Biochimie xxx (2015) 1–8

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Research paper

Amino acid residues Leu135 and Tyr236 are required for RNA binding activity of CFIm25 in *Entamoeba histolytica*

Juan David Ospina-Villa [a], Absalom Zamorano-Carrillo [a, b], Cesar Lopez-Camarillo [c], Carlos A. Castañon-Sanchez [d], Jacqueline Soto-Sanchez [b], Esther Ramirez-Moreno [a, b], Laurence A. Marchat [a, b, *]

[a] *Biotechnology Program, ENMH-IPN, Mexico City, Mexico*
[b] *Molecular Biomedicine Program, ENMH-IPN, Mexico City, Mexico*
[c] *Genomics Sciences Program, UACM, Mexico City, Mexico*
[d] *Subdirección de Enseñanza e Investigación, Hospital Regional de Alta Especialidad de Oaxaca, Oaxaca, Mexico*

ARTICLE INFO

Article history:
Received 28 January 2015
Accepted 22 April 2015
Available online xxx

Keywords:
Polyadenylation
Cleavage factor Im
Mutant
RNA binding
Entamoeba histolytica

ABSTRACT

Pre-mRNA 3′ end processing in the nucleus is essential for mRNA stability, efficient nuclear transport, and translation in eukaryotic cells. In Human, the cleavage/polyadenylation machinery contains the 25 kDa subunit of the Cleavage Factor Im (CFIm25), which specifically recognizes two UGUA elements and regulates the assembly of polyadenylation factors, poly(A) site selection and polyadenylation. In *Entamoeba histolytica*, the protozoan parasite responsible for human amoebiasis, EhCFIm25 has been reported as a RNA binding protein that interacts with the Poly(A) Polymerase. Here, we follow-up with the study of EhCFIm25 to characterize its interaction with RNA. Using *in silico* strategy, we identified Leu135 and Tyr236 in EhCFIm25 as conserved amino acids among CFIm25 homologues. We therefore generated mutant EhCFIm25 proteins to investigate the role of these residues for RNA interaction. Results showed that RNA binding activity was totally abrogated when Leu135 and Tyr236 were replaced with Ala residue, and Tyr236 was changed for Phe. In contrast, RNA binding activity was less affected when Leu135 was substituted by Thr. Our data revealed for the first time -until we know-the functional relevance of the conserved Leu135 and Tyr236 in EhCFIm25 for RNA binding activity. They also gave some insights about the possible chemical groups that could be interacting with the RNA molecule.

1. Introduction

Pre-mRNA 3′ end processing is a crucial event for mRNA maturation in eukaryotic cells, promoting mRNA stability, efficient nuclear transport, and translation [1–3]. This nuclear reaction involves pre-mRNA 3′ end cleavage followed by the addition of a polyadenine tail [4–7]. This process requires the participation of six primary protein complexes in mammals: the Cleavage and Polyadenylation Specificity Factor (CPSF), Cleavage Stimulatory Factor (CstF), Cleavage Factor I and II (CFIm and CFIIm), Poly(A) Polymerase (PAP) and Poly(A) Binding Protein (PABP), together with the RNA Polymerase (RNAP) II [8,9], and additional proteins [10].

CFIm is a heterotetrameric protein complex formed by a deeply intertwined dimer of CFIm25 subunits that interact with two structurally related large subunits [11,12]. Functional assays showed that CFIm25 is essential for the formation of the pre-mRNA 3′ end processing complex, poly(A) site cleavage and poly(A) tail synthesis [13]. CFIm25 stabilizes the interaction of CPSF with poly(A) signal, increases the cleavage rate *in vitro*, and interacts with PAP and PABP [14,15]. CFIm25 also controls the poly(A) site selection [16,17], and governs the alternative polyadenylation [18]. Additionally, CFIm25 binds the splicing factor U2AF65, establishing a functional connection between the different molecular events of mRNA processing [19]. Although it has a conserved NUDIX

Abbreviations: CPSF, Cleavage and Polyadenylation Specificity Factor; CFIm and CFIIm, Cleavage Factor I and II; CstF, Cleavage Stimulatory Factor; IPTG, isopropyl beta-D-thiogalactopyranoside; NUDIX, nucleoside diphosphate linked to another moiety X; PABP, poly(A) binding protein; PAP, poly(A) polymerase; REMSA, RNA electrophoretic mobility shift assay; RNAP II, RNA polymerase II; SELEX, systematic evolution of ligands by exponential enrichment.

* Corresponding author. Sección de Estudios de Posgrado e Investigación, Escuela Nacional de Medicina y Homeopatía del Instituto Politécnico Nacional, Guillermo Massieu Helguera 239, Fracc. La Escalera, CP 07320 México DF, Mexico. Tel.: +52 55 5729 6300x55543.
E-mail addresses: lmarchat@ipn.mx, lmarchat@gmail.com (L.A. Marchat).

http://dx.doi.org/10.1016/j.biochi.2015.04.017
0300-9084/

(nucleoside diphosphate linked to another moiety X) domain, CFIm25 is not a conventional Nudix protein; it is able to bind a specific RNA sequence, but not to perform RNA cleavage, due to the lack of two of the four essential glutamate residues for catalytic function and metal binding [20,21]. SELEX analysis of CFIm25-RNA and the crystallization assays of CFIm25-CFIm68-RNA complexes revealed that CFIm25 homodimer specifically recognizes two UGUA elements, while CFIm68 enhances the RNA-binding affinity of CFIm and facilitates the looping of RNA molecule between the two antiparallel UGUA sites. Phe104, Glu55, and Arg63 in CFIm25 interact with U1, G2, and U3, respectively, whereas A4 is recognized through an intramolecular sugar-edge/Watson-Crick base pair interaction with G2. Phe103 also contributes to stabilizing UGUA through hydrophobic forces. In agreement, mutation of Glu55, Arg63 and Phe103 abrogated RNA binding [13,22–24].

We have previously reported some factors and genomic sequences involved in pre-mRNA 3′ end processing in *Entamoeba histolytica*, the protozoan parasite responsible for human amoebiasis [25–27]. As described for the human protein, the recombinant EhCFIm25 protein interacts with the poly(A) polymerase EhPAP. Moreover, it binds the *EhPgp5* mRNA 3′ UTR in agreement with its possible role in pre-mRNA 3′ end cleavage [28]. Here, we follow-up with the study of EhCFIm25, using *in silico* and *in vitro* strategies to characterize its interaction with RNA. Our results evidenced that mutation of conserved Leu135 and Tyr236 residues affects RNA binding activity of EhCFIm25, demonstrating for the first time -until we know-the functional relevance of these amino acids for CFIm25 protein function.

