



## Par øvelse om Cystisk Fibrose og Bioinformatik

Bioinformatik er analyse af større mængder biologiske data, der er så uoverskuelige, at de kræver beregninger ved hjælp af computere for at give mening. Indenfor biomedicinsk forskning er bioinformatik blevet et vigtigt arbejdsredskab i forståelsen af, hvorfor sygdomme opstår og hvorfor at den medicin man udvikler virker på nogle patienter, men ikke på andre. I skal i dag bruge bioinformatik til at undersøge en sygdom og forstå, hvorfor at alle, der har sygdommen, ikke bliver syge i samme grad.

En af de sygdomme, man har undersøgt ved hjælp af bioinformatiske metoder, er cystisk fibrose. Cystisk fibrose er en af de hyppigst forekommende arvelige sygdomme i Vesteuropa og i Danmark er der ca. 500 personer, der lever med sygdommen. I denne øvelse skal I undersøge de genetiske årsager til, at man får cystisk fibrose.

### **Aktivitet 1: Matrix gruppe arbejde - Hvad er komplikationerne ved cystisk fibrose?**

I skal først undersøge, hvilke komplikationer der er for personer, der har cystisk fibrose. Alle par bliver delt i to grupper.

Gruppe 1 par: I skal undersøge, hvilke lunge komplikationer cystisk fibrose giver.

Gruppe 2 par: I skal undersøge, hvilke fordøjelseskomplicationer cystisk fibrose giver.

Gå ind på <https://cystiskfibrose.dk/cystisk-fibrose/hvad-er-cystisk-fibrose/> og lær mere om sygdommen. Lad dog være med at se videoen om CFTR genet, da I alle skal undersøge genet senere.

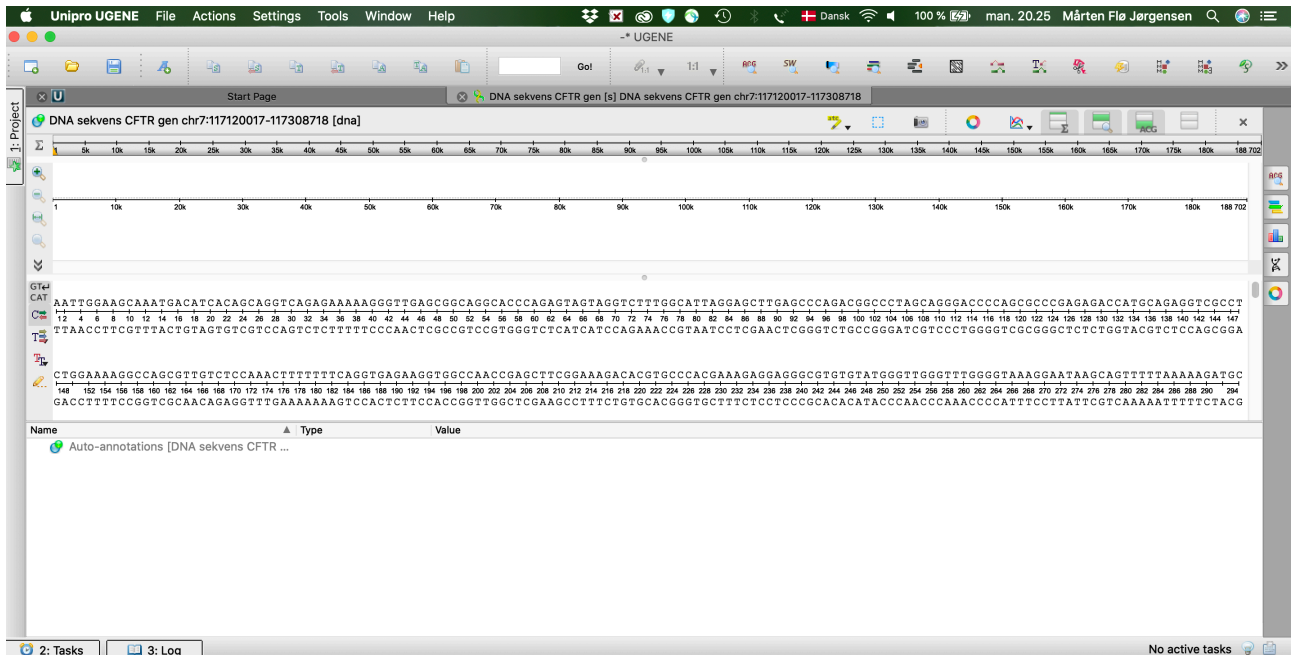
Hver gruppe har 20 minutter til at sætte sig ind i stoffet, og forberede sig på at kunne fortælle om komplikationerne til en anden.

Efter 20 minutter skal I sætte jer i par med en, der er fra den anden gruppe. Fortæl nu til hinanden, hvad I har fundet ud af. Husk at være omhyggelige med jeres forklaringer.

### **Aktivitet 2: CFTR proteinsyntese.**

Forskere har fundet ud af at personer med cystisk fibrose har ændringer i genet CFTR, der står for cystisk fibrosis transmembrane conductance regulator gen. Genet koder for et kanalprotein, hvis funktion er at lade chlorid ioner og hydrogencarbonat ioner (også kaldet bikarbonationer) passere over cellemembraner. For at forstå baggrunden for komplikationerne ved cystisk fibrose, er det vigtigt at kende til den korrekte funktion af det protein, som CFTR genet koder for. Det skal I derfor undersøge nu.

