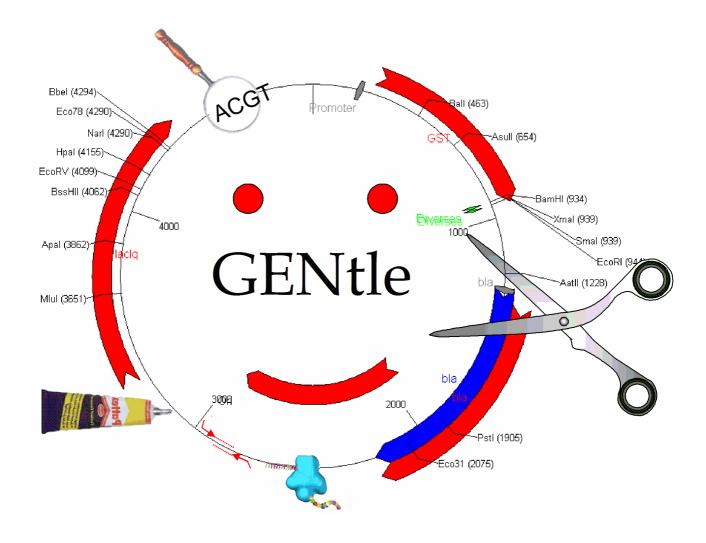
GENtle 1



# Die Anleitung

Version Juni 2004

GENtle 2

# Inhaltsverzeichnis

I. Uber diese Anleitung	3
1. Konventionen.	3
2. Copyright	
3. Kontakt	
II. DNS	4
1. Anzeigen von DNS-Sequenzen	4
2. Toolbar	
3. Detail-Baum.	
4. Kartendarstellung	
5. Sequenzdarstellung.	
III. Peptide	
1. Toolbar	
2. Funktionen	
3. Sequenzdarstellung.	
4. Spezielle Menüs.	
IV. PCR und Primer Design.	
1. Das PCR-Modul	
2. Dialog "Primer bearbeiten"	
3. Dialog "Stille Mutation"	
V. Sequenzierung	
1. Darstellung	
2. Bearbeitung.	
3. Toolbar	
VI. Alignments	
Starten des Alignment-Moduls	
Dialog "Einstellungen"      Das Alignment-Modul	
VII. Rechner	
1. Rechner "Ligation"	
3. Rechner "Protein"	
VIII. Bildbetrachter	
IX. Werkzeuge und Dialoge	
1. Ligation	
2. Einstellungen	
3. Öffnen/Speichern/Datenbank verwalten	
4. Importieren	
5. Sequenz eingeben	
6. Sequenz-Editor	
7. Restriktions-Assistent	
8. Projekte	
9. Drucken	
10. Sequenzdarstellung	
11. Restriktionsenzym-Editor	
X. FAQ	

# I. Über diese Anleitung

#### 1. Konventionen

- Tasten werden wie folgt markiert : Taste
- Menüs, Regler und Buttons werden wie folgt markiert : Menü
- Doppelklick bezieht sich stets auf die linke Maustaste.
- Kontextmenü ist das Menü, das sich beim Klicken mit der rechten Maustaste öffnet.

# 2. Copyright

GENtle ist ©2004 by Magnus Manske, lizensiert unter GPL Diese Anleitung ist ©2004 by Magnus Manske, lizensiert unter GFDL

#### 3. Kontakt

Magnus Manske
AG Klein, Institut für Biochemie, Universität zu Köln
Otto-Fischer-Straße 12-14
50674 Köln
0221 – 470 3222
magnus.manske@uni-koeln.de

#### II. DNS

Im **DNS-Modul** können **DNS-Sequenzen** angezeigt, annotiert und bearbeitet werden.

# 1. Anzeigen von DNS-Sequenzen

DNS-Sequenzen können auf mehrere Arten im DNS-Modul geöffnet werden:

- 1. Öffnen von Datenbankeinträgen (→IX.3).
- 2. Öffnen von Dateien  $(\rightarrow IX.4)$ .
- 3. Manuelle Eingabe ( $\rightarrow IX.5$ ).
- 4. Aus einer Teilsequenz eines anderen DNS-Moduls

# 3

#### 2. Toolbar

Die Toolbar (1) bietet verschiedene Funktionen Abbildung II.1: Das DNS-Modul. und Einstellungen:

- D Neue Sequenz manuell eingeben  $(\rightarrow IX.5)$ .
- œ Sequenz öffnen ( $\rightarrow$ IX.3).
- Sequenz speichern  $(\rightarrow IX.3)$ .
- ➾ Undo (letzte Aktion rückgängig machen).
- \* Ausschneiden (Strg-X).
- Kopieren (Strg-C).
- Einfügen (Strg-V).
- Circulär/linear umschalten (nur für Sequenzen ohne *sticky ends*).
- Open reading frames anzeigen/verbergen.
- Features anzeigen/verbergen.
- Tyr Restriktionen anzeigen/verbergen.
- Nur Kartendarstellung anzeigen.
- Sequenz bearbeiten.
- Zoom Zoomen der Kartendarstellung.

#### 3. Detail-Baum

Die Baumdarstellung (2) links oben im Modul zeigt Eigenschaften der DNS-Sequenz, der Features und der Restriktionsschnittstellen. Features und Restriktionsschnittstellen können per Doppelklick ein- bzw. ausgeblendet werden. Weitere Aktionen sind über das jeweilige Kontextmenü abrufbar.

# 4. Kartendarstellung

Die Kartendarstellung (3) zeigt eine Übersicht der gesamten DNS-Sequenz. Features und Restriktionsschnittstellen werden angezeigt.

#### **Aktionen per Maus**

Aktion mit	Maustaste	Funktion
Hintergrund	Links	Sequenz markieren
	Links (Doppelklick)	Sequenz-Editor öffnen
	Mitte	Sequenz anzeigen
	Rechts	Kontext-Menü
Feature	Links	Verschieben
	Links (Doppelklick)	Feature bearbeiten
	Mitte	Feature-Sequenz markieren und anzeigen
	Rechts	Kontext-Menü
Restriktionsschnittstellen	Links	Verschieben
	Links (Doppelklick)	Enzymliste bearbeiten
	Mitte	Restriktions-Assistent $(\rightarrow \underline{IX.7})$
	Rechts	Kontext-Menü
Open reading frame	Links	ORF-Sequenz markieren
	Links (Doppelklick)	ORF-Sequenz markieren und anzeigen
	Rechts	Kontext-Menü
Gesamte Karte	Links (+Strg-Taste)	Gezoomten Ausschnitt verschieben

#### Kontext-Menü

#### Hintergrund

• Sequenz Bearbeiten Öffnet den Sequenz-Editor  $(\rightarrow \underline{IX.6})$ .

