

Det biologiske alfabet og cellens proteinfabrik – proteinsyntesen

Af cand.scient. Andreas Vedel

Hvis du får udleveret en kasse med 100.000 legoklodser hvor

- der kun findes 20 forskellige slags legoklodser i kassen,
- du skal bygge en masse forskellige legofigurer, og
- hver figur skal være bygget af mellem 100 og 1000 legoklodser og skal indeholde mindst 18 af de 20 forskellige klodser,

så har du en idé om hvordan biologiske celler fremstiller proteiner. I proteiner er “legoklodserne” dog nogle molekyler der hedder aminosyrer.

I denne artikel forklares hvordan biologiske celler vha. livets alfabet bygger proteiner, en proces der også kaldes proteinsyntesen.

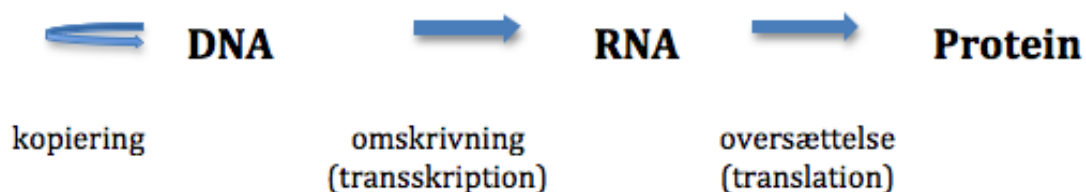
Prøv at forestille dig at du på en ferietur til London beslutter dig for at gå ind på en netcafe for at skrive en email eller opdatere din facebook-profil. Du sætter dig til rette ved tastaturet, men du har glemt at et britisk tastatur mangler de danske bogstaver æ, ø og å. Vil denne forglemmelse betyde at du ikke kan skrive? Nej selvfølgelig ikke, vel? Du vil blot omskrive et ø som “oe”, et å som “aa” og et æ som “ae”.

I alle biologiske celler foregår der en proces der minder lidt om ovenstående eksempel, men hvad er det for en proces, hvorfor foregår den og hvad er denne proces en del af? Der er her tale om den såkaldte *transskription* i cellers *proteinsyntese*, men for at forstå hvordan bogstavhistorien (og legoeksemplet) ovenfor hænger sammen med cellers proteinsyntese, så er vi nødt til at få styr på nogle begreber først.

Biologiens centrale dogme

Hver gang en celle har brug for at producere et hormon (f.eks. insulin, østrogen eller testosteron) eller et enzym (f.eks. et af de mange fordøjelsesenzymer i tarmen) eller et transportprotein (f.eks. hæmoglobin i blodets røde blodlegemer), så går der besked til cellens DNA om at bruge opskriften på det pågældende stof.

I oversigtsform kan processen, der også kaldes for biologiens centrale dogme, illustreres vha. oversigtsfiguren.



Oversigtsfigur

Biologiens centrale dogme. Biologiens centrale dogme omfatter processerne omskrivning og oversættelse.

Kopieringen yderst til venstre i figuren er strengt taget ikke en del af det centrale dogme, men er alligevel medtaget for at illustrere at DNA kan danne nyt DNA “ud fra sig selv” ved hjælp af det såkaldte baseparringsprincip eller “lynås-metoden”.

For at forstå detaljerne i oversigtsfiguren er det nødvendigt at uddybe hvad der gemmer sig bag *molekylerne* DNA, RNA og protein samt *processerne* “kopiering”, “omskrivning” og “oversættelse”.

Oversigtsfiguren antyder at et kemisk stof, DNA, indeholder et budskab, nemlig opskriften på et protein. Budskabet eller proteinopskriften, der ligger i cellens kerne, sendes i en lettere omskrevet udgave ud af cellens kerne vha. det kemiske stof RNA. Dernæst fabrikeres det ønskede protein ud fra opskriften som RNA har afleveret til proteinfabrikken (med navnet ribosom) der ligger udenfor cellens kerne, i det såkaldte cytoplasma.

Desuden antydes det yderst til venstre i figuren at DNA danner nyt DNA ud fra “sig selv”. Kopieringen foregår inden en celle deler sig og bliver til to celler, hvilket også nogle gange omtales som “lynlåsmetoden”. Denne kopiering sikrer at de 2 celler indeholder identiske kopier af DNA fra den oprindelige celle (se figur 2).

