wwPDB X-ray Structure Validation Summary Report (i)

Oct 17, 2021 – 07:18 AM EDT

PDB ID	:	1KZ7
Title	:	Crystal Structure of the DH/PH Fragment of Murine Dbs in Complex with the Placental Isoform of Human Cdc42
Authors	:	Rossman, K.L.; Worthylake, D.K.; Snyder, J.T.; Siderovski, D.P.; Campbell, S.L.; Sondek, J.
Deposited on	:	2002-02-06
Resolution	:	2.40 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 following versions of software and data (see references (i)) were used in the production of this report:

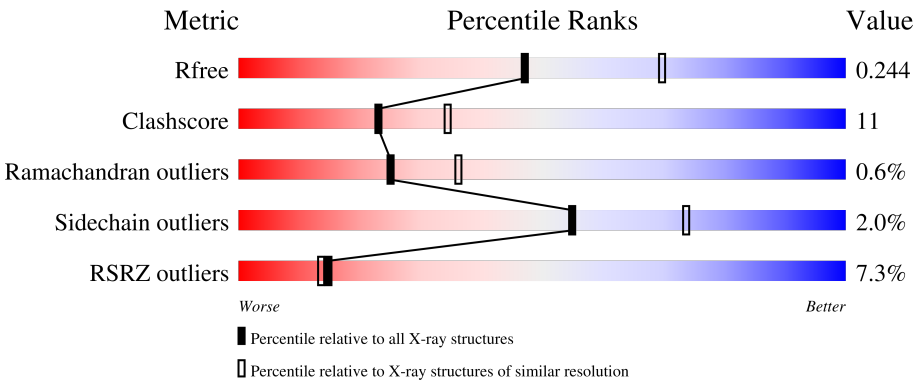
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	353	7%, 69%, 24%, • 5%
1	C	353	7%, 67%, 26%, • 7%
2	B	188	3%, 83%, 17%
2	D	188	11%, 72%, 22%, 6%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8598 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 GUANINE NUCLEOTIDE EXCHANGE FACTOR DBS.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	337	2748	1741	474	512	11	10	0	0	0
1	C	330	2677	1700	457	499	11	10	0	0	0

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	622	MSE	-	SEE REMARK 999	UNP Q64096
A	637	MSE	MET	modified residue	UNP Q64096
A	662	MSE	MET	modified residue	UNP Q64096
A	667	MSE	MET	modified residue	UNP Q64096
A	686	MSE	MET	modified residue	UNP Q64096
A	696	ILE	ASN	SEE REMARK 999	UNP Q64096
A	697	PHE	ILE	SEE REMARK 999	UNP Q64096
A	698	LEU	PRO	SEE REMARK 999	UNP Q64096
A	699	ARG	ALA	SEE REMARK 999	UNP Q64096
A	700	GLU	GLY	SEE REMARK 999	UNP Q64096
A	701	LEU	VAL	SEE REMARK 999	UNP Q64096
A	719	MSE	MET	modified residue	UNP Q64096
A	782	MSE	MET	modified residue	UNP Q64096
A	813	MSE	MET	modified residue	UNP Q64096
A	833	MSE	MET	modified residue	UNP Q64096
A	859	MSE	MET	modified residue	UNP Q64096
A	897	MSE	MET	modified residue	UNP Q64096
A	940	ALA	GLU	SEE REMARK 999	UNP Q64096
A	968	GLU	-	expression tag	UNP Q64096
A	969	HIS	-	expression tag	UNP Q64096
A	970	HIS	-	expression tag	UNP Q64096
A	971	HIS	-	expression tag	UNP Q64096
A	972	HIS	-	expression tag	UNP Q64096
A	973	HIS	-	expression tag	UNP Q64096
A	974	HIS	-	expression tag	UNP Q64096

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	1622	MSE	-	SEE REMARK 999	UNP Q64096
C	1637	MSE	MET	modified residue	UNP Q64096
C	1662	MSE	MET	modified residue	UNP Q64096
C	1667	MSE	MET	modified residue	UNP Q64096
C	1686	MSE	MET	modified residue	UNP Q64096
C	1696	ILE	ASN	SEE REMARK 999	UNP Q64096
C	1697	PHE	ILE	SEE REMARK 999	UNP Q64096
C	1698	LEU	PRO	SEE REMARK 999	UNP Q64096
C	1699	ARG	ALA	SEE REMARK 999	UNP Q64096
C	1700	GLU	GLY	SEE REMARK 999	UNP Q64096
C	1701	LEU	VAL	SEE REMARK 999	UNP Q64096
C	1719	MSE	MET	modified residue	UNP Q64096
C	1782	MSE	MET	modified residue	UNP Q64096
C	1813	MSE	MET	modified residue	UNP Q64096
C	1833	MSE	MET	modified residue	UNP Q64096
C	1859	MSE	MET	modified residue	UNP Q64096
C	1897	MSE	MET	modified residue	UNP Q64096
C	1940	ALA	GLU	SEE REMARK 999	UNP Q64096
C	1968	GLU	-	expression tag	UNP Q64096
C	1969	HIS	-	expression tag	UNP Q64096
C	1970	HIS	-	expression tag	UNP Q64096
C	1971	HIS	-	expression tag	UNP Q64096
C	1972	HIS	-	expression tag	UNP Q64096
C	1973	HIS	-	expression tag	UNP Q64096
C	1974	HIS	-	expression tag	UNP Q64096

