



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 02:57 am GMT

PDB ID : 4GO0
Title : Crystal structure of the c707s mutant of c-terminal domain of 10'formyltetrahydrofolate dehydrogenase in complex with NADPH
Authors : Tsybovsky, Y.
Deposited on : 2012-08-17
Resolution : 3.38 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

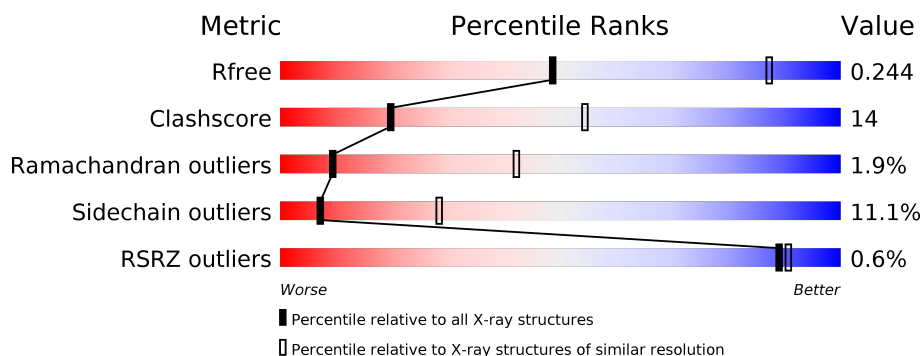
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.38 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}	100719	1241 (3.46-3.30)
Clashscore	112137	1319 (3.46-3.30)
Ramachandran outliers	110173	1298 (3.46-3.30)
Sidechain outliers	110143	1297 (3.46-3.30)
RSRZ outliers	101464	1251 (3.46-3.30)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	517	
1	B	517	
1	C	517	
1	D	517	

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NDP	A	1001	-	-	-	X

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 15488 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 Cytosolic 10-formyltetrahydrofolate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	498	Total	C	N	O	S	0	0	0
			3824	2435	655	715	19			
1	B	498	Total	C	N	O	S	0	0	0
			3824	2435	655	715	19			
1	C	498	Total	C	N	O	S	0	0	0
			3824	2435	655	715	19			
1	D	498	Total	C	N	O	S	0	0	0
			3824	2435	655	715	19			

There are 48 discrepancies between the modelled and reference sequences:

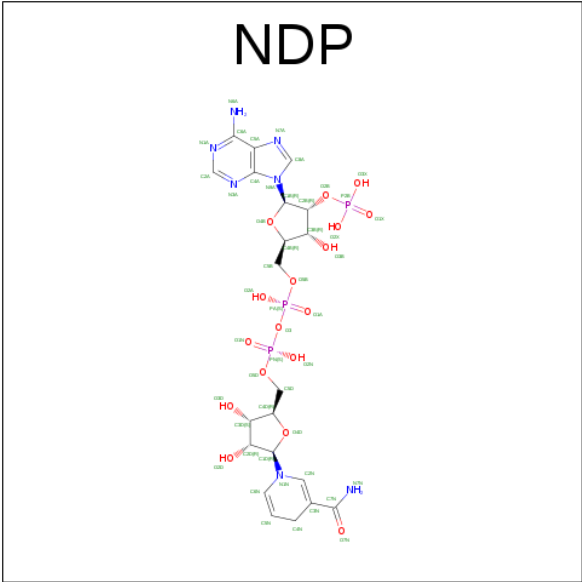
Chain	Residue	Modelled	Actual	Comment	Reference
A	386	MET	-	EXPRESSION TAG	UNP P28037
A	387	ARG	-	EXPRESSION TAG	UNP P28037
A	388	GLY	-	EXPRESSION TAG	UNP P28037
A	389	SER	-	EXPRESSION TAG	UNP P28037
A	390	HIS	-	EXPRESSION TAG	UNP P28037
A	391	HIS	-	EXPRESSION TAG	UNP P28037
A	392	HIS	-	EXPRESSION TAG	UNP P28037
A	393	HIS	-	EXPRESSION TAG	UNP P28037
A	394	HIS	-	EXPRESSION TAG	UNP P28037
A	395	THR	-	EXPRESSION TAG	UNP P28037
A	396	THR	-	EXPRESSION TAG	UNP P28037
A	707	SER	CYS	ENGINEERED MUTATION	UNP P28037
B	386	MET	-	EXPRESSION TAG	UNP P28037
B	387	ARG	-	EXPRESSION TAG	UNP P28037
B	388	GLY	-	EXPRESSION TAG	UNP P28037
B	389	SER	-	EXPRESSION TAG	UNP P28037
B	390	HIS	-	EXPRESSION TAG	UNP P28037
B	391	HIS	-	EXPRESSION TAG	UNP P28037
B	392	HIS	-	EXPRESSION TAG	UNP P28037
B	393	HIS	-	EXPRESSION TAG	UNP P28037
B	394	HIS	-	EXPRESSION TAG	UNP P28037

