



wwPDB X-ray Structure Validation Summary Report i

Mar 1, 2014 – 01:41 AM GMT

PDB ID : 3VWA
Title : Crystal structure of Cex1p
Authors : Nozawa, K.; Ishitani, R.; Nureki, O.
Deposited on : 2012-08-13
Resolution : 2.20 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

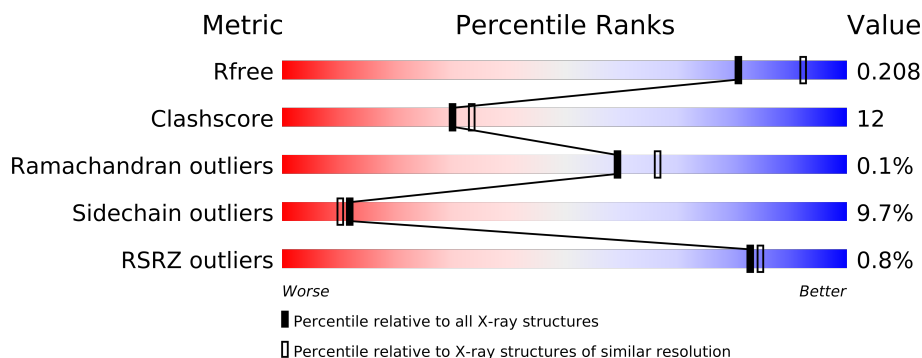
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	FAILED
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.20 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}	66092	2938 (2.20-2.20)
Clashscore	79885	3751 (2.20-2.20)
Ramachandran outliers	78287	3681 (2.20-2.20)
Sidechain outliers	78261	3682 (2.20-2.20)
RSRZ outliers	66119	2939 (2.20-2.20)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	560	
1	B	560	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8822 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Cytoplasmic export protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	531	Total	C	N	O	S	Se	0	0	0
			4184	2702	687	780	8	7			
1	B	529	Total	C	N	O	S	Se	0	0	0
			4180	2699	689	777	8	7			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	5	PRO	-	EXPRESSION TAG	UNP Q12453
A	6	LEU	-	EXPRESSION TAG	UNP Q12453
A	7	HIS	-	EXPRESSION TAG	UNP Q12453
A	8	MSE	-	EXPRESSION TAG	UNP Q12453
B	5	PRO	-	EXPRESSION TAG	UNP Q12453
B	6	LEU	-	EXPRESSION TAG	UNP Q12453
B	7	HIS	-	EXPRESSION TAG	UNP Q12453
B	8	MSE	-	EXPRESSION TAG	UNP Q12453

- Molecule 2 is water.

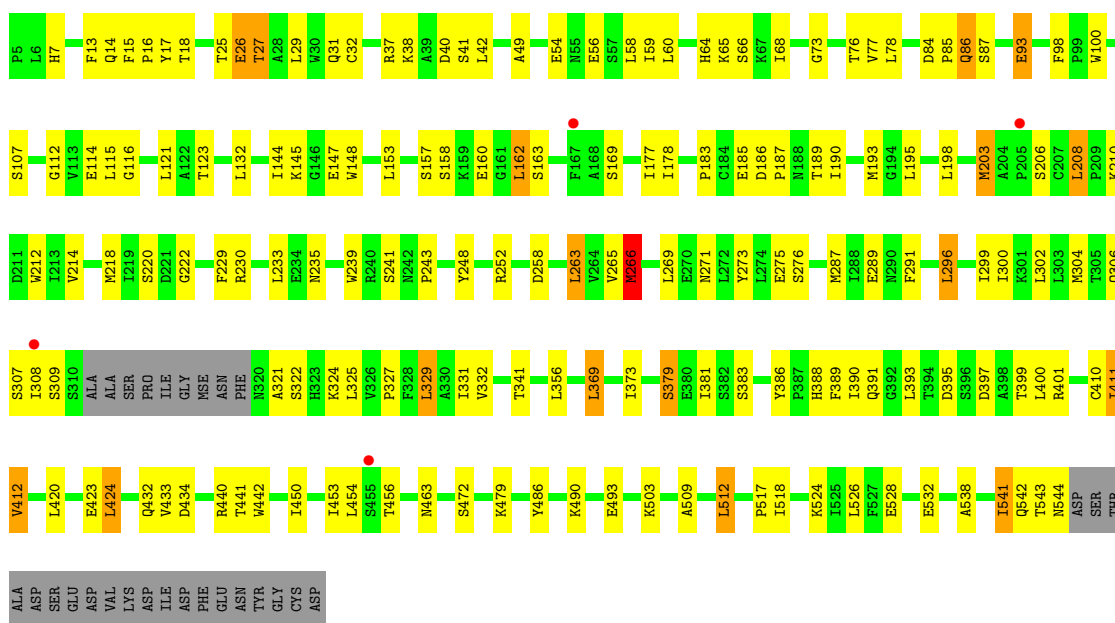
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	229	Total	O	0	0
			229	229		
2	B	229	Total	O	0	0
			229	229		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

