wwPDB X-ray Structure Validation Summary Report (i)

Nov 2, 2021 – 09:04 PM EDT

PDB ID	:	1UIV
Title	:	Crystal structures of the liganded and unliganded nickel binding protein NikA from Escherichia coli (Nickel liganded form)
Authors	:	Heddle, J.; Scott, D.J.; Unzai, S.; Park, S.-Y.; Tame, J.R.H.
Deposited on	:	2003-07-22
Resolution	:	1.95 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 following versions of software and data (see references (1)) were used in the production of this report:

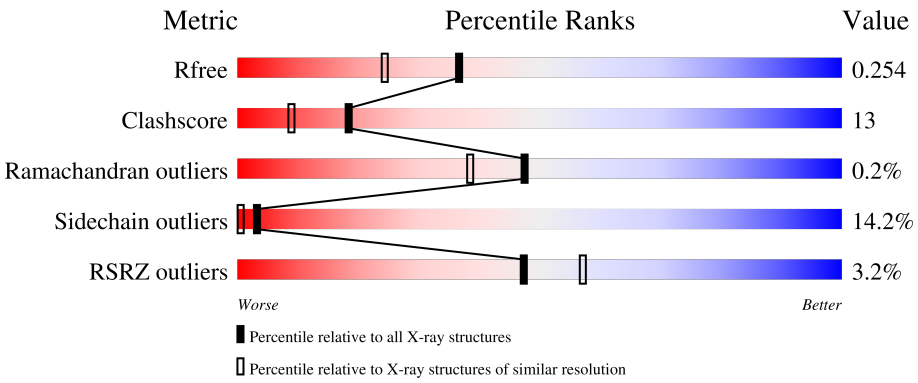
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	A	502	3% 71% 21% 7% ••
1	B	502	3% 67% 23% 7% ••

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8382 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 Nickel-binding periplasmic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	497	3948	2530	668	740	10	0	0	0
1	B	496	3941	2525	667	739	10	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	361	ARG	GLN	engineered mutation	UNP P33590
B	1361	ARG	GLN	engineered mutation	UNP P33590

- Molecule 2 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ni		
2	A	1	1	1	0	0
2	B	1	1	1	0	0

- Molecule 3 is water.

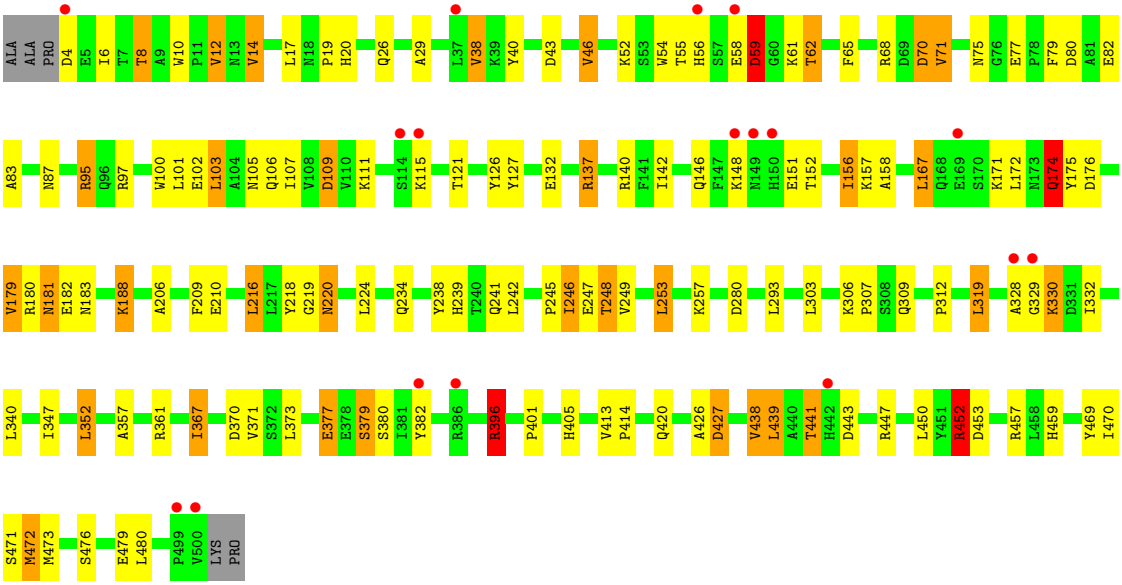
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	258	258	258	0	0
3	B	233	233	233	0	0

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.

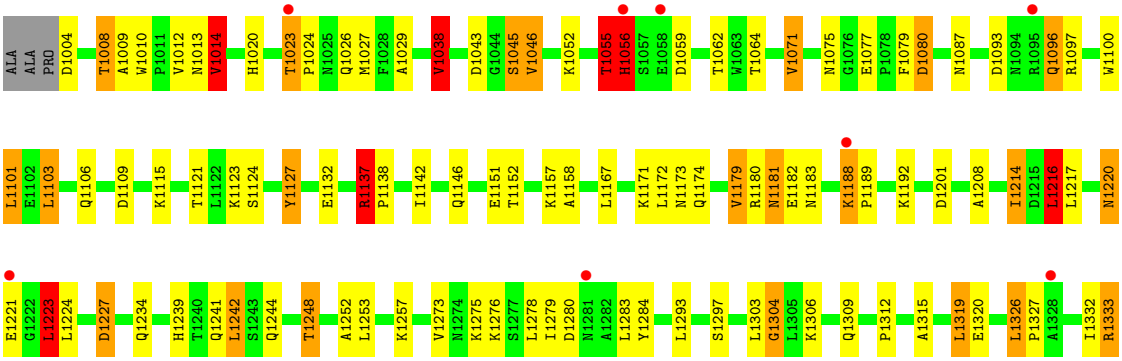
• Molecule 1: Nickel-binding periplasmic protein

