



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 03:15 pm GMT

PDB ID : 4Q1C  
Title : Human dCK C4S-S74E mutant in complex with UDP and the inhibitor 8 {2, 2'-[{4-[(2R)-4-{{[(4,6-diaminopyrimidin-2-yl)sulfanyl]methyl}-5-propyl-2,3-dihydro-1,3-thiazol-2-yl]benzene-1,2-diyl}bis(oxy)]diethanol}  
Authors : Nomme, J.; Lavie, A.  
Deposited on : 2014-04-03  
Resolution : 2.00 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

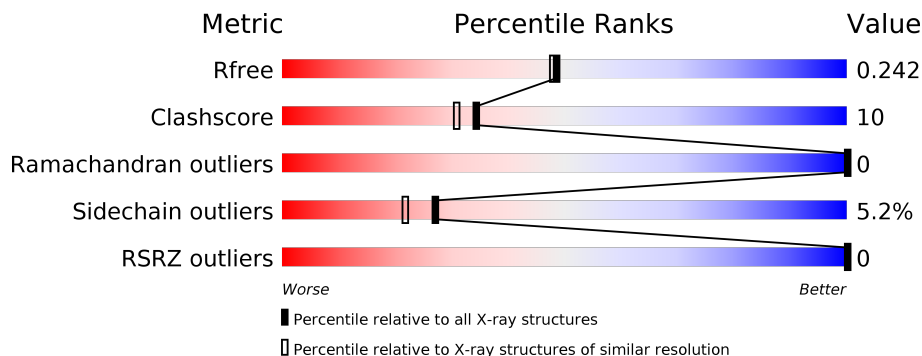
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	100719	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . 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	280	
1	B	280	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3990 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 Deoxycytidine kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	230	Total	C	N	O	S	0	1	0
			1897	1218	313	358	8			
1	B	229	Total	C	N	O	S	0	1	0
			1870	1202	309	352	7			

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP P27707
A	-18	GLY	-	EXPRESSION TAG	UNP P27707
A	-17	SER	-	EXPRESSION TAG	UNP P27707
A	-16	SER	-	EXPRESSION TAG	UNP P27707
A	-15	HIS	-	EXPRESSION TAG	UNP P27707
A	-14	HIS	-	EXPRESSION TAG	UNP P27707
A	-13	HIS	-	EXPRESSION TAG	UNP P27707
A	-12	HIS	-	EXPRESSION TAG	UNP P27707
A	-11	HIS	-	EXPRESSION TAG	UNP P27707
A	-10	HIS	-	EXPRESSION TAG	UNP P27707
A	-9	SER	-	EXPRESSION TAG	UNP P27707
A	-8	SER	-	EXPRESSION TAG	UNP P27707
A	-7	GLY	-	EXPRESSION TAG	UNP P27707
A	-6	LEU	-	EXPRESSION TAG	UNP P27707
A	-5	VAL	-	EXPRESSION TAG	UNP P27707
A	-4	PRO	-	EXPRESSION TAG	UNP P27707
A	-3	ARG	-	EXPRESSION TAG	UNP P27707
A	-2	GLY	-	EXPRESSION TAG	UNP P27707
A	-1	SER	-	EXPRESSION TAG	UNP P27707
A	0	HIS	-	EXPRESSION TAG	UNP P27707
A	9	SER	CYS	ENGINEERED MUTATION	UNP P27707
A	45	SER	CYS	ENGINEERED MUTATION	UNP P27707
A	59	SER	CYS	ENGINEERED MUTATION	UNP P27707
A	74	GLU	SER	ENGINEERED MUTATION	UNP P27707
A	146	SER	CYS	ENGINEERED MUTATION	UNP P27707

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-19	MET	-	EXPRESSION TAG	UNP P27707
B	-18	GLY	-	EXPRESSION TAG	UNP P27707
B	-17	SER	-	EXPRESSION TAG	UNP P27707
B	-16	SER	-	EXPRESSION TAG	UNP P27707
B	-15	HIS	-	EXPRESSION TAG	UNP P27707
B	-14	HIS	-	EXPRESSION TAG	UNP P27707
B	-13	HIS	-	EXPRESSION TAG	UNP P27707
B	-12	HIS	-	EXPRESSION TAG	UNP P27707
B	-11	HIS	-	EXPRESSION TAG	UNP P27707
B	-10	HIS	-	EXPRESSION TAG	UNP P27707
B	-9	SER	-	EXPRESSION TAG	UNP P27707
B	-8	SER	-	EXPRESSION TAG	UNP P27707
B	-7	GLY	-	EXPRESSION TAG	UNP P27707
B	-6	LEU	-	EXPRESSION TAG	UNP P27707
B	-5	VAL	-	EXPRESSION TAG	UNP P27707
B	-4	PRO	-	EXPRESSION TAG	UNP P27707
B	-3	ARG	-	EXPRESSION TAG	UNP P27707
B	-2	GLY	-	EXPRESSION TAG	UNP P27707
B	-1	SER	-	EXPRESSION TAG	UNP P27707
B	0	HIS	-	EXPRESSION TAG	UNP P27707
B	9	SER	CYS	ENGINEERED MUTATION	UNP P27707
B	45	SER	CYS	ENGINEERED MUTATION	UNP P27707
B	59	SER	CYS	ENGINEERED MUTATION	UNP P27707
B	74	GLU	SER	ENGINEERED MUTATION	UNP P27707
B	146	SER	CYS	ENGINEERED MUTATION	UNP P27707

- Molecule 2 is 2,2'-((4-(4-(((4,6-DIAMINOPYRIMIDIN-2-YL)THIO)METHYL)-5-PROPYLTHIAZOL-2-YL)-1,2-PHENYLENE)BIS(OXY))BIS(ETHAN-1-OL) (three-letter code: 2XM) (formula: C<sub>21</sub>H<sub>27</sub>N<sub>5</sub>O<sub>4</sub>S<sub>2</sub>).



