

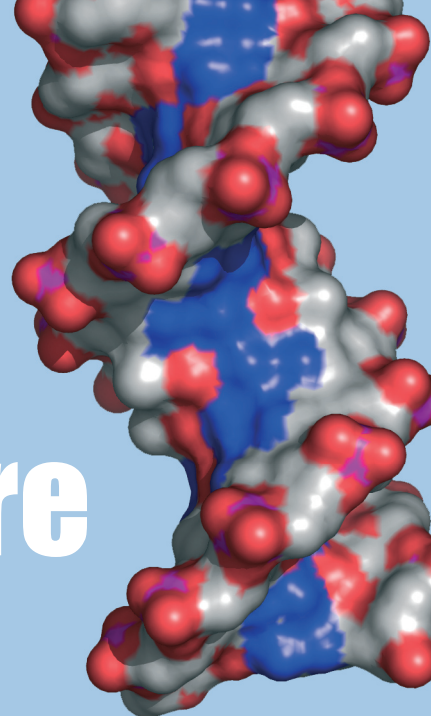
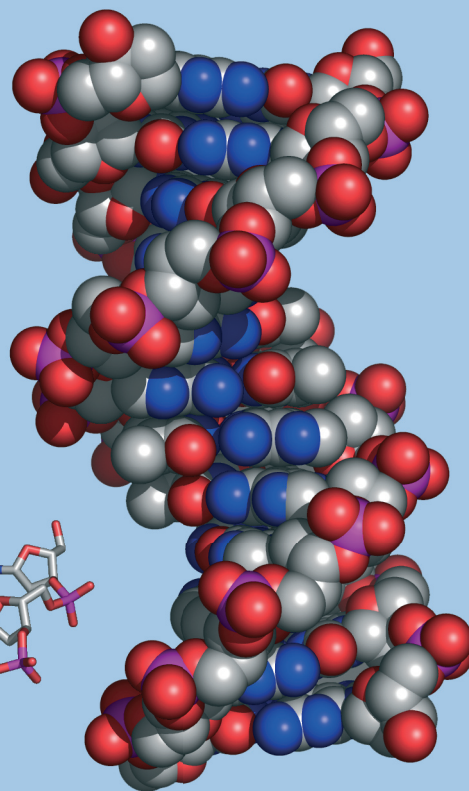
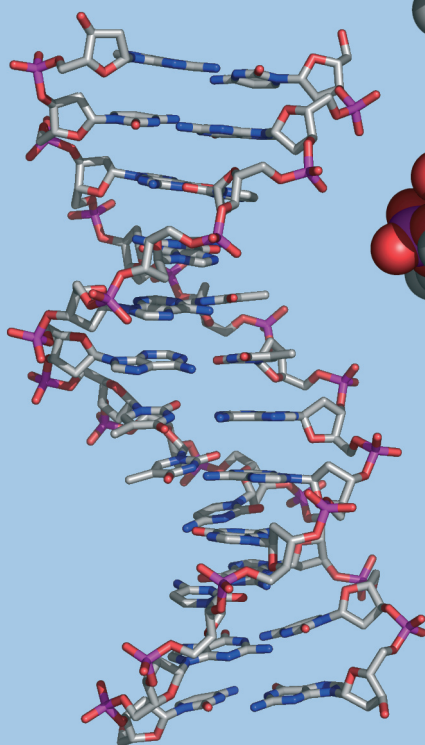
4. ÅRGANG  
NR. 2 / 2006

MATEMATIK **KEMI** I FYSIK *i*

# PERSPETTIV

## Livets molekylære kode

**-kemi og  
biologi mødes**



# DNA – livets molekylære kode

DNA findes i alle levende organismer og indeholder koden til organismen – informationer, der afgør om organismen er en bakterie, en snegl, en hund eller et menneske.

## Kromosomer

DNA er meget lange molekyler, som er spolet op i tætpakkede "garnnøgler", som kaldes kromosomer. Mennesker har to sæt à 23 kromosomer, et sæt fra mor og et fra far. Derfor findes der 46 af disse DNA-molekyler i enhver af menneskecellerne. I vores krop er der ca.  $10^{14}$  celler, som er organiseret i forskelligt væv, fx hud, muskler og knogler.

Selv om hver celle indeholder al den menneskelige genetiske instruktion, er det dog kun en del af denne information, der benyttes i den enkelte celle. Hvert DNA-molekyle, der danner et kromosom, kan opfattes som et sæt af kortere DNA-stykker, der hver især indeholder instruktionerne til fremstilling af netop én komponent til cellen. Disse kortere DNA-stykker kaldes *gener*.

I mennesket findes der ca. 30.000 gener. Alt DNA'et i kromosomerne kaldes samlet for *genomet* - hvilket betyder: den komplette genetiske instruktion for en organisme. Man kan opfatte genomet som et bibliotek, hvor hvert kromosom er en bog, og hver side i denne bog er et gen.