Åbn filen: DNA sekvens CFTR gen.fa i Bioinformatik programmet Ugene. På skærmen vil I se følgende:



I ser en DNA sekvens, begyndende fra venstre. I skal først lave en generel beskrivelse af CFTR genet. Tryk på ikonet til højre med søjlediagrammet:

1. Hvor mange basepar (nucleotider) er DNA sekvensen lang.
2. Hvad er den procentvise fordeling af Adenin (A), Thymin (T), Cytosin (C) og Guanin (G) i genet.
3. Forklar, hvorfor at I bruger bioinformatik, når I besvarer spørgsmål 1+2 og overvej om I kunne besvarer spørgsmålet uden brug af Bioinformatik computerprogrammet.

I skal nu til at undersøge, hvordan der laves et kanalprotein ud fra CFTR genet.

4. Beskriv først de to overordnede trin i proteinsyntese.

Åbn filen: mRNA CFTR gen.fa i Ugene. Der ses en mRNA sekvens, begyndende fra venstre. Sekvensen er blevet annoteret, dvs at man har identificeret funktionen af de forskellige dele. Annoteringen kan ses i tabellen på næste side:

Område	Start pos.	Slut pos.	Længde	Beskrivelse
5' UTR	1	132	132	5' UnTRanslated region
Promoter	4	119	116	Promoter region
Start codon	133	135	3	AUG start codon
RBS	125	136	12	Ribosome Binding Site
CDS	133	4572	4440	CFTR coding region
Stop codon	4573	4575	3	Terminator
3' UTR	4576	6131	1556	3' UnTRanslated region
Poly A hale	6138	6242	105	Poly A tail

- Hvordan kan I se, at det er en mRNA sekvens?
- Beskriv, hvilke områder det færdige mRNA er opbygget af og forklar funktionen af de forskellige dele af mRNA'et. Brug nummeringen af mRNA'et på figuren nedenfor I jeres beskrivelse, der markerer de fire hovedområder CDS, Poly A tail, 3' UTR og 5' UTR.




- Hvad er funktionen af det omdannede guanosintriphosphat, der danner cap strukturen foran mRNA'et.
- Hver gang mRNA'et afkodes af et ribosom, afkortes Poly A halen med 15 adenin nucleotider. Hvor mange CFTR proteiner vil der blive lavet ud fra mRNA'et, når mRNA'et nedbrydes, når der er under 50 adenin nucleotider tilbage.


Det færdige mRNA laves ud fra præRNA, der gennemgår en proces, der kaldes RNA splicing. Resultatet af den proces skal I nu undersøge.

Klik på start page fanebladet. I får nu følgende 4 valgmuligheder:


Welcome to UGENE




Open File(s)



