wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 08:47 PM GMT

PDB ID	:	1M2V
Title	:	Crystal Structure of the yeast Sec23/24 heterodimer
Authors	:	Bi, X.; Corpina, R.A.; Goldberg, J.
Deposited on	:	2002-06-25
Resolution	:	2.75 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
http://wwpdb.org/validation/2016/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:

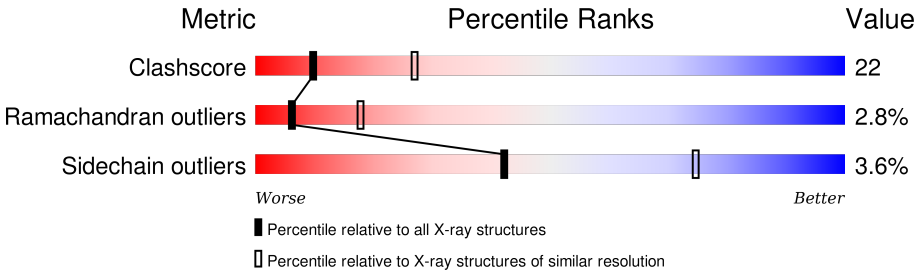
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.75 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))
Clashscore	102246	3829 (2.80-2.72)
Ramachandran outliers	100387	3767 (2.80-2.72)
Sidechain outliers	100360	3770 (2.80-2.72)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	768	55% 34% • 8%
2	B	926	46% 32% • 19%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11577 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 protein transport protein SEC23.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	705	Total	C	N	O	S	0	0	0
			5574	3560	933	1059	22			

- Molecule 2 is a protein called protein transport protein SEC24.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	748	Total	C	N	O	S	0	0	0
			5932	3772	1018	1104	38			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Zn	0	0
			1	1		
3	A	1	Total	Zn	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	34	Total	O	0	0
			34	34		
4	B	35	Total	O	0	0
			35	35		

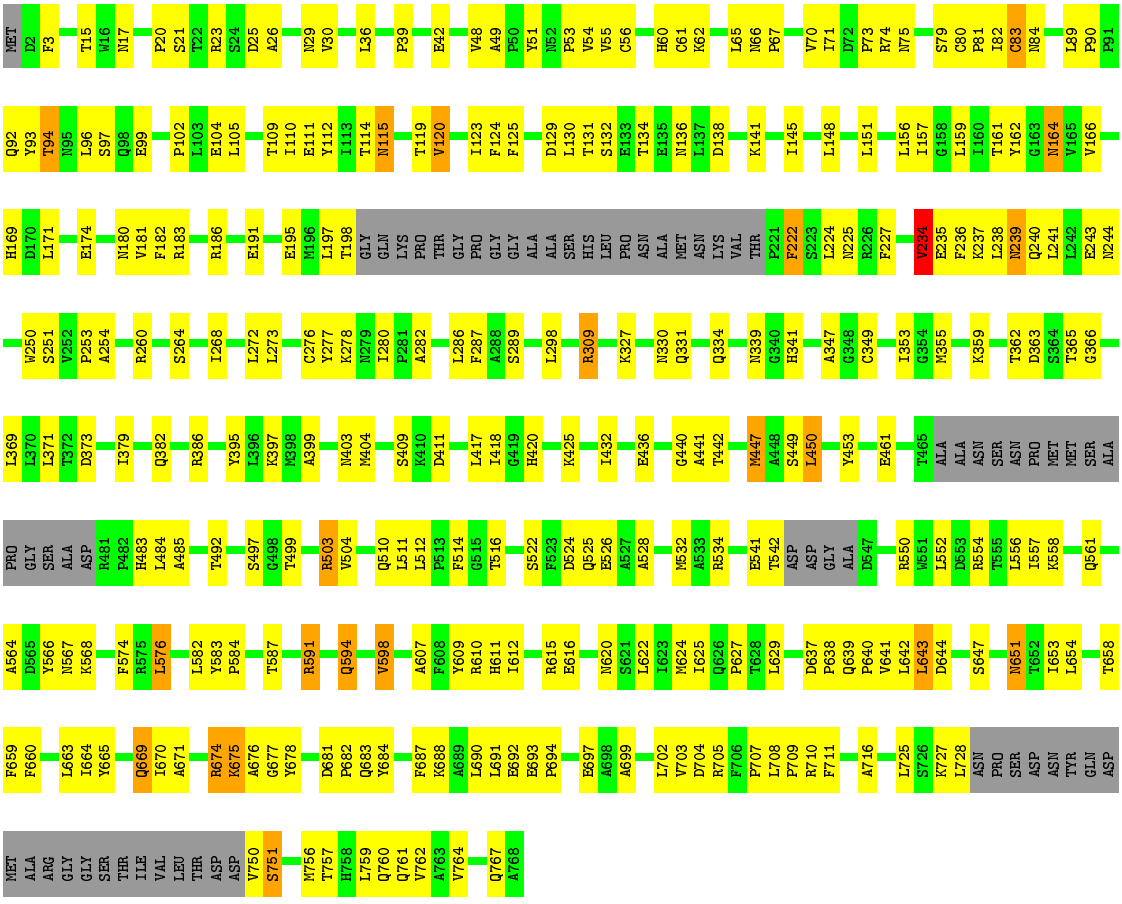
3 Residue-property plots (i)

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.