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	395	THR	-	EXPRESSION TAG	UNP P28037
B	396	THR	-	EXPRESSION TAG	UNP P28037
B	707	SER	CYS	ENGINEERED MUTATION	UNP P28037
C	386	MET	-	EXPRESSION TAG	UNP P28037
C	387	ARG	-	EXPRESSION TAG	UNP P28037
C	388	GLY	-	EXPRESSION TAG	UNP P28037
C	389	SER	-	EXPRESSION TAG	UNP P28037
C	390	HIS	-	EXPRESSION TAG	UNP P28037
C	391	HIS	-	EXPRESSION TAG	UNP P28037
C	392	HIS	-	EXPRESSION TAG	UNP P28037
C	393	HIS	-	EXPRESSION TAG	UNP P28037
C	394	HIS	-	EXPRESSION TAG	UNP P28037
C	395	THR	-	EXPRESSION TAG	UNP P28037
C	396	THR	-	EXPRESSION TAG	UNP P28037
C	707	SER	CYS	ENGINEERED MUTATION	UNP P28037
D	386	MET	-	EXPRESSION TAG	UNP P28037
D	387	ARG	-	EXPRESSION TAG	UNP P28037
D	388	GLY	-	EXPRESSION TAG	UNP P28037
D	389	SER	-	EXPRESSION TAG	UNP P28037
D	390	HIS	-	EXPRESSION TAG	UNP P28037
D	391	HIS	-	EXPRESSION TAG	UNP P28037
D	392	HIS	-	EXPRESSION TAG	UNP P28037
D	393	HIS	-	EXPRESSION TAG	UNP P28037
D	394	HIS	-	EXPRESSION TAG	UNP P28037
D	395	THR	-	EXPRESSION TAG	UNP P28037
D	396	THR	-	EXPRESSION TAG	UNP P28037
D	707	SER	CYS	ENGINEERED MUTATION	UNP P28037

- Molecule 2 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: C₂₁H₃₀N₇O₁₇P₃).