• Sequenz transformieren Sequenz invertiert und/oder komplementär.

• Enzyme Limitieren Zeigt nur noch Enzyme, die höchstens *n*-mal schneiden.

PCR

PCR Startet das PCR-Modul (→<u>IV.</u>).
 VORWÄRTS PCR-Modul mit 3'-Primer.

Rückwärts
 PCR-Modul mit 5'-Primer.

• Beide PCR-Modul mit 3'- und 5'-Primer.

MUTATION PCR-Modul mit Mutagenese-Primern.

Auswahl

Ausschneiden Markierte Sequenz ausschneiden.

• Kopieren Markierte Sequenz in die Zwischenablage kopieren.

ALS NEUE SEQUENZ Ein neues DNS-Modul aus der Markierung erzeugen.

• Enzyme, die in der Markierung schneiden ( $\rightarrow IV.3$ ).

• Auswahl als neues Feature Ein neues Feature aus der Markierung erzeugen.

• Aminosäuren extrahieren Ein neues AS-Modul aus der Markierung erzeugen.

• BLAST DNS BLAST-Suche nach der DNS-Sequenz der Markierung.

BLAST Aminosäuren
 BLAST-Suche nach der AS-Sequenz der Markierung.

Sequenzkarte

• In die Zwischenablage kopieren Kopiert die Sequenzkarte in die Zwischenablage.

Sequenzkarte drucken
 Druckt die Sequenzkarte.

ORFs anzeigen/verbergen Schaltet die Open reading frame-Darstellung an/aus.

ORFs BEARBEITEN ORF-Darstellungsoptionen.

Restriktionsschnittstellen

Restriktionsenzym bearbeiten Restriktionsenzyme entfernen/hinzufügen/organisieren.

Enzym anzeigen/ausblenden Alle Restriktionsschnittstellen des Enzyms anzeigen.

Enzym entfernen Das Restriktionsenzym entfernen.

Restriktions-Site markieren Die Restriktions-Site in der Sequenz markieren.

Restriktions-Site markieren und anzeigen Die Restriktions-Site in der Sequenz anzeigen.

Enzym-Information bei Rebase (Internet-Verbindung!).

Zum Cocktail hinzufügen Enzym zur Restriktion vorbereiten.

Zum Cocktail hinzufügen und schneiden Restriktion mit Enzym vorbereiten und starten.

Features

Feature bearbeiten Das Feature bearbeiten  $(\rightarrow \underline{IX.6})$ 

Feature ausblenden. Das Feature ausblenden.

Feature Löschen Das Feature löschen.

DNS-Sequenz

• Markieren Die DNS-Sequenz des Features markieren.

• Markieren und anzeigen Die DNS-Sequenz des Features anzeigen.

• Kopiere (kodierende) Sequenz Die (kodierende) DNS-Sequenz des Features kopieren.

• Feature als Neue Sequenz Ein neues DNS-Modul aus dem Feature erzeugen.

• BLAST DNS BLAST-Suche nach Feature-DNS.

Aminosäuresequenz

• Kopiere Aminosäuresequenz des Features kopieren.

• ALS NEUEN EINTRAG Die Aminosäuresequenz als neues Peptid-Modul.

• BLAST AMINOSÄUREN BLAST-Suche nach Feature-AS-Sequenz.

**ORFs** 

• Als Neues Feature Aus dem ORF ein neues Feature erzeugen.

DNS-Sequenz

• KOPIERE DNS-SEQUENZ Die DNS-Sequenz des ORFs kopieren.

• Als Neue Sequenz Ein neues DNS-Modul aus dem ORF erzeugen.

• BLAST DNS BLAST-Suche nach ORF-DNS.

Aminosäuresequenz

• Kopiere Aminosäuresequenz Die Aminosäuresequenz des ORFs kopieren.

ALS NEUE SEQUENZ Ein neues Peptid-Modul aus dem ORF erzeugen.

BLAST Aminosäuresequenz.

BLAST-Suche nach ORF-Aminosäuresequenz.

# 5. Sequenzdarstellung

Die Sequenzdarstellung (4) wird von vielen Modulen verwendet. Sie wird hier exemplarisch vorgestellt.

# Darstellungsoptionen

Zusätzlich zur DNS-Sequenz können komplementäre Sequenz, Features, Restriktionsschnittstellen, codierte Aminosäuresequenzen und deren Protease-Schnittstellen dargestellt werden. Alle verfügbaren Darstellungsoptionen finden sich im Menü Ansicht.

- Features können per Menü, Toolbar oder F6 ein- bzw. ausgeblendet werden.
- Restriktionsschnittstellen können per Menü, Toolbar oder F7 ein- bzw. ausgeblendet werden.
- Die komplementäre 3'→5'-Sequenz kann per Menü oder F8 ein- bzw. ausgeblendet werden.
- Aminosäuresequenzen können per Menü oder Tastenkürzel in ihrer Darstellung verändert werden:
  - Strg+0 Aminosäuren ausblenden.
  - Strg+1 Aminosäuren in der gesamten Sequenz mit reading frame 1 anzeigen.
  - Strg+2 Aminosäuren in der gesamten Sequenz mit reading frame 2 anzeigen.
  - Strg+3 Aminosäuren in der gesamten Sequenz mit reading frame 3 anzeigen.
  - Strg+4 Aminosäuren in der gesamten Sequenz mit reading frame -1 anzeigen.
  - Strg+5 Aminosäuren in der gesamten Sequenz mit reading frame -2 anzeigen.
  - Strg+6 Aminosäuren in der gesamten Sequenz mit reading frame -3 anzeigen.
  - Strg+7 Aminosäuren nur bei bekanntem reading frame anzeigen.
  - Strg+8 Aminosäuren ausblenden.
  - Strg+Q Aminosäuren als ein-Buchstaben-Code anzeigen.
  - Strg+W Aminosäuren als drei-Buchstaben-Code anzeigen.

#### Kontext-Menü

Das Kontext-Menü der Sequenzdarstellung entspricht in weiten Teilen dem der Kartendarstellung. Hinzugekommen sind:

- Als Bild kopieren Die Sequenzdarstellung als Bild in die Zwischenablage kopieren.
- ALS BILD SPEICHERN Die Sequenzdarstellung als Bitmap speichern.
- Sequenz drucken Die Sequenzdarstellung (bzw. den ausgewählten Bereich) drucken.

