Cardiac Expression of the Cystic Fibrosis Transmembrane Conductance Regulator Involves Novel Exon 1 Usage to Produce a Unique Amino-terminal Protein*⑤

Wayne L. Davies‡§, Jamie I. Vandenberg§¶, Rana A. Sayeed§, and Ann E. O. Trezise‡‖

From the ‡School of Biomedical Science, University of Queensland, Brisbane, Queensland 4072, Australia, the §Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, United Kingdom, and the ¶Victor Chang Cardiac Research Institute, Sydney, New South Wales 2010, Australia

Cystic fibrosis is caused by mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, which encodes a chloride channel present in many cells. In cardiomyocytes, we report that multiple exon 1 usage and alternative splicing produces four CFTR transcripts, with different 5′-untranslated regions, CFTR$_{TRAD-139}$, CFTR$_{-1C/-1A}$, CFTR$_{-1C}$, and CFTR$_{-1B}$. CFTR transcripts containing the novel upstream exons (exons −1C, −1B, and −1A) represent more than 90% of cardiac expressed CFTR mRNA. Regulation of cardiac *CFTR* expression, in response to developmental and pathological stimuli, is exclusively due to the modulation of *CFTR$_{-1C}$* and *CFTR$_{-1C/-1A}$* expression. Upstream open reading frames have been identified in the 5′-untranslated regions of all CFTR transcripts that, in conjunction with adjacent stem-loop structures, modulate the efficiency of translation initiation at the AUG codon of the main CFTR coding region in CFTR$_{TRAD-139}$ and CFTR$_{-1C/-1A}$ transcripts. Exon −1A, only present in CFTR$_{-1C/-1A}$ transcripts, encodes an AUG codon that is in-frame with the main CFTR open reading frame, the efficient translation of which produces a novel CFTR protein isoform with a curtailed amino terminus. As the expression of this CFTR transcript parallels the spatial and temporal distribution of the cAMP-activated whole-cell current density in normal and diseased hearts, we suggest that CFTR$_{-1C/-1A}$ provides the molecular basis for the cardiac cAMP-activated chloride channel. Our findings provide further insight into the complex nature of *in vivo CFTR* expression, to which multiple mRNA transcripts, protein isoforms, and post-transcriptional regulatory mechanisms are now added.

Cystic fibrosis (CF)[1] is an autosomal recessive disorder that causes severe multisystem disease (1). The cause of CF is mutation of the cystic fibrosis transmembrane conductance regulator gene (*CFTR*) (2–4), which encodes a cAMP-activated, protein kinase A-dependent chloride channel (5, 6) and is a member of the ABC transporter superfamily of genes (7). *CFTR* exhibits spatial and temporal regulation of expression (8–10), accompanied by different transcription start site usage (11) and alternative splicing (12, 13). In addition to epithelial tissues, *CFTR* is expressed in cardiac muscle (14, 15) and neuronal tissues (16, 17), which may contribute to the pathogenesis of CF.

The expression of CFTR transcripts and cAMP-activated chloride currents in the heart have been demonstrated for humans and simians (18), rabbits (19), guinea pigs (20, 21), and cats (22), but are undetectable in murine (23) and canine hearts (24). Previous studies have shown that CFTR transcripts expressed in the heart are alternatively spliced, resulting in the loss of exon 5 (25). We have previously demonstrated that CFTR mRNA is expressed in an epicardial (higher) to endocardial (lower) gradient across the left ventricular free wall (LVFW) of the rabbit heart, coinciding with a 2.5:1 gradient in the cAMP-activated chloride current density in ventricular myocytes (26). Furthermore, this epicardial to endocardial gradient in *CFTR* expression is developmentally regulated, appearing in the first postnatal week (27), and lost in hypertrophic and failing hearts (28). The cardiac distributions of CFTR mRNA and functional channels are consistent with a role in the maintenance of the normal epicardial to endocardial gradients of ventricular repolarization and action potential duration in the heart (14, 26). In addition, the overall reduction and loss of the gradient of *CFTR* expression during hypertrophy could contribute to delayed ventricular repolarization (29) and loss of the gradient of repolarization in hypertrophied hearts (30), both of which are known to be arrhythmogenic (31). Indeed, it has been shown that some CF patients exhibit an increased risk of ventricular arrhythmia (32, 33); however, it is difficult to distinguish between a primary genetic cause due to the loss of *CFTR* expression and secondary effects as a result of pathological manifestations in the pulmonary system.

In this study, we investigated the regulation of temporal, regional, and pathological changes in *CFTR* expression in the

---

[1] The abbreviations used are: CF, cystic fibrosis; CFTR, cystic fibrosis transmembrane conductance regulator; LVFW, left ventricular free wall; ORF, open reading frame; uORF, upstream open reading frame; 5′-UTR, 5′-untranslated region; 5′-RACE, 5′-rapid amplification of cDNA ends; qPCR, quantitative polymerase chain reaction; eGFP, enhanced green fluorescent protein; CHO, Chinese hamster ovary; eBFP, enhanced blue fluorescent protein; uAUG, upstream AUG; gDNA, genomic DNA; LV, left ventricle; RV, right ventricle.

TABLE I
Oligonucleotide and probe sequences used for 5′-RACE PCR and TaqMan qPCR

Primer	Sequence	Description
Anchor primer/F	5′-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3′	Anneals to poly dC-tail
RC2/R	5′-TTTGGTATATGTCTGACAATTCCAGGCGCT-3′	5′-RACE rabbit CFTR exon 2
RC3/R	5′-ACACCTCCGAAGGGCATTATTGAGCTTAGG-3′	5′-RACE rabbit CFTR exon 3
RC4/R	5′-ACAAAGAGTAAGCACAGACCTATGCCCAGG-3′	5′-RACE rabbit CFTR exon 4
RC6/R	5′-CCAGCTCTCTGATCCCTGTACTTCATCATC-3′	5′-RACE rabbit CFTR exon 6
Exon −1C/F	5′-AACACGCTGTTATTCCTCACCTG-3′	TaqMan qPCR CFTR exon −1C
Exon −1B/F	5′-GCTCTAGTGAAGATGGTCTACTTGATGA-3′	TaqMan qPCR CFTR exon −1B
Exon −1A/F	5′-CATCAGAGTTGCACGAATCACAT-3′	TaqMan qPCR CFTR exon −1A
Traditional exon 1/F	5′-GAGAGACCATGCAGAAGTCGC-3′	TaqMan qPCR CFTR traditional exon 1
Exon 2/R	5′-CTGCAGAATCAGCAGAAGGGA-3′	TaqMan qPCR CFTR exon 2
Exon 4/F	5′-GCAGATGAGAATAGCCATGTTCAG-3′	TaqMan qPCR CFTR exon 4
Exon 5/F	5′-CTCCTTTCCAACAACCTGAACAA-3′	TaqMan qPCR CFTR exon 5
Exon 6/R	5′-GAAAGCAAGGCCGCAGAA-3′	TaqMan qPCR CFTR exon 6
Exon 2/F	5′-CTGGACTAGACCGATTTTGAGAAAAGGATACAGACA-3′	TaqMan qPCR CFTRexon 2 probe
Exon 6/F	5′-ACCTCTGCAAGTGACTCTGCTGATGGG-3′	TaqMan qPCR CFTRexon 6 probe

rabbit heart. We show that the majority of CFTR transcripts expressed in the heart initiate at unique transcription start sites and include novel alternative 5′-exons that replace the traditional *CFTR* exon 1. These alternative 5′-exons encode (*a*) a series of short upstream open reading frames (uORFs) in the 5′-untranslated region (5′-UTR), resulting in post-transcriptional regulation of *CFTR* expression, and (*b*) a unique translation initiation codon, in-frame with the main CFTR open reading frame (ORF), which results in a CFTR polypeptide with a distinct amino terminus. Modulation in the levels of these cardiac-specific CFTR transcripts is responsible for the temporal, spatial, and pathological changes in *CFTR* expression observed in the heart. Finally, the distal localization of these cardiac-specific, alternative 5′-exons upstream of the traditional *CFTR* exon 1 suggests the presence of a distinct promoter region directing *CFTR* expression in the heart.

EXPERIMENTAL PROCEDURES

Total RNA Source and Preparation—Triplicate New Zealand White rabbit tissue samples were collected from heart (embryo day 29, neonate day 7, juvenile week 3, adult year 1; left and right ventricular free wall, sham-operated control, and aortic-banded hypertrophy) and adult year 1 duodenum (control). All atrial tissue was removed from each cardiac sample harvested, except for embryo day 29, where whole hearts were used. Atrial tissue contributes less than 10% of the tissue weight to the embryo day 29 samples and will not significantly affect the comparison of *CFTR* expression during development. Moderate cardiac hypertrophy was induced in adult rabbits as previously described (34), resulting in an increased heart weight-to-body weight ratio of $26 \pm 2\%$. Sham-operated animals underwent aortic mobilization, but were not banded. Adult LVFW samples were dissected into three equal parts, along both the apical to basal and epicardial to endocardial axes, and the outer portions located at the apical epicardial, basal epicardial, apical endocardial, and basal endocardial surfaces were taken for further analysis. All procedures were performed in accordance with the United Kingdom Animals (Scientific Procedures) Act of 1986. Fresh tissues were homogenized in guanidinium thiocyanate solution (Fluka), and total RNA was isolated according to Chomczynski and Sacchi (35).

5′-Rapid Amplification of cDNA Ends (5′-RACE)—CFTR transcription start sites were identified by 5′-RACE as previously described (11). Total RNA (2 µg) was reverse transcribed using a CFTR exon 6 reverse primer (Table I). First-strand cDNAs, tailed with dCTP, were subject to two rounds of hemi-nested PCR using reverse primers to exons 4, 3, or 2 and an anchor primer to the dC tail. Amplicons were sequenced directly and subcloned into pLITMUS 28 (New England Biolabs).

Northern Blot Analysis—Rabbit CFTR cDNA (courtesy of Professor Burton Horowitz, University of Nevada School of Medicine, Reno, NV) was random-primer radiolabeled with [α-^{32}P]dCTP (3000 Ci mM, Amersham Biosciences; DECAprime II system, Ambion). Total RNA was electrophoresed on a 1% (w/v) denaturing agarose gel, transferred to a positively charged nylon membrane (GeneScreen Plus, PerkinElmer Life Sciences), incubated with labeled probe, and then washed as previously described (13). Autoradiography was with Hyperfilm MP x-ray film (Amersham Biosciences) for 5 days.

TaqMan Quantitative PCR (qPCR)—Relative *CFTR* expression was quantified using cDNA primed with random hexamer (500 ng) with transcript-specific forward primers (exons −1C, −1B, −1A, 1, 4, and 5), reverse primers and probes (exons 2 and 6), all designed to cross at least one exon-exon boundary (Table I). The 5′-reporter and 3′-quencher dyes used for CFTR probes were 6-carboxyfluorescein and 6-carboxytetramethylrhodamine, respectively, whereas the 18 S rRNA probe used the VIC reporter dye (Applied Biosystems). Each qPCR reaction consisted of 1× TaqMan Universal PCR Master Mix (Applied Biosystems), forward and reverse primers (250 nM), probe (50 nM), and cDNA template (100 ng). Triplicate experimental CFTR and endogenous 18 S rRNA reactions were performed using three independent cDNA samples from different animals. Reporter dye fluorescence was detected using an Applied Biosystems Prism 7700 Sequence Detector, and data were analyzed using Sequence Detector software version 1.6.3 (Applied Biosystems). Quantification of *CFTR* gene expression, relative to 18 S rRNA, was performed using the comparative threshold cycle method according to instructions from the manufacturer (Applied Biosystems).

