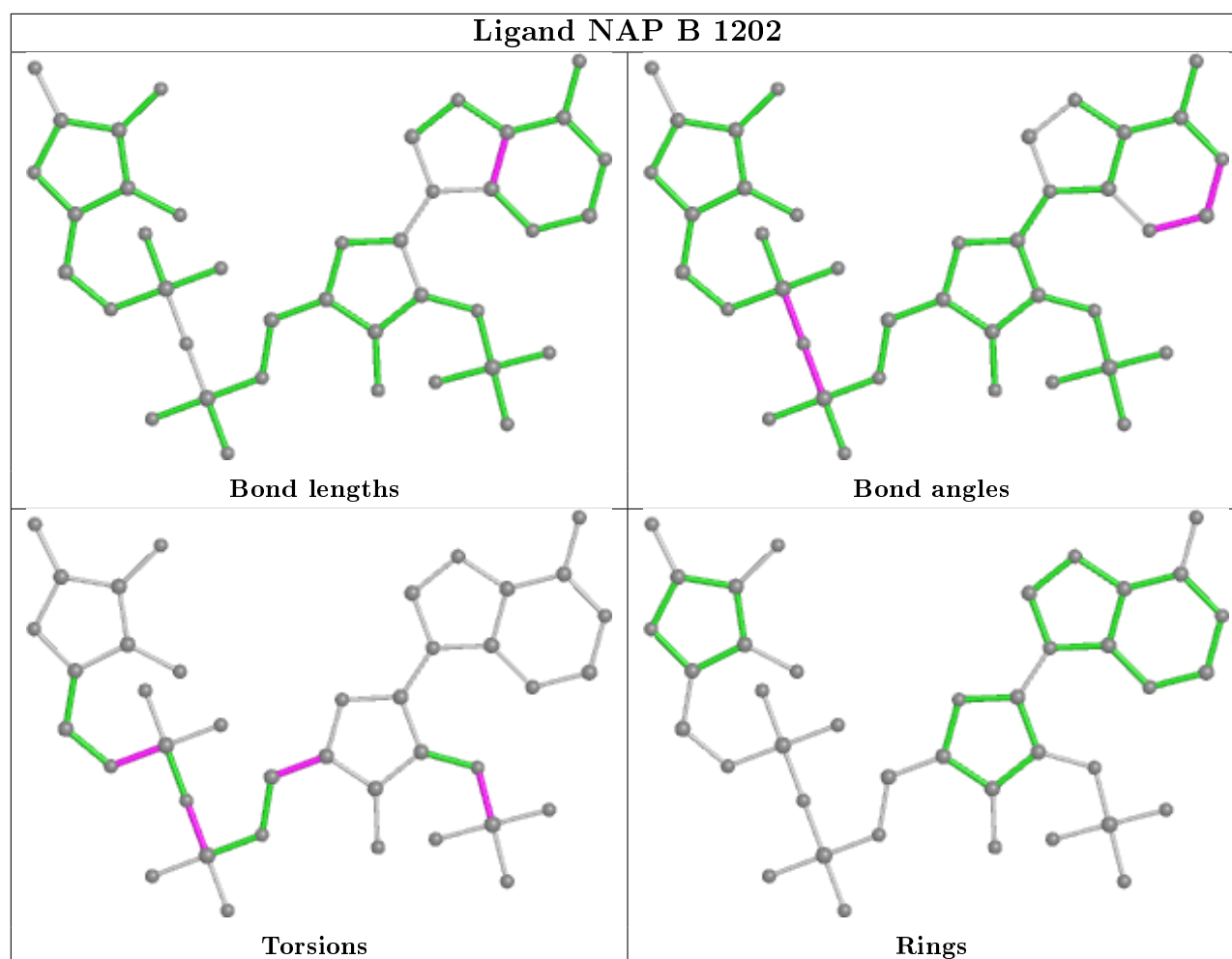
No monomer is involved in short contacts.

