ARTICLE
Journal of Cellular Biochemistry 107:1002–1015 (2009)

Journal of Cellular Biochemistry

Identification and Characterization of the Mitochondrial Targeting Sequence and Mechanism in Human Citrate Synthase

Tsung-Lin Cheng,[1] Ching-Chun Liao,[2] Wen-Hui Tsai,[3,4] Chin-Chih Lin,[2] Chin-Wei Yeh,[2] Chiao-Fang Teng,[2] and Wen-Tsan Chang[1,2*]

[1]*Institute of Basic Medical Sciences, National Cheng Kung University Medical College, Tainan, Taiwan 701, ROC*

[2]*Department of Biochemistry and Molecular Biology, National Cheng Kung University Medical College, Tainan, Taiwan 701, ROC*

[3]*Institute of Clinical Medicine, National Cheng Kung University Medical College, Tainan, Taiwan 701, ROC*

[4]*Department of Pediatrics, Chi Mei Foundation Medical Center, Tainan, Taiwan 710, ROC*

ABSTRACT

Citrate synthase (CS), the first and rate-limiting enzyme of the tricarboxylic acid (TCA) cycle, plays a decisive role in regulating energy generation of mitochondrial respiration. Most mitochondrial proteins are synthesized in the cytoplasm as preproteins with an amino (N)-terminal mitochondrial targeting sequence (MTS) that directs mitochondria-specific sorting of the preprotein. However, the MTS and targeting mechanism of the human CS protein are not fully characterized. The human CS gene is a single nuclear gene which transcribes into two mRNA variants, isoform a (CSa) and b (CSb), by alternative splicing of exon 2. CSa encodes 466 amino acids, including a putative N-terminal MTS, while CSb expresses 400 residues with a shorter N terminus, lacking the MTS. Our results indicated that CSa is localized in the mitochondria and the N-terminal 27 amino acids, including a well-conserved RXY↓(S/A) motif (the RHAS sequence), can efficiently target the enhanced green fluorescent protein (EGFP) into the mitochondria. Furthermore, site-directed mutagenesis analysis of the conserved basic amino acids and serine/threonine residues revealed that the R9 residue is essential but all serine/threonine residues are dispensable in the mitochondrial targeting function. Moreover, RNA interference (RNAi)-mediated gene silencing of the preprotein import receptors, including TOM20, TOM22, and TOM70, showed that all three preprotein import receptors are required for transporting CSa into the mitochondria. In conclusion, we have experimentally identified the mitochondrial targeting sequence of human CSa and elucidated its targeting mechanism. These results provide an important basis for the study of mitochondrial dysfunction due to aberrant CSa trafficking. J. Cell. Biochem. 107: 1002–1015, 2009.

KEY WORDS: CITRATE SYNTHASE (CS); TRICARBOXYLIC ACID (TCA) CYCLE; MITOCHONDRIAL TARGETING SEQUENCE (MTS); PREPROTEIN IMPORT RECEPTORS; RNA INTERFERENCE (RNAi)

The bioenergetic organelle mitochondria are not only essential for energy production in cells, but also important for cell metabolism and apoptosis [Wallace, 1999; DiMauro and Schon, 2003; Orrenius et al., 2007; Oberst et al., 2008]. These organelles are bounded by double-membraned envelope consisting of the outer and inner membranes. Mitochondria contain nearly one thousand different proteins that are localized in four distinct compartments, the outer membrane, the inner membrane, the intermembrane space, and the matrix [Taylor et al., 2003; Prokisch et al., 2004, 2006]. The vast majority of mitochondrial proteins are encoded by nuclear genes and translated on cytosolic ribosomes as preproteins containing either N-terminal targeting sequences or internal targeting sequences, which direct transportation of the preproteins to different mitochondrial compartments [Hurt et al., 1984;

Tsung-Lin Cheng and Ching-Chun Liao contributed equally to this work.

Grant sponsor: National Science Council, Taiwan, ROC; Grant number: NSC 95-2320-B-006-078-MY3; Grant sponsor: Department of Health, Taiwan, ROC; Grant number: DOH-TD-B-111-004.

*Correspondence to: Prof. Wen-Tsan Chang, DPhil, Department of Biochemistry and Molecular Biology, National Cheng Kung University Medical College, No. 1, University Road, Tainan, Taiwan 701, ROC.
E-mail: wtchang@mail.ncku.edu.tw

Received 18 December 2008; Accepted 8 April 2009 • DOI 10.1002/jcb.22200 • © 2009 Wiley-Liss, Inc.
Published online 28 May 2009 in Wiley InterScience (www.interscience.wiley.com).

Sickmann et al., 2003; Dolezal et al., 2006; Neupert and Herrmann, 2007].

After released from the ribosome in the cytosol, the newly synthesized preproteins are immediately bound by molecular chaperones that maintain the preproteins in the unfolded and unassembled states and prevent their misfolding and aggregation [Terada et al., 1995; Mori and Terada, 1998; Terada and Mori, 2000; Young et al., 2003]. These preproteins are then transferred to preprotein import receptors including TOM20 [Goping et al., 1995; Seki et al., 1995; Hanson et al., 1996] and TOM70 [Alvarez-Dolado et al., 1999; Suzuki et al., 2002] at the outer membrane. Generally, the preproteins containing N-terminal targeting signals are first recognized by the preprotein import receptor TOM20 and then by TOM22, while the preproteins containing specific internal targeting signals are primarily recognized by the preprotein import receptor TOM70 and then transferred to TOM22 [Brix et al., 1997, 1999]. They are subsequently targeted to the general import pores consisting of the core component TOM40 and the associated modulators TOM5, TOM6, and TOM7, which initiate translocation of the preproteins across the outer membrane [Neupert, 1997; Schatz, 1997; Bains and Lithgow, 1999]. Once the N-terminal targeting sequences have crossed over the outer membrane, the preproteins are inserted into protein conducting channels in the inner membrane [Neupert, 1997; Pfanner et al., 1997; Koehler, 2004].

CS catalyzes the first reaction of the TCA cycle that regulates mitochondrial respiration and is generally assumed to be the rate-limiting enzyme of the cycle [Suissa et al., 1984]. The mitochondrial CS proteins are encoded in nuclear genomes, translated on cytosolic ribosomes, and then post-translationally transported into the organelle. However, the mechanism by which the human CS protein is targeted to its subcellular location remains unknown. To elucidate the mitochondrial targeting mechanism of the human CS protein, we first examined its subcellular localization in the human cervical carcinoma cell line HeLa, then identified its MTS on the N-terminal region, and analyzed the mitochondrial targeting function of the conserved amino acids in the MTS. To better understand the mitochondrial targeting mechanism, we also examined the efficiency of the mitochondrial targeting of CS under the inhibition of the preprotein import receptors, including TOM20, TOM22, and TOM70, by using sequence-specific shRNAs.

MATERIALS AND METHODS

CELL CULTURE

The human cervical carcinoma derived cell line HeLa and the embryonic kidney epithelial cell line HEK293 were grown and maintained in Dulbecco's Modified Eagle's Medium (DMEM; GIBCO BRL, Gaithersburg, MD) supplemented with 10% heat inactivated fetal calf serum (FCS; BIOLOGICAL INDUSTRIES, Kibbutz Beit Haemek, Israel) and 1% antibiotic/antimycotic solution (GIBCO BRL) in a humidified incubator containing 5% CO_2 at 37°C. Both the cell lines were subcultured two to three times a week after trypsinization with 0.1% trypsin (Biowhittaker Acambrex, Walkersville, MD).

CELL TRANSFECTION

At 24 h before transfection, cells were directly seeded in 6-well plates or in plates which wells contained a sterile glass coverslip, at a density of 1×10^5 cells per well. The cells were transfected with a total of 2 μg of plasmid constructs using Lipofectamine 2000 (Invitrogen, Carlsbad, CA) according to the manufacturer's protocols. After incubation for various time periods, the transfected cells were processed by Western blotting, immunofluorescence staining, or imaging analysis.

RT-PCR OR REAL-TIME RT-PCR ANALYSIS OF THE CS TRANSCRIPTS

Total cellular RNAs were isolated by using TRIZOL™ reagent (GIBCO BRL) according to the manufacturer's protocols. Total RNA extracts (5 μg) were first reversely transcribed into complementary DNAs (cDNAs) by standard reverse transcriptase (RT) reaction with a poly(dT) primer, and then amplified by conventional PCR for 25 cycles or real-time PCR for 40 cycles with CSa- and CSb-specific primers as indicated in Figure 2C and 18S rRNA primers: 18S rRNA-F, 5′-CGGCTACCACATCCAAGGAA-3′; 18S rRNA-R, 5′-GCTGGAA-TTACC GCGGCT-3′ (215 bp product) as described in Dinchuk et al. [2002]. Conventional PCR was performed according to the following protocol: an initial denaturation at 95°C for 5 min; another 25 cycles of denaturation at 95°C for 1 min, annealing at 60°C for 30 s, and extension at 72°C for 30 s; final extension at 72°C for 10 min; last stop and keep at 4°C. The amplified products were analyzed on 1.2% agarose gel electrophoresis and then photographed using an agarose gel imaging and analysis system. The 18S rRNA served as a control for loading. Real-time PCR was carried out on an ABI 7500 Fast Real-Time PCR system (Applied Biosystems, Foster City, CA) with SYBR Green PCR Master Mix (Applied Biosystems), according to the recommended protocol of the manufacturer. Briefly, each reaction was run in triplicate and contained 1 μl of cDNA template along with 50 nM primers in a final reaction volume of 20 μl. Cycling parameters were 95°C for 10 min to activate DNA polymerase, then 40 cycles of 95°C for 15 s and 60°C for 1 min, with a final recording step of 72°C for 30 s to prevent any primer–dimer formation. Melting curves were calculated using Dissociation Curves software (Applied Biosystems) to ensure only a single product was amplified, and samples were also run on a 1.2% agarose gel to confirm specificity. Data analysis was conducted by adapting the ΔΔCt method of relative quantification to estimate copy numbers of CSa and CSb mRNAs. The result for each target mRNA, expressed as n-fold copy number of the target mRNA relative to 18S rRNA, was obtained through the equation $2^{-\Delta\Delta Ct}$, where $\Delta\Delta Ct = (Ct_{CSa} - Ct_{18S\ rRNA}) - (Ct_{CSb} - Ct_{18S\ rRNA})$.