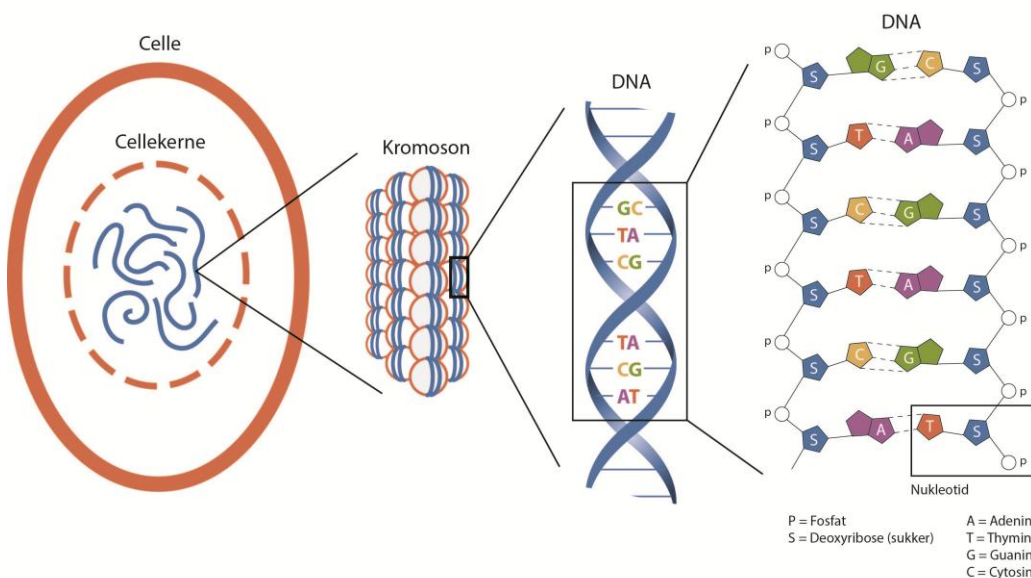
DNA

DNA (fra det engelske **D**eoxyribo**N**ucleic **A**cid der på dansk hedder deoxyribonukleinsyre) er et dobbeltstrengt kædeformet makromolekyle der udgør arvematerialet i alle levende organismer og i mange virus. I organismer med cellekerne(nukleus) findes DNA hovedsageligt i cellekernen, heraf navnet nukleinsyre.

DNA er opbygget af grundenheder, nukleotider, der som perler på en snor er koblet sammen i ofte meget lange dobbeltstrengede kæder. Nukleotiderne består af en sukkerdel, deoxyribose (S), hvortil der er bundet en fosfatgruppe (P) og én base (af fire mulige).

Variationen i DNA er bestemt af 4 forskellige baser: adenin (A), guanin (G), cytosin (C) eller thymin (T). Man omtaler også disse baser som de fire bogstaver i et såkaldt “DNA-alfabet”.

Sukkergrupperne er bundet sammen med fosfatgrupperne, og denne sukker-fosfat-kæde udgør “rygraden” i DNA-strengen. Se figur 1.



Figur 1

Placeringen af kromosomer i cellekernen er vist yderst til venstre. Der “zoomes ind” på DNA i cellekernen ved at bevæge sig fra venstre mod højre i figuren. Kromosomer er opbygget af sammenrullet DNA (og protein). Den kemiske opbygning af DNA er vist i de 2 figurer yderst til højre. De stiplede linjer mellem de 4 bogstaver symboliserer de såkaldte hydrogenbindinger. Disse bindinger er det svageste led mellem de 2 halvdele af det dobbeltstrengede DNA-molekyle. Når “DNA-lynlåsen” åbnes, sker det ved at disse hydrogenbindinger brister. Illustration: Andreas Ahlmann, (Ill. www.reesegrafisk.dk)

Replikation eller "DNA-lynlås-metoden"

DNA er altså opbygget af 2 modsat rettede kæder af nukleotider og det betyder at en base i den ene kæde kan danne nogle svage bindinger (såkaldte hydrogenbindinger) til en base i den anden kæde. Det er dog ikke tilfældigt hvilke baser der "passer til" hinanden, idet adenin (A) altid sidder over for thymin (T) mens guanin (G) altid sidder over for cytosin (C), se figur 1.