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 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, 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	514/1188 (43%)	-0.06	1 (0%) 95 97	15, 40, 71, 106	0
1	B	514/1188 (43%)	0.08	1 (0%) 95 97	18, 44, 76, 107	0
1	C	514/1188 (43%)	0.09	3 (0%) 89 93	21, 47, 76, 100	0
1	D	515/1188 (43%)	-0.01	1 (0%) 95 97	20, 42, 73, 102	0
All	All	2057/4752 (43%)	0.03	6 (0%) 94 97	15, 43, 75, 107	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	946	ALA	3.7
1	C	745	SER	2.3
1	A	745	SER	2.2
1	B	941	GLY	2.1
1	D	746	GLY	2.1
1	C	748	ASP	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 carbohydrates in this entry.

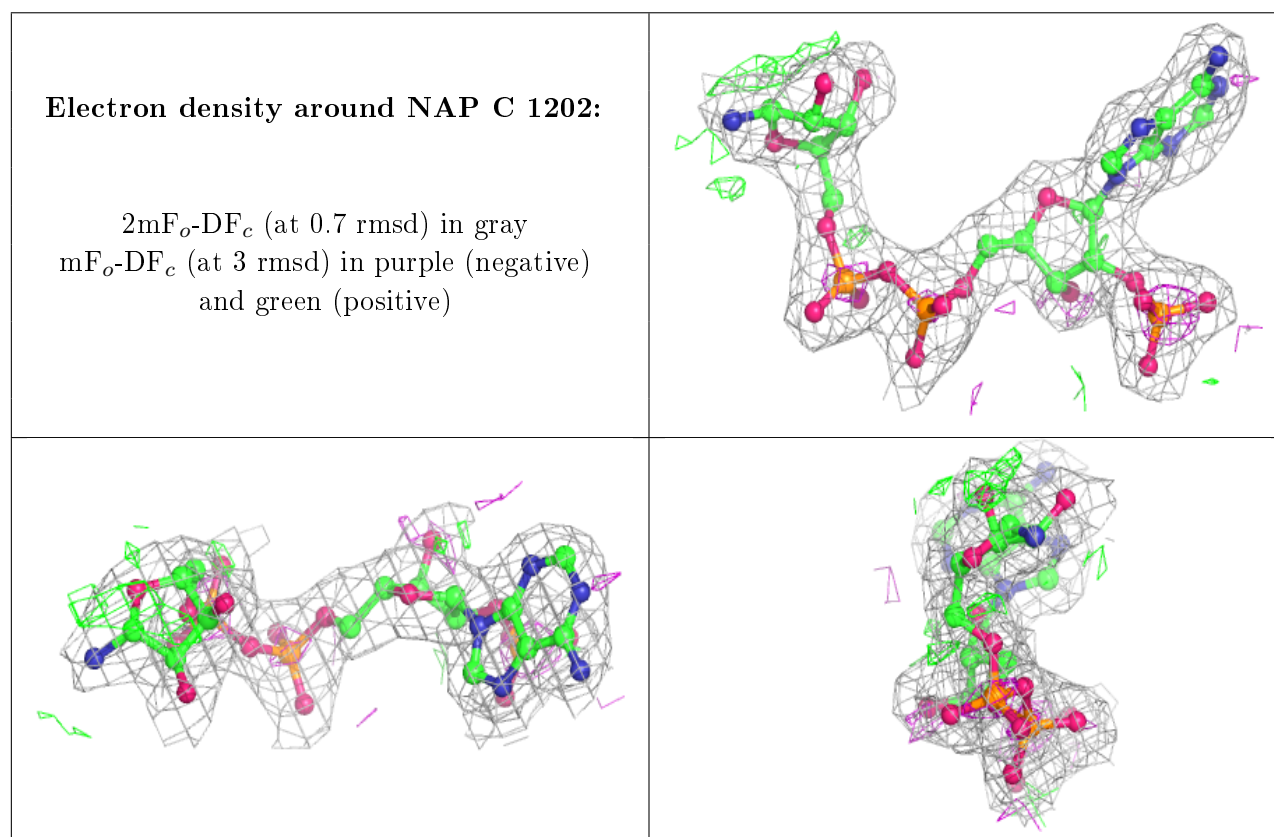
6.4 Ligands [i](#)

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.