Create Sequence

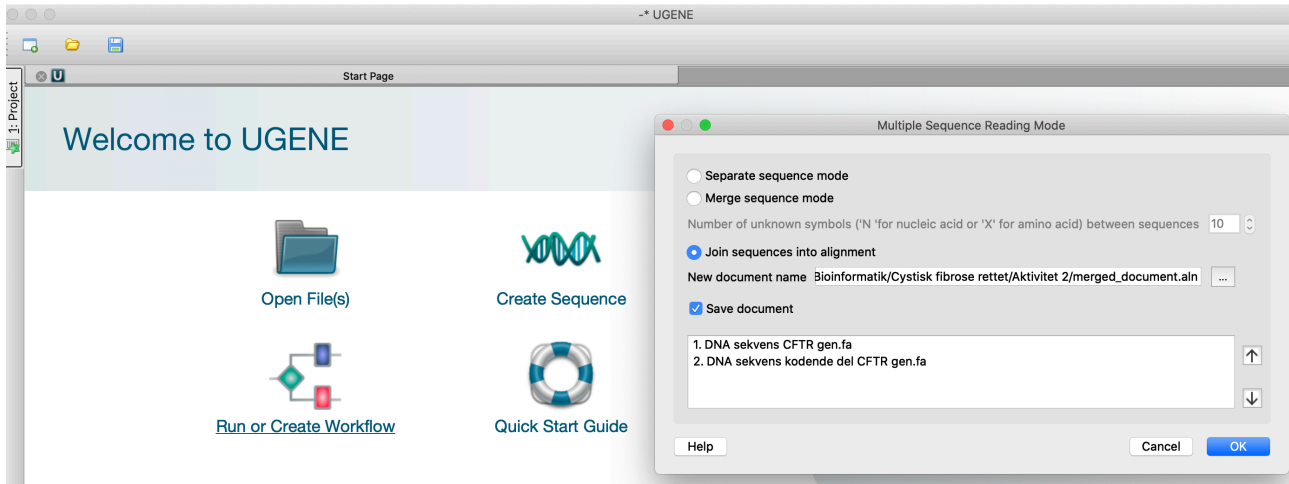


Run or Create Workflow

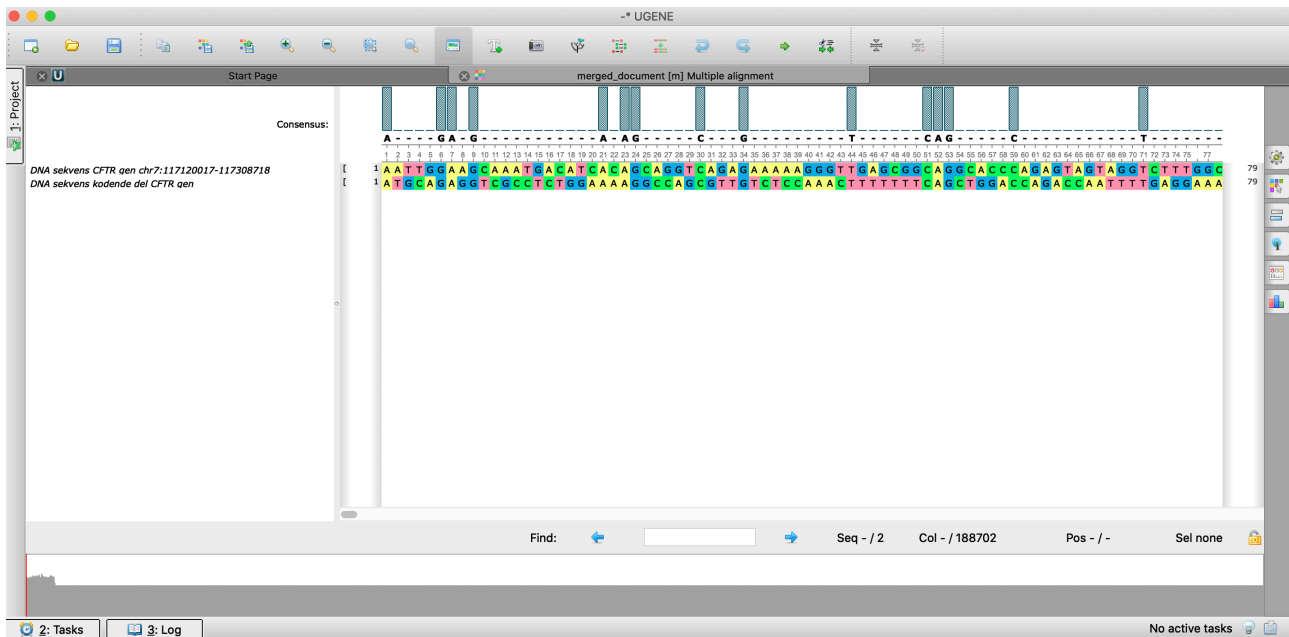


Quick Start Guide

Klik på open file(s) ikonet, og marker følgende to filer: DNA CFTR.fa og kodende DNA CFTR.fa . Tryk på open. I får nu følgende muligheder:

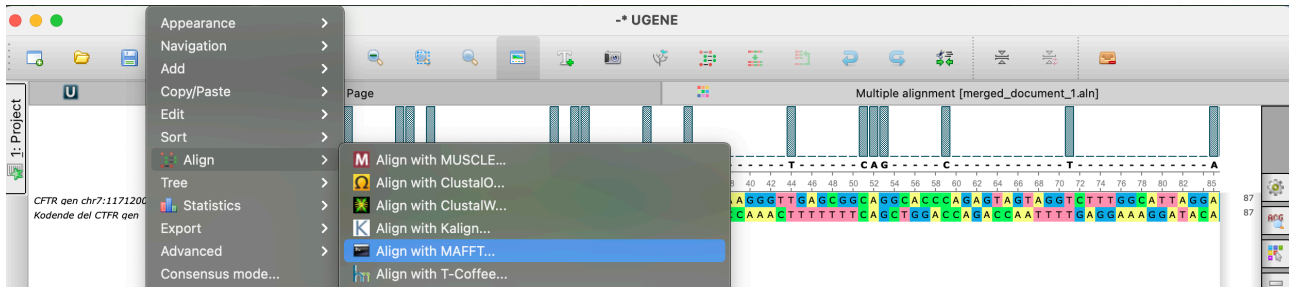


Klik på join sequences into alignment og tryk ok. I får nu et vindue der ser ud på følgende måde:



Sekvensen for CFTR genet og for den kodende del af CFTR genet er nu placeret i samme vindue, men de mangler at blive placeret korrekt under hinanden, således at de samme basepar i genet står under hinanden. Det kalder man at *aligne* sekvenserne og resultat kaldes en *alignment* (fra engelsk at rette ind).

For at lave en parvis alignment af de to sekvenser klikker I på fanebladet actions, align, align with MAFFT som vist på næste side.

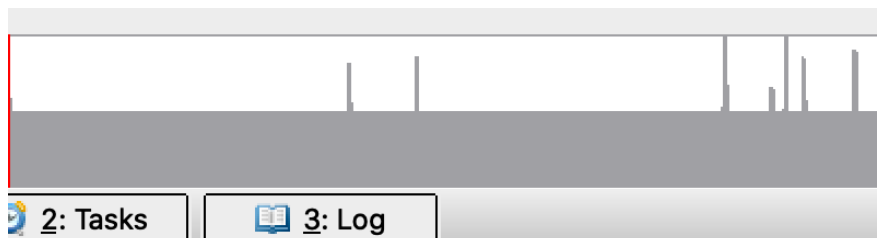


Der findes forskellige matematiske måder at lave alignment på, og I vælger MAFFT programmets metode, fordi det er hurtigt. Programmet vil nu baseret på en udregnings algoritme finde ud af, hvor de to sekvenser er ens og hvor de er forskellige. Det sætter dem sammen så homologe (ens) basepar sættes under hinanden.

I skal nu undersøge, hvor stor en del af CFTR genet der fjernes ved RNA splicing processen. Det gør I ved at identificere, hvor mange introns og hvor mange exons der er i CFTR genet.


9. Forklar, hvad en intron og et exon er. Hvilke dele, introns eller exons, vil derfor findes i både CFTR genets sekvens og CFTR genets kodende del sekvens? Begrund jeres svar.

Da CFTR genet er meget langt, tager det meget tid at scrolle igennem hele genet for at se hvor de to sekvenser er ens. Men nederst i programmet er der en oversigt, der viser hvor ens sekvenserne er i et område. Det ser ud på denne måde:



Hver søjle indikerer at de to sekvenser er ens i det pågældende område. Ved at klikke på søjlen kommer man hen til det pågældende område i sekvens alignment'en.


