Supplementary figures

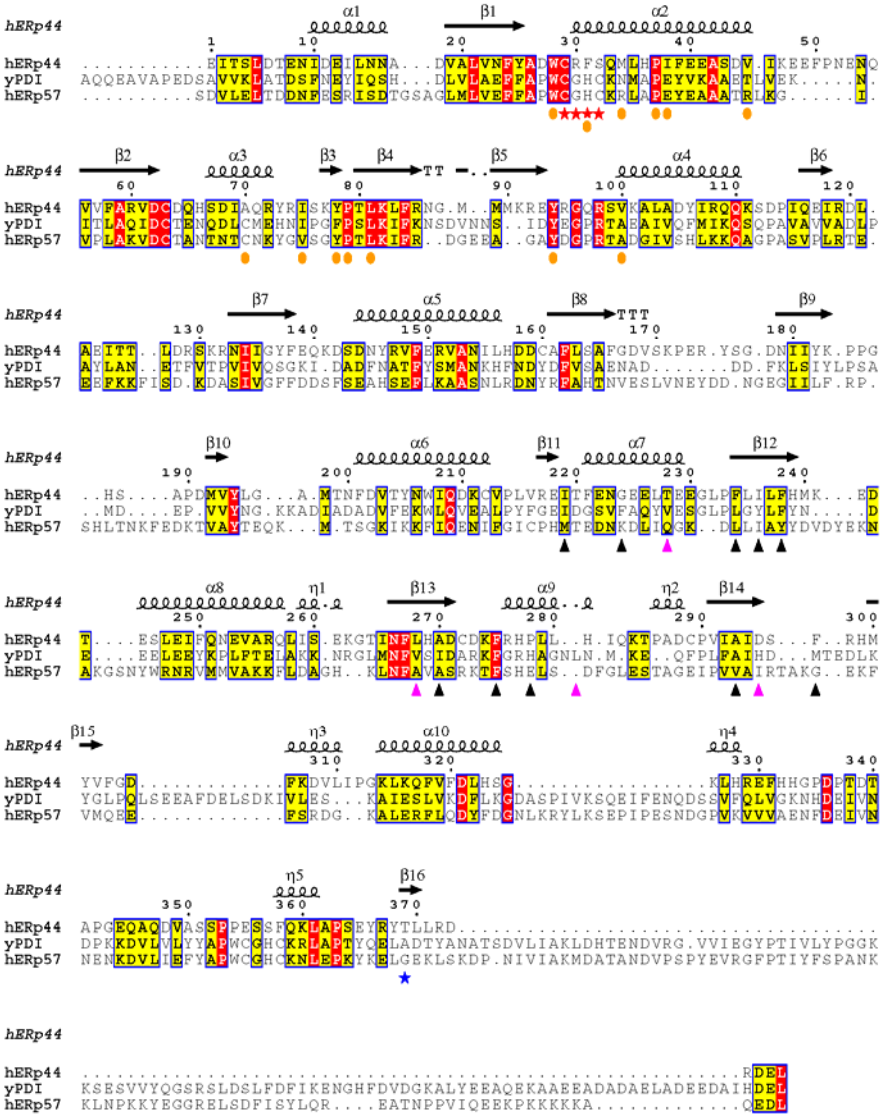

Fig S1: Conserved hydrophobic surfaces in ERp44 and PDI.

Sequence alignment (Thompson *et al*, 1994) based on the structure of hERp44 (PDB code: 2R2J), yPDI (2B5E) and hERp57 (2H8L for domains **b** and **b'**, 2ALB for domain **a**) were adjusted manually, and colored with ESPript 2.1 (Gouet *et al*, 1999). Secondary structure elements are labeled according to hERp44. Identical residues are boxed in red and residues with similarity >50% are in yellow. Black triangles mark the residues forming the hydrophobic pocket in the **b'** domains of both hERp44 and yPDI, and pink triangles for only yPDI. Red stars mark the active site in domain **a**. For hERp44 blue star marks Thr369, which interacts with Cys29, and orange circles represent the residues constituting the hydrophobic patch

surrounding the CRFS motif. hERp44, human ERp44; yPDI, yeast PDI; hERp57, human ERp57.

