

### DNA-Klonierung und *in silico* Analysen

Diese Übung besteht aus zwei Teilen.

Im **Teil A** soll über eine klassische und einfache Methode ein Gen in *Escherichia coli* umklont werden: aus einem "high-copy" Plasmid (pJET1.2), in die Erstklonierungen üblicherweise erfolgen, soll ein "low-copy" Plasmid (pGEX-6P-1) zur heterologen Expression eines Proteins hergestellt werden. Das umzuklonierende Gen kodiert für die Dihydroflavonol-4-reduktase (DFR) (siehe Übung 5) und liegt in 2 Varianten vor, die sich auf DNA und Aminosäureebene unterscheiden. Das Beispiel umfasst alle klassischen Klonierungsschritte: präparativer Restriktionsverdau, Agarosegelelektrophorese, Gelextraktion, Ligation, Transformation von *E. coli*, Selektion, Plasmidpräparation, analytischer Restriktionsverdau und DNA-Sequenzierung. Letzteres erlaubt Ihnen anzugeben, welches *dfr* Gen Sie umklont haben.

**Teil B** beschäftigt sich mit der *in silico* Konstruktion eines Expressionsvektor. Beginnend mit der *in silico* Analyse eines genomischen DNA Fragmentes, werden Sie über die Identifizierung von Regionen mit Aminosäuresequenzähnlichkeiten zwischen ihrer Sequenz und öffentlich zugänglichen Sequenzen, die kodierende Regionen in ihrer DNA Sequenz lokalisieren. Sie werden dann ihre DNA Sequenz in eine cDNA Sequenz umwandeln und Oligonukleotide zur Amplifikation dieser cDNA designen.

### Hinweise:

\* wird von dem/r Betreuer/in ausgegeben

(A) befindet sich im Abzug für die Agarosegelelektrophorese

(B) befindet sich am Betreuertisch

(E) befindet sich im allgemeinen Entnahmebereich auf RT

(G) befindet sich in der Gefrierlade

(K) befindet sich im Kühlschrank

alle weiteren Materialien/Lösungen sind selbst herzustellen oder am Hörer/innenplatz vorhanden

Gerahmte Teile: es ist unter sterilen Bedingungen zu arbeiten!

Doppelt gerahmte Teile: es ist im Abzug zu arbeiten!

*Die Plasmidkarten sind am Ende dieser Vorschrift abgebildet.*

*Die Sequenz der dfr Gene sind im TISS verfügbar.*

### Abkürzungen:

amp, Ampicillin

BLAST, Basic Local Alignment Search Tool

MCS, Multiple cloning site

min, Minuten

RG, Reaktionsgefäß

RT, Raumtemperatur

sb Wasser, steriles (bi-)destilliertes Wasser (selbst herzustellen)

## TEIL A

### 1. Herstellung der Medien

**LB-Medium:** 1 % (w/v) Pepton  
0,5 % (w/v) NaCl  
0,5 % (w/v) Hefeextrakt

- Substanzen in etwas dest. H<sub>2</sub>O lösen
- mit dest. H<sub>2</sub>O auf gewünschtes Volumen auffüllen
- je 5 mL in Epruvetten aliquotieren (Dispenser bei Betreuer/innen erhältlich)
- Epruvetten mit Metallkappen verschließen
- in Korb oder Becherglas stellen
- mit Alufolie (E) abdecken
- autoklavieren
- auf RT oder im Kühlschrank lagern

#### LB/amp-Platten:

- Herstellung von LB-Medium in Schottflasche oder Erlenmeyerkolben wie oben
- 1,5 % (w/v) Agar-Agar zugeben
- schwenken
- Schottflasche mit Schraubkappe verschließen (zuletzt eine Vierteldrehung retour)
- autoklavieren
- auf ca. 50 °C abkühlen lassen/temperieren (Wasserbad)

- Ampicillin zugeben
- Medium in Petrischalen gießen

- Medium fest werden lassen
- Platten können einige Tage auf RT, längerfristig im Kühlschrank gelagert werden

## 2. Restriktionsverdau der Plasmide

Lösungen: Restriktionsenzyme *Bam*HI-HF und *Eco*RI-HF (G)  
 Cut Smart Puffer für Restriktionsenzyme (K)  
 Plasmid pJET1.2-DFR (G)  
 Plasmid pGEX-6P-1 (G)

Geräte: Inkubator oder Thermoblock

Durchführung:

- Die pJET1.2-DFR Plasmide befinden sich in der Plasmidbox und sind durchnummeriert. Verwenden sie eines der Plasmide im Eppendorfgefäßen als Ausgangsplasmid und notieren sie die Nummer.
- Reagenzien falls nötig auftauen
- alle Reagenzien schütteln, kurz abzentrifugieren und auf Eis stellen
- Reaktionsansätze je Plasmid wie folgt pipettieren:

