W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 13, 2017 – 09:29 pm GMT

|		
PDB ID	:	2ZQK
Title	:	Crystal structure of intimin-Tir68 complex
Authors	:	Ma, Y.; Gao, F.; Li, D.-F.; Gao, G.F.
Deposited on	:	2008-08-12
Resolution	:	2.80 Å(reported)

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 2.80 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.
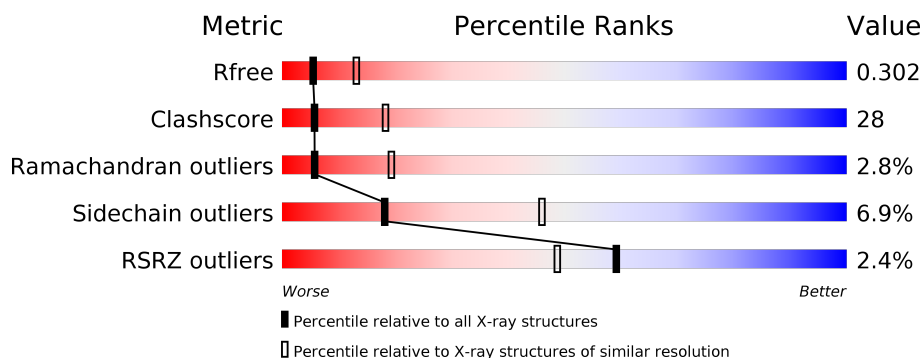
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	2583 (2.80-2.80)
Clashscore	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)
RSRZ outliers	101464	2610 (2.80-2.80)

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 >=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 <=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	189	57% 37% • •
1	B	189	% 47% 43% 7% •
2	C	77	13% 31% 48% • 17%
2	D	77	5% 36% 39% 6% 18%
2	M	77	49% 30% • 17%
2	N	77	44% 31% 6% 18%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 4739 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 Intimin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	183	Total 1410	C 890	N 231	O 284	S 5	0	0	0
1	B	184	Total 1417	C 894	N 232	O 286	S 5	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	INITIATING METHIONINE	UNP P43261
B	0	MET	-	INITIATING METHIONINE	UNP P43261

- Molecule 2 is a protein called Putative translocated intimin receptor protein (Translocated intimin receptor Tir).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	M	64	Total 481	C 289	N 84	O 108	0	0	0
2	N	63	Total 475	C 286	N 83	O 106	0	0	0
2	C	64	Total 481	C 289	N 84	O 108	0	0	0
2	D	63	Total 475	C 286	N 83	O 106	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	0	MET	-	INITIATING METHIONINE	UNP Q7DB77
M	69	LEU	-	EXPRESSION TAG	UNP Q7DB77
M	70	GLU	-	EXPRESSION TAG	UNP Q7DB77
M	71	HIS	-	EXPRESSION TAG	UNP Q7DB77
M	72	HIS	-	EXPRESSION TAG	UNP Q7DB77

Continued on next page...

Continued from previous page...

| Chain	Residue	Modelled	Actual	Comment	Reference
M	73	HIS	-	EXPRESSION TAG	UNP Q7DB77
M	74	HIS	-	EXPRESSION TAG	UNP Q7DB77
M	75	HIS	-	EXPRESSION TAG	UNP Q7DB77
M	76	HIS	-	EXPRESSION TAG	UNP Q7DB77
N	0	MET	-	INITIATING METHIONINE	UNP Q7DB77
N	69	LEU	-	EXPRESSION TAG	UNP Q7DB77
N	70	GLU	-	EXPRESSION TAG	UNP Q7DB77
N	71	HIS	-	EXPRESSION TAG	UNP Q7DB77
N	72	HIS	-	EXPRESSION TAG	UNP Q7DB77
N	73	HIS	-	EXPRESSION TAG	UNP Q7DB77
N	74	HIS	-	EXPRESSION TAG	UNP Q7DB77
N	75	HIS	-	EXPRESSION TAG	UNP Q7DB77
N	76	HIS	-	EXPRESSION TAG	UNP Q7DB77
C	0	MET	-	INITIATING METHIONINE	UNP Q7DB77
C	69	LEU	-	EXPRESSION TAG	UNP Q7DB77
C	70	GLU	-	EXPRESSION TAG	UNP Q7DB77
C	71	HIS	-	EXPRESSION TAG	UNP Q7DB77
C	72	HIS	-	EXPRESSION TAG	UNP Q7DB77
C	73	HIS	-	EXPRESSION TAG	UNP Q7DB77
C	74	HIS	-	EXPRESSION TAG	UNP Q7DB77
C	75	HIS	-	EXPRESSION TAG	UNP Q7DB77
C	76	HIS	-	EXPRESSION TAG	UNP Q7DB77
D	0	MET	-	INITIATING METHIONINE	UNP Q7DB77
D	69	LEU	-	EXPRESSION TAG	UNP Q7DB77
D	70	GLU	-	EXPRESSION TAG	UNP Q7DB77
D	71	HIS	-	EXPRESSION TAG	UNP Q7DB77
D	72	HIS	-	EXPRESSION TAG	UNP Q7DB77
D	73	HIS	-	EXPRESSION TAG	UNP Q7DB77
D	74	HIS	-	EXPRESSION TAG	UNP Q7DB77
D	75	HIS	-	EXPRESSION TAG	UNP Q7DB77
D	76	HIS	-	EXPRESSION TAG	UNP Q7DB77

