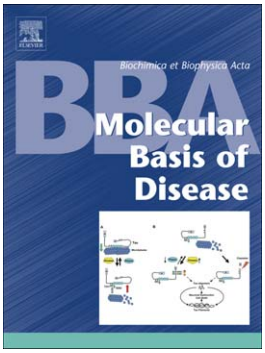
Mutational hotspots in electron transfer flavoprotein underlie defective folding and function in multiple acyl-CoA dehydrogenase deficiency

Bárbara J. Henriques, Peter Bross, Cláudio M. Gomes



Mutational hotspots in electron transfer flavoprotein underlie defective folding and function in multiple acyl-CoA dehydrogenase deficiency

Bárbara J. Henriques, Peter Bross, Cláudio M. Gomes

PII: S0925-4439(10)00152-3
DOI: doi: 10.1016/j.bbadis.2010.07.015
Reference: BBADIS 63140

To appear in: *BBA - Molecular Basis of Disease*

Received date: 25 May 2010
Revised date: 13 July 2010
Accepted date: 16 July 2010

Please cite this article as: Bárbara J. Henriques, Peter Bross, Cláudio M. Gomes, Mutational hotspots in electron transfer flavoprotein underlie defective folding and function in multiple acyl-CoA dehydrogenase deficiency, *BBA - Molecular Basis of Disease* (2010), doi: 10.1016/j.bbadis.2010.07.015

Mutational hotspots in electron transfer flavoprotein underlie defective folding and function in multiple acyl-CoA dehydrogenase deficiency

Bárbara J. Henriques[1], Peter Bross [2], Cláudio M. Gomes[1],*

[1] Instituto Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal

[2] Research Unit for Molecular Medicine, Aarhus University Hospital, Aarhus, Denmark

Corresponding author: C. M. Gomes.
ITQB/UNL. Av. República 127,2780-756 Oeiras (Portugal).
Tel: +351 214469332. Fax: +351 214411277. E-mail: gomes@itqb.unl.pt

ABSTRACT

We have carried out an extensive *in silico* analysis on 18 disease associated missense mutations found in electron transfer flavoprotein (ETF), and found that mutations fall essentially in two groups, one in which mutations affect protein folding and assembly, and another one in which mutations impair catalytic activity and disrupt interactions with partner dehydrogenases. We have further experimentally analyzed three of these mutations, ETFβ-p.Cys42Arg, ETFβ-p.Asp128Asn and ETFβ–p.Arg191Cys, which have been found in homozygous form in patients and which typify different scenarios in respect to the clinical phenotypes. The ETFβ-p.Cys42Arg mutation, related to a severe form of multiple acyl-CoA dehydrogenase deficiency (MADD), affects directly the AMP binding site and intersubunit contacts and impairs correct protein folding. The two other variations, ETFβ-p.Asp128Asn and ETFβ–p.Arg191Cys, are both associated with mild MADD, but these mutations have a different impact on ETF. Although none affects the overall α/β fold topology as shown by far-UV CD, analysis of the purified proteins shows that both have substantially decreased enzymatic activity and conformational stability. Altogether, this study combines *in silico* analysis of mutations with experimental data and has allowed establishing structural hotspots within the ETF fold that are useful to provide a rationale for the prediction of effects of mutations in ETF.

KEYWORDS:

ETF, MADD, Molecular chaperones, Spectroscopy, Protein, Metabolic disorder

Abbreviations used are: MADD, multiple acyl-CoA dehydrogenase deficiency; WT, wild type; ETF, electron transfer flavoprotein; PDB, Protein Data Bank; CD, circular dichroism; SPR, surface Plasmon resonance; T_m, midpoint unfolding temperature, C_m, denaturant midpoint transition concentration

INTRODUCTION

Mutations in the *ETFA/ETFB* or *ETFDH* genes, which encode the two subunits of Electron Transfer Flavoprotein (ETF) and ETF-ubiquinone oxidoreductase (ETF-QO), respectively, cause multiple acyl-CoA dehydrogenase deficiency (MADD), a congenital metabolic disease with a broad clinical expression [1]. The systematic application of newborn screening programs has allowed for a wider coverage of MADD genotypes, that range from missense or nonsense mutations to frameshift. One of the proteins affected in this disorder is ETF, a key enzyme in the mitochondrial fatty acid beta oxidation and amino acid degradation pathways, which transfers electrons from at least 12 dehydrogenases via ETF-QO to the respiratory chain. As in many other metabolic disorders, clear genotype-phenotype relationships remain elusive in MADD: in particular for milder disease forms, it is expected that the mutant proteins are partially functional and that cellular and environmental factors such as temperature and cofactors can modulate disease expression [2]. Progress in this issue has however also been constrained by the fact that only a very limited number of mutations has been characterised upon recombinant expression of the variant proteins. However, substantial steps forward can be achieved if the available data on clinical mutations are merged with a structural and functional analysis of the protein affected. In this paper we have investigated for mutational hotspots on the ETF structure, by analysing the protein for particularly susceptible regions in which MADD-related ETF variations detected in patients map. Complementary experimental validation for the putative mutational hotspots identified was subsequently obtained by carrying out a biochemical and biophysical analysis of purified ETF variants with representative disease-causing mutations. The data gathered by this strategy contribute to a better understanding of the molecular factors underlying functional deficiency and may, in the long term, provide a framework for the design of novel therapies.

MATERIALS AND METHODS

Chemicals All reagents were of the highest purity grade commercially available. Octanoyl-CoA, FAD, AMP and urea were purchased from Sigma. Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from VWR International.

Structural Analysis. The crystallographic structures of human ETF (PDB code: 1efv, [3]) and ETF:MCAD complex (PDB code: 1T9G, [4]) were inspected using PyMOL (DeLano Scientific). Analysis of the molecular interactions, cofactor contacts, topological features and generation of models for the variants was carried out using the WhatIF web server [5] and the PDBsum database [6]. The PolyPhen (Polymorphism Phenotyping) web server was used to predict the effect of amino acid variations on protein structure and function [7]. The input for this server was the UniProt accession number (ETFα: P13804 and ETFβ:P38117), sequence position and the two amino acid variants characterizing the desired mutation.

ETF expression in *E. coli* JM109 cells. The *pK-lac-ETF-αβ3*(C/T) plasmid encoding both ETF subunits (here named pWt), described previously [8], was used as DNA template in the production of the two novel mutants, the ETFβ–p.Cys42Arg and ETFβ–p.Arg191Cys. The mutants were generated using the Quick change mutagenesis kit (Stratagene) as recommended by the manufacturer. Mutagenic primers were 5'-CCATGAACCCCTTC**C**GTGAGATCGCGGTGG-3' for the ETFβ–p.Cys42Arg variant, and 5'-GAGGCTCAACGAGCCC**T**GCTACGCCACGCTG-3' for the ETFβ–p.Arg191Cys mutant (altered bases are underline and in bold type). The second mutagenic primers required

were the reverse complements of those shown. After PCR screening the sections containing the mutation were sequenced to exclude PCR errors and subcloned back into the original plasmid. The ETFβ–D128N mutant, also used in these studies, was already available [9]. *E. coli* JM109 cells transformed with the respective plasmid, pWt, pC42R, pD128N, pR191C, were grown in LB (Luria–Bertani) supplemented with 10 μgml^{-1} kanamycin at 30ºC or 37ºC in a shaking incubator until OD_{532} of 0.5-0.8 was reached. The cells were than induced with 1 mM IPTG for 4 hours. Cells were harvested by centrifugation, re-suspended in 10 mM Hepes, 10% ethylene glycol and 0.5 mM phenylmethylsulphonylfluoride (Roth) in presence of DNase (Applichem) and disrupted in a French press. The soluble extract was obtained by centrifugation at 12,000 g for 30 minutes and used for the ETF activity assays [10] and for SDS-PAGE followed by western blotting analyses according to [11]. An ETF polyclonal antibody raised in rabbits against purified recombinant wild type ETF protein was used to identify ETF proteins and visualized with peroxidase conjugated secondary antibody reaction with 1-step TMB-Blotting (Pierce).