- Molecule 2 is a protein called CDC42 HOMOLOG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	188	1471	943	239	282	7	0	0	0
			Total	C	N	O	S			
2	D	177	1379	888	220	264	7	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	188	SER	CYS	engineered mutation	UNP P60953
D	1188	SER	CYS	engineered mutation	UNP P60953

- Molecule 3 is water.

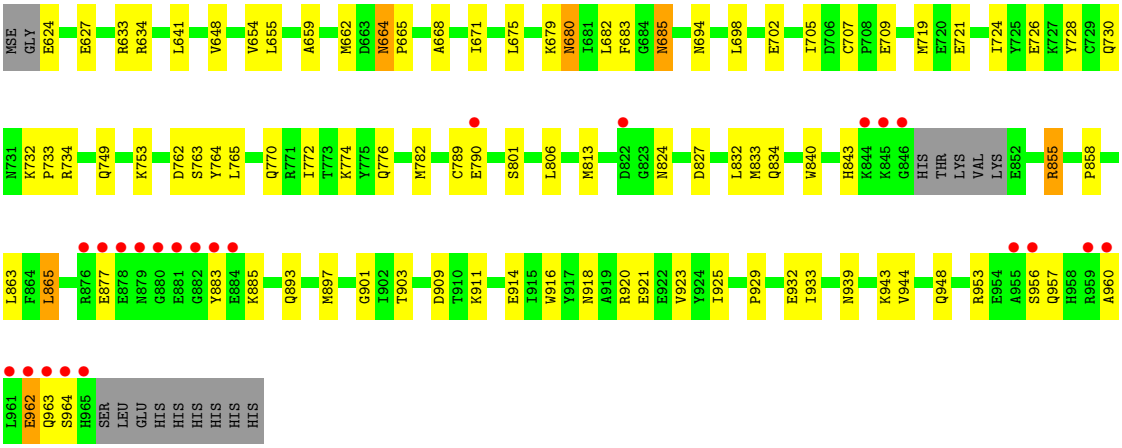
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	92	92	92	0	0
			Total	O		
3	B	97	97	97	0	0
			Total	O		
3	C	87	87	87	0	0
			Total	O		
3	D	47	47	47	0	0

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.

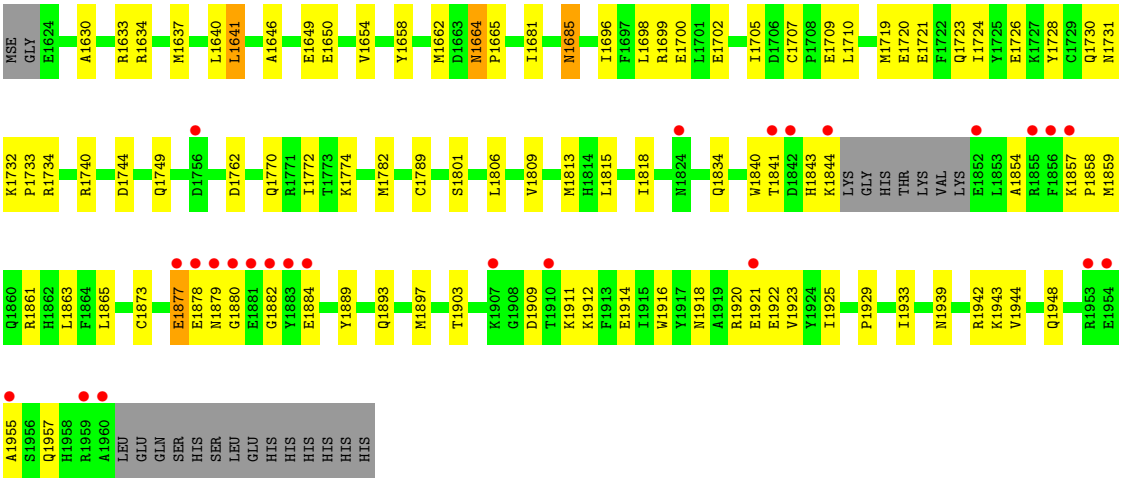
• Molecule 1: GUANINE NUCLEOTIDE EXCHANGE FACTOR DBS