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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.
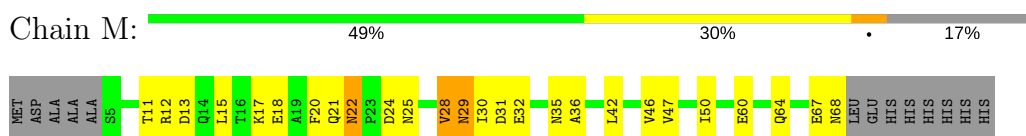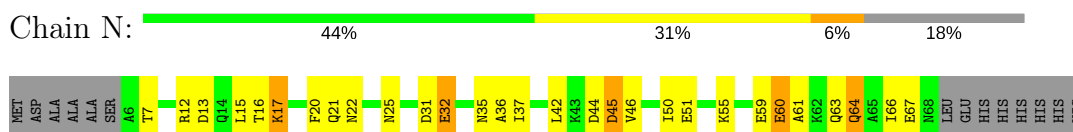
- Molecule 1: Intimin
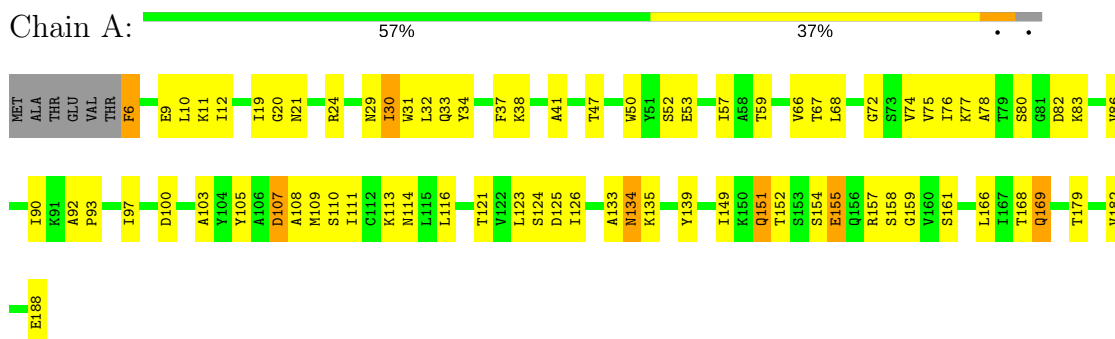
- Molecule 1: Intimin
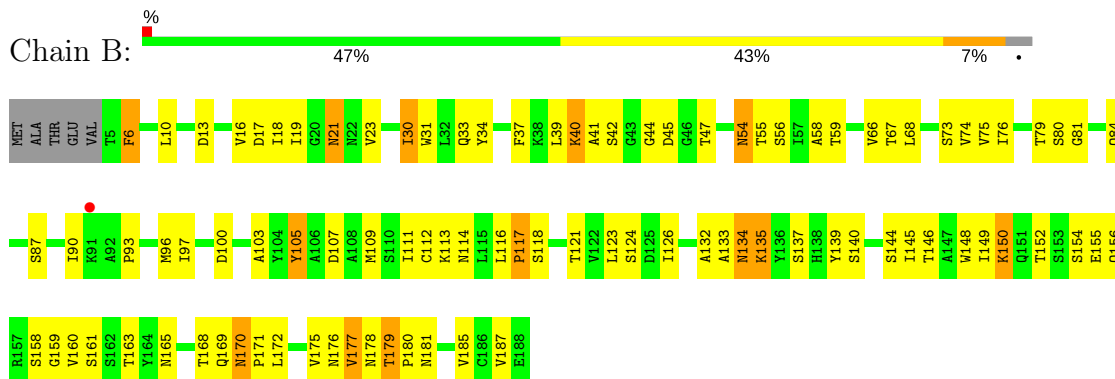
- Molecule 2: Putative translocated intimin receptor protein (Translocated intimin receptor Tir)

