



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 02:52 am GMT

PDB ID : 3M4D  
Title : Crystal structure of the M113N mutant of alpha-hemolysin  
Authors : Montoya, M.; Gouaux, E.  
Deposited on : 2010-03-10  
Resolution : 1.90 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](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
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

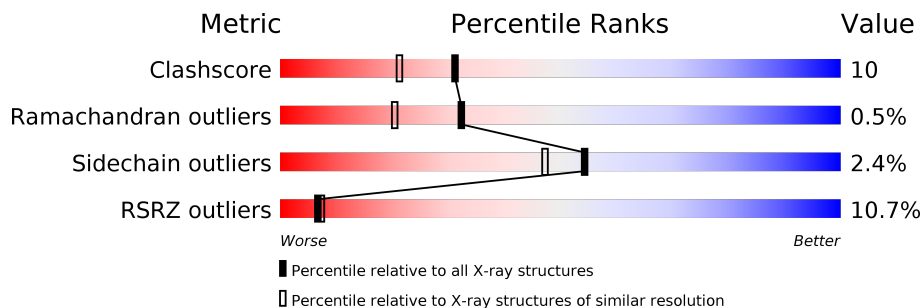
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
Clashscore	112137	5731 (1.90-1.90)
Ramachandran outliers	110173	5669 (1.90-1.90)
Sidechain outliers	110143	5670 (1.90-1.90)
RSRZ outliers	101464	5100 (1.90-1.90)

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	293	<div> <div>11%</div> <div>81%</div> <div>18%</div> </div>
1	B	293	<div> <div>13%</div> <div>81%</div> <div>18%</div> </div>
1	C	293	<div> <div>11%</div> <div>80%</div> <div>18%</div> </div>
1	D	293	<div> <div>13%</div> <div>77%</div> <div>22%</div> </div>
1	E	293	<div> <div>9%</div> <div>83%</div> <div>16%</div> </div>
1	F	293	<div> <div>9%</div> <div>83%</div> <div>15%</div> </div>
1	G	293	<div> <div>9%</div> <div>77%</div> <div>20%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 16762 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 Alpha-hemolysin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	293	Total	C	N	O	S	41	0	0
			2338	1465	402	465	6			
1	B	293	Total	C	N	O	S	45	0	0
			2345	1471	402	466	6			
1	C	293	Total	C	N	O	S	59	0	0
			2345	1471	402	466	6			
1	D	293	Total	C	N	O	S	66	0	0
			2345	1471	402	466	6			
1	E	293	Total	C	N	O	S	49	0	0
			2341	1467	402	466	6			
1	F	293	Total	C	N	O	S	62	0	0
			2345	1471	402	466	6			
1	G	293	Total	C	N	O	S	52	0	0
			2345	1471	402	466	6			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	113	ASN	MET	ENGINEERED	UNP P09616
B	113	ASN	MET	ENGINEERED	UNP P09616
C	113	ASN	MET	ENGINEERED	UNP P09616
D	113	ASN	MET	ENGINEERED	UNP P09616
E	113	ASN	MET	ENGINEERED	UNP P09616
F	113	ASN	MET	ENGINEERED	UNP P09616
G	113	ASN	MET	ENGINEERED	UNP P09616

- Molecule 2 is water.

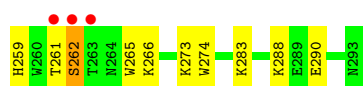
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	45	Total	O	0	0
			45	45		
2	B	52	Total	O	0	0
			52	52		

*Continued on next page...*

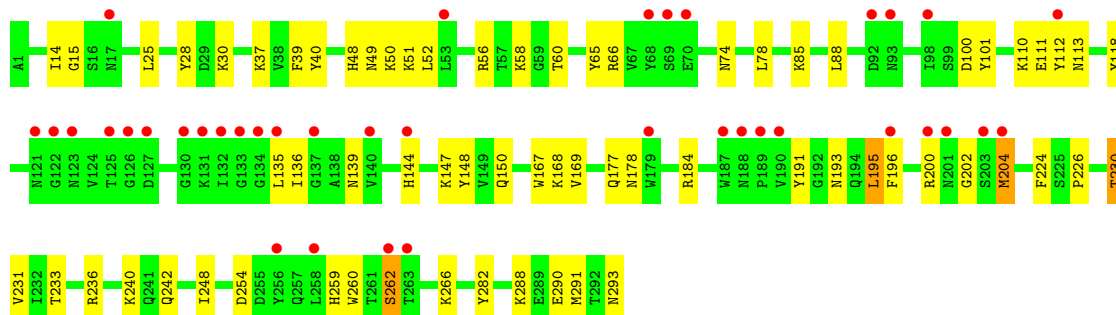
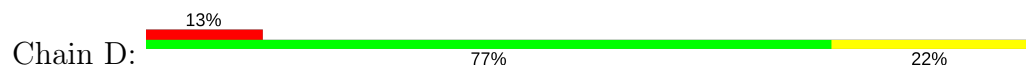
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	52	Total 52	O 52	0	0
2	D	48	Total 48	O 48	0	0
2	E	46	Total 46	O 46	0	0
2	F	54	Total 54	O 54	0	0
2	G	61	Total 61	O 61	0	0

