wwPDB X-ray Structure Validation Summary Report (i)

Nov 12, 2024 – 10:17 AM EST

PDB ID	:	3HYM
Title	:	Insights into Anaphase Promoting Complex TPR subdomain assembly from a CDC26-APC6 structure
Authors	:	Wang, J.; Dye, B.T.; Rajashankar, K.R.; Kurinov, I.; Schulman, B.A.
Deposited on	:	2009-06-22
Resolution	:	2.80 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 (i)) were used in the production of this report:

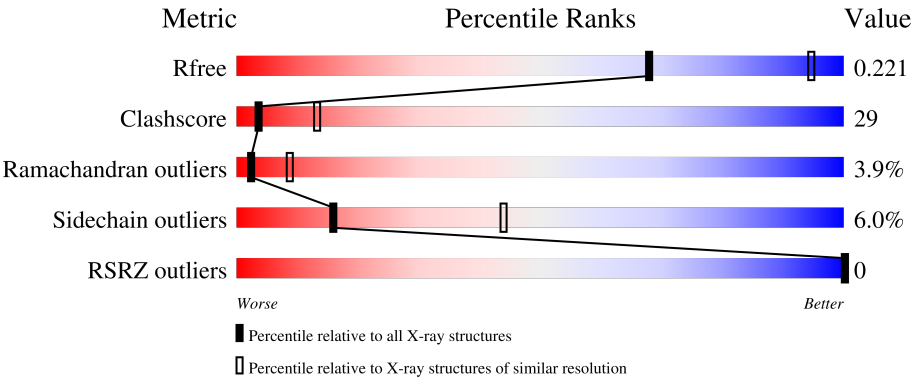
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.80 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}	164625	3657 (2.80-2.80)
Clashscore	180529	4123 (2.80-2.80)
Ramachandran outliers	177936	4071 (2.80-2.80)
Sidechain outliers	177891	4073 (2.80-2.80)
RSRZ outliers	164620	3659 (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 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	29	41% 38% 10% 10%
1	C	29	21% 41% 17% 21%
1	E	29	41% 41% 7% 10%
1	G	29	41% 38% 7% 14%
1	I	29	41% 45% . 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	29	34% 34% 10% 21%
2	B	330	50% 36% 5% 9%
2	D	330	44% 43% 9%
2	F	330	54% 34% 9%
2	H	330	55% 32% 8%
2	J	330	42% 43% 5% 9%
2	L	330	45% 42% 9%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 15517 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 Anaphase-promoting complex subunit CDC26.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	26	Total	C	N	O	Se	0	0	0
			214	134	41	38	1			
1	C	23	Total	C	N	O	Se	0	0	0
			199	125	38	35	1			
1	E	26	Total	C	N	O	Se	0	0	0
			208	131	38	38	1			
1	G	25	Total	C	N	O	Se	0	0	0
			213	133	40	39	1			
1	I	26	Total	C	N	O	Se	0	0	0
			208	131	38	38	1			
1	K	23	Total	C	N	O	Se	0	0	0
			197	124	35	37	1			

- Molecule 2 is a protein called Cell division cycle protein 16 homolog.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	301	Total	C	N	O	S	Se	0	0	0
			2344	1514	383	430	7	10			
2	D	301	Total	C	N	O	S	Se	0	0	0
			2348	1517	384	430	7	10			
2	F	301	Total	C	N	O	S	Se	0	0	0
			2356	1520	389	430	7	10			
2	H	302	Total	C	N	O	S	Se	0	0	0
			2345	1513	382	433	7	10			
2	J	301	Total	C	N	O	S	Se	0	0	0
			2346	1514	385	430	7	10			
2	L	299	Total	C	N	O	S	Se	0	0	0
			2328	1504	383	424	7	10			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	210	GLY	-	expression tag	UNP Q13042

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	211	SER	-	expression tag	UNP Q13042
D	210	GLY	-	expression tag	UNP Q13042
D	211	SER	-	expression tag	UNP Q13042
F	210	GLY	-	expression tag	UNP Q13042
F	211	SER	-	expression tag	UNP Q13042
H	210	GLY	-	expression tag	UNP Q13042
H	211	SER	-	expression tag	UNP Q13042
J	210	GLY	-	expression tag	UNP Q13042
J	211	SER	-	expression tag	UNP Q13042
L	210	GLY	-	expression tag	UNP Q13042
L	211	SER	-	expression tag	UNP Q13042