CONSTRUCTION OF THE CSa-EGFP FUSION PROTEIN EXPRESSION VECTOR

In general, plasmid DNAs were constructed by standard molecular cloning approaches or PCR-based cloning strategies. The full-length cDNA clone (IMAGE Id: 3542491) encoding the human CSa protein was purchased from the Mammalian Gene Collection (MGC; http://mgc.nci.nih.gov/) (Invitrogen). Oligonucleotides used in this study were purchased from local commercial suppliers. The CSa-EGFP fusion protein expression vector pCSa-EGFP was constructed by inserting the full-length CSa coding sequence in-frame into the

pEGFP-N1 vector (Clontech, Palo Alto, CA). The full-length CSa coding sequence was amplified by PCR with the primers CSa-F: cgggatccaccATGGCTTTACTTACTGC GGCCGCC and CSa-R: ccg-ctcgagCCCTGACTTAGAGTCCACAAACTTCATCAG.

CONSTRUCTION OF MITOCHONDRION- AND PEROXISOME-SPECIFIC DsRed EXPRESSION VECTORS

The mitochondria-specific DsRed expression vector pDsRed2-Mito was constructed by inserting the sequence encoding the human cytochrome *c* oxidase (COX) N-terminal 86 amino acids in-frame into the pDsRed2-N1 vector (Clontech). The COX N-terminal 86 amino acids coding sequence was amplified by PCR using the sequence-specific primers COX-F: cccaagcttATGTCCGTCCTGACG-CCGCTG and COX-R: cgggatccCAACGAATG GATCTT GGCGC. The peroxisome-specific DsRed expression vector pDsRed2-Peroxi was constructed by fusing the sequence encoding peroxisomal targeting sequence type 1 (PTS1) SKL residues in-frame into the DsRed2 coding sequence. The full-length DsRed2-SKL coding sequence was amplified by PCR using the sequence-specific primers DsRed2-F: gaagatctGCCTCCTCCGAGAACGTCATCACC and DsRed2-SKL-R: ggaattcCTACAGCTT GGACTTGTACGATCTC-AGGAACAGGTGGTGGCG.

CONSTRUCTION OF EXPRESSION VECTORS CONTAINING EGFP FUSED WITH DIFFERENT LENGTH OF CSa N-TERMINAL SEQUENCES

The pLead(1–19)-EGFP, pLead(1–24)-EGFP, pLead(1–27)-EGFP, and pLead(1–29)-EGFP vectors were constructed by inserting the N-terminal amino acids 1–19, 1–24, 1–27, and 1–29 from CSa in-frame with the pEGFP-N1 vector. CSa N-terminal amino acids 1–19, 1–27, and 1–29 were amplified by PCR with sequence-specific primers Lead(1–19)-F: ctagctagcccacc ATGGCTTTACTTA-CTGCGGCCGCCCGGCTC and Lead(1–19)-R: cccaagcttAAGACAAG ATGCATTCTTGGTTCCCAAGAGCCGGGCGGCC; Lead(1–27)-F: cta-gctagcccaccATGGC TTTACTTACTGCGGCCGCCCGGCTCTTGGGA-ACCAAGAATGCATCT and Lead(1–27)-R: cccaagcttACTGGCAT-GCCGGGCTGCAAGAACAAGACAAGATGCATTCTTGGTTCC CAA; Lead(1–19)-F: cgggatccaccATGGCTTTACTTACTGCGGCCGCCCGG-CTC and Lead (1–29)-R: cccaagcttGGAAGCACTGGCATGCCGGGC. The pLead(1–24)-EGFP vector was constructed by inserting the *Ase*I + FI-*Sph*I-CMV-Lead(1–24) DNA fragment isolated from pCSa-EGFP into the *Ase*I + FI-*Age*I-pEGFP vector prepared from pCSa-EGFP.

CONSTRUCTION OF THE LEAD(1–29)-EGFP AND CSa-EGFP EXPRESSION VECTORS MUTATED IN THE CONSERVED ARGININE/LYSINE AND SERINE/THREONINE RESIDUES BY SITE-DIRECTED MUTAGENESIS

CSa N-terminal amino acids 1–29 with mutations in the conserved arginine9/24 (R9/R24), lysine14 (K14), serine17/27/29 (S17/S27/S29), or/and threonine5/13 (T5/T13) residue(s) were fused in-frame to EGFP (in pEGFP-N1) to create the expression vectors pLead (R9G)-EGFP, pLead(K14G)-EGFP, pLead(R24G)-EGFP, pLead(R9G/K14G)-EGFP, pLead(K14G/R24G)-EGFP, pLead(R9G/R24G)-EGFP, pLead(R9G/K14G/R24G)-EGFP, pLead(S17G/S27G/S29G)-EGFP, pLead (T5G/T13G)-EGFP and pLead(T5G/T13G/S17G/S27G/S29G)-EGFP. These mutated coding sequences were amplified by PCR with sequence-specific primers and templates. The CSa proteins containing various mutations of the conserved R9/R24 or/and K14 residue(s) were fused to EGFP (in pEGFP-N1) to create the expression vectors pCSa(R9G)-EGFP, pCSa(K14G)-EGFP, pCSa(R24G)-EGFP, pCSa(R9G/K14G)-EGFP, pCSa(K14G/R24G)-EGFP, pCSa(R9G/R24G)-EGFP and pCSa (R9G/K14G/R24G)-EGFP.

DESIGN AND CONSTRUCTION OF shRNA EXPRESSION VECTORS DIRECTED AGAINST THE PREPROTEIN IMPORT RECEPTORS TOM20, TOM22, AND TOM70

The shRNA expression vector pSUPER/DsRed was constructed by inserting the DsRed expression cassette from pDsRed2-N1 into pSUPER [Brummelkamp et al., 2002]. Generally, shRNA expression vectors were constructed by inserting an annealed oligonucleotide duplex in the *Bgl*II–*Hin*dIII restriction sites of the pSUPER/DsRed vector. The oligonucleotides used for molecular construction of pshTOM20, pshTOM22, and pshTOM70 were shTOM20-F: gatcccc-GAGCTTGGCTGAAGATGATttcaagagaATCATCTTCAGCCAAGCTCt-ttttggaaa and shTOM20-R:agcttttccaaaaaGAGCTTGGCTGAAGAT-GATtctcttgaaATCATCTTCAGCCAA GCTCggg; shTOM22-F: gatcccc-GTCTTTGAGACGGAGAAGTttcaagagaACTTCTCCGTCTCAAAGAC-ttttggaaa, and shTOM22-R:agcttttccaaaaaGTCTTTGAGACGGAG-AAGT tctcttgaaACTTCTCCGTCTCAAAGACggg; shTOM70-F: gatc-cccGGACATGAACTCTCTTGATttcaagagaATCAAGAGAGTTCATGT-CCtttttggaaa and shTOM70-R: agcttttccaaaaa GGACATGAACTC-TCTTGATtctcttgaaATCAAGAGAGTTCATGTCCggg (underlined are the targeting sequences for TOM20, TOM22, and TOM70).

WESTERN BLOT ANALYSIS

Total protein extracts or subcellular protein fractions were subjected to electrophoresis on 12% SDS–PAGE and then electroblotted onto an Immobilon-P membrane (Millipore, Billerica, MA). The blotted membranes were incubated with mouse anti-CS (D3G4) (Chemicon International, Temecula, CA), anti-β-actin (Sigma Chemical, Saint Louis, MO), anti-α-tubulin (DM1A) (Lab Vision Corp., Fremont, CA), anti-EGFP (B-2), anti-Tom20 (F-10) (Santa Cruz Biotechnology, Santa Cruz, CA), anti-cytochrome *c* (Cyt-*c*) (7H8.2C12) (BD Biosciences Pharmingen, San Diego, CA) monoclonal antibodies, or goat anti-Tom22 (C-15), anti-Tom70 (C-18) polyclonal antibodies (Santa Cruz Biotechnology), and then incubated with horseradish-peroxidase conjugated goat anti-mouse IgG (H + L) (Pierce Biotechnology, Rockford, IL) or donkey anti-goat IgG (Santa Cruz Biotechnology) secondary antibodies. The bands revealed by the specific antibodies were detected by an enhanced chemiluminescence system (Amersham Biosciences, Little Chalfont, Buckinghamshire, UK) according to the manufacturer's guidelines. Subcellular protein extracts of the mitochondrial and cytosolic fractions were isolated as described by Yadava et al. [2004].

IMMUNOFLUORESCENCE STAINING

Cells grown on glass coverslips were fixed with 3% paraformaldehyde and permeabilized with 0.5% Triton X-100. Immunofluorescence staining was performed according to standard procedures. The fixed cells were first stained with anti-CS, anti-TOM20, anti-TOM22, or anti-TOM70 antibody as described above, and then incubated with Alexa Fluor 488 conjugated goat anti-mouse IgG

(H + L) (Invitrogen) or FITC conjugated donkey anti-goat IgG (Santa Cruz Biotechnology) secondary antibodies, as well as co-stained with DAPI. The stained cells were photographed and analyzed under inverted fluorescence microscopy (Olympus IX71; Olympus Co., Tokyo, Japan).

RESULTS

STRUCTURES OF THE HUMAN CS GENE AND PROTEIN

To elucidate in detail the mechanism by which the human CS enzyme is targeted to the mitochondria, it is important to structurally analyze its amino acid sequence to identify potential subcellular localization and targeting sequences. First, database searches of http://www.ncbi.nlm.nih.gov/, http://www.ensembl.org/index.html and http://mgc.nci.nih.gov/ revealed that the human CS gene is a single nuclear gene, which maps to chromosome 12q13.2-q13.3. The CS gene is composed of 12 exons and transcribes into two mRNA variants, the isoform a (CSa; GenBank gene ID: NM_004077) and b (CSb; GenBank gene ID: NM_198324), by alternative splicing of exon 2 (Fig. 1A). CSa contains only 11 exons (lacks exon 2), encodes 466 amino acids that is translated from the first start codon located in exon 1, while CSb contains all 12 exons but encodes a polypeptide with 400 amino acids that is translated from a downstream start codon located in exon 4, resulting in a shorter N-terminus (Fig. 1B). Analysis of the amino acid sequence of CSa indicated that this protein is hypothetically localized in the mitochondrial matrix (http://expasy.org/). The N-terminal portion (27 amino acids) including a well-conserved RXY↓(S/A) motif (the RHAS sequence) is a putative MTS that contains the typical 3 nonconsecutive basic amino acids (R9, K14 and R24) and 4 serine/threonine residues (T5, T13, S17 and S27), but no acidic amino acids [Gavel and von Heijne, 1990; Braun and Schmitz, 1997; Gakh et al., 2002]. Compared with CSa, CSb has a shorter N-terminus lacking 66 amino acids including the N-terminal MTS. Both CSa and CSb contain the same CS domain, including the three conserved amino acids that are essential for the catalytic activity of CS proteins, as identified by Karpusas et al. [1990] (Fig. 1C).

