

## Übungszettel 6

### 2. Proteinsequenzen:

Human Hemoglobin subunit alpha:

MVLSPADKTN VKAAWGKVG A HAGEYGAEALERMFLSFPTTKTYFPHFDLS  
HGSAQVKGHG KKVADALTNA VAHVDDMPNA LSALSDLHAH KLRVDPVNFK  
LLSHCLLVTL AAHLPAEFTP AVHASLKD KFLASVSTVLTSK YR

Human Hemoglobin subunit beta:

MVHLTPEEKS AVTALWGKVN VDEVGGEALG RLLVVYPWTQ RFFESFGDLS  
TPDAVMGNPK VKAHGKKVLG AFSDGLAHL DNLKGT FATLS ELHCDKLHVD  
PENFRLLGNV LVCVLAHHFG KEFTPPVQAA YQKVVAGVAN ALAHKYH

### 3. Unterschiede lokales und globales Alignment:

- bei globalem Alignment Vergleich der ganzen Sequenzen, bei lokalem Alignment nur Vergleich von Teilen der Sequenzen
- globales Alignment sinnvoll bei Sequenzen ähnlicher Länge und vermutlich hoher Übereinstimmung, lokales Alignment kann immer genutzt werden

#### 4.1 Globales Alignment mit voreingestellten Parametern:

Parameter:

```

EMBOSS_001      1 MV-LSPADKTNVKAANGKVGAAHAGEYGAEALERMFLSFPTTKTYFPHF-D      48
                  || |:|.:.:|.|.||| | :..|.|.|||.|:~::~|.|:~::~|..| |
EMBOSS_001      1 MVHLTPEEKSAVTALWGKV--NVDEVGGEALGRLLVVYPWTQRFFESFGD      48

EMBOSS_001     49 LS-----HGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLR      93
                  ||      .|:~::~|.|.||| |..|.~::~:| |:|:~::~~::~~::~|.|:~::~|..||.
EMBOSS_001     49 LSTPDAVMGNPKVKAHGKKVLGAFSDGLAHL DNLKGT FATLS ELHCDKLH      98

EMBOSS_001     94 VDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLKD KFLASVSTVLTSKYR      142
                  |||.||:~::~|.:.|:~::~|.|.~::~|||.|.|.~::~|.|.~::~|.|.~::~|.|.~::~|.
EMBOSS_001     99 VDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH      147

```

Globales Alignment mit voreingestellten Parametern ([https://www.ebi.ac.uk/Tools/services/web/toolresult.ebi?jobId=emboss\\_needle-l20180709-081644-0372-9135947-p1m](https://www.ebi.ac.uk/Tools/services/web/toolresult.ebi?jobId=emboss_needle-l20180709-081644-0372-9135947-p1m))

Matrix: EBLOSUM62

Gap penalty: 10.0

Identity: 65/149 (43.6%)

Similarity: 89/149 (59.7%)

Gaps: 9/149 (6.0%)

Score: 292.5

**4.2 Globales Alignment mit einer anderen Substitution Matrix:**

```

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

EMBOSS_001     49 LSH-----GSAQVKGHGKGVADALTNAVAHVDDMPNALSALSDLHAHKLR      93
                  ||.      |:.|.|.||| |..|.:.|.|.|.|:....|.|.:.|..| |:
EMBOSS_001     49 LSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLNLTGKTFATLSSEHCDKLH      98

EMBOSS_001     94 VDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTISKYR      142
                  |||.||:|:|.:.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
EMBOSS_001     99 VDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH      147

```

Globales Alignment mit einer anderen Substitution Matrix ([https://www.ebi.ac.uk/Tools/services/web/toolresult.ebi?jobId=embooss\\_needle-I20180709-082311-0142-63317176-p1m](https://www.ebi.ac.uk/Tools/services/web/toolresult.ebi?jobId=embooss_needle-I20180709-082311-0142-63317176-p1m))