Co-expression of Molecular chaperones with ETFβ–p.Cys42Arg. Different conditions were used to improve ETFβ–p.Cys42Arg mutant expression, always using LB as growth medium. First, *E. coli* cells transform with pC42R plasmid were expressed at 30ºC and 18ºC. Induction was made with 1 mM IPTG. JM109 *E. coli* cells were co-transformed with the pC42R plasmid and a molecular chaperone plasmid, pOFX-BADSL2 and pOFX-BADKJE2, respectively for the GroES/EL complex and DnaK,J and GprE [12]. Cells transformed with the ETFβ–p.Cys42Arg and with pOFX-BAD2 empty plasmid, control vector, or with the plasmid encoding the respective molecular chaperone system, were grown in LB supplemented with 10 μgml^{-1} kanamycin and 10 μgml^{-1} tetracycline at 30ºC and 18ºC until OD_{532} of 0.5-0.8 was reached. Cells were then induced with 0.2% arabinose and 30 minutes

later with 1mM IPTG. Samples were withdrawn at 1, 3, 6 and 20 hours after induction with IPTG. Cells were harvested by centrifugation and treated with Bug Buster solution (Novagen), soluble and insoluble fractions were analysed by SDS-PAGE and western blot analysis. Soluble fractions were also used in ETF activity assays [10].

Protein Purification and biochemical assays. Soluble extract from 3L culture growth of wild type, and ETFβ–p.Asp128Asn and ETFβ–p.Arg191Cys mutants, was applied on a Q-Sepharose fast flow column (GE Healthcare, 20 ml) previously equilibrated in 10 mM Hepes, 10% ethylene glycol and 0.5 mM phenylmethylsulphonylfluoride (buffer A). The column was washed with 5 volumes buffer A, and bound proteins were eluted by a linear gradient of 0 - 1M NaCl in buffer A. Pure ETF eluted at a salt concentration around 250 mM, and purity was confirmed by SDS/PAGE. Protein concentration was determined using the Bradford assay. Flavin content was determined using the molar extinction coefficient ε^{436nm}=13400 $M^{-1}.cm^{-1}$ reported for FAD bound to ETF [13]. ETF enzyme activity was measured following 2,6-dichlorophenolindophenol (DCPIP) reduction at 600 nm in a coupled assay in which recombinant human MCAD and octanoyl-CoA were employed, as described in [10]. One unit of catalytic activity is defined as nmol of DCPIP reduced per minute, in the conditions used in the assay. All specific activities reported are based on total flavin content. Pure ETF fractions with a 2.5 fold molar excess FAD were fast-frozen using liquid nitrogen and stored -80ºC.

Spectroscopic methods. Before each experiment FAD excess added to buffers as a preservative was removed by extensive washing using ultra filtration/dilution, and all experiments were performed with pure proteins containing full occupancy of FAD site. UV/visible spectra were recorded at room temperature in a Shimadzu UVPC-1601

spectrometer with cell stirring. Fluorescence spectroscopy was performed using a Cary Eclipse instrument. For tryptophan emission excitation wavelength was set at 280 nm, and FAD emission was followed setting excitation wavelength at 436 nm; slits were 5 and 10 nm for excitation and emission, respectively. Typically protein concentration was 1 μM. CD spectra were recorded on a Jasco J-815 spectropolarimeter with a cell holder thermostatically controlled with a Peltier. A quartz polarized 1 mm path length quartz cuvette (Hellma) was used, and protein concentrations were typically 0.1 $mg.ml^{-1}$.

Surface Plasmon Resonance. Surface plasmon resonance (SPR) experiments were performed on a BIAcore™ 2000 Instrument using NTA sensor chip (BIAcore, Inc.). The NTA sensor chip is designed to bind histidine tagged molecules after nickel activation of the surface. The surface was prepared as recommended by the manufacturer. The running buffer consisted of 10 mM Hepes, pH 7.4, 10% ethylene glycol, 50 mM NaCl, 0.005% Tween 20. Recombinant human MCAD with a histidine tag was immobilized in flow cell 2. The dehydrogenase was captured by manually injecting 25μl of a 200 nM solution at a flow rate of 5 $\mu l\ min^{-1}$. Flow cell 1 was used to correct for refractive index changes and nonspecific binding. ETF (ETF-WT, ETFβ–p.Asp128Asn or ETFβ–p.Arg191Cys) in 50 nM solutions was injected over flow cells 1 and 2 at a flow rate of 40 $\mu l\ min^{-1}$, during 90 seconds. All SPR experiments were done at 25ºC.

Thermal stability. Thermal unfolding with a linear temperature increase was followed using circular dichroism (ellipticity variation at 222 nm) and fluorescence spectroscopy, tryptophan emission (λ_{ex} =280 nm; λ_{em} =340 nm), FAD emission (λ_{ex} =436nm; λ_{em} =530 nm) and FRET from tryptophan emission to FAD cofactor (λ_{ex} =280 nm; λ_{em} =530 nm). In all experiments, a

heating rate of 1ºC min^{-1} was used, and temperature was increased from 30 to 90ºC. Data were analysed according to a two-state model, and fits to the transition curves were made using OriginPro8.

Chemical Stability. The denaturation curves were measured diluting ETF (ETF-WT, ETFβ–p.Asp128Asn or ETFβ–p.Arg191Cys) in different urea concentrations and the tryptophan emission spectrum was recorded after 15 minutes of incubation at 25°C. Transition curves were determined plotting the average emission wavelength against urea concentrations. Data were analysed according to a two-state model, and fits to the transition curves were made using OriginPro8.

RESULTS AND DISCUSSION

Mutational hotspots in the ETF structure. A number of genetic defects in the *ETFA/ETFB* gene account for MADD. Most of the variations detected in patients are of the missense type, although frameshift and nonsense mutations are also described [14]. As the effect of modifying a single amino acid in ETF can be rather broad with respect to clinical expression, the possibility to predict whether a point mutation in a particular region of the protein would affect the folding, stability or activity, would be a potentially very useful parameter to establish genotype-phenotype relationships. In order to fill this gap we have set to define a structural mapping of mutations in ETF. For this purpose our strategy was twofold. First, we have compiled a list of known disease-associated missense mutations described for human ETF, whose predicted effects were analysed using a combination of *in silico* tools for mutagenesis predictions and structural analysis (Table 1). Secondly, we have analysed the ETF crystal structure in respect to specific structural factors that, if affected by mutations, are expected to perturb the structural integrity or the biological activity of the protein [3, 4]. In this respect, we have analysed: i. positions that correspond to AMP-protein interactions; ii. residues that are involved in H-bond connections at the ETFα/ETFβ contact surface; iii. highly conserved segments throughout ETF ortologues; iv. amino acids involved in FAD binding or in the stabilization of the FAD moiety, and; v. positions involved in interactions with medium-chain acyl-CoA dehydrogenase (MCAD), both the anchor loop and the FAD domain. At this stage positions involved in the interaction with ETF-QO were not possible to analyze once there is no structural information available on the ETF:ETF-QO complex.