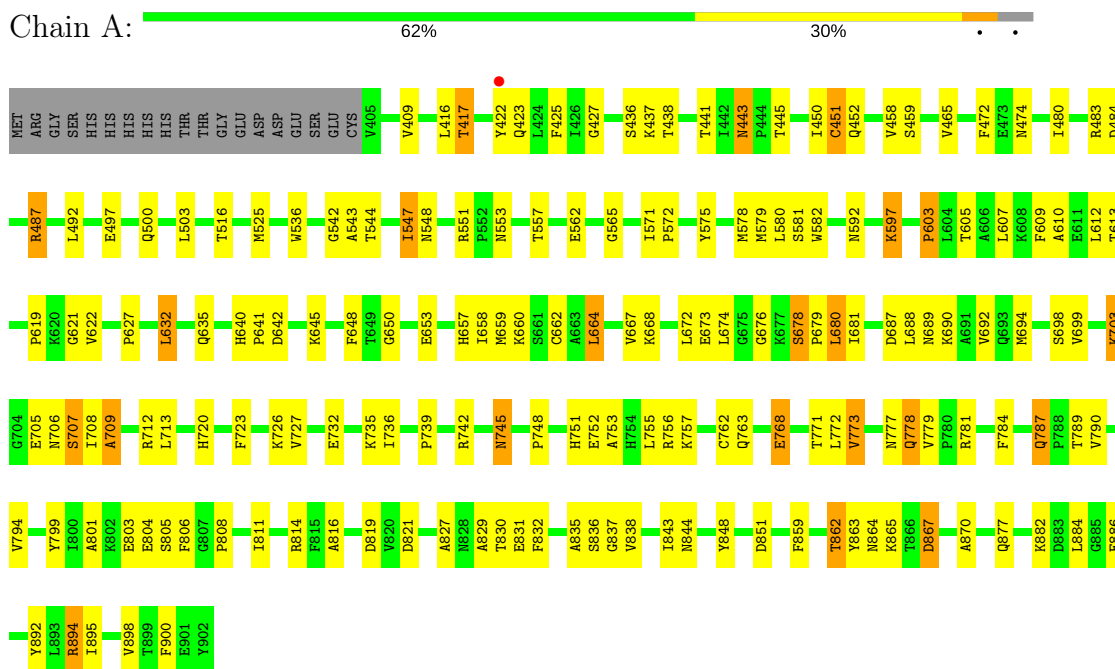


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	C	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	D	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

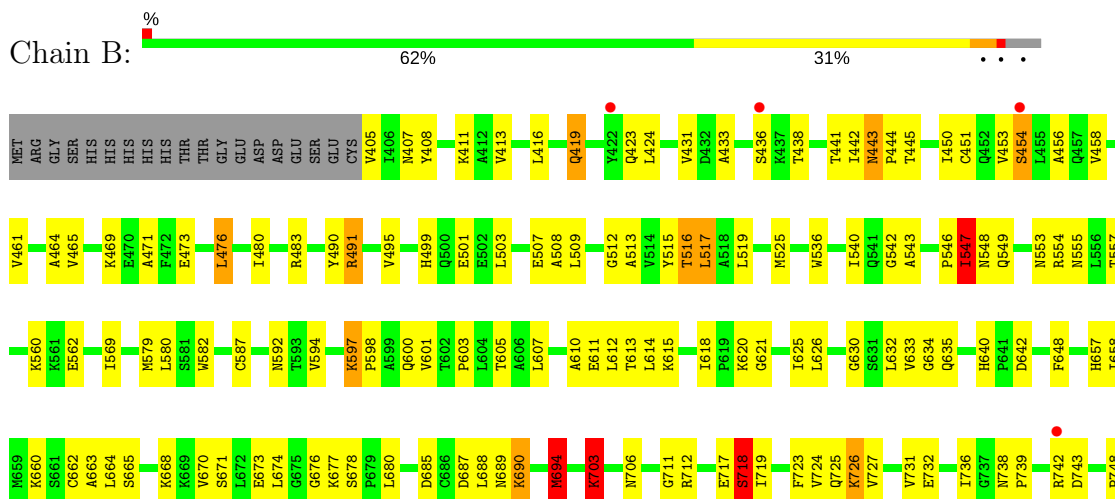
3 Residue-property plots

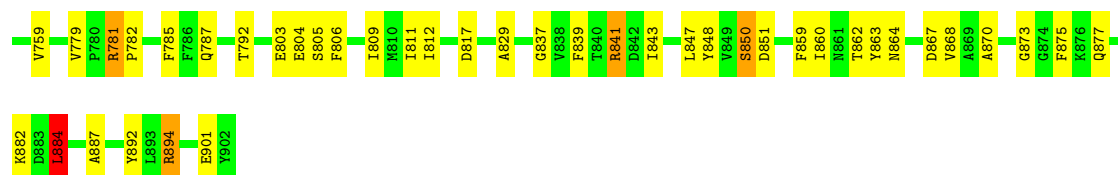
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Cytosolic 10-formyltetrahydrofolate dehydrogenase