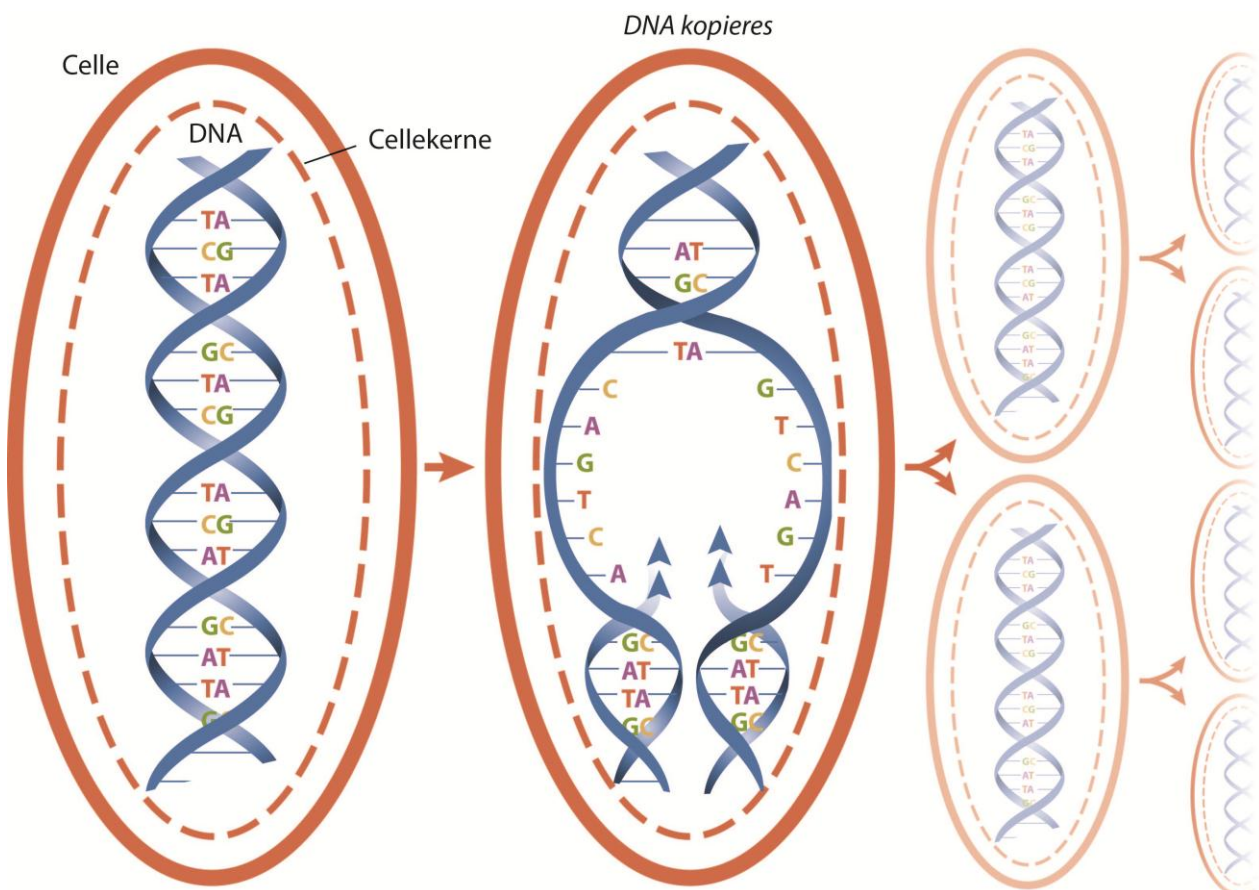
Alle mennesker er begyndt tilværelsen som én celle, nemlig den celle der opstod da en sædcelle befrugtede en ægcelle. En forudsætning for at vi er blevet til et individ opbygget af mange celler der kan læse denne tekst, er at der sket en masse celledelinger siden den første celle blev dannet (den såkaldte zygote). Når celler deler sig, så er det ekstremt vigtigt at hele cellens DNA videregives til de efterfølgende celler fordi alle gener, altså opskriften på organismens bygning, er skrevet i DNA.