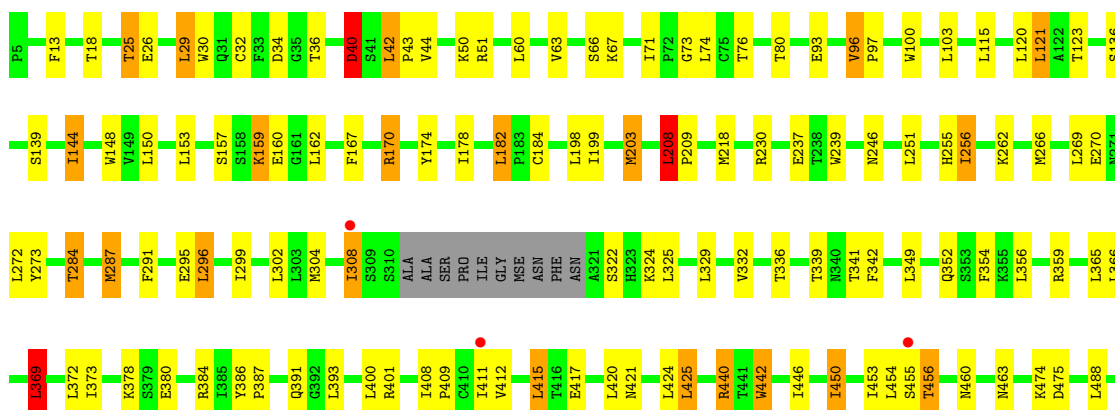
• Molecule 1: Cytoplasmic export protein 1

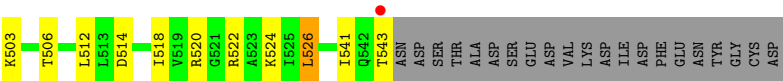
Chain A: 



• Molecule 1: Cytoplasmic export protein 1

Chain B: 





4 Data and refinement statistics

Xtriage (Phenix) failed to run properly - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	34.06Å 182.22Å 112.44Å 90.00° 98.69° 90.00°	Depositor
Resolution (Å)	45.56 – 2.20 45.56 – 2.20	Depositor EDS
% Data completeness (in resolution range)	98.7 (45.56-2.20) 98.8 (45.56-2.20)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.166 , 0.207 0.167 , 0.208	Depositor DCC
R_{free} test set	6784 reflections (11.14%)	DCC
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 21.6	EDS
Estimated twinning fraction	0.575 for H, K, L 0.425 for -H, -K, H+L	Xtriage
Reported twinning fraction	0.575 for H, K, L 0.425 for -H, -K, H+L	Depositor
F_o, F_c correlation	0.95	EDS
Total number of atoms	8822	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

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	0.78	4/4256 (0.1%)	0.88	4/5763 (0.1%)
1	B	0.80	4/4252 (0.1%)	0.92	10/5756 (0.2%)
All	All	0.79	8/8508 (0.1%)	0.90	14/11519 (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	1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	239	TRP	CD2-CE2	7.61	1.50	1.41
1	B	100	TRP	CD2-CE2	7.28	1.50	1.41
1	A	212	TRP	CD2-CE2	6.27	1.48	1.41
1	A	100	TRP	CD2-CE2	5.48	1.48	1.41
1	B	30	TRP	CD2-CE2	5.38	1.47	1.41

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	440	ARG	NE-CZ-NH2	-10.36	115.12	120.30
1	B	440	ARG	NE-CZ-NH1	9.03	124.81	120.30
1	B	218	MSE	CA-CB-CG	-7.88	99.91	113.30
1	A	287	MSE	CA-CB-CG	-6.77	101.80	113.30
1	A	266	MSE	CG-SE-CE	6.58	113.37	98.90

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	40	ASP	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4184	0	4315	100	0
1	B	4180	0	4323	104	1
2	A	229	0	0	8	1
2	B	229	0	0	10	0
All	All	8822	0	8638	204	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 12.

The worst 5 of 204 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:203:MSE:HE3	1:B:208:LEU:HD23	1.33	1.08
1:B:354:PHE:HZ	1:B:369:LEU:HD21	1.20	1.03
1:B:266:MSE:HE1	1:B:269:LEU:CD2	1.91	1.01
1:B:67:LYS:NZ	2:B:722:HOH:O	1.75	1.00
1:B:421:ASN:ND2	1:B:455:SER:OG	1.95	0.99

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:543:THR:O	2:A:636:HOH:O[2_746]	1.89	0.31

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	527/560 (94%)	511 (97%)	15 (3%)	1 (0%)	56	62
1	B	525/560 (94%)	515 (98%)	10 (2%)	0	100	100
All	All	1052/1120 (94%)	1026 (98%)	25 (2%)	1 (0%)	59	66

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	86	GLN

5.3.2 Protein sidechains ⓘ

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	480/502 (96%)	429 (89%)	51 (11%)	10	8
1	B	481/502 (96%)	439 (91%)	42 (9%)	15	14
All	All	961/1004 (96%)	868 (90%)	93 (10%)	12	10

5 of 93 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	456	THR
1	B	29	LEU
1	B	420	LEU
1	A	493	GLU
1	A	532	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	432	GLN
1	B	129	ASN
1	B	352	GLN
1	A	391	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	255	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 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	531/560 (94%)	0.04	4 (0%) 83 85	19, 35, 56, 82	0
1	B	529/560 (94%)	0.02	4 (0%) 83 85	20, 32, 50, 75	0
All	All	1060/1120 (94%)	0.03	8 (0%) 83 85	19, 34, 54, 82	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	543	THR	3.8
1	B	308	ILE	2.8
1	A	308	ILE	2.8
1	A	205	PRO	2.3
1	A	455	SER	2.2

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 carbohydrates 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.