- Molecule 2: Putative translocated intimin receptor protein (Translocated intimin receptor Tir)

• Molecule 2: Putative translocated intimin receptor protein (Translocated intimin receptor Tir)

Chain C:

13%

31% 48% • 17%

MET ASP ALA ALA S5 A6 T7 E8 T11 R12 D13 Q14 L15 T16 K17 F20 Q21 N22 P23 D24 N25 Q26 K27 V28 N29 I30 D31 E32 L33 G34 N35 A36 S39 K43 D44 D45 V46 V47 A48 N49 I50 E51 A54 K55 A57 G58 E59 E60 A61 K62 Q63 Q64 A65

I66 E67 N68 LEU GLU HIS HIS HIS HIS HIS HIS

• Molecule 2: Putative translocated intimin receptor protein (Translocated intimin receptor Tir)

Chain D:

5%

36% 39% 6% 18%

MET ASP ALA ALA SER A6 T7 E8 T9 R12 D13 Q14 L15 T16 K17 F20 Q21 N22 N25 Q26 K27 V28 N29 I30 D31 E32 N35 A36 I37 P38 S39 G40 V41 L42 K43 D44 D45 V46 V47 I50 E51 A54 K55 A56 A57 G58 E59 E60 A61 K62 Q63 I66

E67 N68 LEU GLU HIS HIS HIS HIS HIS HIS
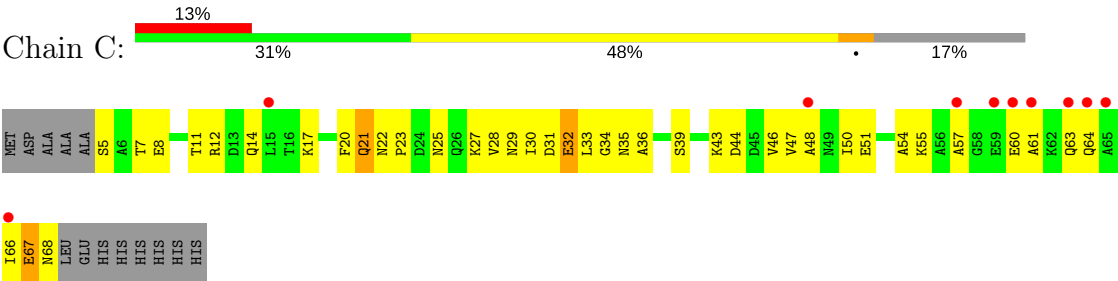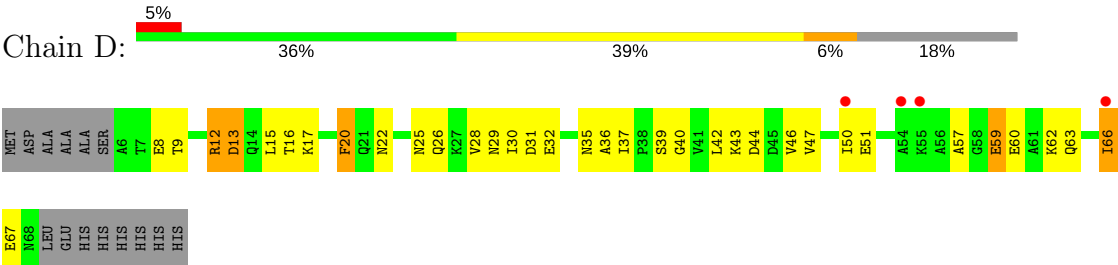
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	116.68Å 42.31Å 182.08Å 90.00° 101.13° 90.00°	Depositor
Resolution (Å)	34.10 – 2.80 34.09 – 2.80	Depositor EDS
% Data completeness (in resolution range)	91.3 (34.10-2.80) 92.9 (34.09-2.80)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	8.51 (at 2.81Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.241 , 0.300 0.245 , 0.302	Depositor DCC
R_{free} test set	2014 reflections (9.83%)	DCC
Wilson B-factor (Å^2)	45.5	Xtriage
Anisotropy	1.109	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 60.4	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	4739	wwPDB-VP
Average B, all atoms (Å^2)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.40% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]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.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