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total O 2 2	0	0
3	B	32	Total O 32 32	0	0
3	C	4	Total O 4 4	0	0
3	D	31	Total O 31 31	0	0
3	E	5	Total O 5 5	0	0
3	F	19	Total O 19 19	0	0
3	G	4	Total O 4 4	0	0
3	H	48	Total O 48 48	0	0
3	I	1	Total O 1 1	0	0
3	J	29	Total O 29 29	0	0
3	K	3	Total O 3 3	0	0
3	L	33	Total O 33 33	0	0

3 Residue-property plots (i)

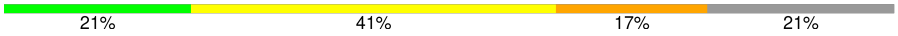
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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.

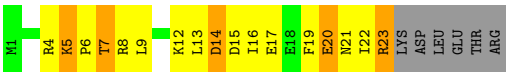
• Molecule 1: Anaphase-promoting complex subunit CDC26

Chain A:

• Molecule 1: Anaphase-promoting complex subunit CDC26

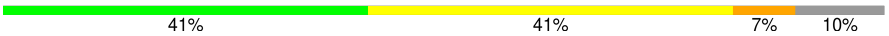
Chain C:

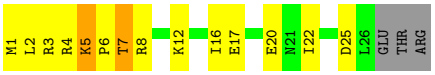




• Molecule 1: Anaphase-promoting complex subunit CDC26

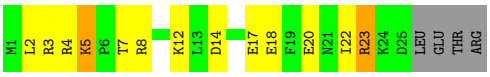
Chain E:





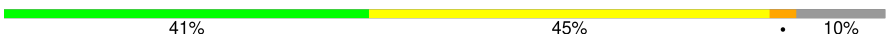
• Molecule 1: Anaphase-promoting complex subunit CDC26

Chain G:



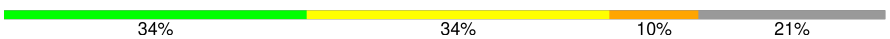
• Molecule 1: Anaphase-promoting complex subunit CDC26

Chain I:



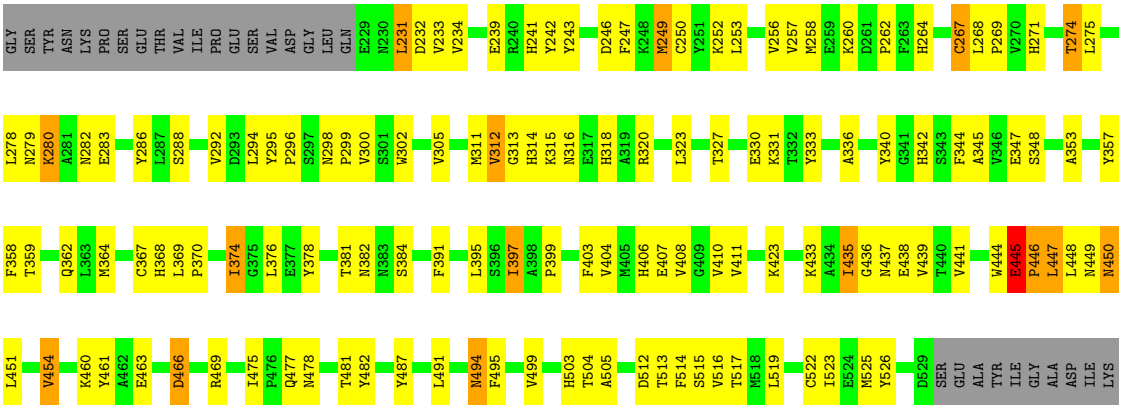
• Molecule 1: Anaphase-promoting complex subunit CDC26

Chain K:



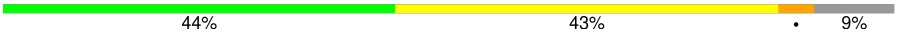
• Molecule 2: Cell division cycle protein 16 homolog