Plasmid-DNA	2 µg
10x CutSmart Puffer	5 µL
<i>Bam</i> HI-HF	1 µL
<i>Eco</i> RI-HF	1 µL
sb Wasser	auf 50 µL

- mit der Pipette gut mischen und kurz abzentrifugieren
- bei 37 °C über Nacht inkubieren

## 3. Agarosegelelektrophorese

Lösungen: 6x Purple Loading Dye (K)  
 DNA-Marker: Gene Ruler 1kb DNA Ladder (K)  
 DNA-Färbemittel SYBR Safe (K)  
 TAE Laufpuffer (A)

Materialien: 250-mL-Erlenmeyerkolben mit Deckel (A)  
 Agarose (E)  
 Wägeschälchen für Agarose (E)  
 Topfhandschuh (A)  
 Agarosegelelektrophoreseschlitten (A)  
 Agarosegelelektrophoresekamm (A)  
 Kleine Wasserwaage (A)  
 Pipette (A)  
 Meßzylinder (A)

Geräte: Mikrowelle (A)  
 Agarosegelelektrophoreselaufkammer mit Deckel (A)  
 Power Supply (A)  
 Gel-Imager

#### Durchführung:

- Entscheiden Sie aufgrund der Tabelle, welche Agarosekonzentration das Gel haben soll

Agarosekonzentration (%)	Größe linearer DNA-Fragmente, für die eine gute Auflösung erreicht wird (kb)
0,5	30 to 1
0,7	12 to 0,8
1,0	10 to 0,5
1,2	7 to 0,4
1,5	3 to 0,2
2,0	<0,2

- die entsprechende Menge Agarose für ein 50 oder 100 mL Gel in das dafür vorgesehene Wägeschälchen einwiegen

- Agarose in den dafür vorgesehenen 250-mL-Kolben geben
- 50 (kleines Gel) oder 100 mL (großes Gel) TAE-Puffer zugeben, schwenken
- Deckel auf Kolben legen
- für ca. 2 min in der Mikrowelle (800 Watt) erhitzen
- vorsichtig schwenken (Achtung: Hitzeverzug!)
- falls Agarose nicht vollständig gelöst sein sollte, noch einmal kurz aufkochen
- die Agarose etwas abkühlen lassen
- Gelschlitten in Gießkammer einspannen und mit Wasserwaage eben ausrichten
- Kamm für breite Gelspuren in den Schlitten stecken
- Agarose mit 5 µL SYBR Safe versetzen, schwenken, in den Schlitten gießen
- eventuell entstandene Luftblasen entfernen (gelben Pipettenspitzen sind hilfreich)
- Gel ca. 20 min aushärten lassen
- gesamte Plasmidverdaue mit entsprechendem Volumen 6x Purple Loading Dye mischen
- überprüfen, ob genug TAE Laufpuffer in der Laufkammer ist
- Gel samt Schlitten in die Gelkammer legen
- vorsichtig Kamm entfernen
- 5 µL DNA-Marker in die erste Geltasche pipettieren
- die Plasmidverdaue in die weiteren Geltaschen laden
- Deckel auf Gelapparatur setzen und richtig (!) mit dem Power Supply verbinden
- Gel für ca. 50 min bei 90 V laufen lassen

- Gel mit dem Schlitten auf den Gel-Imager legen
- falls eine ausreichende Auftrennung der DNA Fragmente erfolgt ist, das Gel ohne Schlitten auf den Gel-Imager legen fotografieren und das Bild speichern
- falls keine ausreichende Auftrennung erfolgt ist, das Gel weiterlaufen lassen
- Gießkammer, Schlitten und Kamm mit dest. H<sub>2</sub>O reinigen und retournieren

#### **4. Gelextraktion**

Lösungen: GeneJET Gel Extraktions Kit (B)

Materialien: Skalpell (A)

Geräte: Gel-Imager (neben A)  
Gel Cutting Guard (A)  
Magnetischer Schalter (A)  
Schutzbrille orange (A)  
Thermoblock

Durchführung: Wenn nicht anders angegeben, bei 13000 x g zentrifugieren.

- Gel ohne Schlitten auf den Gel-Imager legen
- Guard zum Schutz des Imagers auflegen