EXPRESSION OF THE HUMAN CS GENE

To determine the intracellular localization of the human CS protein, we first analyzed its expression in the human cervical carcinoma cell line HeLa and the embryonic kidney epithelial cell line HEK293. Western blot analyses of total protein extracts incubated with an anti-CS-specific antibody revealed that both the cell lines expressed a substantial amount of CS protein (Fig. 2A). In addition, Western blot analyses of cytosolic and mitochondrial fractions incubated with the same antibody showed that CS is found only in the mitochondrial fraction of both cell lines (Fig. 2B). The subcellular markers of the cytosolic and mitochondrial fractions were α-tubulin and cytochrome *c* (Cyt-*c*) proteins, respectively.

The database searches of http://www.ncbi.nlm.nih.gov/, http://www.ensembl.org/index.html, and http://mgc.nci.nih.gov/ showed that the human CS gene transcribes into two mRNA variants, CSa and CSb, which express two polypeptides of 466 and 400 amino acids, respectively (Fig. 1). However, Western blot analyses of total protein extracts revealed that the anti-CS-specific antibody only detected a single protein with a molecular weight of ~50 kDa, suggesting that both HeLa and HEK293 cells express mainly the CSa isoform. To confirm this, total RNA extracts from both the HeLa and HEK293 cells were analyzed by conventional RT-PCR with isoform-specific primers (Fig. 2C). Consistent with the protein expression pattern, both cell lines had a substantial amount of CSa mRNA and an almost undetectable level of the CSb transcript (Fig. 2D). To examine whether HeLa and HEK293 cells express CSb, their total RNA extracts were amplified and analyzed by using real-time RT-PCR. The results revealed that both cell lines express very low or undetectable levels of CSb mRNA (Fig. 2E). Since both the HeLa and HEK293 cells mainly express the CSa isoform, and because only CSa (but not CSb) contains a putative N-terminal MTS, we decided to focus our efforts on the CSa isoform.

INTRACELLULAR LOCALIZATION OF THE HUMAN CSa PROTEIN

To determine whether CSa localizes in the mitochondria, we first examined CSa subcellular localization in the HeLa cells by immunofluorescence staining using an anti-CS-specific antibody. The results showed that CSa was localized in specific intracellular compartment, which completely colocalized with the mitochondrion-specific probe Mito Tracer Red (Fig. 3A). In addition, to directly determine its mitochondrial targeting activity, we constructed a CSa-EGFP fusion protein expression vector (pCSa-EGFP). The pCSa-EGFP-transfected cells exhibited a unique subcellular localization of green fluorescence expression, while the control vector pEGFP-N1-transfected cells displayed an uniform distribution of green fluorescence in the whole cell (Fig. 3B). Western blot analyses of total protein extracts incubated with an anti-EGFP-specific antibody confirmed that the pCSa-EGFP-transfected cells indeed expressed a CSa-EGFP fusion protein (Fig. 3E). To further confirm the mitochondrial localization of the CSa-EGFP fusion protein, we constructed the mitochondrion- and peroxisome-specific reporter vectors pDsRed2-Mito and pDsRed2-Peroxi expressing the reporter protein DsRed. Cells cotransfected with the pCSa-EGFP and pDsRed2-Mito exhibited almost identical subcellular localization patterns of green and red fluorescences, while cells cotransfected with the pCSa-EGFP and pDsRed2 or pDsRed2-Peroxi displayed completely distinguished subcellular localization patterns of green and red fluorescences (Fig. 3C).

Amino acid sequence analysis of the CSa polypeptide indicated that the N-terminal 27 amino acids, including a well-conserved RXY↓(S/A) motif (the RHAS sequence), is a putative MTS (Fig. 1C) [Gavel and von Heijne, 1990; Braun and Schmitz, 1997; Gakh et al., 2002]. To identify the N-terminal MTS in the human CSa protein, we constructed several different length Lead-EGFP expression vectors, including pLead(1–19)-EGFP, pLead(1–24)-EGFP, pLead(1–27)-EGFP, and pLead(1–29)-EGFP, which the EGFPs are N-terminally tagged with CSa N-terminal sequences, 1–19, 1–24, 1–27, and 1–29 amino acids, respectively. Cells transfected with the pLead(1–19)-EGFP exhibited an uniform distribution of green fluorescence in the whole cell, while cells transfected with the pLead(1–24)-EGFP displayed some mitochondrial localization of green fluorescence. In contrast, cells transfected with the pLead(1–27)-EGFP or pLead (1–29)-EGFP exhibited complete mitochondrial localization of

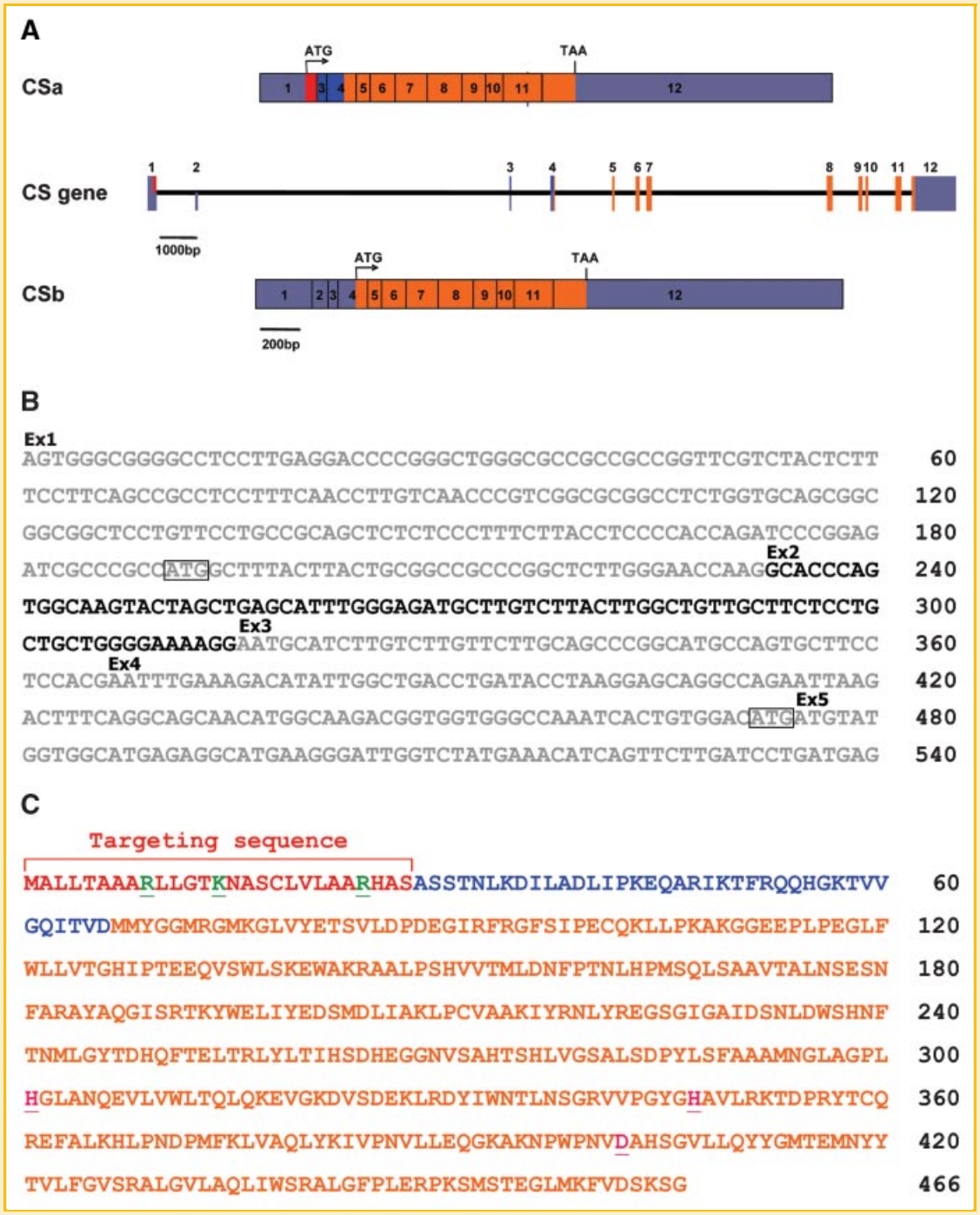

Fig. 1. Analysis of the human CS gene and protein structures. A: Genomic structure and transcript variants of the CS gene. The CS gene is mapped to chromosome 12q13.2-q13.3 locus containing total 12 exons and transcribed into two mRNA variants, isoform a (CSa; GenBank gene ID: NM_004077) and b (CSb; GenBank gene ID: NM_198324), by alternative splicing of exon 2. CSa contains only 11 exons (lacking exon 2) and its start codon locates in exon 1, while CSb contains all 12 exons but its start codon is located in exon 4. B: Partial cDNA sequence of the CSb transcript. CSb contains an additional exon 2 with 83 nucleotides in length (bold letters), which disrupts the translation of exon 1. The start codons located in exons 1 and 4 are boxed. C: Deduced amino acid sequence of the CSa and CSb proteins. CSa consists of 466 amino acids that contain a conserved citrate synthase domain (orange letters) and a putative MTS (red letters) with conserved amino acids, in particular basic residues (green letters and underlined); CSb is composed of 400 amino acids and lacks the N-terminal 66 amino acids, including the predicted MTS. The conserved residues identified by Karpusas et al. [1990] as essential for the catalytic activity of citrate synthases are indicated by pink letters and underlined. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

