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 15, 2017 – 12:51 am GMT

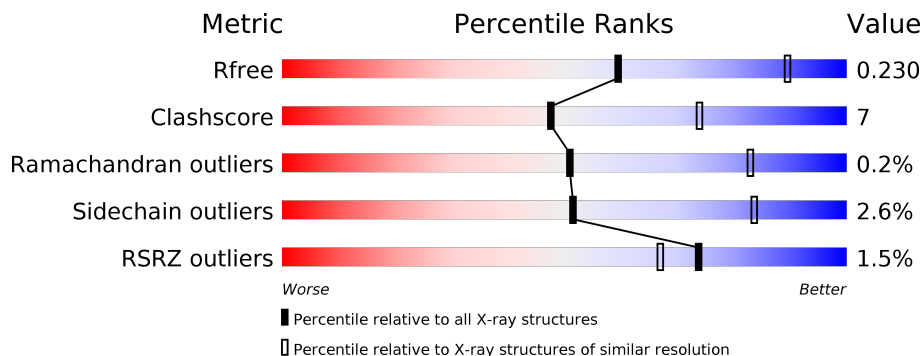|		
PDB ID	:	4NY0
Title	:	Crystal structure of FERM domain of human focal adhesion kinase
Authors	:	Walkiewicz, K.; Arold, S.T.
Deposited on	:	2013-12-10
Resolution	:	2.80 Å(reported)

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	Percentile Ranks	Value
Rfree		0.230
Clashscore		7
Ramachandran outliers		0.2%
Sidechain outliers		2.6%
RSRZ outliers		1.5%

Worse — *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

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	376	74% 13% • 13%
1	B	376	2% 68% 18% • 13%
1	C	376	% 74% 14% 11%
1	D	376	2% 72% 15% 13%

2   Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 10722 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 Focal adhesion kinase 1.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	329	Total	C	N	O	S	0	0	0
			2668	1704	457	496	11			
1	A	329	Total	C	N	O	S	0	0	0
			2668	1704	457	496	11			
1	B	329	Total	C	N	O	S	0	0	0
			2668	1704	457	496	11			
1	C	335	Total	C	N	O	S	0	0	0
			2718	1734	463	510	11			

There are 12 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
D	30	MET	-	EXPRESSION TAG	UNP Q05397
D	85	LEU	PHE	ENGINEERED MUTATION	UNP Q05397
D	181	GLY	TRP	ENGINEERED MUTATION	UNP Q05397
A	30	MET	-	EXPRESSION TAG	UNP Q05397
A	85	LEU	PHE	ENGINEERED MUTATION	UNP Q05397
A	181	GLY	TRP	ENGINEERED MUTATION	UNP Q05397
B	30	MET	-	EXPRESSION TAG	UNP Q05397
B	85	LEU	PHE	ENGINEERED MUTATION	UNP Q05397
B	181	GLY	TRP	ENGINEERED MUTATION	UNP Q05397
C	30	MET	-	EXPRESSION TAG	UNP Q05397
C	85	LEU	PHE	ENGINEERED MUTATION	UNP Q05397
C	181	GLY	TRP	ENGINEERED MUTATION	UNP Q05397

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.
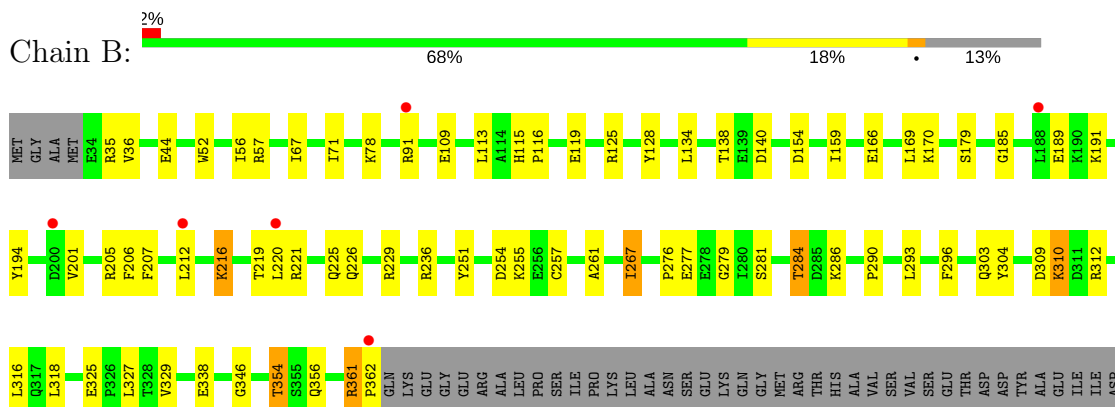
• Molecule 1: Focal adhesion kinase 1

Chain D:

• Molecule 1: Focal adhesion kinase 1

Chain A:

• Molecule 1: Focal adhesion kinase 1

Chain B:

