wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 08:15 PM EDT

PDB ID	:	5U7Q
Title	:	Identification of A New Class of Potent Cdc7 Inhibitors Designed by Putative Pharmacophore Model: Synthesis and Biological Evaluation of 2,3-Dihydroth ieno[3,2-d]pyrimidin-4(1H)-ones
Authors	:	Hoffman, I.D.
Deposited on	:	2016-12-12
Resolution	:	3.15 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:

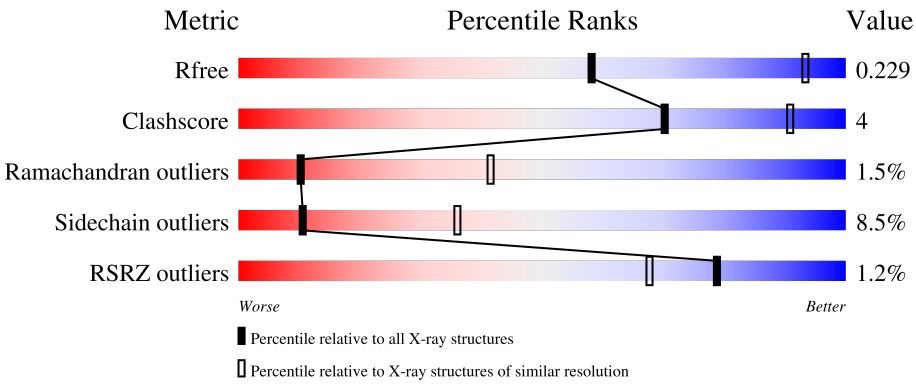
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.15 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	1665 (3.20-3.12)
Clashscore	141614	1804 (3.20-3.12)
Ramachandran outliers	138981	1770 (3.20-3.12)
Sidechain outliers	138945	1769 (3.20-3.12)
RSRZ outliers	127900	1616 (3.20-3.12)

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	395	% 84% 13% • •
1	B	395	% 82% 13% • • •
1	C	395	% 86% 10% • • •
1	D	395	% 81% 12% • • •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12540 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 Rho-associated protein kinase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	388	3147	2018	528	581	20	0	0	0
1	B	385	3121	2004	521	576	20	0	0	0
1	C	390	3156	2023	530	583	20	0	0	0
1	D	383	3106	1996	518	573	19	0	0	0

- Molecule 2 is water.

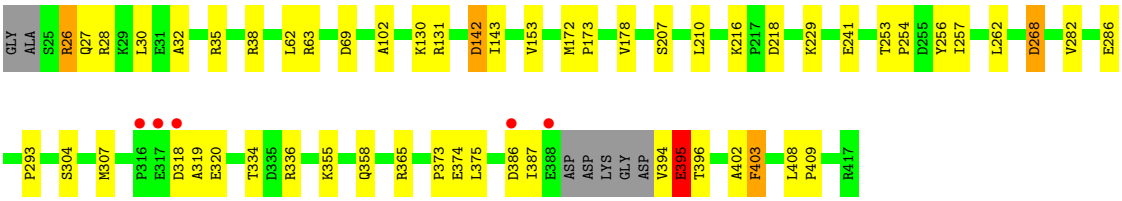
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	2	2	2	0	0
2	B	3	3	3	0	0
2	C	2	2	2	0	0
2	D	3	3	3	0	0

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.

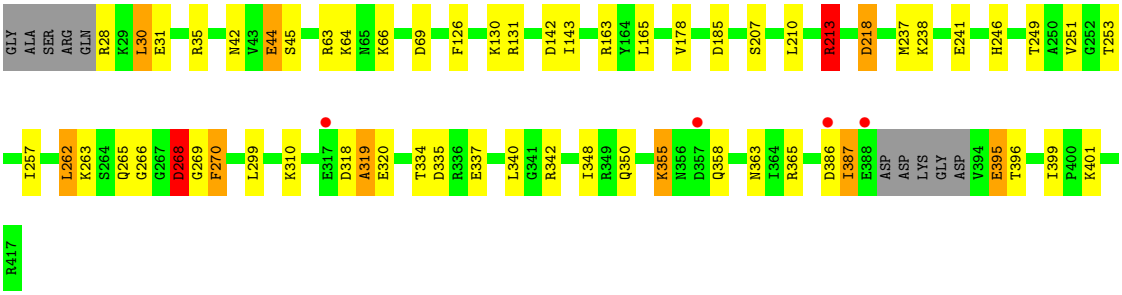
• Molecule 1: Rho-associated protein kinase 2

