



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 03:44 PM GMT

PDB ID : 4D4O  
Title : Crystal Structure of the Kti11 Kti13 heterodimer Spacegroup P64  
Authors : Glatt, S.; Mueller, C.W.  
Deposited on : 2014-10-30  
Resolution : 2.90 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  
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

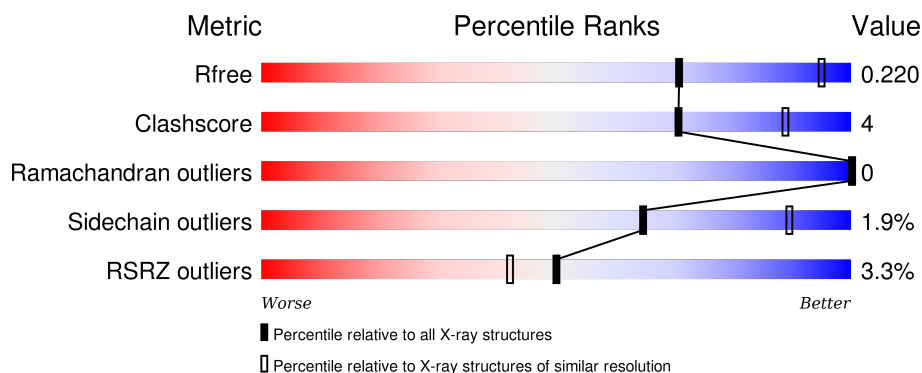
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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))
$R_{free}$	91344	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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	427	<div> <div>4%</div> <div>83%</div> <div>14%</div> <div>•</div> </div>
1	B	427	<div> <div>2%</div> <div>70%</div> <div>7%</div> <div>•</div> <div>22%</div> </div>
1	C	427	<div> <div>16%</div> <div>82%</div> <div>•</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6346 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 PROTEIN ATS1, DIPHTHAMIDE BIOSYNTHESIS PROTEIN 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	414	Total	C	N	O	S	0	0	0
			3185	1987	555	616	27			
1	B	331	Total	C	N	O	S	0	0	0
			2538	1582	464	474	18			
1	C	75	Total	C	N	O	S	0	0	0
			606	384	85	128	9			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	EXPRESSION TAG	UNP P31386
A	0	ALA	-	EXPRESSION TAG	UNP P31386
A	334	GLY	-	LINKER	UNP P31386
A	335	SER	-	LINKER	UNP P31386
A	336	GLY	-	LINKER	UNP P31386
A	337	SER	-	LINKER	UNP P31386
A	338	GLY	-	LINKER	UNP P31386
A	339	SER	-	LINKER	UNP P31386
A	340	GLY	-	LINKER	UNP P31386
A	341	SER	-	LINKER	UNP P31386
A	342	GLY	-	LINKER	UNP P31386
A	343	SER	-	LINKER	UNP P31386
B	-1	GLY	-	EXPRESSION TAG	UNP P31386
B	0	ALA	-	EXPRESSION TAG	UNP P31386
B	334	GLY	-	LINKER	UNP P31386
B	335	SER	-	LINKER	UNP P31386
B	336	GLY	-	LINKER	UNP P31386
B	337	SER	-	LINKER	UNP P31386
B	338	GLY	-	LINKER	UNP P31386
B	339	SER	-	LINKER	UNP P31386
B	340	GLY	-	LINKER	UNP P31386
B	341	SER	-	LINKER	UNP P31386