Merging the data obtained by these two complementary approaches shows that particular regions of the ETF structure are more susceptible to deleterious modifications than others, and the clustering of the known mutations in those structure segments is rather suggestive of structural hotspots for MADD mutations (Fig 1 and Fig S1). This is clearly noted on ETFα, in

which mutations map essentially in two regions: one corresponding to the region involved essentially in interactions with the ETFβ subunit (Glu91-Val165), and the other a conserved region rich in FAD and MCAD interacting residues and inter-subunit interactions (Ala244-Thr295). Clearly mutations in the first region may result in folding defects due to poor dimer assembly or protein destabilisation, as changes in these positions modify the chemistry of buried sites and may result in cavity creation. The latter may arise if an amino acid with a bulkier side chain is replaced by another with a smaller volume (e.g. ETFα-p.Val157Gly) or if a charged residue replaces an hydrophobic one (e.g. ETFα-p.Gly116Arg) (Table 1).

On the other hand, mutations in the second region have an impact on catalysis. This occurs either by disturbing directly the FAD protein interactions, or the second coordination sphere ligands, which are important modulators of the redox properties of the cofactor. In fact, the chemical and catalytic properties of the flavin are tightly controlled by the polypeptide, not only at the level of the residues involved in direct close-range interactions (i.e. within the first coordination sphere), but also by those which are further away from the organic cofactor but that are nevertheless connected to it through hydrogen-bond networks or other weak interactions (i.e. second coordination sphere). Also, this segment comprises positions of the FAD domain that are known to be involved in complex formation with MCAD [15]. Therefore, mutations in this region are expected to affect mostly the biological activity of the protein, although concurrent destabilization may also occur.

In respect to the ETFβ subunit, the scenario is not so clear as a result of the fact that only 5 missense mutations are known in this subunit. Nevertheless, some highly susceptible regions can be defined: those involving two N-terminal segments of ETFβ which harbour simultaneously FAD and AMP binding residues (ETFβ Arg5-Ile20 and ETFβ Val34-Ala45), a region rich in AMP-protein interactions (ETFβ Gly123-Thr134), and segments involved in

interactions with partner proteins, namely the recognition peptide within the anchor domain that constitutes the primary interaction site of partner proteins with ETF (ETFβ Asp184-Lys200) [4], and also a short segment centred around ETFβ–Glu165, an essential residue that destabilises positions compatible with fast inter-protein electron transfer, thus ensuring high complex dissociation rates [15], a strict requirement for subsequent interaction of reduced ETF with ETF:QO.

Analysis of prototypic mutations. In order to provide complementary experimental validation of the putative mutational hotspots identified in ETF, three mutations in ETFβ were selected for subsequent analysis: ETFβ–p.Cys42Arg, ETFβ–p.Asp128Asn and ETFβ–p.Arg191Cys. With the exception of ETFβ–p.Asp128Asn [9, 16], none of these mutations had been previously expressed and purified for a thorough *in vitro* functional and structural characterisation. All three mutations were detected in patients homozygous for the respective mutations allowing to relate the results to the phenotypes. Also, this selection fills a gap in respect to the characterisation of mutations in the ETFβ subunit, in which a lower number of MADD related mutations have been described compared to ETFα. Although an all-embracing division cannot be established, these mutations typify different functional and folding effects. To better predict the structural effect of these individual mutations we have built molecular models of these variants by homology modelling with respect to the wild type protein (Fig. 2). The ETFβ–p.Cys42Arg mutation, associated to a severe clinical expression, is expected to severely affect the interaction with the AMP cofactor, which plays a key role in the assembly of the ETF dimer [17, 18]. The ETFβ–p.Asp128Asn modification, which occurs at an outer layer of the FAD interacting moieties, has been shown to affect the plasticity of the tertiary structure and decrease directly the specific activity [16], and its mutation likely affects

a stabilizing H-bond to ETFβ–Lys11 (Fig. 2). Finally, the ETFβ–p.Arg191Cys mutation is hypothesised to affect catalysis by impairing interaction with the partner dehydrogenase, as this residue is localized in the anchor domain. Also, analysis of the model of the mutant variation shows that removal of ETFβ–Arg191 may destabilize the fold by disrupting electrostatic interactions with the nearby ETFβ–Glu47, ETFβ–Arg51 and ETFβ–Glu54 (Fig. 2). The patients with the latter two mutations display mild forms of the disease, indicating that the modified conformation is nevertheless able to fold and assemble as a dimer, retaining some catalytic proficiency.

Mutations affect folding efficiency. The effect of each mutation on the efficiency of protein folding was investigated by heterologous expression of the different variants and differential analysis of the soluble versus insoluble protein produced, in respect to that of wild type ETF and comparing cell cultures with standardized induction protocols displaying superimposable cell growth characteristics. Therefore, any difference strictly results from distinct expression levels of the heterologously expressed protein. Since it has been previously demonstrated that expression of ETFβ-p.Asp128Asn was improved upon cell growth at 30ºC [9], over expression all of the analysed ETF variants was carried out both at 30 and 37°C (Fig. 3). The results obtained show that while wild type ETF and the mild mutant variants (ETFβ–p.Asp128Asn and ETFβ–p.Arg191Cys) were well expressed as soluble proteins, the severe phenotype mutation (ETFβ–p.Cys42Arg) was not (Fig. 3). The expression level of ETFβ–p.Arg191Cys was 87% of that of the soluble protein detected in the wild type whereas that of ETFβ–p.Asp128Asn was decreased to 48%. This suggests that while the expression of the variant ETFβ–p.Arg191Cys is almost unaffected, the ETFβ–p.Asp128Asn modification has a destabilizing effect on the folding process, making it rather susceptible to misfold and

aggregate, or be degraded. This observation during cell expression agrees with the data obtained *in vitro* using purified ETFβ–p.Asp128Asn for which a higher proteolytic susceptibility was determined [16].

On the other hand, expression of the ETFβ–p.Cys42Arg variant yielded only 10% of ETF protein in the soluble fraction, indicating that this mutation affecting the AMP binding severely impairs ETF folding and assembly. Further expression assays lowering the *E. coli* growth temperature to 18ºC in an attempt to further improve conditions for *in cell* folding did not change the scenario, and even at this low temperature almost all the ETFβ–p.Cys42Arg protein expressed went to the insoluble fraction (not shown). This result agrees with western blot analysis of skin fibroblasts from a patient carrying this mutation that showed no detectable ETF protein [19, 20].

Molecular chaperones partly rescue ETFβ–p.Cys42Arg folding. An approach for the rescue of folding defects resulting from disease-causing mutations is that of stimulating the so called proteostasis network [21]. This involves recruiting the protein quality control machinery and/or the action of small molecule substrates and cofactors as effectors of the folding process. In order to investigate if molecular chaperones could rescue the defective folding resulting from the ETFβ–p.Cys42Arg mutation, we have carried out co-expression experiments with the chaperonins GroEL and GroES (homologs of the human mitochondrial Hsp60 and Hsp10 chaperone system), and the chaperone DnaK (homolog of mammalian Hsp70s) and its cofactors DnaJ and GrpE. This strategy has been shown to result in the effective rescue of proteins with folding defects in other cases [22-25]. The results obtained (Fig. 4) show that the molecular chaperones could in fact partially rescue the protein to the soluble fraction. However, enzymatic activity assays on the soluble fractions showed no

detectable ETF activity, indicating that although molecular chaperones successfully rescue some soluble conformations these are still be incorrectly folded and/or with defective cofactor insertion: presumably these soluble forms are still substantially destabilised and poor incorporation of FAD and/or deficient interaction with the partner dehydrogenases impair activity.