2. Material and methods

2.1. In silico analyzes

Full-length amino acid sequences of the 25 kDa subunit of cleavage factor Im (CFIm25) from different species, including EhCFIm25 (C4M2T1) from *E. histolytica*, were aligned using the ClustalW2 algorithm. The BoxShade server highlighted the identical and similar residues in protein sequence alignment. The 3D model of EhCFIm25 in proximity to the UGUA RNA molecule was predicted by the Geno 3D2 Web Server software using as template the chain A of the crystallographic structure of human CFIm25 in complex with RNA (3MDG_A). The APBS software calculated the electrostatic surface potential of CFIm25 proteins and the visualization was obtained by VMD 1.9.1 program. Distances between selected amino acids and the RNA molecule were determined by the MolBrowser 3.7-2d software.

2.2. Site-directed mutagenesis

Mutagenesis assays were performed as described by Liu and Naismith [29] with some modifications. L135A-F/L135A-R and Y236A-F/Y236A-R primer pairs were used to generate single-site mutagenesis within the *EhCFIm25* gene to substitute Leu135 and Tyr236 by an alanine residue. We also used L135T-F/L135T-R and Y236F-F/Y236F-R primers pairs to replace Leu135 and Tyr236 by Thr and Phe residues, respectively. For each mutagenesis reaction, primers design took into account extended non-overlapping sequences at the 3′ end and primer–primer overlapping sequences at the 5′ end that include the mutation site (Table 1). Mutated *EhCFIm25*L135A*, *EhCFIm25*Y236A*, *EhCFIm25*L135T* and *EhCFIm25*Y236F* genes were PCR amplified from pRSET-*EhCFIm25* plasmid [28] using L135A-F/L135A-R, Y236A-F/Y236A-R, L135T-F/L135T-R and Y236F-F/Y236F-R primer pairs (1 μM each), respectively. A high fidelity Taq DNA polymerase (2.0 U; Invitrogen) was used for the next amplification conditions: 94 °C for 7 min; 12 cycles at 94 °C for 1 min, Tm (no) − 5 °C for 1 min, 72 °C for 10 min; three cycles at 95 °C for 1 min, Tm (o) − 5 °C for 1 min, 72 °C for 10 min; followed by a hold step at 4 °C. After, DNA was treated with DpnI enzyme (0.2 U; Invitrogen) at 37 °C for 4 h to eliminate methylated "parental" DNA strands. Then, competent *Escherichia coli* DH5 α bacteria were independently transformed with pRSET-*EhCFIm25*L135A*, pRSET-*EhCFIm25*Y236A*, pRSET-*EhCFIm25*L135T* and pRSET-*EhCFIm25*Y236F* plasmids, and grown at 37 °C in LB medium containing 100 μg/ml ampicillin. Finally, plasmid DNA was extracted using QIAprep Spin Miniprep Kit (Qiagen) and the presence of each single mutation was confirmed by automated DNA sequencing using the 3500 Genetic Analyzer (Applied Biosystems).

2.3. Expression and purification of recombinant wild type (EhCFIm25) and mutated (EhCFIm25*L135A, EhCFIm25*Y236A, EhCFIm25*L135T and EhCFIm25*Y236F) proteins

Competent *E. coli* BL21 (DE3) pLysS bacteria were transformed with pRSET-*EhCFIm25*, pRSET-*EhCFIm25*L135A*, pRSET-*EhCFIm25*Y236A*, pRSET-*EhCFIm25*L135T* or pRSET-*EhCFIm25*Y236F* plasmids and grown at 37 °C in LB medium containing 100 μg/ml ampicillin and 34 μg/ml chloramphenicol. The expression of recombinant proteins was induced with 1 mM isopropyl beta-D-thiogalactopyranoside (IPTG) at 37 °C for 3 h. Then, cells were centrifuged at 6000 rpm for 20 min, resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0), submitted to 10 sonication cycles of 10 s each with an amplitude of 60%, and kept overnight at 4 °C. After centrifugation at 6000 rpm for 1 h, pellet was discarded and the presence of recombinant proteins in total bacterial extract was verified by 10% SDS-PAGE and Western blot assays using anti-6x-His tag antibodies (Roche) at 1:5000 dilution. Bands were revealed by the ECL Plus Western blotting detection system (Amersham). Later, recombinant proteins were purified by affinity chromatography under nondenaturing conditions. Total bacterial proteins were incubated in batch with 1 ml Ni-NTA Agarose (Qiagen) for 2 h at 4 °C; after washing, recombinant proteins were collected with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0) Protein integrity was confirmed by 10% SDS-PAGE and Western blot assays as described above.

2.4. RNA electrophoretic mobility shift assays (REMSA)

RNA binding assays were performed using a RNA fragment of the *EhPgp5* gene 3′ UTR [25] that contains a cleavage site, two polyadenylation signals, and two U-rich sites. First, we PCR-amplified a DNA fragment from the pBS-PSIII156 plasmid using sense (5′-TAATACGACTCACTATAGGGTGTTAATGACTTAAAAGT-3′) and antisense (5′-TGATTATATAATTATATC-3′) primers, with 0.5 U Taq DNA polymerase (Invitrogen) as follows: 94 °C for 3 min; 30 cycles at 94 °C for 1 min, 40 °C for 1 min, 72 °C for 1 min; plus a final extension step at 72 °C for 20 min. Then, a 100 nt RNA fragment was simultaneously *in vitro* transcribed and labeled using the MEGAshortscript™ T7 Kit (Ambion) and Biotin RNA Labeling Mix (ROCHE) following manufacturer recommendations. A control RNA probe without any label was similarly synthesized. RNA product size and integrity were verified by agarose gel electrophoresis and chemiluminescence assay (Chemiluminescent Nucleic Acid Detection Module – Pierce). Finally, RNA-protein interaction assay was performed using the LightShift Chemiluminescent RNA EMSA kit (Thermo Scientific) with some modifications. Briefly, 100 ng RNA probe and 25 μg wild-type or mutant EhCFIm25 proteins were incubated in binding buffer at room temperature for 20 min. For competition assays, we used a 200-fold molar excess of the unlabeled probe or tRNA. In some assays, proteinase K (1 mg/ml) was

Please cite this article in press as: J.D. Ospina-Villa, et al., Amino acid residues Leu135 and Tyr236 are required for RNA binding activity of CFIm25 in *Entamoeba histolytica*, Biochimie (2015), http://dx.doi.org/10.1016/j.biochi.2015.04.017

Table 1
Oligonucleotides for mutagenesis assays.