Chain A:



• Molecule 1: GUANINE NUCLEOTIDE EXCHANGE FACTOR DBS

Chain C:

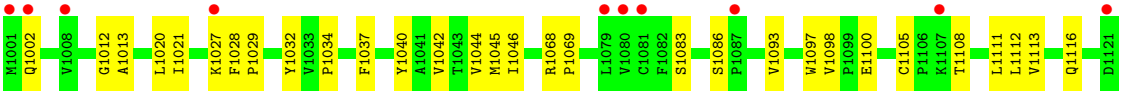


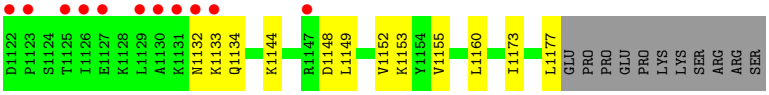
• Molecule 2: CDC42 HOMOLOG

Chain B:

• Molecule 2: CDC42 HOMOLOG

Chain D:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	66.50Å 87.98Å 230.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 2.40 29.51 – 2.40	Depositor EDS
% Data completeness (in resolution range)	91.3 (19.99-2.40) 91.3 (29.51-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.87 (at 2.39Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.197 , 0.239 0.203 , 0.244	Depositor DCC
R_{free} test set	2460 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	36.1	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 46.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8598	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.89% 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	0.34	0/2791	0.58	0/3734
1	C	0.35	0/2718	0.56	0/3639
2	B	0.37	0/1504	0.65	0/2044
2	D	0.32	0/1409	0.61	0/1918
All	All	0.34	0/8422	0.59	0/11335

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	2748	0	2732	68	0
1	C	2677	0	2654	72	0
2	B	1471	0	1494	23	0
2	D	1379	0	1396	28	0
3	A	92	0	0	6	0
3	B	97	0	0	3	0
3	C	87	0	0	5	0
3	D	47	0	0	0	0
All	All	8598	0	8276	189	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 11.

The worst 5 of 189 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:C:1719:MSE:HE2	1:C:1801:SER:HB3	1.41	1.03
1:A:929:PRO:HG2	1:A:933:ILE:HD12	1.55	0.88
1:C:1731:ASN:HD22	1:C:1734:ARG:HH21	1.24	0.85
1:A:685:ASN:H	1:A:685:ASN:HD22	1.24	0.84
1:A:707:CYS:HB3	3:A:2271:HOH:O	1.78	0.82

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	333/353 (94%)	323 (97%)	7 (2%)	3 (1%)	17	25
1	C	326/353 (92%)	307 (94%)	17 (5%)	2 (1%)	25	36
2	B	186/188 (99%)	175 (94%)	10 (5%)	1 (0%)	29	41
2	D	175/188 (93%)	167 (95%)	8 (5%)	0	100	100
All	All	1020/1082 (94%)	972 (95%)	42 (4%)	6 (1%)	25	36

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	668	ALA
1	A	883	TYR
1	A	963	GLN
1	C	1877	GLU
1	C	1884	GLU

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	300/304 (99%)	291 (97%)	9 (3%)	41	61
1	C	290/304 (95%)	282 (97%)	8 (3%)	43	63
2	B	168/168 (100%)	167 (99%)	1 (1%)	86	94
2	D	157/168 (94%)	157 (100%)	0	100	100
All	All	915/944 (97%)	897 (98%)	18 (2%)	55	74

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1744	ASP
1	C	1916	TRP
1	C	1893	GLN
1	A	962	GLU
1	C	1685	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1694	ASN
1	C	1843	HIS
1	C	1788	HIS
1	C	1879	ASN
1	A	843	HIS

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	327/353 (92%)	0.09	23 (7%)	16	15	20, 37, 81, 124	0
1	C	320/353 (90%)	0.19	25 (7%)	13	11	21, 38, 86, 117	0
2	B	188/188 (100%)	-0.24	5 (2%)	54	52	18, 29, 65, 82	0
2	D	177/188 (94%)	0.40	21 (11%)	4	4	22, 43, 89, 98	0
All	All	1012/1082 (93%)	0.11	74 (7%)	15	13	18, 38, 83, 124	0

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1883	TYR	10.7
1	A	880	GLY	10.0
2	D	1001	MET	9.3
1	A	879	ASN	7.8
1	A	883	TYR	7.5

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 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.