## DNA er lange molekyler

De ekstremt lange DNA-molekyler er lange kæder af fire næsten ens kemiske byggeklodser. Disse byggeklodser kaldes *nucleotider*. Navnet henviser til, at DNA findes i cellerkernen – nucleus. Et nucleotid består af tre kemiske molekyldele, fosfat, ribose og en "base", der er en nitrogenholdig molekyldel, som består af én eller to ringe, se figur 1. Der findes fire forskellige "baser", Adenin (A), Thymin (T), Guanin (G) og Cytosin (C), se figur 2. DNA er derfor opbygget af strenge, hvor fosfat fra ét nucleotid er bundet til ribosen af det næste. Herved dannes lange kæder med skiftevis fosfat og ribose, hvorfra basen stikker ud, se figur 3.

Menneskets DNA består af DNA-kæder, der tilsammen er ca. 3.000.000.000 nucleotider lange.

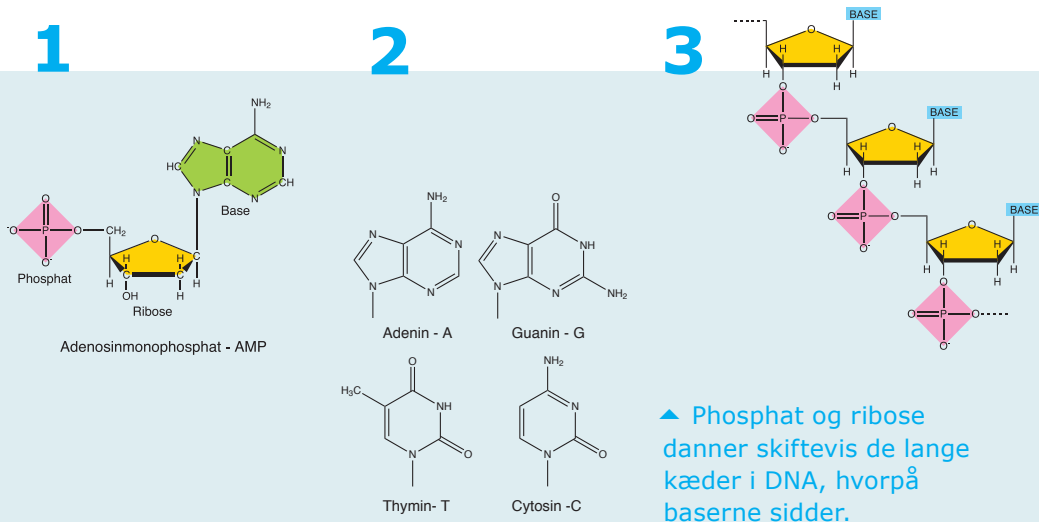
Alle pattedyrs DNA har omtrent samme størrelse, mens fx størrelsen af en bananflues DNA "kun" er 180.000.000 nucleotider. Størrelsen af bakteriers DNA kan variere, men et gennemsnit er ca. 4.000.000 nucleotider.

## DNA-molekylet

Et DNA-molekyle består faktisk af to strenge med A, T, G, og C'er. De to strenge hænger sammen via forskellige kemiske bindinger, men det er især hydrogenbindingerne mellem baserne på de to strenge, der får DNA til at hænge sammen.

Det har vist sig, at nucleotiderne altid findes i par, således vil der altid være A over for T og G over for C. Dette skyldes, at baserne netop kan danne hydrogenbindinger i disse par, mens andre par, fx A-G eller C-C ikke danner hydrogenbindinger. Konsekvensen af dette er, at hvis man kender rækkefølgen, sekvensen af nucleotider i den ene streng, vil man også kende sekvensen i den anden streng. Man kan sige, at den ene streng er en omvendt kopi af den anden, og med den ene DNA-streng kan man lave kopier af den anden, se figur 4.

Dette udnyttes, når cellerne deler sig til to nye celler. DNA'et kopieres til den nye celle ved, at DNA-strengene kopieres hver for sig, og to nye DNA-molekyler dannes. Alt dette foregår ved, at et enzym, DNA-polymerase, aflæser sekvensen på den streng, der kopieres, og opbygger en ny streng, der passer i sekvens. DNA-polymerase enzymet bygger DNA-molekylet op af nucleotider med tre fosfatgrupper. Ved sammenkoblingen bliver de yderste to fosfatgrupper spaltet fra og leverer herved kemisk energi til reaktionen, se figur 5.



► En lille stump af en DNA dobbeltstreng. De to basepar A:T og G:C har henholdsvis to og tre hydrogenbindinger. De nucleofile nitrogen- og oxygenatomer danner hydrogenbindinger med elektrofile hydrogenatomer på modsatte base. Fosfatgrupperne bærer alle én negativ ladning ved neutralt pH, der gør DNA-molekylet meget hydrofilit.

## DNA-helix

DNA-molekylets to strenge snor sig omkring hinanden i en struktur, der ligner en vindeltrappe, hvor de enkelte base-par danner trappetrinene, og fosfat-ribosekæderne danner siderne. Denne struktur kaldes en DNA-helix.