Structure and conformational properties of the ETFβ–p.Asp128Asn and ETFβ–p.Arg191Cys variant proteins – The ETFβ–p.Arg191Cys variant was purified to homogeneity and spectroscopically characterised. The ETFβ–p.Asp128Asn was also purified and characterized as described in [16] and the spectroscopic data from this variant were redrawn from published data [16] in order to allow a direct comparison with the new variant described. The protein fold and secondary structure were analysed using far-UV circular dichroism (Fig. 5). The spectra obtained are typical of folded proteins with a α/β fold, and with the clear α-helix fingerprint with minima at 208 and 222nm. The mutant variants have superimposable spectra with respect to that of wild type ETF, indicating that the mutations do not disrupt the protein fold nor the secondary structure. Trp-emission spectra, which are indicative of tertiary structure interactions, were also recorded. ETF contains two Trp residues, and the fluorescence emission properties of this amino acid ranges typically from 315 to 350 nm and are related to solvent accessibility: a more solvent exposed Trp will emit at higher wavelengths. The results obtained show that both mutants have 10 nm red-shifted maxima of the emitting band, with maximum emission at 330 nm versus 320 nm of the wild type ETF (Fig. S2). This suggests that the structure of the mutant variants is not as constrained as that of the wild type ETF and that they displays a more dynamic tertiary structure. The spectroscopic fingerprint of the catalytic FAD cofactor was also analysed using visible absorption spectroscopy, which is very sensitive to changes in the H-bound interaction

network around the cofactor. From the results obtained we could conclude that the environment of the functional cofactor does not change significantly in any of the mutants, as the two characteristic absorption bands with typical emission maxima at around 376 nm and 436 nm, are retained. Overall, these results indicate that the ETFβ–p.Asp128Asn and ETFβ–p.Arg191Cys variants retain the fold and FAD interactions identical to those of the wild type protein but the differences in Trp fluorescence properties suggest decreased compactness and structural constraints in comparison to the wild type [16]. The latter aspect may have an impact on the enzymatic activity of the mutants as it likely contributes to a less efficient interaction with the partner acyl-CoA dehydrogenases and to an increased protein instability.

Protein instability and MCAD interactions underpin functional deficiency – The specific activity of the ETFβ–p.Asp128Asn and ETFβ–p.Arg191Cys variants was only 30% of that of the wild type protein (400 U/mg *vs.* 1300 U/mg), which shows that, in addition to the lower folding efficiency, these mutant variants also display a decreased specific activity [16]. Since the mutant proteins retain structural features identical to those of the wild type protein under physiological conditions, we have investigated if differences in stability and interaction with the partner dehydrogenases could play a role. The interaction between the ETF variants and the medium chain acyl-CoA dehydrogenase (MCAD) was investigated in preliminary experiments by surface plasmon resonance, using MCAD immobilized in a NTA-sensor chip. The results obtained showed that the ETFβ–p.Asp128Asn and ETFβ–p.Arg191Cys variants interact with MCAD, as an increase in the resonance signal is observed; however, the magnitude of the interaction at equilibrium was only 20-30% of that observed for wild type ETF, which is suggestive of a weaker interaction (Fig. S3).

The effect of the β–p.Asp128Asn and ETFβ–p.Arg191Cys mutations on the conformational stability was also investigated spectroscopically using far-UV CD and fluorescence. Thermal denaturation of the proteins was monitored by following secondary and tertiary structure from far-UV CD and tryptophan and FAD fluorescence emissions (Fig. 6). All the techniques used yielded similar transition curves for each of the proteins analysed, showing that all methods are monitoring the same unfolding event with no noticeable intermediates formed. The analysis showed the same order of relative stabilities of the proteins involved: ETF-WT > ETFβ–p.Arg191Cys > ETFβ–p.Asp128Asn. The thermal stability studies showed that wild type ETF unfolds at an apparent midpoint unfolding temperature (T_m) of 60 ºC, and that ETFβ–p.Arg191Cys is just slightly destabilised (ΔT_m=-3°C, T_m = 57 °C), whereas ETFβ–p.Asp128Asn exhibits a more significant destabilization (ΔT_m =-8°C, T_m = 52 °C). Chemical denaturation was followed using tryptophan emission, and the curves allowed the determination of apparent midpoint urea concentrations (C_m). The obtained values were 2.2, 2.9 and 3.9 M for ETFβ–p.Asp128Asn, ETFβ–p.Arg191Cys, and ETF-WT, respectively (Fig. 6B). Again, the order of stabilities was the same as in thermal denaturations indicating that the ETFβ–p.Asp128Asn variation has a stronger destabilizing effect than ETFβ–p.Arg191Cys. In fact it was already described that ETFβ–p.Asp128Asn exhibits some plasticity in tertiary structure and decreased specific activity as a function of time during incubation at 39°C, in an *in vitro* mimic of a fever episode [16]. These differences among the studied variants can be rationalised from the molecular models made for the ETFβ–p.Asp128Asn and ETFβ–p.Arg191Cys variants (Fig. 2). The replacement of ETFβ–Asp128 propagates to the AMP coordination sphere by affecting an hydrogen bond between Asp128 and Lys11, therefore substantially destabilising the protein fold, as shown by thermal and chemical unfolding studies (Fig. 6). Likewise, the ETFβ–p.Arg191Cys variation also impacts on the conformational stability of the fold as removal of ETFβ–Arg191 disrupts

salt bridges with Glu47 and Glu57 at the nearby α1 helix. This case illustrates how a mutation in a functional region can also have an impact on overall protein stability.

CONCLUSIONS

In this work we have analysed the effects of missense mutations in ETF which correspond to different clinical presentations in MADD patients, combining *in silico* studies with experimental biophysical and biochemical data. Overall, this study showed that ETF mutations map to structural hotspots with respect to subunit interaction, cofactor binding and interaction with partner proteins, and this provides a framework that allows to predict more accurately the probability for a given mutation to affect protein function. Variations leading to severe phenotypes do not map to a particular structural region and a direct correlation with the clinical display can not be generalised, a general problem that has hampered the prediction of mutation effects for many disease-relevant proteins. However, it appears that amino acid changes that lead to drastic alterations of the chemical properties of residues involved in catalytic regions and subunit interactions have a higher probability to result in severe effects. Among these is ETFβ-p.Cys42Arg which results in a protein that is unable to fold: such mutations will be clearly associated with severe clinical phenotypes. Other mutations will affect folding, cofactor binding, and assembly to different degrees, and most of them will be associated with mild phenotypes. The degree of impairment in respect to the clinical display is rather difficult to predict, as different effects can be observed: cofactor binding, interaction with partner proteins and conformational stability. A key aspect arising from our study is the fact that conformational instability of dynamic substructures may be transferred to other parts of the protein, affecting the protein as a whole [26-28] . That is for example the case of ETFβ-p.Asp128Asn variant, in which the modification of a residue which has long range interactions with the cofactor, increases cofactor lability, decreases protein stability and

interaction with the partner dehydrogenase. Likewise, the ETFβ–p.Arg191Cys variation directly affecting the recognition loop also decreases protein stability, as removal of the positively charged Arg will disrupt electrostatic interactions with the acidic residues within helix α_1 (Fig. 2). This effect then propagates to the whole structure, affecting its conformational stability. The observations made for ETF represent a scenario, which is also relevant for many other proteins, as it underpins the necessity to perform and carefully evaluate experimental expression of clinical protein variants in order to collect more data that will contribute to establish general principles that can be used to more accurately predict the functional outcome of genomic variations in metabolic disease.