Primer name	Sequence (5′–3′)	Tm (o)	Tm (no)
L135A-F	GAAATTCAT**GCT**GTTGGAGGAAGATTAAAAATTGGAG	48 °C	42 °C
L135A-R	CCTCCAAC**AGC**ATGAATTTCATCTTTTAAATCCATTG	48 °C	42 °C
L135T-F	GAAATTCAT**ACT**GTTGGAGGAAGATTAAAAATTGGAGAAGATG	56 °C	52 °C
L135T-R	CCTCCAAC**AGT**ATGAATTTCATCTTTTAAATCCATTGATAAAT	44 °C	54 °C
Y236A-F	GAAAAA**GCT**AATATTACACTTTGTTCTATTCC	32 °C	40 °C
Y236A-R	GTAATATT**AGC**TTTTTCAAAATTATTATGTAAAGC	32 °C	42 °C
Y226F–F	TTGAAAAA**TTT**AATATTACACTTTGTTCTATTCCAACTTTAGTT	60 °C	52 °C
Y236F-R	AAAGTGTAATATT**AAA**TTTTTCAAAATTATTATGTAAAGCAAAAAGT	53 °C	54 °C

Underlined, complementary sequences in each pair of primers; bold, mutated codon in *EhCFIm25* gene. o, overlapping regions; no, non-overlapping regions.

added. Protein extract from non-induced bacteria was used as mock control. RNA-protein complexes were resolved at 100 V for 2 h on pre-electrophoresed 6% non-denaturing PAGE and electro-transferred to a Hybond membrane (Amersham Healthcare). RNA-protein complexes were UV cross-linking (254 nm) for 15 min, revealed by the Chemiluminescent Nucleic Acid Detection Module (Pierce) and quantified using the ImageJ processing program [30]. Data corresponding to wild-type EhCFIm25 were taken as 100% and used to determine the RNA affinity of EhCFIm25 mutants.

3. Results

3.1. In silico analyzes suggest that the conserved Leu135 and Tyr236 residues could be relevant for EhCFIm25 RNA binding activity

The only 25 kDa subunit of cleavage factor Im that has been characterized is the human CFIm25 protein (O43809). In the crystallographic structures of CFIm25 bound to the UGUA RNA sequence, 14 amino acids interact directly or indirectly with the

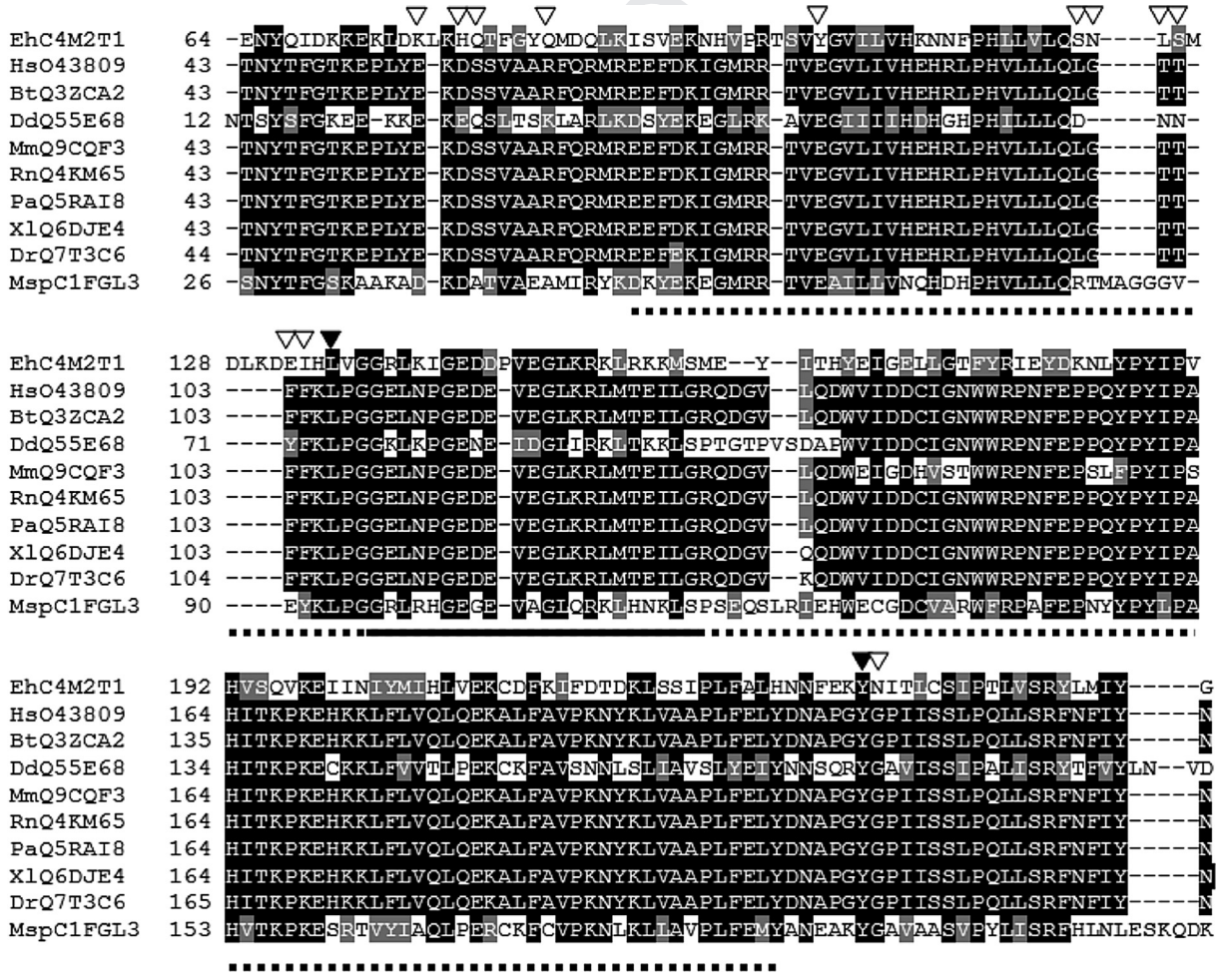

Fig. 1. Representative multiple sequence alignment of CFIm25 proteins from 10 different species. Arrowheads show the 14 amino acids of the human protein that have been demonstrated to interact with the UGUA RNA sequence in the crystallographic structure or REMSA [22–24]. Black box, conserved residue; grey box; homologous residue. Numbers at the left are relative to the first Methionine in each protein. Black arrowheads correspond to the conserved Leu and Tyr residues among all homologous proteins (Leu135 and Tyr236 in EhCFIm25; Leu106 and Tyr208 in human CFIm25). Dotted lane, Nudix domain; plain lane: nudix box. Eh, *Entamoeba histolytica*; Hs, *Homo sapiens*; Bt, *Bos taurus*; Dd, *Dictyostelium discoideum*; Mm, *Mus musculus*; Rn, *Rattus norvegicus*; Pa, *Pongo abelii*; Xl, *Xenopus laevis*; Dr, *Danio rerio*; Msp., *Micromonas sp.* Access number in UniProt Knowledgebase (UniProtKB) database are shown.