Det er en overordentlig genial struktur, hvor ydersiden, der består af fosfat-ribosekæderne, er meget hydrofil og derfor meget opløseligt i vand. Vi skal huske, at de fleste levende organismer består af ca. 80 % vand. Fosfat-ribosekæderne er også meget modstandsdygtige overfor kemisk nedbrydning, hvilket gør DNA stabilt.

Inde i helixen findes baserne, som er beskyttet mod fx kemisk eller enzymatisk nedbrydning, se figur 6. Den enkelte hydrogenbinding er svag, men da der er 2-3 hydrogenbindinger pr. basepar, sikrer den samlede effekt af de mange hydrogenbindinger, at DNA-strengene holdes sammen. Bindingerne kan dog relativt let brydes, hvis temperaturen hæves til ca. 100°C.

## DNA-sekvensen

Det er sekvensen af nucleotider, der indeholder koden eller instruktionerne til at fremstille de molekyler, som er nødvendige i en levende organisme. De fire bogstaver i den genetiske kodebog kan kombineres på uendeligt mange måder.

Tænk blot på, at der findes 16 forskellige sekvenser af to baser (24 muligheder), og en sekvens på 1000 baser kan laves på  $1000^4 = 1.000.000.000.000$  måder. I en *E. coli*-bakterie findes der ca. 4.400 gener med en gennemsnitslængde på ca. 1.000 nucleotider, så der er et astronomisk antal muligheder for at danne disse gener. Det er dog kun nogle ganske få af disse muligheder, der kan bruges for at danne en *E. coli*-bakterie.

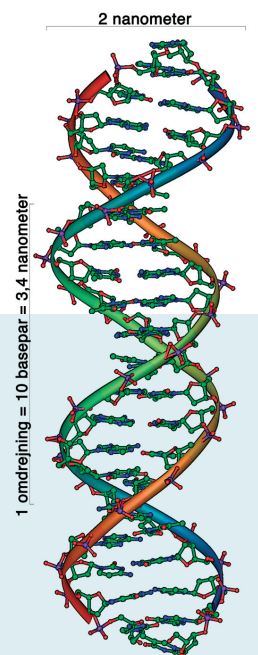
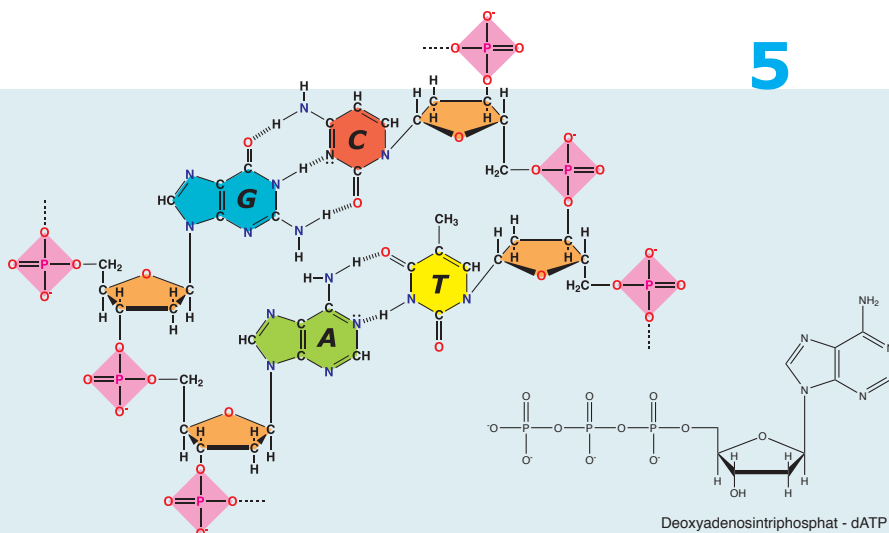
Hvis man sammenligner sekvensen af genomet fra to forskellige *E. coli*-bakterier, vil man finde, at der i gennemsnit er forskelle i 0,1 - 0,2 % af baserne. Det er det samme, man ser, hvis man sammenligner to menneskers genomer. Disse forskelle afgør, om man er høj eller lav; neger eller kineser; blond eller rødhåret, men også om man har en arvelig (genetisk) sygdom, og om man bliver syg eller ej.

Det er derfor yderst interessant at kende sekvensen af det menneskelige genom og dermed kunne sammenligne menneskelige genomer for at finde sammenhænge mellem sekvensvariation og udseende eller sygdomsbillede.

## Det menneskelige genom

Allerede i 1983 begyndte man arbejdet med at sekventere de 3 milliarder nucleotider i det menneskelige genom. På det tidspunkt regnede man ikke med at blive færdig lige med det samme – med datidens teknologi ville det tage over 1.000 år at gøre projektet færdigt. Allerede i 2003 lå den første udgave af hele det menneskelige genom færdigt. Det var den fantastiske teknologiske udvikling, der var skyld i, at det kun tog tyve år at gøre projektet færdigt.