Construction of CFTR-Enhanced Green Fluorescent Protein (eGFP) Fusion Constructs and Cell Culture—CFTR cDNA fragments containing each alternative or traditional exon 1 sequence plus exon 2 were ligated to the eGFP cDNA (Clontech) and the entire fragment subcloned into the pIRESneo mammalian expression vector (Clontech). Chinese hamster ovary (CHO) cells (ATCC, CRL-9096; courtesy of Dr. Yu Lu, University of Cambridge, Cambridge, United Kingdom) were grown in Iscove's modified Dulbecco's medium, 10% (v/v) fetal bovine serum, 4 mM L-glutamine, 18 mM sodium bicarbonate, 0.1 mM hypoxanthine, and 16 µM thymidine (Invitrogen). Triplicate wells of 60% confluent CHO cells were transiently co-transfected with 1 µg each of CFTR-eGFP and control enhanced blue fluorescent protein (eBFP) constructs, using LipofectAMINE according to the instructions from the manufacturer (Invitrogen). After 72 h of incubation, cells were trypsinized and transferred to a Thermo-Fast 96 black PCR plate (Abgene). Enhanced GFP and eBFP fluorescence was measured using a SPECTRAmax GEMINI-XS spectrofluorometer (36, 37). Prior calibration experiments determined the following optimal parameters: for eGFP, $\lambda_{ex} = 472$ nm, $\lambda_{em} = 512$ nm, and $\lambda_{cutoff} = 495$ nm; and for eBFP, $\lambda_{ex} = 378$ nm, $\lambda_{em} = 445$ nm, and $\lambda_{cutoff} = 420$ nm. To correct for differences in transfection efficiency, eGFP fluorescence values were normalized to eBFP signal. Background fluorescence (untransfected cells) was subtracted.

Bioinformatic and Statistical Analyses—Messenger RNA secondary structures were predicted using MFold version 3.1 (38, 39). CFTR amino-terminal hydropathy plots were generated using the Kyte and Doolittle algorithm (40). Protein secondary structures were predicted by PSIPRED version 2.4 (41). The charge distribution for the putative CFTR amino-terminal helix was predicted by helical wheel analyses using MacPlasmap Pro version 3.01. Statistical comparisons were made using Student's *t* tests (unpaired, two-tailed), and a *p* value of less than 0.05 was taken to indicate that a change in transcript level was of statistical significance.

RESULTS

CFTR Transcripts Expressed in Rabbit Heart Include Multiple Alternative 5′-Exons—The major *in vivo* transcription start sites of CFTR transcripts expressed in the heart were identified by 5′-RACE. Typical examples of 5′-RACE products obtained from adult LVFW and duodenal (control) tissues are shown in Fig. 1*A*. Sequencing of 5′-RACE amplicons revealed extensive
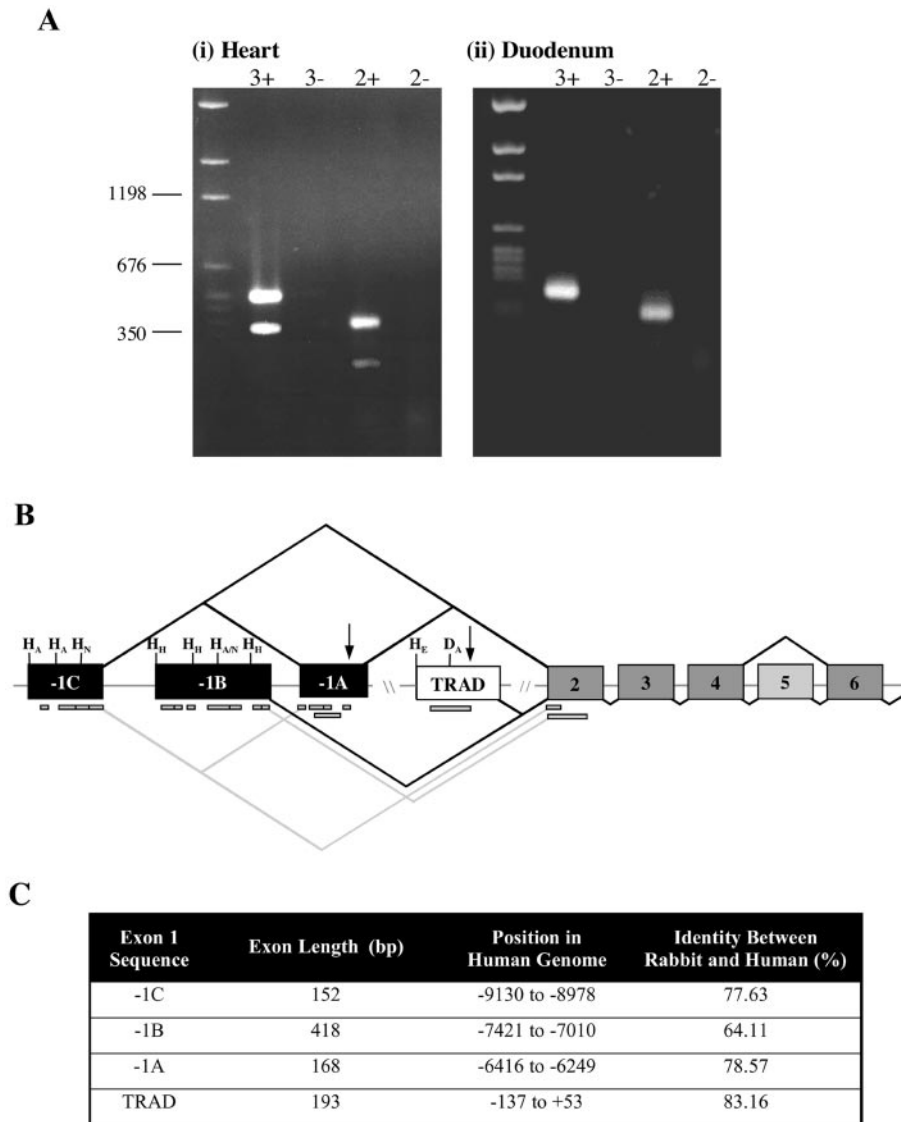
FIG. 1. **Differential transcription start site and alternative exon 1 usage in rabbit tissues.** *A*, representative 5′-RACE amplicons from adult rabbit (*i*) heart and (*ii*) duodenal cDNA samples. Sample order for both gels: *lane 1*, marker; *lane 2*, anchor primer and exon 3 primer amplicon (*3+*); *lane 3*, negative control reaction (cDNA not tailed with terminal deoxynucleotidyltransferase (*No TdT*)). *Lane 4*, anchor primer and exon 2 primer amplicon (*2+*). *Lane 5*, no TdT negative control. *B*, a schematic representation of the rabbit *CFTR* genomic locus showing alternative splicing (*black lines*) and uORFs. Transcription start sites (*black vertical line*) are indicated: *H*, heart; *D*, duodenum; with *subscript* designations for adult (*A*), embryo (*E*), adult hypertrophic heart (*H*); and neonate (*N*). *Vertical arrows* indicate an AUG codon in-frame with exon 2. \ \ represents intron −1A (6 kb in the human genome). // represents intron 1 (24.5 kb in the human genome). Putative uORFs are shown as *light gray boxes* beneath *CFTR* exons −1C to 2, where the 5′ end represents an AUG codon, the 3′ end represents a termination codon, and the intervening *vertical lines* represent internal AUG codons. Junctions between uORFs that cross exon-exon boundaries are shown with *light gray lines*. *C*, summary table showing the degree, the length, and the relative position of identity between each rabbit *CFTR* alternative exon 1 sequence and the human BAC sequence (GenBank™ accession no. AC000111). The relative position of identity is *numbered* with respect to the human *CFTR* translation initiation codon.

variation in transcription start site and exon 1 usage in different tissues and at different developmental stages.

In adult duodenal tissues, a single transcription start site was identified 74 bp upstream of the first base of the translation initiation codon in traditional rabbit CFTR exon 1 ($CFTR_{TRAD-74}$). The nucleotide sequence was identical to the published rabbit CFTR cDNA and promoter sequences (GenBank™ accession nos. OC40227 and X95931). Similarly, CFTR transcripts expressed in embryonic cardiac ventricle initiated in traditional exon 1, but included some 65 additional nucleotides and initiated transcription at position −139 bp ($CFTR_{TRAD-139}$) (Fig. 1*B*; GenBank™ accession no. AY256889). In adult and neonatal cardiac ventricle, as well as hypertrophic heart, multiple *in vivo* transcription start sites were identified (Fig. 1, *A* and *B*). Collectively, sequence analysis revealed three

new alternative 5′-CFTR exons, designated exon −1C, exon −1B, and exon −1A, respectively (GenBank™ accession nos. AY256886–AY256888). Exon −1C and exon −1B are spliced directly to exon 2, with exon −1A subject to alternative splicing between exons −1C and 2 ($CFTR_{-1C}$, $CFTR_{-1B}$, and $CFTR_{-1C/-1A}$ transcripts). All splice donor and acceptor sequences satisfied the GT-AG rule (42). Fig. 1*B* shows the genomic structure and alternative splicing of CFTR transcripts expressed in the heart. In all cases, amplification reactions from control untailed cDNA were negative. Identical CFTR transcription start sites were mapped using three independent samples for all tissue types, confirming the precise location of each identified transcription start site and ensuring artifacts caused by partial mRNA degradation were avoided.

