



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 10:32 PM UTC

PDB ID : 3B71 / pdb_00003b71
Title : CD4 endocytosis motif bound to the Focal Adhesion Targeting (FAT) domain of the Focal Adhesion Kinase
Authors : Garron, M.-L.; Arold, S.T.
Deposited on : 2007-10-30
Resolution : 2.82 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4-5-2 with Phenix2.0
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

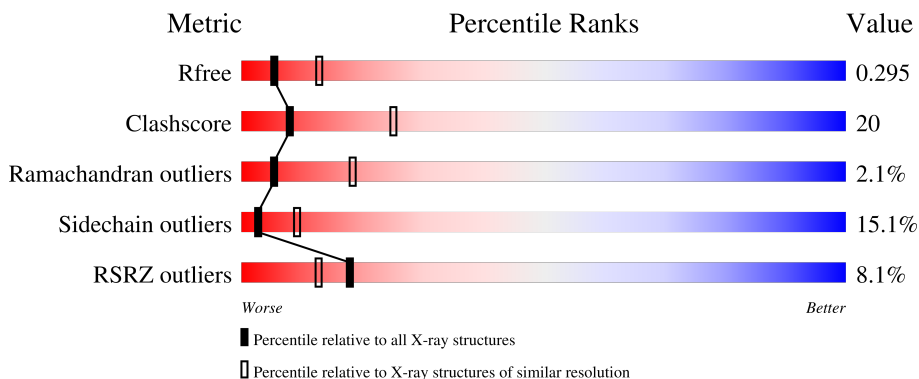
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.82 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}	180053	4591 (2.84-2.80)
Clashscore	190562	5010 (2.84-2.80)
Ramachandran outliers	187476	4916 (2.84-2.80)
Sidechain outliers	187428	4918 (2.84-2.80)
RSRZ outliers	180081	4594 (2.84-2.80)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	162	 49% 28% 17% 3% 2%
1	B	162	 46% 33% 14% 7%
1	C	162	 48% 26% 14% 10% 3%
2	D	23	 17% 13% 17% 52%
2	E	23	 17% 22% 9% 52% 43%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain										
2	F	23	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into four segments: a green segment (9%), a yellow segment (17%), a red segment (13%), and a grey segment (61%). The total length of the bar is 100%. The segments are stacked from left to right in the order: green, yellow, red, grey.</p> <table border="1"><thead><tr><th>Color</th><th>Percentage</th></tr></thead><tbody><tr><td>Green</td><td>9%</td></tr><tr><td>Yellow</td><td>17%</td></tr><tr><td>Red</td><td>13%</td></tr><tr><td>Grey</td><td>61%</td></tr></tbody></table>	Color	Percentage	Green	9%	Yellow	17%	Red	13%	Grey	61%
Color	Percentage												
Green	9%												
Yellow	17%												
Red	13%												
Grey	61%												

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3454 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
			Total	C	N	O	S			
1	A	134	1048	662	178	201	7	0	1	0
1	B	139	1073	680	180	206	7	0	0	0
1	C	139	1073	680	180	206	7	0	0	0

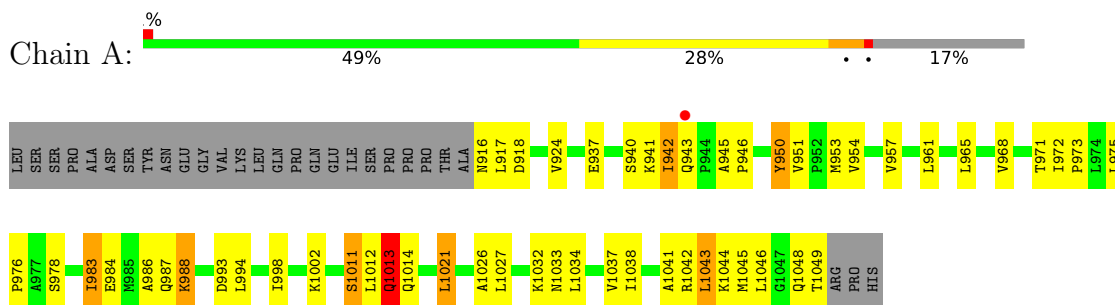
- Molecule 2 is a protein called T-cell surface glycoprotein CD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	11	93	57	19	16	1	0	0	0
2	E	11	93	57	19	16	1	0	0	0
2	F	9	74	46	14	14		0	0	0