*Continued on next page...*

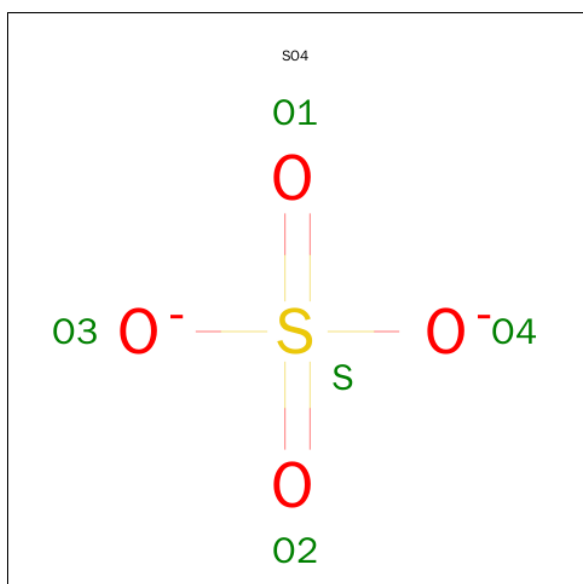
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	342	GLY	-	LINKER	UNP P31386
B	343	SER	-	LINKER	UNP P31386
C	-1	GLY	-	EXPRESSION TAG	UNP P31386
C	0	ALA	-	EXPRESSION TAG	UNP P31386
C	334	GLY	-	LINKER	UNP P31386
C	335	SER	-	LINKER	UNP P31386
C	336	GLY	-	LINKER	UNP P31386
C	337	SER	-	LINKER	UNP P31386
C	338	GLY	-	LINKER	UNP P31386
C	339	SER	-	LINKER	UNP P31386
C	340	GLY	-	LINKER	UNP P31386
C	341	SER	-	LINKER	UNP P31386
C	342	GLY	-	LINKER	UNP P31386
C	343	SER	-	LINKER	UNP P31386

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Fe 1 1	0	0
2	C	1	Total Fe 1 1	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		



ASN	CYS	ASN	LEU	SER
F417	GLY	THR	ASN	VAL
PRO	PRO	VAL	THR	ALA
GLU	GLN	GLU	VAL	VAL
PRO	LYS	SER	SER	ASP
ILE	GLY	PHE	VAL	TYR
ALA	SER	GLY	PHE	VAL
ALA	GLN	ARG	GLY	ALA
ALA	PRO	GLY	THR	MET
ALA	GLY	LEU	GLY	LYS
	LEU	HIS	THR	GLY
	GLN	SER	HIS	ASP
	LEU	SER	SER	PHE
	VAL	GLN	MET	PHE
	GLY	LEU	VAL	MET
	GLN	PHE	ILE	VAL
	GLN	PRO	VAL	VAL
	TYR	GLN	ASP	ASP
	SER	GLU	GLU	GLU
	GLY	ARG	GLY	GLY
	LYS	LEU	LEU	GLY
	PRO	ASP	ARG	ARG
	ARG	PHE	ILE	ILE
	VAL	PRO	VAL	VAL
	PHE	ILE	ILE	HIS
	GLY	VAL	VAL	ALA
	GLY	GLY	GLY	ALA
	CYS	VAL	VAL	SER
	ALA	ALA	ALA	GLY
	THR	THR	THR	ARG
	THR	THR	GLY	LEU
	TRP	SER	SER	PRO
	ILE	GLU	THR	THR
	VAL	HIS	GLU	GLY
	LEU	GLY	HIS	PHE
	GLY	ILE	LEU	GLU
	SER	LEU	LEU	GLY
	GLY	THR	LYS	LEU
	SER	THR	GLN	LYS
	GLY	ALA	GLN	LYS
	SER	ASN	ASN	ARG
	GLY	GLN	GLY	ARG
	SER	GLY	ASN	HIS
	GLY	LYS	LEU	ASN
	S3943	SER	VAL	LEU
	D348	HIS	VAL	VAL
	C368	CYS	LEU	VAL
	P369	TYR	CYS	LEU
	C370	ASN	ASN	CYS
	M381	VAL	MET	MET
	S392	TYR	TRP	TRP
	C393	CYS	ILE	ILE
	Y410	GLY	HIS	HIS
	I415	TRP	LEU	LEU
		GLY	TRP	TRP
		GLU	ASN	ASN
		HIS	ALA	ALA
		GLY	ARG	ARG