Acknowledgments

The work was supported by the Fundação para a Ciência e Tecnologia (FCT/MCTES, Portugal) through research grant PTDC/SAU-GMG/70033/2006 (to C.M.G.), and a PhD fellowship SFRH/BD/29200/2006 (to B.J.H.).

Figures

Figure 1 – Structural features mapped on ETF sequence

Analysis of the ETF structure (PDB: 1efv) allowed the identification of the regions involved in FAD and AMP binding, MCAD interactions, and H-bonding between the ETFα and ETFβ subunits, which are represented as boxed segments in respect to the primary sequence positions. Stars denote MADD mutations [14] and are filled according to the clinical phenotype: severe forms (filled stars) and mild (open stars).

Figure 2 – Structural details around the mutated positions

The cartoons represent magnifications of the ETF structural regions in which the MADD variations are identified, overlaying the mutated and the original residues, as well nearby important . Straight dotted lines represent hydrogen bonds. Structures prepared using PyMOL.

Figure 3 – Expression levels of ETF variants at 30 and 37°C.

Expression of ETF variants at 30 and 37°C was quantified from Western Blot analysis (left) and expressed in a bar graph (right) in respect to that of the wild type protein. See text and materials and methods for details.

Figure 4 – Effect of molecular chaperones on the expression of soluble ETFβ-p.Cys42Arg

The amount of the ETFβ-p.Cys42Arg variant expressed in the soluble (s) and insoluble (i) form was quantified from western blots of cell extracts after inducing for 1 and 6h the co-expression of the molecular chaperones GroEL/ES (left) or DnaK/J/GprE (right). The bar graph shows the relative fraction of soluble (light grey) and insoluble (dark grey) protein. See text and materials and methods for details.

Figure 5 – Spectroscopic analysis (A) far-UV CD (B) Visible absorption

The different variants were analysed in respect to their far-UV CD (A) and visible absorption (B) spectra (the legend is common for both panels): a. ETFβ-p.Asp128Asn, b. ETFβ–WT, c.

ETFβ–p.Arg191Cys. Spectra were slightly offset for clarity of comparison. The data for the ETFβ-p.Asp128Asn variant was redrawn from [16]. See text and materials and methods for details.

Figure 6 – Stability profiles of ETFβ-p.Asp128Asn and ETFβ–p.Arg191Cys variants

The thermal (A) and chemical (B) denaturation profiles of the ETFβ-p.Asp128Asn (inverted triangles) and ETFβ–p.Arg191Cys (open circles) were compared in respect to that of wild type ETF (closed squares). The solid curves represent two-state sigmoid curves from which the apparent midpoint denaturations were determined: melting temperatures (T_m) of 52, 57 and 60°C and midpoint denaturant concentrations (C_m) of 2.2, 2.9 and 3.9M urea for ETFβ–p.Asp128Asn, ETFβ–p.Arg191Cys, and ETF-WT, respectively. See text and materials and methods for details.

References

[1] F.E. Frerman, Goodman, S. I., Scriver, C. R., Beaudet, A. L., Sly, W. S., Valle, D., Childs, B., Kinzler, K. W., and Vogelstein, B. , Defects of Electron Transfer Flavoprotein and Electron Transfer Flavoprotein-Ubiquinone Oxidoreductase: Glutaric Acidemia Type II, The Metabolic & Molecular Basis of Inherited Disease McGrawHill, New York, 2001, pp. 2357-2365.

[2] P. Bross, T.J. Corydon, B.S. Andresen, M.M. Jorgensen, L. Bolund, N. Gregersen, Protein misfolding and degradation in genetic diseases, Hum Mutat 14 (1999) 186-198.

[3] D.L. Roberts, F.E. Frerman, J.J. Kim, Three-dimensional structure of human electron transfer flavoprotein to 2.1-A resolution, Proc Natl Acad Sci U S A 93 (1996) 14355-14360.

[4] H.S. Toogood, A. van Thiel, J. Basran, M.J. Sutcliffe, N.S. Scrutton, D. Leys, Extensive domain motion and electron transfer in the human electron transferring flavoprotein.medium chain Acyl-CoA dehydrogenase complex, J Biol Chem 279 (2004) 32904-32912.

[5] G. Vriend, WHAT IF: a molecular modeling and drug design program, Journal of molecular graphics 8 (1990) 52-56, 29.

[6] R.A. Laskowski, E.G. Hutchinson, A.D. Michie, A.C. Wallace, M.L. Jones, J.M. Thornton, PDBsum: a Web-based database of summaries and analyses of all PDB structures, Trends in biochemical sciences 22 (1997) 488-490.

[7] V. Ramensky, P. Bork, S. Sunyaev, Human non-synonymous SNPs: server and survey, Nucleic Acids Res 30 (2002) 3894-3900.

[8] P. Bross, P. Pedersen, V. Winter, M. Nyholm, B.N. Johansen, R.K. Olsen, M.J. Corydon, B.S. Andresen, H. Eiberg, S. Kolvraa, N. Gregersen, A polymorphic variant in the human electron transfer flavoprotein alpha-chain (alpha-T171) displays decreased thermal stability and is overrepresented in very-long-chain acyl-CoA dehydrogenase-deficient patients with mild childhood presentation, Molecular genetics and metabolism 67 (1999) 138-147.

[9] R.K. Olsen, B.S. Andresen, E. Christensen, P. Bross, F. Skovby, N. Gregersen, Clear relationship between ETF/ETFDH genotype and phenotype in patients with multiple acyl-CoA dehydrogenation deficiency, Hum Mutat 22 (2003) 12-23.

[10] W. Rhead, V. Roettger, T. Marshall, B. Amendt, Multiple acyl-coenzyme A dehydrogenation disorder responsive to riboflavin: substrate oxidation, flavin metabolism, and flavoenzyme activities in fibroblasts, Pediatr Res 33 (1993) 129-135.

[11] P. Bross, B.S. Andresen, V. Winter, F. Krautle, T.G. Jensen, A. Nandy, S. Kolvraa, S. Ghisla, L. Bolund, N. Gregersen, Co-overexpression of bacterial GroESL chaperonins partly overcomes non-productive folding and tetramer assembly of E. coli-expressed human medium-chain acyl-CoA dehydrogenase (MCAD) carrying the prevalent disease-causing K304E mutation, Biochim Biophys Acta 1182 (1993) 264-274.

[12] M.P. Castanie, H. Berges, J. Oreglia, M.F. Prere, O. Fayet, A set of pBR322-compatible plasmids allowing the testing of chaperone-assisted folding of proteins overexpressed in Escherichia coli, Analytical biochemistry 254 (1997) 150-152.

[13] M.C. McKean, J.D. Beckmann, F.E. Frerman, Subunit structure of electron transfer flavoprotein, J Biol Chem 258 (1983) 1866-1870.