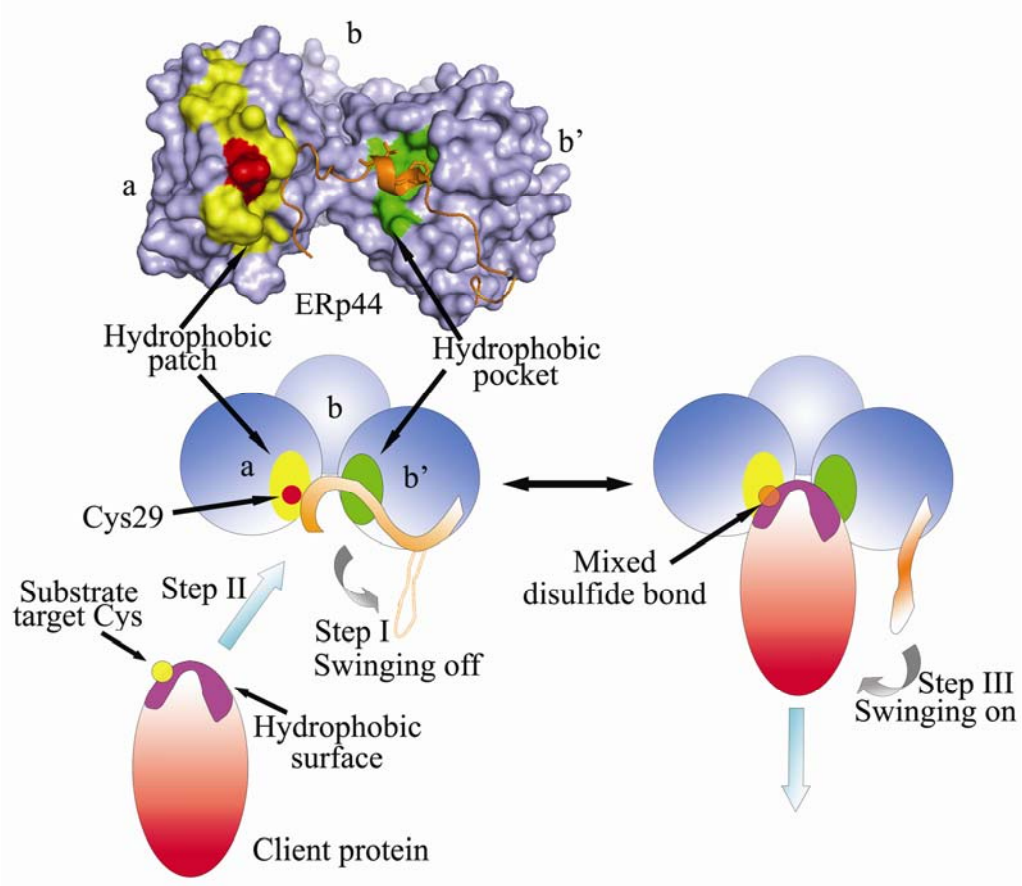

Fig S2: A possible model of the regulatory role of the C-tail in modulating substrate binding.

The top highlights the structure of ERp44, showing the hydrophobic pocket (green) in the **b'** domain, the hydrophobic patch (yellow) surrounding the CRFS motif (red) in the **a** domain, and the C-tail (orange ribbon). The cartoons below depict a schematic representation of substrate binding. The C-tail may swing off (step I) exposing the hydrophobic pocket in domain **b'** and the hydrophobic patch surrounding the CRFS motif in domain **a**. The swinging could be mediated by pH, Ca^{2+}, redox state or by interacting proteins. In step II, client proteins could then interact with ERp44 leaning upon the hydrophobic surfaces and then binding covalently to Cys29. In step III, the tail swings onto the **a** domain, expelling the client protein and blocking the two hydrophobic surfaces, switching ERp44 to a less-active state. Surface representation of hERp44 is made using Pymol (http://pymol.sourseforge.net). hERp44, human ERp44.