Projektet har medført, at mange medicinske problemstillinger nu kan ses fra en genetisk synsvinkel, og måske giver dette mulighed for mere direkte behandlingsmetoder. På langt sigt kan man forvente, at alle kan få lavet en sekvens af deres eget genom, og at det bliver med udgangspunkt heri, at man vil få indblik i, hvad man kan tåle, hvilket arbejde der passer bedst, og om man i fremtiden vil få skavanker eller sygdomme. Det er jo et spændende, men også et skræmmende perspektiv.



► DNA dobbelthelix. Hydrogenbindinger holder baserne sammen parvis og danner "trinnene" i DNA-vindeltrappen. Ribose og fosfat danner kæder, der snor sig omkring hinanden i dobbelthelixer.

# Moderne retsgenetiske metoder

Kun enæggede tvillinger har præcist samme genom-sekvens. For alle andre er genom-sekvensen et ekstremt nøjagtigt fingeraftryk. Selvom vores genom-sekvenser ligner hinanden, er der dog ca. 1 ‰ forskel, og genomets størrelse betyder, at der er ca. 3.000.000 forskelle mellem to forskellige mennesker.

Derfor er DNA blevet et vigtigt værktøj i forbindelse med opklaring af forbrydelser og ved identifikation af ukendte (både levende og døde).

Man kan dog endnu ikke lave en komplet sekvens af et individs genom, men ved at bruge nogle bestemte områder på kromosomerne kan man analysere for ligheder og forskelle imellem forskellige individers DNA.

Der anvendes to forskellige metoder i dag. Den ene går ud på at lave fragmenter af DNA'et ved hjælp af meget specifikke og præcise metoder, den anden går ud på at bestemme sekvensen af korte områder i DNA'et.

## Variation i fragmentstørrelser

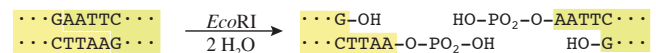
DNA kan skæres i stykker ved hjælp af meget specifikke enzymer, der kun skærer ved ganske bestemte sekvenser. Disse enzymer, der hedder restriktionsenzymer, kommer fra bakterier, hvori de fungerer som forsvar mod fremmed DNA. Hvis en bakterievirus skulle have held med at sende sit DNA ind i bakterien, kan restriktionsenzymene skære det fremmede DNA i stykker.

Disse enzymer kan genkende sekvenser på 4 til 8 nucleotider, og for de fleste enzymer gælder det, at de kun genkender sekvenser, der er palindrome.

Et *palindrom* er et ord (eller en sætning), der læses ens fra begge ender, et eksempel herpå er navnet ANNA (eller sætningen - Sær giraffar i græs). En palindromsekvens kan fx være GAATTC, denne sekvens genkendes af enzymet *EcoRI*, der kommer fra bakterien *E. coli*. Nu kan denne sekvens jo ikke umiddelbart læses ens i begge retninger, men husk på, at DNA er dobbeltstrengt:

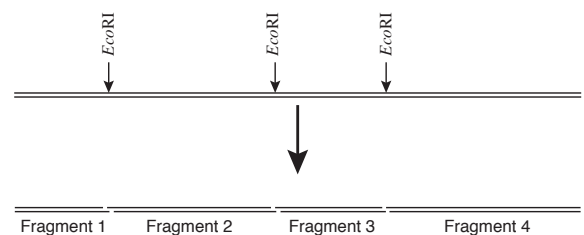
```
...GAATTC...
...CTTAAG...
```

Som det ses, læses man GAATTC fra begge sider. Enzymet vil skære igennem sekvensen, dog ikke på midten, men imellem G og A. *EcoRI* enzymet hydrolyserer bindingen mellem fosfat og ribose i begge strenge:



Der findes også enzymer, der skærer midt igennem sekvensen, og enzymer, der skærer i den modsatte ende i sekvensen.

Hvis man isolerer DNA fra menneskeceller, kan man skære dette med fx *EcoRI*. DNA'et vil så blive skåret i et antal stykker, der svarer til antal *EcoRI*-sekvenser plus én.



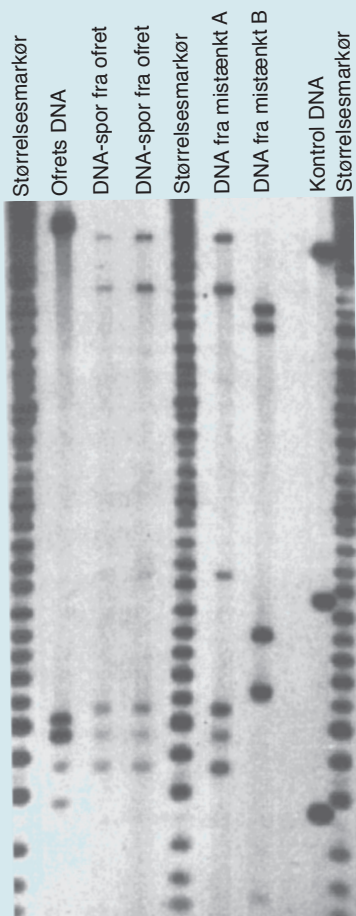
▲ Hvis et DNA-molekyle skæres tre steder af fx enzymet *EcoRI*, resulterer dette i 4 DNA-fragmenter ( $n+1$  fragmenter).