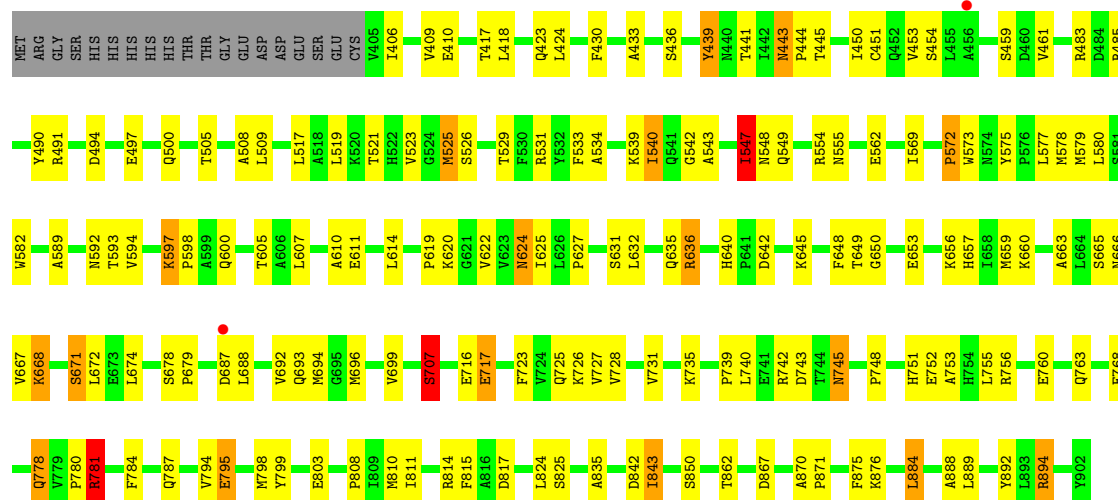
- Molecule 1: Cytosolic 10-formyltetrahydrofolate dehydrogenase





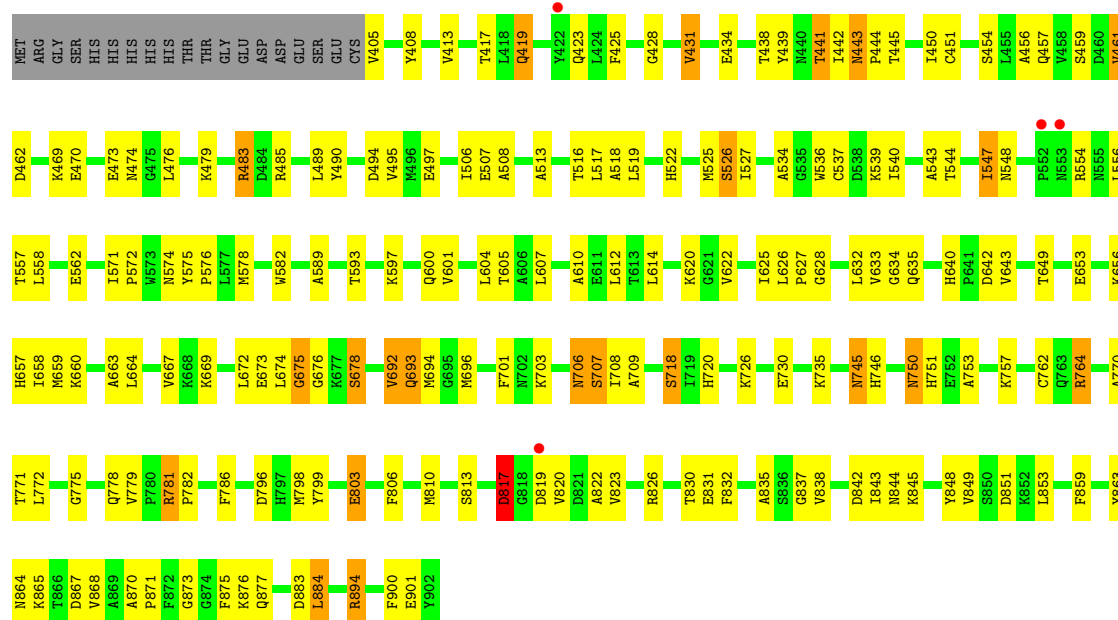
• Molecule 1: Cytosolic 10-formyltetrahydrofolate dehydrogenase