Please cite this article in press as: J.D. Ospina-Villa, et al., Amino acid residues Leu135 and Tyr236 are required for RNA binding activity of CFIm25 in *Entamoeba histolytica*, Biochimie (2015), http://dx.doi.org/10.1016/j.biochi.2015.04.017

different ribonucleotides, namely, Glu55, Asp57, Ser58, Arg63, Glu81, Leu99, Gly100, Thr101, Thr102, Phe103, Phe104, Leu106, Tyr208, and Gly209. Among them, Glu55, Arg63, and Phe103 are required for RNA-binding activity in REMSA [22–24]. Multiple sequence alignment of 50 homologous CFIm25 proteins from various organisms showed that only two of these 14 residues, corresponding to Leu135 and Tyr236 in EhCFIm25 (Leu106 and Tyr208 in human CFIm25), are conserved among *Homo sapiens*, *E. histolytica* and other organisms. This suggests that they may be key residues for RNA binding. Leu135 is located within the Nudix domain (residues 91–230 in EhCFIm25 corresponding to residues 77–202 in human CFIm25) at two residues upstream the Nudix box; while Tyr236 remains outside this characteristic domain in both proteins (Fig. 1). The three-dimensional model of *E. histolytica* protein in proximity to the UGUA RNA molecule confirms that EhCFIm25 and the human CFIm25 used as a template share a similar folding with an RMSD value of 1.68 Å between them. Their tertiary structures include the characteristic canonical Nudix fold. Notably, amino acids residues corresponding to Leu135 in EhCFIm25 are localized in turns in both molecules. Tyr208 is at the beginning of helix-4 in the human protein while the corresponding Tyr236 is in preceding turn in EhCFIm25 (Fig. 2A). Interestingly, both Leu106 and Leu135 are located in an internal negatively charged cavity within human and *E. histolytica* protein, respectively. In turn, the aromatic and polar Tyr208 and Tyr236 residues belong to a positively charged region on the surface of both proteins (Fig. 2B). Moreover, both Leu135 and Tyr236 residues are close to U2 of the UGUA molecule (lower than 10 Å, distances from selected atoms of principal and lateral chains), indicating that these residues may be able to make contacts with the RNA fragment (Fig. 2C).

3.2. Conserved Leu135 and Tyr236 residues are essential for EhCFIm25 RNA-binding activity

To gain insights into the functional relevance of Leu135 and Tyr236 for EhCFIm25 RNA binding activity, we produced a total of four mutants. Each amino acid was first replaced by an Ala residue through site-directed mutagenesis assays to shorten the lateral chain. Specifically, the hydrophobic Leu135 was replaced by a residue that has the same polarity but a shorter side chain. The change of the hydrophilic Tyr236 for Ala causes the loss of the aromatic ring of the side chain. Then, Leu135 and Tyr236 were replaced by Thr and Phe, respectively, which changed the polarity of each residue without significantly modifying their lateral chain (Fig. 3A). Site-directed mutagenesis assays were performed as described and *E. coli* bacteria were transformed with mutant constructions. After purification of plasmid DNA, sequencing of pRSET-*EhCFIm25*L135A* and pRSET-*EhCFIm25*Y236A* plasmids confirmed that CTT and TAC codons, respectively, were replaced by GCT codon, without any other changes throughout the gene sequence. The CTT and TAC codons were successfully replaced by ACT and TTT codons in pRSET-*EhCFIm25*L135T* and pRSET-*EhCFIm25*Y236F* plasmids, respectively (Fig. 3B). The recombinant EhCFIm25 (wild-type), EhCFIm25*L135A, EhCFIm25*Y236A, EhCFIm25*L135T and EhCFIm25*Y236F (mutant) proteins with a 6x-His tag at the N-terminus were expressed in *E. coli* BL21 (DE3) plysS strain in the presence of 1 mM IPTG for 3 h and purified by Ni-NTA affinity chromatography

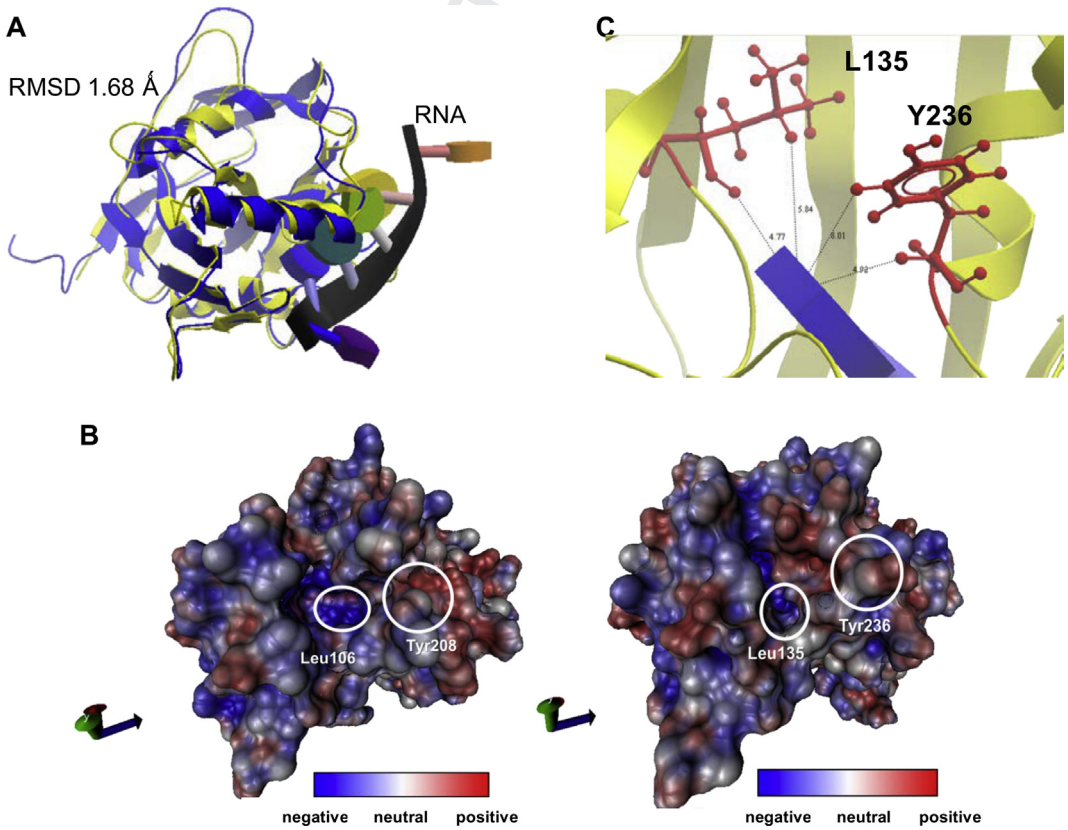

Fig. 2. Analyzes of the three-dimensional model of EhCFIm25. (A) Comparison of ribbon diagrams of EhCFIm25 3D model (yellow) and human CFIm25 3D structure (blue) in proximity to the UGUA RNA molecule (black). EhCFIm25 3D model was predicted by Geno 3D2 Web Server software using chain A of the human CFIm25 (3MDG_A) as template. (B) Electrostatic potential surface representation of human (left) and *E. histolytica* (right) CFIm25 proteins. Positions of Leu135 and Tyr236 in EhCFIm25, and Leu106 and Tyr208 in human CFIm25, are indicated. (C) Magnification showing the relative position of the amino acids Leu135 and Tyr236 in EhCFIm25 (yellow) in proximity to the UGUA RNA molecule (blue). Distances to U2 of the UGUA RNA molecule are in Armstrong (Å).