#### **Editiermodus**

Zwischen Darstellungs- und Editiermodus wird über die Toolbar oder F2 umgeschaltet. Dabei wird die Sequenzdarstellung während des Bearbeitens der Übersichtlichkeit halber maximiert. Je nach Modul können nur bestimmte Zeichen eingegeben werden, z.B. nur A, C, G oder T im DNS-Modul. Wird eine andere Taste betätigt, erfolgt eine Nachfrage, ob andere Tasten vorübergehend erlaubt werden sollen.

Der Text-Cursor kann ähnlich einer Textverarbeitung bewegt werden. Löschen von Zeichen funktioniert ebenso. Zwischen den Modi "Einfügen" und "Überschreiben" kann mit der Einfg-Taste umgeschaltet werden, wobei sich der Cursor sichtbar verändert. In manchen Modulen (z.B. PCR, Sequenzierung) ist der Modus "Überschreiben" fest eingestellt; hier kann auch nicht gelöscht werden.

# III. Peptide

Das Peptid- oder Protein-Modul ermöglicht das Darstellen, Annotieren, Analysieren und Bearbeiten von Aminosäuresequenzen.

#### 1. Toolbar

Funktionen der Toolbar (1):

- Neue Sequenz manuell eingeben  $(\rightarrow \underline{IX.5})$ .
- Sequenz öffnen  $(\rightarrow IX.3)$ .
- Sequenz speichern  $(\rightarrow \underline{IX.3})$ .
- Ausschneiden (Strg-X).
- Kopieren (Strg-C).
- Einfügen (Strg-V).

PLOT Zeigt einen Plot in der Sequenz an.

Horizontal Zeigt die Aminosäuresequenz als eine lange Zeile an.

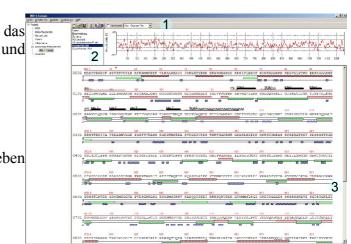


Abbildung III.1: Das Peptid-/Protein-Modul

#### 2. Funktionen

Eine Liste mit Funktionen (2):

- Daten Basisdaten der Aminosäuresequenz.
- Beschreibung der Aminosäuresequenz.
- Schematische Darstellung der Aminosäuresequenz.
- AS-GEWICHT Plot der Aminosäuresequenz nach Aminosäuregewicht.
- AS-Pl Plot der Aminosäuresequenz nach isoelektrischem Punkt.
- Hydrophobizität Plot der Aminosäuresequenz nach Hydrophobizität.
- Chou-Fasman-Plot Sekundärstrukturvorhersage nach Chou-Fasman-Methode.

# 3. Sequenzdarstellung

Die Sequenzdarstellung (3) entspricht weitgehend der des DNS-Moduls. Dabei wird die Aminosäuresequenz bearbeitet; die DNS-Darstellung entfällt offensichtlich. Beim Wechsel in den Editiermodus wird die Sequenzdarstellung *nicht* vergrößert.

Das Kontext-Menü entspricht weitgehend dem der Sequenzdarstellung im DNS-Modul.

# 4. Spezielle Menüs

• Bearbeiten/Photometer-Analyse Die Daten des aktuellen Peptids werden in das Rechner-Modul

"Protein" (→<u>VII. 3.</u>) eingetragen.

Bearbeiten/DNS 'zurückübersetzen'
 Unter Verwendung von IUPAC-Codes wird eine DNS-

Sequenz erzeugt, die für das Peptid kodiert.

# IV. PCR und Primer Design

In diesem Modul können Primer erstellt und "virtuelle PCRs" durchgeführt werden.

#### 1. Das PCR-Modul

#### Starten des PCR-Moduls

Das PCR-Modul kann aus dem DNS-Modul gestartet werden, entweder über das Menü Werkzeuge/PCR, oder über das Kontext-Menü in der Sequenz oder der DNS-Karte. Wenn eine Sequenz markiert ist, können automatisch Primer am Anfang bzw. am Ende der markierten Sequenz erstellt werden. Wenn die markierte Sequenz genau drei Nukleotide lang ist, können wahlweise auch Mutagenese-Primer erstellt werden. Anfangs-, End- und Mutageneseprimer können nur über das Kontext-Menü erstellt werden. Solche automatisch erstellten Primer sind in keiner Weise optimiert und bedürfen manueller Nachbearbeitung.

#### **Toolbar**

Die Toolbar (1) des PCR-Moduls bietet folgende Funktionen:

- Neuen Primer manuell eingeben. Dieser kann dann zur virtuellen PCR hinzugefügt werden (→IX.5).
- Arr Dateien (Primer) öffnen ( $\rightarrow$ IX.3).
- Primer hinzufügen. Nur gegenwärtig geöffnete Primer können zur virtuellen PCR hinzugefügt werden. Hinzugefügte Primer annealen automatisch in der optimalen Position und Richtung an der Sequenz.
- Primer exportieren. Basierend auf der Auswahl in der Primer-Liste wird ein neuer Primer generiert, aber noch nicht gespeichert.

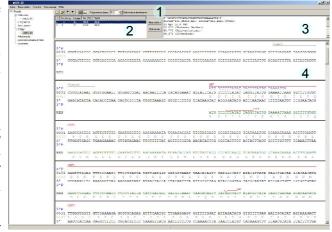


Abbildung IV.1: Das PCR-/Primer Design-Modul.

- Editiermodus (=F2).
- Features aus-/einblenden.
- Laufweite der (virtuellen) Polymerase. Dies ist das Äquivalent zur Elongationszeit am (realen) PCR-Gerät. Beim Start des PCR-Moduls und beim Hinzufügen von Primern wird automatisch eine optimale Laufweite geschätzt; diese kann dann bei Bedarf manuell geändert werden.

#### **Primer-Liste**

Die Primer-Liste (2) befindet sich unter der Toolbar. Sie zeigt alle in der PCR verwendeten Primer, zusammen mit einigen Eckdaten wie Länge, Schmelztemperatur und Richtung, in welcher der Primer gegenwärtig an die Sequenz annealt.

Wird ein Primer aus der Liste ausgewählt, werden weitere Details des Primers (Primersequenz, selfannealing etc.) im rechten Textfeld (3) dargestellt. Ein Doppelklick auf einen Primer in der Primer-Liste markiert diesen in der Template-Sequenz und zeigt ihn dort an.