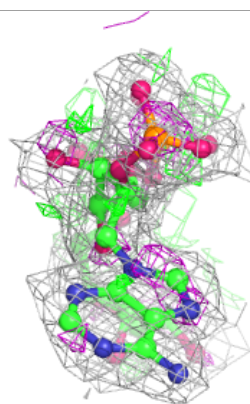
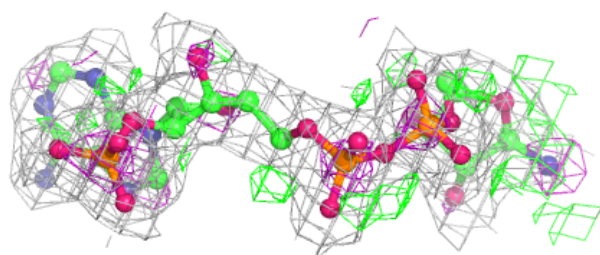
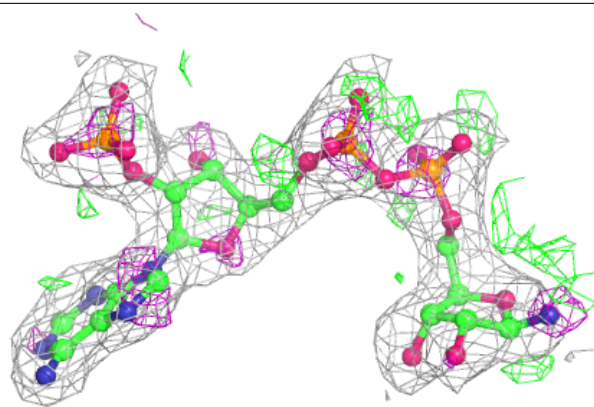
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	PNS	B	1201	21/22	0.95	0.12	21,27,39,57	0
2	PNS	A	1201	21/22	0.97	0.12	19,25,41,60	0
2	PNS	C	1201	21/22	0.97	0.15	22,27,37,58	0
2	PNS	D	1201	21/22	0.97	0.11	20,23,36,55	0
3	NAP	C	1202	40/48	0.98	0.11	16,20,25,27	0
3	NAP	A	1202	40/48	0.99	0.10	10,14,20,23	0
3	NAP	D	1202	40/48	0.99	0.10	14,16,20,20	0
3	NAP	B	1202	40/48	0.99	0.09	12,15,23,26	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.

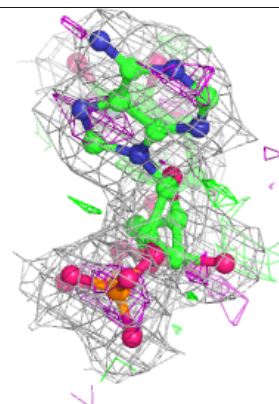
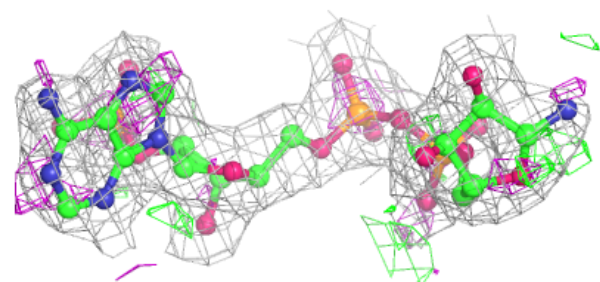
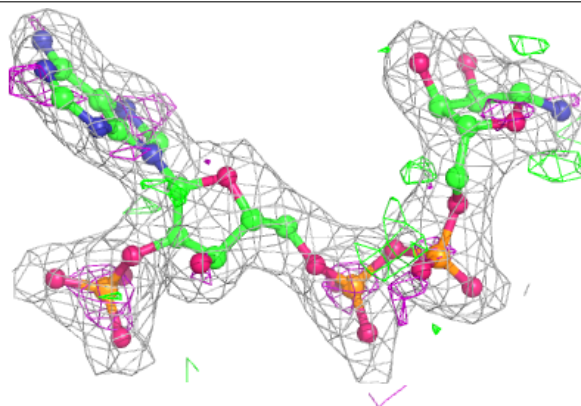