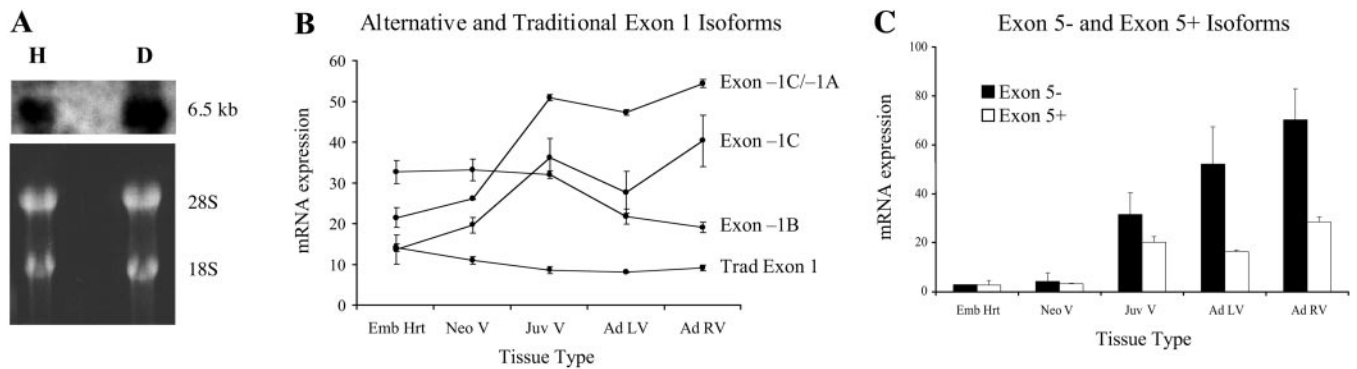Transcripts that included exon −1A contain an AUG codon

FIG. 2. **Rabbit *CFTR* expression in the developing heart.** *A*, Northern blot analysis of CFTR mRNA in rabbit adult heart ventricle (*H*) and duodenum (*D*) (*top panel*). Corresponding ethidium bromide-stained agarose gel with 30 μg of total RNA/lane (*bottom panel*). *B* and *C*, quantitative PCR analysis of *CFTR* expression during cardiac development. *B*, the relative expression levels of each alternative CFTR transcript, $CFTR_{-1C}$, $CFTR_{-1B}$, $CFTR_{-1C/-1A}$, and traditional $CFTR_{TRAD-139}$, were measured using forward primers specific to exon 1 of each transcript and a common reverse primer and TaqMan probe to CFTR exon 2. *C*, the relative expression levels of the exon 5 alternatively spliced CFTR transcripts, $CFTR_{EXON5-}$ (*black*) and $CFTR_{EXON5+}$ (*white*), were measured using forward primers in exons 4 and 5 and a common reverse primer and TaqMan probe to CFTR exon 6. Reaction conditions were such that the exon 4 primer only amplified the CFTR exon 5 minus transcript (52). All expression levels were measured relative to 18 S rRNA. In all cases *error bars* equal one standard deviation. *Hrt*, heart; *Emb*, embryo day 29; *Neo*, neonate day 7; *Juv*, juvenile week 3; *V*, both left and right ventricles; *Ad*, adult; *LV*, left ventricle; *RV*, right ventricle.

in-frame with the main CFTR ORF. Translation initiation from this AUG codon would result in CFTR protein with a novel amino-terminal region. Exons −1B and −1C did not contain an AUG codon in-frame with the main CFTR ORF. However, many upstream AUGs (uAUGs) were identified in the 5′-UTR of CFTR transcripts including exons −1C, −1B, or −1A, and a single uAUG codon was identified in the 5′-UTR of $CFTR_{TRAD-139}$. These uAUGs defined the start of several uORFs, which were followed by translation termination codons and thus encoded short peptides. The presence of short uORFs in the 5′-UTRs of eukaryotic mRNA transcripts is indicative of post-transcriptional and translational regulation of gene expression (43). Data base comparisons of each putative short polypeptide sequence revealed no homology to known protein sequences.

Genomic DNA Structure and Transcriptional Control of CFTR Expression in the Heart—Rabbit genomic DNA (gDNA) spanning *CFTR* exon −1C to exon −1A, was amplified by long range PCR and the sequence compared with the human and mouse gDNA sequences (GenBank[TM] accession nos. AC000111 and AF162137; see Fig. 1*C* and supplemental data (available in the on-line version of this article) showing dot matrix comparisons of rabbit, mouse, and human sequences). We also compared *CFTR* traditional exon 1 and flanking sequences across the three species. In both exonic and intronic gDNA regions, human and rabbit sequences showed a high level of identity for both the *CFTR* exon −1C to −1A region and traditional exon 1 (Fig. 1*C*). However, only *CFTR* traditional exon 1 sequences showed any similarity between the mouse and the other species. The human homologues of rabbit *CFTR* exons −1C to −1A spanned a region located 6–9 kb upstream of the traditional human *CFTR* exon 1. Further, the short uORFs and the rabbit *CFTR* exon −1A AUG codon, that is in-frame with the main CFTR ORF, are conserved in the human genome. The presence of homologues of the rabbit alternative exon 1 sequences in the human genome, but not the mouse, is consistent with expression of *CFTR* in the hearts of many species, including the rabbit (Fig. 3*A*) (15) and human (18), but not in the mouse (tested by RNA *in situ* hybridization and two rounds of reverse transcription-polymerase chain reaction; data not shown). Only 2 kb of genomic DNA sequence upstream of rabbit *CFTR* traditional exon 1 is available, and the alternative *CFTR* exons −1A to −1C lie beyond that. Long range PCR between *CFTR* traditional exon 1 and exons −1A to −1C, using rabbit gDNA template, was unsuccessful. Therefore, the promoter region

controlling the expression of *CFTR* exons −1C, −1B, and −1A is likely to be distally located and cardiac-specific.

Greater than 90% of CFTR Transcripts Expressed in the Heart Initiate at Exon −1C or −1B—Total CFTR mRNA expression in rabbit adult LVFW was similar to the level of *CFTR* expression detected in duodenal epithelium. A 6.5 kb CFTR transcript was detected in both tissues (Fig. 2*A*), which is consistent with the known CFTR transcript size (3). Past studies have also shown alternative splicing of exon 5 (25), and changes in total *CFTR* expression during heart development (27) and in cardiac hypertrophy (28). To investigate the biological relevance of the multiple alternative CFTR transcripts, we measured the relative expression level of individual CFTR transcripts in rabbit heart tissue from embryo day 29, neonate day 7, juvenile week 3, and normal adult left ventricle (LV) and right ventricle (RV).

We used TaqMan quantitative PCR (qPCR) to quantify each distinct CFTR transcript: $CFTR_{-1C}$, $CFTR_{-1B}$, $CFTR_{-1C/-1A}$, $CFTR_{TRAD-139}$, $CFTR_{EXON5+}$, and $CFTR_{EXON5-}$. In all cases, amplicons were less than 150 bp, the standard curve plots showed a very high correlation coefficient ($R^2 > 0.99$, $p < 0.01$), and amplification efficiencies were close to 100%. The expression levels of each CFTR transcript were measured in arbitrary relative expression units and normalized to 18 S rRNA. We compared *CFTR* expression in whole heart (atria and ventricles) from embryo day 29, and ventricular tissues from neonate day 7, juvenile week 3, and normal adult year 1 LV and RV. Atrial tissue contributes less than 10% of the tissue weight to the embryo day 29 samples and will not significantly affect the comparison of *CFTR* expression during development.

The predominantly expressed CFTR transcript in early rabbit cardiac development was $CFTR_{-1B}$, with lower levels of the $CFTR_{-1C}$ and $CFTR_{-1C/-1A}$ transcripts (Fig. 2*B*). As development proceeds the expression of $CFTR_{-1B}$ transcripts decreased ($p < 0.05$), whereas expression of both $CFTR_{-1C}$ and $CFTR_{-1C/-1A}$ transcripts increased and became the predominant CFTR transcripts in the adult rabbit heart ($p < 0.05$ and $p < 0.05$, respectively). In contrast, the expression of $CFTR_{TRAD-139}$ transcripts was low and static. In adult cardiac tissue, greater than 90% of total CFTR mRNA transcripts include the novel alternative exons (exon −1C, exon −1B, and exon −1C/−1A) described here. This contrasts with duodenum, where more than 95% of CFTR transcripts initiate from traditional CFTR exon 1.

Analysis of CFTR exon 5 alternative splicing showed equal levels of exon 5+ and exon 5- CFTR transcripts at embryonic and neonatal stages (Fig. 3*C*). In hearts from juvenile and adult animals, there were differential increases in expression of both $CFTR_{EXON5+}$, and $CFTR_{EXON5-}$ isoforms. This resulted in $CFTR_{EXON5-}$ transcripts being 3-fold more abundant than $CFTR_{EXON5+}$ transcripts in the LVFW of adult hearts ($p < 0.05$). This also contrasts with CFTR expression in adult duodenum, where over 90% of CFTR transcripts include exon 5.

The Expression of CFTR Transcripts Initiating at Exon −1C Is Primarily Responsible for the Epicardial to Endocardial Gradient across the LVFW—Epicardial (higher) to endocardial (lower) gradients across the rabbit LVFW have been shown for both cAMP-activated chloride currents and *CFTR* expression (26). Quantitative analysis of the differential distribution of each CFTR transcript in the adult LVFW shows that $CFTR_{-1C}$ and $CFTR_{-1C/−1A}$ transcripts, as well as alternatively spliced $CFTR_{EXON5-}$ transcripts, are responsible for the epicardial to endocardial gradient of *CFTR* expression and function in the left ventricle (Fig. 3, *A* (*i* and *iii*) and *B*, $p < 0.05$ in all cases). In contrast, there were no substantial differences in epicardial *versus* endocardial expression of $CFTR_{-1B}$, $CFTR_{TRAD-139}$, or $CFTR_{EXON5+}$ transcripts (Fig. 3, *A* (*ii* and *iv*) and *B*).

This work has also identified a second, perpendicular gradient in *CFTR* expression along the apical to basal axis of the left ventricle (Fig. 3, *A* and *B*). All three cardiac-specific CFTR transcripts ($CFTR_{-1C}$, $CFTR_{-1B}$, and $CFTR_{-1C/−1A}$) and both exon 5 alternatively spliced forms contribute to the apical to basal gradient ($p < 0.05$ in all cases). Collectively, these data show that *CFTR* expression is distributed in a radial pattern across the LVFW: highest at the apical epicardial surface and decreasing radially to the lowest point at the basal endocardial surface.

Cardiac Hypertrophy Causes the Loss of the Epicardial to Endocardial Gradient, but Does Not Alter the Apical to Basal Gradient of CFTR Expression—Cardiac ventricular hypertrophy is associated with a loss of repolarizing ion currents, including CFTR (28), and prolongation of the ventricular action potential duration (44). Here we show that the loss of the epicardial to endocardial *CFTR* expression gradient, in hypertrophic hearts, is caused by the preferential down-regulation of $CFTR_{-1C}$ and $CFTR_{-1C/−1A}$ transcripts (Fig. 3*A*; $p < 0.05$ in all cases), whereas $CFTR_{-1B}$ and $CFTR_{TRAD-139}$ transcripts were unaffected (Fig. 3*A* (*ii* and *iv*)). Although cardiac hypertrophy resulted in the loss of the epicardial to endocardial *CFTR* expression gradient, the apical to basal gradient was unaffected.

Cardiac hypertrophy also differentially affected the expression of exon 5 alternatively spliced CFTR transcripts (Fig. 3*B*). The establishment of cardiac hypertrophy leads to a loss of the epicardial to endocardial gradient of $CFTR_{EXON5-}$ transcripts, whereas $CFTR_{EXON5+}$ transcripts remain evenly distributed across the epicardial to endocardial axis. Further, both exon 5 alternatively spliced transcripts contribute to the apical to basal gradient in *CFTR* expression. The overall increase in *CFTR* expression during development reflects a preferential accumulation of $CFTR_{-1C}$ and $CFTR_{-1C/−1A}$ transcripts and a shift in exon 5 alternative splicing, such that $CFTR_{EXON5-}$ transcripts predominate. This process is reversed in cardiac hypertrophy with preferential loss of $CFTR_{-1C}$, $CFTR_{-1C/−1A}$, and $CFTR_{EXON5-}$ transcripts.