- magnetischen Schalter anbringen damit das Gerät auch ohne Kameraabdeckung gestartet werden kann
- Schutzbrille aufsetzen
- Transillumination starten
- zügig mit dem Skalpell die richtigen Banden ausschneiden (*dfr*-Gen und Zielplasmid)
- Gelstücke in ein 2-mL-RG geben
- Imager ausschalten, mit dest. H<sub>2</sub>O reinigen und in ursprüngliche Konfiguration zurückführen
- alle verwendeten Materialien (Schlitten, Skalpell etc.) mit dest. H<sub>2</sub>O reinigen und retournieren
- Gelstücke im 2-mL-RG abwägen (vorher tarieren!)
- das gleiche Volumen (in µL) an Binding Buffer des Gelstückgewichtes (in µg) zugeben
- 10 min bei 50 - 60 °C inkubieren (oder bis sich das Gel vollständig gelöst hat)
- Bis zu 800 µL der Lösung auf ein Säulchen auftragen und 1 min zentrifugieren
- Durchfluss verwerfen
- 700 µL Wash Buffer auf die Säule geben, 1 min zentrifugieren
- Durchfluss verwerfen
- die leere Säule noch einmal 1 min zentrifugieren
- Säule in ein neues, deckelloser 1,5-mL-RG setzen
- 20 µL Elution Buffer zugeben, 1 min zentrifugieren
- Eluat in neues 1,5-mL-RG transferieren
- DNA-Konzentration mit dem NanoDrop messen (kontaktieren Sie dazu eine/n Betreuer/in)
- weiterverwenden oder auf -20 °C aufbewahren

## 5. Ligation

Lösungen: 10x T4 Reaction Buffer (G)  
T4 DNA Ligase (G)

### Durchführung:

- Reagenzien auftauen, schütteln, kurz abzentrifugieren, auf Eis stellen
- in ein 1,5-mL-RG pipettieren:
 

geschnittenes Zielplasmid	µL
geschnittene Insert-DNA (Gen)	µL
10x T4 Reaction Buffer	2 µL
sb Wasser	auf 20 µL Gesamtvolumen
T4 Ligase	1 µL
- mit der Pipette gut mischen und kurz abzentrifugieren
- 1 h bei RT inkubieren

### Formel zur Berechnung der benötigten Menge an Insert:

Das molare Verhältnis Vektor zu Insert soll für diese Sticky end Ligation 1:3 sein, wobei zwischen 20-50 ng Plasmid eingesetzt werden sollen und das Gesamtvolumen des Ligationsansatzes 20 µl nicht überschreiten darf.

Um die benötigte ng Menge an Insert für die Ligationsreaktion auszurechnen, kann folgende Formel verwendet werden:

$$[\text{ng des Zielplasmids}] \times [\text{kb Insert}] / [\text{kb Vektor}] \times (\text{Insert:Vektor Verhältnis}) = \text{ng benötigtes Insert}$$

Größe des *dfr* Gens: 1,026 kb

Größe des pGEX-6P-1: 4,984 kb

## 6. Transformation des Ligationsansatzes in *E. coli* TOP10

Lösungen: kompetente *E. coli* TOP10 Zellen\*

Materialien: Parafilm (E)

Geräte: Thermoblock  
Inkubator

Durchführung:

- ein Aliquot kompetenter *E. coli* TOP10 Zellen auf Eis auftauen lassen (ca. 10 min)
- 10  $\mu$ L des Ligationsansatzes dazu pipettieren und vorsichtig mischen
  - 30 min auf Eis inkubieren
  - Hitzeschock für exakt 2 min bei 42 °C (Thermoblock)
  - zurück auf Eis stellen
  - 300  $\mu$ L LB-Medium dazu mischen
  - Zellen bei 37 °C 30 min regenerieren (Inkubator)
  - 133  $\mu$ L der Bakteriensuspension auf 3 LB/amp-Platten mit Drigalskyspatel ausstreichen
- Platten bei 37 °C über Nacht inkubieren
  - Am Morgen Platten aus dem Inkubator nehmen, mit Parafilm verschließen und bei 4 °C aufbewahren

## 7. Anzucht von *E. coli* Einzelkolonien

Lösungen: Ampicillin (1000-fach) (K)

Materialien: Zahnstocher (E)

Geräte: Schüttelinkubator

Durchführung:

- fünf 5-mL-Aliquote LB-Medium mit Ampicillin versetzen
  - mit je einer *E. coli* Einzelkolonie beimpfen, dazu können sterile gelbe Pipettenspitzen oder sterile Zahnstocher verwendet werden
- bei 37 °C und 200 rpm über Nacht im Schüttelinkubator inkubieren

## 8. Isolierung der Plasmid-DNA

Lösungen: GeneJET Plasmid Miniprep Kit (B)

Durchführung: Wenn nicht anders beschrieben, bei 13000 x g zentrifugieren.

- 2 mL jeder *E. coli* Kultur in jeweils ein 2-mL-RG überführen
- 2 min bei 6800 x g zentrifugieren
- Überstand verwerfen
- weitere 2 mL jeder Kultur in das jeweilige RG leeren
- 2 min bei 6800 x g zentrifugieren
- Falls noch Kulturflüssigkeit vorhanden, Ernte nochmals wiederholen
- nach letztem Zentrifugieren den Überstand komplett abnehmen
- Zellpellet in 250  $\mu$ L Resuspension Solution durch Schütteln resuspendieren
- 250  $\mu$ L Lysis Solution dazu pipettieren
- das RG 4-6 mal schwenken
- 350  $\mu$ L Neutralization Solution dazu pipettieren
- das RG 4-6 mal schwenken