10. Undersøg, hvor mange exons CFTR genet har i alt?

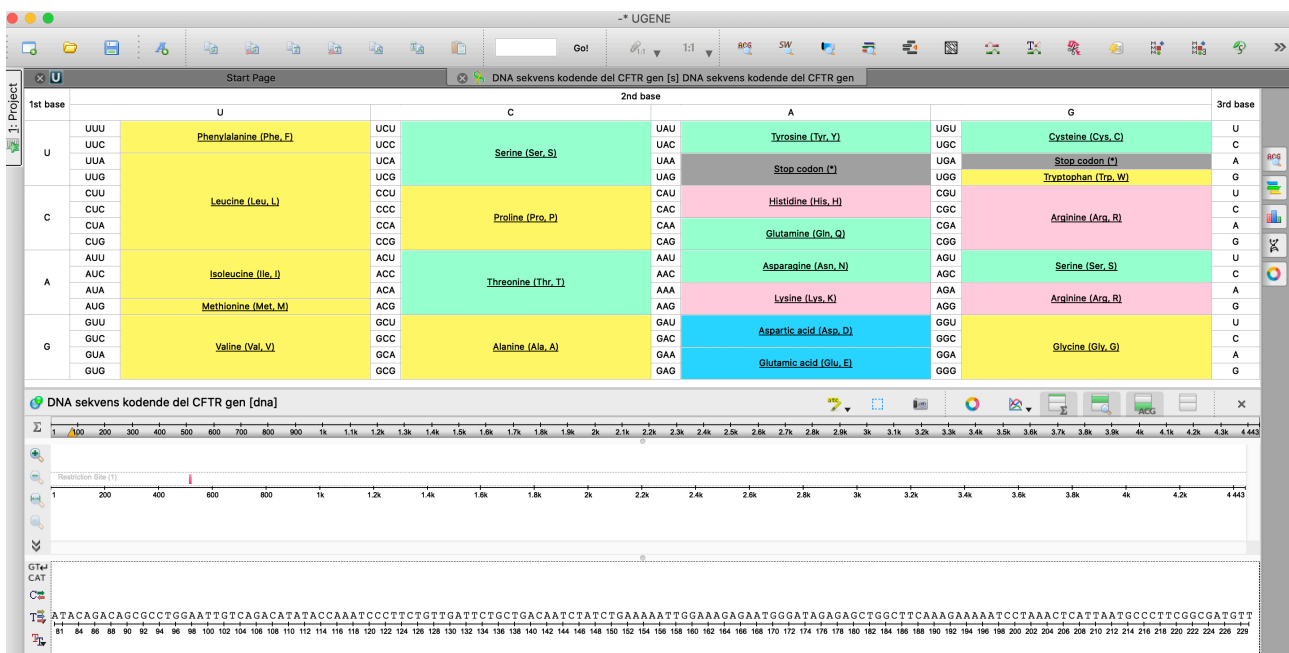
Åben filen: kodende DNA CFTR.fa. Marker sekvensen ved at højreklikke med markøren på sekvensen og vælge hele sekvensområdet - det vil automatisk være valgt. Tryk GO. Analysér sekvensen ved at klikke på  ikonet til højre.

11. Hvor mange basepar er den kodende sekvens?
12. Beregn, hvor mange % den kodende sekvens udgør af det samlede CFTR gen sekvens basepar. Brug jeres svar fra spørgsmål 1 til beregningen.

13. Overvej, hvorfor det er en fordel for eukaryote organismer, at de laver RNA splicing (har introns og exons), selvom længden af det enkelte gen derved bliver meget længere og koster mere energi at lave?

Det færdige mRNA transporteres ud til ribosomerne i cytoplasma, hvor en SSU rRNA del og en LSU rRNA del kobler sig sammen om mRNA'et og danner et funktionelt ribosom, hvor dannelsen af proteinet foregår. I skal nu undersøge, hvordan CFTR protein dannes, og hvilke strukturer der er med til at stabilisere proteinet.


Brug igen filen: kodende DNA CFTR gen.fa. Klik på  symbolet. Den genetiske kode bliver nu synlig.



The screenshot shows the UGENE software interface. At the top, there's a menu bar and a toolbar. Below that, a window titled "DNA sekvens kodende del CFTR gen [s]" displays the genetic code table. The table is organized by the first base (U, C, A, G) and the second base (U, C, A, G). Each cell in the table shows a codon and its corresponding amino acid. For example, UUU codes for Phenylalanine (Phe, F), UCU for Serine (Ser, S), and UGU for Cysteine (Cys, C). The table also includes stop codons (UAA, UAG, UGA) and the start codon (AUG). Below the genetic code table, there's a window titled "DNA sekvens kodende del CFTR gen [dna]" showing the DNA sequence of the CFTR gene. The sequence is displayed in a text area with a scale bar at the top indicating the position of each base pair (1 to 4443).