Cellen sikrer at der sker en nøjagtig kopiering af DNA ved

- 1) at det dobbeltstrengede DNA-molekyle adskilles på langs, lidt ligesom når en lynlås lynes op,
- 2) og derefter indsættes der et A over for T, eller et C over for et G eller vice versa ... Se figur 2

Denne kopiering (kaldes også replikation) af DNA vha. hver af de gamle dele af "DNA-lynlåsen" sikrer altså at der dannes 2 nøjagtige kopier af den oprindelige DNA-streng.

Kopieringen (der i virkeligheden foretages af en del forskellige enzymer) sørger for at der indsættes hhv. et G over for C eller et T over for A. Denne kopiering kaldes også for baseparringsprincippet eller baseparringsreglen.



Figur 2

Kopiering (replikation) af en celleds DNA sker før cellen deler sig og bliver til to, hver med samme DNA. Kopieringen sker ved at DNA-strengen deler sig på langs ligesom en åbnet lynlås. Hver af de gamle dele af "DNA-lynlåsen" sikrer at der dannes 2 nøjagtige kopier af den oprindelige DNA-streng ved hjælp af baseparringsprincippet.

Illustration: Andreas Ahlmann, (Ill. www.reesegrafisk.dk)

RNA

RNA (fra det engelske **RiboNucleic Acid** der på dansk hedder ribonukleinsyre) adskiller sig fra DNA på tre væsentlige punkter.

- 1) Som navnet antyder, er sukkergruppen i RNA en ribose (R) i modsætning til en deoxyribose i DNA. Der er tale om en meget lille kemisk forskel på hhv. ribose og deoxyribose.
- 2) De enkelte grundenheder i RNA, nukleotiderne, er således opbygget af en fosfatgruppe (P), en ribose (i stedet for deoxyribose i DNA), hvortil er bundet en af fire forskellige baser: adenin (A), guanin (G), cytosin (C) eller uracil (U). RNA indeholder således basen uracil (U) i stedet for thymin (T), men de to baser er kemisk næsten ens. "RNA-alfabetet" adskiller sig altså fra "DNA-alfabetet" ved at én base (T) er udskiftet med en anden base (U).
- 3) Selvom RNA-molekylets kemiske opbygning minder meget om DNA-molekylets, er RNA, som hovedregel, enkeltstrengt, mens DNA er dobbeltstrengt. Se figur 3.

DNA	RNA
Dobbeltstrengt	Enkeltstrengt
Nukleotider opbygget af P, S og en baserne A, G, C eller T	Nukleotider opbygget af P, R og en baserne A, G, C eller U
A "passer sammen med" T G "passer sammen med" C	A "passer sammen med" U G "passer sammen med" C

Figur 3

Oversigt over forskelle og ligheder mellem DNA og RNA. Sammenlign evt. ovenstående med figur 1.

Omskrivning (transskription)

Når celler danner RNA, så omtales denne proces som en omskrivning eller transskription fra DNA til RNA. Omskrivningen til RNA er, som i tastatureksemplet vi indledte denne artikel med, en omskrivning af et budskab fra et alfabet til et andet alfabet, nemlig fra et "DNA-alfabet" til et "RNA-alfabet".

Ligesom man altså kan erstatte et "ø" med "oe", så kan T (i DNA) altså erstattes med et U (i RNA); men hvordan foregår denne omskrivning?

Det minder lidt om "lynlåsmetoden" som blev beskrevet ovenfor, idet DNA-molekylet igen "lynes op", men i dette tilfælde dannes der ikke en identisk kopi ud fra hver af de to halvdele af "DNA-lynlåsen", men kun ud fra den ene halvdel. Den halvdel af DNA der benyttes til at lave en omskrevet RNA-kopi, kaldes for skabelon-DNA. Processen er vist i oversigtsform nedenfor (figur 4).