Hvis man sammenligner resultatet heraf med et tilsvarende fra et andet menneske, vil man se, at der ikke er lige mange fragmenter. Altså har vi ikke lige mange *EcoRI*-sekvenser i vores genom-sekvens. Dette skyldes variationen i vores genom-sekvens, og at der en gang imellem vil være en *EcoRI*-sekvens, hvor én af baserne er ændret. Sekvensen er nu ikke længere en sekvens, der kan genkendes af *EcoRI*, fx:

```
...GAATCC...
...CTTAGG...
```

Vi har nu en metode, hvor vi kan bruge variationen i DNA'et til at se forskel på forskellige mennesker. Hvis man bruger flere enzymer, får man en høj nøjagtighed, som er tilstrækkelig til at bevise, at en mistænkt ikke er den rette gerningsmand. Det modsatte - at bevise, at en mistænkt er gerningsmanden, kræver dog større nøjagtighed, end denne metode kan yde.

# bruger DNA



▲ En kvinde er blevet voldtaget, der er to mistænkte. DNA-prøver af ofret selv, DNA-spor på ofret og af de to mistænkte udtages. DNAet skæres med enzymer og analyseres ved hjælp af gelelektroforese. Ved at sammenligne mønstrene fra DNA-sporene med de mistænkte DNA er det tydeligt, hvem der er den skyldige. Kvindens DNA er med i analysen, som bevis på at DNA-sporene ikke stammer fra hende selv.



## STR (Short Tandem Repeats)

I dag findes der mere specifikke analyser, der udnytter, at vores genom indeholder områder, hvor korte sekvenser på 2-5 baser (fx AATG) repeteres mange gange. I forskellige individer er disse korte sekvenser repeteret et forskelligt antal gange.

Man kan kopiere disse områder ved hjælp af DNA-polymerase-enzymet og måle deres størrelse. Ved at sammenligne disse størrelser fra DNA, som er fundet på et gerningssted og fra en eller flere mistænkte, kan man med meget høj nøjagtighed bevise en mistænks uskyld, men også med meget høj sandsynlighed bevise, om en mistænkt er skyldig. FBI benytter et system med 13 forskellige områder på forskellige kromosomer. Den største nøjagtighed, der kan opnås ved denne analyse, er  $1:10^{17}$ .

Det betyder, at sandsynligheden for at udpege den rigtige er ekstrem stor, når man husker på, at hele Jordens befolkning pt. er ca. 6 milliarder mennesker.

## DNA sekvens metoden

Næsten alle celler indeholder et organel, der kaldes et *mitokondrie*. Disse organeller har stor betydning for energiomsætningen i cellen. Hvert af disse mitokondrier har sit eget DNA i form af et lille kromosom. Da mitokondrierne stammer fra ægcellen og derfor kun nedarves fra moderen til barnet, kan sekvenser på dets kromosom anvendes til verificering af familieforhold på mødrenes side.

Det mandlige Y-kromosom nedarves fra far til søn og kan derfor ligeledes anvendes til påvisning af familieforhold på faderens side, fra far til drengebørn og mellem brødre.

I begge tilfælde laver man en bestemmelse af DNA-sekvensen af et lille stykke af kromosomet fra de implicerede familiemedlemmer. Sekvenserne sammenlignes, og identiske sekvenser indikerer et familieforhold. Denne type analyse blev benyttet til at verificere, at det faktisk var resterne af den henrettede russiske tzarfamilie, som blev fundet i et massegravsted for nogle år siden.

# Hvorfor er vi ikke ens?

Hvorfor er vi mennesker så forskellige? Svaret findes i variationen i vores genom-sekvens. Tre millioner forskelle findes der mellem to mennesker. Af disse har nogle betydning for udseendet, andre for vores muligheder for at modstå påvirkninger fra miljøet. Nogle resulterer i sygdomme, mens andre ingen betydning har.

## Udviklingshistorie – en lang række mutationer

Hvor kommer alle disse forskelle fra – stammer vi ikke fra den samme forfader? Tja, det ved vi faktisk ikke rigtigt, men mange analyser af vores genom-sekvens peger på, at vi stammer fra en fælles forfader (-moder) et sted i Afrika.

Udviklingen af denne forfader til det moderne menneske skyldes, at der til stadighed opstår mutationer i genomet. Det vil sige, at en eller flere baser i genom-sekvensen ændrer sig, og at denne ændring videreføres til afkommet. Hvis afkommet overlever, så det kan videreføre ændringen til sit afkom, har mutationen været en succes, men hvis afkommet dør, har mutationen været en fiasko og forsvinder igen.