Ein ausgewählter Primer kann

- über die Toolbar exportiert werden
- bearbeitet werden (→Dialog "Primer bearbeiten")
- entfernt werden

#### Sequenz

Die Sequenzdarstellung (4) umfasst folgende Zeilen:

- 1. Features des Templates (diese Zeile kann ausgeblendet werden).
- 2. 5'-Primer.
- 3. Template-DNS-Sequenz  $(5'\rightarrow 3')$ .
- 4. Template-Aminosäure-Sequenz. Der reading frame kann wie im DNS-Modul eingestellt werden.
- 5. Template-DNS-Sequenz (komplementär,  $3'\rightarrow 5'$ ).
- 6. 3'-Primer.
- 7. Restriktionsschnittstellen der resultierenden Sequenz.
- 8. Resultierende Sequenz (in grün,  $5'\rightarrow 3'$ ).
- 9. Resultierende Aminosäure-Sequenz. Der reading frame entspricht dem der Template-Sequenz.

#### Primersequenzen

Nur die beiden Primersequenzen (2. & 6. Zeile) können markiert und bearbeitet werden, wobei der "Überschreiben"-Modus fest eingestellt ist. Nukleotide können mit Hilfe der Leertaste gelöscht werden. Ein Druck auf die "Punkt" (1)-Taste kopiert das jeweilige Nukleotid aus der Template-Sequenz.

Nukleotide, die der Template-Sequenz entsprechen, werden blau dargestellt, "Mismatches" rot. Ist ein "leerer" Bereich der Primer-Sequenz ausgewählt, kann über das Kontext-Menü Auswahl als Neuen Primer der entsprechende Bereich der Template-DNS als neuer Primer hinzugefügt werden. Eingeben einer Nucleotidsequenz in einen leeren Bereich der Primer-Sequenz erzeugt *keinen* neuen Primer.

#### Resultierende Sequenzen

Resultierende DNS- und Aminosäure-Sequenzen können über das Kontext-Menü in die Zwischenablage kopiert oder als neue DNS-/Aminosäure-Sequenz erzeugt werden. Beim Kopieren bleibt *nur* die Sequenz erhalten, Features und Restriktionsenzyme gehen verloren.

# Spezielle Menüs

 Restriktionsschnittstelle rechts/links von der Markierung einfügen Öffnet einen Auswahl-Dialog, in dem ein Restriktionsenzym gewählt werden kann. Die entsprechende Schnittstelle wird dann rechts bzw. links von der aktuellen Markierung eingefügt.

# 2. Dialog "Primer bearbeiten"

Dieser Dialog wird aufgerufen, indem ein Primer aus der →Primer-Liste markiert wird, und dann der Dialog über Bearbeiten gestartet wird. Er dient der computergestützten Optimierung von Primern.

Rechts werden in einem Textfeld detaillierte Information zum aktuellen Primer dargestellt. Dies entspricht der Anzeige der Primer-Liste.

Links finden sich mehrere Grenzwerte, innerhalb derer automatisch Primer-Variationen erzeugt, untersucht und bewertet werden. Einstellbar sind:

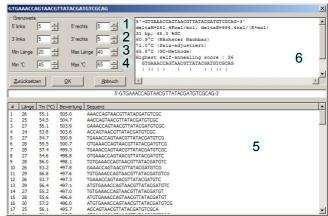


Abbildung IV.2: Dialog "Primer bearbeiten"

- Die Variation des 5'-Endes des Primers nach links bzw. rechts. (1)
- Die Variation des 3'-Endes des Primers nach links bzw. rechts. (2)
- Die minimale bzw. maximale Länge des Primers. (3)
- Die minimale bzw. maximale Schmelztemperatur des Primers. (4)

Bei jeder Änderung dieser Werte erfolgt eine sofortige Neuberechnung möglicher Primer unter Berücksichtigung der neuen Grenzwerte. Die berechneten Primer werden intern bewertet, und, nach absteigender Bewertung sortiert, zusammen mit einigen Eckdaten in einer Liste in der unteren Hälfte des Dialogs (5) angezeigt. Ein Doppelklick auf einen dieser Einträge legt die ausgewählte Variante als neuen Primer fest; dieser wird dann in der mittleren Zeile des Dialogs angezeigt, und das Textfeld (6) zeigt detaillierte Daten des neuen Primers an.

Mit Zurücksetzen wird der ursprünglich in der Primer-Liste ausgewählte Primer wieder hergestellt.

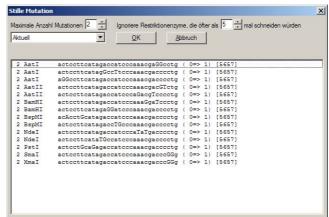
Mit OK wird der aktuelle Primer in das PCR-Modul übernommen. Er ersetzt dabei den ursprünglich in der Primer-Liste ausgewählten Primer.

Abbruch kehrt zum PCR-Modul zurück, ohne dieses zu ändern. Änderungen aus dem Dialog werden dabei verworfen.

# 3. Dialog "Stille Mutation"

Dieser Dialog wird aufgerufen, indem ein Primer (oder ein Teil eines Primers) in der Sequenz markiert wird. Der Dialog wird dann das Kontext-Menü STILLE über MUTATION aufgerufen.

Der Dialog hilft bei der Suche nach stillen Mutationen. Das sind Mutationen, welche eine neue Restriktionsschnittstelle einführen, die resultierende Aminosäuresequenz jedoch nicht verändern. Stille Mutationen sind hilfreich, um den Erfolg einer PCR zu überprüfen. Dabei wird die während der PCR erzeugte DNS mit dem Abbildung IV.3: Dialog "Stille Mutation"



gewählten Restriktionsenzym geschnitten. Die Anzahl bzw. Länge der resultierenden DNS-Fragmente zeigt, ob durch die PCR eine neue Restriktionsschnittstelle erzeugt wurde, was den Erfolg der PCR nachweist.

Die Suche nach stillen Mutationen kann im Dialog wie folgt begrenzt werden:

- Maximale Anzahl Mutationen. Da Primer möglichst gut zu ihrer Zielsequenz passen sollen, ist die Zahl der eingeführten Mutationen möglichst gering zu halten.
- Ignoriere Restriktionsenzyme, die öfter als X mal schneiden würden. Wird der gesamte Vektor von einem Restriktionsenzym in viele Fragmente zerschnitten, ist eine zusätzliche Restriktionsschnittstelle nur schwer oder gar nicht zu erkennen.
- Die Gruppe der Enzyme, für die stille Mutationen gesucht werden, kann bestimmt werden. Alle Enzyme zu durchsuchen kann auf langsamen Computern lange dauern, und bringt viele "gute" Resultate, die in der Praxis oft am Fehlen der Enzyme scheitern. Daher ist beim Start des Dialogs die Gruppe der aktuellen Enzyme ausgewählt.