The Translation Initiation Codon in Exon −1A Supports CFTR Protein Production—In the absence of traditional CFTR exon 1, Carroll and co-workers (45) have provided evidence for translation initiation from downstream AUG codons, for example those present in exons 3 and 4. Of the CFTR transcripts expressed in the heart, only $CFTR_{TRAD-139}$ and $CFTR_{-1C/−1A}$

transcripts contain an AUG codon upstream of exon 2 and in-frame with the main CFTR ORF. Both translation initiation codons, in exon −1A and traditional exon 1, equally match an optimal Kozak consensus sequence. Fig. 4*A* shows a comparison of the sequences surrounding each AUG codon and the Kozak consensus sequence (46). AUG codons present in exons 3 and 4 all showed significantly lower identity to the Kozak consensus sequence, suggesting translation initiation at AUG codons in exons 3 and 4 would be less efficient than at AUG codons in either exon −1A or traditional exon 1. Exon 3 and 4 AUG codons are the first AUG codons in-frame with the main CFTR ORF present in $CFTR_{-1C}$ and $CFTR_{-1B}$ transcripts.

To investigate the potential for translation of each alternative CFTR transcript identified in this study, fusion constructs were produced that linked the amino-terminal alternative exon 1 sequences, plus exon 2, to the cDNA encoding an enhanced *Aequorea victoria* eGFP, each under the control of a cytomegalovirus promoter. An internal ribosomal entry site element, located downstream of the eGFP translation termination codon, artificially stabilized the expressed CFTR-eGFP transcripts, thus limiting any undesired variation in translation efficiency resulting from differences in mRNA stability. Independent constructs were transiently transfected in CHO cells, and GFP fluorescence, indicating protein expression, was measured with a spectrofluorometer.

This experiment showed that efficient translation initiation does occur from the AUG codon identified in CFTR exon −1A, with an identical level of expressed protein produced by translation from the AUG codon in traditional CFTR exon 1 (Fig. 4*B*). There was no significant protein production from exon −1C and exon −1B fusion constructs. Protein expression from the AUG codons in either exon −1A or traditional exon 1 were 8-fold lower than eGFP alone ($p < 0.01$), suggesting that there may be elements within the 5′-UTRs of either $CFTR_{-1C/−1A}$ or $CFTR_{TRAD-139}$ transcripts that may modulate the efficiency of translation initiation.

Upstream ORFs and 5′-UTR Secondary Structure Post-transcriptionally Modulate CFTR Expression—The presence of uAUG codons, distinct from the main ORF initiating methionine, in the 5′-UTR of most eukaryotic genes is unusual (47, 48). Inspection of the 5′-UTRs of the $CFTR_{-1C}$, $CFTR_{-1B}$, $CFTR_{-1C/−1A}$, and $CFTR_{TRAD-139}$ transcripts identified 16 putative translation initiation codons associated with uORFs. In the $CFTR_{-1C/−1A}$ transcript, we identified five uORFs in the 5′-UTR, whereas $CFTR_{TRAD-139}$ encoded one uORF in the 5′-UTR. McCarthy and co-workers (43) have demonstrated that uORFs and mRNA secondary structure, such as stem-loops, can act alone or in combination to regulate translation initiation efficiency at downstream AUG codons.

Although both the $CFTR_{TRAD-139}$/eGFP and the $CFTR_{-1C/−1A}$/eGFP constructs supported translation, the efficiency of translation was reduced compared with eGFP alone. With uORFs and adjacent stem-loop secondary structures in both these CFTR 5′-UTRs, we shortened the 5′-UTRs to remove or reduce the number of uORFs and measured translation from the AUG of the main CFTR ORF. The $CFTR_{-1C/−1A}$/eGFP construct was truncated to $CFTR_{-1A}$/eGFP, reducing the number of uORFs from 5 to 3. Also, the 5′-UTR of the $CFTR_{TRAD-139}$/eGFP construct was truncated by 65 nucleotides, removing the uORF and producing a 5′-UTR typical of CFTR transcripts expressed in duodenum ($CFTR_{TRAD-74}$/eGFP).

Reduction in the number of uORFs located upstream of the in-frame AUG present in exon −1A caused a statistically significant ($p < 0.05$, Student's *t* test) 1.5-fold increase in protein production (Fig. 5*A*). However, the continuing presence of three uORFs is a likely reason for the still relatively low level of
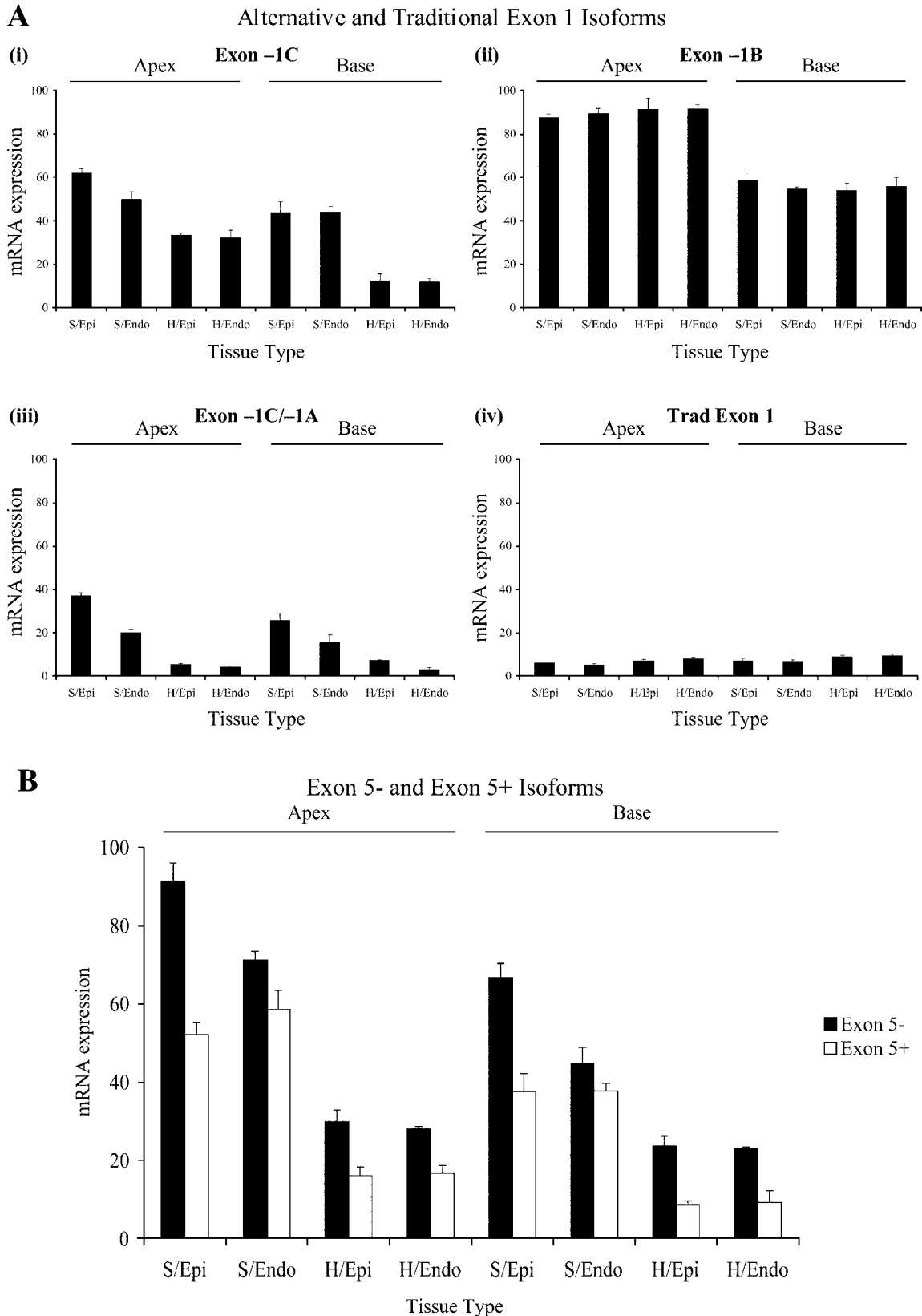
FIG. 3. **The effect of hypertrophy on rabbit *CFTR* expression across the LVFW.** This figure shows a comparison of *CFTR* expression in four distinct regions of the LVFW (epicardial apex, epicardial base, endocardial apex, and endocardial base) from animals with surgically induced left ventricular hypertrophy and in sham-operated animals. *A*, the relative expression levels of each alternative CFTR transcript, $CFTR_{-1C}$ (*i*), $CFTR_{-1B}$ (*ii*), and $CFTR_{-1C/-1A}$ (*iii*), and traditional (*iv*) $CFTR_{TRAD-139}$ transcripts, were measured using forward primers specific to exon 1 of each transcript and a common reverse primer and TaqMan probe to CFTR exon 2. *B*, the relative expression levels of the exon 5 alternatively spliced CFTR transcripts, $CFTR_{EXON5-}$ (*black*) and $CFTR_{EXON5+}$ (*white*), were measured using forward primers in exons 4 and 5 and a common reverse primer and TaqMan probe to CFTR exon 6. Reaction conditions were such that the exon 4 primer only amplified the CFTR exon 5 minus transcript (52). All expression levels were measured relative to 18 S rRNA. In all cases *error bars* equal one standard deviation. *S*, sham-operated control heart tissue; *H*, hypertrophic heart tissue; *Epi*, epicardium; *Endo*, endocardium.

A

AUG$_{position}$	Flanking sequence			Identity (%)
Kozak consensus	GCCGCC	$\boxed{\text{A}}$ $\boxed{\text{G}}$ (−3)	CC**AUG**$\boxed{\text{G}}$ (+1 +4)	(+5 +6) ($\underline{\text{A}}$ ($\underline{\text{U}}$) ($\underline{\text{C}}$	100.00 (100.00)
AUG$_{-1A}$	GAUCCA	C	AC**AUG**$\boxed{\text{G}}$	($\underline{\text{C}}$ A)	53.85 (53.33)
AUG$_1$	CGAGAG	$\boxed{\text{A}}$	CC**AUG**C	($\underline{\text{A}}$ G)	53.85 (53.33)
AUG$_3$	UGGAGA	U	UU**AUG**U	(U C)	23.08 (20.00)
AUG$_{4.1}$	CAUAUU	$\boxed{\text{G}}$	GA**AUG**C	($\underline{\text{A}}$ G)	30.77 (33.33)
AUG$_{4.2}$	GGAAUG	C	AG**AUG**A	(G A)	30.77 (26.67)
AUG$_{4.3}$	AGAAUA	$\boxed{\text{G}}$	CC**AUG**U	(U C)	46.15 (40.00)