Chain A:



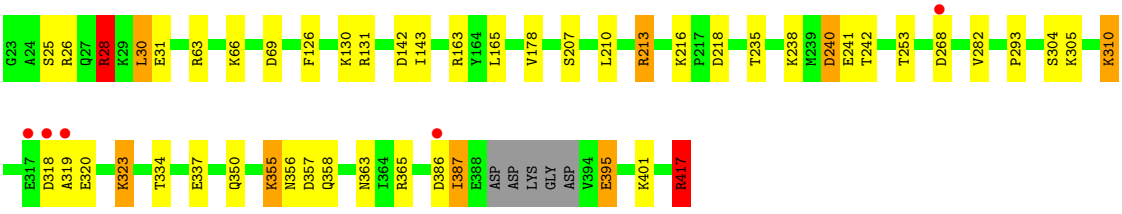
• Molecule 1: Rho-associated protein kinase 2

Chain B:



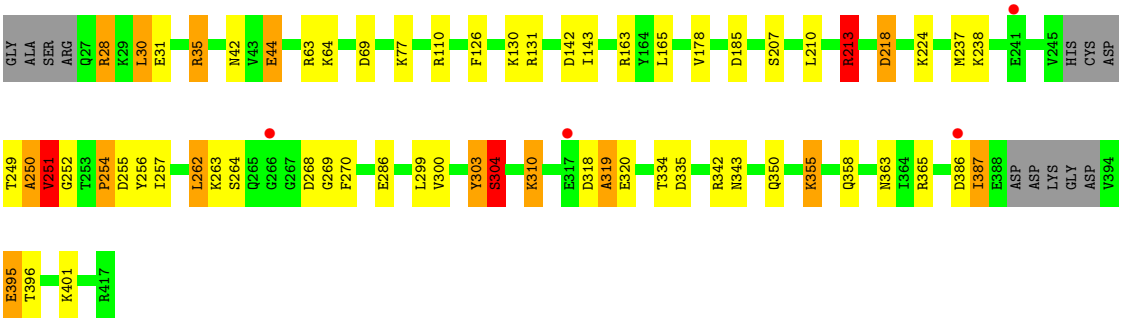
• Molecule 1: Rho-associated protein kinase 2

Chain C:



• Molecule 1: Rho-associated protein kinase 2

Chain D:


GLY
ALA
SER
ARG
Q27
R28
K29
L30
E31
R35
N42
V43
E44
R63
K64
D69
K77
R110
F126
K130
R131
D142
I143
R163
Y164
L165
V178
D185
S207
L210
R213
D218
K224
M237
K238
E241
V245
HIS
CYS
ASP
T249
A250
V251
G252
T253
P254
D255
Y256
I257
L262
K263
S264
Q265
G266
G267
D268
G269
F270
E286
L299
V300
Y303
S304
K310
E317
D318
A319
E320
T334
D335
R342
N343
Q350
K355
Q358
N363
I364
R365
D386
I387
E388
ASP
ASP
LYS
GLY
ASP
V394
E395
T396
K401
R417

