



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:01 AM GMT

PDB ID : 2Z6B  
Title : Crystal Structure Analysis of (gp27-gp5)3 conjugated with Fe(III) protoporphyrin  
Authors : Koshiyama, T.; Yokoi, N.; Ueno, T.; Kanamaru, S.; Nagano, S.; Shiro, Y.; Arisaka, F.; Watanabe, Y.  
Deposited on : 2007-07-28  
Resolution : 3.11 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  
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

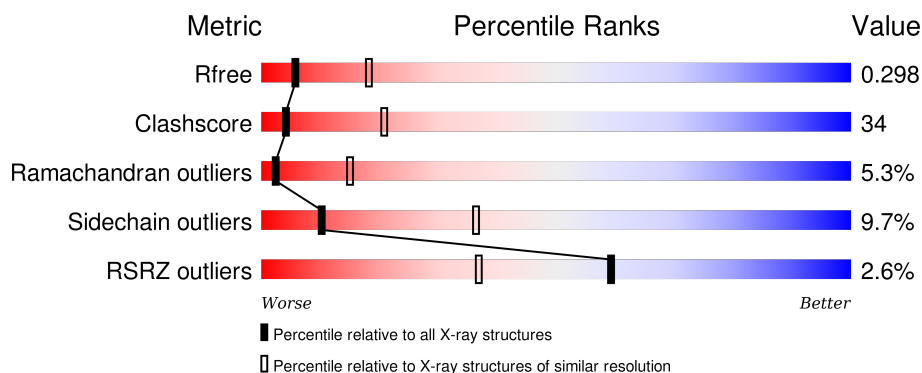
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.11 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	1112 (3.16-3.08)
Clashscore	102246	1218 (3.16-3.08)
Ramachandran outliers	100387	1175 (3.16-3.08)
Sidechain outliers	100360	1175 (3.16-3.08)
RSRZ outliers	91569	1114 (3.16-3.08)

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	584	<div> <div></div> <div>55%</div> <div>37%</div> <div>6%</div> </div>
2	D	391	<div> <div>4%</div> <div>39%</div> <div>44%</div> <div>11%</div> <div>6%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7389 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 Tail-associated lysozyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	573	Total	C	N	O	S	0	0	0
			4273	2645	744	862	22			

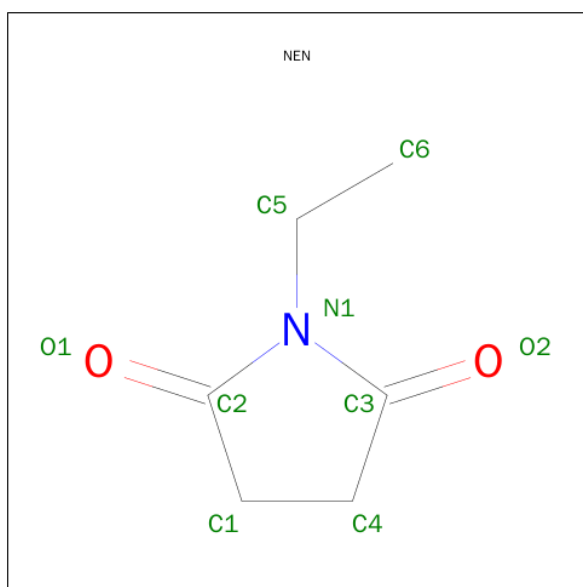
There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	CYS	ASN	ENGINEERED	UNP P16009
A	351	LEU	SER	ENGINEERED	UNP P16009
A	576	SER	-	EXPRESSION TAG	UNP P16009
A	577	VAL	-	EXPRESSION TAG	UNP P16009
A	578	ASP	-	EXPRESSION TAG	UNP P16009
A	579	HIS	-	EXPRESSION TAG	UNP P16009
A	580	HIS	-	EXPRESSION TAG	UNP P16009
A	581	HIS	-	EXPRESSION TAG	UNP P16009
A	582	HIS	-	EXPRESSION TAG	UNP P16009
A	583	HIS	-	EXPRESSION TAG	UNP P16009
A	584	HIS	-	EXPRESSION TAG	UNP P16009

- Molecule 2 is a protein called Baseplate structural protein Gp27.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	368	Total	C	N	O	S	0	0	0
			2786	1779	453	537	17			

- Molecule 3 is 1-ETHYL-PYRROLIDINE-2,5-DIONE (three-letter code: NEN) (formula:  $C_6H_9NO_2$ ).

