



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:32 PM GMT

PDB ID : 4CJD  
Title : Crystal structure of Neisseria meningitidis trimeric autotransporter and vaccine antigen NadA  
Authors : Malito, E.; Biancucci, M.; Spraggon, G.; Bottomley, M.J.  
Deposited on : 2013-12-19  
Resolution : 2.06 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](mailto: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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

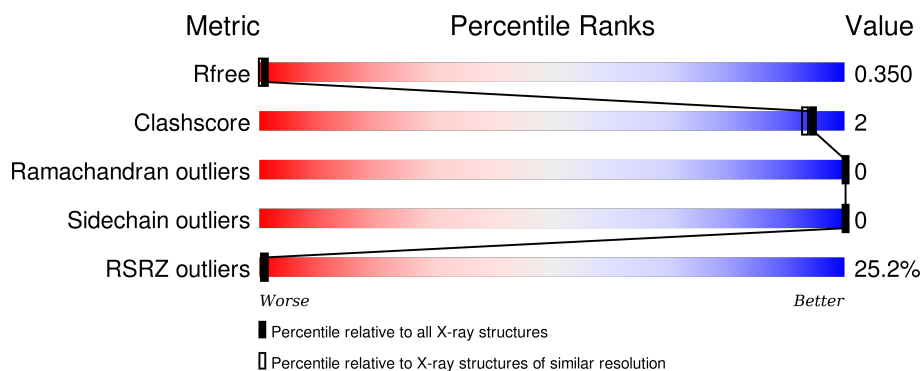
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.06 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))
$R_{free}$	91344	1799 (2.08-2.04)
Clashscore	102246	1910 (2.08-2.04)
Ramachandran outliers	100387	1893 (2.08-2.04)
Sidechain outliers	100360	1893 (2.08-2.04)
RSRZ outliers	91569	1802 (2.08-2.04)

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  $\geq 3$ , 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	208	<div> <div>15%</div> <div>57%</div> <div>.</div> <div>41%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 1009 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 NADA.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	123	Total	C	N	O	0	0	0
			905	552	155	198			

There are 29 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	MET	-	EXPRESSION TAG	UNP A0ELI2
A	23	LYS	-	EXPRESSION TAG	UNP A0ELI2
A	221	ALA	-	EXPRESSION TAG	UNP A0ELI2
A	222	SER	-	EXPRESSION TAG	UNP A0ELI2
A	223	LYS	-	EXPRESSION TAG	UNP A0ELI2
A	224	HIS	-	EXPRESSION TAG	UNP A0ELI2
A	225	HIS	-	EXPRESSION TAG	UNP A0ELI2
A	226	HIS	-	EXPRESSION TAG	UNP A0ELI2
A	227	HIS	-	EXPRESSION TAG	UNP A0ELI2
A	228	HIS	-	EXPRESSION TAG	UNP A0ELI2
A	229	HIS	-	EXPRESSION TAG	UNP A0ELI2
A	40	GLY	ASN	CONFLICT	UNP A0ELI2
A	41	VAL	SER	CONFLICT	UNP A0ELI2
A	.	-	GLY	DELETION	UNP A0ELI2
A	61	GLU	ASN	CONFLICT	UNP A0ELI2
A	64	GLN	LYS	CONFLICT	UNP A0ELI2
A	65	PRO	ILE	CONFLICT	UNP A0ELI2
A	66	LYS	THR	CONFLICT	UNP A0ELI2
A	67	GLY	ARG	CONFLICT	UNP A0ELI2
A	68	ARG	LYS	CONFLICT	UNP A0ELI2
A	69	PRO	THR	CONFLICT	UNP A0ELI2
A	104	GLU	ASP	CONFLICT	UNP A0ELI2
A	132	ARG	GLY	CONFLICT	UNP A0ELI2
A	140	THR	LYS	CONFLICT	UNP A0ELI2
A	144	LYS	GLN	CONFLICT	UNP A0ELI2
A	146	GLU	ALA	CONFLICT	UNP A0ELI2
A	149	GLU	ALA	CONFLICT	UNP A0ELI2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	152	GLY	GLU	CONFLICT	UNP A0ELI2
A	215	LYS	ALA	CONFLICT	UNP A0ELI2

- Molecule 2 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	4	Total I 4 4	0	0