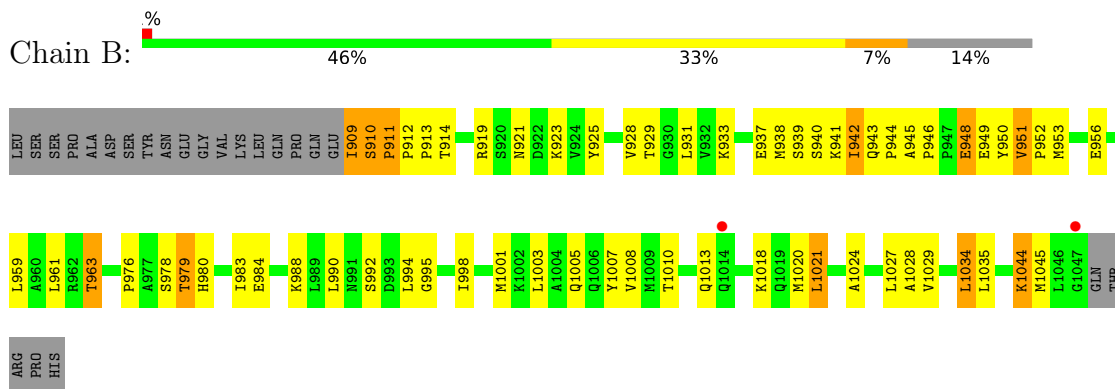
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

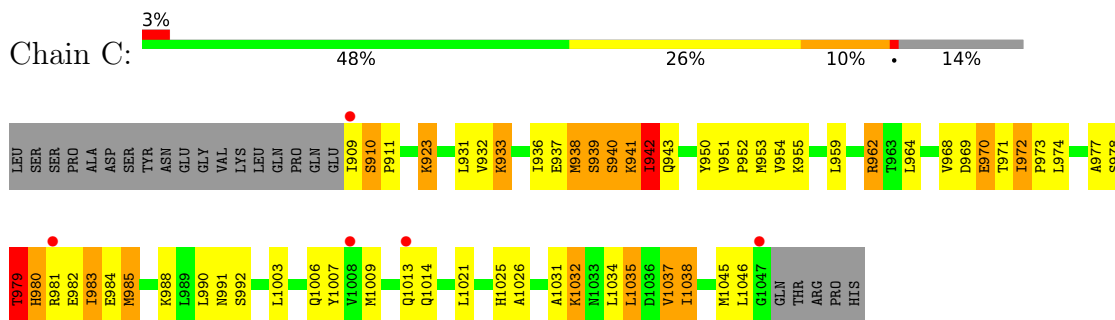
- Molecule 1: Focal adhesion kinase 1



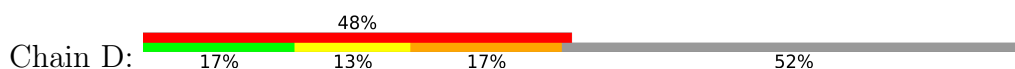
- Molecule 1: Focal adhesion kinase 1

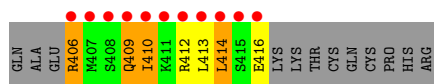


- Molecule 1: Focal adhesion kinase 1

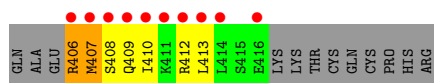
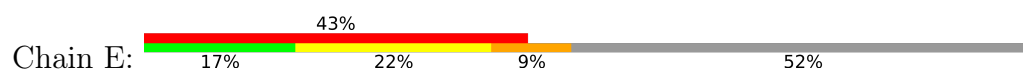