Oversigt - Transskriptionen	
Kodende DNA	ATG TCT CGT TTC
Skabelon-DNA	TAC AGA GCA AAG
RNA	AUG UCU CGU UUC

Figur 4

Sammenhængen mellem sammensætningen af "bogstaver" i DNA ("kodende DNA") og det RNA som transkriptionsprocessen har omskrevet DNA til. Ved sammenligning ses at "det kodende DNA" svarer til RNA, bortset fra at T er udskiftet med U. For at sikre at bogstavsammensætningen i RNA er en omskrevet version af det kodende DNA, bruges den anden halvdel af DNA-strengen (det såkaldte "skabelon-DNA") som udgangspunkt for omskrivningen (transskriptionen).

Fidusen er at cellen benytter den ene del af "DNA-lynlåsen" til at lave en RNA-udgave. Processen sikrer altså at RNA (hvor T er erstattet med U) er en omskrevet kopi af den ene af de to DNA-halvdele. Derfor kaldes processen også en omskrivning eller en transskription.

Hver gang skabelon-DNA-delen indeholder et C, indsættes et G i RNA, og hver gang skabelon-DNA-delen indeholder et G, indsættes et C i RNA. Dette sikres ved at C, som nævnt tidligere, kun "passer sammen med" G.

Hver gang skabelon-DNA-delen indeholder et T indsættes et A i RNA, og hver gang skabelon-DNA-delen indeholder et A, indsættes et U i RNA. Dette sikres ved at A altid "passer sammen med" T og U altid "passer sammen med" A.

Ved at betragte figur 4 kan det konstateres at det dannede RNA er en kopi af det såkaldt kodende DNA, bortset fra at T er erstattet med U. Med "kodende DNA" understreges at den ene halvdel af den dobbelte DNA-streng indeholder en kode. Det viser sig nemlig at DNA indeholder en opskrift på sammensætningen eller rækkefølgen af de byggeklodser som proteiner er sammensat af.

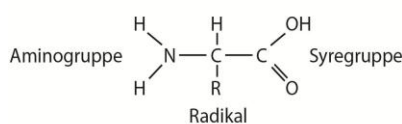
Proteiner og legoklodser

Den danske legetøjsgigant LEGO havde på et tidspunkt et reklameslogan der hed: "It's a new toy every day" (Det er et nyt stykke legetøj hver dag). Enhver der har leget med legoklodser, ved at man med ganske få forskellige slags legoklodser kan lave mange forskellige slags biler, huse, figurer etc.

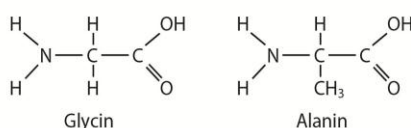
Når biologiske celler fremstiller proteiner, så minder det lidt om bygge med lego, men i proteiner er "legoklodserne" nogle kemiske stoffer, nemlig aminosyrer. Disse aminosyrer er nogle relativt små molekyler, og der findes 20 forskellige slags i cellen. Proteiner består således af en lang række af aminosyrer der er sat sammen i forlængelse af hinanden som perler på en snor. Forskellige proteiner er simpelthen forskellige rækkefølger, længder og kombinationer af de 20 forskellige aminosyrer. Se figur 5.

Ved at sammensætte disse 20 forskellige aminosyrer forskelligt kan der bygges en enorm mængde forskellige proteiner; og det er faktisk hvad en biologisk celle gør når den laver proteiner. Denne proces kaldes translation (se figur 6 samt nedenstående forklaring).