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	91.19Å 146.46Å 112.42Å 90.00° 96.34° 90.00°	Depositor
Resolution (Å)	30.00 – 3.15 29.83 – 3.12	Depositor EDS
% Data completeness (in resolution range)	97.9 (30.00-3.15) 98.1 (29.83-3.12)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.68 (at 3.11Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.205 , 0.235 0.203 , 0.229	Depositor DCC
R_{free} test set	2611 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	67.7	Xtriage
Anisotropy	0.082	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 37.4	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	0.93	EDS
Total number of atoms	12540	wwPDB-VP
Average B, all atoms (Å^2)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.57% 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	0.58	0/3225	0.85	8/4357 (0.2%)
1	B	0.55	0/3199	0.86	4/4323 (0.1%)
1	C	0.55	0/3234	0.84	5/4369 (0.1%)
1	D	0.57	0/3182	0.88	8/4298 (0.2%)
All	All	0.56	0/12840	0.86	25/17347 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	2
1	D	0	2
All	All	0	4

There are no bond length outliers.

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	270	PHE	CB-CG-CD1	9.17	127.22	120.80
1	D	213	ARG	NE-CZ-NH1	9.08	124.84	120.30
1	B	213	ARG	NE-CZ-NH1	8.72	124.66	120.30
1	D	304	SER	N-CA-CB	8.71	123.57	110.50
1	B	270	PHE	CB-CG-CD2	-8.61	114.78	120.80

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	246	HIS	Peptide
1	B	269	GLY	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	D	251	VAL	Peptide
1	D	252	GLY	Peptide

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	3147	0	3074	31	0
1	B	3121	0	3048	17	0
1	C	3156	0	3082	19	0
1	D	3106	0	3039	25	0
2	A	2	0	0	0	0
2	B	3	0	0	0	0
2	C	2	0	0	0	0
2	D	3	0	0	1	0
All	All	12540	0	12243	88	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 4.

The worst 5 of 88 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:A:172:MET:CE	1:A:229:LYS:HD2	1.72	1.19
1:A:172:MET:HE1	1:A:229:LYS:CD	1.82	1.08
1:A:153:VAL:HG22	1:A:172:MET:CE	2.09	0.83
1:A:153:VAL:HG22	1:A:172:MET:HE3	1.62	0.82
1:A:172:MET:HE1	1:A:229:LYS:HD2	0.89	0.81

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	384/395 (97%)	368 (96%)	11 (3%)	5 (1%)	12	44
1	B	381/395 (96%)	364 (96%)	13 (3%)	4 (1%)	15	51
1	C	386/395 (98%)	371 (96%)	10 (3%)	5 (1%)	12	44
1	D	377/395 (95%)	359 (95%)	9 (2%)	9 (2%)	6	30
All	All	1528/1580 (97%)	1462 (96%)	43 (3%)	23 (2%)	10	41

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	268	ASP
1	B	395	GLU
1	C	268	ASP
1	C	395	GLU
1	D	250	ALA

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	342/346 (99%)	331 (97%)	11 (3%)	39	70
1	B	339/346 (98%)	299 (88%)	40 (12%)	5	21
1	C	342/346 (99%)	313 (92%)	29 (8%)	10	36
1	D	337/346 (97%)	301 (89%)	36 (11%)	6	25
All	All	1360/1384 (98%)	1244 (92%)	116 (8%)	10	36

5 of 116 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	131	ARG
1	D	358	GLN
1	C	337	GLU
1	D	355	LYS
1	D	238	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	107	GLN
1	D	363	ASN
1	B	363	ASN
1	C	107	GLN
1	C	358	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	388/395 (98%)	-0.29	5 (1%)	77	66	38, 66, 100, 141	0
1	B	385/395 (97%)	-0.26	4 (1%)	82	73	41, 69, 108, 141	0
1	C	390/395 (98%)	-0.27	5 (1%)	77	66	42, 71, 107, 136	0
1	D	383/395 (96%)	-0.23	4 (1%)	82	73	46, 72, 110, 138	0
All	All	1546/1580 (97%)	-0.26	18 (1%)	79	68	38, 70, 108, 141	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	386	ASP	2.9
1	C	317	GLU	2.8
1	A	317	GLU	2.6
1	B	388	GLU	2.5
1	B	317	GLU	2.4

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.