FIG. 4. **Efficiency of translation of each CFTR transcript expressed in rabbit heart.** *A*, comparison of putative translation start codons in exons −1A, 1, 3, and 4 to the Kozak consensus sequence. Important nucleotides surrounding the translation initiation codon (*bold*) at positions −3 and +4 (*boxed*) and +5 and +6 (*underlined*) are indicated. Percentage homology over positions −9 to +4 (*no brackets*) and −9 to +6 (*brackets*). *B*, protein expression from CFTR-eGFP fusion constructs bearing traditional and alternative exon 1 sequences joined to CFTR exon 2 and eGFP ORF (minus ATG). Positive control: eGFP (*ATG+*), control eGFP with translation initiation codon intact. Negative control: eGFP (*ATG−*), control eGFP with translation initiation codon removed. *CHO blank*, untransfected cells. # indicates a statistically significant ($p <$ 0.01) difference in protein expression from transfected cells *versus* untransfected cells (*CHO blank*). Variation in transfection efficiency was normalized by co-transfection with a control plasmid expressing blue fluorescent protein (Clontech).
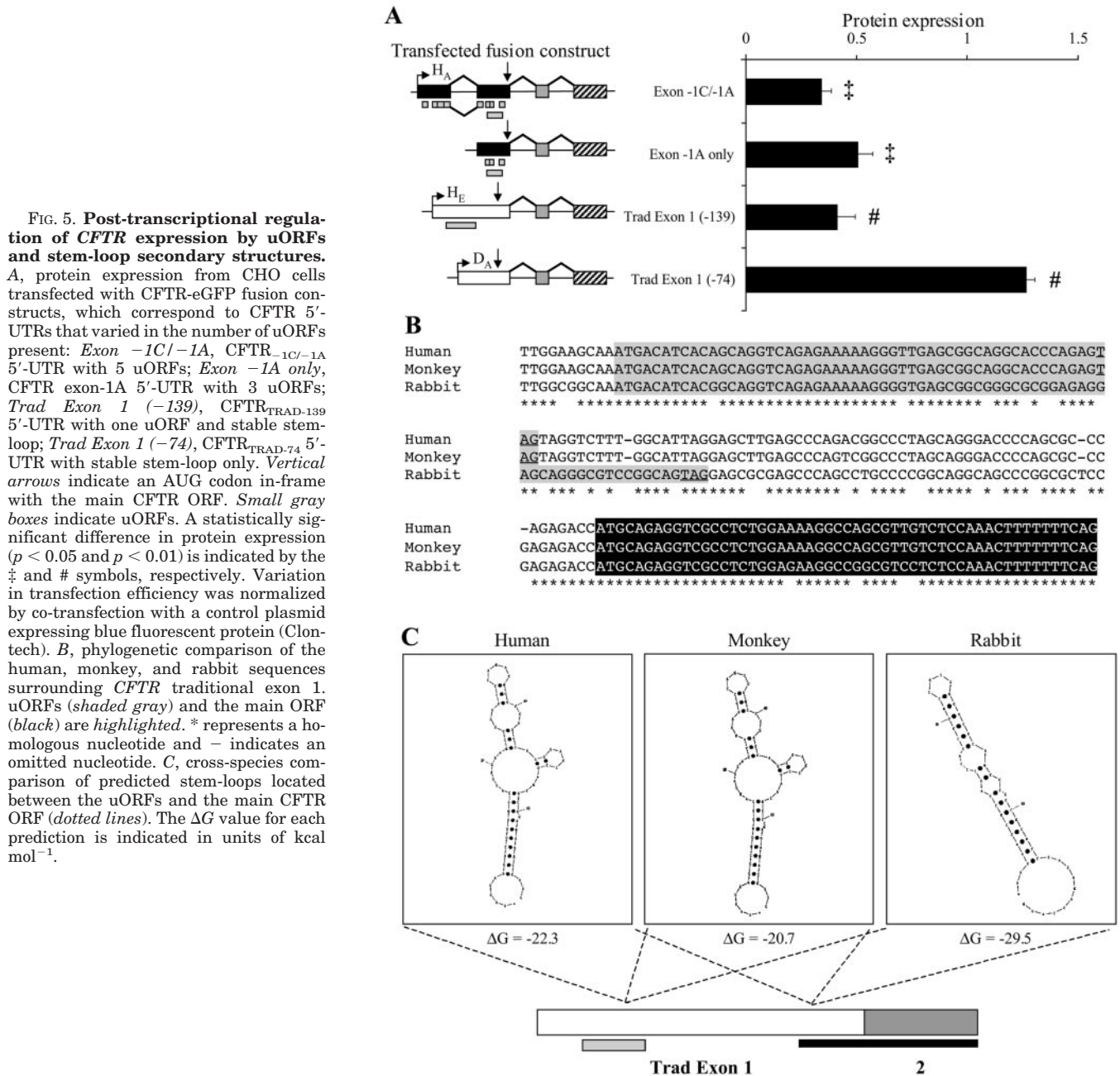
B
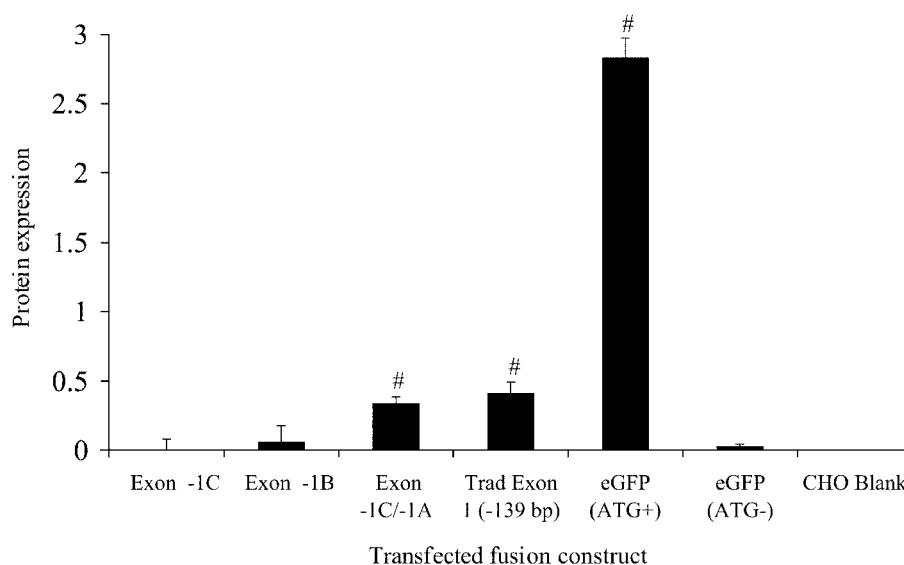
protein production. The removal of the only uORF in the 5′-UTR of traditional exon 1 resulted in a 3-fold increase in translation initiation at the AUG of the main CFTR ORF ($p <$ 0.01), but this was still lower than that observed for control eGFP transfections.

The involvement of secondary structure within the 5′-UTR of CFTR traditional exon 1 was investigated as a possible explanation for the still reduced translation efficiency of the CFTR$_{TRAD-74}$/eGFP construct, even in the absence of the uORF. A mRNA secondary structure prediction program (MFold) was used to analyze the 5′-UTR sequences for traditional exon 1, up to positions equivalent to −139 bp, from human (GenBank^{TM} accession no. AC000111), monkey (GenBank^{TM} accession no. X95930), and rabbit (GenBank^{TM} accession no. X95931). A conserved uORF was identified in the 5′-UTR of traditional CFTR exon 1 of all three species (Fig. 5*B*). Computational prediction of mRNA secondary structure formation identified classic stem-loop structures localized between the uORF termination codon and the main ORF start codon. In all species, the calculated stabilities (ΔG values) of the stem-loop secondary structures were in the range of −20 to −30 kcal mol^{-1} (Fig. 5*C*). Messenger RNA stem-loop structures localized to the 5′-UTR, with similar calculated stabilities, have been shown to independently reduce translation initiation at a downstream

AUG codon (49–51), but have a stronger inhibitory effect when localized immediately downstream of an uORF (43). Also, small differences in the sequence immediately surrounding the AUG codon, compared with the Kozak consensus, may contribute to the efficiency of translation.

Translation of CFTR$_{-1C/-1A}$ mRNA Results in CFTR Protein with a Unique Amino Terminus—As CFTR exon −1A splices directly to exon 2, the 17 amino acids encoded by traditional exon 1 are omitted during translation of CFTR$_{-1C/-1A}$ transcripts and 2 other amino acids, encoded by exon −1A, constitute the new CFTR amino terminus (Fig. 6*A*). The efficiency of translation of both the CFTR$_{TRAD-139}$ and CFTR$_{-1C/-1A}$ transcripts is indistinguishable (Fig. 4*B*). As CFTR$_{-1C/-1A}$ transcripts are ~6-fold more abundant in adult heart, than CFTR$_{TRAD-139}$ transcripts, it is likely that the majority of CFTR protein expressed in the heart is the curtailed amino terminus isoform reported here (CFTR$_{-1C/-1A}$ protein). We have used computational protein secondary structure and hydropathy analysis to predict possible differential functional roles of CFTR$_{-1C/-1A}$ protein.

Hydropathy analyses over the first 150 amino acids of both full-length and truncated CFTR isoforms were performed using the Kyte and Doolittle algorithm (40). The full-length CFTR amino terminus is predominantly hydrophilic, which is consist-

FIG. 5. **Post-transcriptional regulation of *CFTR* expression by uORFs and stem-loop secondary structures.** *A*, protein expression from CHO cells transfected with CFTR-eGFP fusion constructs, which correspond to CFTR 5′-UTRs that varied in the number of uORFs present: *Exon −1C/−1A*, CFTR$_{-1C/-1A}$ 5′-UTR with 5 uORFs; *Exon −1A only*, CFTR exon-1A 5′-UTR with 3 uORFs; *Trad Exon 1 (−139)*, CFTR$_{TRAD-139}$ 5′-UTR with one uORF and stable stem-loop; *Trad Exon 1 (−74)*, CFTR$_{TRAD-74}$ 5′-UTR with stable stem-loop only. *Vertical arrows* indicate an AUG codon in-frame with the main CFTR ORF. *Small gray boxes* indicate uORFs. A statistically significant difference in protein expression ($p < 0.05$ and $p < 0.01$) is indicated by the ‡ and # symbols, respectively. Variation in transfection efficiency was normalized by co-transfection with a control plasmid expressing blue fluorescent protein (Clontech). *B*, phylogenetic comparison of the human, monkey, and rabbit sequences surrounding *CFTR* traditional exon 1. uORFs (*shaded gray*) and the main ORF (*black*) are *highlighted*. * represents a homologous nucleotide and − indicates an omitted nucleotide. *C*, cross-species comparison of predicted stem-loops located between the uORFs and the main CFTR ORF (*dotted lines*). The ΔG value for each prediction is indicated in units of kcal mol^{-1}.

ent with previous work (3) and a predicted cytoplasmic localization. However, we report here the identification of a particularly hydrophobic region, corresponding to amino acids 11–26 of the amino-terminal end and encoded across the exon 1/exon 2 boundary. This amino-terminal hydrophobic region is completely absent in the curtailed CFTR$_{-1C/-1A}$ isoform (first described by Davies (Ref. 52); Fig. 6B). In contrast, the CFTR amino acid region involved in binding to syntaxin 1A (53, 54) was found to be entirely hydrophilic in nature (Fig. 6B, *shaded zone*), encoded by exons 2 and 3, and so still present in the amino-terminal truncated CFTR$_{-1C/-1A}$ isoform.