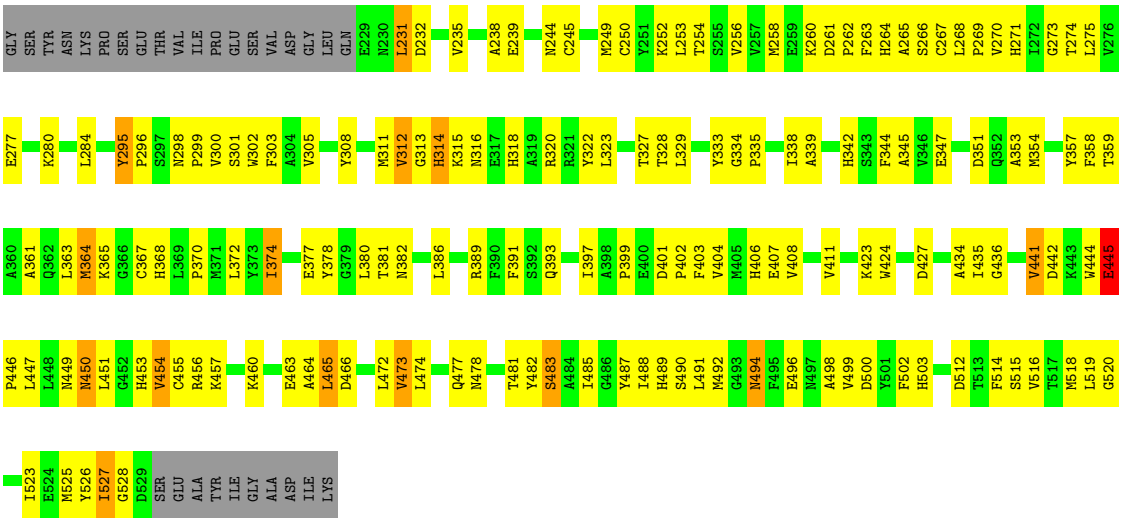
Chain B:



• Molecule 2: Cell division cycle protein 16 homolog

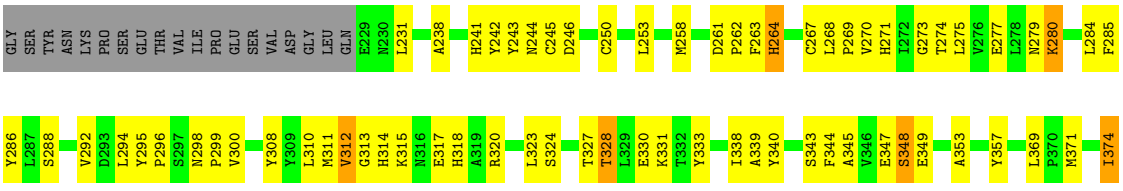
Chain D:

