

# ITBC BI1 Python 1 en Basiskennis voor Bioinformatici

## Weektaak 3

---

## OWE1 Python 1 en Basiskennis voor Bioinformatici

### Taak 3 – Analyse van de eiwitsequentie ten opzichte van de DNA sequentie

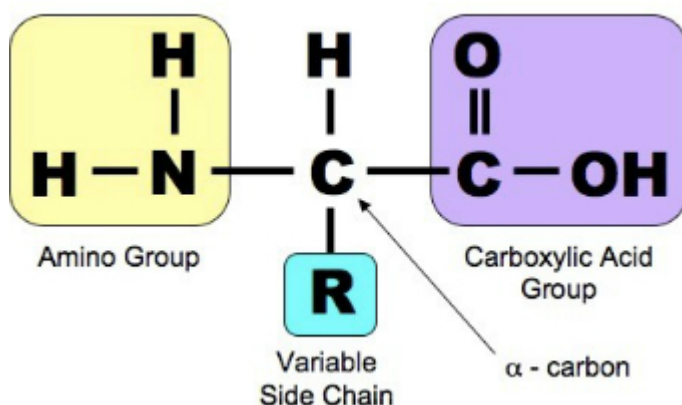
### Doelstellingen

Na afloop van deze taak kun je:

- In python eenvoudige scripts schrijven om eiwit sequentiedata te analyseren.
- De opzet van een script uitwerken in pseudocode.
- Het script testen.

### Casus

We hebben vorige week gekeken naar genoom en mRNA sequenties. Deze week gaan we door naar de eiwitsequenties. Het ribosoom leest de genetische code van het mRNA en koppelt voor elk codon een aminozuur aan de groeiende eiwitketen. De genetische code bestaat uit codons, steeds setjes van 3 letters mRNA. De 64 ( $3^4$ ) verschillende codons coderen voor 20 aminozuren. Deze aminozuren zijn opgebouwd uit een aminogroep, een centraal koolstofatoom met daaraan een restgroep en een carboxylgroep.



Afbeelding: Aminozuur structuur. <http://study.com/academy/lesson/how-amino-acids-form-protein.html>, geraadpleegd op 8-6-2017.

Het eerste codon noemen we het startcodon. Dit codon (ATG) codeert voor het aminozuur Methionine. Het laatste codon van noemen we het stopcodon. Dat codon codeert niet voor een aminozuur, maar het is een teken voor het ribosoom om het proces van aminozuren koppelen te stoppen. Als het eiwit klaar is, lezen we het eiwit van de N-terminus naar de C-terminus.

De 20 aminozuren hebben allemaal een andere restgroep, waardoor elk aminozuur zijn eigen specifieke eigenschappen heeft. Een voorbeeld van een eigenschap is lading. De aminozuren arginine (R) en Lysine (K) hebben een positieve lading in de cel bij een pH van ongeveer 7. De aminozuren asparaginezuur (D) en glutaminezuur (E) hebben dan juist een negatieve lading.

## Opdracht: Python scripts voor analyse van eiwit sequenties

Het ribosoom vertaalt de mRNA sequentie naar eiwit, maar dat doet hij niet zomaar. Vergelijk de lengte van de eiwit sequentie met de lengte van de mRNA sequentie.

- Maak een opzet voor je script in pseudocode
- Schrijf een python script dat:
  - toetst of de invoer sequentie DNA of eiwit is.
  - de lengte van een eiwit sequentie of van een mRNA sequentie bepaald.
  - de lengte van een eiwit sequentie en een mRNA sequentie met elkaar vergelijkt.
  - het aantal aminozuren D, E, R, K in de eiwitsequentie telt.
- Test je script.
- Sla je script remote op in GIT-classroom group assignment.

## Resultaat

Aan het eind van deze taak heb je

- een python script dat eiwit sequenties of mRNA sequenties kan analyseren.
- een opzet van het script in pseudocode.
- een vergelijking van de lengte van de mRNA sequenties ten opzichte van de bijbehorende eiwitsequenties van de 5 organismen.
- een vergelijking van de hoeveelheid D, E, R, K, aminozuren in de eiwitsequenties van de 5 organismen ten opzichte van de netto lading van het eiwit.
- alle data en vergelijkingen overzichtelijk opgeslagen in Excel.
- de python scripts opgeslagen in GIT-classroom.

# Presentatie

Presenteer je resultaten aan de tutorgroep. De tutor zal per projectgroep een student aanwijzen om het woord te doen. Van iedere toehoorder/student wordt verwacht dat hij kritische vragen stelt ten aanzien van de inhoud en/of de vorm van de presentatie.

Vergeet niet de bronvermelding in je presentatie (zie APA richtlijnen onder Algemene informatie op OnderwijsOnline).

## Middelen

- Campbell et al. Biology, a global approach. Pearson, 11<sup>th</sup> edition, 2017, concept 5.4 blz 123-126.

## Nabespreking

In de nabespreking en tijdens de presentaties komen de volgende vragen aan de orde:

- Hoe kan je de gevonden resultaten verklaren?
- Wat bleek moeilijker dan wel makkelijker dan verwacht?
- Wat zou je de volgende keer anders doen?
- Hoe verliep de samenwerking / communicatie?

Copyright @ HAN, I. Paffen