wwPDB X-ray Structure Validation Summary Report (i)

Aug 30, 2023 – 09:09 AM EDT

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)

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 (i) 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 (1)) were used in the production of this report:

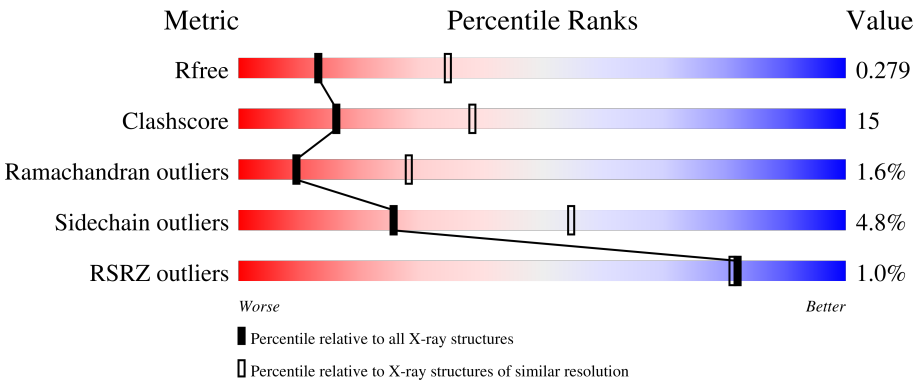
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (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 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 <=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	357	% 65% 30% • •
1	B	357	% 67% 30% • •
1	C	357	% 67% 28% • •
1	D	357	2% 67% 31% •
2	E	16	62% 25% 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	16	62% 12% 25%
2	G	16	25% 31% 6% 38%
2	H	16	38% 38% 6% 19%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 11254 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 EspG.

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

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

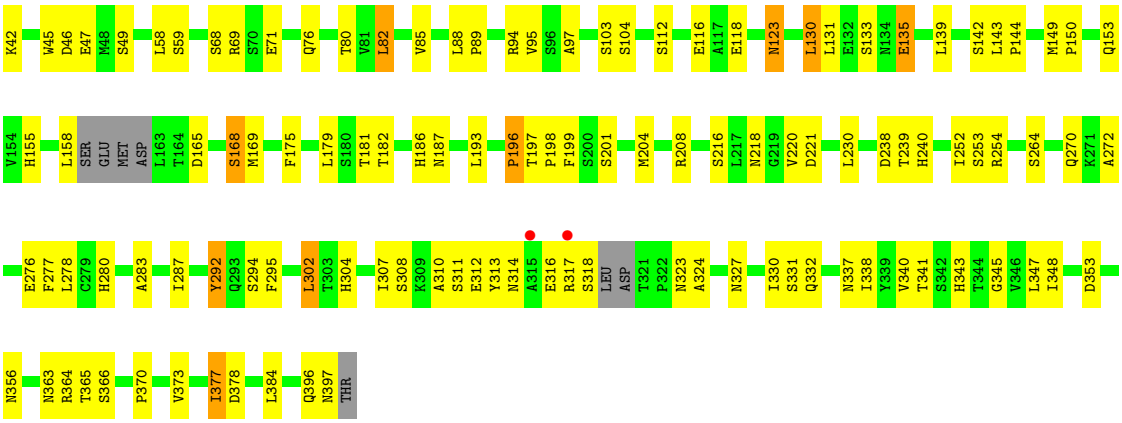
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	E	14	111	73	16	22	0	0	0
2	F	12	97	64	14	19	0	0	0
2	G	10	84	57	11	16	0	0	0
2	H	13	104	68	15	21	0	0	0

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.