**Parameter:**

Matrix: EPAM90

Gap penalty: 10.0

Identity: 65/149 (43.6%)

Similarity: 89/149 (59.7%)

Gaps: 9/149 (6.0%)

Score: 321.5

**4.3 Globales Alignment mit einer anderen GAP OPEN penalty:**

```

EMBOSS_001      1 -MVLSPADKTNVKAANGKVGAGAHAGEYGAEALERMFLSFPTTKTYFPHF--      47
                  :. |:|:|.:.|.|.|.||| | :..|.|.|||.|:....|.|.:.|..| |
EMBOSS_001      1 MVHLTPEEKSAVTALWGKV--NVDEVGGEALGRLLVVYPWTQRFFESFGD      48

EMBOSS_001     48 ----DLSHGSAQVKGHGKGVADALTNAVAHVDDMPNALSALSDLHAHKLR      93
                  |...|:|.|.|.||| |..|.:.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
EMBOSS_001     49 LSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLNLTGKTFATLSSEHCDKLH      98

EMBOSS_001     94 VDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTISKYR      142
                  |||.||:|:|.:.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
EMBOSS_001     99 VDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH      147

```

Globales Alignment mit einer anderen GAP OPEN penalty ([https://www.ebi.ac.uk/Tools/services/web/toolresult.ebi?jobId=embooss\\_needle-I20180709-082753-0775-83853101-p1m](https://www.ebi.ac.uk/Tools/services/web/toolresult.ebi?jobId=embooss_needle-I20180709-082753-0775-83853101-p1m))

**Parameter:**

Matrix: EBLOSUM62

Gap penalty: 25.0

Identity: 61/149 (40.9%)

Similarity: 87/149 (58.4%)

Gaps: 9/149 (6.0%)

Score: 260.0

#### 4.4 Lokales Alignment mit voreingestellten Parametern:

```

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

EMBOSS_001     51  ----HGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDDLHAHKLVRDP      96
                   .|:|.:.|.|.|||  |..|.:.:.|:|.:.:.:.|.|.:.|.|.  |||
EMBOSS_001     52  PDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFTATLSELHCDKLHVDP     101

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

```

Lokales Alignment mit voreingestellten Parametern ([https://www.ebi.ac.uk/Tools/services/web/toolresult.ebi?jobId=embooss\\_water-I20180709-083518-0704-97922751-p1m](https://www.ebi.ac.uk/Tools/services/web/toolresult.ebi?jobId=embooss_water-I20180709-083518-0704-97922751-p1m))

##### Parameter:

Matrix: EBLOSUM62

Gap penalty: 10.0

Identity: 63/145 (43.4%)

Similarity: 88/145 (60.7%)

Gaps: 8/145 (5.5%)

Score: 293.5

##### Vergleich mit Alignment aus Vorlesung:

Ein Vergleich mit dem, in der Vorlesung gezeigten Alignment ist schwierig, da dort multiple Sequenzen verglichen wurden und nicht nur die zwei gegebenen. Allerdings ist zu sehen, dass bei dem Alignment der Vorlesung wesentlich mehr Gaps eingefügt wurden.

**Gap Penalty:** Strafe, die einer Gap bei einem Alignment von Sequenzen zugeteilt wird.

**BLOSUM62 Matrix:** Bei der BLOSUM 62 Matrix werden alle Sequenzen mit, in diesem Fall, mehr als 62% Sequenzidentität zu einer Sequenz zusammengefasst.

**PAM90 Matrix:** Berechnung der Wahrscheinlichkeitswerte für die Matrix:  $M_n = (M_1)^n$

##### Quellen:

[https://www.informatik.hu-berlin.de/de/forschung/gebiete/wbi/teaching/archive/ws0506/bioinformatik-bpi/05\\_pam.pdf](https://www.informatik.hu-berlin.de/de/forschung/gebiete/wbi/teaching/archive/ws0506/bioinformatik-bpi/05_pam.pdf) (09.07.2018; 20:48 Uhr)