Wird eine dieser Einstellungen verändert, werden alle passenden stillen Mutationen sofort neu berechnet. Die Liste der stillen Mutationen zeigt in jeder Zeile

- die Anzahl der nötigen Mutationen zur Einführung dieser Schnittstelle,
- den Namen des Restriktionsenzyms,
- die mutierte Sequenz, wobei die originalen Nukleotide als Kleinbuchstaben und die Mutationen als Großbuchstaben dargestellt werden,

- die Anzahl der Schnittstellen dieses Enzyms, vor und nach den Mutationen, in der Form "vorher => nachher", sowie
- die (bei erfolgreicher PCR) entstehenden DNS-Fragmente (in eckigen Klammern).

Ein Doppelklick auf eine der Zeilen (bzw. die Auswahl einer Zeile und das Betätigen von OK) übernehmen die Mutation(en) der ausgewählten Zeile; ausserdem wird das ausgewählte Restriktionsenzym, sofern nötig, zur Anzeige der PCR hinzugefügt.

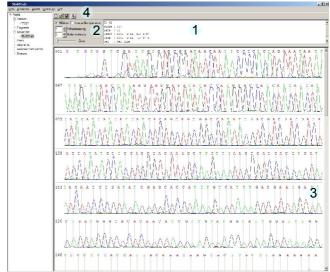
# V. Sequenzierung

Sequenzierungs-Modul kann Sequenzierungs-Dateien im ABI-Format lesen, darstellen und bietet die Möglichkeit, diese zu bearbeiten. Es werden Informationen über die Sequenzierung angezeigt (1).

# 1. Darstellung

Folgende Darstellung-Optionen (2) Sequenz (3) werden angeboten:

- HILFSLINIEN sind graue, senkrechte Linien, die erkannte Nukleotide mit den Peaks der Grafik zu verbinden suchen.
  - Invers & komplementär kann einen sequenzierten  $3' \rightarrow 5'$ -Strang die entsprechende 5'→3'-Sequenz umwandeln. Dadurch kann Abbildung V.1: Das Sequenzierungs-Modul das Alignment mit der "erwarteten" Sequenz aligned werden.



- Höhenskalierung gibt an, wie viele Textzeilen für die Darstellung der Grafik genutzt werden sollen.
- Breitenskalierung gibt an, wie dicht die einzelnen Peaks aufeinander folgen sollen.
- Zoom vergrößert die Darstellung, damit auch kleinere Peaks gesehen werden können.
- Ein Textfenster zeigt Details der Sequenzierung an.

# 2. Bearbeitung

Im Editiermodus können Nukleotide ersetzt werden; ein Einfügen oder Entfernen von Nukleotiden ist nicht möglich. Beim Speichern der Sequenzierung wird nur die Sequenz gespeichert, die Grafik geht verloren.

#### 3. Toolbar

Funktionen der Toolbar (4):

- D Neue Sequenz manuell eingeben  $(\rightarrow IX.5)$ .
- projection (Control of the Control o Dateien öffnen  $(\rightarrow IX.3)$ .
- Sequenz speichern. Sequenzierungs-Dateien werden nur als Sequenz in der Datenbank gespeichert; die Daten (Peaks) werden *nicht* mit gespeichert ( $\rightarrow IX.3$ ).
- Markierung bzw. komplette Sequenz als Text in die Zwischenablage kopieren.

Horizontal Zeigt die Sequenzierung als eine lange Zeile an.

# VI. Alignments

Das Alignment-Modul ermöglicht die Berechnung und Darstellung von DNS- und Aminosäure-Sequenz-Alignments.

# 1. Starten des Alignment-Moduls

Das Modul kann über das Menü Werkzeuge/Alignment oder das Tastenkürzel Strg-Gaufgerufen werden.

Beim Start des Moduls erscheint der Einstellungs-Dialog, der auch später über den entsprechenden Button aufgerufen werden kann. Beim Import eines Alignments aus einer Datei (über Datei/Importieren) wird dieser Dialog nicht aufgerufen, da bereits ein Alignment vorliegt, das nur noch dargestellt werden muss.

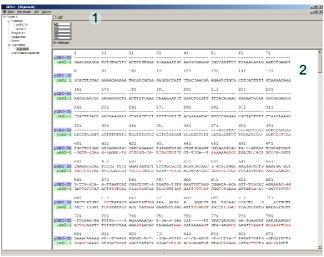


Abbildung VI.1: Das Alignment-Modul

# 2. Dialog "Einstellungen"

Der Einstellungs-Dialog wird entweder automatisch oder über den entsprechenden Button aufgerufen. Hier können die zu alignenden Sequenzen und deren initiale Reihenfolge, sowie der zu verwendende Alignment-Algorithmus und dessen Parameter festgelegt werden. Als Algorithmen stehen zur Verfügung:

- 1. **Clustal-W**. Dieser Algorithmus ist standardmäßig ausgewählt und benutzt ein externes Programm zur Berechnung. Clustal-W erzeugt *globale Alignments* (Sequenzen, die etwa gleich lang sind) von hoher Qualität.
- 2. **Smith-Waterman**. Ein interner Algorithmus für *lokale Alignments* (mehrere kurze Sequenzen werden gegen eine längere aligned). Dabei ist die erste Sequenz die lange "Master"-Sequenz, gegen die alle anderen Sequenzen aligned werden. Der Algorithmus ist schneller als Clustal-W, eignet sich aber nur für einfache Alignments, insbesondere zur Überprüfung von Sequenzierungen.
- 3. **Needlemann-Wunsch**. Ein interner Algorithmus für globale Alignments. Der Algorithmus ist schneller, aber qualitativ schlechter als Clustal-W. Wie bei Smith-Waterman werden alle Sequenzen gegen die erste Sequenz aligned, *nicht* untereinander. Der Algorithmus eignet sich zum schnellen Vergleich mehrerer, (sehr) ähnlicher Sequenzen.

Achtung: Das Beenden des Dialogs mit OK führt zur Neuberechnung des Alignments. Alle manuell im Alignment vorgenommenen Änderungen (eingefügte Lücken etc.) gehen dabei verloren!

# 3. Das Alignment-Modul

# Toolbar (1)

Neue Sequenz manuell eingeben  $(\rightarrow IX.5)$ .

Dateien öffnen  $(\rightarrow IX.3)$ .

Alignment speichern  $(\rightarrow IX.3)$ .

HORIZONTAL Zeigt das Alignment als eine lange Zeile an. Das kann die Übersichtlichkeit beim

Alignment vieler Sequenzen erhöhen.

Mittlere Maustaste Aktion beim Betätigen der mittleren Maustaste. Alle Aktionen können auch in der

Sequenz über das Kontext-Menü ausgewählt werden.

Einstellungs-Dialog.

#### Menü "Ansicht"

• Fett Stellt alle Zeichen fett dar.

