



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 01:56 AM GMT

PDB ID : 3NT9  
Title : CRYSTAL STRUCTURE OF LSSmKate1 red fluorescent proteins with large Stokes shift  
Authors : Malashkevich, V.N.; Piatkevich, K.; Almo, S.C.; Verkhusha, V.  
Deposited on : 2010-07-03  
Resolution : 1.99 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

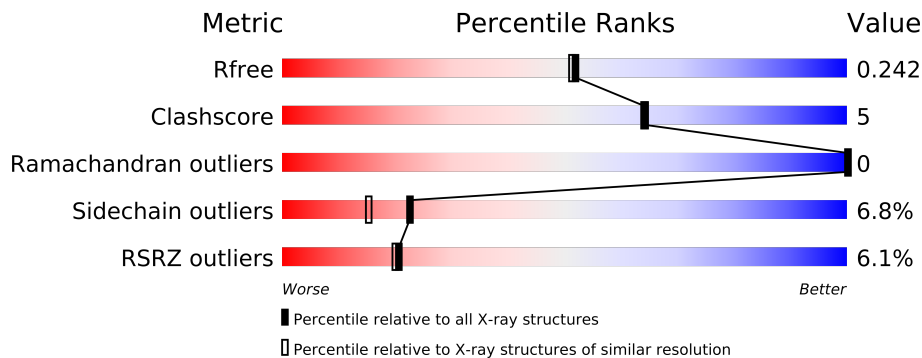
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.99 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))
$R_{free}$	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	243	
1	B	243	
1	C	243	
1	D	243	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7632 atoms, of which 0 are hydrogen and 0 are deuterium.

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 LSSmKate1 red fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	224	Total	C	N	O	S	0	2	0
			1808	1148	308	341	11			
1	B	224	Total	C	N	O	S	0	1	0
			1797	1142	304	340	11			
1	C	224	Total	C	N	O	S	0	1	0
			1795	1143	301	340	11			
1	D	224	Total	C	N	O	S	0	1	0
			1795	1143	301	340	11			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	136	Total	O	0	0
			136	136		
2	B	126	Total	O	0	0
			126	126		
2	C	78	Total	O	0	0
			78	78		
2	D	97	Total	O	0	0
			97	97		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.18Å 72.18Å 226.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.81 – 1.99 19.71 – 1.99	Depositor EDS
% Data completeness (in resolution range)	97.3 (19.81-1.99) 97.3 (19.71-1.99)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.79 (at 1.99Å)	Xtriage
Refinement program	REFMAC 5.5.0089	Depositor
R, $R_{free}$	0.182 , 0.232 0.192 , 0.242	Depositor DCC
$R_{free}$ test set	3843 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	35.2	Xtriage
Anisotropy	0.058	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 29.6	EDS
Estimated twinning fraction	0.057 for h,-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 76295 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7632	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.23% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

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

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	0.94	4/1825 (0.2%)	0.81	3/2461 (0.1%)
1	B	0.85	0/1814	0.78	2/2447 (0.1%)
1	C	0.79	0/1803	0.74	1/2432 (0.0%)
1	D	0.77	0/1803	0.74	1/2432 (0.0%)
All	All	0.84	4/7245 (0.1%)	0.77	7/9772 (0.1%)

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	0	1
1	B	0	1
1	C	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	62	PHE	C-O	6.35	1.35	1.23
1	A	145	GLU	CD-OE1	-5.91	1.19	1.25
1	A	145	GLU	CD-OE2	-5.42	1.19	1.25
1	A	221	TYR	CD1-CE1	-5.37	1.31	1.39

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	102	LEU	CA-CB-CG	6.87	131.10	115.30
1	A	146	MET	CG-SD-CE	-6.83	89.28	100.20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	62	PHE	CA-C-O	-6.79	105.85	120.10
1	D	159	ASP	CB-CG-OD1	5.58	123.32	118.30
1	B	154	LEU	CA-CB-CG	5.55	128.06	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	75	GLY	Peptide
1	B	75	GLY	Peptide
1	C	75	GLY	Peptide

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1808	0	0	12	0
1	B	1797	0	0	11	0
1	C	1795	0	0	6	0
1	D	1795	0	0	10	0
2	A	136	0	0	4	0
2	B	126	0	0	3	0
2	C	78	0	0	0	0
2	D	97	0	0	1	0
All	All	7632	0	0	38	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

