Aufgabe 2

HBA_HUMAN Hemoglobin subunit alpha (P69905)

```
MVLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHG
KKVADALTNAVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTP
AVHASLDKFLASVSTVLTSKYR
```

HBB_Human Hemoglobin subunit beta (P68871)

```
MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPK
VKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHFG
KEFTPPVQAAYQKVVAGVANALAHKYH
```

Aufgabe 3

- Globales Alignment: alle Sequenzen werden gleichzeitig komplett betrachtet und vergliechen. Die Sequenzen müssen in etwa gleicher Länge sein. Ein globales Alignment wäre nur sinnvoll, wenn die betrachtenden Sequenzen einander über ihre gesamte Länge ähnlich sind.

- Lokales Alignment: es werden nur Teilsequenzen betrachtet. In jeder Sequenz wird nach dem am besten zum anderen Sequenz passenden Sequenzstück gesucht und nur diese Teilstücke werden aligniert. Die zu alignierenden Sequenzen müssen nicht über ihre ganze Länge ähnlich sein.

Aufgabe 4A

4.1a. globals alignment mit voreingestellten Parametern

```
 Program: needle
# Rundate: Tue 10 Jul 2018 19:30:55
# Commandline: needle
#    -auto
#    -stdout
#    -asequence emboss_needle-I20180710-193053-0816-52163436-p2m.asequence
#    -bsequence emboss_needle-I20180710-193053-0816-52163436-p2m.bsequence
#    -datafile EBLOSUM62
#    -gapopen 10.0
#    -gapextend 0.5
#    -endopen 10.0
#    -endextend 0.5
#    -aformat3 pair
#    -sprotein1
#    -sprotein2
# Align_format: pair
# Report_file: stdout
########################################

#=======================================
#
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# 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
#
#
#=======================================

EMBOSS_001         1 MV-LSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHF-D
48
                     || |:|.:|:.|.|.||||  :..|.|.|||.|:.:.:|.|:.:|..| |
EMBOSS_001         1 MVHLTPEEKSAVTALWGKV--NVDEVGGEALGRLLVVYPWTQRFFESFGD
48

EMBOSS_001        49 LS-----HGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLR
93
                     ||     .|:.:||.|||||..|.::.:||:|::....:.||:||..||.
EMBOSS_001        49 LSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLH
98

EMBOSS_001        94 VDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR    142
                     |||.||:||.:.:|:..||.|...||||.|.|:..|.:|.|:..|..||.
EMBOSS_001        99 VDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH    147
```

4. 2a. Globales Alignment mit BLOSUM90 Matrix

```
Program: needle
# Rundate: Tue 10 Jul 2018 19:16:33
# Commandline: needle
#    -auto
#    -stdout
#    -asequence emboss_needle-I20180710-191630-0641-7037160-p1m.asequence
#    -bsequence emboss_needle-I20180710-191630-0641-7037160-p1m.bsequence
#    -datafile EBLOSUM90
#    -gapopen 10.0
#    -gapextend 0.5
#    -endopen 10.0
#    -endextend 0.5
#    -aformat3 pair
#    -sprotein1
#    -sprotein2
# Align_format: pair
# Report_file: stdout
########################################

#=======================================
#
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# Matrix: EBLOSUM90
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 149
# Identity:      65/149 (43.6%)
# Similarity:    83/149 (55.7%)
# Gaps:           9/149 ( 6.0%)
# Score: 311.5
#
#
#=======================================
```

```
EMBOSS_001         1 MV-LSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHF-D
48
                     || |:|.:|:.|.|.||||  ...|.|.|||.|:...:|.|:.:|..| |
EMBOSS_001         1 MVHLTPEEKSAVTALWGKV--NVDEVGGEALGRLLVVYPWTQRFFESFGD
48

EMBOSS_001        49 LSH-----GSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLR
93
                     ||.     |..:||.|||||..|.::..||.|::....:.||:||..||.
EMBOSS_001        49 LSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLH
98

EMBOSS_001        94 VDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR    142
                     |||.||:||...|...||.|...||||.|:|:..|..|.|:..|..||.
EMBOSS_001        99 VDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH    147
```

