

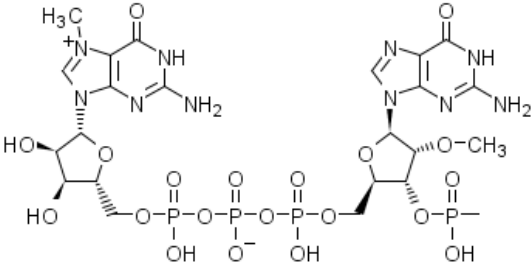


11392

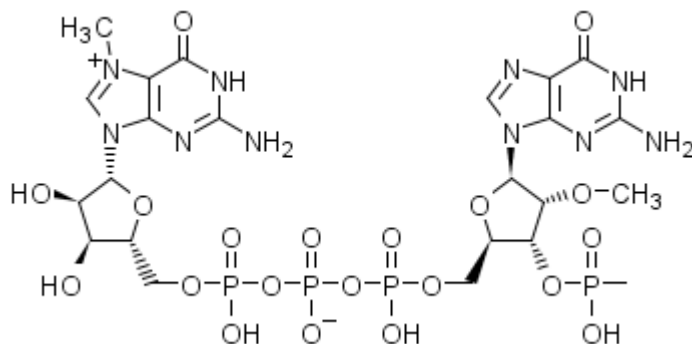
### Schematic



### Table of features

Structural Element	Description	Position in Sequence
Cap Structure		1-2
5' UTR	<u>GGAC...CACG</u>	3-140
Start Codon hCFTR	<b>AUG</b>	141-143
Stop Codon hCFTR	<b>UAA</b>	4581-4583
3' UTR	<u>CGGG...AGCU</u>	4584-4688
PolyA tail	(A) <sub>x</sub> , x=200-1000*	4689-ff

**5'- capping structure = G<sup>1</sup>G<sup>2</sup>: m<sup>7</sup>G<sup>+</sup>-(5'→5')-ppp-Gm-(3'→5')**





**1-2** : cap m<sup>7</sup>G<sup>+</sup>-(5'→5')-ppp-Gm; **3-140** : 5'UTR; **141-143** : start codon; 144-4580 : CFTR; **4581-4583** : stop codon; 4584-4688 : 5'UTR; 4689-... : polyA tail (A)<sub>n</sub> (n = 200-1000).