Det er altså kun de succesfulde mutationer, vi kan finde i vores genom i dag. Nogle mutationer er så succesfulde, at de forbedrer individets muligheder for at overleve, og det er disse mutationer, som udvikler en race. Med Darwins ord: *“Survival of the fittest”*.

## Mutationer – hvordan opstår de?

Hvordan opstår disse mutationer? Der findes mange årsager til, at der sker ændringer i kromosomerne, UV-lys fra Solen, giftige kemikalier fra miljøet og kopieringsfejl, når cellerne deler sig. Der er dog én betingelse, som skal være opfyldt, for at mutationen kan viderebringes i slægten: at mutationen er sket på kromosomerne i kønscellerne. Alle andre mutationer i andre celler end kønscellerne vil ikke videregives og vil derfor forsvinde igen, når individet dør. Det betyder dog ikke, at disse mutationer ikke kan have stor indflydelse på netop dette individs livsbetingelser.

Mutationer kan opstå, hvis DNA bliver bestrålet af UV-lys eller radioaktiv stråling. Et meget kendt eksempel er en reaktion, hvor T-baser, som sidder ved siden af hinanden i sekvensen, krydsbinder sig til hinanden ved bestråling og derfor ikke længere kan danne hydrogenbindinger med A-baser på den modsatte DNA-streng. Når dette bestrålede DNA kopieres, bliver DNA-polymerase-enzymet forvirret og kan sætte en eller flere forkerte baser ind i den nye kopi af DNA'et. Se figuren til højre.

Kemikalier, som ligner baserne, kan binde til DNA'et og forskyde afstandene mellem base-trinnene i dobbelt-helixen. Dette kan også forvirre polymerase-enzymet til at sætte en ekstra base ind i sekvensen.

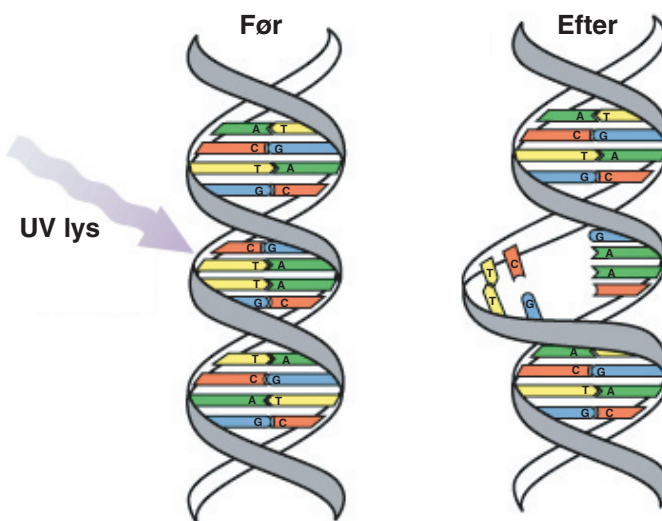
DNA-polymerase enzymet har en indbygget korrekturlæsnings-funktion, som i langt de fleste tilfælde kan finde og reparere fejl i DNA-sekvensen på den ene DNA-streng. Man skal dog huske, at enzymet jo ikke ved hvilken streng, der er den rigtige, og reparationen kan derfor betyde, at det er den "rigtige" base, der bliver rettet til den muterede.

Der findes mange muligheder for, at der kan ske fejl og opstå mutationer i vores DNA, og det er faktisk overraskende, hvor få fejl der får lov til at ske.

## Det menneskelige genom

Nogle af de forskelle, vi har i vores genom, påvirker vores udseende og måske vores tilpasning til det miljø, vi lever i. Andre forskelle findes uden for generne i områder af genomet, der formentlig ikke har en funktion. Nogle få sidder i gener, der medfører arvelige sygdomme, som enten er til stede ved fødslen eller først viser sig senere.

Et af perspektiverne med sekventeringen af det menneskelige genom er at kunne finde og forstå disse mutationer, og måske endda kunne sætte ind med modforanstaltninger, der kan reducere eller helt fjerne de ubehagelige konsekvenser af den arvelige sygdom.



To T'er ved siden af hinanden i DNA-sekvensen kan reagere og binde til hinanden, hvis de udsættes for UV-lys eller Røntgen-stråling. Dette kan resultere i fejl, når DNA'et efterfølgende kopieres. Herved kan der opstå mutationer.

# DNA sekventering

**DNA**-sekvensbestemmelse gør brug af metoden dideoxysekventering, der er udviklet af nobelpristageren Frederic Sanger. Det er samme metode, der er udnyttet til bestemmelse af sekvensen af det menneskelige genom.

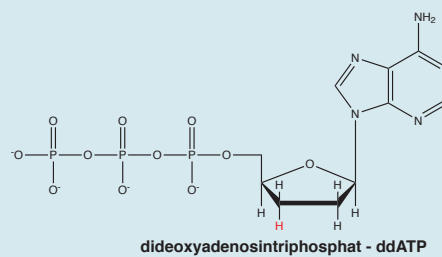
Metoden efterligner cellernes kopiering af DNA'et, idet man bruger enzymet DNA-polymerase.