CFTR amino-terminal secondary structure was predicted by PSIPRED. Coinciding with the amino-terminal hydrophobic region described above (amino acids 11–26) is a putative helical structure (predicted with a high level of confidence) that spans the boundary between traditional exon 1 and exon 2 (Fig. 6C). Helical wheel analysis of this putative, hydrophobic helix predicts clustering of three positively charged residues (Fig. 6D). The removal of the amino-terminal half of this putative helix,

as would be the case in the CFTR$_{-1C/-1A}$ isoform, is accompanied by a marked reduction in the confidence of prediction of the remainder of the helix. Both secondary structure and helical wheel predictions were used by Naren and co-workers (54) in the identification of a helix that has been confirmed to interact with syntaxin 1A and regulate CFTR channel activity.

DISCUSSION

The distribution and control of *CFTR* gene expression is tightly regulated by temporal, spatial, and tissue-specific mechanisms (8–10). Multiple *in vivo* transcription start sites have been localized within 2 kb of the AUG translation initiation codon in traditional CFTR exon 1 (3, 11, 55). This is the first study showing differential CFTR *in vivo* transcription start site usage in the heart that is regulated by temporal, spatial, and pathophysiological stimuli. CFTR transcripts expressed in the heart initiate at three distinct exon 1 sequences (exon −1C, exon −1B, and traditional exon 1 initiating at −139 bp), each of which splice directly to exon 2, with exon −1A
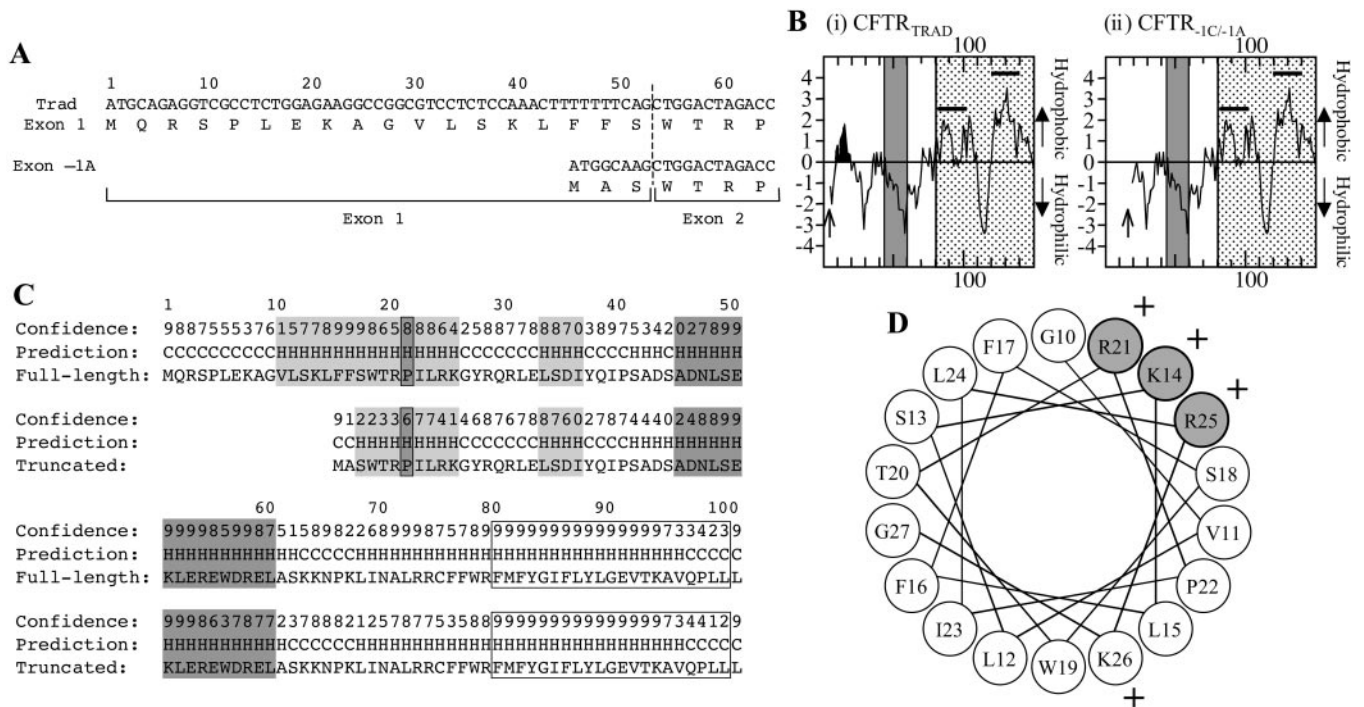
FIG. 6. **Characterization of CFTR protein isoforms encoded by CFTR$_{-1C/-1A}$ and CFTR$_{TRAD-139}$ transcripts expressed in rabbit heart.** *A*, comparison of CFTR amino-terminal polypeptide sequences encoded by CFTR$_{-1C/-1A}$ and CFTR$_{TRAD-139}$ transcripts. The polypeptide sequence is identical for both CFTR ORFs downstream of the exon 1 to exon 2 boundary (*vertical dotted line*). *B*, hydropathy plot analysis of the first 150 residues for (*i*) CFTR$_{TRAD}$ and (*ii*) CFTR$_{-1C/-1A}$ polypeptides. The first two transmembrane domains (*horizontal bar*) of the first membrane-spanning region are shown (*dotted*). Also, the location of the putative helix involved in the CFTR/syntaxin 1A interaction is indicated (*dark gray shaded region*) (54). The hydrophobic region in CFTR$_{TRAD}$ (*black*) is *highlighted*. *Open arrows* denote the translation initiation codons identified in CFTR$_{TRAD}$ and CFTR$_{-1C/-1A}$. *C*, comparison of protein secondary structure predictions for full-length and truncated CFTR amino termini. Regions of the CFTR protein are *shaded* to indicate: the predicted helical structures encoded within exon −1A and traditional exon 1 (*light gray*); a conserved proline residue (*dark gray, boxed*); a putative helix involved in the CFTR/syntaxin 1A interaction (*dark gray*) and transmembrane domains (*white, boxed*). Protein secondary structure (*C*, coil; *H*, helix) and confidence of predictions (*0*, low; *9*, high) are indicated. *D*, helical wheel plot of the amino-terminal region containing the putative helix identified in the CFTR$_{TRAD}$ protein. The cluster of positively charged residues present on one surface of the helix is indicated (*gray circle*).

alternatively spliced between exon −1C and exon 2. In the heart, the only previously identified alternative splicing of CFTR transcripts involved the differential exclusion of exon 5 (12). Exons −1C to −1A are themselves distributed over 2.5 kb of the rabbit genome and are located distal to the traditional *CFTR* exon 1.

Comparative phylogenetic analysis identified human homologues of both the rabbit exon −1C to −1A and traditional exon 1 *CFTR* regions, localized 10 kb upstream of human *CFTR* and at traditional exon 1, respectively. However, comparison with murine gDNA only revealed homology to traditional *CFTR* exon 1. These findings are consistent with species-specific, molecular, and functional distributions of *CFTR* expression in the heart, identified in humans (18) and rabbits (15, 25) but not in mice (23). Collectively, these findings suggest that *CFTR* transcription from exons −1C to −1A is cardiac-specific and controlled by a distinct, and previously unidentified, promoter. Additionally, this work provides support for the development of a rabbit model of cystic fibrosis, to allow investigation of some pathophysiological features of human CF disease that are not present in murine models, such as cardiac involvement in CF (56).

Spatial, developmental, and pathophysiological signals regulate *CFTR* expression in the heart, with the development of a left ventricular epicardial to endocardial gradient during the late fetal and neonatal periods, and loss of this gradient following hypertrophic stimuli (26–28). We have also identified an apical (higher) to basal (lower) gradient in *CFTR* expression that is unaffected by hypertrophy. This defines a radial pattern of *CFTR* expression across the left ventricle, correlating well

with the spread of repolarization throughout the left ventricle. These findings further support the view that the cAMP-stimulated chloride current, encoded by *CFTR*, contributes to ventricular repolarization and differential action potential duration throughout the heart. The expression of multiple CFTR mRNA transcripts in the heart raises questions of differential regulation of specific transcripts by various signals.

During cardiac development there is a preferential accumulation of CFTR$_{-1C}$, CFTR$_{-1C/-1A}$, and CFTR$_{EXON5-}$ transcripts. Concomitantly, there is a preferential loss of CFTR$_{-1B}$ transcripts. Additionally, the epicardial to endocardial gradient is primarily the result of CFTR$_{-1C}$, CFTR$_{-1C/-1A}$, and CFTR$_{EXON5-}$ transcripts, whereas all cardiac-specific CFTR transcripts contribute to the apical to basal ventricular *CFTR* gradient. In contrast, CFTR transcripts initiated at traditional exon 1 are present at very low, static levels (less than 10%). Analysis of left ventricular *CFTR* gradients in the hypertrophic heart demonstrated a loss of the epicardial to endocardial gradient but not the apical to basal gradient, culminating in an overall decrease in the radial gradient of ventricular *CFTR* expression. Again, individual CFTR transcripts were differentially regulated, with a preferential loss of CFTR$_{-1C}$, CFTR$_{-1C/-1A}$, and CFTR$_{EXON5-}$ transcripts, but no change in CFTR$_{-1B}$ and CFTR$_{TRAD-139}$ expression. The overall increase in *CFTR* expression during heart development, and loss in cardiac hypertrophy is consistent with the "re-expression of fetal gene program hypothesis" to explain global gene expression changes in cardiac hypertrophy (57, 58). However, our findings extend that hypothesis to include differential effects on individual CFTR transcripts. This demonstration of the

differential regulation of individual CFTR transcripts suggests further complexity in the regulation of *CFTR* expression, with mechanisms that allow transcript-specific interpretation of spatial, temporal, and pathologic signals.

We have identified *CFTR* regulatory mechanisms that involve both alternative splicing, and differential transcription start site and exon 1 usage. An important consequence of differential exon 1 usage is the generation of four CFTR transcripts with distinct 5′-UTRs, all of which encode one or more uORFs. There is increasing evidence that post-transcriptional regulation of gene expression, through modulation of mRNA stability and translation initiation, is achieved through 5′-UTR encoded elements, such as uORFs and stem-loop secondary structures (43, 48, 59). Although the *cis*-elements controlling post-transcriptional regulation of gene expression were first identified in yeast, similar mechanisms are now known in mammalian cells. To date over two-thirds of identified mammalian genes that encode uORFs in their 5′-UTRs are proto-oncogenes (60–63). However, uORFs have also been identified in a few genes with functions unrelated to cell growth control, the *S-adenosylmethionine decarboxylase* gene (64, 65) and *Huntingtin* gene (66). To these examples we now add the *CFTR* gene. Further, the causative mutations in the inherited diseases of familial melanoma (67, 68) and thrombocythemeia (69) create or abolish uAUGs that result in dramatic alterations in steady-state mRNA and protein levels.