- 5 min zentrifugieren
- Überstand vorsichtig in Saulchen pipettieren
- 1 min zentrifugieren
- 500 µL Wash Solution in das Saulchen pipettieren
- 1 min zentrifugieren
- noch einmal 500 µL Wash Solution in das Saulchen pipettieren
- 1 min zentrifugieren
- leeres Saulchen noch einmal 1 min zentrifugieren
- Saule in ein neues, deckelloes 1,5-mL-RG geben
- mit 50 µL Elution Buffer exakt auf den Saulchenboden pipettieren
- 1 min zentrifugieren
- Eluat in neues RG transferieren
- benennen Sie Ihre Plasmide; ublich ware:
  - ein p fur Plasmid
  - ein Hinweis auf die Plasmid-"Backbone"
  - ein Hinweis auf das Insert
  - und/oder Ihre Initialen
  - eine laufende Nummer
  - zB pGEX-DFR-1 oder pAMA1
- DNA-Konzentrationen mit dem NanoDrop messen (kontaktieren Sie dazu eine/n Betreuer/in)
- weiterverwenden oder auf -20 °C aufbewahren

## 9. Kontrollverdau (Restriktionsanalyse)

Losungen: Restriktionsenzyme *PstI*-HF, *BamHI*-HF und *EcoRI*-HF (G)  
 CutSmart Puffer  
 6x Purple Loading Dye (K)  
 DNA-Marker: Gene Ruler 1kb DNA Ladder (K)  
 DNA-Farbemittel SYBR Safe (K)  
 TAE Laufpuffer (A)

Materialien: 250-mL-Erlenmeyerkolben mit Deckel (A)  
 Agarose (E)  
 Wageschalchen fur Agarose (E)  
 Topfhandschuh (A)  
 Agarosegelelektrophoreseschlitten (A)  
 Agarosegelelektrophoresekamm (A)  
 Kleine Wasserwaage (A)  
 Pipette (A)

Gerate: Mikrowelle (A)  
 Agarosegelelektrophoreselaufkammer mit Deckel (A)  
 Power Supply (A)  
 Gel-Imager

Fuhren Sie 2 Kontrollverdaue durch: einer soll das korrekte Plasmid linearisieren (*PstI*-HF) und der zweite daraus das Insert wieder herausschneiden (*BamHI*-HF und *EcoRI*-HF). Stellen Sie fur beide Kontrollverdaue je einen Mastermix her.

*Anzahl der Proben je Mastermix:*

__ Plasmid-DNA Proben	__
Aufschlag	1
Gesamtanzahl	__

Mastermix pro Kontrollverdau (sb Wasser dient zum Auffüllen des Volumens):

Reagens	Konz.	Endkonz./-menge	1 Probe [ $\mu$ L]	__ Proben [ $\mu$ L]
CutSmart Puffer	10x	1x		
<i>Pst</i> I-HF	-	-	0,5	
Plasmid DNA-Probe		ca. 500 ng		
sb Wasser	-	-		
Gesamt	-	-	20	

Reagens	Konz.	Endkonz./-menge	1 Probe [ $\mu$ L]	__ Proben [ $\mu$ L]
CutSmart Buffer	10x	1x		
<i>Bam</i> HI-HF	-	-	0,5	
<i>Eco</i> RI-HF	-	-	0,5	
Plasmid DNA-Probe		ca. 500 ng		
sb Wasser	-	-		
Gesamt	-	-	20	

Durchführung:

- pro Mastermix ein 1,5-mL-RG auf Eis bereitstellen
- Reagenzien und Plasmid-DNA Proben auftauen, schütteln, kurz abzentrifugieren, auf Eis stellen
- Mastermixe herstellen: die oben berechneten Mengen der Reagenzien (außer der Proben) in das jeweilige RG pipettieren, gut schütteln und kurz abzentrifugieren
- die entsprechenden Volumina der jeweiligen Mastermixe in 1,5-mL-RG vorlegen
- die entsprechenden Volumina Plasmid-DNA Proben zu beiden vorgelegten Mastermixen mischen
- 60 min bei 37 °C inkubieren
- alle Ansätze mit entsprechendem Volumen 6x-Ladepuffer versetzen
- je 15  $\mu$ L auf einem 0,8%igen Agarosegel analysieren (siehe dazu Punkt 3. Agarosegelelektrophorese)
- Gellauf: 50 min, 90 V
- Gel am Gel-Imager mit Kamerahood fotografieren und Bild speichern

## 10. Sequenzierung

Durchführung:

- ein Plasmid mit korrektem Restriktionsmuster auswählen
- falls erforderlich diese Plasmid-DNA auftauen, schütteln, kurz abzentrifugieren
- 700 – 1200 ng der Plasmid-DNA in ein 1,5-mL-RG pipettieren
- mit sb Wasser auf ein Endvolumen von 12  $\mu$ L auffüllen
- das RG bei einem/r Betreuer/in abgeben
- das Sequenzierergebnis wird per e-mail an Sie übermittelt
- Kontrollieren Sie die Sequenz des *dfr*-Gens und identifizieren Sie um welches es sich gehandelt hat. Die Sequenzen der verschiedenen *dfr*-Gene werden ihnen dazu zur Verfügung gestellt.