4. 3a. globales Alignment mit anderer GAP OPEN (hier: 100) und BLOSUM62-Matrix

```
Program: needle
# Rundate: Tue 10 Jul 2018 19:20:06
# Commandline: needle
#    -auto
#    -stdout
#    -asequence emboss_needle-I20180710-192004-0122-93440791-p2m.asequence
#    -bsequence emboss_needle-I20180710-192004-0122-93440791-p2m.bsequence
#    -datafile EBLOSUM62
#    -gapopen 100.0
#    -gapextend 0.5
#    -endopen 10.0
#    -endextend 0.5
#    -aformat3 pair
#    -sprotein1
#    -sprotein2
# Align_format: pair
# Report_file: stdout
########################################

#=======================================
#
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# Matrix: EBLOSUM62
# Gap_penalty: 100.0
# Extend_penalty: 0.5
#
# Length: 147
# Identity:      44/147 (29.9%)
# Similarity:    62/147 (42.2%)
# Gaps:           5/147 ( 3.4%)
# Score: 146.0
#
#
#=======================================

EMBOSS_001         1 -----MVLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFP
45
                           .....|.....................:......|..:.....
EMBOSS_001         1 MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLS
50
```

```
EMBOSS_001         46 HFDLSHGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLRVD
95
                      ..|...|:.:||.|||||..|.::.:||:|::....:.||:||..||.||
EMBOSS_001         51 TPDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVD
100

EMBOSS_001         96 PVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR    142
                      |.||:||.:.|:..||.|...||||.|.|:..|.:|.|:..|..||.
EMBOSS_001        101 PENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH    147
```

globules Alignment hit GAP OPEN (hier: 100) und BLOSUM90

```
Program: needle
# Rundate: Tue 10 Jul 2018 19:22:49
# Commandline: needle
#    -auto
#    -stdout
#    -asequence emboss_needle-I20180710-192247-0717-38015925-p1m.asequence
#    -bsequence emboss_needle-I20180710-192247-0717-38015925-p1m.bsequence
#    -datafile EBLOSUM90
#    -gapopen 100.0
#    -gapextend 0.5
#    -endopen 10.0
#    -endextend 0.5
#    -aformat3 pair
#    -sprotein1
#    -sprotein2
# Align_format: pair
# Report_file: stdout
########################################

#=======================================
#
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# Matrix: EBLOSUM90
# Gap_penalty: 100.0
# Extend_penalty: 0.5
#
# Length: 148
# Identity:      52/148 (35.1%)
# Similarity:    67/148 (45.3%)
# Gaps:           7/148 ( 4.7%)
# Score: 134.5
#
#
#=======================================

EMBOSS_001          1 MVLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHF---
47
                        ......................|.|.|||.|:...:|.|:.:|..|
EMBOSS_001          1 -MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDL
49

EMBOSS_001         48 ---DLSHGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLRV
94
                         |...|..:||.|||||..|.::..||.|::....:.||:||..||.|
EMBOSS_001         50 STPDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHV
```

```
99

EMBOSS_001        95 DPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR    142
                     ||.||:||...|...||.|...||||.|:|:..|..|.|:..|..||.
EMBOSS_001       100 DPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH    14
```

4. 4a. locales Alignment hit voreingestellten Parametern

```
Program: water
# Rundate: Tue 10 Jul 2018 19:26:37
# Commandline: water
#    -auto
#    -stdout
#    -asequence emboss_water-I20180710-192536-0024-52845378-p1m.asequence
#    -bsequence emboss_water-I20180710-192536-0024-52845378-p1m.bsequence
#    -datafile EBLOSUM62
#    -gapopen 10.0
#    -gapextend 0.5
#    -aformat3 pair
#    -sprotein1
#    -sprotein2
# Align_format: pair
# Report_file: stdout
########################################

#=======================================
#
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# 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 LSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHF-DLS-
50
                     |:|.:|:.|.|.||||  :..|.|.|||.|:.:.:|.|:.:|..| |||
EMBOSS_001         4 LTPEEKSAVTALWGKV--NVDEVGGEALGRLLVVYPWTQRFFESFGDLST
51

EMBOSS_001        51 ----HGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLRVDP
96
                         .|:.:||.|||||..|.::.:||:|::....:.||:||..||.|||
EMBOSS_001        52 PDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDP
101

EMBOSS_001        97 VNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKY    141
                     .||:||.:.|:..||.|...||||.|.|:..|.:|.|:..|..||
EMBOSS_001       102 ENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKY    146
```