**Sequence / Séquence / Secuencia**

<u>GGACAGAU</u> CG	<u>CCUGGAG</u> ACG	<u>CCAUCCAC</u> GC	<u>UGUUUUG</u> ACC	<u>UCCAUA</u> GAAAG	50
<u>ACACCGGG</u> AC	<u>CGAUCCAG</u> CC	<u>UCCGCGG</u> CCG	<u>GGAACGG</u> UGC	<u>AUUGGA</u> ACGC	100
<u>GGAUUCC</u> CCG	<u>UGCCAAG</u> AGU	<u>GACUCAC</u> CGU	<u>CCUUGAC</u> ACG	<b>AUG</b> CAACGCU	150
CUCCUCUUGA	AAAGGCCUCG	GUGGUGUCCA	AGCUCUUCUU	CUCGUGGACU	200
AGACCCAUC	UGAGAAAGG	GUACAGACAG	CGCUUGGAGC	UGUCCGAU	250
CUAUCAAAUC	CCUUCGUGG	ACUCCGCGGA	CAACCUGUCC	GAGAAGCUCG	300
AGAGAGAAUG	GGACAGAGAA	CUCGCCUCA	AGAAGAACC	GAAGCUGAU	350
AAUGCGCUUA	GGCGGUGCU	UUUCUGGCGG	UUCAUGUUCU	ACGGCAUCUU	400
CCUCUACCG	GGAGAGGUC	CCAAGGCCGU	GCAGCCCCUG	UUGCUGGGAC	450
GGAUUAUUG	CUCCUACGAC	CCCACAACA	AGGAAGAAAG	AAGCAUCGCU	500
AUCUACUUG	GCAUCGGUCU	GUGCCUGCU	UUCAUCGUCC	GGACCCUCU	550
GUUGCAUCC	GCUAUUUUCG	GCCUGCAUCA	CAUUGGCAUG	CAGAUGAGAA	600
UUGCCAUGUU	UUCCUGAUC	UACAAGAAA	CUCUGAAGCU	CUCGAGCCGC	650
GUGCUUGACA	AGAUUCCA	CGGCCAGCUC	GUGUCCUGC	UCUCCAACAA	700
UCUGAACAA	UUCGACGAGG	GCCUCGCCU	GGCCACUUC	GUGUGGAUCG	750
CCCCUCUGCA	AGUGGCGCU	CUGAUGGGCC	UGAUCUGGGA	GCUGCUGCAA	800
GCCUCGGCAU	UCUGUGGGCU	UGGAUUCUG	AUCGUGCUGG	CACUGUCCA	850
GGCCGGACUG	GGGCGGAUGA	UGAUGAAGUA	CAGGGACCAG	AGAGCCGGAA	900
AGAUUCCGA	ACGGCUGGUG	AUCACUUCGG	AAAUGAUCGA	AAACAUCAG	950
UCAGUGAAGG	CCUACUGCUG	GGAAGAGGCC	AUGGAAAAGA	UGAUUGAAAA	1000
CCUCGGCAA	ACCGAGCUGA	AGCUGACCCG	CAAGGCCGCU	UACGUGCGCU	1050
AUUUCAACUC	GUCCGCUUUC	UUUCUUC	GGUUCUUCGU	GGUGUUUCUC	1100
UCCGUGCUCC	CCUACGCCU	GAUUAAGGGA	AUCAUCCUCA	GGAAGAUUCU	1150
CACCACCAU	UCCUUCUGUA	UCGUGCUCCG	CAUGGCCGUG	ACCCGGCAGU	1200
UCCCAUGGGC	CGUGCAGACU	UGGUACGACU	CCCUGGGAGC	CAUUAACAAG	1250
AUCCAGGACU	UCCUUCAAAA	GCAGGAGUAC	AAGACCCUCG	AGUACAACCU	1300
GACUACUACC	GAGGUCGUGA	UGGAAAACGU	CACCGCCUUU	UGGGAGGAGG	1350
GAUUUGGCGA	ACUGUUCGAG	AAGGCCAAGC	AGAACAACAA	CAACCGCAAG	1400
ACCUCGAACG	GUGACGACUC	CCUCUUCUUU	UCAACUUC	GCCUGCUCGG	1450
GACGCCGUG	CUGAAGGACA	UUAACUUC	GAUCGAAAGA	GGACAGCUCC	1500
UGGCGGUGGC	CGGAUCGACC	GGAGCCGGAA	AGACUCCCU	GCUGAUGGUG	1550
AUCAUGGGAG	AGCUUGAACC	UAGCGAGGGA	AAGAUC	ACUCCGGCCG	1600
CAUCAGCUUC	UGUAGCCAGU	UUUCCUGGAU	CAUGCCC	ACCAUUAAGG	1650
AAAACAUCAU	CUUCGGCGUG	UCCUACGAUG	AAUACCGCUA	CCGGUCCGUG	1700
AUCAAAAGCCU	GCCAGCUGGA	AGAGGAU	UCAAGUUCG	CGGAGAAAGA	1750
UAACAUCGUG	CUGGGCGAAG	GGGUUAU	CUUGUCGGG	GGCCAGCGG	1800
CUAGAAUCUC	GCUGGCCAGA	GCCGUGUAU	AGGACGCCGA	CCUGUAUCUC	1850
CUGGACUCCC	CCUUCGGAUA	CCUGGACGUC	CUGACCGAAA	AGGAGAUUCU	1900
CGAAUCGUGC	GUGUGCAAGC	UGAUGGCUAA	CAAGACUCGC	AUCCUCGUGA	1950
CCUCCAAAAU	GGAGCACCUG	AAGAAGGCAG	ACAAGAUUCU	GAUUCUGCAU	2000
GAGGGGUCCU	CCUACUUUUA	CGGCACCUUC	UCGGAGUUC	AGAACUUGCA	2050
GCCCCGACUUC	UCAUCGAAGC	UGAUGGGUUG	CGACAGCUUC	GACCAGUUCU	2100
CCGCCGAAAG	AAGGAACUCG	AUCCUGACGG	AAACCUUGCA	CCGCUUCUCU	2150