Chain C: 64% 28%



• Molecule 1: Cytosolic 10-formyltetrahydrofolate dehydrogenase

Chain D: 60% 32%



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	258.42Å 194.13Å 96.97Å 90.00° 108.82° 90.00°	Depositor
Resolution (Å)	52.50 – 3.38 48.83 – 3.38	Depositor EDS
% Data completeness (in resolution range)	98.2 (52.50-3.38) 98.2 (48.83-3.38)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.27 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.195 , 0.252 0.194 , 0.244	Depositor DCC
R_{free} test set	3161 reflections (5.37%)	DCC
Wilson B-factor (Å ²)	57.0	Xtriage
Anisotropy	0.797	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 26.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	15488	wwPDB-VP
Average B, all atoms (Å ²)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% 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: NDP

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	0.83	1/3898 (0.0%)	0.92	1/5274 (0.0%)
1	B	0.84	1/3898 (0.0%)	0.89	3/5274 (0.1%)
1	C	0.89	5/3898 (0.1%)	0.92	3/5274 (0.1%)
1	D	0.87	4/3898 (0.1%)	0.91	4/5274 (0.1%)
All	All	0.86	11/15592 (0.1%)	0.91	11/21096 (0.1%)

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	A	0	1

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	490	TYR	CE2-CZ	7.58	1.48	1.38
1	D	490	TYR	CE2-CZ	7.51	1.48	1.38
1	D	803	GLU	CG-CD	6.33	1.61	1.51
1	A	762	CYS	CB-SG	-6.16	1.71	1.82
1	B	490	TYR	CE2-CZ	6.04	1.46	1.38

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	687	ASP	CB-CG-OD1	-7.59	111.47	118.30
1	C	687	ASP	CB-CG-OD2	6.62	124.26	118.30
1	D	781	ARG	NE-CZ-NH2	-6.23	117.19	120.30
1	D	554	ARG	NE-CZ-NH2	-6.08	117.26	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	694	MET	CG-SD-CE	6.07	109.91	100.20

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	705	GLU	Peptide

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	3824	0	3851	121	0
1	B	3824	0	3851	133	0
1	C	3824	0	3851	117	0
1	D	3824	0	3851	123	0
2	A	48	0	26	3	0
2	B	48	0	26	4	0
2	C	48	0	26	2	0
2	D	48	0	26	7	0
All	All	15488	0	15508	443	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 14.

The worst 5 of 443 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:497:GLU:O	1:A:500:GLN:HG2	1.58	1.03
1:A:676:GLY:O	1:A:804:GLU:HG3	1.66	0.95
1:B:443:ASN:C	1:B:443:ASN:HD22	1.75	0.90
1:B:443:ASN:ND2	1:B:445:THR:H	1.69	0.88
1:C:593:THR:HG22	1:C:622:VAL:HG13	1.54	0.87

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	496/517 (96%)	452 (91%)	35 (7%)	9 (2%)	10	43
1	B	496/517 (96%)	449 (90%)	37 (8%)	10 (2%)	9	41
1	C	496/517 (96%)	436 (88%)	51 (10%)	9 (2%)	10	43
1	D	496/517 (96%)	428 (86%)	58 (12%)	10 (2%)	9	41
All	All	1984/2068 (96%)	1765 (89%)	181 (9%)	38 (2%)	9	42

