WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2023 – 07:48 PM EDT

PDB ID	:	1NKP
Title	:	Crystal structure of Myc-Max recognizing DNA
Authors	:	Nair, S.K.; Burley, S.K.
Deposited on	:	2003-01-03
Resolution	:	1.80 Å(reported)

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

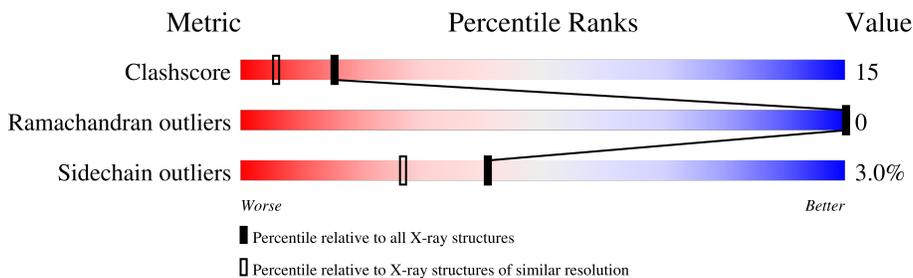|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 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	Percentile Ranks	Value
Clashscore		15
Ramachandran outliers		0
Sidechain outliers		3.0%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	F	19	74% 21% 5%
1	G	19	84% 16%
1	H	19	79% 21%
1	J	19	79% 21%
2	A	88	60% 39% •
2	D	88	63% 32% 6%
3	B	83	71% 27% •
3	E	83	73% 24% •

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 4906 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 DNA chain called 5'-D(*CP*GP*AP*GP*TP*AP*GP*CP*AP*CP*GP*TP*GP*CP*TP*AP*CP*TP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	F	19	Total	C	N	O	P	0	0	0
			385	184	71	112	18			
1	G	19	Total	C	N	O	P	0	0	0
			385	184	71	112	18			
1	H	19	Total	C	N	O	P	0	0	0
			385	184	71	112	18			
1	J	19	Total	C	N	O	P	0	0	0
			385	184	71	112	18			

- Molecule 2 is a protein called Myc proto-oncogene protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	88	Total	C	N	O	S	0	0	0
			733	455	145	131	2			
2	D	83	Total	C	N	O	S	0	0	0
			705	440	138	126	1			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	897	GLY	-	cloning artifact	UNP P01106
A	898	HIS	-	cloning artifact	UNP P01106
A	899	MET	-	cloning artifact	UNP P01106
A	982	GLY	-	insertion	UNP P01106
A	983	GLY	-	insertion	UNP P01106
A	984	CYS	-	insertion	UNP P01106
D	497	GLY	-	cloning artifact	UNP P01106
D	498	HIS	-	cloning artifact	UNP P01106
D	499	MET	-	cloning artifact	UNP P01106
D	582	GLY	-	insertion	UNP P01106
D	583	GLY	-	insertion	UNP P01106

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	584	CYS	-	insertion	UNP P01106

- Molecule 3 is a protein called Max protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	83	Total 686	C 414	N 143	O 127	S 2	0	0	0
3	E	81	Total 661	C 399	N 139	O 121	S 2	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	282	GLY	-	insertion	UNP P61244
B	283	GLY	-	insertion	UNP P61244
B	284	CYS	-	insertion	UNP P61244
E	782	GLY	-	insertion	UNP P61244
E	783	GLY	-	insertion	UNP P61244
E	784	CYS	-	insertion	UNP P61244

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	F	62	Total 62	O 62	0	0
4	G	67	Total 67	O 67	0	0
4	H	60	Total 60	O 60	0	0
4	J	65	Total 65	O 65	0	0
4	A	91	Total 91	O 91	0	0
4	B	75	Total 75	O 75	0	0
4	D	82	Total 82	O 82	0	0
4	E	79	Total 79	O 79	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.
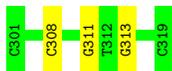
Note EDS was not executed.