It is estimated that less than 10% of eukaryotic mRNAs have an uORF in their 5′-UTR; however, very few have been investigated (48, 60). In general, uORFs lead to destabilization of the mRNA of the main ORF, secondary to disruption of ribosome scanning and reduced translation initiation at the main ORF (43). All CFTR transcripts expressed in the heart have the necessary 5′-UTR elements, uORFs with adjacent stem-loop structures, to allow post-transcriptional mechanisms to contribute to the regulation of *CFTR* expression. We have shown that the uORF, probably acting in concert with the adjacent stem-loop, encoded in $CFTR_{TRAD-139}$ transcripts functions to reduce translation initiation efficiency at the downstream AUG of the main CFTR ORF. In addition, the uORFs encoded in exon −1C have a similar effect. With reduced translation resulting in mRNA destabilization (43), it is probable that differential stability of CFTR transcripts in the heart may be a key factor governing spatial, temporal, and pathological changes in *CFTR* expression. Indeed, in cardiac tissues, different levels of *CFTR* expression were measured over exons 1–2 compared with exons 4–6 for all CFTR transcripts, suggesting that the 5′-UTR elements involved in modulating translation efficiency from the main AUG codon may also have a role in modulating differential stability along the CFTR transcript. However, the precise control of *in vivo* CFTR transcript stability remains unclear.

The discovery that uORFs in CFTR 5′-UTRs modulate CFTR protein production has implications beyond the expression of *CFTR* in the heart, as most mouse tissues express CFTR transcripts that include an uORF (11). Also, the CFTR transcripts expressed in human fetal lung, but not adult lung, include an uORF (11), allowing the possibility that post-transcriptional mechanisms may contribute to the large changes in *CFTR* expression that occur during lung development. We have identified a conserved stable stem-loop structure in the 5′-UTR of traditional CFTR exon 1 that would be included in the majority of CFTR transcripts expressed in all tissues. Similar stem-loop structures have been shown to alter translation efficiency (49, 50). Whether 5′-UTR encoded secondary structures lead to widespread post-transcriptional regulation of *CFTR* expression is under further investigation.

The cardiac potassium channels encoded by the *HERG* and *KvLQT1* genes are subject to alternative exon 1 usage, resulting in distinct amino-terminal protein isoforms (70). Similarly, we have shown that $CFTR_{-1C/-1A}$ transcripts contain an AUG codon that is in-frame with the main CFTR ORF and directs translation of an unique CFTR protein isoform, bearing a 15-amino acid truncation at the amino terminus. $CFTR_{-1B}$ and $CFTR_{-1C}$ transcripts do not encode an in-frame AUG; however, translation may initiate from downstream AUG codons in exons 3 and 4 (45). The substantially higher levels of $CFTR_{-1C/-1A}$ transcripts suggest that the majority of CFTR protein expressed in the heart is the truncated amino-terminal isoform.

The regulation of CFTR protein trafficking and channel activity has been shown to involve the amino terminus. Hydrophilic residues of the H3 domain of syntaxin 1A physically interact with a cytoplasmic amino-terminal helix of CFTR, encoded in exons 2 and 3 (71, 72). Furthermore, an exon 1-encoded diphenylalanine motif (Phe-16/Phe-17) regulates CFTR membrane trafficking (73). It has been suggested that loss of CFTR exon 1 would remove both motifs (55). However, although the diphenylalanine motif would be removed, the CFTR domain that interacts with syntaxin 1A would remain as it is encoded within exons 2 and 3. It was also suggested that the CFTR domain that interacts with syntaxin 1A is hydrophobic in nature and encoded at the exon 1/2 boundary (55). However, hydrophobicity analysis demonstrates that the syntaxin 1A interacting domain is in fact hydrophilic (Fig. 6*B*) (52). Our analysis does indeed identify a CFTR amino-terminal hydrophobic region (residues 11–26) encoded at the exon 1/2 boundary, first described by Davies (52). This hydrophobic region, truncated by loss of exon 1, has a predicted helical structure and may affect CFTR function or subcellular location by its presence or absence. The putative helical structure of the CFTR amino terminus (residues 11–26) is supported by the recent identification, from the crystal structure (4.5 Å), of an α helix of similar length and location within the amino terminus (residues 10–21) of the ABC transporter homologue, MsbA, from *Escherichia coli* (74).

We, and others, have established an absolute correlation between CFTR mRNA expression, and the species-specific distribution of cardiac cAMP-activated chloride currents (14), present in rabbits (15, 26), guinea pigs (20, 21), and monkeys and humans (18). Antisense oligodeoxynucleotide inhibition studies have confirmed *CFTR* as the molecular basis of the *in vivo* cAMP-activated chloride current present in epithelial cells and cardiomyocytes. Primary cultures of sweat duct epithelial cells (75) and ventricular cardiomyocytes (25), as well as pancreatic duct (76) and colonic and tracheal epithelial cell lines (77), were treated with antisense oligodeoxynucleotides that bind up to 23 nucleotides surrounding the traditional CFTR exon 1 translation initiation codon, present only in the $CFTR_{TRAD-74}$ and $CFTR_{TRAD-139}$ transcripts. This resulted in maximal inhibition (greater than 90%) of the endogenous cAMP-activated chloride currents present in all cells of epithelial origin. In contrast, there was only a partial inhibition (∼40%) of the endogenous cAMP-activated chloride currents in ventricular cardiomyocytes, despite doubling the concentration of antisense oligodeoxynucleotide. In all cases, sense oligodeoxynucleotides had no effect on the cAMP-activated chloride current and neither sense nor antisense oligodeoxynucleotides had any effect on calcium-activated chloride conductance. Our finding, that $CFTR_{TRAD-139}$ mRNA accounts for only a small proportion of the CFTR protein-coding transcripts present in the heart, provides an explanation for the partial inhibition of the cAMP-activated chloride current in ventricular myocytes described by Hart *et al.* (25). The cardiac-specific $CFTR_{-1C/-1A}$

transcript we have identified will not bind the antisense oligodeoxynucleotides used by Hart *et al.* (25), and their findings support our conclusion that CFTR$_{-1C/-1A}$ mRNA codes the majority of CFTR protein and cAMP-activated chloride current present in the heart. Further, it is specifically the differential expression of the CFTR$_{-1C/-1A}$ transcript that is absolutely correlated with differences in the whole cell current density of cAMP-activated chloride currents in epicardial *versus* endocardial ventricular myocytes (26), and the loss of cAMP-activated chloride conductance in cardiac hypertrophy (78).

Overall, this work identifies multiple new levels at which *CFTR* expression is regulated *in vivo*. This is the first study to show the *in vivo* post-transcriptional regulation of *CFTR* expression through modulation of translation initiation efficiency by 5′-UTR encoded elements. Furthermore, this is probably a widespread mechanism regulating *CFTR* expression. We are also the first to show that, through alternative exon 1 usage, a unique isoform of CFTR protein is generated and is likely to be the major CFTR isoform present in the heart. This novel curtailed form of CFTR protein, CFTR$_{-1C/-1A}$, is missing functionally important amino-terminal motifs and is predicted to display unique subcellular localization and activity.
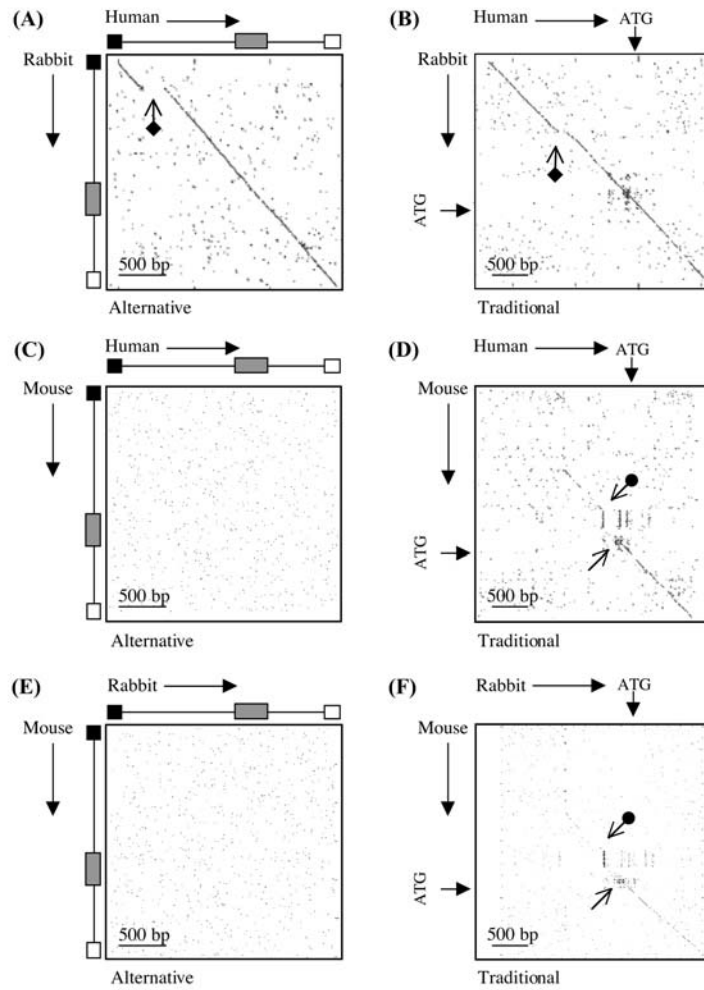
REFERENCES