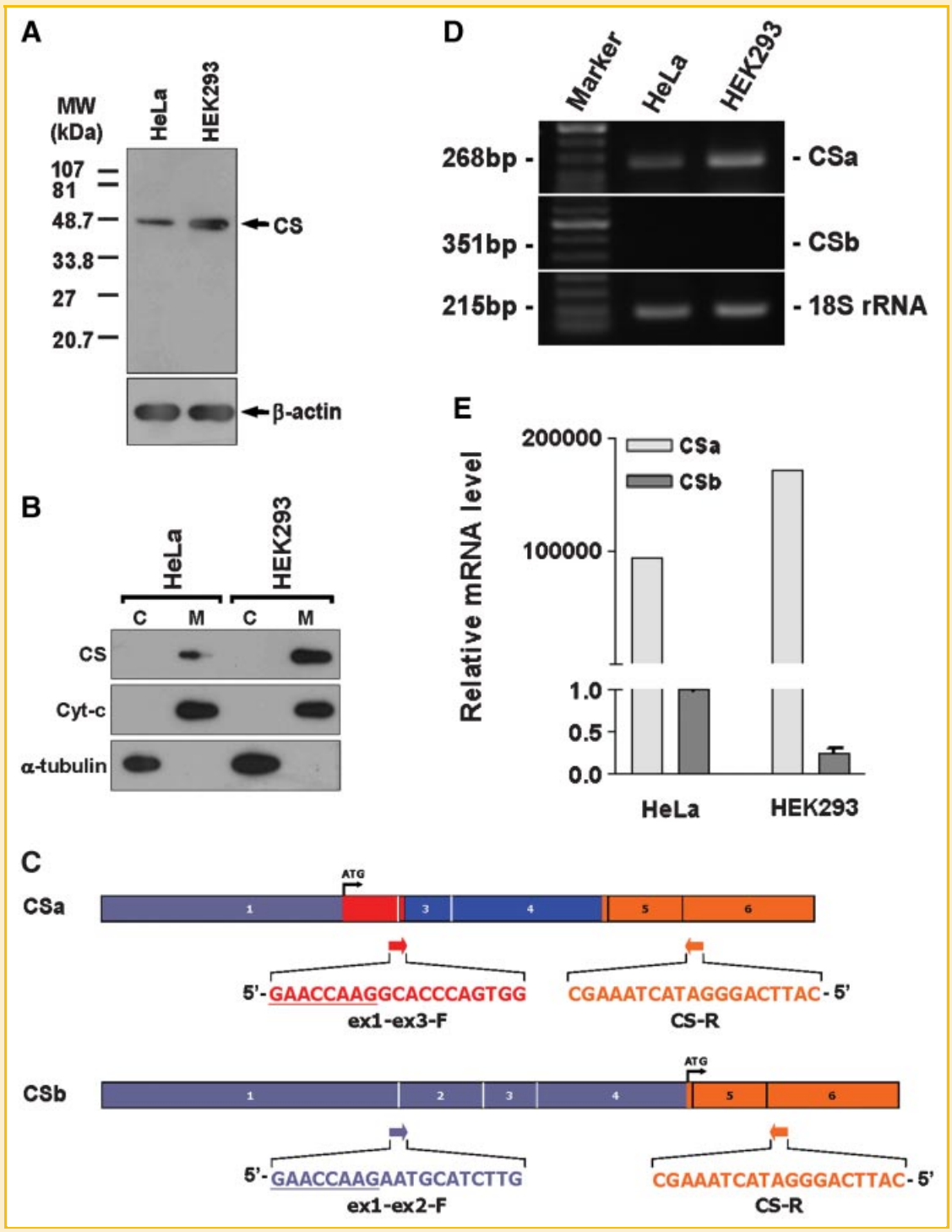

Fig. 2. Analysis of the human CS gene expression. A,B: Western blot analysis of the CS protein. Total protein extracts (A) and protein extracts from mitochondrial and cytosolic fractions (B) of HeLa and HEK293 cells were subjected to electrophoresis on 12% SDS–PAGE and then transferred onto a membrane; the blot was then incubated with an anti-CS antibody. β-Actin served as control for protein loading. Cytochrome *c* (Cyt-c) and α-tubulin served as markers of mitochondrial and cytosolic (M and C) fractions, respectively. C: Design of the CS isoform-specific primers. The upstream CSa- and CSb-specific primers were ex1-ex3-F (which contains 8 nucleotides of exon 1 in 3′ position, and 11 nucleotides of exon 3 in 5′ position) and ex1-ex2-F (which contains 8 nucleotides of exon 1 in 3′ position, and 11 nucleotides of exon 2 in 5′ position), respectively. Both the CSa and CSb isoforms shared the same downstream primer CS-R. D,E: RT-PCR and real-time RT-PCR analyses of the CS transcripts. Total RNA extracts from HeLa and HEK293 were first reversely transcribed into cDNAs by standard reverse transcriptase reaction with poly(dT) primer and then amplified by (D) conventional PCR for 25 cycles or (E) real-time PCR for 40 cycles with CSa- and CSb-specific primers, as well as 18S rRNA-specific primers. The conventional PCR amplified products were analyzed on 1.2% agarose gel electrophoresis and then photographed using an agarose gel imaging and analysis system. The 18S rRNA served as a control for loading. Data analysis of real-time PCR was conducted by adapting the ΔΔCt method of relative quantification to estimate copy numbers of CSa and CSb mRNAs. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

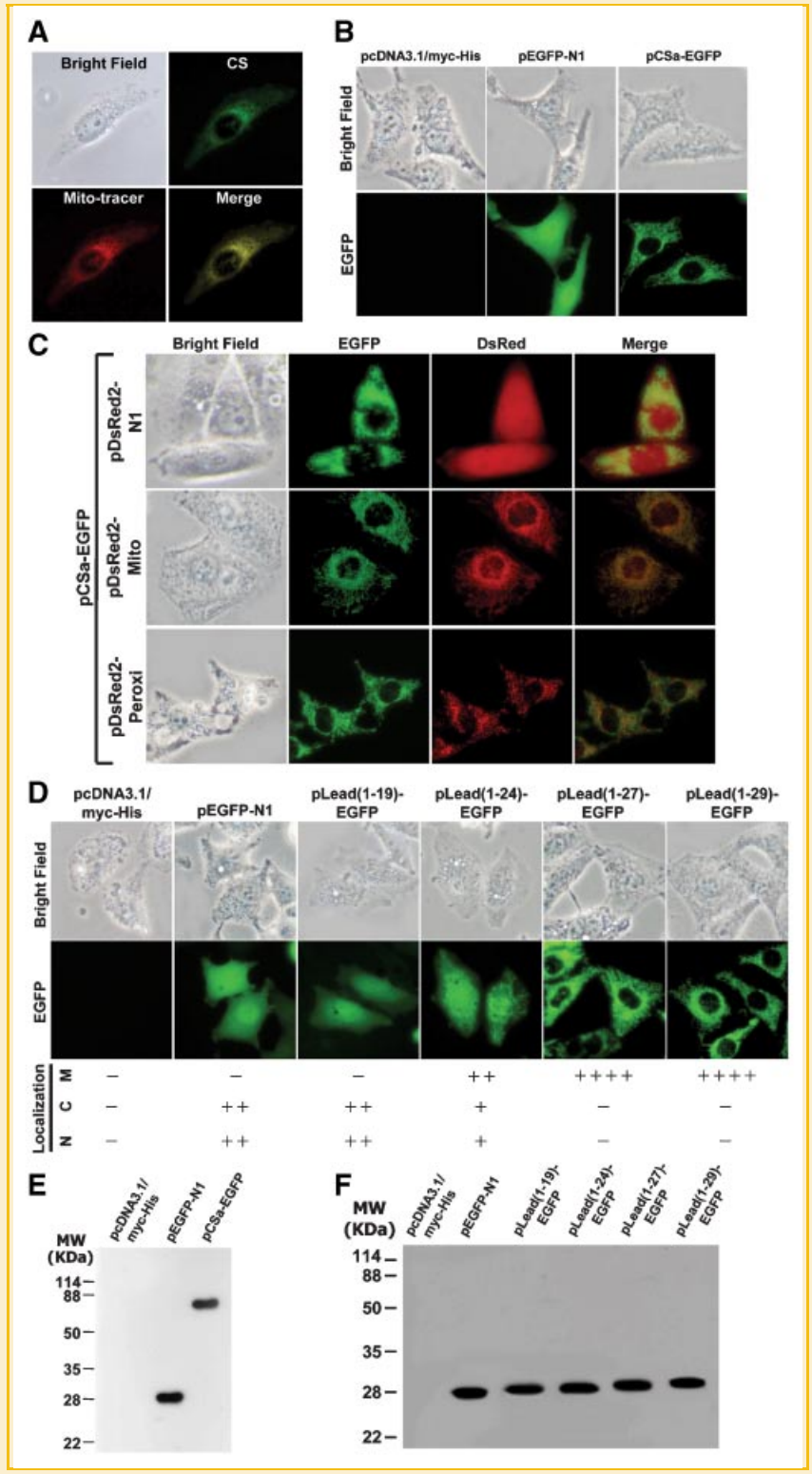

Fig. 3. Analysis of the human CSa protein localization. A: Immunofluoresence staining of the CSa protein. HeLa cells were fixed and stained with an anti-CS-specific antibody, and then co-stained with the mitochondrial-specific probe Mito Tracer Red. B: Intracellular localization of the CSa-EGFP fusion protein. HeLa cells were transfected with pCSa-EGFP expression construct as indicated and then photographed under inverted fluorescence microscopy after 48 h of incubation. C: Co-localization of the CSa-EGFP fusion protein and the mitochondria- or peroxisome-specific reporter DsRed protein. HeLa cells were co-transfected with pCSa-EGFP and pDsRed2-Mito or pDsRed2-Peroxi expression constructs as indicated and then photographed under inverted fluorescence microscopy after 48 h of incubation. D: Identification of the MTS in CSa protein. HeLa cells were transfected with the different size pLead-EGFP expression constructs as indicated, then photographed after 48 h of incubation; the mitochondrial targeting activity was estimated by visualization of cells under inverted fluorescence microscopy. The –, +, ++, +++, and ++++ represent mitochondrial targeting efficiencies of approximately 0%, 20%, 40%, 60% and 80%, respectively. E,F: Western blot analysis of the CSa-EGFP and different Lead-EGFP fusion proteins. The (E) pCSa-EGFP- and (F) different pLead-EGFP-transfected cells as indicated were lyzed after 48 h of incubation and total protein extracts were subjected to electrophoresis on 12% SDS–PAGE, and transferred onto a membrane; the blot was then incubated with an anti-EGFP-specific antibody. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

green fluorescence (Fig. 3D). Western blot analyses of total protein extracts probed with an anti-EGFP-specific antibody confirmed that the transfected cells indeed expressed Lead(1–19)-EGFP, Lead(1–24)-EGFP, Lead(1–27)-EGFP, or Lead(1–29)-EGFP fusion proteins of correct sizes (Fig. 3F).