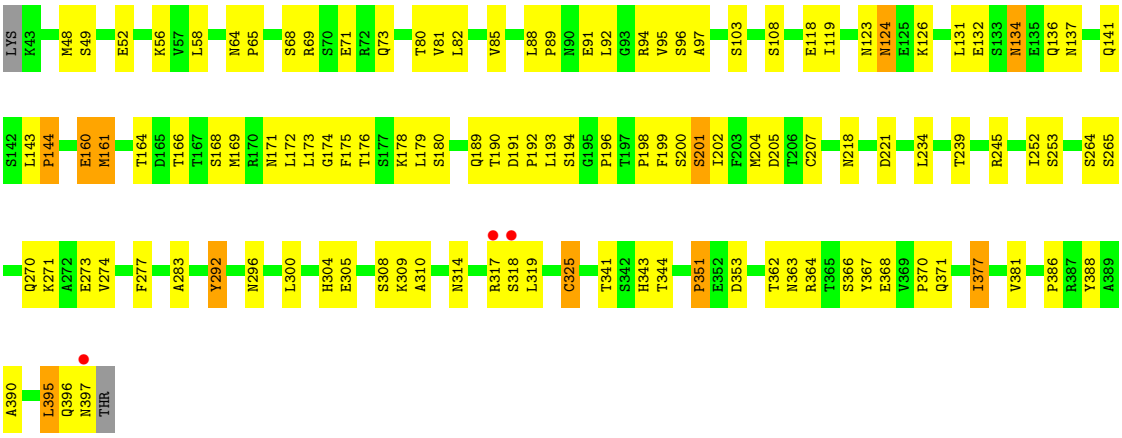
• Molecule 1: EspG

Chain A:



• Molecule 1: EspG

Chain B:



• Molecule 1: EspG

Chain C:

K42 M48 S49 E52 K56 S68 R69 R72 Q76 V81 L82 L88 P89 L92 V95 S96 A97 S108 V115 E132 S133 N134 E135 Q136 N137 R138 L139 L140 Q141 S142 L143 P144 M149 P150 Q153 L158 SER GLU MET ASP L163 T164 D165 T166

T167 S168 M169 L172 L173 G174 F175 T176 S177 K178 A188 Q189 P192 L193 P196 T197 P198 F199 S200 S201 I202 F203 M204 C207 R208 G209 N212 A224 Q227 K228 L229 L230 R231 T239 E263 S264 S265 F277 E282 A283 A284 T285 A286 S289 A290 F291 Y292

Q293 V297 P298 A299 L300 H304 E305 R306 I307 S308 K309 A310 E316 ARG SER LEU ASP THR P322 N323 A324 C325 I326 N327 I330 S331 Q332 S333 I338 T341 S342 H343 T344 G345 V346 L347 I348 M349 P355 N356 R364 T365 S366 Y367 I377 D378 S382 P386 R387

Y388 L395 GLN ASN THR

- Molecule 1: EspG

Chain D: 2% 67% 31%

K42 A51 E52 K53 L54 F55 K56 V57 L58 S59 F60 Y67 S68 R69 S70 E71 R72 L82 L88 P89 V95 S96 L105 S112 E113 L114 R120 N123 K126 I127 T128 V129 L130 L131 E132 N137 S142 L143 P144 I145 D146 Y151 H155 R156 A157 L158

SER GLU MET ASP L163 T167 N171 F175 L179 S180 T181 T182 L183 I184 L193 T197 P198 F199 S200 S201 I202 F203 M204 D205 R208 S216 L217 N218 G219 V220 L230 L234 T239 H240 S241 S242 P243 T244 R245 N246 V247 I248 I252 I260 A261 R262 E263

S264 S265 Q270 K271 A272 E273 E276 S289 Y292 Q293 S294 V297 P298 A299 L300 H304 E305 R306 I307 S311 A315 E316 R317 SER LEU D320 N323 A324 S329 H343 V346 M349 A350 P351 G359 T362 M363 R364 T365 S366 Y367 I377 M380