# TEIL B

## 11. Identifizierung der kodierenden Region in einer DNA Sequenz mit BLAST

Im ersten Schritt werden Regionen mit Ähnlichkeiten in der Aminosäuresequenz zwischen ihrer Sequenz und öffentlich zugänglichen Sequenzen identifiziert. Dazu wird ihre DNA Sequenz mit Aminosäuresequenzen der Datenbank am NCBI ([National Center for Biotechnology Information](https://www.ncbi.nlm.nih.gov/)) verglichen. Zum Vergleich verwenden Sie das BLAST (Basic Local Alignment Search Tool) Programm. Unter den verschiedenen BLAST Programmen wählen sie dabei das blastx Programm aus. Blastx übersetzt ihre Sequenz in alle 6 Leserahmen und vergleicht diese zur Proteindatenbank am NCBI.

Materialien: DNA Sequenz  
Abbildung Expressionsvektor  
Sequenz Expressionsvektor  
Gene Runner

Durchführung:

- Gehen Sie auf die NCBI Webseite: <https://www.ncbi.nlm.nih.gov/>
- Suchen Sie das BLAST Programm und klicken Sie es an
- Suchen Sie das blastx Programm und klicken Sie es an
- Kopieren Sie ihre DNA Sequenz im FASTA Format in das offene Fenster: Enter Query Sequence.  
Das FASTA Format sieht so aus:  
>definition line (=Name ihrer Sequenz)  
catgagagatctagtcagctgat (ihre Sequenz selbst).
- Klicken Sie auf den BLAST Button, um ihren Sequenz abzuschicken. Das Resultat erscheint dann nach ein paar Sekunden bis Minuten.

Im Fenster Blast Results sehen Sie dann die Verteilung der Treffer (Hits) in Bezug auf ihre Query Sequenz. Eine Farbcode gibt dabei an wie ähnlich die Sequenzen zu ihrer übersetzten Aminosäure Query Sequenz sind. Diese farbigen Balken entsprechen in der Regel den Exons ihres Gens und sind meist durch dünne schwarze Striche unterbrochen, die den Introns entsprechen. Aus dieser Abbildung lässt sich somit bestimmen aus wie vielen Exons bzw Introns ihr Gen ungefähr besteht. Außerhalb der Exons ihres Gens befinden sich weitere nicht konservierte Regionen, das sind die Promoter- und Terminatorregion. Unter diesen Hits befinden sich die Sequences producing significant alignments, die überspringen Sie und gehen direkt zu Alignments.

Überprüfen sie zuerst in welcher Orientierung sich das Gen in ihrer DNA befindet. Ihre Sequenz ist der Query. Die Sequenzen der NCBI Datenbank sind das Sbjct (Subject). Bei der Überprüfung achten Sie darauf, dass die Nummerierung der Aminosäuren der Proteine der Datenbank immer in der richtigen Orientierung vom N- zum C-Terminus angegeben wird. Ihre Query Sequenz orientiert sich aber an der Proteinsequenz des Sbjt. Da ihr Gen in der DNA in 2 Orientierungen (+ oder - Strang) vorkommen kann, ist es für die weitere Analyse notwendig, das Gen in der richtigen Orientierung zu haben. Sind ihre Frames (Leserahmen) positiv, also +1, +2 oder +3, ist ihre DNA Sequenz in der richtigen Orientierung. Sind ihre Leserahmen -1, -2, -3, dann müssen sie ihre Sequenz mit Hilfe eines DNA Programmes umdrehen.

Durchführung:

- Überprüfen Sie die Orientierung ihres Gens an Hand des BLAST Ergebnisses
- Falls sich das Gen in der falschen Orientierung befindet, drehen sie die DNA Sequenz um und wiederholen den blastx Vorgang mit dem Gegenstrang

## 12. Änderung der Orientierung der DNA

Durchführung:

- Installieren Sie das DNA Programm Gene Runner (ein aktueller Link wird ihnen am Kursanfang zur Verfügung gestellt)

- Öffnen Sie den Gene Runner. Gehen sie zu Go to File → New → Nucleic acid sequence
- Kopieren Sie ihre DNA Sequenz und fügen sie die DNA Sequenz ein
- Gehen Sie zu Edit → Sequence Orientation → Klicken sie auf Reverse and Complement.
- Ihre Sequenz ist somit umgedreht und komplementiert. Sie können die obere Sequenz mittels Copy and Paste in das blastx Fenster einfügen. Modifizieren Sie die Sequenz noch so, dass Sie dem FASTA Format entspricht.
- Wiederholen sie nun den BLAST wie oben angegeben.

