wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 27, 2014 – 01:08 PM GMT

PDB ID	:	3PCS
Title	:	Structure of EspG-PAK2 autoinhibitory Ialpha3 helix complex
Authors	:	Tomchick, D.R.; Alto, N.M.; Selyunin, A.S.
Deposited on	:	2010-10-21
Resolution	:	2.86 Å(reported)

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

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.86 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		0.283
Clashscore		5
Ramachandran outliers		1.6%
Sidechain outliers		4.8%
RSRZ outliers		1.2%

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
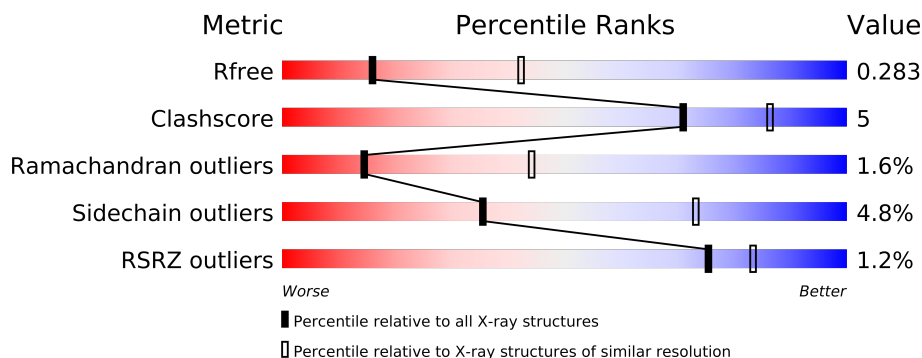
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1524 (2.90-2.82)
Clashscore	79885	1879 (2.90-2.82)
Ramachandran outliers	78287	1824 (2.90-2.82)
Sidechain outliers	78261	1827 (2.90-2.82)
RSRZ outliers	66119	1526 (2.90-2.82)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	357	
1	B	357	
1	C	357	
1	D	357	
2	E	16	
2	F	16	
2	G	16	
2	H	16	

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 11254 atoms, of which 0 are hydrogen and 0 are deuterium.

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 EspG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	350	Total	C	N	O	S	0	0	0
			2719	1688	478	539	14			
1	B	355	Total	C	N	O	S	0	0	0
			2757	1709	482	551	15			
1	C	345	Total	C	N	O	S	0	0	0
			2678	1666	468	530	14			
1	D	348	Total	C	N	O	S	0	0	0
			2704	1680	474	536	14			

- Molecule 2 is a protein called Serine/threonine-proteinkinase PAK 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	14	Total	C	N	O	0	0	0
			111	73	16	22			
2	F	12	Total	C	N	O	0	0	0
			97	64	14	19			
2	G	10	Total	C	N	O	0	0	0
			84	57	11	16			
2	H	13	Total	C	N	O	0	0	0
			104	68	15	21			

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 errors 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.
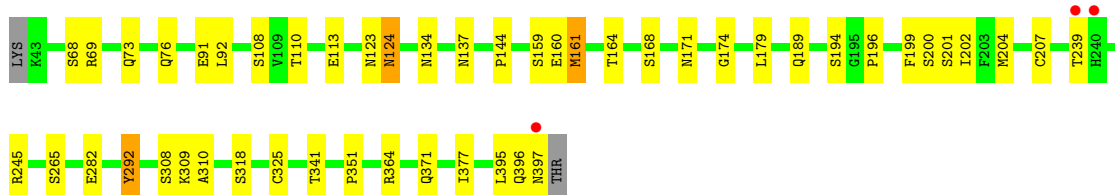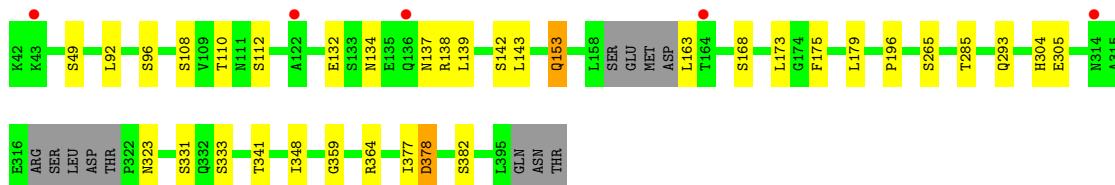
- Molecule 1: EspG

Chain A:
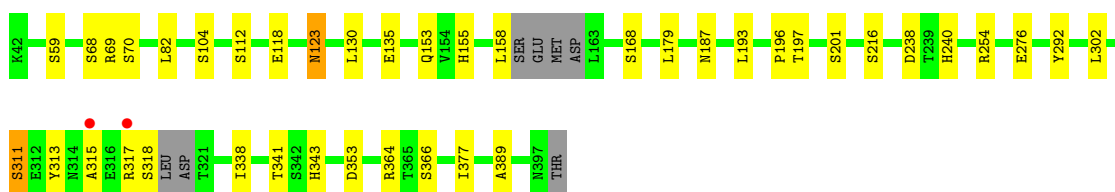
- Molecule 1: EspG
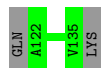
Chain B:

- Molecule 1: EspG

Chain C:

- Molecule 1: EspG

Chain D:

• Molecule 2: Serine/threonine-proteinkinase PAK 2

Chain E:

• Molecule 2: Serine/threonine-proteinkinase PAK 2

Chain F:

• Molecule 2: Serine/threonine-proteinkinase PAK 2

Chain G:

• Molecule 2: Serine/threonine-proteinkinase PAK 2

Chain H:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	86.72Å 104.60Å 191.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.50 – 2.86 40.50 – 2.78	Depositor EDS
% Data completeness (in resolution range)	99.9 (40.50-2.86) 99.1 (40.50-2.78)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.25 (at 2.77Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.6.4_486)	Depositor
R, R_{free}	0.191 , 0.281 0.191 , 0.283	Depositor DCC
R_{free} test set	2221 reflections (5.03%)	DCC
Wilson B-factor (Å^2)	63.3	Xtriage
Anisotropy	0.452	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 28.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Outliers	1 of 44193 reflections (0.002%)	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11254	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 3.50% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

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	0.45	0/2767	0.66	0/3753
1	B	0.42	0/2807	0.63	0/3811
1	C	0.37	0/2726	0.58	0/3697
1	D	0.40	0/2752	0.58	0/3733
2	E	0.51	0/112	0.65	0/152
2	F	0.35	0/98	0.63	0/132
2	G	0.45	0/85	0.54	0/114
2	H	0.34	0/105	0.63	0/142
All	All	0.41	0/11452	0.61	0/15534

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2719	0	0	13	0
1	B	2757	0	12	17	0
1	C	2678	0	0	11	0
1	D	2704	0	0	10	0
2	E	111	0	8	0	0
2	F	97	0	5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	G	84	0	4	3	0
2	H	104	0	6	0	0
All	All	11254	0	35	51	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

The worst 5 of 51 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:179:LEU:O	1:B:364:ARG:NH2	2.26	0.68
1:D:179:LEU:O	1:D:364:ARG:NH2	2.33	0.62
1:B:161:MET:CG	1:B:161:MET:O	2.52	0.57
1:B:137:ASN:OD1	1:B:137:ASN:O	2.25	0.55
1:C:378:ASP:OD1	1:C:378:ASP:N	2.40	0.55

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	344/357 (96%)	312 (91%)	28 (8%)	4 (1%)	19	54
1	B	353/357 (99%)	314 (89%)	29 (8%)	10 (3%)	8	27
1	C	339/357 (95%)	313 (92%)	21 (6%)	5 (2%)	15	46
1	D	342/357 (96%)	308 (90%)	30 (9%)	4 (1%)	19	54
2	E	12/16 (75%)	11 (92%)	1 (8%)	0	100	100
2	F	10/16 (62%)	10 (100%)	0	0	100	100
2	G	8/16 (50%)	7 (88%)	1 (12%)	0	100	100
2	H	11/16 (69%)	10 (91%)	1 (9%)	0	100	100
All	All	1419/1492 (95%)	1285 (91%)	111 (8%)	23 (2%)	14	44

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	124	ASN
1	B	377	ILE
1	C	134	ASN
1	D	351	PRO
1	D	377	ILE

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	310/317 (98%)	291 (94%)	19 (6%)	26	59
1	B	315/317 (99%)	300 (95%)	15 (5%)	35	73
1	C	305/317 (96%)	289 (95%)	16 (5%)	32	69
1	D	308/317 (97%)	299 (97%)	9 (3%)	55	88
2	E	13/15 (87%)	13 (100%)	0	100	100
2	F	11/15 (73%)	11 (100%)	0	100	100
2	G	10/15 (67%)	9 (90%)	1 (10%)	11	30
2	H	12/15 (80%)	10 (83%)	2 (17%)	3	8
All	All	1284/1328 (97%)	1222 (95%)	62 (5%)	35	73

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	245	ARG
1	C	49	SER
1	D	329	SER
1	B	341	THR
1	C	96	SER

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 chains 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	350/357 (98%)	-0.44	2 (0%)	86	91	26, 60, 104, 159	0
1	B	355/357 (99%)	-0.42	3 (0%)	83	89	39, 68, 129, 180	0
1	C	345/357 (96%)	-0.22	5 (1%)	72	80	41, 77, 126, 149	0
1	D	348/357 (97%)	-0.23	8 (2%)	57	66	44, 73, 142, 183	0
2	E	14/16 (87%)	-0.05	0	100	100	61, 84, 99, 102	0
2	F	12/16 (75%)	-0.03	0	100	100	64, 86, 111, 130	0
2	G	10/16 (62%)	0.87	0	100	100	95, 124, 142, 143	0
2	H	13/16 (81%)	-0.61	0	100	100	74, 92, 105, 130	0
All	All	1447/1492 (96%)	-0.32	18 (1%)	75	83	26, 70, 129, 183	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	316	GLU	6.9
1	B	397	ASN	3.8
1	D	240	HIS	3.3
1	A	317	ARG	3.2
1	D	151	TYR	3.0

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.