EFFECTS OF THE MUTATED CONSERVED ARGININE/LYSINE OR SERINE/THREONINE RESIDUES IN MTS-MEDIATED MITOCHONDRIAL TARGETING OF THE CSa PROTEIN

The N-terminal MTS contains three nonconsecutive basic amino acids (R9, K14, and R24) and four serine/threonine residues (T5, T13, S17, and S27), but no acidic amino acids (asparate/glutamate) (Fig. 1C). To analyze in detail the mitochondrial targeting function of these conserved arginine/lysine residues, we constructed several mutated Lead(1–29)-EGFP expression vectors containing single, double, or triple mutation(s) in the conserved arginine/lysine residues. Compared with cells transfected with the wild-type pLead(1–29)-EGFP, cells transfected with the single arginine or lysine mutation construct pLead(R9G)-EGFP, pLead(K14G)-EGFP or pLead(R24G)-EGFP exhibited normal mitochondrial localization of green fluorescence. In addition, cells transfected with the double arginine/lysine mutations construct pLead(K14G/R24G)-EGFP also displayed normal mitochondrial localization of green fluorescence, while cells transfected with the pLead(R9G/K14G)-EGFP or pLead(R9G/R24G)-EGFP exhibited some mislocalization of green fluorescence to the cytosols and nuclei. In contrast, cells transfected with the triple arginine/lysine mutations construct pLead(R9G/K14G/R24G)-EGFP displayed strong mislocalization of green fluorescence to the cytosols and nuclei (Fig. 4A). Western blot analyses of total protein extracts incubated with an anti-EGFP-specific antibody confirmed that the transfected cells expressed the Lead(1–29)-EGFP, Lead(R9G)-EGFP, Lead(K14G)-EGFP, Lead(R24G)-EGFP, Lead(R9G/K14G)-EGFP, Lead(K14G/R24G)-EGFP, Lead(R9G/R24G)-EGFP or Lead(R9G/K14G/R24G)-EGFP fusion proteins with similar sizes (data not shown).

To further confirm these results, we constructed several mutated CSa-EGFP expression vectors that also contained single, double, or triple mutation(s) in the conserved arginine/lysine residues. Consistent with the above arginine/lysine mutation analyses, cells tranfected with the pCSa(R9G)-EGFP, pCSa(K14G)-EGFP, pCSa(R24G)-EGFP, or pCSa(K14G/R24G)-EGFP exhibited normal mitochondrial localization of green fluorescence, while cells transfected with the pCSa(R9G/K14G)-EGFP, pCSa(R9G/R24G)-EGFP or pCSa(R9G/K14G/R24G)-EGFP displayed some to strong mislocalization of green fluorescence to the cytosols and nuclei (Fig. 4B). Western blot analyses of total protein extracts incubated with an anti-EGFP-specific antibody confirmed that the transfected cells indeed expressed the CSa-EGFP, CSa(R9G)-EGFP, CSa(K14G)-EGFP, CSa(R24G)-EGFP, CSa(R9G/K14G)-EGFP, CSa(K14G/R24G)-EGFP, CSa(R9G/R24G)-EGFP, or CSa(R9G/K14G/R24G)-EGFP fusion proteins with similar sizes (data not shown).

To further investigate the mitochondrial targeting function of the conserved basic amino acids (arginine/lysine), especially the R9 residue in the MTS, we constructed several additionally mutated Lead(R9G/K14G/R24G)-EGFP expression vectors with arginine substitutions in the A6, A8, L10, or L11 residues. Compared with cells transfected with the wild-type pLead(1–29)-EGFP, cells transfected with the pLead(R9G/K14G/R24G)-EGFP, pLead(A6R/R9G/K14G/R24G)-EGFP, or pLead(A8R/R9G/K14G/R24G)-EGFP exhibited strong mislocalization of green fluorescence to the cytosols and nuclei (Fig. 4C). In contrast, cells transfected with the pLead(R9G/L10R/K14G/R24G)-EGFP or pLead(R9G/L11R/K14G/R24G)-EGFP displayed weak and strong localization of green fluorescence to the mitochondria (Fig. 4C). Western blot analyses of total protein extracts incubated with an anti-EGFP-specific antibody confirmed that the transfected cells indeed expressed the Lead(R9G/K14G/R24G)-EGFP, Lead(A6R/R9G/K14G/R24G)-EGFP, Lead(A8R/R9G/K14G/R24G)-EGFP, Lead(R9G/L10R/K14G/R24G)-EGFP, or Lead(R9G/L11R/K14G/R24G)-EGFP fusion proteins with similar sizes (data not shown). In addition, Western blot analyses of cytosolic and mitochondrial fractions incubated with the same antibody confirmed that Lead(1–29)-EGFP mainly localized in the mitochondria but Lead(A6R/R9G/K14G/R24G)-EGFP localized in the cytosol (Fig. 4D).

To completely elucidate the mitochondrial targeting function of the conserved serine/threonine residues, we further constructed several mutated Lead(1–29)-EGFP expression vectors that contained mutations in all three serines, or in two threonines, or in all three serines and two threonines. Compared with cells transfected with the wild-type pLead(1–29)-EGFP, cells transfected with the pLead(S17G/S27G/S29G)-EGFP, pLead(T5G/T13G)-EGFP, or pLead(T5G/T13G/S17G/S27G/S29G)-EGFP exhibited normal mitochondrial localization of green fluorescence (Fig. 4E). Western blot analyses of total protein extracts incubated with an anti-EGFP-specific antibody confirmed that the transfected cells indeed expressed the Lead(1–29)-EGFP, Lead(S17G/S27G/S29G)-EGFP, Lead(T5G/T13G)-EGFP, or Lead(T5G/T13G/S17G/S27G/S29G)-EGFP fusion proteins with similar sizes (data not shown).

EFFECTS OF THE SHRNA-INDUCED KNOCKDOWN OF PREPROTEIN IMPORT RECEPTORS TOM20, TOM22, AND TOM70 ON MITOCHONDRIAL TARGETING OF THE CSa PROTEIN

To easily identify the cells transfected with shRNA expression vectors, we constructed an efficient shRNA expression vector pSUPER/DsRed, which could express the reporter protein DsRed under the control of the cytomegalovirus (CMV) promoter and simian virus 40 (SV40) poly(A) tail. This vector could be clearly distinguished from the vectors expressing EGFP by inverted fluorescence microscopy. To efficiently inhibit the expression of human preprotein import receptors, we designed and screened three highly effective RNAi targeting sequences against TOM20, TOM22, and TOM70 by using a robust and comparative siRNA validating system [Hung et al., 2006]. These three RNAi targeting sequences were then cloned into the pSUPER/DsRed vector to generate the pshTOM20, pshTOM22 and pshTOM70 constructs.

To evaluate these three selected shRNA expression vectors, we examined their silencing activities in HeLa cells. HeLa cells were transfected with pshTOM20, pshTOM22, and pshTOM70. At 48 h post-transfection, the expression levels of TOM20, TOM22, or TOM70 in the total protein extracts were examined by Western blot analysis. Compared with cells transfected with the pSUPER/DsRed, cells transfected with the pshTOM20, pshTOM22 or pshTOM70