### 13. Übersetzung der DNA Sequenz und Darstellung als DNA/Protein Alignment

Es sind nun alle Informationen vorhanden, um ein DNA Protein Sequenzalignment zu machen. Prinzipiell geht das auch im Gene Runner, nur lässt sich dieses Alignment in den neueren Versionen nicht mehr exportieren und abspeichern (Gehen Sie auf View → Translation und übersetzen Sie Frame 1-3). Als Alternative stehen z.B. Programme wie EMBOSS Sixpack zur Verfügung.

#### Durchführung:

- Gehen sie am EMBL-EBI auf das Tool EMBOSS Sixpack ([https://www.ebi.ac.uk/Tools/st/emboss\\_sixpack](https://www.ebi.ac.uk/Tools/st/emboss_sixpack)).
- Kopieren Sie ihre DNA Sequenz und fügen Sie die DNA Sequenz ein. Gehen Sie auf Step 2 Select Parameters bei REVERSE auf no (Sie bekommen dann nur 3 Leserahmen)
- Drücken Sie submit
- Sie erhalten nun ihre Sequenz mit den 3 Leserahmen. Klicken Sie auf Download Sequence File und kopieren Sie die ganze DNA Protein Sequenz und fügen Sie alles in ein Word Dokument ein. Verwenden Sie eine Schriftart bei der jedes Zeichen die gleiche Größe hat wie z.B. Courier. Im Falle des EMBOSS Sixpack Files befindet sich die Aminosäure über dem ersten Nucleotid des Codons (Triplett). Stops in ihrer Sequenz sind durch \* gekennzeichnet.
- Suchen Sie nun den Anfang ihres Proteins in ihrem blastx Ergebnis. Schauen Sie sich dafür die erste Zahl der Aminosäuren der sbjt Line der verschiedenen Reading Frames an. Die kodierende Region in ihrem Sbjct line beginnt idealerweise mit Aminosäure 1, wenn ihr erstes Exon also mit dem Start ATG/Methionin beginnt. Falls aber der Anfang des Proteins nicht konserviert ist, müssen Sie ihr DNA/Protein Sequenzalignment upstream nach einem ATG/Methionin absuchen (zur Vereinfachung befindet sich das Methionin im diesem Fall immer im selben Leserahmen). Nehmen Sie das erste Methionin im Leserahmen als ihr Start Methionin.
- Im nächsten Schritt wird nun das Ende des ersten Leserahmens in ihrem DNA/Proteinalignment gesucht. Anhand des blastx Ergebnisses können Sie nun die Sequenz bis zum Ende des Exons/Leserahmens nachverfolgen. Markieren Sie das Ende des Leserahmens im File.
- Auf das Exon folgt logischerweise ein Intron. Der Anfang eines Introns beginnt in den meisten Fällen (90-95%) mit der Sequenz GT. Es kommt aber vor, dass sich beim Übergang von Exon zu Intron mehrere GTs befinden, die in Frage kommen könnten. Sie müssen daher das richtige finden. Dabei müssen Sie jenes GT nehmen, das am besten zu ihrem blastx Resultat korreliert. Der Anfang des Introns beginnt daher meist nach dem Bereich, der für die letzten konservierten Aminosäuren kodiert. Markieren Sie das GT.
- Gehen Sie dann zum Ende des Introns. Das Intron endet nach einem AG, gefolgt vom zweiten Exon. Nehmen Sie wieder ihren blastx Vergleich, suchen das zweite Exon und suchen sie die ersten Aminosäuren des Exons in ihrem DNA/Protein Alignment. Markieren Sie mögliche passende AGs. Das richtige AG befindet sich meist kurz vor dem Bereich, der für die ersten konservierten Aminosäuren des zweiten Exons kodiert. Nach dem AG kann es zu einem Wechsel des Leserahmens kommen. Vergleichen sie dazu die Leserahmen im blastx Resultat.
- Wenn Sie nun die Exon/Intron Grenzen gefunden haben, überprüfen Sie ob die Grenzen korrekt sind. Das machen Sie durch die Überprüfung der Translation der entstehenden Sequenz. Es kann vorkommen, dass z.B. ein Intron ein Triplett unterbricht. 1-2 Nucleotide befinden sich dann vor dem Intron, also vor der GT Sequenz, während 2 oder 1 Nucleotid(e) nach dem AG folgen, bevor das erste vollständige Triplett anschließt. Falls kein vollständiges Triplett entsteht, müssen Sie Anfang und Ende ihres Introns noch einmal überprüfen und nach Alternativen für ihr gewähltes GT oder AG suchen. Achten Sie dabei aber immer auf die konservierten AA, die erhalten werden müssen.
- Weiters ist darauf zu achten, dass die Spaltung eines Triplets durch ein Intron auch Konsequenzen auf die Aminosäuresequenz haben kann und eventuell eine Änderung der AA

bewirken kann. Überprüfen Sie daher ob das neue Triplet eine Änderung der AA zur Folge hat und korrigieren Sie es, wenn nötig, im File.