Please cite this article in press as: J.D. Ospina-Villa, et al., Amino acid residues Leu135 and Tyr236 are required for RNA binding activity of CFIm25 in *Entamoeba histolytica*, Biochimie (2015), http://dx.doi.org/10.1016/j.biochi.2015.04.017

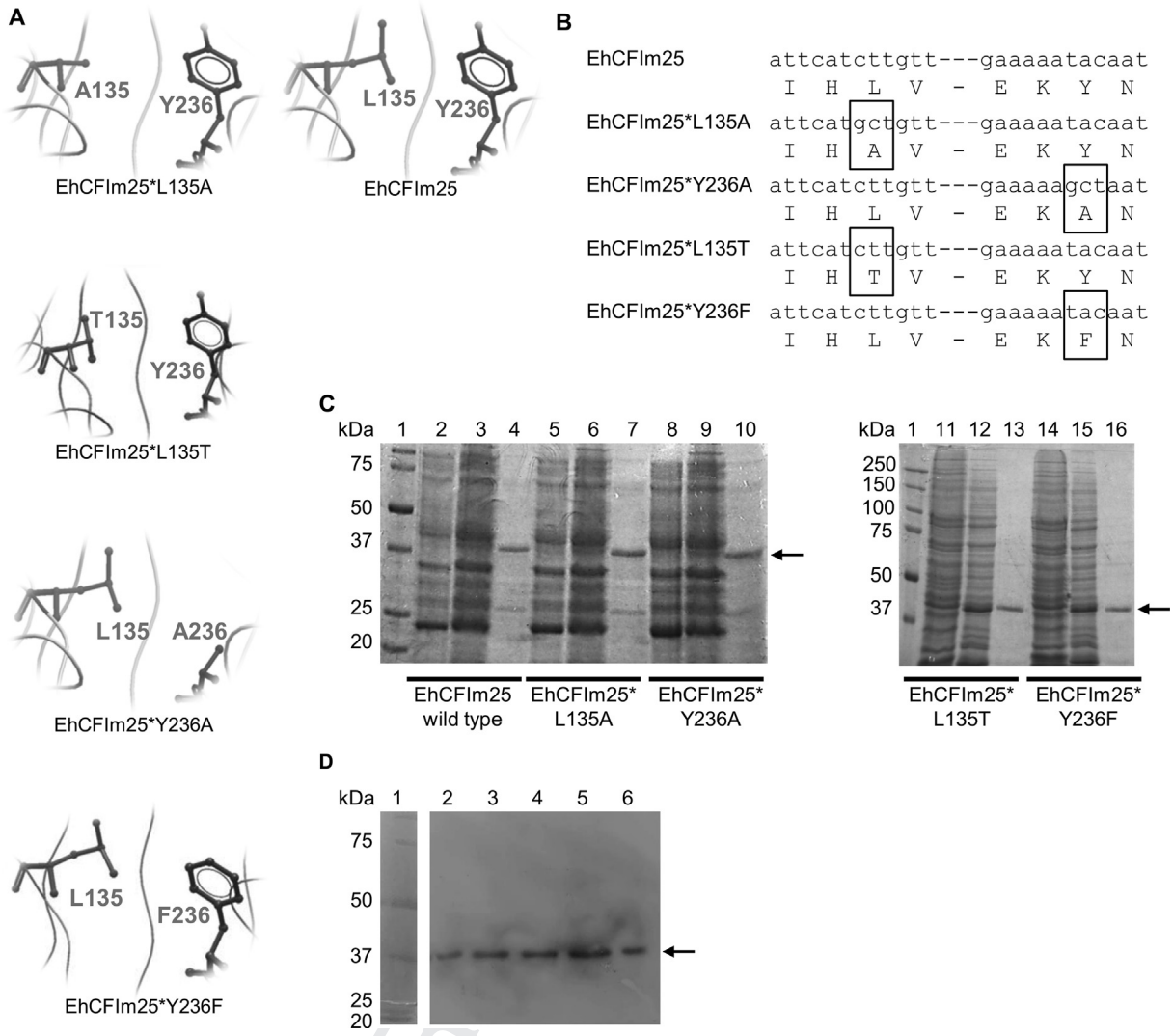

Fig. 3. Expression and purification of wild type EhCFIm25 and mutated EhCFIm25*L135A, EhCFIm25*Y236A, EhCFIm25*L135T and EhCFIm25*Y236F proteins. (A) Magnification showing the relative positions of the amino acids Ala135, Thr135, Ala236 and Phe236 in mutant proteins in comparison with those of Leu135 and Tyr236 in wild-type EhCFIm25. (B) Comparison of nucleotides (upper line) and amino acids (lower line) sequence of wild type and mutant *EhCFIm25* genes and predicted proteins. Mutated residues are indicated in box. (C) Expression and purification of EhCFIm25 proteins. Recombinant proteins were expressed in bacteria treated with 1 mM IPTG and purified by chromatography affinity through Ni-NTA columns. Lane 1, Molecular weight markers; lanes 2, 5, 8, 11 and 14, bacterial extracts before addition of IPTG; lanes 3, 6, 9,12 and 15, bacterial extracts at 3 h after addition of IPTG; lanes 4, 7, 10, 13 and 16, purified recombinant proteins. (D) Immunodetection of purified recombinant proteins using anti-6x-His tag antibodies (Roche) in Western blot assays. Lane 1, Molecular markers; lane 2, EhCFIm25; lane 3, EhCFIm25*L135A; lane 4, EhCFIm25*Y236A; lane 5, EhCFIm25*L135T; lane 6, EhCFIm25*Y236F. Arrow, recombinant EhCFIm25 proteins.

under nondenaturing conditions (Fig. 3C). A unique band with the expected molecular weight was recognized by monoclonal anti-6x-His tag antibodies in Western blot assays, validating the identity of recombinant proteins (Fig. 3D).

To determine whether Leu135 and Tyr236 are necessary for RNA binding activity, we next performed REMSA using a biotin-labeled RNA fragment that corresponds to 100 nt of the 3′ UTR of the *EhPgp5* gene [25] as a probe with recombinant wild-type and mutant EhCFIm25 proteins (Fig. 4A). It has been previously reported that EhCFIm25 can bind the full length 3′ UTR of *EhPgp5* [28]. Here, we showed that purified EhCFIm25 was able to bind a smaller region of the *EhPgp5* 3′ UTR that contains two polyadenylation signals, two U-rich elements and one poly (A) site (Fig. 4B, lane 2). RNA-protein complex was specifically competed by a 350-fold molar excess of the same unlabeled transcript, while it was maintained in the presence of tRNA, used as unspecific competitor (lanes 3 and 4), showing the specificity of the interaction between EhCFIm25 and the 3′ UTR fragment used as probe. In addition, complex disappeared in the control experiment using Proteinase K (lane 5), confirming the presence of the protein within the complex. Finally, no complex was observed using mock fraction, which indicates that complex formation was due to the recombinant EhCFIm25 protein, and not to contaminating proteins from *E. coli* (lane 6). Interestingly, mutant proteins with amino acid substitutions at Leu135 and Tyr236 showed distinct degrees of RNA-binding activities. No complex was detected when EhCFIm25*L135A and EhCFIm25*Y236A proteins were used in REMSA experiment (Fig. 4C, lanes 2 and 3). Similarly, the EhCFIm25*Y236F protein was not able to bind RNA (lane 5), whereas EhCFIm25*L135T remains able to form a slight RNA complex (lane 4). Notably, replacement of Leu135 by Thr led to about 35% of wild-type EhCFIm25 activity (Fig. 4D).