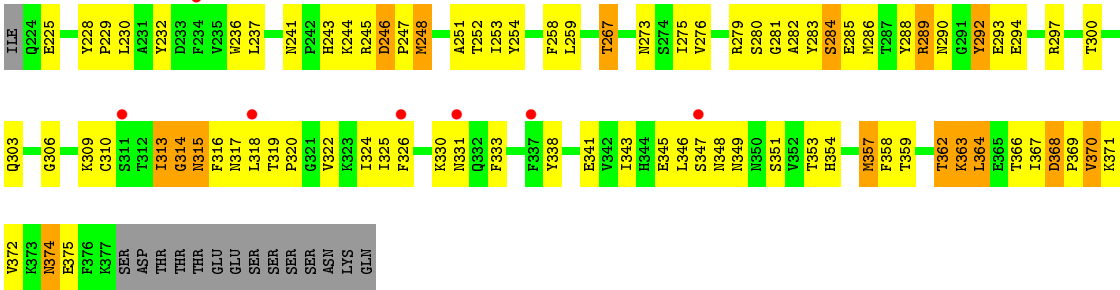


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			9	6	1	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	188	Total	O	0	0
			188	188		
4	D	133	Total	O	0	0
			133	133		





## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	138.54Å 138.54Å 389.12Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.64 – 3.11 45.04 – 3.11	Depositor EDS
% Data completeness (in resolution range)	98.0 (29.64-3.11) 97.7 (45.04-3.11)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.36 (at 3.12Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.243 , 0.315 0.230 , 0.298	Depositor DCC
$R_{free}$ test set	1286 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	60.9	Xtriage
Anisotropy	0.077	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 92.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	2 of 25758 reflections (0.008%)	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	7389	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.30% 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 [i](#)

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

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

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	0.39	0/4355	0.67	0/5931
2	D	0.35	0/2853	0.60	0/3894
All	All	0.37	0/7208	0.64	0/9825

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	4273	0	4022	248	0
2	D	2786	0	2542	218	0
3	A	9	0	7	0	0
4	A	188	0	0	4	0
4	D	133	0	0	1	0
All	All	7389	0	6571	458	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 34.

The worst 5 of 458 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)
2:D:97:VAL:HG12	2:D:98:ASP:H	1.05	1.10
1:A:38:GLN:H	1:A:38:GLN:NE2	1.56	1.04
2:D:313:ILE:HD12	2:D:313:ILE:H	1.28	0.97
1:A:375:LEU:HD22	1:A:455:LYS:HG2	1.46	0.94
1:A:232:ASN:HB3	1:A:233:PRO:HD3	1.50	0.93

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	571/584 (98%)	474 (83%)	78 (14%)	19 (3%)	5	26
2	D	364/391 (93%)	288 (79%)	45 (12%)	31 (8%)	1	5
All	All	935/975 (96%)	762 (82%)	123 (13%)	50 (5%)	2	15

5 of 50 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	258	SER
1	A	293	ASN
1	A	321	LYS
1	A	355	ALA
2	D	32	ALA

### 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	448/494 (91%)	402 (90%)	46 (10%)	9	32
2	D	282/350 (81%)	257 (91%)	25 (9%)	12	42
All	All	730/844 (86%)	659 (90%)	71 (10%)	10	36

5 of 71 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	442	THR
1	A	519	THR
2	D	310	CYS
1	A	466	ASN
1	A	487	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	469	ASN
2	D	5	GLN
2	D	303	GLN
1	A	493	ASN
2	D	25	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

1 ligand is 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	$\# Z  > 2$
3	NEN	A	607	1	9,9,9	1.17	0	12,12,12	4.39	9 (75%)

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
3	NEN	A	607	1	-	1/2/15/15	0/1/1/1

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	607	NEN	C3-N1-C2	-8.18	108.92	112.97
3	A	607	NEN	O1-C2-C1	-4.32	116.00	127.09
3	A	607	NEN	C5-N1-C3	-3.17	119.89	123.32
3	A	607	NEN	O2-C3-C4	-2.66	120.25	127.09
3	A	607	NEN	C4-C3-N1	2.32	110.42	107.93

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	607	NEN	C6-C5-N1-C3

There are no ring outliers.

No monomer is involved in short contacts.

## 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 [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	573/584 (98%)	-0.08	7 (1%) 81 65	9, 40, 85, 161	0
2	D	368/391 (94%)	0.17	17 (4%) 36 17	25, 65, 112, 128	0
All	All	941/975 (96%)	0.02	24 (2%) 59 37	9, 49, 106, 161	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	352	ALA	5.1
2	D	99	SER	4.8
2	D	98	ASP	4.7
1	A	349	SER	4.4
1	A	351	LEU	4.3

### 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](#)

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
3	NEN	A	607	9/9	0.81	0.26	-	103,107,109,111	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.