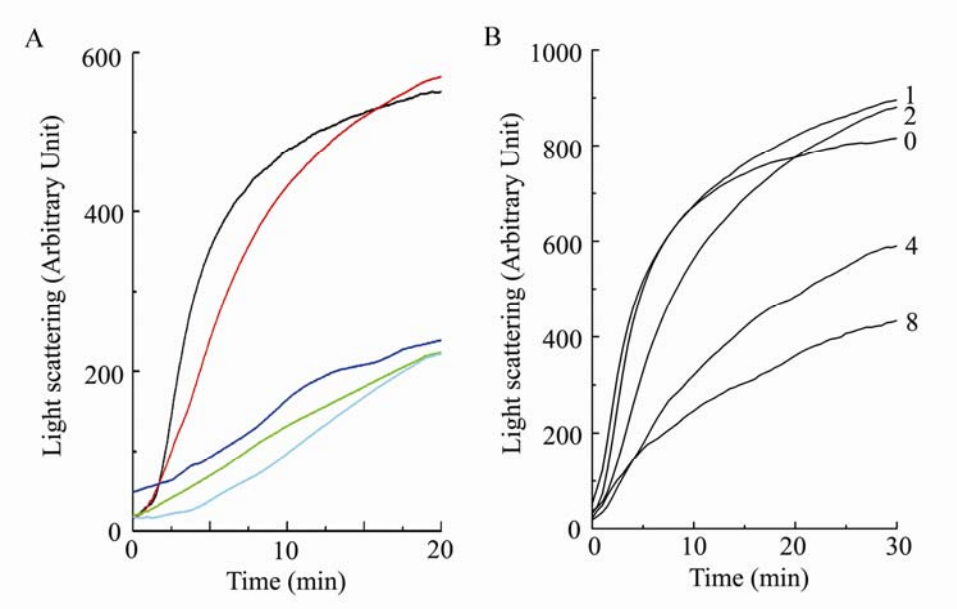

Fig S3: **Removal of the flexible C-tail rescues chaperone activity of hERp44.**

Time courses of light-scattering changes produced by the aggregation of denatured rhodanese during refolding upon dilution in the absence (black) and presence of hERp44 (red), hERp44Δ(331-377) (green), hERp44-T369A (dark blue) or hPDI (light blue) at a molar ratio to rhodanese of 8 (**A**) or in the presence of hERp44Δ(331-377) at different molar ratios as indicated (**B**). The tail mutants suppressed rhodanese aggregation almost as efficiently as PDI. Both the rate and extent of rhodanese aggregation decreased with increase of the concentrations of hERp44Δ(331-377).

hERp44, human ERp44; yPDI, yeast PDI.