5 of 38 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	772	LEU
1	B	436	SER
1	B	718	SER
1	C	707	SER
1	C	781	ARG

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	409/426 (96%)	365 (89%)	44 (11%)	7	30
1	B	409/426 (96%)	363 (89%)	46 (11%)	7	28
1	C	409/426 (96%)	368 (90%)	41 (10%)	9	33
1	D	409/426 (96%)	358 (88%)	51 (12%)	5	22
All	All	1636/1704 (96%)	1454 (89%)	182 (11%)	7	28

5 of 182 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	759	VAL
1	C	554	ARG
1	D	726	LYS
1	B	803	GLU
1	C	436	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	443	ASN
1	C	693	GLN
1	D	745	ASN
1	C	548	ASN
1	C	720	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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

4 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
2	NDP	A	1001	-	43,52,52	1.72	5 (11%)	49,80,80	2.06	8 (16%)
2	NDP	B	1001	-	43,52,52	1.63	5 (11%)	49,80,80	1.90	6 (12%)
2	NDP	C	1001	-	43,52,52	1.66	5 (11%)	49,80,80	1.85	5 (10%)
2	NDP	D	1001	-	43,52,52	1.70	5 (11%)	49,80,80	1.91	10 (20%)

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	NDP	A	1001	-	-	0/30/77/77	0/5/5/5
2	NDP	B	1001	-	-	0/30/77/77	0/5/5/5
2	NDP	C	1001	-	-	1/30/77/77	0/5/5/5
2	NDP	D	1001	-	-	0/30/77/77	0/5/5/5

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1001	NDP	C2A-N1A	2.23	1.38	1.33
2	D	1001	NDP	C2A-N1A	2.51	1.38	1.33
2	B	1001	NDP	C2N-C3N	2.53	1.42	1.34
2	A	1001	NDP	C2A-N1A	2.57	1.38	1.33
2	A	1001	NDP	C2N-C3N	2.68	1.42	1.34

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1001	NDP	N3A-C2A-N1A	-9.69	120.42	128.86
2	C	1001	NDP	N3A-C2A-N1A	-9.42	120.65	128.86
2	A	1001	NDP	N3A-C2A-N1A	-9.39	120.68	128.86
2	D	1001	NDP	N3A-C2A-N1A	-9.03	120.99	128.86
2	B	1001	NDP	O4B-C1B-C2B	-3.39	100.66	106.59

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1001	NDP	PN-O5D-C5D-C4D

There are no ring outliers.

4 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	NDP	3	0
2	B	1001	NDP	4	0
2	C	1001	NDP	2	0
2	D	1001	NDP	7	0

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	498/517 (96%)	-0.20	1 (0%) 94 96	38, 52, 68, 78	0
1	B	498/517 (96%)	-0.19	4 (0%) 86 88	38, 53, 68, 87	0
1	C	498/517 (96%)	-0.33	2 (0%) 92 93	33, 49, 61, 74	0
1	D	498/517 (96%)	-0.27	4 (0%) 86 88	35, 51, 67, 80	0
All	All	1992/2068 (96%)	-0.25	11 (0%) 89 90	33, 51, 67, 87	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	422	TYR	2.9
1	B	422	TYR	2.7
1	D	552	PRO	2.5
1	D	553	ASN	2.3
1	B	454	SER	2.2

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
2	NDP	A	1001	48/48	0.92	0.25	2.05	62,72,92,94	0
2	NDP	C	1001	48/48	0.92	0.27	1.93	51,65,81,82	0
2	NDP	D	1001	48/48	0.95	0.21	0.31	51,62,80,83	0
2	NDP	B	1001	48/48	0.94	0.20	0.02	64,71,86,87	0

6.5 Other polymers [i](#)

There are no such residues in this entry.