Please cite this article in press as: J.D. Ospina-Villa, et al., Amino acid residues Leu135 and Tyr236 are required for RNA binding activity of CFIm25 in *Entamoeba histolytica*, Biochimie (2015), http://dx.doi.org/10.1016/j.biochi.2015.04.017

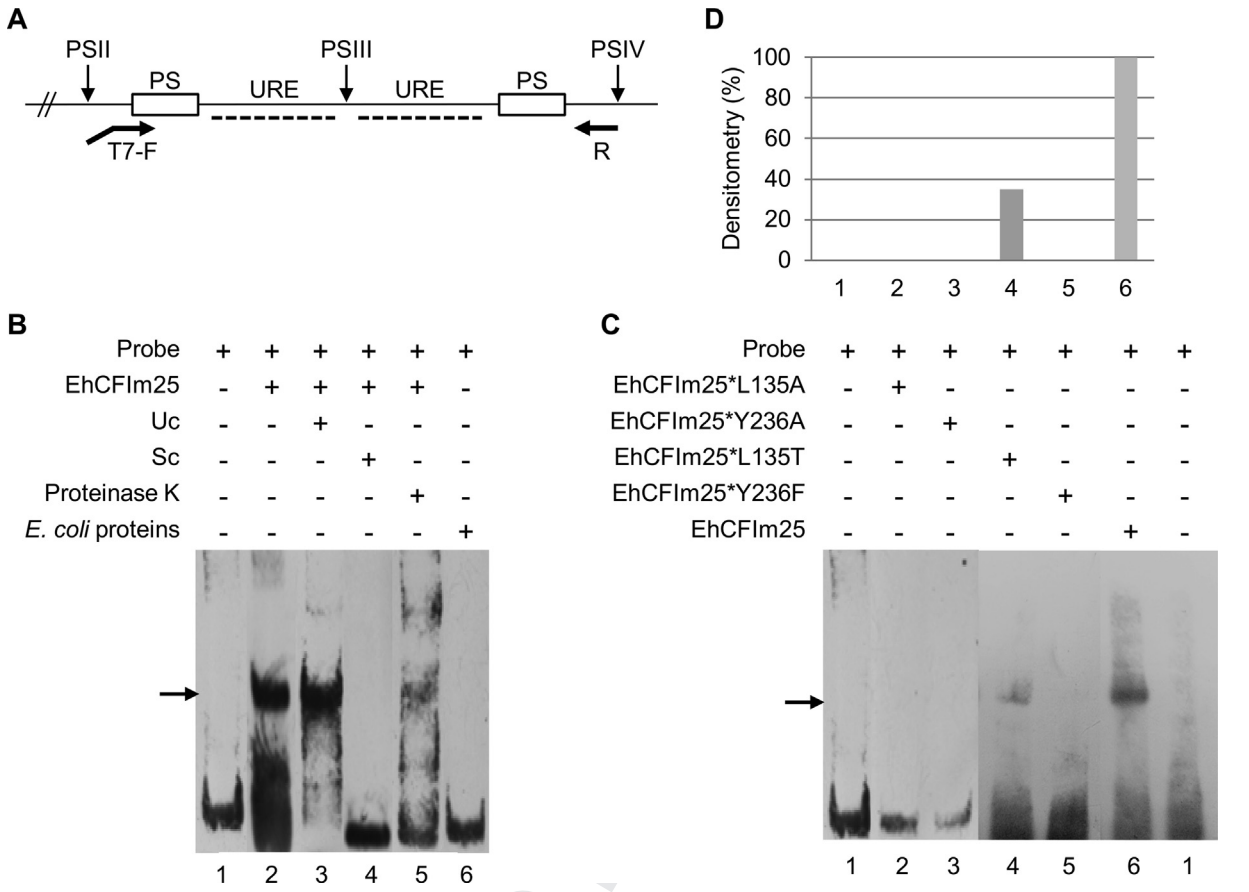

Fig. 4. RNA-protein interaction assays. (A) Molecular characteristics of the *EhPgp5* 3′UTR fragment used as RNA probe [25]. PS: Polyadenylation signal; PSII, PSIII, PSIV: Polyadenylation site I, II and III, PS: Polyadenylation sites; URE: U-rich element; T7-F: forward T7 oligonucleotide; R: reverse oligonucleotide. (B and C) REMSA. Recombinant wild type (B) and mutant (C) EhCFIm25 proteins (20 μg) were incubated with biotin-labeled RNA probe (5 ng/μl); RNA-protein complexes were resolved through PAGE and revealed by chemiluminescence assays. (B) Lane 1, free probe; lane 2; probe + EhCFIm25; lane 3, probe + EhCFIm25 + unspecific competitor (Uc); lane 4, probe + EhCFIm25 + specific competitor (Sc); lane 5, probe + EhCFIm25 + proteinase K (1 μg); lane 6, probe + mock fraction. (C) Lane 1, free probe; lane 2; probe + EhCFIm25*L135A; lane 3, probe + EhCFIm25*Y236A; lane 4, probe + EhCFIm25*L135T; lane 5, probe + EhCFIm25*Y236F; lane 6, probe + wild type EhCFIm25 protein, as control. (D) Representative densitometry analysis of RNA-protein complexes from (C). Pixels corresponding to the RNA-protein complex formed with the wild-type EhCFIm25 protein were taken as 100%.

4. Discussion

The formation of poly(A) tail at pre-mRNA 3′ end is essential for mRNA export, stability, and translation to functional proteins, influencing virtually all aspects of RNA metabolism in eukaryotic cells [1–7]. In human, the 25 kDa subunit of the heterotetrameric CFIm complex is an essential component for the poly(A) site selection in 3′UTR and the generation of mRNAs with poly(A) tails of different sizes [31]. The protozoan *E. histolytica* has a EhCFIm25 protein that is able to bind on its own the full-length *EhPgp5* mRNA 3′ UTR, forming two RNA-protein complexes in REMSA. This RNA sequence has all the pre-mRNA 3′ UTR *cis*-regulatory sequences described in this parasite, including various PAS, poly(A) sites, U-rich tracts, and A-rich elements that control the alternative polyadenylation of the *EhPgp5* mRNA [27,28]. Here, we showed that EhCFIm25 is able to form a single complex with a shorter region of *EhPgp5* mRNA 3′ UTR that only contains one cleavage site surrounded by two polyadenylation signals and two U-rich sites. This confirms that EhCFIm25 could be an important regulator of the cleavage/polyadenylation reaction in *E. histolytica*, as it has been described for the human protein [13–16].