14. Oversæt de 5 første codoner til aminosyrer. Skriv dem både med tre-bogstavs og et-bogstavs kode.

15. Hvilken aminosyre koder startcodonet for?

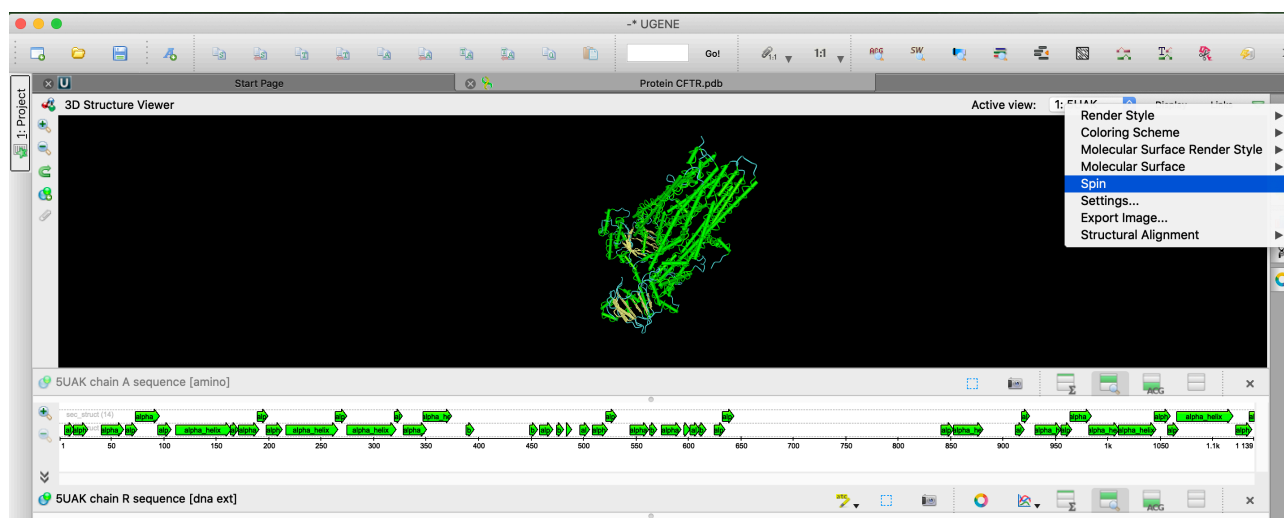
Bioinformatik programmet kan selvfølgelig gøre det meget hurtigere for jer. Markér filen ved at højreklikke som før, og vælg hele sekvensen. Tryk på ikonet  og hele aminosyresekvensen vil komme til syne over DNA sekvensen. I skal nu analysere proteinsekvensen, som er kopieret over i en ny fil.


Åben filen: Aminosyre CFTR.fa

16. Hvor mange aminosyrer består CFTR proteinet af?

Programmer som Ugene kan modellere 3D strukturen af et protein ud fra aminosyrerækkefølgen. Men at forudse 3D strukturen af proteiner er meget vanskeligt, og kræver meget regnekraft, da der skal tages højde for forskellige bindingstyper styrke, forskellige aminosyre sidegrupper orientering m.m. Programmer som Ugene er derfor ikke særligt gode til at forudse 3D strukturen af proteiner. En langt bedre måde er at sammenligne aminosyre sekvensen mod kendte protein strukturer i protein foldnings databaser. I skal nu prøve at undersøge resultatet af en sådan undersøgelse.

Åbn filen: Protein CFTR.pdb. I får nu et skærmbillede som dette:



Klik på display, og vælg spin. Proteinet drejes nu rundt. Ved at trække i proteinet med cursoren, kan I vende og dreje det som I vil. I kan også slå spin fra, og blot flytte rundt med proteinet med cursoren, hvis I synes det er lettere at se. I kan også forstørre proteinet, hvis det gør det lettere for jer ved at trykke på  ikonet.

17. Se på 3D strukturen af proteinet. Beskriv, hvilken tertiær struktur proteinet har. Her er det vigtigt at I har roteret proteinet mange gange, og i forskellige vinkler for at I er sikre på formen.

18. Hvilke sekundære strukturer indgår i proteinet og hvordan er de afbildet i programmet?

I er nu nået frem til at der er lavet et færdigt CFTR protein. Placeringen af de forskellige områder i proteinet har betydning for proteinets funktion, som I skal lære om i aktivitet 3. Men inden I går videre til aktivitet 3, skal I besvare et par refleksionsspørgsmål.

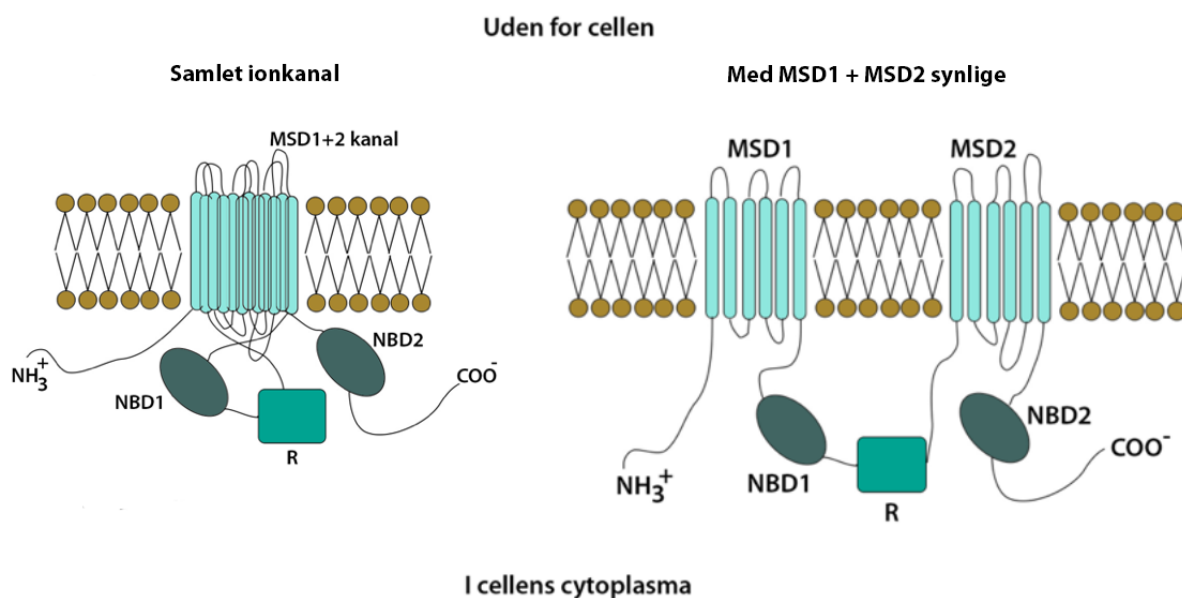
19. Hvilke dele af undersøgelsen af CFTR proteinsyntesen ville ikke have været mulig at lave uden brug af bioinformatik?