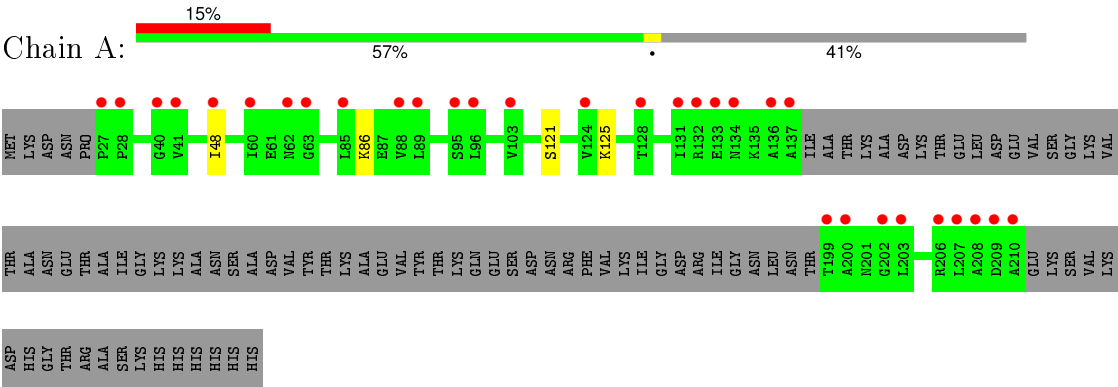
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	100	Total O 100 100	0	0

### 3 Residue-property plots i

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: NADA



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.23Å 51.23Å 577.94Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.16 – 2.06 48.16 – 2.06	Depositor EDS
% Data completeness (in resolution range)	99.3 (48.16-2.06) 99.4 (48.16-2.06)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.13 (at 2.05Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.182 , 0.207 0.363 , 0.350	Depositor DCC
$R_{free}$ test set	955 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	34.8	Xtriage
Anisotropy	0.403	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 44.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	3 of 19087 reflections (0.016%)	Xtriage
$F_o, F_c$ correlation	0.81	EDS
Total number of atoms	1009	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.41% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

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

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	0.65	0/912	0.61	0/1237

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	905	0	876	3	0
2	A	4	0	0	0	0
3	A	100	0	0	0	1
All	All	1009	0	876	3	1

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 2.

All (3) 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:48:ILE:CG2	1:A:86:LYS:HE3	2.42	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:48:ILE:HG22	1:A:86:LYS:HE3	1.97	0.47
1:A:121:SER:O	1:A:125:LYS:HD2	2.19	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:2081:HOH:O	3:A:2088:HOH:O[3_555]	2.17	0.03

## 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	119/208 (57%)	117 (98%)	2 (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	95/167 (57%)	95 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no



such sidechains identified.

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	123/208 (59%)	1.68	31 (25%) <b>1</b> <b>1</b>	12, 30, 101, 121	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	27	PRO	13.4
1	A	209	ASP	9.3
1	A	208	ALA	9.1
1	A	137	ALA	7.8
1	A	136	ALA	7.2
1	A	199	THR	5.8
1	A	207	LEU	5.5
1	A	28	PRO	5.4
1	A	131	ILE	5.3
1	A	210	ALA	5.1
1	A	132	ARG	4.7
1	A	63	GLY	4.0
1	A	88	VAL	3.7
1	A	206	ARG	3.4
1	A	134	ASN	3.1
1	A	200	ALA	2.9
1	A	133	GLU	2.8
1	A	48	ILE	2.7
1	A	62	ASN	2.6
1	A	203	LEU	2.6
1	A	96	LEU	2.6
1	A	41	VAL	2.5
1	A	124	VAL	2.5
1	A	103	VAL	2.5
1	A	202	GLY	2.4
1	A	85	LEU	2.3
1	A	89	LEU	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	128	THR	2.1
1	A	60	ILE	2.0
1	A	95	SER	2.0
1	A	40	GLY	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. 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	IOD	A	1001	1/1	1.00	0.52	-	34,34,34,34	0
2	IOD	A	1002	1/1	1.00	0.36	-	17,17,17,17	0
2	IOD	A	1003	1/1	0.94	0.62	-	64,64,64,64	0
2	IOD	A	1004	1/1	0.32	0.80	-	83,83,83,83	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.