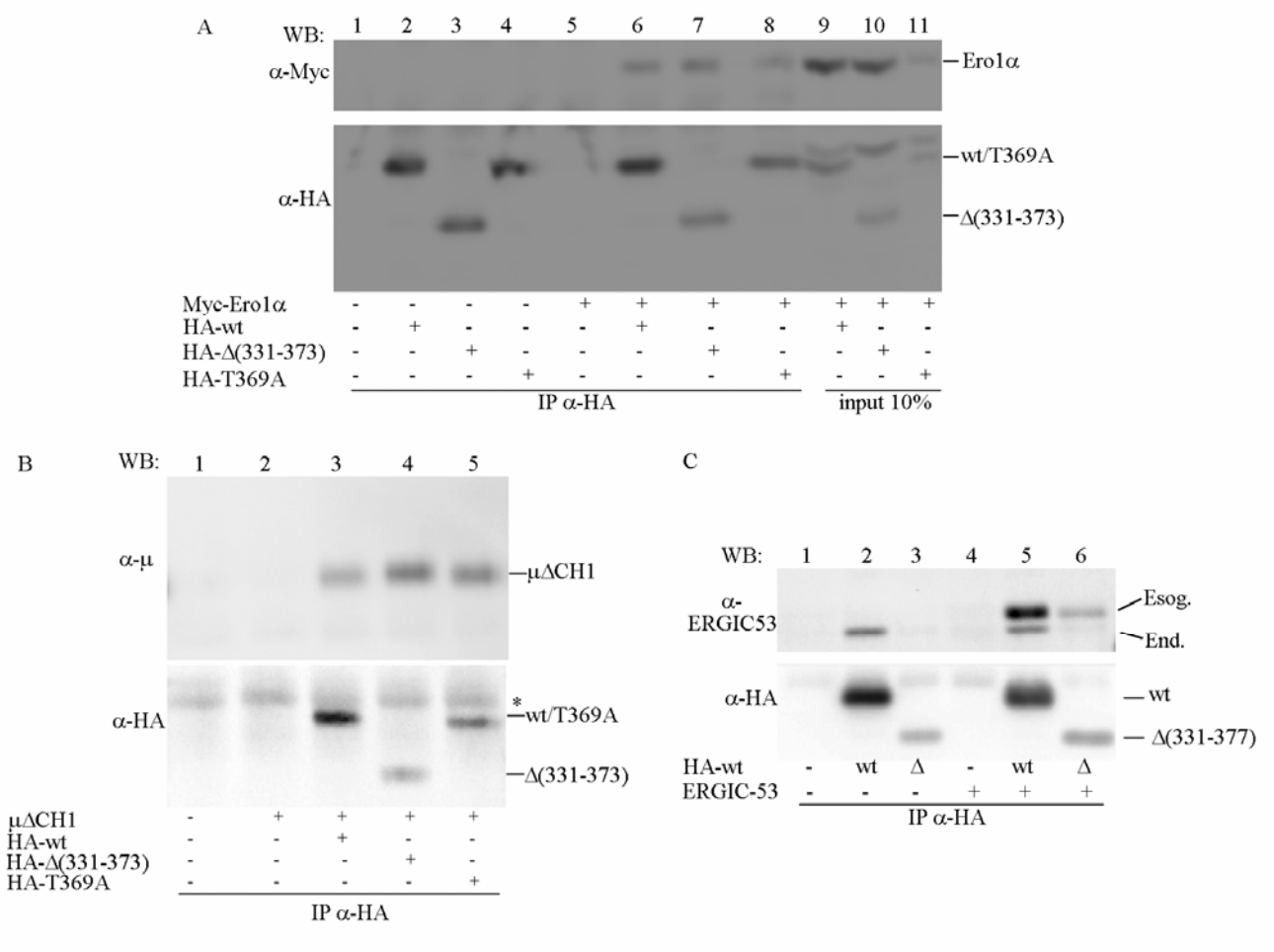

Fig S4: Deletion of the C-tail does not alter hERp44 binding specificities

HeLa cells transfected so as to co-express hERp44, hERp44Δ(331-377) or hERp44T369A and Ero1α (**A**), μΔCH1 (**B**) or ERGIC-53 (**C**) were lysed and immunoprecipitated with anti-HA antibodies. After washing, immunoprecipitates were resolved by SDS-12% PAGE under reducing conditions, transferred to nitrocellulose and decorated with suitable antibodies to reveal associations (Anelli *et al*, 2003; Anelli *et al*, 2002; Anelli *et al*, 2007).

A. The immunoprecipitates from the total lysates of 10^6 HeLa transfected as indicated, were washed thoroughly and resolved by SDS-12% PAGE under reducing conditions. Blots were hybridized with anti-Myc to reveal co-immunoprecipitated Ero1α, or with anti-HA antibodies to follow exogenous hERp44. Lanes 9-11 (input) show an aliquot of the lysates (10^5 cells) before immunoprecipitation, to monitor the expression levels in the different transfectants. The

efficiency of co-immunoprecipitation was determined by densitometric analyses of the Ero1α and hERp44 bands. Note that the tail ERp44 mutants interact more strongly with Ero1α than wild-type hERp44 does. hERp44Δ(331-373) co-precipitated a larger fraction of Ero1α than hERp44T369A. Similar results were obtained in two independent experiments performed with the same protocol.

B. The experiment shown, a representative of three similar ones, was performed according to a protocol similar to the one described in panel A, using HeLa transfectants expressing hERp44 variants and a known substrate, Ig-μΔCH1 chains (Anelli *et al*, 2003; Anelli *et al*, 2007). Blots were hybridized with anti-μ and anti-HA antibodies. The band indicated with an asterisk consists of the heavy chain of the anti-HA antibody used for immunoprecipitation, that is recognized by the secondary antibody used for detection. As described for Ero1α, more μΔCH1 chains are co-precipitated by the tail mutants than by wild-type hERp44. hERp44Δ(331-373) was slightly more efficient than hERp44T369A in co-precipitating μΔCH1.

C. After cross-linking with DSP, a step that increases the efficiency of ERGIC-53/hERp44 co-precipitation (Anelli *et al*, 2007), the lysates from HeLa cells transfected as indicated were immunoprecipitated with anti-HA, and blots decorated with anti-ERGIC-53 or anti-HA antibodies. Taking into account the different expression levels of wt hERp44 and hERp44Δ(331-373), the results clearly indicate that similar amounts of ERGIC-53 (either endogenous or myc-tagged, glycosylated overexpressed molecules, Anelli *et al*, 2007) are co-precipitated, in line with the assumption that the association between the two proteins does not involve their substrate-binding domains and C29 (Anelli *et al*, 2007). To visualize ERGIC-53, anti-ERp58 (Neve *et al*, 2005) was generously provided by Drs. E. Neve and R. Petterson (Stockholm, Sweden).

hERp44, human ERp44; wt, wild-type hERp44; Δ(331-373), hERp44Δ(331-373),T369, hERp44T369; Δ, hERp44Δ(331-373) ; IP, immunoprecipitation; WB, western-blot; Myc, c-myc epitope tag; HA, hemagglutinin epitope tag; DSP, Di(N-succinimidyl) 3,3'-dithiodipropionate Dithiobis (succinimidyl propionate); Esog, exogenous; End, endogenous.