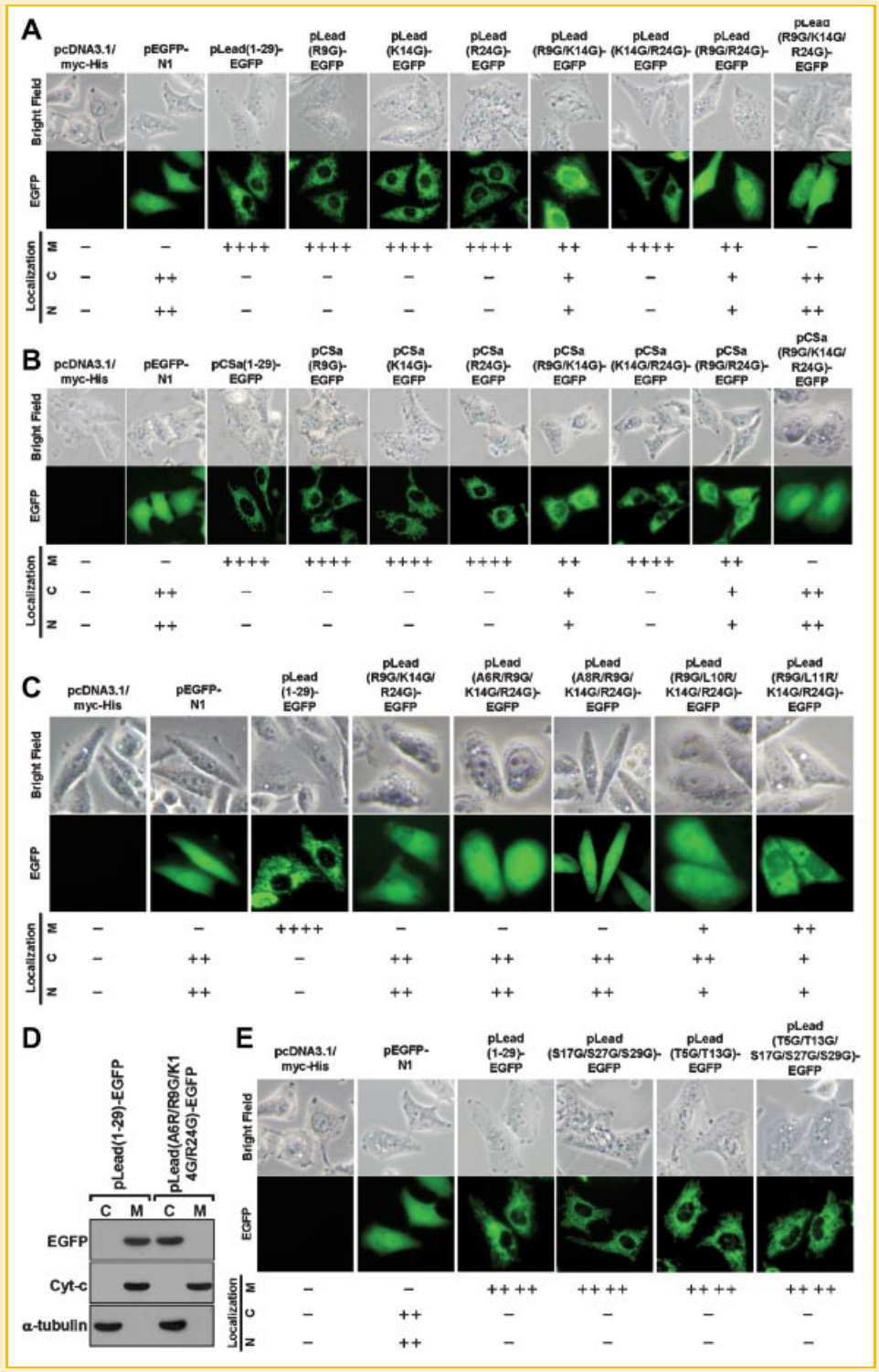

Fig. 4. Analysis of the mutated conserved amino acids in MTS-mediated mitochondrial targeting of the CSa protein. A,B: Effects of the mutated conserved arginine/lysine residues in the mitochondrial targeting of Lead(1–29)-EGFP and CSa-EGFP fusion proteins. HeLa cells were transfected with (A) Lead(1–29)-EGFP and (B) CSa-EGFP expression constructs containing various mutations as indicated, then photographed. The mitochondrial targeting activity was evaluated under inverted fluorescence microscopy after 48 h of incubation. C: Effects of the mutated amino acids surrounding the R9G residue in the mitochondrial targeting of Lead(R9G/K14G/R24G)-EGFP fusion protein. HeLa cells were transfected with Lead(R9G/K14G/R24G)-EGFP expression constructs containing various substitutions as indicated, then photographed; the mitochondrial targeting activity was estimated under inverted fluorescence microscopy after 48 h of incubation. D: Western blot analysis of the Lead(1–29)-EGFP and Lead(A6R/R9G/K14G/R24G)-EGFP fusion proteins. The pLead(1–29)-EGFP- and pLead(A6R/R9G/K14G/R24G)-EGfP-transfected cells were lyzed after 48 h of incubation and protein extracts from mitochondrial and cytosolic fractions were subjected to electrophoresis on 12% SDS–PAGE, and transferred onto a membrane; the blot was then incubated with an anti-EGFP-specific antibody. Cytochrome *c* (Cyt-*c*) and α-tubulin served as markers of mitochondrial and cytosolic (M and C) fractions, respectively. E: Effects of the mutated conserved serine/threonine residues in the mitochondrial targeting of Lead(1–29)-EGFP fusion protein. HeLa cells were transfected with Lead(1–29)-EGFP expression constructs containing various mutations as indicated, then photographed; the mitochondrial targeting activity was evaluated under inverted fluorescence microscopy after 48 h of incubation. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

exhibited a strong inhibition of TOM20, TOM22 or TOM70 expression (Fig. 5A). After 48 h of incubation, the transfected cells were also analyzed directly by immunofluorescence staining. In contrast to cells transfected with the pSUPER/DsRed, cells transfected with the pshTOM20, pshTOM22 or pshTOM70 and marked by the DsRed red fluorescence displayed almost complete inhibition of the TOM20, TOM22, or TOM70 expression (Fig. 5E–G). To analyze the inhibitory effects of pshTOM20, pshTOM22, or psTOM70 at different incubation times, HeLa cells were transfected with pshTOM20, pshTOM22, or pshTOM70. At 6, 12, 24, 36, 48, 60, and 72 h post-transfection, the expression levels of TOM20, TOM22, or TOM70 in the total protein extracts were examined by Western blot analysis. Compared with cells transfected with the pSUPER/DsRed, cells transfected with the pshTOM20, pshTOM22, or pshTOM70 exhibited a similar time-dependent inhibition of the TOM20, TOM22, or TOM70 expression (Fig. 5B–D). The highest inhibition on TOM20, TOM22, or TOM70 expression was observed after 48 h of incubation.

To determine the function of preprotein import receptors in mitochondrial targeting of the human CSa protein, we first analyzed the mitochondrial targeting efficiency of the Lead(1–29)-EGFP fusion protein under RNAi-mediated inhibition of the preprotein import receptors. HeLa cells were co-transfected with pLead(1–29)-EGFP and pshTOM20, pshTOM22 or pshTOM70 expression constructs and then photographed under inverted fluorescence microscopy at 24, 48, and 72 h post-transfection. At 24 and 48 h post-transfection, cells co-transfected with the pLead(1–29)-EGFP and pSUPER/DsRed, pshTOM20, pshTOM22 or pshTOM70 and labeled by the DsRed red fluorescence exhibited almost normal mitochondrial localization of the Lead(1–29)-EGFP fusion protein (data not shown). In contrast, after 72 h of incubation, cells co-transfected with the pLead(1–29)-EGFP and pshTOM20, pshTOM22 or pshTOM70 and marked by the DsRed red fluorescence displayed strong mislocalization of the Lead(1–29)-EGFP fusion protein to the cytosols and nuclei (Fig. 6A). To confirm the results obtained from the above co-transfection experiments, we further performed post-transfection experiments. HeLa cells were first transfected with pshTOM20, pshTOM22, or pshTOM70 expression construct for 48 h. At 48 h post-transfection, the transfected cells were re-transfected with pLead(1–29)-EGFP. After additional 24 h incubation, the double-transfected cells were photographed under an inverted fluorescence microscope. In contrast to cells transfected with the pLead(1–29)-EGFP only and to cells transfected with the pSUPER/DsRed, cells double-transfected with the pLead(1–29)-EGFP and pshTOM20, pshTOM22 or pshTOM70 and labeled with the DsRed red fluorescence exhibited strong mislocalization of the Lead(1–29)-EGFP fusion protein to the cytosols and nuclei (data not shown).

To further confirm the results obtained from above co- and post-transfection experiments, we directly examined the mitochondrial targeting efficiency of CSa under RNAi-induced silencing of the preprotein import receptors at different incubation times. HeLa cells were transfected with pshTOM20, pshTOM22, or pshTOM70. At 48 h post-transfection, the transfected cells were fixed and stained with anti-CS-specific antibody, as well as co-stained with DAPI. After 48 h of incubation, cells transfected with the pshTOM20, pshTOM22, or pshTOM70 and marked by the DsRed red fluorescence displayed strong mislocalization of the CSa protein to the cytosols and nuclei, when compared with cells transfected with the pSUPER/DsRed (Fig. 6B).

In general, during the transport of preproteins into the mitochondria, preproteins containing an N-terminal targeting signal are first recognized by TOM20 and then delivered to TOM22, while preproteins containing an internal targeting signal are primarily recognized by TOM70 and subsequently transferred to TOM22 [Brix et al., 1997, 1999]. In the above experiments, however, cells transfected with the pshTOM20, pshTOM22 or pshTOM70 exhibited strong mislocalization of Lead(1–29)-EGFP and CSa to the cytosols and nuclei (Fig. 6A,B). These results clearly indicated that not only TOM20 and TOM22 but also TOM70 are required for the transport of CSa into the mitochondria. To investigate the possible function of TOM70 in the translocation of preproteins containing an N-terminal targeting signal into the mitochondria, we examined the subcellular localization of TOM20 and TOM22 under RNAi-triggered knockdown of the preprotein import receptor TOM70. HeLa cells were transfected with pshTOM70. At 48 h post-transfection, the transfected cells were fixed and stained with anti-TOM20-, anti-TOM22- or anti-TOM70-specific antibody, as well as co-stained with DAPI. Compared with cells transfected with the pSUPER/DsRed, cells transfected with the pshTOM70 and labeled by the DsRed fluorescence exhibited a strong mislocalization of both the preprotein import receptors TOM20 and TOM22 into the nuclei (Fig. 6C,D).

DISCUSSION

The data presented here have experimentally shown that the single human CS gene could express two mRNA variants, CSa and CSb. Both HeLa and HEK293 cells express mainly the CSa isoform but very low or undetectable levels of the CSb isoform. In particular, we demonstrated that the human CSa isoform is indeed localized in the mitochondria, and that the N-terminal 27 amino acids are responsible for the mitochondrial targeting function. The conserved basic amino acid R9 residue is essential but all serine/threonine residues are dispensable for the mitochondrial targeting of CSa. In addition, all three preprotein import receptors including TOM20, TOM22, and TOM70 are required for efficient transportation of the CSa isoform into the mitochondria.

Most nuclear-encoded mitochondrial proteins contain a cleavable N-terminal MTS that directs mitochondrial targeting of the protein; the N-terminal MTS is cleaved off by matrix processing proteases at a well-conserved RXY↓(S/A) motif [Gavel and von Heijne, 1990; Braun and Schmitz, 1997; Gakh et al., 2002]. These N-terminal MTSs are typically 15–30 amino acids in length including 3–5 nonconsecutive basic amino acid (arginine/lysine) residues, often with several serine/threonine residues but no acidic amino acid (asparate/glutamate) residues. In their molecular structure, these MTSs are able to form strong basic amphipathic α-helices that are essential for efficient mitochondrial transportation [Hurt and van Loon, 1986; von Heijne, 1986; Roise and Schatz, 1988; Horwich, 1990]. The human CSa isoform contains a typical