Mono Monochrom-Darstellung (keine Farben).

• Konserviert Stellt einen Punkt '.' dar, falls die Base/Aminosäure der in der 1. Sequenz entspricht.

Normal Standard-Darstellung

• INVERTIERT Weiße Schrift auf farbigem Hintergrund (Ausnahme: Lücken)

STÄRKE DES ALIGNMENTS [nicht implementiert]
 STÄRKE DES ALIGNMENTS (INVERTIERT) [nicht implementiert]
 ÄHNLICHKEIT [nicht implementiert]
 FREQUENZ [nicht implementiert]
 FEATURES [nicht implementiert]
 RNA [nicht implementiert]

IDENTITÄT Markiert gleiche/ähnliche Basen/Aminosäuren in einer zusätzlichen "Sequenz".

# Sequenzdarstellung (2)

Die Sequenzdarstellung des Alignments kann über das Kontextmenü angepasst werden. Dabei wird das Alignment *nicht* neu berechnet, lediglich die Darstellung wird verändert.

- · Zeilen können umgeordnet werden.
- Features der Sequenz können angezeigt bzw. ausgeblendet werden. Standardmäßig werden die Features der ersten Sequenz angezeigt, die der anderen Sequenzen werden ausgeblendet.
- Lücken können eingefügt bzw. entfernt werden (in einer Zeile bzw. in allen anderen).

Ein Doppelklick auf eine Base/Aminosäure (*nicht* auf eine Lücke) öffnet das entsprechende Fenster und markiert die Stelle dort. Das kann z.B. bei der Überprüfung von Sequenzierungen hilfreich sein. Sequenzen können in der Alignment-Darstellung *nicht* direkt bearbeitet werden. Dazu müssen die Orignalsequenz bearbeitet und ein neues Alignment errechnet werden.

#### VII. Rechner

Das Rechner-Modul kann über das Menü Werkzeuge/Rechner aufgerufen werden. Es umfasst mehrere spezialisierte Rechner für typische molekularbiologische Aufgaben. Dabei werden generell vom Benutzer einzugebende Werte blau unterlegt, sowie die Ergebnisse der Berechnung in roter Schrift dargestellt.

# 1. Rechner "Ligation"

Dieser Rechner ermittelt die zu pipettierenden Mengen von Insert und Vektor für eine Ligation. Einzugeben sind die Länge von Insert bzw. Vektor, deren jeweils vorliegende Konzentration, sowie das gewünschte Verhältnis von Insert zu Vektor (typischerweise 4:1 oder 5:1).

Ebenfalls anzugeben ist die gewünschte Gesamtmasse der ligierten Moleküle. Ist keine bestimmte Gesamtmasse erforderlich, kann über diesen Wert auch das Ergebnis dahingehend beeinflußt werden, dass "pipettierbare" Mengen entstehen (z.B. 2µl statt 0.34µl).

# 2. Rechner "DNS-Konzentration"

Aus Photometerwerten der Absorption einer DNS-Lösung bei 260 bzw. 280 nm können Reinheit und Menge der DNS errechnet werden. Wurde eine verdünnte DNS-Lösung gemessen, kann über die Angabe der Verdünnung die DNS-Menge der Stammlösung errechnet werden; für unverdünnte DNS-Lösung ist ein Verhältnis 1:1 (Eingabe "1") zu verwenden.

Bei der Messung von einzelsträngiger DNS/RNA bzw. Oligonukleotiden ist der entsprechende Korrekturfaktor einzugeben.

#### 3. Rechner "Protein"

Aus Photometerwerten der Absorption einer Protein-Lösung bei 250 bzw. 280 nm können Protein-Reinheit und Konzentration berechnet werden. Die Daten können manuell eingetragen werden, oder aus dem Peptid-Modul (→III.) über das Menü Bearbeiten/Photometer-Analyse automatisch generiert werden.

# VIII. Bildbetrachter

Das Bildbetrachter-Modul kann über das Menü Werkzeuge/Bildbetrachter aufgerufen werden. Es dient zur Darstellung von digitalen Bildern, vornehmlich Gel-Bildern. Unterstützt wird momentan das IMG-Format von BioRad.

Über den Button in der linken oberen Ecke des Moduls kann das Verzeichnis gewechselt werden. Alle Dateien des Verzeichnisses werden in der Liste unter dem Button angezeigt. Das Auswählen einer Datei (einfacher Klick!) zeigt das entsprechende Bild an. Über das Kontext-Menü (im Bild klicken) kann das Bild in die Zwischenablage kopiert, als .BMP-Datei gespeichert oder ausgedruckt werden.

#### Bekannte Fehler:

- Möglicherweise wirkt das Bild bei der Darstellung gerastert. Auf einem Ausdruck erscheint es aber trotzdem normal.
- Wurde das Bild in der BioRad-IMG-Datei mit Beschriftungen versehen, werden diese möglicherweise leicht versetzt angezeigt.
- Wird die Bilddarstellung von einem anderen Fenster überlagert, verschwindet die Darstellung. Erneutes Auswählen der Datei stellt das Bild wieder korrekt dar.

# IX. Werkzeuge und Dialoge

# 1. Ligation

Der Ligations-Dialog dient zur virtuellen Ligation zweier oder mehrerer DNS-Fragmente. Er kann über das Menü Werkzeuge/Ligation oder Strg-L erreicht werden.

Die linke Seite das Dialogs zeigt DNS-Sequenzen, die ligiert werden können. Einige davon werden zu Beginn automatisch ausgewählt, die Auswahl kann aber manuell verändert werden.

Die rechte Seite des Dialogs zeigt die Liste aller möglichen Produkte der Ligation dar, basierend auf der Auswahl auf der linken Seite. *Bei vielen möglichen Ligationsprodukten kann die Neuberechnung länger dauern*. Ausgewählte Ligationsprodukte werden durch den Button Ligieren als neue DNS-Module erzeugt.

Bei der Berechnung möglicher Ligationsprodukte wird jedes zur Verfügung stehende Fragment höchstens einmal verwendet. Bei circulären Ligationsprodukten werden zwei Formen ("A-B" und "B-A") angezeigt, die sich nicht in der Sequenz, jedoch im Layout unterscheiden. Im Zweifelsfall sollten beide erzeugt, verglichen, und dann eines gelöscht werden.

# 2. Einstellungen

Über das Menü Werkzeuge/Einstellungen können verschiedene Einstellung für das Programm gewählt werden.

Sprache Momentan werden Englisch und Deutsch unterstützt.

Bessere Anzeige zu langsam ist.

Zeige Sequenzname Titel einer DNS-Sequenz in der Kartendarstellung zeigen.
 Zeige Sequenzlänge Länge einer DNS-Sequenz in der Kartendarstellung zeigen.