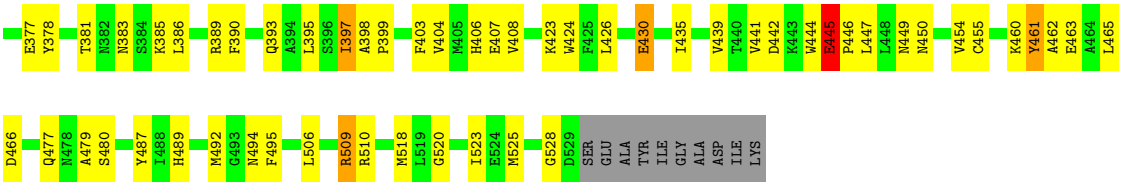




• Molecule 2: Cell division cycle protein 16 homolog

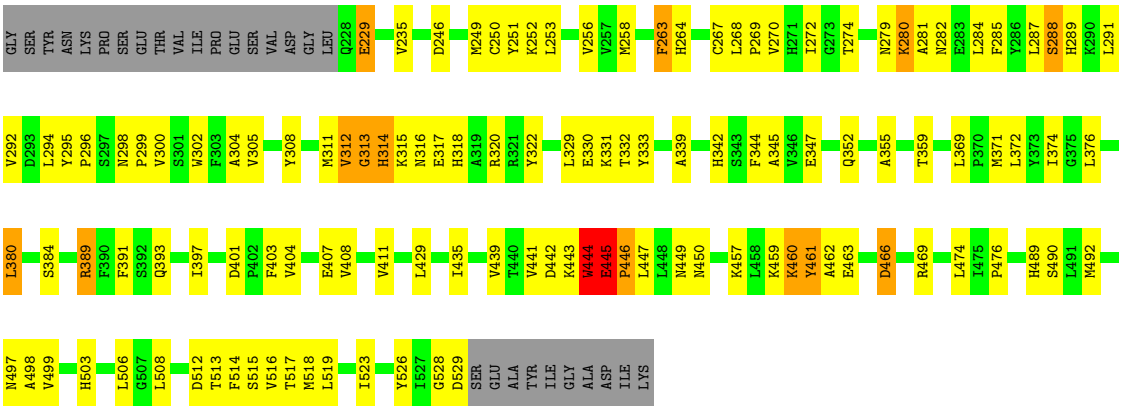
Chain F:





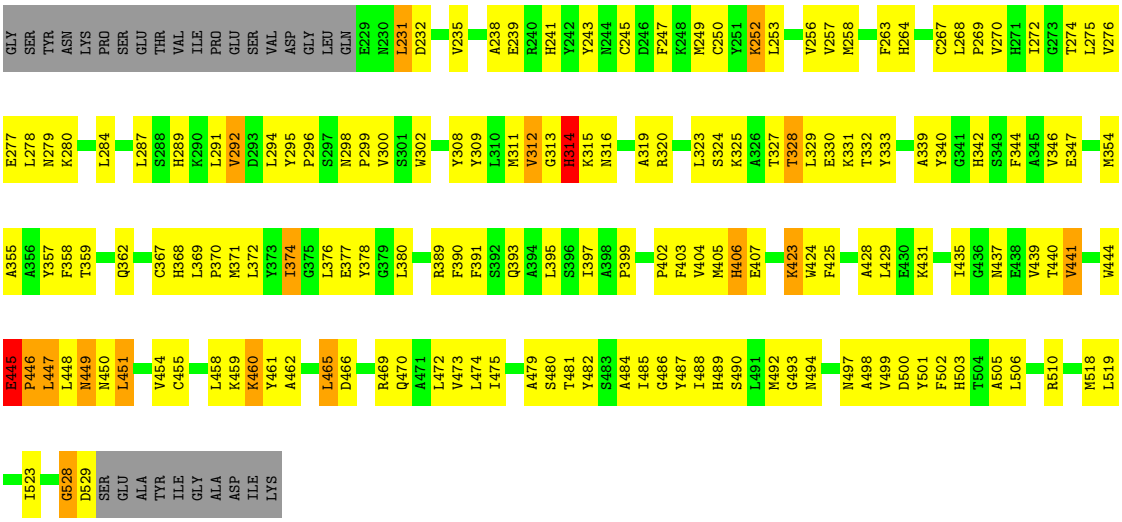
• Molecule 2: Cell division cycle protein 16 homolog

Chain H:



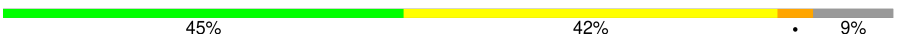
• Molecule 2: Cell division cycle protein 16 homolog

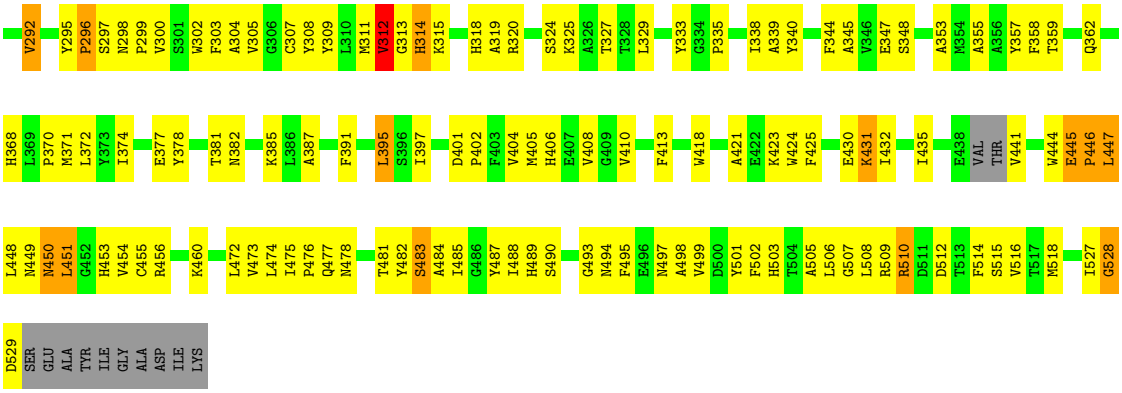
Chain J:



• Molecule 2: Cell division cycle protein 16 homolog

Chain L:




V292
Y295 P296 S297 N298 P299 V300 S301 W302 F303 A304 V305 G306 C307 Y308 Y309 L310 M311 V312 G313 H314 K315
H318 A319 R320
S324 K325 A326 T327 T328 L329
Y333 G334 P335
I338 A339 Y340
F344 A345 V346 E347 S348
A353 M354 A355 A356 Y357 F358 T359
Q362
H368 L369 P370 M371 L372 Y373 I374
E377 Y378
T381 N382
K385 L386 A387
F391
L395 S396 I397
D401 P402 F403 V404 M405 H406 E407 V408 G409 V410
F413
W418
A421 E422 K423 W424 F425
E430 K431 I432
I435
E438 VAL THR V441
W444 E445 P446 L447
L448 N449 N450 L451 G452 H453 V454 C455 R456
K460
L472 V473 L474 I475 P476 Q477 N478
T481 Y482 S483 A484 I485 G486 Y487 I488 H489 S490
G493 N494 F495 E496 N497 A498 V499 D500 Y501 F502 H503 T504 A505 L506 G507 L508 R509 R510 D511 D512 T513 F514 S515 V516 T517 M518
I527 G528
D529 SER GLU ALA TYR ILE GLY ALA ASP ILE LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, α, β, γ	301.90Å 301.90Å 80.17Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.40 – 2.80 49.40 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.3 (49.40-2.80) 98.7 (49.40-2.80)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$ [1]	3.13 (at 2.81Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.188 , 0.220 0.194 , 0.221	Depositor DCC
R_{free} test set	4939 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å^2)	47.2	Xtriage
Anisotropy	0.288	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 46.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.40$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	0.459 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	15517	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.22% 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	0.37	0/215	0.58	0/286
1	C	0.27	0/200	0.49	0/265
1	E	0.32	0/209	0.50	0/279
1	G	0.33	0/214	0.58	0/284
1	I	0.28	0/209	0.51	0/279
1	K	0.33	0/198	0.56	0/263
2	B	0.31	0/2401	0.47	0/3254
2	D	0.32	0/2405	0.48	0/3258
2	F	0.29	0/2413	0.46	0/3268
2	H	0.31	0/2402	0.46	0/3258
2	J	0.32	0/2403	0.49	0/3257
2	L	0.32	0/2384	0.49	0/3230
All	All	0.31	0/15653	0.48	0/21181

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	214	0	218	20	0
1	C	199	0	212	30	0
1	E	208	0	207	18	0
1	G	213	0	220	17	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	I	208	0	207	18	0
1	K	197	0	205	25	0
2	B	2344	0	2194	127	0
2	D	2348	0	2205	163	0
2	F	2356	0	2216	110	0
2	H	2345	0	2178	120	0
2	J	2346	0	2194	157	0
2	L	2328	0	2181	155	0
3	A	2	0	0	0	0
3	B	32	0	0	6	0
3	C	4	0	0	0	0
3	D	31	0	0	7	0
3	E	5	0	0	0	0
3	F	19	0	0	1	0
3	G	4	0	0	0	0
3	H	48	0	0	6	0
3	I	1	0	0	0	0
3	J	29	0	0	6	0
3	K	3	0	0	2	0
3	L	33	0	0	6	0
All	All	15517	0	14437	850	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 29.