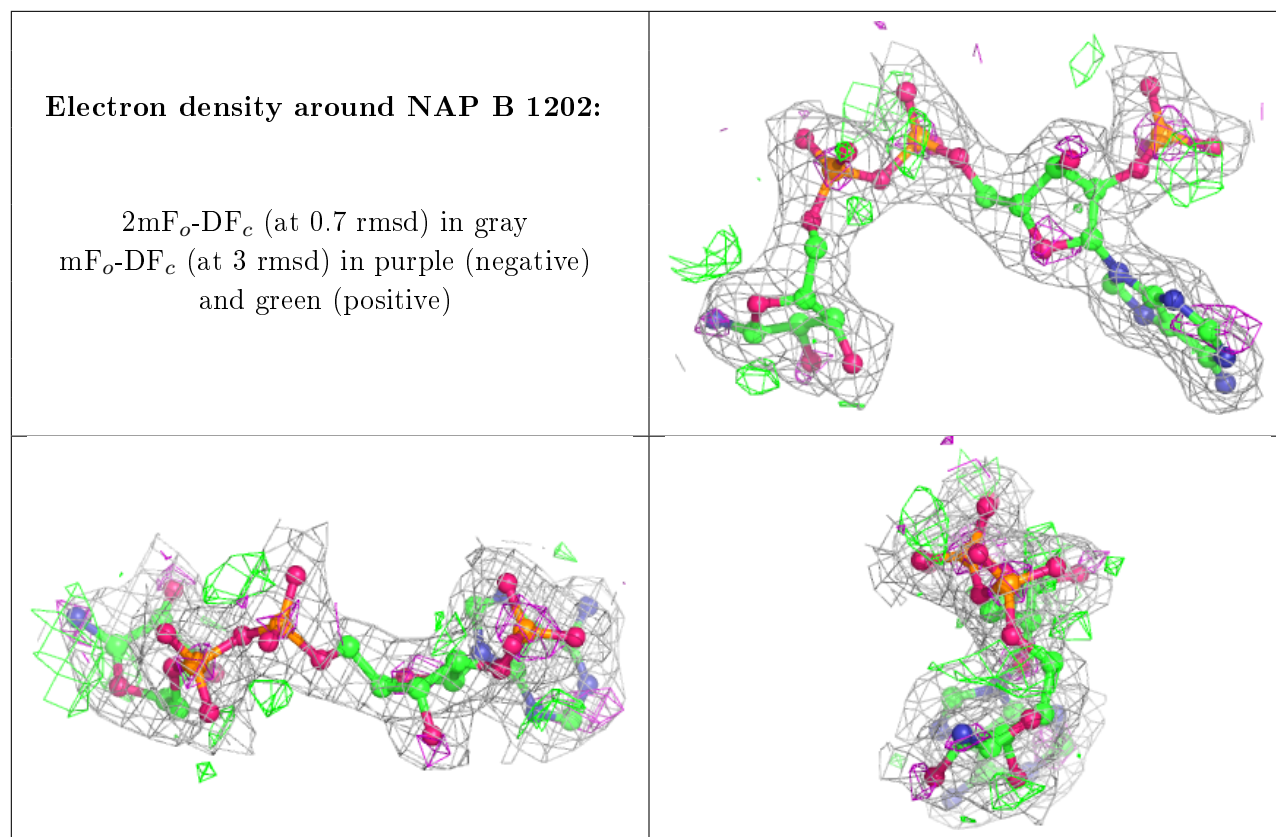
Electron density around NAP A 1202:

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

**Electron density around NAP D 1202:**

$2mF_o-DF_c$ (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.