• Letztes Projekt laden Beim Programmstart das letzte bearbeitete Projekt laden.

Metafile-Format
 Bilder f
ür die Zwischenablage im Metafile-Format.

SPLASHSCREEN GENtle-Bild beim Programmstart.

Auf neue Version prüfen Internet-Abfrage, ob eine neue GENtle-Version verfügbar ist.

• Interne Hilfe Internes Hilfe-Fenster (sonst: Standard-Web-Browser).

# 3. Öffnen/Speichern/Datenbank verwalten

Dieser Dialog kann wie folgt erreicht werden:

- Über 🔁 oder 🖫 in der Toolbar
- Über Strg-O oder Strg-S
- Über die Menüs Datei/Öffnen oder Datei/Speichern
- Über das Menü Werkzeuge/Datenbank verwalten

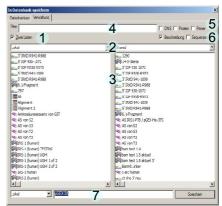
# Verwaltung

Hier werden die Einträge in den Datenbanken angezeigt. Einträge können gesucht, umbenannt, kopiert, verschoben, gelöscht und ggf. gespeichert werden.

#### Zwei Listen

Über Zwei Listen (1) kann zwischen einer und zwei parallelen Datenbank-Ansichten umgeschaltet werden. Die zwei-Listen-Ansicht kann benutzt werden, um Einträge zu kopieren und zu verschieben; ausserdem wird in beiden angezeigten Datenbanken gesucht. Die Datenbank kann aus der Drop-Down-Box (2) über der jeweiligen Liste (3) ausgewählt werden. Standardmäßig sind die lokale und die Standard-Datenbank ausgewählt.

In der zwei-Listen-Ansicht können Einträge verschoben werden, indem sie mit der Maus in der einen Liste "gegriffen" und in der anderen Liste "fallengelassen" werden. Wird während des Abbildung IX.1: Datenbank verwalten. "Fallenlassens" Strg gedrückt gehalten, wird der Eintrag kopiert.



#### Suchen

Über den Filter (4) können Einträge effizient gesucht werden. Ein oder mehrere Suchbegriffe können dort eingetragen werden; Suchbegriffe werden durch Leerzeichen getrennt, und nur Einträge, die alle Suchbegriffe beinhalten, werden angezeigt.

Standardmäßig werden alle Typen von Einträgen durchsucht und angezeigt. Über DNS, Protein und PRIMER (5) kann man die Suche auf die angewählten Typen einschränken.

Standardmäßig wird auch die Sequenz-Beschreibung durchsucht; dies kann durch Abwählen von Beschreibung unterbunden werden. Zusätzlich kann auch die Sequenz selbst durchsucht werden (6).

#### Öffnen

Einen einzelnen Eintrag kann man durch einen Doppelklick öffnen. Mehrere Einträge können, als Gruppe durch Ziehen eines Rahmens mit der linken Maustaste oder durch Auswahl mehrerer Einträge bei gedrückter Strg-Taste, ausgewählt und über das Kontext-Menü Öffnen oder durch drücken der Eingabetaste (Return oder Enter) geöffnet werden.

Über das Kontext-Menü können Einträge umbenannt oder permanent gelöscht werden.

#### Speichern

Soll eine Sequenz gespeichert werden, sind unten im Dialog neue Elemente (7) zu sehen:

- Eine Drop-Down-Box mit der Datenbank, in der die Sequenz gespeichert werden soll. Bei einer neuen Sequenz ist die Standard-Datenbank ausgewählt, sonst die Datenbank, aus der die Sequenz ursprünglich geladen wurde.
- Der Name, unter dem die Sequenz gespeichert werden soll.
- Der Speichern-Button.

Bestehende Einträge werden (mit Nachfrage) überschrieben, wenn die bestehende Sequenz identisch ist. Sollte die zu speichernde Sequenz sich in ihrer Nukleotid- bzw. Aminosäureabfolge von der bestehenden unterscheiden, muss sie unter einem anderen Namen gespeichert werden, oder die bestehende Sequenz muss zuerst über das Kontext-Menü gelöscht werden.

Wird ein Eintrag in der/einer der Liste(n) ausgewählt, ändert sich der Name der zu speichernden Sequenz entsprechend. Das ist hilfreich, wenn der neue Name der Sequenz einem bestehenden ähneln soll.

#### **Datenbank**

Hier können ganze Datenbanken neu erstellt oder existierende hinzugefügt werden. Außerdem können Einträge für Datenbanken (nicht die Datenbanken selbst!) entfernt (Entfernen) oder zur Standard-Datenbank erklärt werden (ALS STANDARD).

Datei-Datenbanken (Sqlite) und MySQL-Datenbanken können über die entsprechenden Hinzuffügen-Buttons zur Liste der Datenbanken hinzugefügt werden. Bei Datei-Datenbanken wird dabei nur der Dateiname benötigt; MySQL-Datenbanken benötigen mehr Informationen. Datei-Datenbanken sind einfach anzulegen (Neu) und zwecks Transport oder Backup zu duplizieren (kopieren der Datei genügt), werden aber bei großen Datenbeständen oder beim Zugriff über Netzwerk zunehmend langsamer. MySQL-Datenbanken sind komplizierter anzulegen und zu verwalten, bieten aber einen Geschwindigkeitsvorteil. Die lokale Datenbank ist eine Datei-Datenbank. Sie wird beim ersten Programmstart automatisch angelegt und kann nicht entfernt werden.

#### 4. Importieren

Der Importieren-Dialog kann über das Menü Datei/Importieren oder Strg-I erreicht werden. Über ihn können Sequenzen aus Dateien in GENtle geladen werden. Standardmäßig prüft GENtle jede der ausgewählten Dateien auf ihr Format (mehrere Dateien können durch das Gedrückthalten von Strg ausgewählt werden). Es kann aber auch ein Format vorgegeben werden. Folgende Formate können importiert werden:

- GenBank
- GenBank XML
- Clone
- ABI/AB1 (Sequenzierer-Format)
- FASTA
- PDB (3D-Format, Import als annotierte Sequenz)
- Eine Reihe weiterer, einfacher Formate

# 5. Sequenz eingeben

Der Eingabe-Dialog kann über das Menü Datei/Sequenz eingeben oder über Strg-N erreicht werden. Über ihn können Sequenzen manuell eingegeben werden. Es können Titel der Sequenz, Format/Typ, sowie die Sequenz selbst eingegeben werden. An Formaten stehen zur Auswahl:

- DNS
- Aminosäuren
- GenBank
- GenBank-XML
- Primer

Bei den Formaten DNS und Aminosäuren werden nicht-passende Zeichen (z.B. Ziffern) automatisch entfernt. Handelt es sich bei der Sequenz um einen Primer, *muss* als Typ auch Primer angegeben werden (*nicht* DNS).