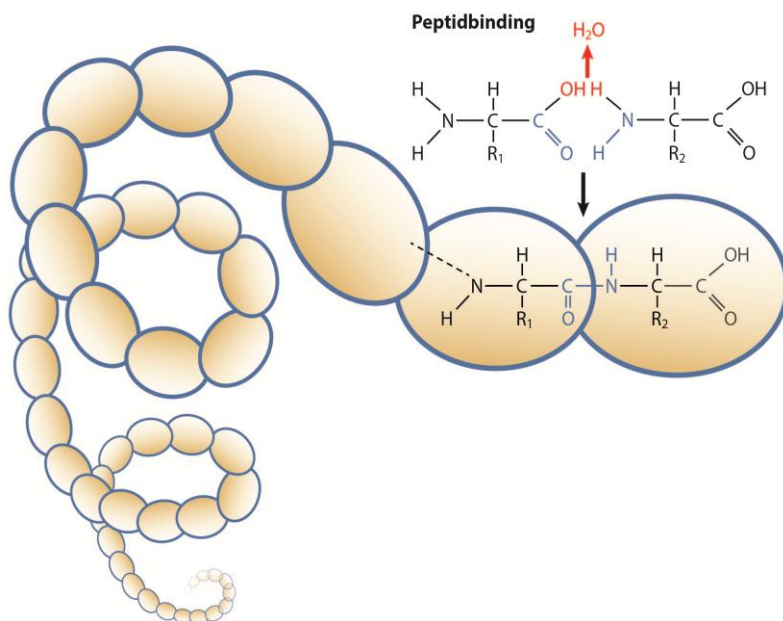
Aminosyre-struktur



Eksempler



Peptidbinding

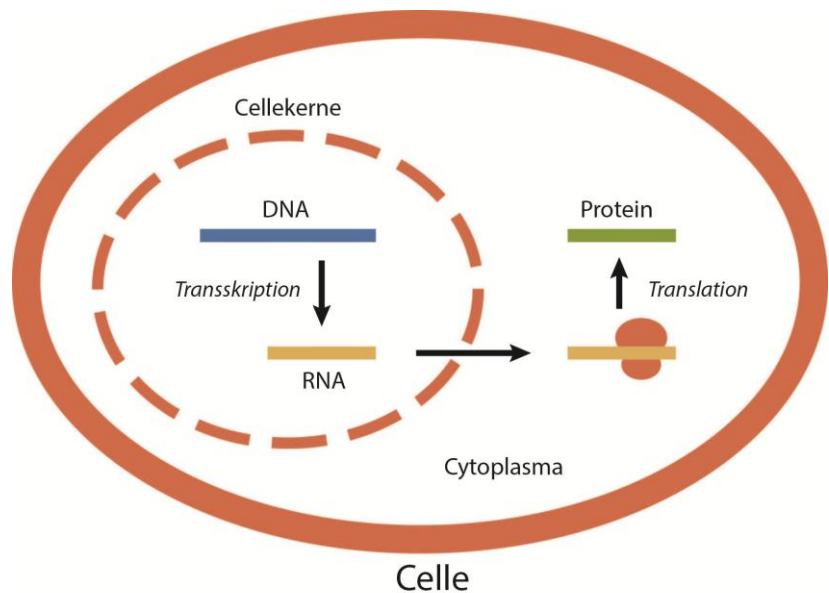


Figur 5

Proteiner er opbygget af molekylære "byggeklodser" der hedder aminosyrer. Navnet skyldes at aminosyrer indeholder både en aminogruppe (-NH₂) og en syregruppe (-COOH). Øverst til venstre i figuren er vist en generel formel for aminosyrers opbygning. Øverst til højre er vist 2 eksempler på aminosyrer.

Celler danner proteiner ved at syregruppen i en aminosyre hæftes sammen med amino-gruppen i en anden aminosyre. Sammenhæftningen mellem de to aminosyrer kaldes en peptidbinding (vist med blå). Ved dannelse af en peptidbinding fraspaltes et vandmolekyle (vist med rødt). Ved til stadighed at hæfte nye aminosyrer på én ad gangen (vha. peptidbindinger) kan cellen danne proteiner med den sammensætning og længde som tripletterne i DNA (og RNA) foreskriver. Se også figur 6, 7 og 8.

Illustration: Andreas Ahlmann, (Ill. www.reesegrafisk.dk)



Figur 6

Oversigtsfigur over proteinsyntesen. Først foretages en omskrivning (transskription) fra DNA til RNA i cellekernen. Når RNA er transporteret ud af cellekernen, sker oversættelsen (translationen) fra RNA til protein i cellens cytoplasma. Dette sker vha. et såkaldt ribosom der fungerer som både aflæsningsenhed og samlefabrik. Illustration: Andreas Ahlmann, (Ill. www.reeseografisk.dk)

Translation

Vi mangler nu kun at forklare hvordan cellen faktisk sammensætter det enkelte protein ud fra de 20 forskellige aminosyrer der er tilgængelig for cellen.

Hvordan kan DNA (og RNA) indeholde information om hvilke af de 20 forskellige aminosyrer (svarende til legoklodser) som skal indsættes i proteinets perlerække?