1. Welsh, M., Tsui, L., Boat, T., and Beaudet, A. (1995) *The Metabolic Basis of Inherited Diseases*, McGraw-Hill, New York
2. Kerem, B., Rommens, J., Buchanan, J., Markiewicz, D., Cox, T., Chakravarti, A., Buchwald, M., and Tsui, L. (1989) *Science* **245,** 1073–1080
3. Riordan, J., Rommens, J., Kerem, B., Alon, N., Rozahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J., Drumm, M., Iannuzzi, M., Collins, F., and Tsui, L. (1989) *Science* **245,** 1066–1073
4. Rommens, J., Iannuzzi, M., Kerem, B., Drumm, M., Melmer, G., Dean, M., Rozmahel, R., Cole, J., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Riordan, J., Tsui, L., and Collins, F. (1989) *Science* **245,** 1059–1065
5. Anderson, M., Gregory, R., Thompson, S., Souza, D., Paul, S., Mulligan, R., Smith, A., and Welsh, M. (1991) *Science* **251,** 202–205
6. Bear, C., Kartner, N., Bridges, J., Jensen, T., Ramjeesingh, M., and Riordan, J. (1992) *Cell* **68,** 809–818
7. Hyde, S., Emsley, P., Hartshorn, M., Mimmack, M., Gileadi, U., Pearce, S., Gallagher, M., Gill, D., Hubbard, R., and Higgins, C. (1990) *Nature* **346,** 362–365
8. Trezise, A., and Buchwald, M. (1991) *Nature* **353,** 434–437
9. Trezise, A., Romano, P., Gill, D., Hyde, S., Sepulveda, F., Buchwald, M., and Higgins, C. (1992) *EMBO J.* **11,** 4291–4303
10. Trezise, A., Chambers, J., Wardle, C., Goud, S., and Harris, A. (1993) *Hum. Mol. Genet.* **2,** 213–218
11. White, N., Higgins, C., and Trezise, A. (1998) *Hum. Mol. Genet.* **3,** 363–369
12. Horowitz, B., Tsung, S., Hart, P., Levesque, P., and Hume, J. (1993) *Am. J. Physiol.* **264,** H2214–H2220
13. Trezise, A., Buchwald, M., and Higgins, C. (1993) *Hum. Mol. Genet.* **2,** 801–802
14. Gadsby, D., Nagel, G., and Hwang, T. (1995) *Annu. Rev. Physiol.* **57,** 387–416
15. Levesque, P., Hart, P., Hume, J., Kenyon, J., and Horowitz, B. (1992) *Circ. Res.* **71,** 1002–1007
16. Mulberg, A., Wiedner, E., Bao, X., Marshall, J., Jefferson, D., and Altschuler, S. (1994) *Neuroreport* **5,** 1684–1688
17. Weyler, R., Yurko-Mauro, K., Rubenstein, R., Kollen, W., Reenstra, W., Altschuler, S., Egan, M., and Mulberg, A. (1999) *Am. J. Physiol.* **277,** C563–C571
18. Warth, J., Collier, M., Hart, P., Geary, Y., Gelband, C., Chapman, T., Horowitz, B., and Hume, J. (1996) *Cardiovasc. Res.* **31,** 615–624
19. Takano, M., and Noma, A. (1992) *Pflugers Arch.* **420,** 223–226
20. Bahinski, A., Nairn, A., Greengard, P., and Gadsby, D. (1989) *Nature* **340,** 718–721
21. James, F., Tominaga, T., Okada, Y., and Tominaga, M. (1996) *Circ. Res.* **79,** 201–207
22. Zhang, K., Barrington, P., Martin, R., and Eick, R. (1994) *Circ. Res.* **75,** 133–143
23. Levesque, P., and Hume, J. (1995) *Cardiovasc Res.* **29,** 336–343
24. Sorota, S., Siegal, M., and Hoffman, B. (1991) *J. Mol. Cell. Cardiol.* **23,** 1191–1198
25. Hart, P., Warth, J., Levesque, P., Collier, M., Geary, Y., Horowitz, B., and Hume, J. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 6343–6348
26. Wong, K., Trezise, A., Bryant, S., Hart, G., and Vandenberg, J. (1999) *Am. J. Physiol.* **277,** H1403–H1409
27. Wong, K., Trezise, A., and Vandenberg, J. (2000) *Mech. Dev.* **94,** 195–197
28. Wong, K., Trezise, A., Crozatier, B., and Vandenberg, J. (2000) *Biochem. Biophys. Res. Commun.* **278,** 144–149
29. Hart, G. (1994) *Cardiovasc. Res.* **28,** 933–946
30. Antzelevitch, C., Sicouri, S., Lukas, A., Nesterenko, V., Liu, D., and DiDiego, J. (1995) *Cardiac Electrophysiology: From Cell to Bedside*, W. B. Saunders, New York
31. Marban, E. (1999) *Molecular Basis of Cardiovascular Disease*, W. B. Saunders, New York
32. Cheron, G., Paradis, K., Steru, D., Demay, G., and Lenoir, G. (1984) *Acta Paediatr. Scand.* **73,** 697–700
33. Sullivan, M., Moss, R., Hindi, R., and Lewiston, N. (1986) *Chest* **90,** 239–242
34. Guarnieri, C., Muscari, C., and Caldarera, C. (1985) *Adv. Myocardiol.* **5,** 191–199
35. Chomczynski, P., and Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159
36. Lydford, S., and Giller, T. (2001) *MAXline Application Note 44; Measurement of Green Fluorescent Protein in the SPECTRAmax GEMINI Spectrofluorometer*, Molecular Devices, United Kingdom
37. McGown, E. (1999) *MAXline Application Note 30*; *Selecting Excitation and Emission Wavelengths Using the SPECTRAmax GEMINI Microplate Spectrofluorometer: Basic Principles*, Molecular Devices, United Kingdom
38. Mathews, D., Sabina, J., Zuker, M., and Turner, D. (1999) *J. Mol. Biol.* **288,** 911–940
39. Zuker, M., Mathews, D., and Turner, D. (1999) *RNA Biochemistry and Biotechnology,* Kluwer Academic Publishers, Dordrecht, Netherlands
40. Kyte, J., and Doolittle, R. (1982) *J. Mol. Biol.* **157,** 105–132
41. Jones, D. (1999) *J. Mol. Biol.* **292,** 195–202
42. Mount, S. (1982) *Nucleic Acids Res.* **14,** 1319–1324
43. Vilela, C., Ramirez, C., Linz, B., Rodrigues-Pousada, C., and McCarthy, J. (1999) *EMBO J.* **18,** 3139–3152
44. Tomaselli, G., Beuckelmann, D., Calkins, H., Berger, R., Kessler, P., Lawrence, J., Kass, D., Feldman, A., and Marban, E. (1994) *Circulation* **90,** 2534–2539
45. Carroll, T., Morales, M., Fulmer, S., Allen, S., Flotte, T., Cutting, G., and Guggino W. (1995) *J. Biol. Chem.* **270,** 11941–11946
46. Kozak, M. (1991) *J. Cell Biol.* **115,** 887–903
47. Kozak, M. (2000) *Genomics* **70,** 396–406
48. Pesole, G., Mignone, F., Gissi, C., Grillo, G., Licciulli, F., and Liuni, S. (2001) *Gene* (*Amst.*) **276,** 73–81
49. Kozak, M. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 2850–2854
50. Kozak, M. (1989) *Mol. Cell. Biol.* **9,** 5134–5142
51. Vega Laso, M., Zhu, D., Sagliocco, F., Brown, A., Tuite, M., and McCarthy, J. (1993) *J. Biol. Chem.* **268,** 6453–6462
52. Davies, W. (2002) *Molecular and Functional Investigation of the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) in Rabbit Heart.* Ph.D. thesis, University of Cambridge, Cambridge, United Kingdom
53. Naren, A., Quick, M., Collawn, J., Nelson, D., and Kirk, K. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 10972–10977
54. Naren, A., Cormet-Boyaka, E., Fu, J., Villain, M., Blalock, J., Quick, M., and Kirk, K. (1999) *Science* **286,** 544–548
55. Mouchel, N., Broackes-Carter, F., and Harris, A. (2003) *Hum. Mol. Genet.* **12,** 759–769
56. Vuillaumier, S., Kaltenboeck, B., Lecointre, G., Lehn, P., and Denamur, E. (1997) *Mol. Biol. Evol.* **14,** 372–380
57. Chien, K., Knowlton, K., Zhu, H., and Chien, S. (1991) *FASEB J.* **5,** 3037–3046
58. Chien, K., Zhu, H., Knowlton, K., Miller-Hance, W., Van-Bilsen, M., O'Brien, T., and Evans, S. (1993) *Annu. Rev. Physiol.* **55,** 77–95
59. Guhaniyogi, J., and Brewer, G. (2001) *Gene* (*Amst.*) **265,** 11–23
60. Kozak, M. (1987) *Nucleic Acids Res.* **15,** 8125–8148
61. Wang, X., and Rothnagel, J. (2001) *J. Biol. Chem.* **276,** 1311–1316
62. Child, S., Miller, M., and Geballe, A. (1999) *J. Biol. Chem.* **274,** 24335–24341
63. Jin, X., Turcott, E., Englehardt, S., Mize, G., and Morris, D. (2003) *J. Biol. Chem.* **278,** 25716–25721
64. Hill, J., and Morris, D. (1993) *J. Biol. Chem.* **268,** 726–731
65. Mize, G., Ruan, H., Low, J., and Morris, D. (1998) *J. Biol. Chem.* **273,** 32500–32505
66. Lee, J., Park, E., Couture, G., Harvey, I., Garneau, P., and Pelletier, J. (2002) *Nucleic Acids Res.* **30,** 5110–5119
67. Liu, L., Dilworth, D., Gao, L., Monzon, J., Summers, A., Lassam, N., and Hogg, D. (1999) *Nat. Genet.* **21,** 128–132
68. Mendell, J., and Dietz, H. (2001) *Cell* **107,** 411–414
69. Wiestner, A., Schlemper, R., van der Maas, A., and Skoda, R. (1998) *Nat. Genet.* **18,** 49–52
70. Roden, D., Balser, J., George, A., and Anderson, M. (2002) *Annu. Rev. Physiol.* **64,** 431–475
71. Ganeshan, R., Di, A., Nelson, D., Quick, M., and Kirk, K. (2003) *J. Biol. Chem.* **278,** 2876–2885
72. Naren, A., and Kirk, K. (2000) *News Physiol. Sci.* **15,** 57–61
73. Olenych, S., and Teem, J. (1999) *Pediatr. Pulmonol. Suppl.* **19,** 185
74. Chang, G., and Roth, C. (2001) *Science* **293,** 1793–1800
75. Sorscher, E., Kirk, K., Weaver, M., Jilling, T., Blalock, J., and LeBoeuf, R. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 7759–7762
76. Kopelman, H., Gauthier, C., and Bornstein, M. (1993) *J. Clin. Invest.* **91,** 1253–1257
77. Wagner, J., McDonald, T., Nghiem, P., Lowe, A., Schulman, H., Gruenert, D., Stryer, L., and Gardner, P. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 6785–6789
78. Wallis, W., Cooklin, M., Sheridan, D., and Fry, C. (2001) *Arch. Physiol. Biochem.* **109,** 117–126

Legend for Supplemental Figure.

Cross-species analysis of *CFTR* in the human, rabbit and murine genomes. Dot plot analyses showing identity between human, rabbit and murine (GenBank: AC000111 and AF162137) genomic sequences, over regions spanning both the alternative and traditional *CFTR* exon 1 loci. The long closed black arrows represent a 5' to 3' direction. In **(A)**, **(C)** and **(E)**, the alternatively spliced exon 1 sequences are coded: exon –1C, black; exon –1B, grey; exon –1A, white. The traditional CFTR ATG codon is indicated in **(B)**, **(D)** and **(F)**, by a short closed black arrow. Important structural elements are represented by open arrows, as follows: GC rich region, no tail; Pu.Py stretch, circular tail; and an indel unit, diamond-shaped tail. Horizontal bars represent a scale of 500 bp.

(A) Human → / Rabbit ↓ / Alternative / 500 bp

(B) Human → ATG / Rabbit ↓ / ATG → / Traditional / 500 bp

(C) Human → / Mouse ↓ / Alternative / 500 bp

(D) Human → ATG / Mouse ↓ / ATG → / Traditional / 500 bp

(E) Rabbit → / Mouse ↓ / Alternative / 500 bp

(F) Rabbit → ATG / Mouse ↓ / ATG → / Traditional / 500 bp

SUPPLEMENTAL FIGURE

Cardiac Expression of the Cystic Fibrosis Transmembrane Conductance Regulator Involves Novel Exon 1 Usage to Produce a Unique Amino-terminal Protein
Wayne L. Davies, Jamie I. Vandenberg, Rana A. Sayeed and Ann E. O. Trezise

Access the most updated version of this article at doi: 10.1074/jbc.M313628200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 71 references, 34 of which can be accessed free at
http://www.jbc.org/content/279/16/15877.full.html#ref-list-1