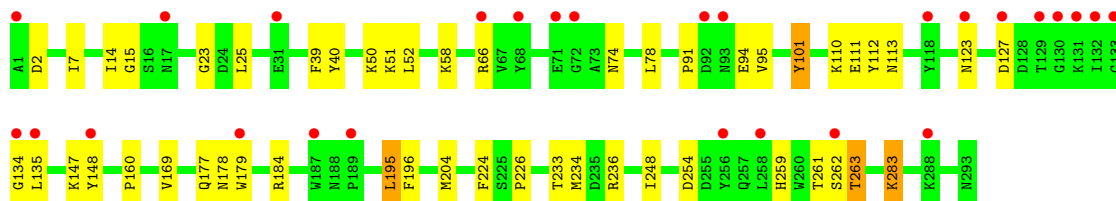
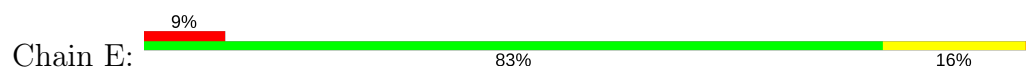




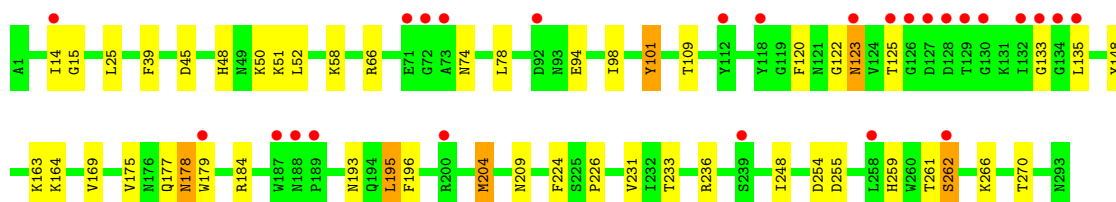
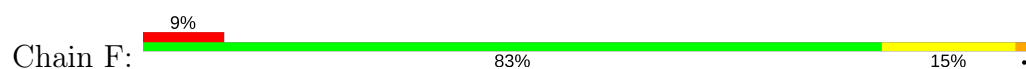
• Molecule 1: Alpha-hemolysin



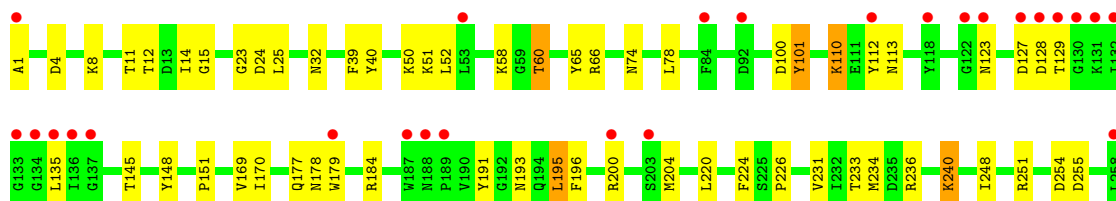
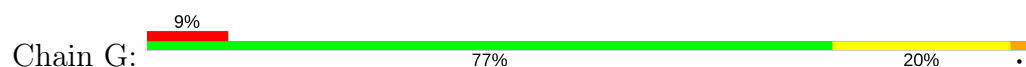
• Molecule 1: Alpha-hemolysin



