wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 2, 2022 – 06:34 pm GMT

PDB ID	:	7Q94
Title	:	Crystal Structure of Agrobacterium tumefaciens NADQ, DNA complex.
Authors	:	Cianci, M.; Minazzato, G.; Heroux, A.; Raffaelli, N.; Sorci, L.; Gasparrini, M.
Deposited on	:	2021-11-11
Resolution	:	4.30 Å(reported)

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

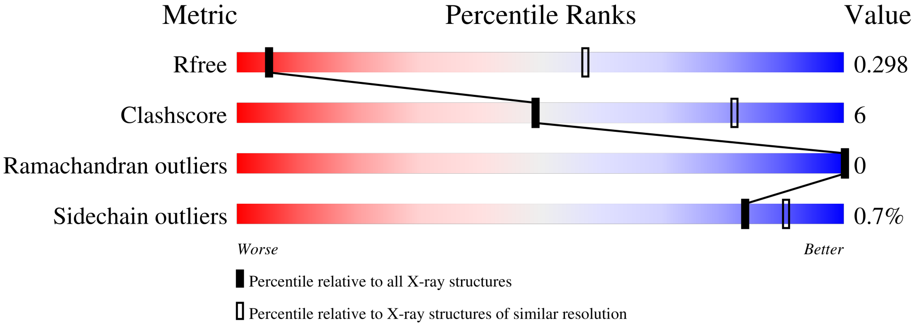
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 4.30 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	1014 (4.80-3.80)
Clashscore	141614	1077 (4.80-3.80)
Ramachandran outliers	138981	1029 (4.80-3.80)
Sidechain outliers	138945	1012 (4.80-3.80)

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%

Mol	Chain	Length	Quality of chain
1	A	336	
1	B	336	
2	C	32	
3	D	32	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5614 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 NADQ transcription factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	267	Total 2188	C 1385	N 402	O 397	S 4	0	0	0
1	B	264	Total 2163	C 1373	N 395	O 391	S 4	0	0	0

There are 74 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-35	MET	-	initiating methionine	UNP A9CG24
A	-34	ARG	-	expression tag	UNP A9CG24
A	-33	GLY	-	expression tag	UNP A9CG24
A	-32	SER	-	expression tag	UNP A9CG24
A	-31	HIS	-	expression tag	UNP A9CG24
A	-30	HIS	-	expression tag	UNP A9CG24
A	-29	HIS	-	expression tag	UNP A9CG24
A	-28	HIS	-	expression tag	UNP A9CG24
A	-27	HIS	-	expression tag	UNP A9CG24
A	-26	HIS	-	expression tag	UNP A9CG24
A	-25	GLY	-	expression tag	UNP A9CG24
A	-24	MET	-	expression tag	UNP A9CG24
A	-23	ALA	-	expression tag	UNP A9CG24
A	-22	SER	-	expression tag	UNP A9CG24
A	-21	MET	-	expression tag	UNP A9CG24
A	-20	THR	-	expression tag	UNP A9CG24
A	-19	GLY	-	expression tag	UNP A9CG24
A	-18	GLY	-	expression tag	UNP A9CG24
A	-17	GLN	-	expression tag	UNP A9CG24
A	-16	GLN	-	expression tag	UNP A9CG24
A	-15	MET	-	expression tag	UNP A9CG24
A	-14	GLY	-	expression tag	UNP A9CG24
A	-13	ARG	-	expression tag	UNP A9CG24
A	-12	ASP	-	expression tag	UNP A9CG24
A	-11	LEU	-	expression tag	UNP A9CG24