H416
F417
PRO
GLU
PRO
ILE
ALA
ALA
ALA
ALA

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	151.27Å 151.27Å 107.14Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	75.64 – 2.90 82.94 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.4 (75.64-2.90) 99.4 (82.94-2.90)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.53 (at 2.91Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.198 , 0.219 0.197 , 0.220	Depositor DCC
$R_{free}$ test set	1545 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	83.8	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 16.1	EDS
Estimated twinning fraction	0.034 for h,-h-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 30909 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6346	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% 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: FE, SO4

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	0.22	0/3261	0.39	0/4426
1	B	0.22	0/2598	0.39	0/3525
1	C	0.42	0/621	0.46	0/841
All	All	0.25	0/6480	0.40	0/8792

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	3185	0	3009	32	0
1	B	2538	0	2438	19	0
1	C	606	0	541	6	0
2	A	1	0	0	0	0
2	C	1	0	0	1	0
3	A	5	0	0	0	0
3	B	10	0	0	2	0
All	All	6346	0	5988	53	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 4.

The worst 5 of 53 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:B:107:ARG:NH1	3:B:1336:SO4:O3	2.12	0.82
1:B:38:VAL:O	1:B:321:ARG:NH2	2.14	0.78
1:A:224:VAL:HB	1:A:235:TRP:HB3	1.67	0.75
1:B:224:VAL:HB	1:B:235:TRP:HB3	1.73	0.71
1:A:294:TRP:HH2	1:C:370:CYS:HB3	1.60	0.66

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	410/427 (96%)	383 (93%)	27 (7%)	0	100	100
1	B	327/427 (77%)	312 (95%)	15 (5%)	0	100	100
1	C	73/427 (17%)	70 (96%)	3 (4%)	0	100	100
All	All	810/1281 (63%)	765 (94%)	45 (6%)	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	344/351 (98%)	340 (99%)	4 (1%)	78	94

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	270/351 (77%)	263 (97%)	7 (3%)	54	85
1	C	69/351 (20%)	67 (97%)	2 (3%)	50	83
All	All	683/1053 (65%)	670 (98%)	13 (2%)	65	89

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	160	LEU
1	B	165	SER
1	B	281	GLU
1	B	118	PHE
1	B	259	LEU

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 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 ⓘ

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 for Mogul analysis.

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	SO4	A	1420	-	4,4,4	0.62	0	6,6,6	0.59	0
3	SO4	B	1336	-	4,4,4	0.39	0	6,6,6	0.23	0
3	SO4	B	1337	-	4,4,4	0.23	0	6,6,6	0.16	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
3	SO4	A	1420	-	-	0/0/0/0	0/0/0/0
3	SO4	B	1336	-	-	0/0/0/0	0/0/0/0
3	SO4	B	1337	-	-	0/0/0/0	0/0/0/0

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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1336	SO4	1	0
3	B	1337	SO4	1	0

## 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	414/427 (96%)	0.54	19 (4%) 36 30	20, 48, 90, 174	0
1	B	331/427 (77%)	0.53	7 (2%) 67 62	20, 43, 86, 102	0
1	C	75/427 (17%)	0.44	1 (1%) 79 78	24, 46, 81, 89	0
All	All	820/1281 (64%)	0.53	27 (3%) 50 42	20, 46, 87, 174	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	176	ASP	4.2
1	A	382	PHE	3.8
1	A	34	ASP	3.5
1	B	201	ILE	3.3
1	A	108	VAL	3.1

### 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	FE	A	501	1/1	0.99	0.19	-0.19	45,45,45,45	0
3	SO4	B	1337	5/5	0.89	0.17	-0.85	82,82,82,82	0
2	FE	C	501	1/1	0.98	0.17	-1.26	47,47,47,47	0
3	SO4	B	1336	5/5	0.89	0.15	-	114,114,114,114	0
3	SO4	A	1420	5/5	0.93	0.10	-	107,107,107,107	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.