- Molecule 2: T-cell surface glycoprotein CD4

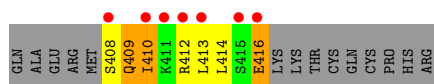




- Molecule 2: T-cell surface glycoprotein CD4



- Molecule 2: T-cell surface glycoprotein CD4



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	88.07Å 220.47Å 97.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.00 – 2.82 29.00 – 2.82	Depositor EDS
% Data completeness (in resolution range)	93.8 (29.00-2.82) 93.8 (29.00-2.82)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.33 (at 2.80Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.241 , 0.307 0.233 , 0.295	Depositor DCC
R_{free} test set	459 reflections (1.96%)	wwPDB-VP
Wilson B-factor (Å ²)	69.3	Xtrriage
Anisotropy	0.214	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 43.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	3454	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.66% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	1.06	3/1064 (0.3%)	1.21	8/1439 (0.6%)
1	B	1.13	4/1087 (0.4%)	1.28	6/1474 (0.4%)
1	C	0.96	1/1087 (0.1%)	1.23	5/1474 (0.3%)
2	D	0.69	0/92	1.14	0/119
2	E	0.87	0/92	0.93	0/119
2	F	1.13	0/73	1.11	1/95 (1.1%)
All	All	1.04	8/3495 (0.2%)	1.23	20/4720 (0.4%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	951	VAL	CA-CB	-7.39	1.50	1.54
1	C	979	THR	CA-CB	5.98	1.57	1.53
1	A	954	VAL	CA-CB	-5.73	1.47	1.54
1	B	911	PRO	CA-C	5.62	1.54	1.51
1	B	1024	ALA	CA-CB	-5.48	1.44	1.53

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	911	PRO	CA-C-N	-8.27	111.86	120.38
1	B	911	PRO	C-N-CA	-8.27	111.86	120.38
1	C	980	HIS	N-CA-C	-7.36	104.28	113.55
1	B	910	SER	N-CA-C	6.71	119.03	108.55
1	A	946	PRO	CA-C-N	6.46	126.39	119.28

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	1048	0	1096	26	0
1	B	1073	0	1127	47	0
1	C	1073	0	1127	47	0
2	D	93	0	104	10	0
2	E	93	0	104	10	0
2	F	74	0	82	4	0
All	All	3454	0	3640	139	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 20.

The worst 5 of 139 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:950:TYR:HA	1:B:953:MET:HE2	1.37	1.03
1:C:942:ILE:HG23	1:C:943:GLN:N	1.82	0.94
1:C:977:ALA:HA	1:C:980:HIS:NE2	1.84	0.91
1:C:1003:LEU:O	1:C:1006:GLN:HB2	1.71	0.91
1:B:950:TYR:HA	1:B:953:MET:CE	2.00	0.90

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	133/162 (82%)	123 (92%)	8 (6%)	2 (2%)	8 26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	137/162 (85%)	125 (91%)	11 (8%)	1 (1%)	18	45
1	C	137/162 (85%)	115 (84%)	16 (12%)	6 (4%)	2	6
2	D	9/23 (39%)	9 (100%)	0	0	100	100
2	E	9/23 (39%)	3 (33%)	6 (67%)	0	100	100
2	F	7/23 (30%)	4 (57%)	3 (43%)	0	100	100
All	All	432/555 (78%)	379 (88%)	44 (10%)	9 (2%)	5	18

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	942	ILE
1	C	978	SER
1	A	1013	GLN
1	C	970	GLU
1	C	923	LYS

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	118/142 (83%)	106 (90%)	12 (10%)	7	22
1	B	121/142 (85%)	107 (88%)	14 (12%)	5	17
1	C	121/142 (85%)	97 (80%)	24 (20%)	1	4
2	D	11/22 (50%)	7 (64%)	4 (36%)	0	0
2	E	11/22 (50%)	9 (82%)	2 (18%)	2	5
2	F	9/22 (41%)	6 (67%)	3 (33%)	0	0
All	All	391/492 (80%)	332 (85%)	59 (15%)	3	9

5 of 59 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	933	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	E	407	MET
1	C	955	LYS
2	E	406	ARG
1	C	1046	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	927	ASN
1	B	1006	GLN
1	C	991	ASN
1	C	1005	GLN
2	D	409	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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, 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(Å ²)	Q<0.9
1	A	134/162 (82%)	0.03	1 (0%) 84 78	20, 36, 48, 53	1 (0%)
1	B	139/162 (85%)	-0.15	2 (1%) 73 64	26, 35, 43, 49	0
1	C	139/162 (85%)	0.33	5 (3%) 46 37	26, 37, 48, 56	0
2	D	11/23 (47%)	4.75	11 (100%) 0 0	53, 54, 58, 58	11 (100%)
2	E	11/23 (47%)	3.26	10 (90%) 0 0	40, 44, 48, 48	11 (100%)
2	F	9/23 (39%)	3.25	7 (77%) 0 0	34, 36, 38, 40	9 (100%)
All	All	443/555 (79%)	0.33	36 (8%) 18 13	20, 36, 50, 58	32 (7%)

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	407	MET	6.7
2	D	411	LYS	6.0
2	E	407	MET	5.9
2	D	415	SER	5.8
2	D	414	LEU	5.5

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.