Eller mere præcist: Hvordan kan et alfabet der kun består af 4 "bogstaver" (baserne A, T, C og G i DNA og/eller A, U, C og G i RNA), skabe tusindvis af forskellige proteiner i vores celler?

Ved at sammensætte de 4 "bogstaver" i "ord" bestående af 3 "bogstaver" er det muligt at danne 64 forskellige kombinationer eller "ord". Hver af disse "ord", der også kaldes en triplet eller en codon, indeholder koden for én af de 20 forskellige aminosyrer. Det betyder til gengæld at flere forskellige tripletter kan kode for den samme aminosyre da der jo er 44 ekstra tripletter til rådighed udover de 20 der egentlig er brug for. Sammenhængen mellem RNA-tripletter og de enkelte byggeklodser, aminosyrer, er vist i figur 7.

I figur 6 ses en oversigtstabel med eksempler på både transskriptionen og translationen. Ved sammenligning med figur 4 ses at tabellen er udvidet med en ekstra vandret række nederst: en lille stump af et protein opbygget af fire forskellige aminosyrer methionin (met), serin (ser), arginin (arg) og phenylalanin (phe) svarende til de fire kodende DNA-tripletter.

Oversigt - Transskription og translation				
Kodende DNA	ATG	TCT	CGT	TTC
Skabelon DNA	TAC	AGA	GCA	AAG
RNA	AUG	UCU	CGU	UUC
Protein	met	ser	arg	phe

Figur 7

Ved sammenligning med figur 4 ses at tabellen er udvidet med en ekstra vandret række nederst: en lille stump af et protein opbygget af de fire aminosyrer methionin (met), serin (ser), arginin (arg) og phenylalanin (phe) svarende til de fire kodende DNA-tripletter. Sammenlign evt. RNA-tripletterne med oversigten i figur 8.

I figur 8 er vist sammenhængen mellem alle 64 forskellige RNA-tripletter og de 20 forskellige aminosyrer. Denne sammenhæng kaldes for den genetiske kode eller det genetiske alfabet.

1. bogstav	2. bogstav				3. bogstav
	U	C	A	G	
U	<i>phe</i>	<i>ser</i>	<i>tyr</i>	<i>cys</i>	U
U	<i>phe</i>	<i>ser</i>	<i>tyr</i>	<i>cys</i>	C
U	<i>leu</i>	<i>ser</i>	<i>stop</i>	<i>stop</i>	A
U	<i>leu</i>	<i>ser</i>	<i>stop</i>	<i>trp</i>	G
C	<i>leu</i>	<i>pro</i>	<i>his</i>	<i>arg</i>	U
C	<i>leu</i>	<i>pro</i>	<i>his</i>	<i>arg</i>	C
C	<i>leu</i>	<i>pro</i>	<i>gln</i>	<i>arg</i>	A
C	<i>leu</i>	<i>pro</i>	<i>gln</i>	<i>arg</i>	G
A	<i>ile</i>	<i>thr</i>	<i>asn</i>	<i>ser</i>	U
A	<i>ile</i>	<i>thr</i>	<i>asn</i>	<i>ser</i>	C
A	<i>ile</i>	<i>thr</i>	<i>lys</i>	<i>arg</i>	A
A	<i>met</i> <i>(start)</i>	<i>thr</i>	<i>lys</i>	<i>arg</i>	G
G	<i>val</i>	<i>ala</i>	<i>asp</i>	<i>gly</i>	U
G	<i>val</i>	<i>ala</i>	<i>asp</i>	<i>gly</i>	C
G	<i>val</i>	<i>ala</i>	<i>glu</i>	<i>gly</i>	A
G	<i>val</i>	<i>ala</i>	<i>glu</i>	<i>gly</i>	G

Figur 8 – Oversigt over det genetiske alfabet i RNA.

Hver af de 64 triletter koder for én bestemt aminosyre, men flere triletter kan kode for den samme aminosyre. Koden for aminosyren phe (øverste venstre hjørne af tabellen) skrives vha. tripletten UUU. Koden for aminosyren gly (nederste højre hjørne af tabellen) skrives vha. tripletten CGG. Bemærk at der optræder både en start-triplet og flere stop-tripletter. Når cellens aflæsningsværktøj (ribosomerne) skal begynde translationen, sker det ved start-tripletten, og tilsvarende stopper translationen når ribosomerne møder en stop-triplet.