[14] P.D. Stenson, E.V. Ball, K. Howells, A.D. Phillips, M. Mort, D.N. Cooper, The Human Gene Mutation Database: providing a comprehensive central mutation database for molecular diagnostics and personalized genomics, Human genomics 4 (2009) 69-72.

[15] H.S. Toogood, A. van Thiel, N.S. Scrutton, D. Leys, Stabilization of non-productive conformations underpins rapid electron transfer to electron-transferring flavoprotein, J Biol Chem 280 (2005) 30361-30366.

[16] B.J. Henriques, J.V. Rodrigues, R.K. Olsen, P. Bross, C.M. Gomes, Role of flavinylation in a mild variant of multiple acyl-CoA dehydrogenation deficiency: a molecular rationale for the effects of riboflavin supplementation, J Biol Chem 284 (2009) 4222-4229.

[17] K. Sato, Y. Nishina, K. Shiga, In vitro assembly of FAD, AMP, and the two subunits of electron-transferring flavoprotein: an important role of AMP related with the conformational change of the apoprotein, J Biochem (Tokyo) 121 (1997) 477-486.

[18] K. Sato, Y. Nishina, K. Shiga, In vitro refolding and unfolding of subunits of electron-transferring flavoprotein: characterization of the folding intermediates and the effects of FAD and AMP on the folding reaction, J Biochem (Tokyo) 120 (1996) 276-285.

[19] A. Curcoy, R.K. Olsen, A. Ribes, V. Trenchs, M.A. Vilaseca, J. Campistol, J.H. Osorio, B.S. Andresen, N. Gregersen, Late-onset form of beta-electron transfer flavoprotein deficiency, Molecular genetics and metabolism 78 (2003) 247-249.

[20] M. Schiff, R. Froissart, R.K. Olsen, C. Acquaviva, C. Vianey-Saban, Electron transfer flavoprotein deficiency: functional and molecular aspects, Molecular genetics and metabolism 88 (2006) 153-158.

[21] E.T. Powers, R.I. Morimoto, A. Dillin, J.W. Kelly, W.E. Balch, Biological and chemical approaches to diseases of proteostasis deficiency, Annual review of biochemistry 78 (2009) 959-991.

[22] P. Bross, C. Jespersen, T.G. Jensen, B.S. Andresen, M.J. Kristensen, V. Winter, A. Nandy, F. Krautle, S. Ghisla, L. Bolundi, et al., Effects of two mutations detected in medium chain acyl-CoA dehydrogenase (MCAD)-deficient patients on folding, oligomer assembly, and stability of MCAD enzyme, J Biol Chem 270 (1995) 10284-10290.

[23] D. Salazar, L. Zhang, G.D. deGala, F.E. Frerman, Expression and characterization of two pathogenic mutations in human electron transfer flavoprotein, J Biol Chem 272 (1997) 26425-26433.

[24] B.S. Andresen, P. Bross, S. Udvari, J. Kirk, G. Gray, S. Kmoch, N. Chamoles, I. Knudsen, V. Winter, B. Wilcken, I. Yokota, K. Hart, S. Packman, J.P. Harpey, J.M. Saudubray, D.E. Hale, L. Bolund, S. K›lvraa, N. Gregersen, The molecular basis of medium-chain acyl-CoA dehydrogenase (MCAD) deficiency in compound heterozygous patients: Is there correlation between genotype and phenotype?, Hum.Mol.Genet. 6 (1997) 695-707.

[25] T. Majtan, L. Liu, J.F. Carpenter, J.P. Kraus, Rescue of cystathionine beta-synthase (CBS) mutants with chemical chaperones: purification and characterization of eight CBS mutant enzymes, J Biol Chem 285 (2010) 15866-15873.

[26] A.R. Correia, S. Adinolfi, A. Pastore, C.M. Gomes, Conformational stability of human frataxin and effect of Friedreich's ataxia-related mutations on protein folding, The Biochemical journal 398 (2006) 605-611.

[27] A.R. Correia, C. Pastore, S. Adinolfi, A. Pastore, C.M. Gomes, Dynamics, stability and iron-binding activity of frataxin clinical mutants, The FEBS journal 275 (2008) 3680-3690.

[28] C. Kayatekin, J.A. Zitzewitz, C.R. Matthews, Zinc binding modulates the entire folding free energy surface of human Cu,Zn superoxide dismutase, Journal of molecular biology 384 (2008) 540-555.

Table 1 – Mutations in ETF related to MADD.

	PSIC score	Structural rationale	Ref
ETF α			
p.Leu95Val	0.085	Near inter subunit H-bonded region.	[27]
p.Gly116Arg	2.667	Changes at buried site: hydrophobicity, overpacking and charge.	[22, 28]
p.Arg122Lys	1.507	Change at buried site hydrophobicity. Closest contact is ETFβ-Gln146. Affects subunits interface.	[19]
p.Phe144Ser	1.781	Changes at buried site: hydrophobicity and cavity creation.	[19]
p.Val157Gly	2.346	Closest contact is ETFβ-Ser223. Cavity creation. Increase in flexibility. Affects subunits interface.	[29]
p.Val165Ala	1.373	Near inter subunit H-bonded region.	[19]
p.Leu212Pro	2.798	Change at buried site hydrophobicity.	[19]
p.Arg249Cys	3.472	Disruption of FAD binding site. Stabilizes FAD semiquinone. Contacts with MCAD.	[19].
p.Gly255Val	2.988	Maps at FAD interacting region.	[30]
p.Thr266Met	2.834	Disruption of FAD binding site. Hydrophobicity change at buried site. Contacts with MCAD.	[7, 22, 28]
p.Gly267Arg	2.982	Maps at FAD interacting region. Charged residue inserted in turn. Contacts with MCAD.	[30]
ETF β			
p.Ala17Pro	1.612	Maps at FAD and AMP interacting regions. Maps nearby ETFα-Ile284.	[31]
p.Cys42Arg	3.343	Disruption of AMP binding site. Changes at buried site: overpacking, hydrophobicity and charge.	[18, 19]
p.Asp128Asn	2.438	Charge change at buried site. Maps at AMP interacting regions. Maps nearby ETFα-Tyr149.	[7, 15]
p.Arg164Gln	2.492	Charge change at buried site. Near inter subunit H-bonded region. Maps nearby ETFα-Asn118.	[32]
p.Arg191Cys	3.515	Removal of charged residue. Contacts with MCAD.	[19]
The PSIC score is the position-specific independent count parameter computed by PolyPhen [7].			