Først syntetiseres et lille stykke DNA – en primer, der er ca. 20 nucleotider lang. Primeren kan base-parre til det område, der ønskes sekventeret (target-DNA). Der findes synteseapparater, der kan lave små stykker DNA, se figur 7.

Target-DNA'et opvarmes, hvorved samtlige hydrogenbindinger mellem baserne brydes, og DNA-strengene går fra hinanden. Herefter tilsættes primeren, og temperaturen sænkes (Fig. 7, A). Primeren finder hurtigt det sted på target-DNA'et, hvor den kan lave den optimale base-parring ved hydrogenbinding og danne et lille stykke dobbelthelix.

Efterfølgende tilsættes DNA-polymerase og de fire nucleotider (Fig. 7, B). Alle disse har tre fosfatgrupper (nucleotidtriphosphat). Polymerasen starter ved primeren og syntetiserer herfra den modsatte kopi af target-DNA'et. Polymerasen kan indbygge ca. 1.000 nucleotider pr. minut!

Frederic Sanger udviklede nogle modificerede nucleotider, som netop mangler den hydroxygruppe, der skal reagere med fosfatgruppen på det næste nucleotid i DNA-sekvensen. Disse modificerede nucleotider (dideoxynucleotidtriphosphat, ddNTP, se nedenstående figur) virker som stopklods ved DNA syntesen.



Hvis man tilsætter en lille smule ddATP til reaktionen fra før, vil syntesen af en del af kopierne standse ved et A.

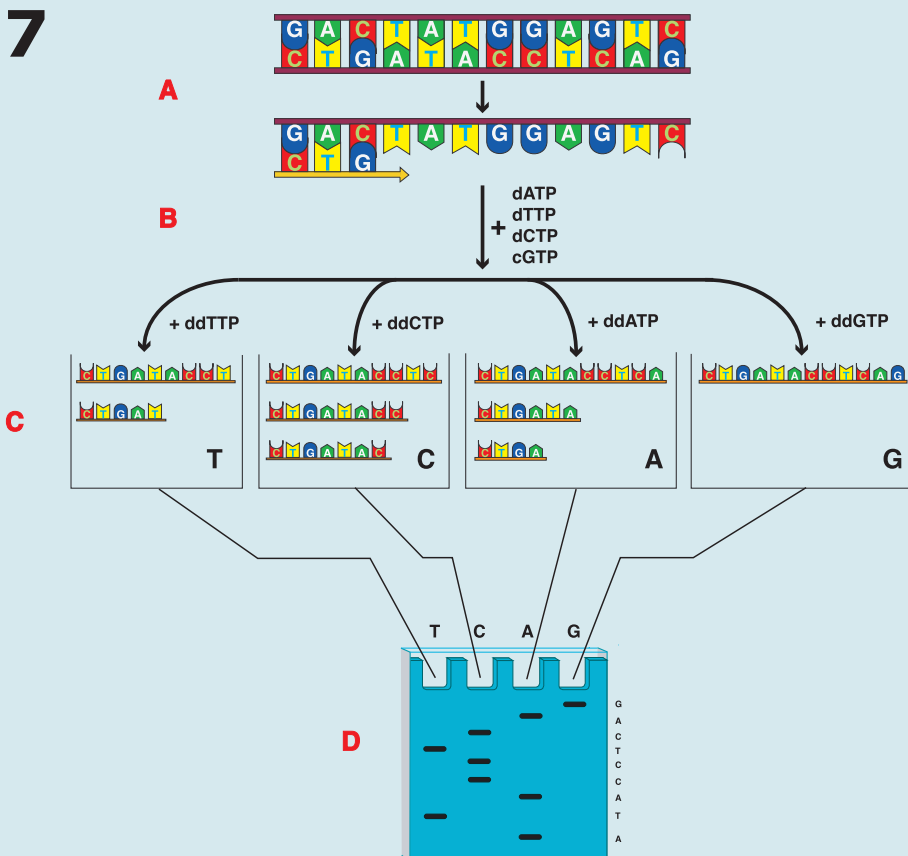
Deler man sin reaktion fra før i fire portioner og tilsætter hver portion en smule af én af de fire ddNTP'er, vil der i hver portion dannes fragmenter, som standser henholdsvis ved A, T, G og C.