The overall protein architecture prediction of EhCFIm25-UGUA complex is nearly identical to that of the previously published human CFIm25-UGUA complex. EhCFIm25 is composed of a central domain encompassing residues 77–202, which adopts a α/β/α fold common to all Nudix proteins. Importantly, the Nudix domain contains the majority of the residues involved in RNA binding activity of human CFIm25 [22–24]. However, the conserved residues located outside the Nudix domain, such as Leu106 and Tyr208 in human CFIm25 (Leu135 and Tyr236 in EhCFIm25), may also be relevant for protein functions. Indeed, the main-chain amide of Leu106 interacts with U1 through an intermolecular hydrogen bond that involves stabilization of O4 of the ribose via a glycerol molecule in the crystal structure of the human CFIm25/UGUA complex. U1 is further stabilized by stacking of the uracil base with the plane formed by the peptide bond between Tyr208 and Gly209 [22–24]. However, the functional relevance of Leu106 and Tyr208 has not been demonstrated in REMSA. Congruently with their possible role in RNA binding, the corresponding Leu135 and Tyr236 residues in EhCFIm25 are in proximity to the UGUA molecule in our 3D structure model. Moreover, Leu135 is located within a negative cavity formed by polar amino acids that is known as group of ambience and is thought to avoid water entrance to strengthen weak interactions between the protein and its substrate [32]. In addition, the location of Tyr208 and Tyr236 in the external part of the protein could contribute to the formation of hydrogen bonds with RNA to increase substrate binding and promote substrate positioning [33]. The independent shortening of lateral chain in each amino acids entirely abolished RNA binding ability of EhCFIm25, which suggests the participation of these particular

Please cite this article in press as: J.D. Ospina-Villa, et al., Amino acid residues Leu135 and Tyr236 are required for RNA binding activity of CFIm25 in *Entamoeba histolytica*, Biochimie (2015), http://dx.doi.org/10.1016/j.biochi.2015.04.017

chemical groups in RNA-protein interaction. Congruently, a slight RNA affinity remained when the hydrophobic Leu135 was changed to Thr, without affecting the lateral chain significantly. In contrast, the Tyr236Phe change also suppressed RNA binding, indicating that the OH group of the phenol ring at position 236 is essential for RNA interaction. Importantly, all four single point mutation variant forms share the same predicted folding as wild-type EhCFIm25 protein (data not shown), showing that protein variants are properly folded. Taken altogether, our results suggest that Leu135 mainly interacts with the UGUA molecule through its lateral chain, while Tyr236 may create different types of bonds to interact with RNA. Additionally, we cannot discard the participation of other residues in RNA interaction, including those corresponding to Glu55, Arg63, and Phe103 in the human protein.

5. Conclusions

Using EhCFIm25 of *E. histolytica* as an experimental model, this work showed for the first time -until we know-the functional relevance of the conserved Leu135 and Tyr236 (Leu106 and Tyr208 in human CFIm25) for RNA binding activity. Further analyses of EhCFIm25 variants folding by circular dichroism and molecular dynamics currently in progress, will help to understand better how L135 and Y236 interact with RNA.

Disclosure

All authors agree with the version that is submitted here and declare no conflict of interest. The work described has not been published previously and it is not under consideration for publication elsewhere. The funding sources had no involvement study design, in the collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the article for publication.

Acknowledgments

We greatly thank Dr. Huanting Liu, University of St Andrews, UK, for his technical support for mutagenesis assays. This work was supported by Mexican grants from CONACyT (178550) and SIP-IPN (20150720). Absalom Zamorano-Carrillo, Esther Ramirez-Moreno and Laurence A. Marchat are supported by COFAA-IPN. Juan David Ospina-Villa is a scholarship recipient from Mexican BEIFI-IPN and CONACyT programs.

References

[1] A. Jacobson, S.W. Peltz, Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells, Annu. Rev. Biochem. 65 (1996) 693–739, http://dx.doi.org/10.1146/annurev.bi.65.070196.003401.

[2] M. Wickens, P. Anderson, R.J. Jackson, Life and death in the cytoplasm: messages from the 3′ end, Curr. Opin. Genet. Dev. 7 (1997) 220–232, http://dx.doi.org/10.1016/S0959-437X(97)80132-3.

[3] N.L. Garneau, J. Wilusz, C.J. Wilusz, The highways and byways of mRNA decay, Nat. Rev. Mol. Cell Biol. 8 (2007) 113–126, http://dx.doi.org/10.1038/nrm2104.

[4] D.F. Colgan, J.L. Manley, Mechanism and regulation of mRNA polyadenylation, Genes Dev. 11 (November 1, 1997) 2755–2766, http://dx.doi.org/10.1101/gad.11.21.2755.

[5] J. Zhao, L. Hyman, C. Moore, Formation of mRNA 3′ ends in eukaryotes: mechanism, regulation and interrelationships with other steps in mRNA synthesis, Microbiol. Mol. Biol. Rev. 63 (1999) 405–445.

[6] M. Edmonds, A history of poly(A) sequences: from formation to factors to function, Prog. Nucleic Acid Res. Mol. Biol. 71 (2002) 285–389, http://dx.doi.org/10.1016/S0079-6603(02)71046-5.

[7] S. Millevoi, S. Vagner, Molecular mechanisms of eukaryotic pre-mRNA 3′ end processing regulation, Nucleic Acids Res. 38 (2010) 2757–2774, http://dx.doi.org/10.1093/nar/gkp1176.

[8] S. McCracken, N. Fong, E. Rosonina, K. Yankulov, G. Brothers, D. Siderovski, D.L. Bentley, 5′-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II, Genes Dev. 11 (24) (1997) 3306–3318, http://dx.doi.org/10.1101/gad.11.24.3306.

[9] Y. Hirose, J.L. Manley, RNA polymerase II is an essential mRNA polyadenylation factor, Nature 395 (1998) 93–96, http://dx.doi.org/10.1038/25786.

[10] Y. Shi, D.C. Di Giammartino, D. Taylor, A. Sarkeshik, W.J. Rice, J.R. Yates, J.L. Manley, Molecular architecture of the human pre-mRNA 3′ processing complex, Mol. Cell 33 (3) (2009) 365–376, http://dx.doi.org/10.1016/j.molcel.2008.12.028.

[11] H. De Vries, U. Rüegsegger, W. Hübner, A. Friedlein, H. Langen, W. Keller, Human pre-mRNA cleavage factor IIm contains homologs of yeast proteins and bridges two other cleavage factors, EMBO J. 19 (21) (2000) 5895–5904, http://dx.doi.org/10.1093/emboj/19.21.5895.

[12] U. Rüegsegger, K. Beyer, W. Keller, Purification and characterization of human cleavage factor Im involved in the 3′ end processing of messenger RNA precursors, J. Biol. Chem. 271 (1996) 6107–6113, http://dx.doi.org/10.1074/jbc.271.11.6107.