Figure 1

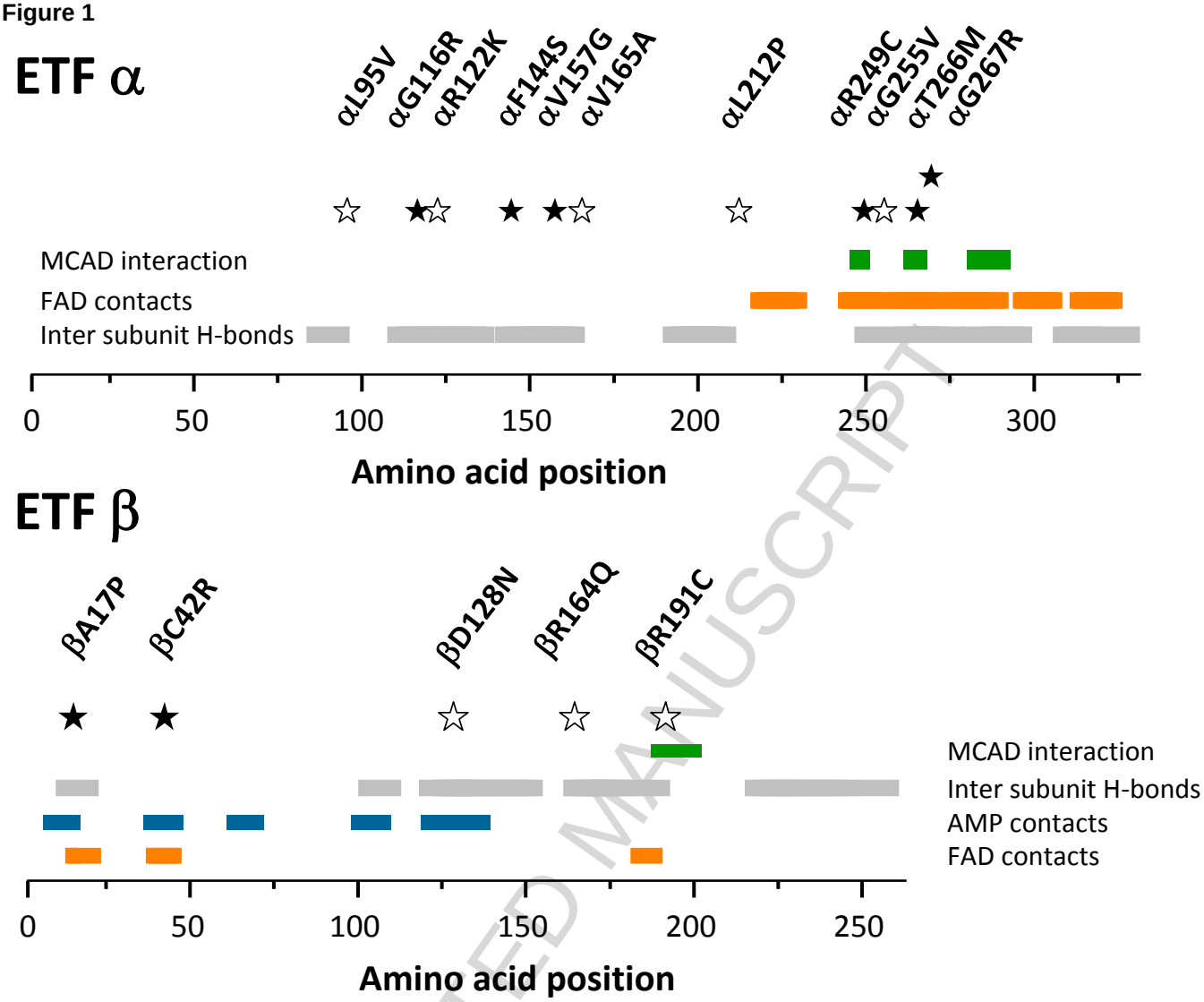


Figure 2

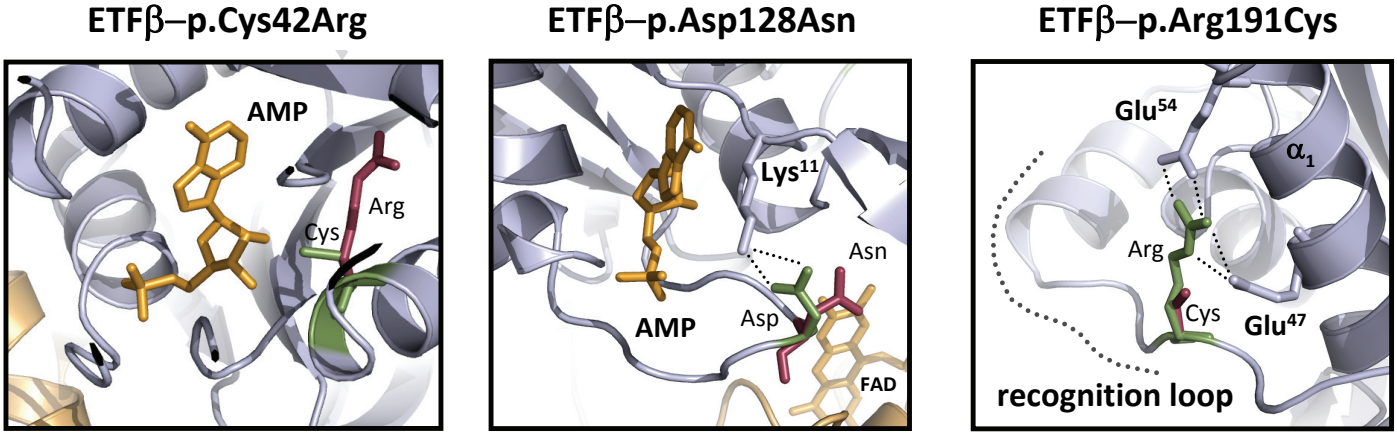


Figure 3