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	0.46	0/1439	0.74	0/1956
1	B	0.36	0/1446	0.63	0/1966
2	C	0.30	0/483	0.57	0/653
2	D	0.33	0/477	0.57	0/645
2	M	0.38	0/483	0.62	0/653
2	N	0.41	0/477	0.62	0/645
All	All	0.39	0/4805	0.65	0/6518

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	1410	0	1379	56	0
1	B	1417	0	1386	93	0
2	C	481	0	462	42	0
2	D	475	0	457	34	0
2	M	481	0	462	25	0
2	N	475	0	457	25	0
All	All	4739	0	4603	262	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 28.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

The worst 5 of 262 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:A:155:GLU:HG2	1:A:161:SER:HA	1.50	0.93
1:B:96:MET:HG2	1:B:97:ILE:H	1.35	0.90
2:D:28:VAL:HG11	2:D:36:ALA:HB1	1.56	0.87
1:B:172:LEU:HD13	2:D:36:ALA:HB2	1.57	0.84
2:D:22:ASN:HB3	2:D:25:ASN:HD22	1.43	0.82

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	181/189 (96%)	151 (83%)	26 (14%)	4 (2%)	8	26
1	B	182/189 (96%)	148 (81%)	26 (14%)	8 (4%)	3	9
2	C	62/77 (80%)	49 (79%)	12 (19%)	1 (2%)	11	36
2	D	61/77 (79%)	56 (92%)	3 (5%)	2 (3%)	4	15
2	M	62/77 (80%)	55 (89%)	7 (11%)	0	100	100
2	N	61/77 (79%)	52 (85%)	7 (12%)	2 (3%)	4	15
All	All	609/686 (89%)	511 (84%)	81 (13%)	17 (3%)	6	19

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	177	VAL
1	A	110	SER
1	B	133	ALA
1	B	150	LYS
2	D	12	ARG

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	158/163 (97%)	146 (92%)	12 (8%)	15	40
1	B	159/163 (98%)	149 (94%)	10 (6%)	21	51
2	C	51/61 (84%)	49 (96%)	2 (4%)	37	71
2	D	50/61 (82%)	45 (90%)	5 (10%)	9	26
2	M	51/61 (84%)	47 (92%)	4 (8%)	15	39
2	N	50/61 (82%)	47 (94%)	3 (6%)	22	54
All	All	519/570 (91%)	483 (93%)	36 (7%)	18	46

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	N	17	LYS
1	B	17	ASP
2	D	44	ASP
2	N	64	GLN
1	B	21	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
2	N	35	ASN
1	B	33	GLN
2	D	35	ASN
2	N	64	GLN
1	B	21	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 ⓘ

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, 95th 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(Å²)	Q<0.9
1	A	183/189 (96%)	-0.18	0	100	100	31, 54, 80, 97	0
1	B	184/189 (97%)	0.04	1 (0%)	90	88	50, 80, 109, 126	0
2	C	64/77 (83%)	0.75	10 (15%)	2	1	73, 107, 136, 153	0
2	D	63/77 (81%)	0.60	4 (6%)	21	13	70, 102, 130, 140	0
2	M	64/77 (83%)	-0.29	0	100	100	42, 61, 85, 91	0
2	N	63/77 (81%)	-0.29	0	100	100	39, 58, 102, 122	0
All	All	621/686 (90%)	0.04	15 (2%)	59	49	31, 71, 121, 153	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	63	GLN	5.3
2	C	66	ILE	4.8
2	C	64	GLN	3.2
2	C	60	GLU	3.1
2	C	65	ALA	2.6

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

6.5 Other polymers ⓘ

There are no such residues in this entry.