

# Full wwPDB X-ray Structure Validation Report (i)

#### Aug 15, 2023 – 03:57 AM EDT

PDB ID : 1T6V

Title: Crystal structure analysis of the nurse shark new antigen receptor (NAR)

variable domain in complex with lysozyme

Authors: Stanfield, R.L.; Dooley, H.; Flajnik, M.F.; Wilson, I.A.

Deposited on : 2004-05-07

Resolution : 1.70 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.orgA user guide is available at

https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

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 : 2.35

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$ 

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

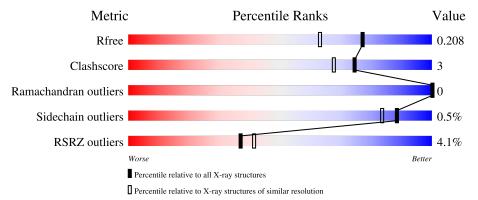
Validation Pipeline (wwPDB-VP) : 2.35

# 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 1.70 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	Similar resolution
Metric	$(\# \mathrm{Entries})$	$(\#  ext{Entries},  ext{ resolution range}( ext{Å}))$
$R_{free}$	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	L	129	5% 92%	8%				
1	M	129	89%	11%				
2	N	113	89%	9%				
2	О	113	86%	12%				



# 2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3961 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 Lysozyme C.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	T.	129	Total	С	N	О	S	0	0	0
1	П	125	1000	613	193	184	10		U	O
1	M	129	Total	С	N	O	S	0	0	0
1	1V1	129	1000	613	193	184	10	0		0

• Molecule 2 is a protein called novel antigen receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	112	Total	С	N	О	S	0	0	0
	11	112	835	505	148	176	6	U		
2	0	112	Total	С	N	О	S	0	0	0
		112	835	505	148	176	6	U		U

• Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	N	1	Total Cl 1 1	0	0
3	О	1	Total Cl 1 1	0	0

• Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	L	69	Total O 69 69	0	0
4	N	78	Total O 78 78	0	0
4	M	78	Total O 78 78	0	0
4	О	64	Total O 64 64	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: Lysozyme C

Chain L:

92%

8%

Molecule 1: Lysozyme C

Chain M:

89%

11%

• Molecule 2: novel antigen receptor

Chain N:

89%

• Molecule 2: novel antigen receptor

Chain O:

88%

• Molecule 2: novel antigen receptor



# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants	75.79Å 113.81Å 58.44Å	Donositon
a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.10 - 1.70	Depositor
Resolution (A)	46.28 - 1.70	EDS
% Data completeness	96.3 (46.10-1.70)	Depositor
(in resolution range)	96.4 (46.28-1.70)	EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$< I/\sigma(I) > 1$	2.72 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.1.9999	Depositor
D D	0.194 , 0.221	Depositor
$R, R_{free}$	0.206 , 0.208	DCC
$R_{free}$ test set	2750 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	8.3	Xtriage
Anisotropy	0.350	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.38, 46.8	EDS
L-test for twinning <sup>2</sup>	$ < L > = 0.49, < L^2> = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	3961	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The analyses of the Patterson function reveals a significant off-origin peak that is 64.19 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 8.6132e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.

<sup>&</sup>lt;sup>2</sup>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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: CL

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	Boı	nd lengths	Bond angles		
IVIOI	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	L	1.00	1/1020 (0.1%)	1.02	0/1379	
1	M	0.97	0/1020	1.02	3/1379 (0.2%)	
2	N	1.02	0/845	1.10	7/1139 (0.6%)	
2	О	1.09	0/845	1.20	9/1139 (0.8%)	
All	All	1.02	1/3730 (0.0%)	1.08	$19/5036 \ (0.4\%)$	

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	${f Z}$	Observed(Å)	$\operatorname{Ideal}( ext{\AA})$
1	L	20	TYR	CD2-CE2	-5.10	1.31	1.39

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$Observed(^o)$	$Ideal(^{o})$
1	M	68	ARG	NE-CZ-NH2	-11.18	114.71	120.30
2	О	8	ARG	NE-CZ-NH2	-9.34	115.63	120.30
2	О	2	ARG	NE-CZ-NH1	7.90	124.25	120.30
2	О	8	ARG	NE-CZ-NH1	7.13	123.87	120.30
2	N	72	ASP	CB-CG-OD2	7.07	124.66	118.30
2	N	106	ASP	CB-CG-OD2	6.64	124.28	118.30
2	N	93	ASP	CB-CG-OD1	6.42	124.08	118.30
2	О	38	ARG	NE-CZ-NH2	-6.24	117.18	120.30
2	N	69	ARG	NE-CZ-NH2	-6.15	117.22	120.30
1	M	52	ASP	CB-CG-OD1	6.00	123.70	118.30
2	Ο	54	ARG	NE-CZ-NH2	-6.00	117.30	120.30
2	Ο	100	ARG	NE-CZ-NH2	-5.73	117.44	120.30
2	Ο	2	ARG	NE-CZ-NH2	-5.65	117.48	120.30
2	N	2	ARG	NE-CZ-NH1	5.41	123.00	120.30
1	M	114	ARG	NE-CZ-NH1	5.40	123.00	120.30
2	О	106	ASP	CB-CG-OD2	5.30	123.07	118.30

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\operatorname{Observed}(^{o})$	$\operatorname{Ideal}({}^{o})$
2	N	2	ARG	NE-CZ-NH2	-5.25	117.67	120.30
2	N	8	ARG	NE-CZ-NH2	-5.24	117.68	120.30
2	О	101	TYR	CG-CD1-CE1	-5.14	117.19	121.30

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	L	1000	0	959	6	0
1	M	1000	0	959	8	0
2	N	835	0	792	4	0
2	O	835	0	792	7	0
3	N	1	0	0	0	0
3	О	1	0	0	0	0
4	L	69	0	0	0	0
4	M	78	0	0	1	0
4	N	78	0	0	0	0
4	О	64	0	0	0	0
All	All	3961	0	3502	23	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 3.