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-10	TYR	-	expression tag	UNP A9CG24
A	-9	ASP	-	expression tag	UNP A9CG24
A	-8	ASP	-	expression tag	UNP A9CG24
A	-7	ASP	-	expression tag	UNP A9CG24
A	-6	ASP	-	expression tag	UNP A9CG24
A	-5	LYS	-	expression tag	UNP A9CG24
A	-4	ASP	-	expression tag	UNP A9CG24
A	-3	HIS	-	expression tag	UNP A9CG24
A	-2	PRO	-	expression tag	UNP A9CG24
A	-1	PHE	-	expression tag	UNP A9CG24
A	0	THR	-	expression tag	UNP A9CG24
A	1	VAL	-	expression tag	UNP A9CG24
B	-35	MET	-	initiating methionine	UNP A9CG24
B	-34	ARG	-	expression tag	UNP A9CG24
B	-33	GLY	-	expression tag	UNP A9CG24
B	-32	SER	-	expression tag	UNP A9CG24
B	-31	HIS	-	expression tag	UNP A9CG24
B	-30	HIS	-	expression tag	UNP A9CG24
B	-29	HIS	-	expression tag	UNP A9CG24
B	-28	HIS	-	expression tag	UNP A9CG24
B	-27	HIS	-	expression tag	UNP A9CG24
B	-26	HIS	-	expression tag	UNP A9CG24
B	-25	GLY	-	expression tag	UNP A9CG24
B	-24	MET	-	expression tag	UNP A9CG24
B	-23	ALA	-	expression tag	UNP A9CG24
B	-22	SER	-	expression tag	UNP A9CG24
B	-21	MET	-	expression tag	UNP A9CG24
B	-20	THR	-	expression tag	UNP A9CG24
B	-19	GLY	-	expression tag	UNP A9CG24
B	-18	GLY	-	expression tag	UNP A9CG24
B	-17	GLN	-	expression tag	UNP A9CG24
B	-16	GLN	-	expression tag	UNP A9CG24
B	-15	MET	-	expression tag	UNP A9CG24
B	-14	GLY	-	expression tag	UNP A9CG24
B	-13	ARG	-	expression tag	UNP A9CG24
B	-12	ASP	-	expression tag	UNP A9CG24
B	-11	LEU	-	expression tag	UNP A9CG24
B	-10	TYR	-	expression tag	UNP A9CG24
B	-9	ASP	-	expression tag	UNP A9CG24
B	-8	ASP	-	expression tag	UNP A9CG24
B	-7	ASP	-	expression tag	UNP A9CG24
B	-6	ASP	-	expression tag	UNP A9CG24

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	LYS	-	expression tag	UNP A9CG24
B	-4	ASP	-	expression tag	UNP A9CG24
B	-3	HIS	-	expression tag	UNP A9CG24
B	-2	PRO	-	expression tag	UNP A9CG24
B	-1	PHE	-	expression tag	UNP A9CG24
B	0	THR	-	expression tag	UNP A9CG24
B	1	VAL	-	expression tag	UNP A9CG24

- Molecule 2 is a DNA chain called DNA binding region (31-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	31	630	304	112	184	30	0	0	0

- Molecule 3 is a DNA chain called DNA binding region (31-MER).

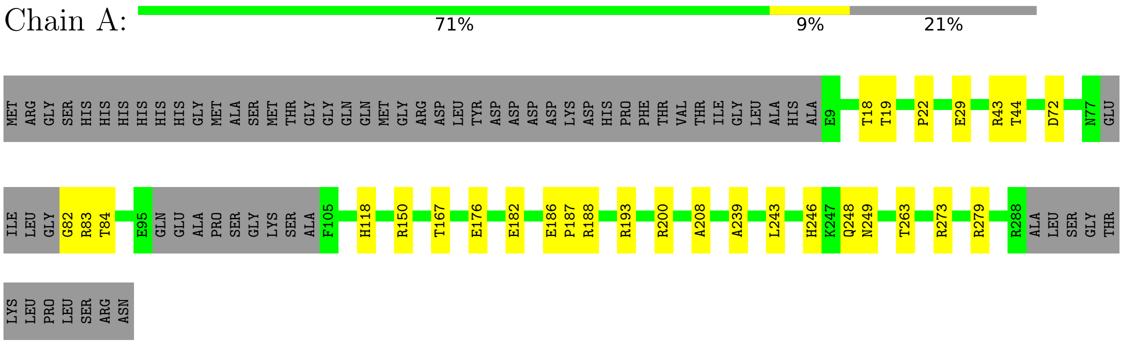
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	D	31	633	305	115	183	30	0	0	0

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. 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. 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.