[13] K.M. Brown, et al., A mechanism for the regulation of pre-mRNA 3′ processing by human cleavage factor Im, Mol. Cell 12 (6) (2003) 1467–1476, http://dx.doi.org/10.1016/S1097-2765(03)00453-2.

[14] S. Dettwiler, C. Aringhieri, S. Cardinale, W. Keller, S.M.L. Barabino, Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein-protein interactions, and subcellular localization, J. Biol. Chem. 279 (34) (2004) 35788–35797, http://dx.doi.org/10.1074/jbc.M403927200.

[15] H. Kim, Y. Lee, Interaction of poly(A) polymerase with the 25-kDa subunit of cleavage factor I, Biochem. Biophys. Res. Commun. 289 (2001) 513–518, http://dx.doi.org/10.1006/bbrc.2001.5992.

[16] T. Kubo, T. Wada, Y. Yamaguchi, A. Shimizu, H. Handa, Knock-down of 25 kDa subunit of cleavage factor Im in Hela cells alters alternative polyadenylation within 3′-UTRs, Nucleic Acids Res. 34 (2006) 6264–6271, http://dx.doi.org/10.1093/nar/gkl794.

[17] A.R. Gruber, G. Martin, W. Keller, M. Zavolan, Cleavage factor Im is a key regulator of 3′ UTR length, RNA Biol. 9 (12) (2012) 1405–1412, http://dx.doi.org/10.4161/rna.22570.

[18] C.P. Masamha, Z. Xia, J. Yang, T.R. Albrecht, M. Li, A.B. Shyu, W. Li, E.J. Wagner, CFIm25 links alternative polyadenylation to glioblastoma tumour suppression, Nature 510 (7505) (2014) 412–416, http://dx.doi.org/10.1038/nature13261.

[19] S. Millevoi, C. Loulergue, S. Dettwiler, S.Z. Karaa, W. Keller, M. Antoniou, S. Vagner, An interaction between U2AF 65 and CF Im links the splicing and 3′ end processing machineries, EMBO J. 25 (20) (2006) 4854–4864, http://dx.doi.org/10.1038/sj.emboj.7601331.

[20] M.J. Bessman, D.N. Frick, S.F. O'Handley, The MutT proteins or "Nudix" hydrolases, a family of versatile, widely distributed, "housecleaning" enzymes, J. Biol. Chem. 271 (1996) 25059–25062, http://dx.doi.org/10.1074/jbc.271.41.25059.

[21] W. Ranatunga, E.E. Hill, J.L. Mooster, E.L. Holbrook, U. Schulze-Gahmen, et al., Structural studies of the Nudix hydrolase DR1025 from *Deinococcus radiodurans* and its ligand complexes, J. Mol. Biol. 339 (2004) 103–116, http://dx.doi.org/10.1016/j.jmb.2004.01.065.

[22] Q. Yang, Structural basis of UGUA recognition by the Nudix protein CFIm25 and implications for a regulatory role in mRNA 3′ processing, Proc. Natl. Acad. Sci. U S A 107 (22) (2010) 10062–10067, http://dx.doi.org/10.1073/pnas.1000848107.

[23] Q. Yang, M. Coseno, G.M. Gilmartin, S. Doublié, Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping, Structure 19 (2011) 368–377, http://dx.doi.org/10.1016/j.str.2010.12.021.

[24] H. Li, S. Tong, X. Li, H. Shi, Z. Ying, et al., Structural basis of pre-mRNA recognition by the human cleavage factor Im complex, Cell Res. 21 (2011) 1039–1051, http://dx.doi.org/10.1038/cr.2011.67.

[25] C. López-Camarillo, J.P. Luna-Arias, L.A. Marchat, E. Orozco, EhPgp5 mRNA stability is a regulatory event in the *Entamoeba histolytica* multidrug resistance phenotype, J. Biol. Chem. 278 (2003) 11273–11280, http://dx.doi.org/10.1074/jbc.M211757200.

[26] C. López-Camarillo, M.L. García-Hernández, L.A. Marchat, J.P. Luna-Arias, O. Hernández de la Cruz, L. Mendoza, E. Orozco, *Entamoeba histolytica* EhDEAD1 protein is a conserved DEAD-box RNA helicase with ATPase and ATP-dependent RNA unwinding activities, Gene 414 (2008) 19–31, http://dx.doi.org/10.1016/j.gene.2008.01.024.

[27] C. López-Camarillo, O.N. Hernández de la Cruz, J.G. Vivas, J.F. Retana, M.P. Valdez, I.L. Rosas, E. Álvarez-Sánchez, L.A. Marchat, Recent insights in pre-mRNA 3-end processing signals and proteins in the protozoan parasite *Entamoeba histolytica*, Infect. Disord. Drug Targets 10 (4) (2010) 258–265, http://dx.doi.org/10.2174/187152610791591575.

[28] M. Pezet-Valdez, J. Fernández-Retana, J.D. Ospina-Villa, M.E. Ramírez-Moreno, E. Orozco, et al., The 25 kDa subunit of cleavage factor Im is a RNA-binding protein that interacts with the poly(A) polymerase in *Entamoeba histolytica*, PLoS One 8 (6) (2013) e67977, http://dx.doi.org/10.1371/journal.pone.0067977.

[29] H. Liu, J.H. Naismith, An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol, BMC Biotechnol. 8 (2008) 91, http://dx.doi.org/10.1186/1472-6750-8-91.

[30] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH image to ImageJ: 25 years of image analysis, Nat. Methods 9 (7) (2012) 671–675, http://dx.doi.org/10.1038/nmeth.2089.

Please cite this article in press as: J.D. Ospina-Villa, et al., Amino acid residues Leu135 and Tyr236 are required for RNA binding activity of CFIm25 in *Entamoeba histolytica*, Biochimie (2015), http://dx.doi.org/10.1016/j.biochi.2015.04.017

[31] H. Fukumitsu, H. Soumiya, S. Furukawa, Knockdown of pre-mRNA cleavage factor Im 25 kDa promotes neurite outgrowth, Biochem. Biophys. Res. Commun. 425 (4) (2012) 848–853, http://dx.doi.org/10.1016/j.bbrc.2012.07.164.
[32] H. Yin, G. Feng, G.M. Clore, G. Hummer, J.C. Rasaiah, Water in the polar and nonpolar cavities of the protein interleukin-1β, J. Phys. Chem. 114 (49) (2010) 16290–16297, http://dx.doi.org/10.1021/jp108731r.
[33] T.L. Davis, J.R. Walker, A. Allali-Hassani, S.A. Parker, B.E. Turk, S. Dhe-Paganon, Structural recognition of an optimized substrate for the ephrin family of receptor tyrosine kinases, FEBS J. 276 (16) (2009) 4395–4404, http://dx.doi.org/10.1111/j.1742-4658.2009.07147.x.

Please cite this article in press as: J.D. Ospina-Villa, et al., Amino acid residues Leu135 and Tyr236 are required for RNA binding activity of CFIm25 in *Entamoeba histolytica*, Biochimie (2015), http://dx.doi.org/10.1016/j.biochi.2015.04.017