All (23) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	$egin{array}{c} { m Clash} \\ { m overlap} \ ({ m \AA}) \end{array}$
2:O:18:LEU:HG	2:O:19:THR:N	1.97	0.79
2:O:10:VAL:HG11	2:O:18:LEU:CD1	2.13	0.78
2:O:10:VAL:HG11	2:O:18:LEU:HD13	1.69	0.74
1:M:21:ARG:NE	4:M:195:HOH:O	2.21	0.72
1:M:78:ILE:HD12	1:M:83:LEU:HD21	1.76	0.66
1:M:121:GLN:O	1:M:125:ARG:HG3	2.07	0.55
2:N:8:ARG:HD2	2:O:89:GLY:O	2.07	0.53

Continued on next page...



Continued from previous page...

Atom-1	Atom-2	Interatomic	Clash
Atom-1	Atom-2	${ m distance}({ m \AA})$	$overlap(\AA)$
2:N:46:GLU:CD	2:N:94:TYR:CE1	2.81	0.53
1:M:111:TRP:CD1	1:M:115:CYS:HB2	2.45	0.52
1:L:66:ASP:HB3	1:L:80:CYS:SG	2.51	0.51
1:M:21:ARG:NH1	1:M:21:ARG:HG2	2.27	0.48
1:L:111:TRP:CH2	1:L:116:LYS:HG3	2.49	0.48
2:O:18:LEU:HD21	2:O:20:ILE:HG13	1.97	0.47
1:M:15:HIS:HB3	1:M:92:VAL:HG11	1.98	0.45
2:N:46:GLU:OE1	2:N:94:TYR:CE1	2.69	0.45
2:O:38:ARG:NH2	2:O:76:GLU:O	2.46	0.45
1:L:62:TRP:CZ3	1:L:73:ARG:HD3	2.51	0.45
1:L:113:ASN:OD1	1:M:15:HIS:HE1	2.00	0.44
2:O:39:LYS:HD3	2:O:44:GLY:HA2	1.99	0.43
1:L:111:TRP:CD1	1:L:115:CYS:HB2	2.54	0.43
1:L:62:TRP:HB2	1:L:63:TRP:CD1	2.55	0.42
2:N:36:TRP:CZ3	2:N:83:CYS:HB3	2.55	0.41
1:M:63:TRP:O	1:M:76:CYS:HB2	2.21	0.40

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	Perce	$_{ m ntiles}$
1	L	127/129~(98%)	127 (100%)	0	0	100	100
1	M	127/129~(98%)	126 (99%)	1 (1%)	0	100	100
2	N	110/113 (97%)	108 (98%)	2 (2%)	0	100	100
2	О	110/113 (97%)	107 (97%)	3 (3%)	0	100	100
All	All	474/484 (98%)	468 (99%)	6 (1%)	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	L	105/105~(100%)	105 (100%)	0	1	100	100	
1	M	105/105 (100%)	105 (100%)	0	1	100	100	
2	N	91/91 (100%)	90 (99%)	1 (1%)		73	63	
2	О	91/91 (100%)	90 (99%)	1 (1%)		73	63	
All	All	392/392 (100%)	390 (100%)	2 (0%)		88	83	

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	N	66	PHE
2	О	18	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	L	41	GLN
1	L	46	ASN
1	L	113	ASN
1	M	15	HIS
1	M	41	GLN
1	M	46	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 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,  $95^{th}$  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 $>$	$\#\mathrm{RSRZ}{>}2$	$OWAB(Å^2)$	Q<0.9
1	L	129/129~(100%)	0.03	7 (5%) 25 28	6, 9, 26, 35	0
1	M	129/129 (100%)	-0.14	3 (2%) 60 65	5, 9, 18, 39	0
2	N	112/113 (99%)	-0.01	3 (2%) 54 58	6, 10, 21, 28	0
2	О	112/113 (99%)	0.17	7 (6%) 20 22	6, 12, 31, 42	0
All	All	482/484 (99%)	0.01	20 (4%) 37 41	5, 10, 24, 42	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	О	42	GLY	7.0
1	M	128	ARG	6.7
1	M	129	LEU	6.2
1	L	70	PRO	5.2
1	L	71	GLY	4.9
2	О	44	GLY	4.6
1	L	128	ARG	4.6
2	O	45	ASN	4.1
2	О	43	GLU	3.9
2	O	41	SER	3.2
2	N	42	GLY	2.9
1	L	117	GLY	2.7
2	N	44	GLY	2.7
2	О	40	LYS	2.4
1	L	67	GLY	2.3
1	L	69	THR	2.2
2	N	94	TYR	2.2
2	О	47	GLU	2.1
1	L	126	GLY	2.1
1	M	121	GLN	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 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,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
3	CL	N	114	1/1	1.00	0.07	6,6,6,6	0
3	CL	О	114	1/1	1.00	0.05	8,8,8,8	0

### 6.5 Other polymers (i)

There are no such residues in this entry.