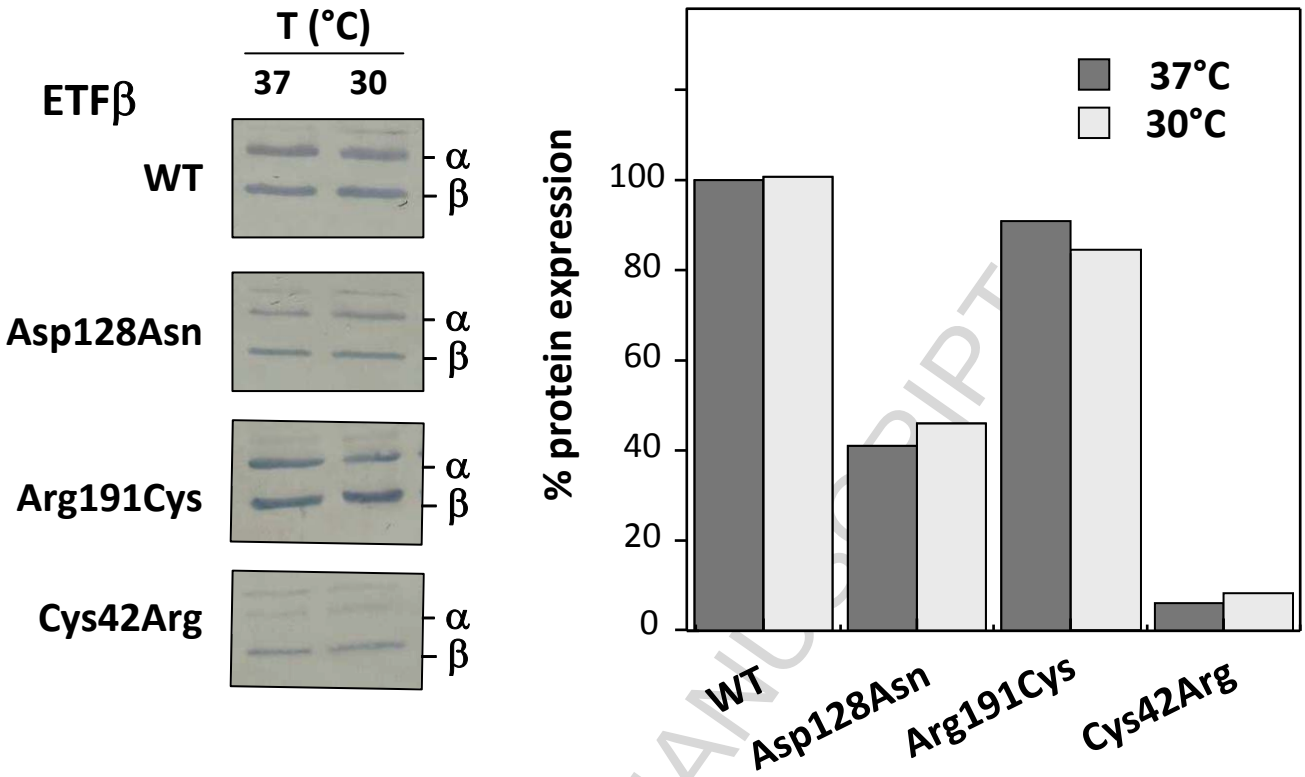


Figure 4

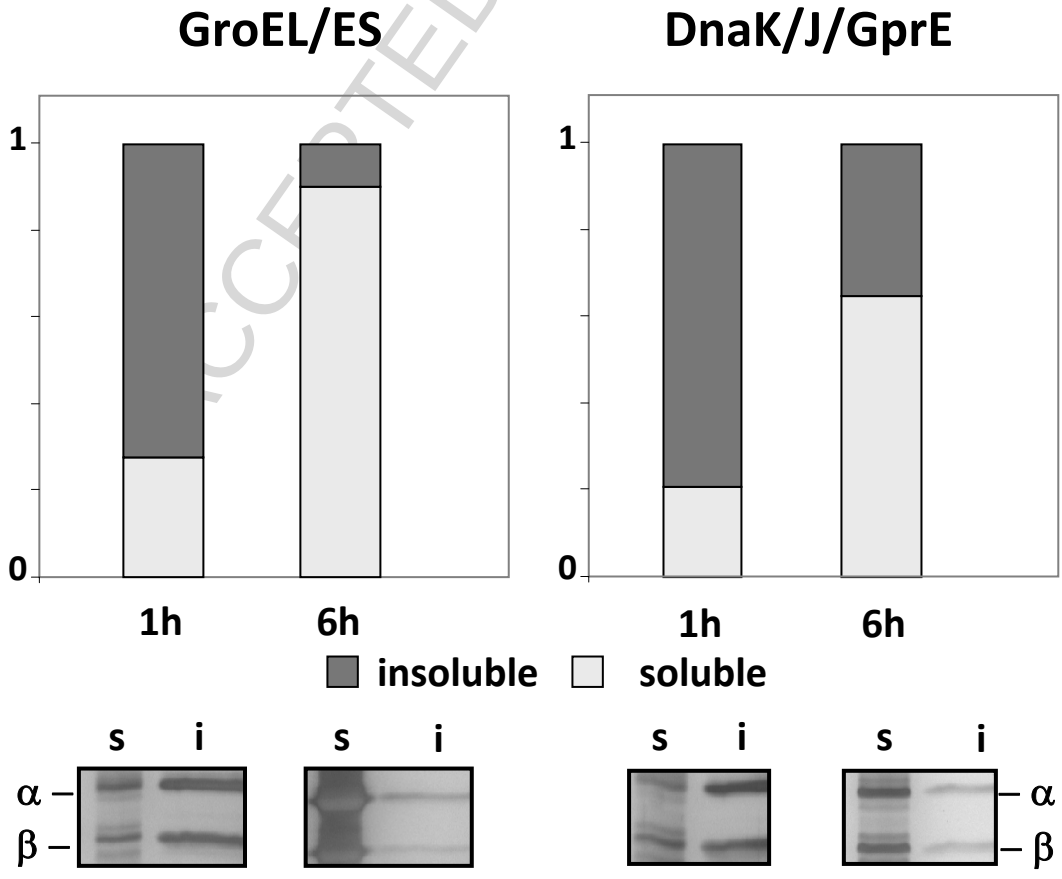


Figure 5

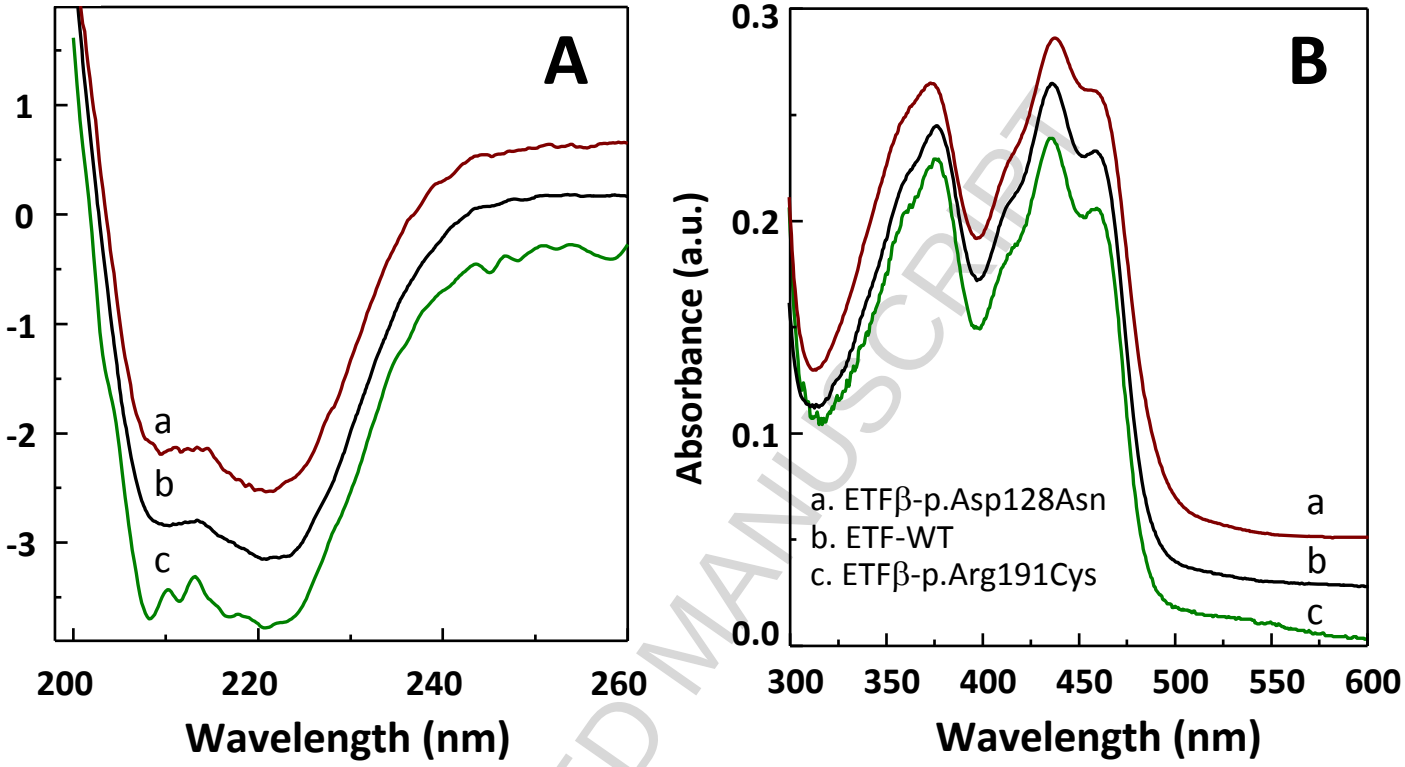


Figure 6