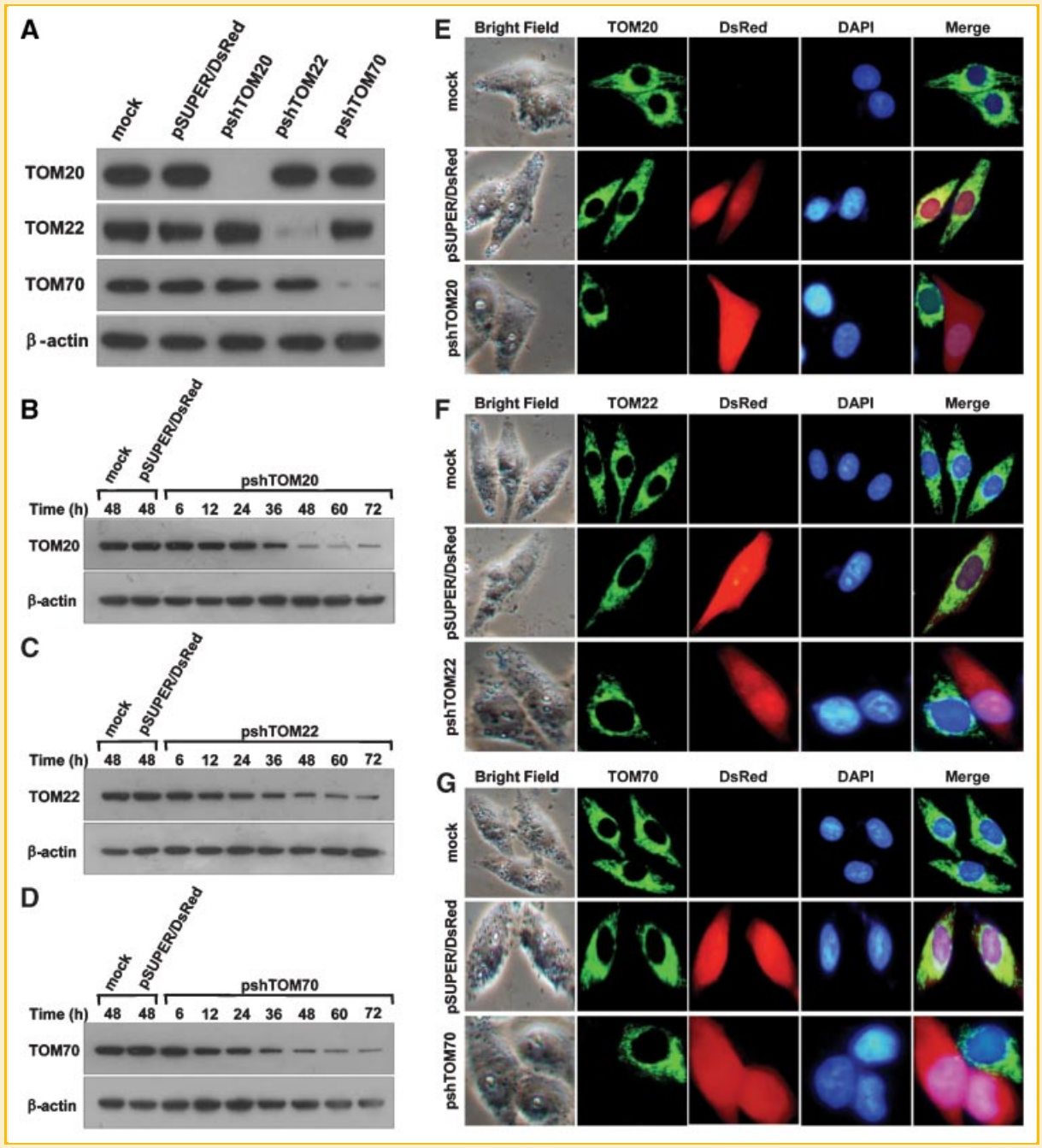

Fig. 5. Effects of the shRNA-induced inhibition of preprotein import receptors. A: Western blot analysis of the shRNA-induced knockdown of the preprotein import receptors. HeLa cells were transfected with the pshTOM20, pshTOM22, or pshTOM70 expression vector as indicated. At 48 h post-transfection, the transfected cells were lyzed and total protein extracts were subjected to electrophoresis on 12% SDS–PAGE, and transferred onto a membrane; the blot was then incubated with anti-TOM20-, anti-TOM22- or anti-TOM70-specific antibody. B–D: Western blot analysis of the shRNA-induced inhibition of the preprotein import receptors at different incubation times. At various times after transfection as indicated, the transfected cells were lyzed and total protein extracts were subjected to electrophoresis on 12% SDS–PAGE, and transferred onto a membrane; the blot was then incubated with (B) anti-TOM20-, (C) anti-TOM22- or (D) anti-TOM70-specific antibody. E–G: Immunofluoresence staining of the shRNA-induced silencing of the preprotein import receptors. At 48 h post-transfection, the transfected cells were fixed and stained with (E) anti-TOM20-, (F) anti-TOM22- or (G) anti-TOM70-specific antibody, and then co-stained with DAPI. The shRNA-transfected cells were marked with red fluorescence by reporter DsRed expression; the nuclei were labeled with blue fluorescence by DAPI staining. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

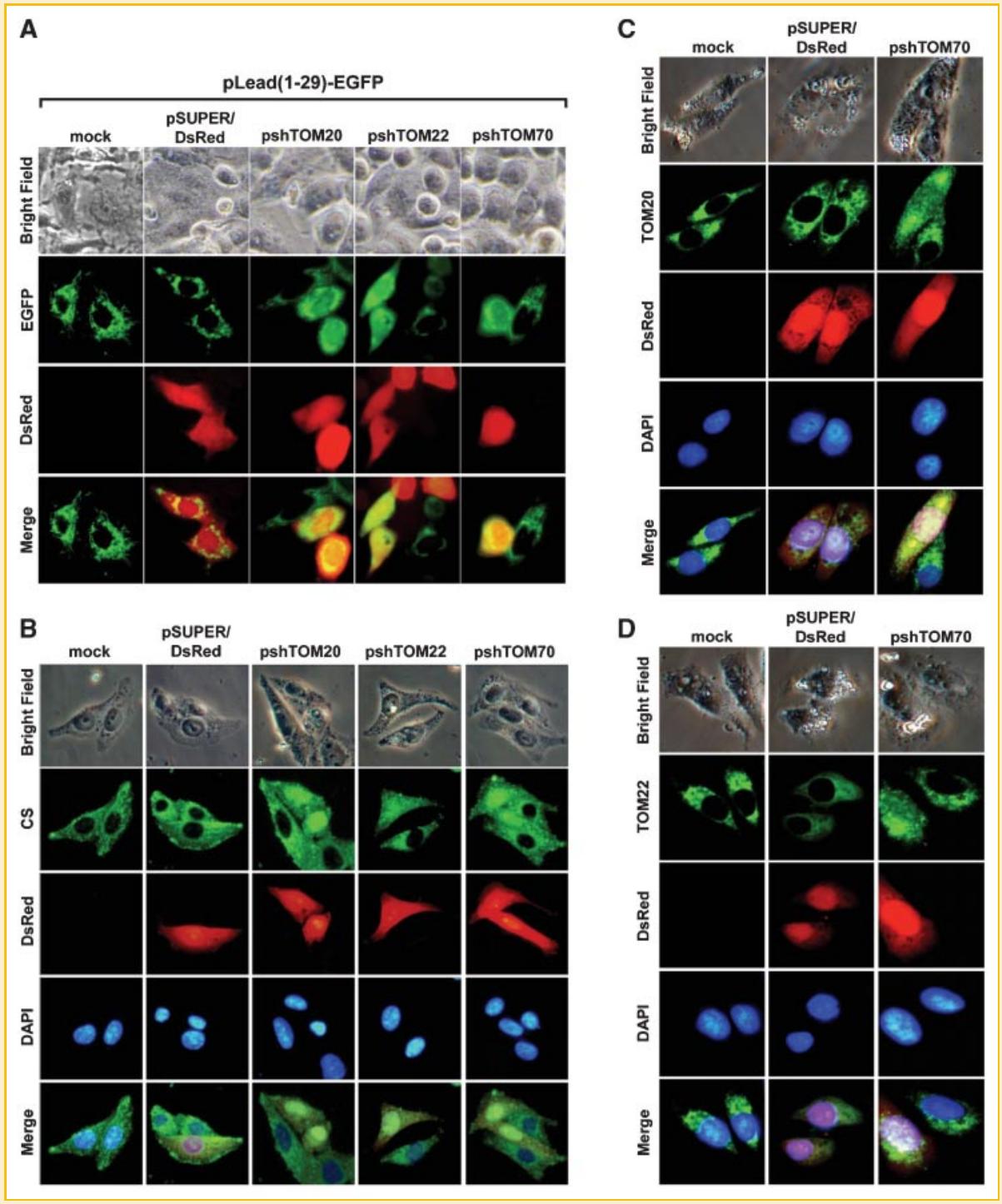

Fig. 6. Analysis of the shRNA-induced knockdown of preprotein import receptors on mitochondrial targeting of the CSa protein. A: Effects of the shRNA-induced inhibition of preprotein import receptors on mitochondrial targeting of the Lead(1–29)-EGFP fusion protein. HeLa cells were co-transfected with pLead(1–29)-EGFP and pshTOM20, pshTOM22 or pshTOM70 expression constructs as indicated. At 72 h post-transfection, the transfected cells were photographed under inverted fluorescence microscopy. The shRNA-transfected cells were labeled with red fluorescence by reporter DsRed expression. B: Effects of the shRNA-induced silencing of preprotein import receptors on mitochondrial targeting of the CSa protein. HeLa cells were transfected with the pshTOM20, pshTOM22, or pshTOM70 expression construct as indicated. At 48 h post-transfection, the transfected cells were fixed and stained with an anti-CS-specific antibody, and then co-stained with DAPI. The shRNA-transfected cells were marked with red fluorescence by reporter DsRed expression; the nuclei were labeled with blue fluorescence by DAPI staining. C,D: Effects of the pshTOM70-induced knockdown of TOM70 on mitochondrial targeting of the TOM20 and TOM22 proteins. HeLa cells were transfected with the pTOM70 expression construct as indicated. At 48 h post-transfection, the transfected cells were fixed and stained with (C) anti-TOM20- or (D) anti-TOM22-specific antibody, and then co-stained with DAPI. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

MTS of 27 amino acids in length consisting of a well-conserved RXY ↓ (S/A) motif, the RHAS sequence, three nonconsecutive basic amino acids (R9, K14, and R24) and four serine/threonine residues (T5, T13, S17, and S27), but no acidic amino acid [Gavel and von Heijne, 1990; Braun and Schmitz, 1997; Gakh et al., 2002]. Furthermore, molecular simulation analysis indicated that the N-terminal 27 amino acids (MALLTAAARLLGTKNASCLVLAARHAS) could form a strong basic amphipathic α-helice (chhhhhhhhh-hcccchhhhhhhhhccc) (http://expasy.org/). It is generally believed that the conserved arginine/lysine and serine/threonine residues are required for efficient transportation of nuclear encoded proteins to the mitochondria. However, the site-directed mutagenesis analysis revealed that among the conserved basic amino acids, only the R9 residue is essential and not all serine/threonine residues are required for efficient targeting of the CSa to the mitochondria. In addition, specific arginine substitutions of the Lead(R9G/K14G/R24G)-EGFP fusion protein revealed that the L11R substitution particularly could enhance transportation of the Lead(R9G/L11R/K14G/R24G)-EGFP fusion protein into the mitochondria. These analyses demonstrated that the proposed characteristics, in particular, the three to five nonconsecutive basic amino acids of the N-terminal MTSs plays the most important function in the translocation of nuclear-encoded proteins to the mitochondria. In contrast, the serine/threonine residues may not be directly involved in mitochondrial targeting or play a role as enhancers in promoting targeting efficiency.