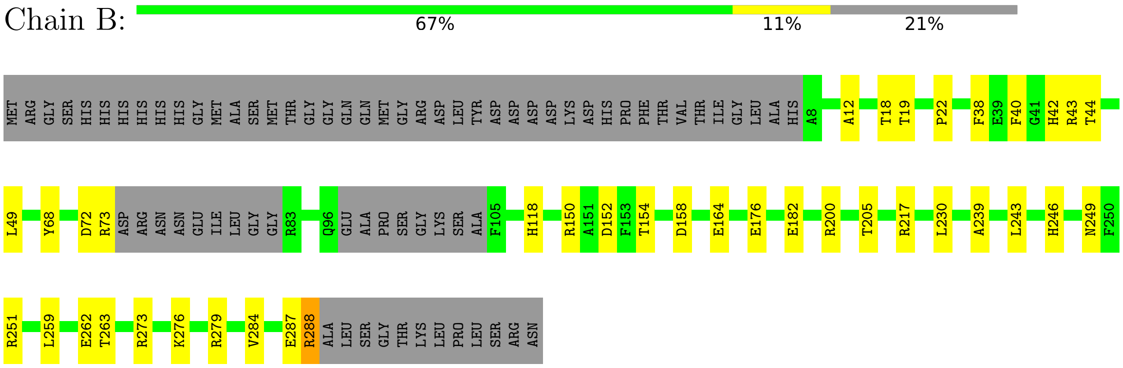
- Molecule 1: NADQ transcription factor

Chain A:



- Molecule 1: NADQ transcription factor

Chain B:



- Molecule 2: DNA binding region (31-MER)

Chain C:

- Molecule 3: DNA binding region (31-MER)

Chain D:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	104.27Å 104.27Å 152.15Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	77.65 – 4.30 77.65 – 4.30	Depositor EDS
% Data completeness (in resolution range)	99.7 (77.65-4.30) 99.7 (77.65-4.30)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.43 (at 4.30Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.254 , 0.300 0.253 , 0.298	Depositor DCC
R_{free} test set	331 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å^2)	169.0	Xtriage
Anisotropy	0.156	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.095 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5614	wwPDB-VP
Average B, all atoms (Å^2)	177.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.88% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

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

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	0.24	0/2239	0.50	0/3026
1	B	0.23	0/2214	0.51	0/2993
2	C	0.50	0/704	0.94	0/1081
3	D	0.50	0/710	0.93	0/1094
All	All	0.32	0/5867	0.65	0/8194

There are no bond length outliers.

There are no bond angle outliers.

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	2188	0	2132	19	0
1	B	2163	0	2113	28	0
2	C	630	0	350	12	0
3	D	633	0	353	19	0
All	All	5614	0	4948	60	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 6.

The worst 5 of 60 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:186:GLU:HG2	1:A:187:PRO:HD3	1.52	0.91
1:A:248:GLN:OE1	3:D:20:DA:N6	2.26	0.69
1:B:246:HIS:HB3	1:B:249:ASN:HB2	1.76	0.68
1:B:273:ARG:NH2	3:D:6:DA:O4'	2.28	0.67
1:B:152:ASP:HB2	1:B:158:ASP:HB2	1.79	0.64

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	261/336 (78%)	244 (94%)	17 (6%)	0	100	100
1	B	258/336 (77%)	242 (94%)	16 (6%)	0	100	100
All	All	519/672 (77%)	486 (94%)	33 (6%)	0	100	100

There are no Ramachandran outliers to report.

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	224/278 (81%)	223 (100%)	1 (0%)	91	94
1	B	221/278 (80%)	219 (99%)	2 (1%)	78	88
All	All	445/556 (80%)	442 (99%)	3 (1%)	84	90

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	188	ARG
1	B	72	ASP
1	B	288	ARG

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

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 ⓘ

There are no ligands in this entry.

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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.