20. Er der noget i proteinsyntesen af CFTR kanalproteinet, der har overrasket jer?

### Aktivitet 3. Undersøgelse af de genetiske årsager til cystisk fibrose.

I denne aktivitet skal I undersøge, hvorfor at CFTR kanalproteinets funktion nogle gange er hæmmet eller helt mangler. For at kunne forstå det, skal I først se på, hvordan CFTR proteinet virker.

I har undersøgt CFTR proteinets struktur, og det kan skematisk vises sådan her:



Område	Start pos.	Slut pos.	Længde	Beskrivelse
MSD1	81	350	270	MSD1 1-6 transmembran alfa-helixer
NBD1	434	583	150	Cytoplasma nukleotid binding site (ATP-ase)
R region	590	831	242	Regulatorisk binding site (cAMP aktivering)
MSD2	860	1149	290	MSD2 1-6 transmembran alfa helixer
NBD2	1219	1382	164	NBD2 cytoplasma binding site (ATP'ase)


MSD1+2 danner selve ionkanalen, der altså består af 12 alfa-helix strukturer, der går igennem cellemembranen.

CFTR kanalproteinet kan være åbent eller lukket. Om det er åbent eller lukket afhænger af de tre regulatorområder, NBD1, R og NBD2. R området er det overordnede regulatoriske område, hvor forskellige signalstoffer fra cellen kan binde sig, f.eks cyklisk AMP. Hvis det sker, gives der signal til CFTR kanalproteinet om at åbne sig. Det kræver dog også, at der spaltes ATP, hvilket foregår i begge de regulatoriske områder NBD1 og NBD2. Disse tre områder styrer altså, om CFTR proteinet er lukket eller åbent.

1. Overvej hvorfor, at det er smart at der kan åbnes og lukkes for CFTR kanalproteinet.

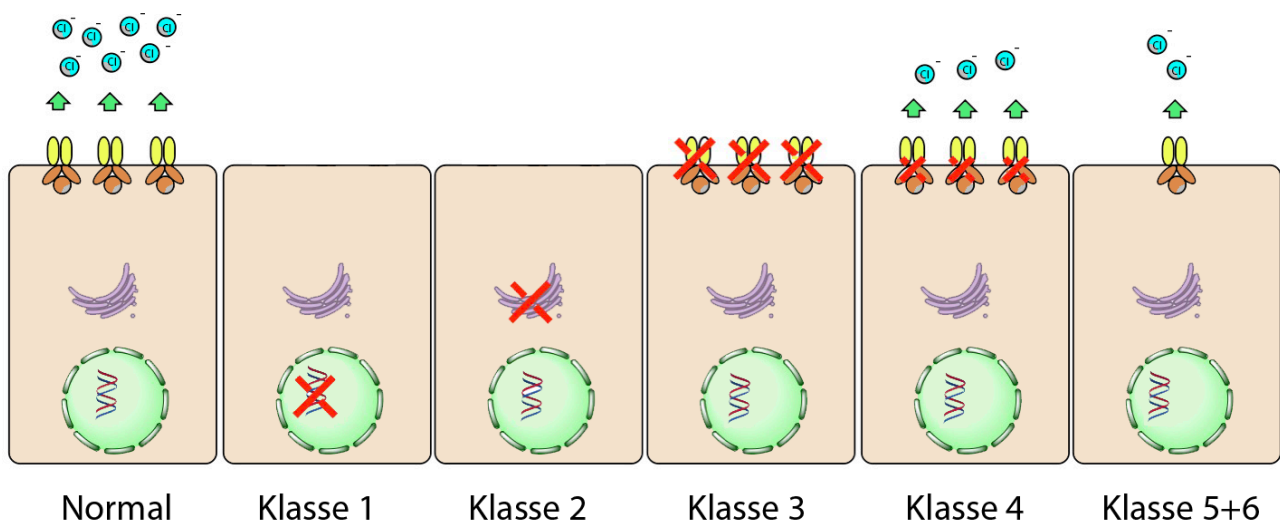
Man har undersøgt 3D strukturen af CFTR genet med ATP og cAMP tilsat. Hvis man blander ATP + cAMP molekyler med CFTR proteinet, sker der en ændring af 3D strukturen. Det skal I undersøge



nu. Åben først filen protein CFTR.pdb, dernæst protein CFTR+ATP.pdb. Klik tilbage på fanebladet protein CFTR.pdb. Klik på  ikonet til venstre, klik på den lille trekant ud for protein CFTR+ATP.pdb og klik på den linje der kommer frem. Tryk ok. I vil nu have de to proteiner i 2 vinduer ved siden af hinanden. Placer proteinerne ens ved at rykke dem med cursoren, indtil de har ionkanalen opad.

2. Hvilken forskel gør det for CFTR proteinets tertiære struktur at ATP + cAMP tilsættes?
3. Overvej, hvorfor at CFTR protein først bliver aktivt som kanalprotein, når der er tilsat ATP + cAMP ud fra ændringen i proteinets tertiære form?

Cystisk fibrose skyldes mutationer i CFTR genet. Man har identificeret mere end 2000 forskellige mutationer i CFTR genet, som påvirker dannelsen af CFTR proteinet på forskellig måde. Mutationerne har man inddelt i 6 forskellige klasser, alt efter hvilken effekt de har på dannelsen af CFTR kanalproteinet.



I en normal celle vil der dannes CFTR protein i ribosomerne, ligesom I har undersøgt I aktivitet 2. Det dannede protein vil, hvis det er foldet i den rigtige 3D struktur, transporteres via det endoplasmatiske retikulum ud til cellemembranen, hvor det indlejres i cellemembranen og virker som et kanalprotein.