The worst 5 of 38 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:67:TYR:OH	1:C:92:ARG:NH2	2.26	0.69
1:C:21:ASN:ND2	1:C:21:ASN:C	2.48	0.67
1:A:201:ARG:NH2	2:A:258:HOH:O	2.34	0.61
1:A:159:ASP:OD2	1:B:159:ASP:OD2	2.23	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:29:GLU:OE2	1:C:42:ARG:NH1	2.39	0.56

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	
1	A	221/243 (91%)	215 (97%)	6 (3%)	0	100	100
1	B	220/243 (90%)	216 (98%)	4 (2%)	0	100	100
1	C	219/243 (90%)	214 (98%)	5 (2%)	0	100	100
1	D	219/243 (90%)	214 (98%)	5 (2%)	0	100	100
All	All	879/972 (90%)	859 (98%)	20 (2%)	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	
1	A	193/207 (93%)	182 (94%)	11 (6%)	29	21
1	B	192/207 (93%)	184 (96%)	8 (4%)	40	34
1	C	191/207 (92%)	177 (93%)	14 (7%)	20	13
1	D	191/207 (92%)	172 (90%)	19 (10%)	11	6
All	All	767/828 (93%)	715 (93%)	52 (7%)	22	15

5 of 52 residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	C	112	ASP
1	C	197	ARG
1	D	186	ASN
1	C	119	VAL
1	C	151	ASP

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

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

6 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
1	NRQ	A	63	1	24,24,25	6.85	3 (12%)	30,32,34	3.05	9 (30%)
1	NRQ	B	63	1	24,24,25	4.54	5 (20%)	30,32,34	2.81	7 (23%)
1	NRQ	C	63[A]	-	7,8,25	0.51	0	8,10,34	0.39	0
1	NRQ	C	63[B]	-	7,8,25	0.64	0	8,10,34	0.65	0
1	NRQ	D	63[A]	-	7,8,25	0.34	0	8,10,34	0.58	0
1	NRQ	D	63[B]	-	7,8,25	0.49	0	8,10,34	0.47	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	NRQ	A	63	1	-	0/9/31/32	0/2/2/2
1	NRQ	B	63	1	-	0/9/31/32	0/2/2/2

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	NRQ	C	63[A]	-	-	0/0/0/32	0/1/1/2
1	NRQ	C	63[B]	-	-	0/0/0/32	0/1/1/2
1	NRQ	D	63[A]	-	-	0/0/0/32	0/1/1/2
1	NRQ	D	63[B]	-	-	0/0/0/32	0/1/1/2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	63	NRQ	O3-C3	32.87	1.34	1.11
1	B	63	NRQ	O3-C3	18.10	1.23	1.11
1	B	63	NRQ	CA3-C3	-10.74	1.33	1.48
1	B	63	NRQ	C2-N3	-3.84	1.31	1.39
1	A	63	NRQ	CA2-C2	-3.71	1.44	1.48

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	63	NRQ	O2-C2-CA2	-11.58	124.27	130.96
1	B	63	NRQ	O2-C2-CA2	-11.28	124.44	130.96
1	A	63	NRQ	CA2-C2-N3	6.52	107.15	103.44
1	B	63	NRQ	CA2-C2-N3	6.19	106.96	103.44
1	A	63	NRQ	CA3-N3-C1	4.67	131.05	124.10

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates 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 ⓘ

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	224/243 (92%)	0.21	5 (2%) 59 59	20, 33, 52, 72	0
1	B	224/243 (92%)	0.30	10 (4%) 32 31	19, 35, 53, 71	0
1	C	224/243 (92%)	0.54	18 (8%) 12 12	24, 42, 65, 78	0
1	D	224/243 (92%)	0.57	22 (9%) 8 7	23, 44, 72, 95	0
All	All	896/972 (92%)	0.41	55 (6%) 21 20	19, 38, 64, 95	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	184	ALA	6.2
1	D	185	LYS	5.7
1	C	185	LYS	5.3
1	D	186	ASN	5.1
1	C	206	ASP	4.6

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

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	NRQ	C	63[B]	8/24	0.15	1.35	20,25,31,32	8
1	NRQ	C	63[A]	8/24	0.15	1.35	28,29,34,35	8
1	NRQ	D	63[A]	8/24	0.14	0.84	32,33,34,34	8
1	NRQ	D	63[B]	8/24	0.14	0.84	27,32,34,35	8
1	NRQ	B	63	23/24	0.14	0.73	29,31,34,37	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	NRQ	A	63	23/24	0.12	0.59	25,29,33,35	0

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.

### 6.5 Other polymers ⓘ

There are no such residues in this entry.