

Aufgabe 2:

Human Hemoglobin subunit alpha:

MVLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHFDLS-
HGSAQVKGHG KKVADALTNVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLS-
HCLLVTLAAHLPAEFTP AVHASLDKFLASVSTVLTSKYR

Human Hemoglobin subunit beta:

MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPDA-
VMGNPK VKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLL-
GNVLVCVLAHHFG KEFTPPVQAAYQKVVAGVANALAHKYH

Aufgabe 3:

- 1) Globales Alignment vergleicht die vollständigen Sequenzen miteinander, während beim lokalen Alignment nur Teile der Sequenzen verglichen werden.
- 2) Das Globale Alignment wird genutzt, wenn die Sequenzen die gleiche Länge und starke Sequenzhomologien aufweisen. Beim Lokalen Alignment ist dies nicht notwendig.

Aufgabe 4:

Globales Alignment mit voreingestellten Parametern:

Matrix: EBLOSUM62

Gap_penalty: 10.0

Extend_penalty: 0.5

Length: 149

Identity: 65/149 (43.6%)

Similarity: 90/149 (60.4%)

Gaps: 9/149 (6.0%)

Score: 292.5

Globales Alignment mit einer anderen GAP OPEN penalty:

Matrix: EBLOSUM62

Gap_penalty: 20.0

Extend_penalty: 0.5

Length: 149

Identity: 61/149 (40.9%)

Similarity: 87/149 (58.4%)

Gaps: 9/149 (6.0%)

Score: 270.0

```
#=====
EMBOSS_001      1 -MVLSPADKTNVKAAGKVGAGHAGEYGAEALERMFLSFPTTKTYFPHF-- 47
                  :.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
EMBOSS_001      1 MVHLTPEEKSAVTALWGKV--NVDEVGGEALGRLLVVYPWTQRFFESFGD 48

EMBOSS_001     48 ----DLSHGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLR 93
                  |...|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
EMBOSS_001     49 LSTPDAMGPNPKVKAHGKKVLGAFSDGLAHLDNLKGTFTLSELHCDKLH 98

EMBOSS_001     94 VDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR 142
                  |||.||:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
EMBOSS_001     99 VDPENFRLLGNVLVCLAHFHGKEFTPPVQAAYQKVVAGVANALAHKYH 147

#-----
#-----
```

Abbildung 3: Globales Alignment mit einer anderen GAP OPEN penalty.
(https://www.ebi.ac.uk/Tools/services/web/toolresult.ebi?jobId=emboss_needle-I20180710-143055-0087-10527696-p1m, 10.07.18, 15:34 Uhr)

Lokales Alignment mit voreingestellten Parameter:

Matrix: EBLOSUM62

Gap_penalty: 10.0

Extend_penalty: 0.5

Length: 145

Identity: 63/145 (43.4%)

Similarity: 88/145 (60.7%)

Gaps: 8/145 (5.5%)

Score: 293.5

```

#
#=====

EMBOSS_001      3  LSPADKTNVKAANGKVGAGHAGEYGAEALERMFLSFPTTKTYFPHF-DLS-      50
                  |:|.:|:|.|.|||  :..|.|.|||.|:~.:|.|:~.:|.|||
EMBOSS_001      4  LTPEEKSAVTALWGKV--NVDEVGGEALGRLLVVYPWTQRFFESFGDLST      51

EMBOSS_001      51  ---HGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSSDLHAHKLRVDP      96
                  .|:~.:|:|.|||~.:|.|:~.:|:~.:~.:|.|:~.:|.|||
EMBOSS_001      52  PDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFTATLSELHCDKLHVDP      101

EMBOSS_001      97  VNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTISKY      141
                  .||:|:|.~.:|:~.:|:~.:|:~.:|:~.:|.|.~.:|.~.:|.~.:|
EMBOSS_001     102  ENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKY      146

#-----
#-----

```

Abbildung 4: Lokales Alignment mit voreingestellten Parameter.
(https://www.ebi.ac.uk/Tools/services/web/toolresult.ebi?jobId=embooss_water-120180710-143631-0511-31291125-p1m, 10.07.18, 15:59)

Vergleich: Vergleicht man die Matrizes mit denen aus der Vorlesungsfolie 9, so fällt auf, dass auf den Vorlesungsfolien eine höhere Übereinstimmung vorzufinden ist. Dies wurde durch Einfügen von Gaps erreicht.

BLOSUM62: Unter einer BLOSUM versteht man eine Substitutionsmatrix, die verwendet wird um Sequenzen zu vergleichen. BLOSUM mit höheren Nummern werden verwendet, um ähnliche Sequenzen miteinander zu vergleichen, wohingegen die mit niedrigeren Zahlen verwendet werden, um stark unterschiedliche Sequenzen zu vergleichen. Die BLOSUM62 wird so genannt, da alle Sequenzen mit mindestens 62 % gleicher Sequenzidentität als eine Sequenz dargestellt werden.

PAM30: Unter einer PAM30 versteht man, die berechnete Substitutionsmatrix für zwei Sequenzen die 30 PAM (1 PAM: Einheit für den Abstand von Proteinsequenzen → eine Sequenz wurde in eine andere Sequenz überführt mit X Punktmutationen pro 100 AS) entfernt sind. Dabei muss darauf geachtet werden, dass damit nicht 30 Punktmutationen pro 100 AS gemeint sind, da Doppelmutationen, Rückmutationen etc. mit beachtet werden müssen.

Gap Penalty: Verringert den Score beim Einfügen von Gaps. Ist die Gap Penalty höher wird der Score schlechter, da höhere negative Zahlen einberechnet werden. Deswegen ist der Score in Abbildung 3 auch wesentlich schlechter als der in Abbildung 1.

Quellen:

- <http://www.uky.edu/Classes/BIO/520/BIO520WWW/blosum62.htm>
(10.07.18, 16:45 Uhr)
- [file:///C:/Users/hanna/AppData/Local/Packages/Microsoft.MicrosoftEdge_8wekyb3d8bbwe/TempState/Downloads/05_pam%20\(1\).pdf](file:///C:/Users/hanna/AppData/Local/Packages/Microsoft.MicrosoftEdge_8wekyb3d8bbwe/TempState/Downloads/05_pam%20(1).pdf)
(10.07.18, 17:00 Uhr)