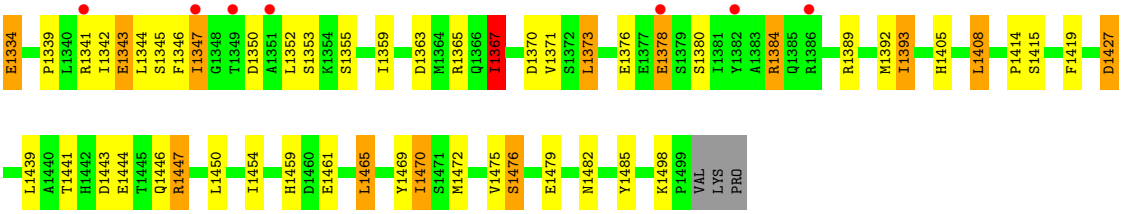
Chain A: 3% 71% 21% 7%



• Molecule 1: Nickel-binding periplasmic protein

Chain B: 3% 67% 23% 7%




E1334
P1339
L1340
R1341
I1342
E1343
L1344
S1345
F1346
I1347
G1348
T1349
D1350
A1351
L1352
S1353
K1354
S1355
I1359
D1363
M1364
R1365
Q1366
T1367
D1370
V1371
S1372
L1373
E1376
E1377
E1378
S1379
S1380
T1381
Y1382
A1383
R1384
Q1385
R1386
R1389
M1392
T1393
H1405
L1408
P1414
S1415
F1419
D1427
L1439
A1440
T1441
H1442
D1443
E1444
T1445
Q1446
R1447
L1450
I1454
H1459
D1460
E1461
L1465
Y1469
I1470
S1471
M1472
V1475
S1476
E1479
N1482
Y1485
K1498
P1499
VAL
LYS
PRO

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	69.72Å 192.81Å 75.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.95 19.90 – 1.94	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.00-1.95) 93.7 (19.90-1.94)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	8.27 (at 1.94Å)	Xtriage
Refinement program	REFMAC 5.1.19	Depositor
R, R_{free}	0.199 , 0.254 0.201 , 0.254	Depositor DCC
R_{free} test set	3602 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	22.0	Xtriage
Anisotropy	0.337	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 48.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8382	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NI

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	0.99	5/4051 (0.1%)	1.21	36/5518 (0.7%)
1	B	1.00	1/4044 (0.0%)	1.14	38/5508 (0.7%)
All	All	0.99	6/8095 (0.1%)	1.17	74/11026 (0.7%)

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	1	1
1	B	0	2
All	All	1	3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1444	GLU	CD-OE1	9.53	1.36	1.25
1	A	479	GLU	CD-OE2	7.66	1.34	1.25
1	A	174	GLN	CB-CG	7.65	1.73	1.52
1	A	452	ARG	NE-CZ	-5.93	1.25	1.33
1	A	452	ARG	CG-CD	5.68	1.66	1.51

The worst 5 of 74 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	452	ARG	NE-CZ-NH2	-24.79	107.90	120.30
1	A	4	ASP	CB-CA-C	19.10	148.59	110.40
1	A	452	ARG	NE-CZ-NH1	15.58	128.09	120.30
1	B	1303	LEU	C-N-CA	-12.21	96.66	122.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1447	ARG	NE-CZ-NH2	-12.12	114.24	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	4	ASP	CA

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	328	ALA	Peptide
1	B	1055	THR	Peptide
1	B	1304	GLY	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	3948	0	3900	105	0
1	B	3941	0	3891	107	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	258	0	0	4	0
3	B	233	0	0	16	0
All	All	8382	0	7791	208	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 13.

The worst 5 of 208 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:B:1315:ALA:HB3	1:B:1367:ILE:CD1	1.61	1.26
1:A:330:LYS:HB3	1:A:330:LYS:NZ	1.55	1.11
1:A:472:MET:HE3	1:A:472:MET:HA	1.32	1.08
1:A:330:LYS:HB3	1:A:330:LYS:HZ3	1.07	1.06

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:234:GLN:HG2	1:B:1304:GLY:HA3	1.42	0.98

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	495/502 (99%)	483 (98%)	11 (2%)	1 (0%)	47	38
1	B	494/502 (98%)	480 (97%)	13 (3%)	1 (0%)	47	38
All	All	989/1004 (98%)	963 (97%)	24 (2%)	2 (0%)	47	38

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	157	LYS
1	B	1378	GLU

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	422/425 (99%)	361 (86%)	61 (14%)	3	0
1	B	421/425 (99%)	362 (86%)	59 (14%)	3	0
All	All	843/850 (99%)	723 (86%)	120 (14%)	3	0

5 of 120 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	452	ARG
1	B	1393	ILE
1	B	1080	ASP
1	B	1384	ARG
1	B	1482	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1302	ASN
1	B	1459	HIS
1	B	1309	GLN
1	B	1420	GLN
1	A	309	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	497/502 (99%)	0.11	17 (3%)	45	55	13, 25, 42, 54	0
1	B	496/502 (98%)	0.10	15 (3%)	50	59	14, 24, 42, 53	0
All	All	993/1004 (98%)	0.11	32 (3%)	47	57	13, 24, 42, 54	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	382	TYR	5.3
1	B	1058	GLU	3.8
1	B	1351	ALA	3.8
1	B	1378	GLU	3.7
1	A	149	ASN	3.6

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, 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(Å^2)	Q<0.9
2	NI	A	503	1/1	1.00	0.02	34,34,34,34	0
2	NI	B	1503	1/1	1.00	0.05	36,36,36,36	0

6.5 Other polymers (i)

There are no such residues in this entry.