Note EDS was not executed.

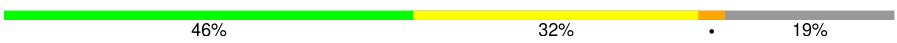
- Molecule 1: protein transport protein SEC23

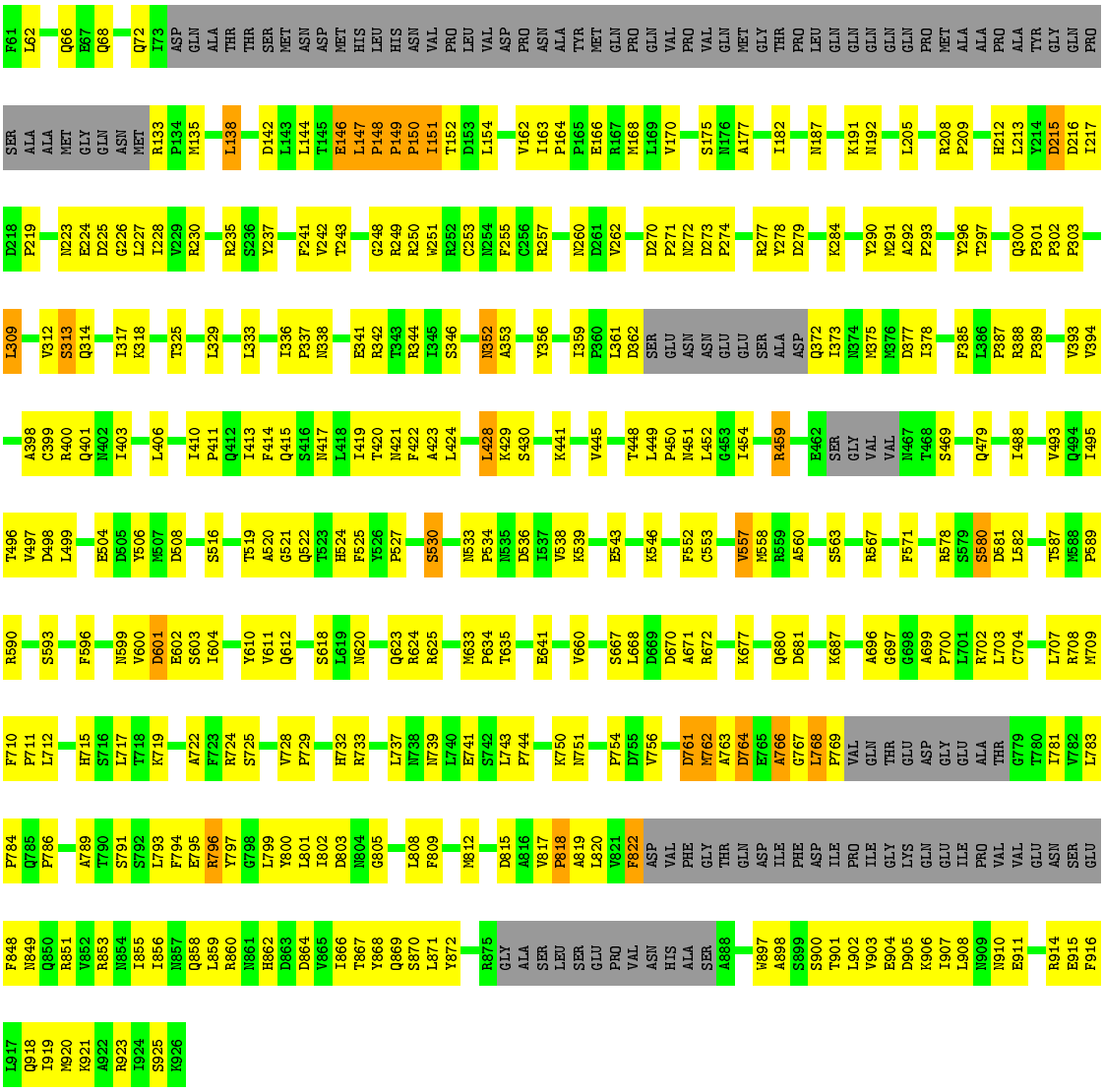
Chain A:



- Molecule 2: protein transport protein SEC24

Chain B:




F61 L62 Q66 E67 Q68 Q72 I73
SER ALA ALA MET GLY GLN ASN MET R133 P134 M135 L138 D142 L143 L144 T145 E146 L147 P148 P149 P150 I151 T152 D153 L154 V162 I163 P164 P165 E166 R167 M168 L169 V170 S175 N176 A177 I182 N187 K191 N192 L205 R208 P209 H212 L213 Y214 D215 D216 I217
D218 P219 N223 E224 D225 G226 L227 I228 V229 R230 R235 S236 Y237 F241 V242 T243 G248 R249 R250 W251 R252 C253 N254 F255 C256 R257 M260 D261 V262 D270 P271 N272 D273 P274 R277 Y278 D279 K284 Y290 M291 A292 P293 Y296 T297 Q300 P301 P302 P303
L309 V312 S313 Q314 I317 K318 T325 L329 L333 I336 P337 N338 E341 R342 T343 R344 I345 S346 N352 A353 Y356 I359 P360 L361 D362 SER GLU ASN ASN GLU GLU SER ALA ASP Q372 I373 N374 M375 H376 D377 I378 F385 L386 P387 R388 P389 V393 V394
A398 C399 R400 Q401 N402 I403 L406 I410 P411 Q412 I413 F414 Q415 S416 N417 L418 I419 T420 N421 F422 A423 L424 L428 K429 S430 K441 V445 T448 L449 P450 N451 L452 G453 I454 R459 E462 SER GLY VAL VAL N467 T468 S469 Q479 I488 V493 Q494 I495
T496 V497 D498 L499 E504 D505 Y506 M507 D508 S516 T519 A520 G521 Q522 T523 H524 F525 Y526 P527 S530 N533 P534 N535 D536 I537 V538 K539 E543 K546 F552 C553 V557 M558 R559 A560 S563 R567 F571 R578 S579 S580 D581 L582 T587 M588 P589
R590 S593 F596 N599 V600 D601 E602 S603 I604 Y610 V611 Q612 S618 L619 N620 Q623 R624 R625 M633 P634 T635 E641 V660 S667 L668 D669 D670 A671 R672 K677 Q680 D681 K687 A696 G697 G698 A699 P700 L701 R702 L703 C704 L707 R708 M709
F710 P711 L712 H715 S716 L717 T718 K719 A722 F723 R724 S725 V728 P729 H732 R733 L737 N738 N739 L740 E741 S742 L743 P744 K750 N751 P754 D755 V756 D761 M762 A763 D764 E765 A766 G767 L768 P769 VAL GLN THR GLU ASP GLY GLU ALA THR G779 T780 I781 V782 L783
P784 Q785 P786 A789 T790 S791 S792 L793 F794 E795 R796 Y797 G798 L799 Y800 L801 I802 D803 N804 G805 L808 F809 M812 D815 A816 V817 P818 A819 L820 V821 F822 ASP VAL PHE GLY THR GLN ASP ILE PHE ASP ILE PRO ILE GLY LYS GLN GLU ILE PRO VAL GLU ASN SER GLU
F848 N849 Q850 R851 V852 R853 N854 I855 I856 N857 Q858 L859 R860 N861 H862 D863 D864 V865 I866 T867 Y868 Q869 S870 L871 Y872 R875 GLY ALA SER LEU SER GLU PRO VAL ASN HIS ALA SER A888 W897 A898 S899 S900 T901 L902 V903 E904 D905 K906 I907 L908 N909 N910 E911 R914 E915 F916
L917 Q918 I919 M920 K921 A922 R923 I924 S925 K926

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	90.31Å 126.37Å 180.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 2.75	Depositor
% Data completeness (in resolution range)	91.2 (19.96-2.75)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.210 , 0.257	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	11577	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	0.40	0/5706	0.66	1/7765 (0.0%)
2	B	0.39	0/6054	0.67	1/8214 (0.0%)
All	All	0.39	0/11760	0.66	2/15979 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	677	GLY	N-CA-C	6.93	130.42	113.10
2	B	768	LEU	CA-CB-CG	5.33	127.57	115.30

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	5574	0	5475	230	0
2	B	5932	0	5958	282	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	34	0	0	5	0
4	B	35	0	0	3	0
All	All	11577	0	11433	510	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 22.

The worst 5 of 510 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:B:907:ILE:HG22	2:B:908:LEU:HG	1.35	1.08
2:B:147:LEU:HB3	2:B:148:PRO:HD2	1.16	1.08
2:B:147:LEU:CB	2:B:148:PRO:HD2	1.97	0.95
1:A:651:ASN:H	1:A:651:ASN:HD22	0.96	0.94
1:A:359:LYS:HD2	1:A:607:ALA:HB1	1.49	0.92

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	695/768 (90%)	631 (91%)	52 (8%)	12 (2%)	11 32
2	B	734/926 (79%)	651 (89%)	55 (8%)	28 (4%)	4 11
All	All	1429/1694 (84%)	1282 (90%)	107 (8%)	40 (3%)	6 18

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	115	ASN
1	A	120	VAL
1	A	234	VAL
1	A	254	ALA
1	A	678	TYR

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	621/668 (93%)	595 (96%)	26 (4%)	36	68
2	B	672/819 (82%)	652 (97%)	20 (3%)	48	80
All	All	1293/1487 (87%)	1247 (96%)	46 (4%)	42	75

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	647	SER
1	A	683	GLN
2	B	761	ASP
1	A	651	ASN
1	A	674	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	760	GLN
2	B	352	ASN
2	B	857	ASN
2	B	72	GLN
2	B	372	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.