Beispiel AGTGNNNNNNNNNAGCC. Das Triplet AGT kodiert für Serin. Aber durch die Entfernung der Intronsequenz (unterstrichen) entsteht ein neues Triplet nämlich ACC, welches für Threonin kodiert.

- In weiterer Folge werden nun die anderen Introns/Exons lokalisiert, bis zum Ende der kodierenden Sequenz. Speichern Sie das Dokument am besten in Kopie ab.
- Machen Sie nun das DNA Proteinalignment. Entfernen Sie alle nicht übersetzten Leserahmen und AA bis nur mehr die DNA und die kodierte Aminosäuresequenz übrigbleibt. Achten Sie darauf, dass Sie durch das Löschen nicht die Aminosäuren verschieben.
- Extrahieren Sie nun die Aminosäuresequenz
- Extrahieren Sie nun den kodierenden Teil der cDNA Sequenz beginnend beim ersten ATG bis zum Stop Codon.
- Kontrolle der erhaltenen Sequenzen  
Überprüfen Sie ihre cDNA Sequenz. Es muss ein durchgehender Leserahmen ohne Stops vorhanden sein. Das können Sie z.B. im Gene Runner machen. Die übersetzte DNA können Sie dann noch mit einem BLAST Programm vergleichen. Dazu gehen Sie statt auf blastx auf PROTEIN BLAST (blastp). Klicken PROTEIN BLAST an und dann auf Align two or more sequences. Fügen Sie die Aminosäuresequenzen in die 2 Fenster ein und klicken auf BLAST. Die Sequenzen sollten dann 100% identisch sein.

## 14. Design von Oligonukleotiden für die Amplifikation der cDNA und Klonierung in einen Expressionsvektor

Basierend auf ihrer cDNA Sequenz werden nun Oligonukleotide für die *in silico* Amplifikation der cDNA designt, um die cDNA dann in einen Expressionsvektor über Restriktionsschnittstellen zu insertieren. Die Herstellung des Vektors erfolgt dabei einmal über eine konventionelle Klonierung basierend auf Restriktionsverdau und Ligation sowie auf einer Rekombinationsklonierung.

Durchführung:

### A. Konventionelle Klonierung

- Überprüfen Sie an Hand der erhaltenen Abbildung ihres Expressionsvektors, welche Restriktionsenzyme in der Multiple Cloning Site (MCS) schneiden.
- Überprüfen Sie welche Restriktionsenzyme der MCS in der cDNA nicht schneiden. Öffnen Sie dafür den Gene Runner und gehen Sie zu Analysis → Nucleic Acid → Restriction Sites. Bei Output format klicken Sie nur List non-cutters an. Vergleichen Sie welche der Restriktionsenzyme, die in der MCS schneiden, nicht in der cDNA schneiden.
- Restriktionsenzyme, die im Expressionsvektor in der MCS und nicht in der cDNA schneiden sind geeignet. Davon wählen Sie 2 aus und nehmen Sie am besten ein Restriktionsenzym das am 5' Ende und eines das am 3' Ende der MCS schneidet. Wichtig ist auch die Reihenfolge der Restriktionsenzyme in der MCS zu beachten, damit die cDNA richtig orientiert ist.
- Primer werden, wie für DNA üblich, immer in 5' zu 3' Richtung geschrieben. Der sogenannte Forward Primer beginnt beim Start Methionin (ATG) und dann folgen weitere 16-18 Nukleotide des Anfangs der cDNA. Um die cDNA in ein Plasmid zu klonieren, wird noch zusätzlich die Sequenz für eine Restriktionsenzymsschnittstelle an den Anfang des Primers gehängt: im Beispiel GAATTC für *EcoRI*.

Beispiel: 5'-GAATTC-ATG-gagctagatagtaatgata-3'  
(*EcoRI* Schnittstelle-StartATG-cDNA spezifische Sequenz)

Der Reverse Primer wird analog zusammengestellt. Er beginnt mit dem letzten Teil der kodierenden Sequenz einschließlich Stop Codon gefolgt von der gewählten Restriktionsschnittstelle. Der Reverse Primer würde dann natürlich in die falsche Richtung laufen. Daher muss die Sequenz umgedreht werden und die komplementäre Sequenz gebildet werden (Reverse and complement).

Beispiel: 5'-gggtatatatatgtat-TAG-GGATCC-3'.  
(cDNA spezifische Sequenz-StopTAG-*BamHI* Schnittstelle)

Reverse and Complement: 5' GGATCC-CTA-atatatacccc-3'

Nun hängt es noch von der Art ihres Expressionsvektors ab, ob die Oligonukleotide noch weiter verändert werden müssen. Expressionsvektoren enthalten oft Tags die N oder C-Terminal an das Protein angehängt werden. Daher ist zu beachten, dass das Protein und die Tags im selben Leserahmen sind.