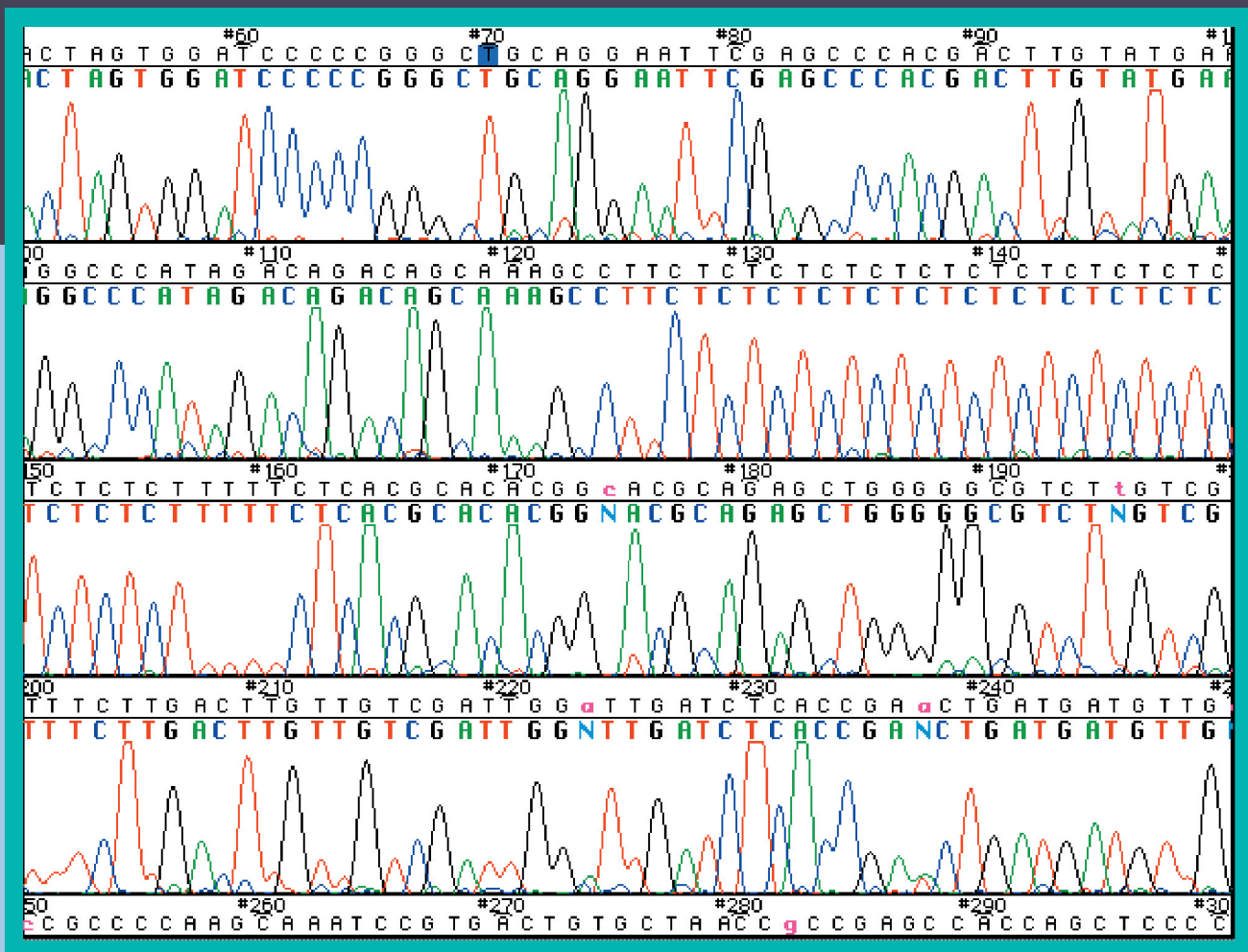
Da der syntetiseres mange kopier i hver portion, vil der i hver portion findes DNA-fragmenter af alle mulige længder, som dog alle standser på den base, der tilsat som dideoxystopklods (Figur 7, C).

Efterfølgende analyseres alle fragmenternes størrelse (længde) ved gel-elektroforese.

Denne analyse skiller fragmenterne efter størrelse, hvor de mindste bevæger sig hurtigst og dermed længst. Ved at placere de fire reaktionsprodukter ved siden af hinanden i gelen vil man efter elektroforesen se et mønster af bånd på gelen.

Hvert bånd repræsenterer et fragment, der standser ved en given base. Læses båndene fra den korteste ende, kan DNA-sekvensen aflæses (prøv om du kan læse sekvensen på Fig. 7, D).





Moderne metoder anvender fluorescerende ddNTP'er i fire forskellige farver. Alle fire reaktioner kan derfor laves på én gang og analyseres optisk (ovenfor kan du se, hvordan et resultat ser ud, efter at det er blevet analyseret af en computer).

Da det er muligt at sekventere mere end 1.000 baser pr. target-DNA og ved hjælp af robotter at analysere op mod 400 sekvens-reaktioner på én gang, kan man altså løse sekvensen af næsten en halv million baser pr. analyse!

## i PERSPEKTIV

Udgivet af Fysikforlaget med støtte fra Undervisningsministeriets tips/lottomidler og af Birch & Krogboe Fonden

Redaktion: Niels Elbrønd Hansen

Layout: Mette Qvistorff

Produktionsgruppe:  
Inge Kaufmann (fagredaktør),  
Kim Kusk Mortensen og  
Jakob Schiødt.

[www.perspektiv.fys.dk](http://www.perspektiv.fys.dk)

Tryk: Budolfi Tryk, Aalborg

Oplag: 9.000