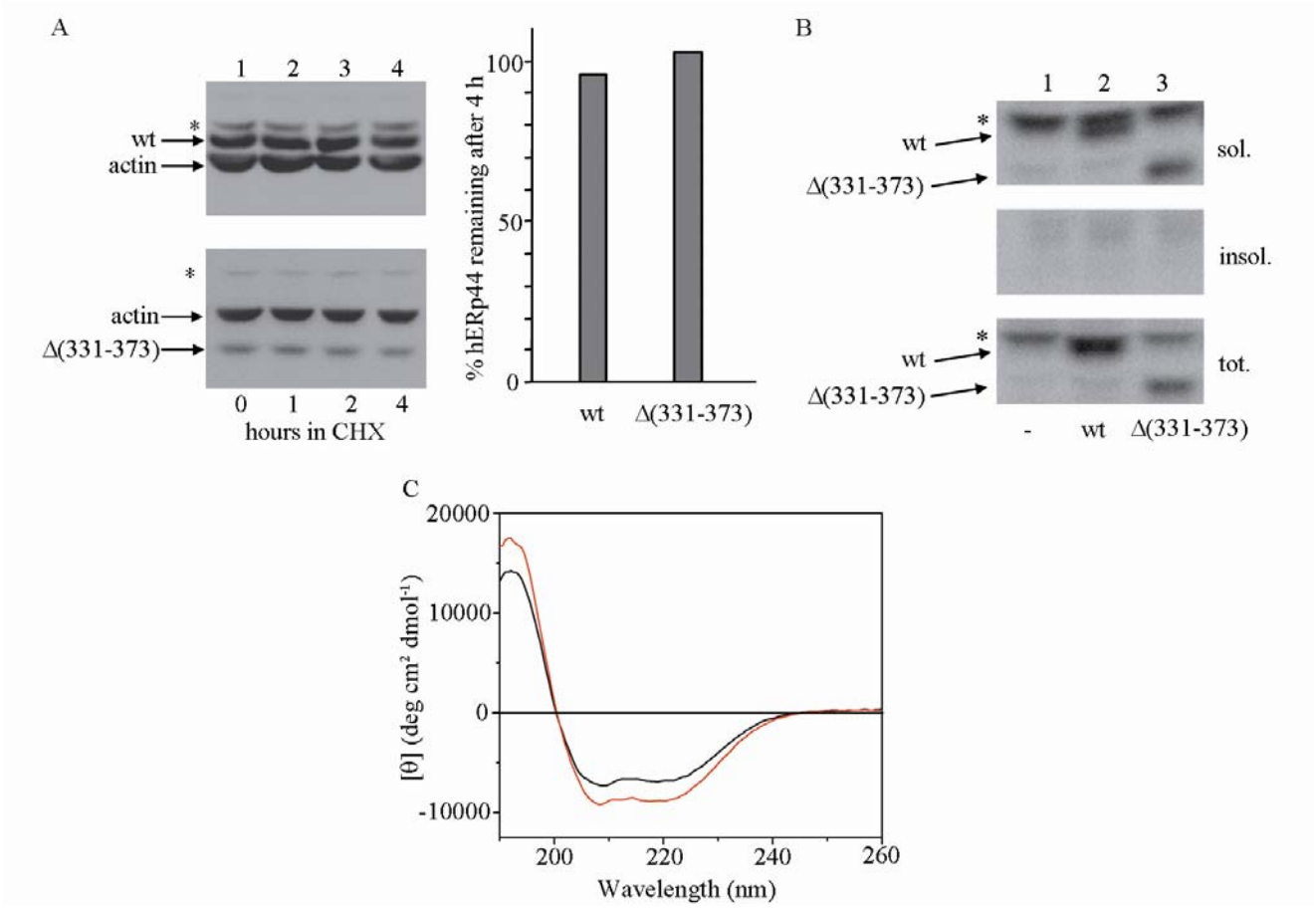

Fig S5: **Deletion of the C-tail does not cause gross misfolding of hERp44.**

A. Stability. HeLa cells expressing hERp44 or hERp44Δ(331-373) were treated with cycloheximide (CHX) for 1, 2 or 4 h as indicated at 48 h after transfection. Aliquots of cell lysates were resolved under reducing conditions in SDS-12% PAGE, blotted and decorated with monoclonal anti-HA and α-actin antibodies (left panels). The intensity of the hERp44 bands was determined by densitometry and normalized against actin. Asterisks mark anti-HA reactive background bands. The plot in the right panel shows the percentage of hERp44 remaining in cells after 4 h of CHX treatment. The intensity of short-lived molecules (i.e. Ik-Ba, orphan Ig subunits) decreased significantly in similar assays (Mancini *et al*, 2000; Cenci *et al*, 2006; our unpublished results).

B. Detergent solubility. HeLa cells expressing hERp44 or hERp44Δ(331-373) were lysed in NP-40 (upper two panels) or SDS (lower panel) at 48 h after transfection. NP-40 soluble and

insoluble fractions were separated by centrifugation, and aliquots corresponding to the same number of cells resolved by SDS-12% PAGE, blotted and decorated with anti-HA antibody. Note that neither hERp44 nor the hERp44Δ(331-373) mutants are detectable in the insoluble fraction when over-expressed in HeLa cells. Asterisks mark anti-HA reactive background bands.

C. Circular dichroism spectra. The far UV circular dichroism spectra of hERp44 (black) and hERp44Δ(331-377) (red) were determined at 20°C with 8 scans accumulated. The very similar spectra indicated that the removal of the C-tail does not significantly alter secondary structures.