Ved en klasse I mutation dannes der ikke et funktionelt CFTR protein pga. fejl i CFTR genets kode. Ved en klasse 2 mutation dannes der et CFTR protein, men det bliver foldet i en forkert 3D struktur, og vil derfor ikke blive transporteret op til cellemembranen. Ved en klasse 3 mutation er CFTR protein foldet korrekt, og transporteres op til cellemembranen, men proteinets ionkanal kan ikke åbnes. Ved klasse 4 mutationer er proteinets ion kanal misformet, så hastigheden hvormed chloridioner kan passere er lavere og ved klasse 5+6 mutationer produceres der ikke nok CFTR



protein pga. fejl i genekspressionen af CFTR genet eller for hurtigt nedbrydning af proteinet. Læs mere på <https://cystiskfibrose.dk/cystisk-fibrose/mutationsklasser-ved-cystisk-fibrose/>)

4. Mutationsklasse 1-3 giver cystisk fibrose med de alvorligste komplikationer. Overvej hvorfor?


I skal nu analysere 3 almindeligt forekommende mutationer i CFTR genet, og undersøge, hvilken effekt de har på det færdige CFTR protein.

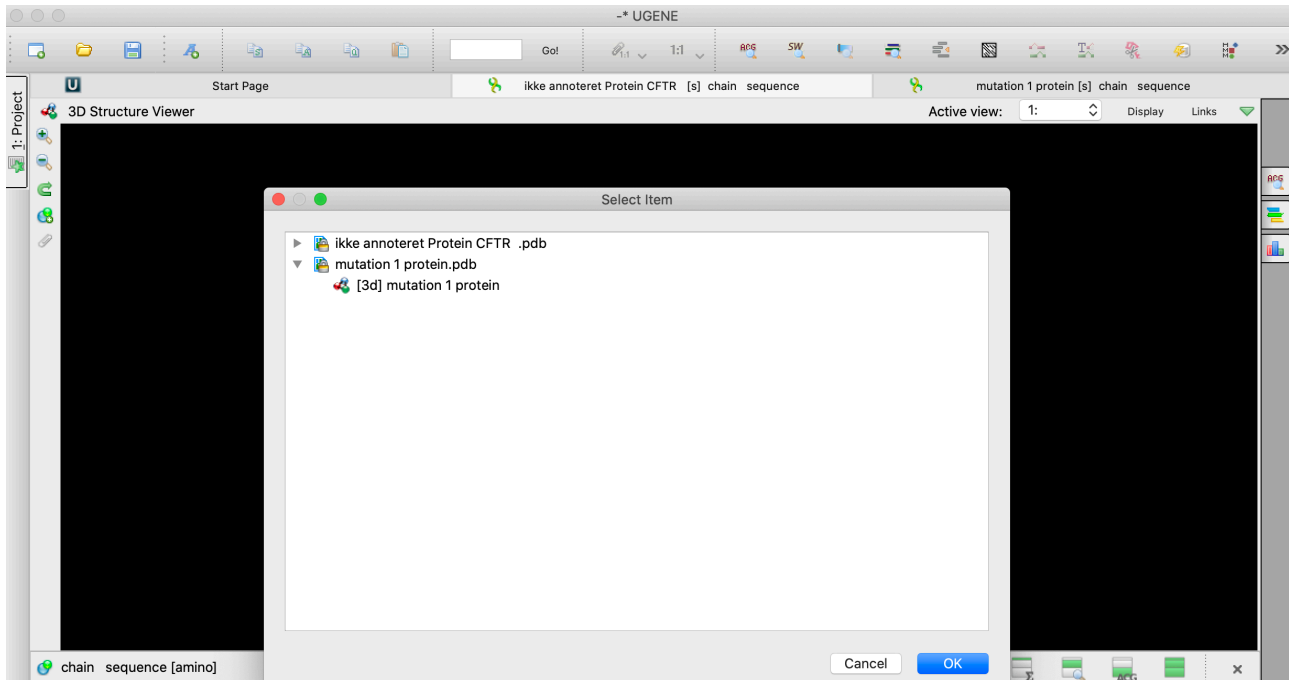
Vælg kodende del DNA CFTR.fa og mutation 1 DNA.fa filerne og åben dem sammen. Lav en alignment som beskrevet på side 3+4.

5. Sammenlign de to sekvenser. Hvilken mutation er der sket og ved hvilket basepar er mutationen placeret?

Åben nu aminosyresekvens aminosyre CFTR.fa og mutation1 aminosyre.fa filerne sammen. Lav igen en alignment som beskrevet på side 3+4, men vælg denne gang Muscle i stedet for MAFFT.

6. Hvilken forskel er der mellem de to aminosyresekvenser?
7. Se på annotationslisten på side 9 for CFTR proteinet. Ligger mutationen i et område af proteinet, hvor det påvirker proteinets funktion? Hvis ja, på hvilken måde?

Åben protein ikke annoteret CFTR.pdb i Ugene. Åben dernæst Mutation 1 protein.pdb filen. Klik tilbage på protein ikke annoteret CFTR.pdb. Klik på  ikonet til venstre, og vælg mutation 1.pdb ved at markere den - se dialog boks nedenfor. Tryk ok. I vil nu have de to proteiner i 2 vinduer ved siden af hinanden.