UUGGAAGGCG	ACGCCCCUGU	GUCAUGGACC	GAGACUAAGA	AGCAGAGCUU	2200
CAAGCAGACC	GGGGAAUUCG	GCGAAAAGAG	GAAGAACAGC	AUCUUGAACC	2250
CCAUUAACUC	CAUCCGCAAG	UUCUCAAUUCG	UGCAAAAAGAC	GCCACUGCAG	2300
AUGAACGGCA	UUGAGGAGGA	CUCCGACGAA	CCCCUUGAGA	GGCGCCUGUC	2350
CCUGGUGCCG	GACAGCGAGC	AGGGAGAAGC	CAUCCUGCCU	CGGAUUUCCG	2400
UGAUCUCCAC	UGGUCCGACG	CUCCAAGCCC	GGCGGCGGCA	GUCCGUGCUG	2450
AACCUGAUGA	CCCACAGCGU	GAACCAGGGC	CAAAACAUIUC	ACCGCAAGAC	2500
UACCGCAUCC	ACCCGGAAAG	UGUCCUGGC	ACCUCAAGCG	AAUCUUAACG	2550
AGCUCGACAU	CUACUCCCGG	AGACUGUCGC	AGGAAACCGG	GCUCGAAAUU	2600
UCCGAAGAAA	UCAACGAGGA	GGAUCUGAAA	GAGUGCUUCU	UCGACGAUUA	2650
GGAGUCGAUA	CCCGCCGUGA	CGACUUGGAA	CACUUAUCUG	CGGUACAUCA	2700
CUGUGCACAA	GUCAUUGAUC	UUCGUGCUGA	JUUGGUGCCU	GGUGAUUUUC	2750
CUGGCCGAGG	UCGCGGCCUC	ACUGGUGGUG	CUCUGGCUGU	UGGGAAACAC	2800
GCCUCUGCAA	GACAAGGGAA	ACUCCACGCA	CUCGAGAAAC	AACAGCUAUG	2850
CCGUGAUUAU	CACUUCACC	UCCUCUUAU	ACGUGUUCUA	CAUCUACGUC	2900
GGAGUGGCGG	AUACCCUGCU	CGCGAUGGGU	UUCUUCAGAG	GACUGCCGCU	2950
GGUCCACACC	UUGAUCACCG	UCAGCAAGAU	UCUUCACCAC	AAGAUGUJGC	3000
AUAGCGUGCU	GCAGGCCCCC	AUGUCCACCC	UCAACACUCU	GAAGGCCGGA	3050
GGCAUUCUGA	ACAGAUUCUC	CAAGGACAUC	GCUAUCCUGG	ACGAUCUCCU	3100
GCCGCUUACC	AUCUUUGACU	UCAUCCAGCU	GCUGCUGAUC	GUGAUUGGAG	3150
CAAUCGCAGU	GGUGGCGGUG	CUGCAGCCUU	ACAUUUUCGU	GGCCACUGUG	3200
CCGGUCAUUG	UGGCGUUCAU	CAUGCUGCGG	GCCUACUJCC	UCCAAACCAG	3250
CCAGCAGCUG	AAGCAACUGG	AAUCCGAGGG	ACGAUCCCCC	AUCUUCACUC	3300
ACCUUGUGAC	GUCGUUGAAG	GGACUGUGGA	CCUCCGGGGC	UUUCGGACGG	3350
CAGCCCUACU	UCGAAACCCU	CUUCCACAAG	GCCCUGAACC	UCCACACCGC	3400
CAAUJGGUUC	CUGUACCUGU	CCACCCUGCG	GUGGUUCCAG	AUGCGCAUCG	3450
