SUPPLEMENTARY MATERIALS AND METHODS

Cloning and expression

Recombinant *Saccharomyces cerevisae* Cep3p was produced in *E. coli*. The coding sequences for full-length Cep3p and Cep3pΔ were cloned into pET28 vector using NdeI and EcoRI restriction sites, incorporating an N-terminal His-tag sequence and subsequent thrombin cleavage site. Each final construct was transformed into BL21(DE3) codon$^+$ cells. Recombinant Cep3p was expressed in TB medium containing 25mg/l kanamycin using 0.5mM IPTG, while Cep3pΔ was expressed using Overnight Express TB medium (Novagen) containing 25mg/l kanamycin. The cells were harvested at 4ºC by centrifugation at 3,300 x g prior to re-suspension in 60mM Tris pH 8.0, 0.5M NaCl, 0.5% Tween 20, 8% (w/v) glycerol, 5mM β-mercaptoethanol, with 1 protease inhibitor cocktail tablet (Roche) per 40ml. Recombinant protein was extracted by sonication on ice for a total of 210 seconds. Cell debris was removed by 26,000 x g centrifugation at 4ºC prior to protein precipitation with 65% ammonium sulphate. The precipitant was recovered by 20,000 x g centrifugation at 4ºC and re-suspended in Ni-run buffer (60mM Tris pH 8.0, 0.5M NaCl, 50mM imidazole, 5mM β-mercaptoethanol). The tagged protein was isolated using a 5ml HisTrap FF column (GE Healthcare) pre-equilibrated with Ni-run buffer, on which the sample was washed with Ni-run buffer containing 10mM $MgCl_2$ and 5mM ATP, prior to elution with strip buffer (60mM Tris pH 8.0, 0.5M NaCl, 100mM EDTA, 5mM β-mercaptoethanol). Each sample was further purified by size exclusion chromatography using a Superdex 200 16/60 prep grade column (GE Healthcare) pre-equilibrated with gel-filtration buffer (50mM Tris pH 8.5, 200mM NaCl, 5mM β-mercaptoethanol). The N-terminal His-tag was cleaved using thrombin

(GE Healthcare) at 4ºC, Cep3pΔ was then passing through a 1ml HisTrap FF column (GE Healthcare) pre-equilibrated with Ni-run buffer. Size exclusion chromatography was then performed with gel-filtration buffer containing 1μM $ZnCl_2$ (Cep3p) or 1mM EDTA (Cep3pΔ). The final samples were either frozen in 4% (w/v) glycerol for EMSA binding studies or concentrated to 5.5mg/ml (Cep3pΔ) for crystallization.

Selenomethionine incorporation into recombinant Cep3pΔ (SeMet-Cep3pΔ) was achieved using the same construct as previously described transformed into the $B834^{+}$ plus *E. coli* strain. Recombinant SeMet-Cep3pΔ was expressed using minimal medium containing L-selenomethionine and 25mg/l kanamycin. Extraction and purification of SeMet-Cep3pΔ was achieved as previously described for Cep3pΔ.

Electrophoretic mobility shift assays

EMSA binding assays were performed using the EMSA Accessory Kit (Novagen) substituting the 4x binding buffer for 80mM HEPES pH 8.0, 400mM KCl, 10% (w/v) glycerol. All DNA used was purchased from Sigma GENOSYS, re-suspended in annealing buffer (23mM Tris pH 7.5, 230mM NaCl) and annealed with its complementary strand by heating at 65ºC for 15 minutes before allowing to cool using a water bath. FAM6 fluorescently labelled native CDEIII dsDNA (5'-AAATATTAGTGTATTTGATTTCCGAAAGTTAAA-3') probe was used to observe DNA-protein binding. Four unlabelled competitor versions of the DNA were designed with mutations at the proposed Cep3p binding sites; native CDEIII (5'-AAATATTAGTGTATTTGATTTCCGAAAGTTAAA-3'), site 1 mutant (5'-AAATATTAGTGTATTTGATTTGATAAAGTTAAA-3'), site 2 mutant (5'-AAATATTAGAACTTTTGATTTCCGAAAGTTAAA-3'), site 1 and 2 mutant (5'-AAATATTAGAACTTTTGATTTGATAAAGTTAAA-3'). Each binding reaction

was setup as per the instructions with 4.2μM fluorescent probe incubated with either Cep3pΔ (4.2, 8.3, 12.7μM), Cep3p (4.2, 8.3, 12.7μM), or Cep3p (12.7μM) and competitor DNA (3, 6, or 12x probe concentration) at 16ºC for 1 hour before loading onto a 6% non-denaturing DNA retardation gel (at 4ºC) pre-run using 0.5x TBE for 1 hour at 3mA. Samples were then run for 80-90 minutes at 6mA before imaging using a Storm 860 Molecular Imager. Analysis was carried out using ImageQuant.