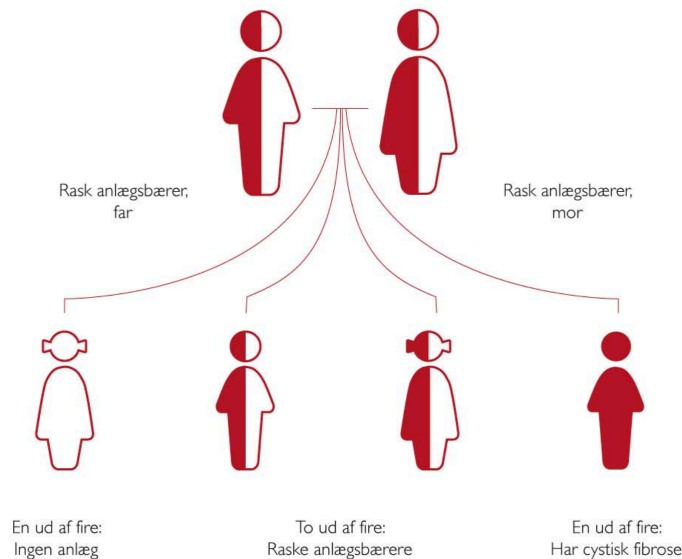


8. Sammenlign de to proteiners 3D struktur. Har mutationen medført en stor ændring i proteinets 3D form?

Gentag spørgsmål 5-8 for mutation 2 og 3 (Mutation 2 DNA.fa og Mutation 3 DNA.fa filerne) og notér svarene under de respektive spørgsmål.

9. Angiv om de undersøgte mutationer er en punktmutation, en deletion eller en insertion, en nonsense mutation (for tidligt stopcodon) eller en frameshift mutation (forrykket læseramme)
10. Hvilken mutation vil høre til mutationsklasse I? Begrund jeres svar.

#### Aktivitet 4. Nedarvning af cystisk fibrose gener og vejledning til bærere af CFTR mutationer.



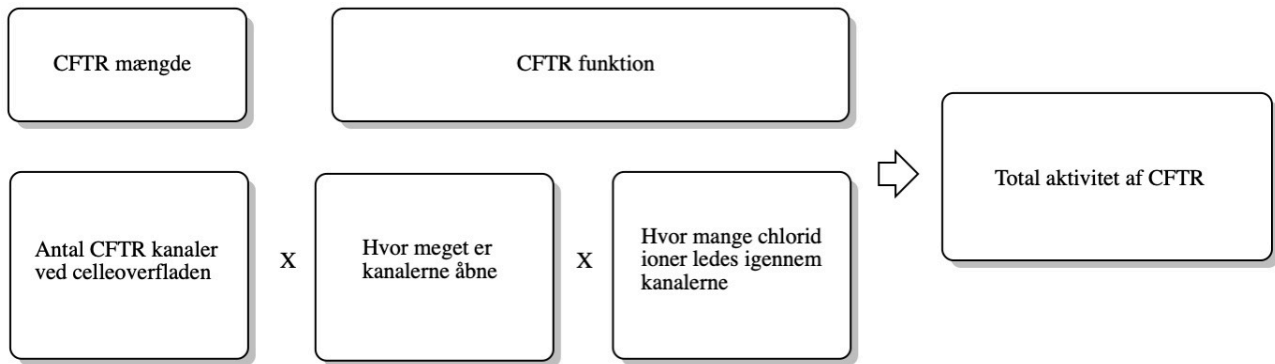
© Cystisk Fibrose foreningen

Cystisk fibrose er et eksempel på en arvelig sygdom, der nedarves autosomt recessivt.

1. Hvad betyder det, at sygdommen er autosomal recessiv.
2. Forklar, hvorfor at sygdommen nedarves som vist på figuren ovenfor.
3. Sygdomme, der skyldes defekte proteiner eller enzymer nedarves ofte recessivt. Hvorfor?

I skal nu forberede en samtale med en familie, hvis barn har cystisk fibrose. I ved, at der findes mange forskellige mutationer, der kan føre til cystisk fibrose og at de kan deles i 6 overordnede mutationsklasser. Derfor er det langt fra altid at en person, der har cystisk fibrose, har den samme mutation på begge kromosomer. I skal nu vejlede en familie, der for nyligt har fået et barn med cystisk fibrose. Moren og faren er efterfølgende blevet testet, og det har vist sig at moren har mutationen delta P508 (mutation 1, tilhører mutationsklasse II) og faren er bærer af mutationen Arg334Trp (tilhører mutationsklasse 4)

Hvor slemt påvirket af cystisk fibrose en person bliver, afhænger af hvordan den totale aktivitet af CFTR kanal proteinerne er. Det kan stilles op på følgende måde:



4. Vurdér, hvor slemt barnet vil være påvirket, ud fra forældrenes mutationstyper.
5. Planlæg nu, hvad vil I fortælle til familien om udsigterne for komplikationer for barnet. Husk at begrunde jeres svar.

To af de medikintyper, man bruger til behandling af cystisk fibrose er Ivacaftor® og Lumacaftor®. Ivacaftor® virker ved at udvide ionkanalerne i CFTR proteiner, så der kan strømme flere chloridioner igennem. Lumacaftor® virker ved at rette 3D strukturen af misfoldede CFTR proteiner.

6. Hvilken behandling vil I foreslå til behandling af barnet med cystisk fibrose? Begrund jeres svar.
7. Vil Lumacaftor® kunne bruges i behandling af mutationsklasse 1 CFTR mutationer? Hvorfor eller hvorfor ikke?

I har nu lært om, hvorfor at cystisk fibrose opstår. Men hvordan er det at leve med en genetisk sygdom som cystisk fibrose? Det skal i afslutningsvis sætte jer ind i. Gå ind på <https://cystiskfibrose.dk/cystisk-fibrose/personlige-historier/?filter=at-leve-med-cystisk-fibrose> og vælg en personlig historie. Din makker skal vælge en anden historie end dig. Læs den personlige historie.

Snak med din makker om det du har læst - er der noget der har overrasket jer?

Stor tak til Cystisk Fibrose foreningen for tilladelse til at bruge figur om nedarvning side 13.