GLU
GLU
ASP

• Molecule 1: Focal adhesion kinase 1

Chain C:
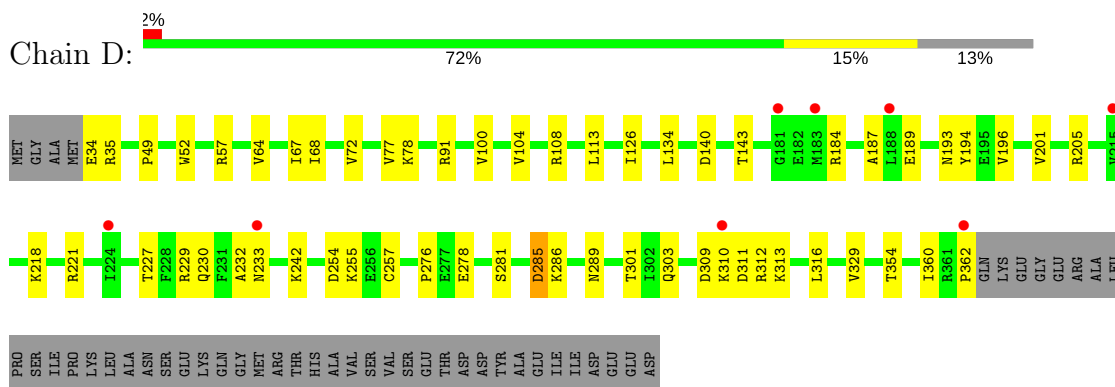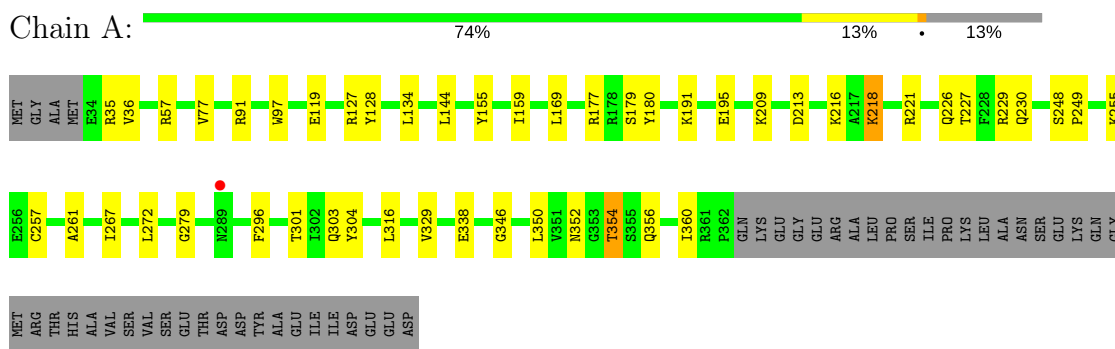
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	71.96Å 152.50Å 95.34Å 90.00° 93.62° 90.00°	Depositor
Resolution (Å)	95.15 – 2.80 95.15 – 2.80	Depositor EDS
% Data completeness (in resolution range)	97.0 (95.15-2.80) 97.1 (95.15-2.80)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.76 (at 2.82Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.193 , 0.230 0.193 , 0.230	Depositor DCC
R_{free} test set	2478 reflections (5.06%)	DCC
Wilson B-factor (Å^2)	74.8	Xtriage
Anisotropy	0.024	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 39.8	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	10722	wwPDB-VP
Average B, all atoms (Å^2)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.82% 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	0.53	0/2726	0.68	0/3680
1	B	0.48	0/2726	0.68	0/3680
1	C	0.52	0/2775	0.70	0/3745
1	D	0.51	0/2726	0.68	0/3680
All	All	0.51	0/10953	0.69	0/14785

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	2668	0	2663	31	0
1	B	2668	0	2663	53	0
1	C	2718	0	2704	27	0
1	D	2668	0	2663	42	0
All	All	10722	0	10693	150	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 7.

The worst 5 of 150 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:218:LYS:HG2	1:D:221:ARG:HH12	1.16	1.10
1:B:255:LYS:NZ	1:B:257:CYS:SG	2.39	0.95
1:B:361:ARG:HH11	1:B:361:ARG:HG2	1.28	0.95
1:D:218:LYS:HG2	1:D:221:ARG:NH1	1.83	0.91
1:D:255:LYS:NZ	1:D:257:CYS:SG	2.43	0.90

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	327/376 (87%)	304 (93%)	23 (7%)	0	100	100
1	B	327/376 (87%)	304 (93%)	22 (7%)	1 (0%)	44	77
1	C	331/376 (88%)	302 (91%)	27 (8%)	2 (1%)	28	62
1	D	327/376 (87%)	305 (93%)	22 (7%)	0	100	100
All	All	1312/1504 (87%)	1215 (93%)	94 (7%)	3 (0%)	51	83

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	185	GLY
1	C	185	GLY
1	C	46	ASN

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	294/334 (88%)	287 (98%)	7 (2%)	54	85
1	B	294/334 (88%)	283 (96%)	11 (4%)	39	73
1	C	300/334 (90%)	291 (97%)	9 (3%)	46	80
1	D	294/334 (88%)	290 (99%)	4 (1%)	71	92
All	All	1182/1336 (88%)	1151 (97%)	31 (3%)	51	83

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	267	ILE
1	B	309	ASP
1	C	354	THR
1	B	277	GLU
1	B	310	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	225	GLN
1	A	303	GLN
1	B	289	ASN
1	B	303	GLN

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	329/376 (87%)	0.13	1 (0%)	93	92	40, 64, 98, 121	0
1	B	329/376 (87%)	0.16	6 (1%)	69	60	48, 79, 127, 142	0
1	C	335/376 (89%)	0.20	5 (1%)	74	67	42, 75, 126, 140	0
1	D	329/376 (87%)	0.21	8 (2%)	59	49	42, 66, 125, 135	0
All	All	1322/1504 (87%)	0.18	20 (1%)	74	67	40, 71, 124, 142	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	188	LEU	4.7
1	C	188	LEU	3.9
1	D	310	LYS	3.7
1	B	188	LEU	3.7
1	C	200	ASP	3.2

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.

6.5 Other polymers ⓘ

There are no such residues in this entry.