Stop Codons sind meist im Expressionsvektor nach der MCS enthalten. Die Benutzung des eigenen Stop Codons der cDNA Sequenz ist aber vorzuziehen, da dann keine zusätzliche AA an C-Terminus des Proteins hängen. Falls aber ein Tag an den C-Terminus gehängt wird, muss das Stop Codon entfernt werden, da sonst dieser Tag nicht übersetzt wird. Ist am C-Terminus ein Tag, muss auch darauf geachtet werden, dass der Leserahmen des Proteins und des Tags einen gemeinsamen Leserahmen bilden. Es kann daher sein, dass nach dem letzten für eine Aminosäure kodierenden Codons noch zusätzlich 1-2 Nucleotide eingefügt werden müssen, um einen durchgehenden Leserahmen zu erhalten.

Expressionsvektoren können aber auch N-terminal Tags an das Protein knüpfen. Auch dann ist darauf zu achten, dass Sie im richtigen Leserahmen bleiben. Das zu exprimierende Gen muss daher im selben Leserahmen wie der Tag sein. Auch hier kann es nötig sein, 1-2 zusätzliche Nucleotide in den Primer einzufügen.

Beispiel Forward Primer: 5'-GAATTC- N<sub>1-2</sub>-ATG-gagctagatagtaaatgata-3'

Beispiel Reverse Primer 5' GAATTC-N<sub>1-2</sub>-atatataccccgtgctcgt-3'

- Überprüfen Sie daher an Hand ihres Expressionsvektors welche Arten von Tags Sie haben und wie sich die Klonierung ihres cDNA Konstrukts in den Expressionsvektor auf den Leserahmen auswirkt. Adaptieren Sie dann die Sequenz des Primers entsprechend. Verwenden Sie dabei aber keine Ns, sondern A, G C, oder T.

Bevor Primer bestellt werden, macht es Sinn noch folgendes zu beachten:

- Zusätzliche Nucleotide an den Enden der Oligonucleotide sind oft notwendig, damit die Restriktionsenzyme verlässlich schneiden können. Dafür kann man entsprechende Listen der Hersteller konsultieren.
- Die Annealing Temperatur der Primer sollte möglichst identisch sein. Die Temperatur wird dabei aus der Sequenz des cDNA Teils des Primers berechnet. Um halbwegs gleiche Temperaturen zu erreichen, kann daher der cDNA Sequenz Anteil der Primer noch verkürzt oder verlängert werden.
- Sekundärstruktur der Primer: dazu gibt es verschiedene Online Programme bzw kann auch der Gene Runner verwendet werden. Basierend auf den Ergebnissen kann die DNA Sequenz der Primer noch entsprechend modifiziert werden ohne natürlich die Aminosäure-zusammensetzung des kodierenden Bereichs des Gens zu ändern.

## B. Rekombinationsklonierung

- Basierend auf der cDNA Sequenz und der Sequenz des Expressionsvektors werden Oligonukleotide für die Rekombinationsklonierung designt. Ein Beispiel ist das In-fusion System von Clontech.
- Öffnen Sie die Website <http://bioinfo.clontech.com/infusion/convertPcrPrimersInit.do> und folgen Sie den Anweisungen auf der Website.
- Bei der Rekombinationsklonierung ist nur eine Restriktionsschnittstelle im Expressionsvektor nötig ist und es spielt auch keine Rolle ob das Restriktionsenzym im cDNA Fragment schneidet. Wählen sie also eine oder zwei Restriktionsschnittstellen.
- Beachten sei aber wieder, dass ihre Sequenzen in frame mit ihren Tags sein müssen.

## 15. Abgabe

- Abgabe eines DIN A4 Blattes mit der Nummer des Ausgangsplasmids, den Gelbildern der präparativen Plasmidverdaue sowie der Kontrollverdaue und das ausgewertete Sequenzierungsergebnis (welches *dfr* Gen haben Sie umkloniert?). Wählen Sie außerdem ein

weiteres Enzym oder eine Enzymkombination aus, das/die theoretisch ein sinnvolles Schnittmuster ergeben würde und geben Sie den/die Name/n des/der Enzyms/e sowie das zu erwartende Schnittmuster (= Größen der theoretisch entstehenden Fragmente in bp) an.

- Abgabe der folgenden Ergebnisse des Teils B als doc und pdf per email an ([bernhard.seiboth@tuwien.ac.at](mailto:bernhard.seiboth@tuwien.ac.at)):
  1. Ihre Query DNA Sequenz
  2. Nukleotid/Aminosäure Alignment
  3. Protein Sequenz und cDNA Sequenz
  4. Name des Expressionsvektors und der 2 Restriktionsenzyme
  5. Forward und Reverse Primer für Klonierung mittels Restriktionsschnitt und Ligation
  6. Forward und Reverse Primer für Klonierung durch das In-fusion System

**Plasmidkarten:**