• Molecule 1: 5'-D(*CP*GP*AP*GP*TP*AP*GP*CP*AP*CP*GP*TP*GP*CP*TP*AP*CP*TP*C)-3'

Chain F:

• Molecule 1: 5'-D(*CP*GP*AP*GP*TP*AP*GP*CP*AP*CP*GP*TP*GP*CP*TP*AP*CP*TP*C)-3'

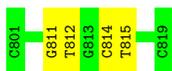Chain G:

• Molecule 1: 5'-D(*CP*GP*AP*GP*TP*AP*GP*CP*AP*CP*GP*TP*GP*CP*TP*AP*CP*TP*C)-3'

Chain H:

• Molecule 1: 5'-D(*CP*GP*AP*GP*TP*AP*GP*CP*AP*CP*GP*TP*GP*CP*TP*AP*CP*TP*C)-3'

Chain J:

• Molecule 2: Myc proto-oncogene protein

Chain A:

G897 H898 K902 R903 R904 T905 H906 N907 V908 Q912 R913 R914 N915 E916 R919 Q927 I928 L931 N934 P938 K939 V940 V941 K945 A946 T947 L951 S952 V953 E956 E957 Q958 K959 L960 I961 S962 E963 K969 R970 R971 E972 Q973 L974 K975 H976 Q980 C984

• Molecule 2: Myc proto-oncogene protein

Chain D:

63% 32% 6%

GLY HIS M499 R504 T505 H506 E510 R514 R519 D526 Q527 I528 P529 E530 L531 N534 E535 K536 A537 P538 V541 K545 A546 Y549 E557 I561 S562 L566 L567 R568 K569 R570 R571 E572 L581 GLY GLY CYS

• Molecule 3: Max protein

Chain B:

71% 27% •

D202 K203 R212 K213 R214 R215 K219 H223 S224 L225 V229 L232 E248 Y249 Y252 M253 R254 R255 I264 D265 K268 R269 Q270 N271 A272 L273 Q276 R279 C284

• Molecule 3: Max protein

Chain E:

73% 24% •

ASP LYS R704 K713 R714 R715 D716 H717 F722 S738 Q741 K745 I750 Q751 H758 T759 H760 D765 D766 R769 Q770 N771 L774 E775 R779 A780 L781 C784

4 Data and refinement statistics ⓘ

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	39.24Å 45.13Å 86.48Å 87.91° 84.61° 71.50°	Depositor
Resolution (Å)	20.00 – 1.80	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-1.80)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.219 , 0.263	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4906	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

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	F	0.41	0/431	0.77	0/663
1	G	0.37	0/431	0.75	0/663
1	H	0.37	0/431	0.74	0/663
1	J	0.38	0/431	0.72	0/663
2	A	0.31	0/740	0.48	0/985
2	D	0.31	0/711	0.48	0/947
3	B	0.30	0/695	0.49	0/927
3	E	0.30	0/670	0.47	0/896
All	All	0.34	0/4540	0.61	0/6407

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	F	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	F	114	DC	Sidechain

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	F	385	0	215	5	0
1	G	385	0	215	3	0
1	H	385	0	215	5	0
1	J	385	0	215	3	0
2	A	733	0	771	39	0
2	D	705	0	750	33	0
3	B	686	0	681	31	0
3	E	661	0	649	20	0
4	A	91	0	0	6	0
4	B	75	0	0	6	0
4	D	82	0	0	12	0
4	E	79	0	0	6	0
4	F	62	0	0	2	0
4	G	67	0	0	0	0
4	H	60	0	0	3	0
4	J	65	0	0	0	0
All	All	4906	0	3711	119	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 15.

The worst 5 of 119 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)
2:A:908:VAL:HB	4:A:287:HOH:O	1.51	1.08
2:A:927:GLN:HE21	3:B:254:ARG:NH2	1.52	1.06
2:A:927:GLN:HE21	3:B:254:ARG:HH22	1.09	0.93
2:D:534:ASN:ND2	2:D:537:ALA:HB2	1.89	0.88
3:B:229:VAL:HG21	3:B:232:LEU:HD12	1.61	0.82

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	
2	A	86/88 (98%)	84 (98%)	2 (2%)	0	100	100
2	D	81/88 (92%)	80 (99%)	1 (1%)	0	100	100
3	B	81/83 (98%)	77 (95%)	4 (5%)	0	100	100
3	E	79/83 (95%)	76 (96%)	3 (4%)	0	100	100
All	All	327/342 (96%)	317 (97%)	10 (3%)	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	
2	A	80/80 (100%)	78 (98%)	2 (2%)	47	34
2	D	78/80 (98%)	76 (97%)	2 (3%)	46	32
3	B	73/73 (100%)	70 (96%)	3 (4%)	30	16
3	E	69/73 (94%)	67 (97%)	2 (3%)	42	29
All	All	300/306 (98%)	291 (97%)	9 (3%)	41	27

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

Mol	Chain	Res	Type
3	E	722	PHE
3	E	774	LEU
3	B	225	LEU
3	B	279	ARG
2	D	526	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
3	E	760	HIS
3	E	770	GLN
2	D	506	HIS
2	D	507	ASN
2	D	515	ASN

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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.