wt, wild-type hERp44; Δ(331-373), hERp44Δ(331-373); hERp44, human ERp44; HA, hemagglutinin epitope tag; sol, soluble; insol, insoluble; tot, total.

Supplementary Table S1: Data collection, phasing and refinement statistics

Data collection	
Space group	$P3_121$
Cell dimensions a, b, c (Å) α, β, γ (°)	 83.5, 83.5, 123.2 90, 90, 120
Wavelength (Å)	0.9790
Resolution (Å)[1]	50-2.6 (2.69 – 2.60)
R_{merge}[1]	0.054 (0.394)
$< I > / <\sigma I>$[1]	26.7 (2.5)
Completeness (%)[1]	99.5 (96.6)
Redundancy[1]	6.3 (3.9)
Phasing	
Se atoms found	8 (total 8)
Initial phases	
Resolution (Å)	50-2.6
FOM	0.31
Centric	0.06
Acentric	0.35
Density modified phases	
Resolution (Å)	50.0-2.6
FOM[2]	0.6 (0.65)
Refinement	
Resolution (Å)	50-2.6
No. reflections	14,916
R_{work}/R_{free} (%)	21.7/26.3
rmsd. Bonds (Å)	0.016
rmsd. Angles (º)	1.95
No. of protein atoms	2,703
No. of ligand atoms	14
No. of waters	116
B-factors (Å^2)	
Total	79.8

Domain **a**	82.2
Domain **b**	86.0
Domain **b'**	75.0
C-tail	74.0
Solvent	75.8
Ramachandran statistics (Lovell *et al*, 2003)	
Residues in favored regions (%)[3]	93.4 (299/320)
Residues in allowed regions (%)[3]	98.4 (315/320)
Outliers (%)[3]	1.6 (5/320)

[1]Highest-resolution shell is shown in parentheses.
[2]The improved figure of merit after using OASIS-2004 (Yao *et al*, 2006) is shown in parentheses.
[3]The exact number of residues over the number of all residues in the model is shown in parentheses.