Targeting cytoplasmically synthesized preproteins to the mitochondrial matrix first involves the specific recognition of the N-terminal MTSs by the preprotein import receptor TOM20 [Brix et al., 1997, 1999]. With the help of another preprotein import receptor, TOM22, the preproteins are then targeted to the general import pores that initiate translocation of the preproteins across the outer membrane. After crossing the outer membrane, the preproteins subsequently bind to the TIM23 complex, which prevents the preprotein from releasing into the intermembrane space. The human CSa isoform contains a typical N-terminal MTS that is presumably recognized by TOM20 and then interacts with TOM22. Therefore, both TOM20 and TOM22 may play essential roles during the translocation of CSa to the mitochondrial matrix [Yamano et al., 2008]. Consistent with this mechanism, knockdown of the preprotein import receptors TOM20 or TOM22 resulted in mislocalization of the CSa protein into the cytosol and the nucleus. However, silencing of the preprotein import receptor TOM70, which primarily recognizes the preproteins containing specific internal targeting sequences, also caused mislocalization of the CSa isoform into the cytosol and nucleus, indicating that this preprotein import receptor is also required for mitochondrial targeting of the human CSa protein. Since the TOM complex is a multisubunit machinery consisting of preprotein import receptors and a general import pore, inhibition of the preprotein import receptor TOM70 could induce the dysfunction of the TOM complex or impair the formation of the TOM complex. Immunofluorescence staining of the preprotein import receptors revealed that knockdown of TOM70 resulted in disarrangement of both TOM20 and TOM22 to the nuclei. Therefore, silencing TOM70 could not only directly affect mitochondrial targeting of the preproteins containing the specific internal targeting sequences, but also indirectly affect the targeting of proteins containing the typical N-terminal MTSs.

In summary, our study clearly demonstrate that the human CSa isoform is localized in the mitochondria, and that the N-terminal 27 amino acids including the conserved basic R9 residue as well as all three preprotein import receptors TOM20, TOM22, and TOM70 are essential for the transport of CSa to the mitochondria. This study identifies and characterizes for the first time the mitochondrial targeting sequence and also provides the mechanistic basis for mitochondrial translocation of the human CSa protein. This report also elucidates for the first time the targeting mechanism of a nuclear-encoded mitochondrial protein involved in the human TCA cycle. Together, our results provide an important basis for the future study of mitochondrial dysfunction due to CSa trafficking. However, to fully elucidate the mechanism by which the human CSa isoform is transported to the mitochondria, it is still necessary to systemically analyze all the components involved in targeting mitochondrial proteins.

ACKNOWLEDGMENTS

This work was supported by the grants from National Science Council, Taiwan, ROC (NSC 95-2320-B-006-078-MY3) and Department of Health, Taiwan, ROC (DOH-TD-B-111-004).

REFERENCES

Alvarez-Dolado M, Gonzalez-Moreno M, Valencia A, Zenke M, Bernal J, Munoz A. 1999. Identification of a mammalian homologue of the fungal Tom70 mitochondrial precursor protein import receptor as a thyroid hormone-regulated gene in specific brain regions. J Neurochem 73:2240–2249.

Bains G, Lithgow T. 1999. The Tom channel in the mitochondrial outer membrane: Alive and kicking. Bioessays 21:1–4.

Braun HP, Schmitz UK. 1997. The mitochondrial processing peptidase. Int J Biochem Cell Biol 29:1043–1045.

Brix J, Dietmeier K, Pfanner N. 1997. Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70. J Biol Chem 272:20730–20735.

Brix J, Rudiger S, Bukau B, Schneider-Mergener J, Pfanner N. 1999. Distribution of binding sequences for the mitochondrial import receptors Tom20, Tom22, and Tom70 in a presequence-carrying preprotein and a non-cleavable preprotein. J Biol Chem 274:16522–16530.

Brummelkamp TR, Bernards R, Agami R. 2002. A system for stable expression of short interfering RNAs in mammalian cells. Science 296:550–553.

DiMauro S, Schon EA. 2003. Mitochondrial respiratory-chain diseases. N Engl J Med 348:2656–2668.

Dinchuk JE, Focht RJ, Kelley JA, Henderson NL, Zolotarjova NI, Wynn R, Neff NT, Link J, Huber RM, Burn TC, Rupar MJ, Cunningham MR, Selling BH, Ma J, Stern AA, Hollis GF, Stein RB, Friedman PA. 2002. Absence of post-translational aspartyl β-hydroxylation of epidermal growth factor domains in mice leads to developmental defects and an increased incidence of intestinal neoplasia. J Biol Chem 277:12970–12977.

Dolezal P, Likic V, Tachezy J, Lithgow T. 2006. Evolution of the molecular machines for protein import into mitochondria. Science 313:314–318.

Gakh O, Cavadini P, Isaya G. 2002. Mitochondrial processing peptidases. Biochim Biophys Acta 1592:63–77.

Gavel Y, von Heijne G. 1990. Cleavage-site motifs in mitochondrial targeting peptides. Protein Eng 4:33–37.

Goping IS, Millar DG, Shore GC. 1995. Identification of the human mitochondrial protein import receptor, huMas20p. Complementation of delta mas20 in yeast. FEBS Lett 373:45–50.

Hanson B, Nuttal S, Hoogenraad N. 1996. A receptor for the import of proteins into human mitochondria. Eur J Biochem 235:750–753.

Horwich A. 1990. Protein import into mitochondria and peroxisomes. Curr Opin Cell Biol 2:625–633.

Hung CF, Lu KC, Cheng TL, Wu RH, Huang LY, Teng CF, Chang WT. 2006. A novel siRNA validation system for functional screening and identification of effective RNAi probes in mammalian cells. Biochem Biophys Res Commun 346:707–720.

Hurt EC, van Loon APGM. 1986. How proteins find mitochondria and intramitochondrial compartments. Trends Biochem Sci 11:204–207.

Hurt EC, Pesold-Hurt B, Schatz G. 1984. The cleavable prepiece of an imported mitochondrial protein is sufficient to direct cytosolic dihydrofolate reductase into the mitochondrial matrix. FEBS Lett 178:306–310.

Karpusas M, Branchaud B, Remington SJ. 1990. Proposed mechanism for the condensation reaction of citrate synthase: 1.9-Å structure of the ternary complex with oxaloacetate and carboxymethyl coenzyme A. Biochemistry 29:2213–2219.

Koehler CM. 2004. New developments in mitochondrial assembly. Annu Rev Cell Dev Biol 20:309–335.

Mori M, Terada K. 1998. Mitochondrial protein import in animals. Biochim Biophys Acta 1403:12–27.

Neupert W. 1997. Protein import into mitochondria. Annu Rev Biochem 66:863–917.

Neupert W, Herrmann JM. 2007. Translocation of proteins into mitochondria. Annu Rev Biochem 76:723–749.

Oberst A, Bender C, Green DR. 2008. Living with death: The evolution of the mitochondrial pathway of apoptosis in animals. Cell Death Differ 15:1139–1146.

Orrenius S, Gogvadze V, Zhivotovsky B. 2007. Mitochondrial oxidative stress: Implications for cell death. Annu Rev Pharmacol Toxicol 47:143–183.

Pfanner N, Craig EA, Honlinger A. 1997. Mitochondrial preprotein translocase. Annu Rev Cell Dev Biol 13:25–51.

Prokisch H, Scharfe C, Camp DG II, Xiao W, David L, Andreoli C, Monroe ME, Moore RJ, Gritsenko MA, Kozany C, Hixson KK, Mottaz HM, Zischka H, Ueffing M, Herman ZS, Davis RW, Meitinger T, Oefner PJ, Smith RD, Steinmetz LM. 2004. Integrative analysis of the mitochondrial proteome in yeast. PLoS Biol 2:e160.

Prokisch H, Andreoli C, Ahting U, Heiss K, Ruepp A, Scharfe C, Meitinger T. 2006. MitoP2: The mitochondrial proteome database—Now including mouse data. Nucleic Acids Res 34:D705–D711.

Roise D, Schatz G. 1988. Mitochondrial presequences. J Biol Chem 263:4509–4511.

Schatz G. 1997. Just follow the acid chain. Nature 388:121–122.

Seki N, Moczko M, Nagase T, Zufall N, Ehmann B, Dietmeier K, Schafer E, Nomura N, Pfanner N. 1995. A human homolog of the mitochondrial protein import receptor Mom19 can assemble with the yeast mitochondrial receptor complex. FEBS Lett 375:307–310.

Sickmann A, Reinders J, Wagner Y, Joppich C, Zahedi R, Meyer HE, Schonfisch B, Perschil I, Chacinska A, Guiard B, Rehling P, Pfanner N, Meisinger C. 2003. The proteome of Saccharomyces cerevisiae mitochondria. Proc Natl Acad Sci USA 100:13207–13212.

Suissa M, Suda K, Schatz G. 1984. Isolation of the nuclear yeast genes for citrate synthase and fifteen other mitochondrial proteins by a new screening method. EMBO J 3:1773–1781.

Suzuki H, Maeda M, Mihara K. 2002. Characterization of rat TOM70 as a receptor of the preprotein translocase of the mitochondrial outer membrane. J Cell Sci 115:1895–1905.

Taylor SW, Fahy E, Zhang B, Glenn GM, Warnock DE, Wiley S, Murphy AN, Gaucher SP, Capaldi RA, Gibson BW, Ghosh SS. 2003. Characterization of the human heart mitochondrial proteome. Nat Biotechnol 21:281–286.

Terada K, Mori M. 2000. Human DnaJ homologs dj2 and dj3, and bag-1 are positive cochaperones of hsc70. J Biol Chem 275:24728–24734.

Terada K, Ohtsuka K, Imamoto N, Yoneda Y, Mori M. 1995. Role of heat shock cognate 70 protein in import of ornithine transcarbamylase precursor into mammalian mitochondria. Mol Cell Biol 15:3708–3713.

von Heijne G. 1986. Mitochondrial targeting sequences may form amphiphilic helices. EMBO J 5:1335–1342.

Wallace DC. 1999. Mitochondrial diseases in man and mouse. Science 283:1482–1488.

Yadava N, Houchens T, Potluri P, Scheffler IE. 2004. Development and characterization of a conditional mitochondrial complex I assembly system. J Biol Chem 279:12406–12413.

Yamano K, Yatsukawa Y, Esaki M, Hobbs AE, Jensen RE, Endo T. 2008. Tom20 and Tom22 share the common signal recognition pathway in mitochondrial protein import. J Biol Chem 283:3799–3807.

Young JC, Hoogenraad NJ, Hartl FU. 2003. Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. Cell 112:41–50.