- Molecule 3 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula:  $\text{C}_9\text{H}_{14}\text{N}_2\text{O}_{12}\text{P}_2$ ).



- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	56	Total 56	O 56	0	0
4	B	53	Total 53	O 53	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.66Å 68.66Å 119.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.98 – 2.00 27.35 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.1 (25.98-2.00) 98.1 (27.35-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.09 (at 1.99Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.205 , 0.238 0.209 , 0.242	Depositor DCC
$R_{free}$ test set	1821 reflections (5.24%)	DCC
Wilson B-factor (Å <sup>2</sup> )	44.1	Xtriage
Anisotropy	0.093	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 16.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.479 for h,-k,-l	Xtriage
Reported twinning fraction	0.504 for H, K, L 0.496 for -H, K, -L	Depositor
Outliers	0 of 36947 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3990	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.27% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: UDP, 2XM

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	0.62	0/1946	0.81	3/2636 (0.1%)
1	B	0.63	0/1918	0.77	0/2600
All	All	0.62	0/3864	0.79	3/5236 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	244	ASP	CB-CA-C	7.82	126.03	110.40
1	A	230	GLU	CB-CA-C	-6.16	98.07	110.40
1	A	188	ARG	NE-CZ-NH1	5.11	122.85	120.30

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	1897	0	1817	38	0
1	B	1870	0	1786	34	0
2	A	32	0	27	2	0
2	B	32	0	27	1	0
3	A	25	0	11	0	0
3	B	25	0	11	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	56	0	0	5	0
4	B	53	0	0	8	0
All	All	3990	0	3679	72	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 10.

The worst 5 of 72 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:45:SER:O	4:B:434:HOH:O	1.97	0.83
1:A:209:HIS:O	4:A:443:HOH:O	1.95	0.82
1:B:145:GLU:OE2	4:B:415:HOH:O	2.00	0.79
1:B:48:TRP:O	4:B:434:HOH:O	2.01	0.77
1:B:140:ASN:ND2	4:B:422:HOH:O	2.22	0.72

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	227/280 (81%)	213 (94%)	14 (6%)	0	100	100
1	B	226/280 (81%)	216 (96%)	10 (4%)	0	100	100
All	All	453/560 (81%)	429 (95%)	24 (5%)	0	100	100

There are no Ramachandran outliers to report.

### 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	204/256 (80%)	192 (94%)	12 (6%)	23	17
1	B	199/256 (78%)	190 (96%)	9 (4%)	32	27
All	All	403/512 (79%)	382 (95%)	21 (5%)	27	22

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	227	TYR
1	A	241	ASP
1	B	211	LYS
1	A	226	ASP
1	B	217	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	75	GLN
1	A	156	GLN
1	B	113	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry

4 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
2	2XM	A	301	-	31,34,34	2.16	3 (9%)	32,45,45	1.73	5 (15%)
3	UDP	A	302	-	21,26,26	0.86	1 (4%)	22,40,40	2.04	2 (9%)
2	2XM	B	301	-	31,34,34	1.89	3 (9%)	32,45,45	1.34	5 (15%)
3	UDP	B	302	-	21,26,26	0.82	0	22,40,40	1.83	3 (13%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	2XM	A	301	-	-	0/17/20/20	0/3/3/3
3	UDP	A	302	-	-	0/12/32/32	0/2/2/2
2	2XM	B	301	-	-	0/17/20/20	0/3/3/3
3	UDP	B	302	-	-	0/12/32/32	0/2/2/2

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	301	2XM	CBF-SAW	-8.32	1.62	1.73
2	A	301	2XM	CBF-SAW	-7.92	1.62	1.73
2	A	301	2XM	CBE-SAW	-7.25	1.60	1.74
2	B	301	2XM	CBE-CBD	-4.20	1.34	1.42
2	A	301	2XM	CBE-CBD	-4.07	1.34	1.42

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	301	2XM	CAJ-CAO-CBE	-4.46	99.51	115.44
3	B	302	UDP	C4'-O4'-C1'	-3.51	106.03	109.77

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	301	2XM	N1-C2-N3	-3.34	120.84	126.82
2	B	301	2XM	N1-C2-N3	-2.83	121.77	126.82
2	B	301	2XM	CAJ-CAO-CBE	-2.16	107.73	115.44

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	2XM	2	0
2	B	301	2XM	1	0
3	B	302	UDP	2	0

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	230/280 (82%)	0.01	0 100 100	29, 30, 30, 30	0
1	B	229/280 (81%)	0.03	0 100 100	29, 30, 30, 31	0
All	All	459/560 (81%)	0.02	0 100 100	29, 30, 30, 31	0

There are no RSRZ outliers to report.

### 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 carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	2XM	B	301	32/32	0.95	0.13	0.44	28,29,29,30	0
3	UDP	A	302	25/25	0.97	0.12	0.29	28,29,30,30	0
2	2XM	A	301	32/32	0.96	0.12	-0.05	28,29,30,30	0
3	UDP	B	302	25/25	0.97	0.12	-0.45	28,29,30,30	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.