Supplementary Methods

Structure determination

SAD data were processed with the HKL2000 suite of programs (Otwinowski & Minor, 1997). Eight selenium atoms were located using SHELXD (Schneider & Sheldrick, 2002) (best solution, CCall=53.0, CCweak=29.5, PATFOM=26.1). SOLVE (Terwilliger, 2000) was used to calculate the initial phases to 2.6 Å with a mean figure of merit (FOM) of 0.31. Density modification through solvent flattening (solvent content 0.55) and phase extension were performed by using RESOLVE (Terwilliger, 2000) to yield an interpretable 2.6 Å electron density map with FOM of 0.6. The initial model was built manually using O (Jones *et al*, 1991), and then was used to improve the phase by using OASIS-2004 (Yao *et al*, 2006) (KMI=0.01) and RESOLVE (Terwilliger, 2000) (final FOM=0.65). COOT (Emsley & Cowtan, 2004) was then utilized to complete the model building coupled to interactive refinement with Refmac (Murshudov *et al*, 1997). Refmac5.0 (Murshudov *et al*, 1997) was used to refine the model with the least squares method, coupled to the refinement of the TLS tensors. The whole model was separated into four TLS groups: domain **a**, **b**, **b'**, C-tail. The final model was refined to 2.6 Å with Rwork 21.7% and Rfree 26.3%. Residues 332-350 and 373-377 were not traceable due to lack of electron density, and residues 1-2, 51-53, 119-130 and 170-177 are

highly flexible with blurry and poor associated density, therefore all these fragments were excluded from the final model. Structure determination statistics are shown in Table S1.

Determination of mixed disulfides with endogenous proteins in cells. For HA-hERp44 and HA-hERp44Δ(331-373), the C-terminal RDEL sequence was left in place to prevent secretion, and an HA tag was inserted immediately after the endogenous cleaved leader sequence (Anelli *et al*, 2002). The tag does not influence the function of hERp44 in any of the available assays (T. Anelli, S. Ceppi, L. Bergamelli & R. Sitia, unpublished data). The PCR product of HA-hERp44 was re-inserted in pcDNA3.1(-) using the restriction sites XhoI-Acc65I. The construct was entirely re-sequenced before use in transfection experiments.
Primers for PCR amplification of HA-hERp44Δ(331-373) from HA-hERp44 were as follows:
TA2S (CAACTCGAGCGTTACCATGCATCCTGCC);
P44DtailC (CACGGTACCTTAAAGCTCATCTCGTTCTCTGTGCAGTTTTCCAGAA).

References

Anelli T, Alessio M, Bachi A, Bergamelli L, Bertoli G, Camerini S, Mezghrani A, Ruffato E, Simmen T, Sitia R (2003) Thiol-mediated protein retention in the endoplasmic reticulum: the role of ERp44. *EMBO J* **22**: 5015-5022

Anelli T, Alessio M, Mezghrani A, Simmen T, Talamo F, Bachi A, Sitia R (2002) ERp44, a novel endoplasmic reticulum folding assistant of the thioredoxin family. *EMBO J* **21**: 835-844

Anelli T, Ceppi S, Bergamelli L, Cortini M, Masciarelli S, Valetti C, Sitia R (2007) Sequential steps and checkpoints in the early exocytic compartment during secretory IgM biogenesis. *EMBO J* **26**: 4177-4188

Cenci S, Mezghrani A, Cascio P, Bianchi G, Cerruti F, Fra A, Lelouard H, Masciarelli S, Mattioli L, Oliva L, Orsi A, Pasqualetto E, Pierre P, Ruffato E, Tagliavacca L, Sitia R.(2006) Progressively impaired proteasomal capacity during terminal plasma cell differentiation. *EMBO J* **25**:1104-1113.

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

Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**: 305-308

Jones TA, Zou JY, Cowan SW, Kjeldgaard (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47**: 110-119

Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **50**: 437-450

Mancini R, Fagioli C, Fra AM, Maggioni C, Sitia R (2000) Degradation of unassembled soluble Ig subunits by cytosolic proteasomes: evidence that retrotranslocation and degradation are coupled events. *FASEB J* **14**: 769-78.

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**: 240-255

Neve EP, Lahtinen U, Pettersson RF (2005) Oligomerization and interacellular localization of the glycoprotein receptor ERGIC-53 is independent of disulfide bonds. *J Mol Biol* **354**: 556-68.

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. In *Macromolecular Crystallography, part A*, Carter Jr. CW, Sweet RM (eds), Vol. 276, pp 307-326. New York: Academic Press

Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* **58**: 1772-1779

Terwilliger TC (2000) Maximum-likelihood density modification. *Acta crystallographica* **56**: 965-972

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-4680

Yao DQ, Huang S, Wang JW, Gu YX, Zheng CD, Fan HF, Watanabe N, Tanaka I (2006) SAD phasing by OASIS-2004: case studies of dual-space fragment extension. *Acta Crystallogr D Biol Crystallogr* **62**: 883-890