L384 Q385 P386 A389 A390 S391 Y394 L395 GLN ASN THR

- Molecule 2: Serine/threonine-protein kinase PAK 2

Chain E: 62% 25% 12%

GLN A122 V123 L124 D125 V126 L127 K128 V135 LYS

- Molecule 2: Serine/threonine-protein kinase PAK 2

Chain F: 62% 12% 25%

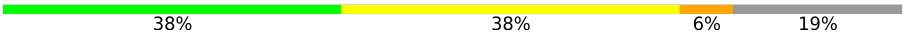
GLN A122 V123 V126 N133 THR VAL LYS

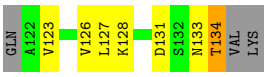
- Molecule 2: Serine/threonine-protein kinase PAK 2

Chain G: 25% 31% 6% 38%

- Molecule 2: Serine/threonine-protein kinase PAK 2

Chain H:





4 Data and refinement statistics (i)

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 1.6.4_486	Depositor
R, R_{free}	0.191 , 0.281 0.185 , 0.279	Depositor DCC
R_{free} test set	2224 reflections (5.03%)	wwPDB-VP
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 , 50.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	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.

[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 (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	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 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	2719	0	2686	89	0
1	B	2757	0	2714	90	0
1	C	2678	0	2648	78	0
1	D	2704	0	2671	78	0
2	E	111	0	110	3	0
2	F	97	0	94	1	0
2	G	84	0	83	7	0
2	H	104	0	101	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	11254	0	11107	325	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 15.

The worst 5 of 325 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:204:MET:HG3	1:B:239:THR:CG2	1.98	0.93
1:C:153:GLN:HE21	1:C:153:GLN:HA	1.35	0.90
1:D:82:LEU:HG	1:D:95:VAL:HG11	1.57	0.86
1:C:367:TYR:HB3	1:C:386:PRO:HA	1.61	0.82
1:B:189:GLN:HE22	1:C:305:GLU:HB2	1.46	0.79

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	344/357 (96%)	312 (91%)	28 (8%)	4 (1%)	13 35
1	B	353/357 (99%)	314 (89%)	29 (8%)	10 (3%)	5 16
1	C	339/357 (95%)	313 (92%)	21 (6%)	5 (2%)	10 30
1	D	342/357 (96%)	308 (90%)	30 (9%)	4 (1%)	13 35
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%)	9 28

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 (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	310/317 (98%)	291 (94%)	19 (6%)	18	43
1	B	315/317 (99%)	300 (95%)	15 (5%)	25	55
1	C	305/317 (96%)	289 (95%)	16 (5%)	23	51
1	D	308/317 (97%)	299 (97%)	9 (3%)	42	72
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%)	7	21
2	H	12/15 (80%)	10 (83%)	2 (17%)	2	5
All	All	1284/1328 (97%)	1222 (95%)	62 (5%)	25	55

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

Mol	Chain	Res	Type
1	B	245	ARG
1	D	265	SER
1	C	92	LEU
1	D	245	ARG
2	G	132	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	255	HIS
1	D	332	GLN
1	C	141	GLN
1	C	153	GLN
1	C	189	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 monosaccharides 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	350/357 (98%)	-0.46	2 (0%)	89	89	26, 60, 104, 159	0
1	B	355/357 (99%)	-0.41	3 (0%)	86	85	39, 68, 129, 180	0
1	C	345/357 (96%)	-0.26	3 (0%)	84	84	41, 77, 126, 149	0
1	D	348/357 (97%)	-0.22	7 (2%)	65	62	44, 73, 142, 183	0
2	E	14/16 (87%)	-0.04	0	100	100	61, 84, 99, 102	0
2	F	12/16 (75%)	-0.12	0	100	100	64, 86, 111, 130	0
2	G	10/16 (62%)	0.98	0	100	100	95, 124, 142, 143	0
2	H	13/16 (81%)	-0.60	0	100	100	74, 92, 105, 130	0
All	All	1447/1492 (96%)	-0.33	15 (1%)	82	81	26, 70, 129, 183	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	316	GLU	6.3
1	D	240	HIS	4.0
1	D	151	TYR	3.4
1	B	397	ASN	3.2
1	C	164	THR	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 monosaccharides 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.