Hvis der er interesse for at komme lidt mere i dybden med hvordan både replikation, transskription og translation foregår i cellen, så kan det anbefales at studere to pædagogiske animationer på <http://biokemibogen.dk/animationer/>

Fascinerende kompleksitet

Som det fremgår af ovenstående (og forhåbentligt pædagogiske) forklaring, så er cellers evne til at danne proteiner en på én gang både fascinerende og kompleks proces. Det er dog på sin plads at understrege at der er meget mere at sige om proteiner og kompleksiteten i celler, hvilket vi kort vil skitsere vha. tre eksempler.

Cellers tredimensionelle kompleksitet kan illustreres ved at sammenligne en celle med en storby. Med København som eksempel vil cellekernen fylde hele den indre by, og membranen rundt om hele cellen ville ligge ca. 10 km fra centrum. Et protein ville være på størrelse med

en personbil, og proteinets bevægelsesfrihed inde i cellen vil ikke være større end bilens er i København. Hertil skal lægges hele den øvrige infrastruktur (nedgravede kabler, kloaknet, metro- og togsinker etc.) samt alle menneskers transport og aktivitet i løbet af bare nogle få timer.

Menneskets forskellige celler producerer mere end 100.000 forskellige proteiner, og det sker vel at mærke vha. ca. 25.000 gener i de enkelte celler. En del af forklaringen på dette paradoks er bl.a. at samme gen, altså en bestemt rækkefølge af baser (basesekvens), kan resultere i forskellige proteiner ved at sammensætte mindre områder af genets basesekvens i forskellige kombinationer. Hertil kommer at gener kan være "slukket" eller "tændt" – eller mange forskellige tilstande mellem disse to muligheder. Dette lidt underlige faktum skyldes at generne er i stand til at reagere på det aktuelle miljø inden i og uden for cellen ved at opbygge og kontrollere en bred vifte af molekylære kontrolmekanismer, en slags miniature-kontrolcomputere.

Et er at danne den rigtige rækkefølge af aminosyrer i et protein vha. proteinsyntesen, men foldningen af proteinets tredimensionelle struktur skal styres, så de nydannede proteiner ikke kommer galt af sted. En løselig beregning kan vise at et mindre protein der består ca. 125 aminosyrer, kan danne mere end 10^{50} forskellige konformationer. Selv om man forestiller sig at molekylet kunne afprøve 10^{13} af disse muligheder pr. sekund, så ville det tage mere end 10^{30} år at nå igennem hovedparten. I cellen sker den korrekte foldning på ca. et minut vha. af bl.a. nogle molekylære "livvagter" der beskytter det nydannede protein mod uønsket indblanding fra andre molekyler i cellen eller uønsket tiltrækning mellem forskellige områder af proteinet.

Desuden afleveres det endnu ikke-foldede protein i særlige områder der sørger for at lede proteinet til det rigtige sted i cellen. Dette sker vha. særlige transportproteiner der er i stand til "at se" om det nydannede protein "hører hjemme" i det pågældende område af cellen.

Der skal altså en by af Københavns størrelse til for at få en stærkt underdrevet idé om en celled kompleksitet og dannelsen samt cellens håndtering af proteiner som altså er styret af proteiner, hvilket gør at kompleksiteten øges voldsomt.

Litteratur

Carey, Nessa: The epigenetics revolution (Icon books 2011)

Egebo, Lone Als: Genetikbogen – genetik, genteknologi og evolution (Nucleus 2003)

Hoffmeyer, Jesper: Biosemiotik – En afhandling om livets tegn og tegnenes liv (Forlaget Ries 2005).

Lennox, John C.: Guds bøddel (Origo 2012)

Torp, Kresten Cæsar: Biokemibogen – liv, funktion, molekyle (Nucleus 2010), se også

<http://biokemibogen.dk/animationer/>

Denne artikel findes i en journalistisk bearbejdet og tredelt artikel i bladet Origo 126, et temanummer om bioinformatik, marts 2013. Red.