The worst 5 of 850 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)
2:J:489:HIS:HB2	2:J:498:ALA:HB2	1.42	1.00
1:C:17:GLU:HA	1:C:20:GLU:HG3	1.46	0.98
2:D:445:GLU:H	2:D:446:PRO:CD	1.79	0.95
2:B:445:GLU:H	2:B:446:PRO:CD	1.78	0.95
2:D:445:GLU:H	2:D:446:PRO:HD3	1.34	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	24/29 (83%)	21 (88%)	3 (12%)	0	100 100
1	C	21/29 (72%)	17 (81%)	4 (19%)	0	100 100
1	E	24/29 (83%)	21 (88%)	2 (8%)	1 (4%)	2 8
1	G	23/29 (79%)	21 (91%)	2 (9%)	0	100 100
1	I	24/29 (83%)	21 (88%)	3 (12%)	0	100 100
1	K	21/29 (72%)	17 (81%)	3 (14%)	1 (5%)	2 6
2	B	299/330 (91%)	255 (85%)	34 (11%)	10 (3%)	3 11
2	D	299/330 (91%)	260 (87%)	28 (9%)	11 (4%)	2 9
2	F	299/330 (91%)	255 (85%)	29 (10%)	15 (5%)	1 5
2	H	300/330 (91%)	259 (86%)	28 (9%)	13 (4%)	2 7
2	J	299/330 (91%)	253 (85%)	34 (11%)	12 (4%)	2 8
2	L	295/330 (89%)	237 (80%)	45 (15%)	13 (4%)	2 7
All	All	1928/2154 (90%)	1637 (85%)	215 (11%)	76 (4%)	2 9

5 of 76 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	433	LYS
2	B	435	ILE
2	B	438	GLU
2	B	445	GLU
2	B	447	LEU

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	22/28 (79%)	19 (86%)	3 (14%)	3 10
1	C	22/28 (79%)	17 (77%)	5 (23%)	0 2
1	E	21/28 (75%)	18 (86%)	3 (14%)	2 9
1	G	23/28 (82%)	20 (87%)	3 (13%)	3 11
1	I	21/28 (75%)	19 (90%)	2 (10%)	7 22
1	K	22/28 (79%)	19 (86%)	3 (14%)	3 10
2	B	238/270 (88%)	222 (93%)	16 (7%)	13 38
2	D	239/270 (88%)	227 (95%)	12 (5%)	20 51
2	F	240/270 (89%)	231 (96%)	9 (4%)	28 62
2	H	237/270 (88%)	224 (94%)	13 (6%)	18 47
2	J	238/270 (88%)	225 (94%)	13 (6%)	18 47
2	L	236/270 (87%)	224 (95%)	12 (5%)	20 51
All	All	1559/1788 (87%)	1465 (94%)	94 (6%)	16 44

5 of 94 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	389	ARG
2	J	374	ILE
2	H	445	GLU
1	I	22	ILE
2	J	445	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
2	L	264	HIS
2	L	477	GLN
2	L	271	HIS
2	L	449	ASN
2	L	503	HIS

5.3.3 RNA

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

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, 95^{th} 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	25/29 (86%)	-2.05	0	100	100	33, 54, 78, 119	0
1	C	22/29 (75%)	-2.04	0	100	100	41, 53, 75, 80	0
1	E	25/29 (86%)	-2.08	0	100	100	38, 54, 89, 119	0
1	G	24/29 (82%)	-2.11	0	100	100	38, 56, 81, 119	0
1	I	25/29 (86%)	-2.04	0	100	100	36, 54, 81, 121	0
1	K	22/29 (75%)	-2.06	0	100	100	36, 51, 71, 82	0
2	B	291/330 (88%)	-2.00	0	100	100	27, 54, 96, 196	0
2	D	291/330 (88%)	-1.94	0	100	100	32, 56, 99, 141	0
2	F	291/330 (88%)	-1.97	0	100	100	24, 54, 98, 170	0
2	H	292/330 (88%)	-1.95	0	100	100	36, 56, 99, 140	0
2	J	291/330 (88%)	-1.91	0	100	100	35, 58, 115, 201	0
2	L	289/330 (87%)	-1.89	0	100	100	35, 58, 106, 152	0
All	All	1888/2154 (87%)	-1.95	0	100	100	24, 56, 103, 201	0

There are no RSRZ outliers to report.

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 monosaccharides 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.