Crystallisation, data collection and processing

Cep3pΔ crystals were obtained using the hanging drop vapour diffusion method with a 2μl drop containing equal amounts of mother liquor (100mM sodium cacodylate pH 6.5, 0.2M potassium thiocyanate, and 12-15% (w/v) PEG 4000) and protein solution at 16°C. The protein crystallised in space group $P4_32_12$, with unit cell dimensions a = b = 83.7 Å, c = 231.4 Å, with one monomer in the asymmetric unit (molecular weight 66.31 kDa). Diffraction data were collected at cryogenic temperatures from a single crystal at the European Synchrotron Radiation Facility (ESRF) on beamline ID29 (wavelength 1.0000 Å). Crystals were cryoprotected with 23% PEG 400 before flash freezing at 100K prior to data collection. Mercury and silver derivatives were prepared by soaking crystals in mother liquor containing 10mM thimerosal for 160 minutes or 5mM silver nitrate for 125 minutes before cryoprotecting and flash freezing for data collection. SeMet-Cep3pΔ crystals were obtained under differing mother liquor conditions (100mM MES pH 6.5, 10.5% PEG 20,000). The protein crystallised in spacegroup $P4_32_12$, with unit cell dimensions (a = b = 84.9 Å, c = 231.3 Å). Single-wavelength Anomalous Dispersion (SAD) method was used to collect phase information from a single crystal at the ESRF on beamline ID29 (wavelength 0.9793 Å). The data sets were processed using MOSFLM (Steller et al, 1997) before

scaling and merging using SCALA (CCP4, 1994). Statistics of data collection are summarised in table 1.

Phase determination, model building and refinement

Phasing was carried out using SAD and MIR methodology. A total of 12 Se positions were determined from the SAD data which along with 2 Hg and 1 Ag positions from the derivatives, treating the SeMet-Cep3pΔ dataset as native, was sufficient to phase the data using the program SOLVE (Terwilliger & Berendzen, 1999). Phases were extended to 2.5 Å against the native data and maps solvent flattened using RESOLVE (Terwilliger, 2000). Automated tracing was performed using BUCCANEER (Cowtan, 2006) and RESOLVE with the consensus model rebuilt manually using COOT (Emsley & Cowtan, 2004). The structure was refined using CNS (Brünger et al, 1998) Cartesian dynamics protocol between rounds of manual rebuilding. Once the structure was nearing completion individual isotropic B-factors were also refined before well-defined water molecules and other solvent ligands were added. The final model of Cep3pΔ contains 4318 protein atoms, 78 water molecules, 1 β-mercaptoethanol and 1 cacodylate molecule with R-factor 22.0% (R-free 24.6%). A summary of the refinement statistics and stereochemistry analysis is given in table1. Structural comparisons to the protein database were carried out using the DALI server (Holm & Sander, 1993). Electrostatic surface calculations and representation were performed using the CCP4 molecular graphics package (Potterton et al, 2004).

Table 1. Crystallographic statistics for the structure determination.

	Native	**Se-met**	**Hg**	**Ag**
Resolution range (Å)	42-2.5	42-3.2	28-3.5	30-3.8
Completeness (%)	99.6	99.7	99.9	99.8
R_{symm} (overall / highest) (%)	6.5 (24.7)	9.0 (29.6)	8.0 (31.4)	10.3 (29.5)
I/(σ)I (overall / highest)	8.2 (3.0)	6.4 (2.5)	8.6 (2.4)	7.1 (2.6)
R_{deriv} (%)	-	-	28	26
No. of sites	-	12	2	1
Overall mean figure of merit		0.47		
R-factor (%)	22.0			
R_{free} (%)	24.6			
R.m.s.d. bond lengths (Å)	0.010			
R.m.s.d. bond angles (deg)	1.3			

REFERENCES

Brünger A, Adams P, Clore G, DeLano W, Gros P, Grosse-Kunstleve R, Jiang J, Kuszewski J, Nilges M, Pannu N, Read R, Rice L, Simonson T, Warren G (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54:** 905-921

Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* **62:** 1002-1011

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126-2132

Holm L, Sander C (1993) Protein structure comparison by alignment of distance matrices. *J Mol Biol* **233:** 123-138

Potterton L, McNicholas S, Krissinel E, Gruber J, Cowtan K, Emsley P, Murshudov G, Cohen S, Perrakis A, Noble M (2004) Developments in the CCP4 molecular-graphics project. *Acta Crystallogr D Biol Crystallogr* **60:** 2288-2294

Steller I, Bolotovsky R, Rossmann MG (1997) An Algorithm for Automatic Indexing of Oscillation Images using Fourier Analysis. *Journal of Applied Crystallography* **30:** 1036-1040

Terwilliger T (2000) Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr* **56:** 965-972

Terwilliger T, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* **55:** 849-861