• Molecule 1: Alpha-hemolysin



• Molecule 1: Alpha-hemolysin





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	151.15Å 134.59Å 132.90Å 90.00° 90.78° 90.00°	Depositor
Resolution (Å)	20.00 – 1.90 19.88 – 1.87	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-1.90) 98.7 (19.88-1.87)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.32 (at 1.87Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.240 , 0.265 0.236 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	21.8	Xtriage
Anisotropy	0.218	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 50.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.007 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	16762	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.41% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.32	0/2389	0.61	0/3233
1	B	0.32	0/2397	0.63	0/3244
1	C	0.32	0/2397	0.64	0/3244
1	D	0.33	0/2397	0.63	0/3244
1	E	0.32	0/2393	0.61	0/3238
1	F	0.32	0/2397	0.61	0/3244
1	G	0.33	0/2397	0.64	0/3244
All	All	0.32	0/16767	0.62	0/22691

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 [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	2338	0	2260	49	0
1	B	2345	0	2267	56	0
1	C	2345	0	2267	63	0
1	D	2345	0	2267	63	0
1	E	2341	0	2257	36	0
1	F	2345	0	2267	45	0
1	G	2345	0	2267	69	0
2	A	45	0	0	4	0
2	B	52	0	0	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	52	0	0	1	0
2	D	48	0	0	2	0
2	E	46	0	0	1	0
2	F	54	0	0	1	0
2	G	61	0	0	1	0
All	All	16762	0	15852	324	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 10.

The worst 5 of 324 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:G:123:ASN:HB3	1:G:135:LEU:HB3	1.39	1.01
1:F:123:ASN:HB3	1:F:135:LEU:HB3	1.45	0.96
1:C:160:PRO:HD2	1:D:60:THR:HG21	1.51	0.93
1:B:123:ASN:HB3	1:B:135:LEU:HB3	1.52	0.90
1:A:56:ARG:NH2	1:G:12:THR:HG21	1.87	0.88

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	291/293 (99%)	280 (96%)	9 (3%)	2 (1%)	25	13
1	B	291/293 (99%)	276 (95%)	13 (4%)	2 (1%)	25	13
1	C	291/293 (99%)	281 (97%)	9 (3%)	1 (0%)	44	34
1	D	291/293 (99%)	279 (96%)	11 (4%)	1 (0%)	44	34
1	E	291/293 (99%)	280 (96%)	11 (4%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	291/293 (99%)	278 (96%)	11 (4%)	2 (1%)	25	13
1	G	291/293 (99%)	277 (95%)	12 (4%)	2 (1%)	25	13
All	All	2037/2051 (99%)	1951 (96%)	76 (4%)	10 (0%)	32	20

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	262	SER
1	B	262	SER
1	C	262	SER
1	B	123	ASN
1	F	262	SER

### 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	258/259 (100%)	253 (98%)	5 (2%)	62	57
1	B	259/259 (100%)	254 (98%)	5 (2%)	62	57
1	C	259/259 (100%)	251 (97%)	8 (3%)	45	36
1	D	259/259 (100%)	253 (98%)	6 (2%)	56	49
1	E	258/259 (100%)	252 (98%)	6 (2%)	56	49
1	F	259/259 (100%)	253 (98%)	6 (2%)	56	49
1	G	259/259 (100%)	251 (97%)	8 (3%)	45	36
All	All	1811/1813 (100%)	1767 (98%)	44 (2%)	54	47

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	101	TYR
1	E	25	LEU
1	G	195	LEU
1	D	110	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	204	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 56 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	64	GLN
1	D	242	GLN
1	G	89	GLN
1	D	74	ASN
1	D	113	ASN

### 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](#)

There are no ligands in this entry.

## 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, 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	293/293 (100%)	0.59	33 (11%) 6 6	14, 24, 52, 69	13 (4%)
1	B	292/293 (99%)	0.56	37 (12%) 4 4	13, 24, 54, 66	12 (4%)
1	C	293/293 (100%)	0.47	32 (10%) 6 7	14, 24, 51, 65	16 (5%)
1	D	292/293 (99%)	0.49	38 (13%) 4 4	13, 24, 50, 60	19 (6%)
1	E	293/293 (100%)	0.45	27 (9%) 10 11	13, 24, 50, 61	17 (5%)
1	F	293/293 (100%)	0.54	26 (8%) 10 11	13, 24, 52, 66	19 (6%)
1	G	292/293 (99%)	0.41	27 (9%) 10 11	14, 23, 46, 65	16 (5%)
All	All	2048/2051 (99%)	0.50	220 (10%) 7 7	13, 24, 52, 69	112 (5%)

The worst 5 of 220 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	179	TRP	8.4
1	A	132	ILE	7.5
1	G	179	TRP	7.2
1	D	179	TRP	6.8
1	A	179	TRP	6.3

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.