AGAUGAUUUU	CGUCAUCUUC	UUCAUCGCGG	UCACAUUCAU	CAGCAUCCUG	3500
ACUACCGGAG	AGGGAGAGGG	ACGGGUCGGA	AUAAUCCUGA	CCCUCGCCAU	3550
GAACAUUAUG	AGCACCCUGC	AGUGGGCAGU	GAACAGCUCG	AUCGACGUGG	3600
ACAGCCUGAU	GCGAAGCGUC	AGCCGCGUGU	UCAAGUUCAU	CGACAUGCCU	3650
ACUGAGGGAA	AACCCACUAA	GUCCACUAAG	CCCUACAAAA	AUGGCCAGCU	3700
GAGCAAGGUC	AUGAUCAUCG	AAAACUCCCA	CGUGAAGAAG	GACGAUUAUU	3750
GGCCUCCGG	AGGUCAAUUG	ACCGUGAAGG	ACCUGACCGC	AAAGUACACC	3800
GAGGGAGGAA	ACGCCAUUCU	CGAAAACAUC	AGCUUCUCCA	UUUCGCCGGG	3850
ACAGCGGGUC	GGCCUUCUCG	GGCGGACCGG	UUCGCGGAAG	UCAACUCUGC	3900
UGUCGGCUUU	CCUCCGGCUG	CUGAAUACCG	AGGGGGAAAU	CCAAAUUGAC	3950
GGCGUGUCUU	GGGAUCCAUC	UACUCUGCAG	CAGUGGCGGA	AGGCCUUCGG	4000
CGUGAUCCCC	CAGAAGGUGU	UCAUCUUCUC	GGGUACCUUC	CGGAAGAACC	4050
UGGAUCCUUA	CGAGCAGUGG	AGCGACCAAG	AAUUCUGGAA	GGUCGCCGAC	4100
GAGGUCGGCC	UGCGUCUCCG	GAUUGAACAA	JUUCUUGGAA	AGCUGGACUU	4150
CGUGCUCGUC	GACGGGGGAU	GUGUCCUGUC	GCACGGACAU	AAGCAGCUCA	4200
UGUGCCUCGC	ACGGUCCGUG	CUCUCCAAGG	CCAAGAUUCU	GCUGCUGGAC	4250
GAACCUUCGG	CCCACCUUGA	UCCGGUCACC	UACCAGAUCA	UCAGGAGGAC	4300
CCUGAAGCAG	GCCUUUGCCG	AUJGCACCGU	GAUUCUCUGC	GAGCACCGCA	4350
UCGAGGCCAU	GCUGGAGUGC	CAGCAGUJCC	UGGUAUCGA	GGAGAACAAG	4400
GUCCGCCAAU	ACGACUCCAUC	UCAAAAAGCUC	CUCAACGAGC	GGUCGCUGUU	4450
CAGACAAGCU	AUUUCACCGU	CCGAUAGAGU	GAAGCUCUUC	CCGCAUCGGA	4500
ACAGCUCAAA	GUGCAAUUCG	AAGCCGCAGA	UCGCAGCCUU	GAAGGAAGAG	4550
ACUGAGGAAG	AGGUGCAGGA	CACCCGGCUU	<b>UAACGGGUGG</b>	<u>CAUCCUGUG</u>	4600



**WHO**  
**International Nonproprietary Names Programme**

7/2020

ACCCUCCCC AGUGCCUCUC CUGGCCUGG AAGUUGCCAC UCCAGUGCCC 4650  
ACCAGCCUUG UCCUAAUAAA AUUAAGUUGC AUCAAGCU--polyA tail 4688