# 6. Sequenz-Editor

Im Sequenz-Editor können Eigenschaften der Sequenz, Features, sowie die zu verwendenden Restriktionsenzyme und Proteasen angegeben und bearbeitet werden. Alle Änderungen werden erst durch OK in die Sequenz übernommen.

# Eigenschaften

Hier können Sequenzname und -Beschreibung geändert werden. Handelt es sich um eine DNS-Sequenz, können hier auch *sticky ends* angegeben werden.

#### **Features**

Hier werden die Features der Sequenz angezeigt. Es können Features hinzugefügt, gelöscht und

verändert werden. Änderungen an einem Feature müssen nicht bestätigt werden, es genügt, ein anderes Feature auszuwählen oder OK zu drücken.

Die Einstellung Leseraster ist nur verfügbar, wenn als Typ "Kodierende Sequenz" ausgewählt wurde. Wenn der Punkt Kodierend angewählt ist, wird dir Sequenz  $5' \rightarrow 3'$  gelesen (Standard), andernfalls  $3' \rightarrow 5'$ . Auswahl Löschen bezieht sich nur auf die *Auswahl* eines Features; um ein Feature zu löschen, Feature Löschen benutzen.

Über Feature Bearbeiten können verschiedene Darstellungs-Varianten des Features gewählt werden:

• FÜLLFARBE Die Farbe, mit der das Feature in der Kartendarstellung gezeigt wird.

• Darstellung in Sequenz Die Form des Features in der Sequenzdarstellung.

• Offset benutzen Aminosäure-Offset in der Sequenzdarstellung.

# Restriktionsenzyme

Auf der linken Seite werden alle aktuell in der Sequenz verwendeten Restriktionsenzyme angezeigt. Auf der rechten Seite sind alle definierten Enzym-Gruppen angezeigt. Die Enzyme in der ausgewählten Gruppe werden rechts unten angezeigt. Enzyme können zwischen den Listen über die mittig angeordneten Buttons bewegt werden.

#### **Proteasen**

Hier können Proteasen ausgewählt werden, deren Schnittstellen dann in der Sequenz angezeigt werden. Neue Proteasen können zur Liste hinzugefügt werden.

• Beispiel: Thermolysin

• Sequenz: ! DE | AFILMV

• Erklärung: "nicht D oder E", "dann Schnitt", "dann A, F, I, L, M oder V"

• Beispiel: Proline-endopeptidase

• Sequenz: HKR, P | ! P

• Erklärung: "H, K oder R", "dann P", "dann Schnitt", "dann nicht P"

#### 7. Restriktions-Assistent

Der Restriktions-Assistent kann über das Menü Werkzeuge/Restriktions-Assistent oder über die Auswahl einer Restriktionsschnittstelle in der Kartendarstellung einer DNA-Sequenz durch die mittlere Maustaste gestartet werden ( $\rightarrow II.4$ ). Im letzten Fall wird das ausgewählte Enzym automatisch markiert.

Auf der linken Seite des Dialogs werden die verfügbaren Enzyme angezeigt, abhängig von den Einstellungen unter Gruppe und Unterauswahl. Die Liste kann nach Enzym oder Anzahl der Schnitte durch Klicken auf den entsprechenden Titel sortiert werden. Ist ein Enzym ausgewählt, werden die Schnittpositionen und Fragmente unten links angezeigt.

Die rechte Liste zeigt den Restriktions-"Cocktail". Mit den Enzymen im "Cocktail" wird später die Restriktion durchgeführt. Die dabei entstehenden Fragmente werden rechts unten angezeigt. Enzyme können durch Zum Cocktail hinzugefügt und durch Restriktionsenzym entfernen aus dem "Cocktail" gelöscht werden. Über Fertig gelang man zurück zur Sequenz, wobei der "Cocktail" gespeichert wird; Abbruch verläßt den Restriktions-Assistenten und verwirft die vorgenommenen Änderungen.

Starte Restriktion erzeugt die Fragmente als neue DNA-Module, wobei die Originalsequenz erhalten bleibt. Keine Fragmente unter \_\_ Basenpaare erstellen limitiert dabei die Fragmente auf eine Mindestgröße; diese Option ist abschaltbar.

# 8. Projekte

Ein Projekt ist eine Sammlung von Sequenzen, die zusammengehören. Projekte können über Datei/Projekt speichern bzw. F12 gespeichert oder über Datei/Projekt Laden bzw. F11 geladen werden. Je nach Einstellungen ( $\rightarrow$ IX.2) wird das zuletzt verwendete Projekt beim Start von GENtle automatisch geöffnet.

Projekte speichern nur die Namen/Datenbanken der Sequenzen des Projekts, *nicht* die Sequenzen selbst. Wird eine der Sequenzen des Projekts verschoben, umbenannt oder gelöscht, wird diese Sequenz beim nächsten Öffnen des Projekts nicht mit geöffnet werden. Ein entsprechender Hinweis wird dabei angezeigt.

#### 9. Drucken

Sequenz- und Kartendarstellungen können über deren Kontext-Menü gedruckt werden. Über das Menü Datei/Bericht drucken kann auch eine Übersichtsseite gedruckt werden, auf der sowohl die Kartendarstellung als auch eine Liste der Features und Daten der Sequenz dargestellt werden. Der Bericht umfasst *nicht* die Sequenz selbst.

# 10. Sequenzdarstellung

- Die gesamte Sequenz kann mit Strg-A oder über das Menü Bearbeiten/Alles markieren ausgewählt werden.
- Mit Strg-F oder über das Menü Bearbeiten/Suchen kann innerhalb einer Sequenz gesucht werden.

# 11. Restriktionsenzym-Editor

Dieser Dialog dient zur Verwaltung von Restriktionsenzymen. Er entspricht dem Restriktionsenzym-Teil des Sequenz-Editors  $(\rightarrow \underline{IX.6})$ .

# X. FAQ

Frequently asked questions – Häufig gestellte Fragen

- O: Warum will GENtle sich dauernd mit dem Internet verbinden?
- A: Eine Internetverbindung muss für BLAST- und Rebase-Suchen vorhanden sein.
- A: Beim Start von GENtle wird per Internet überprüft, ob eine neue Version von GENtle verfügbar ist. Dies kann über Werkzeuge/Einstellungen unterbunden werden.
- Q: Warum kann ich bei einer ausgewählten DNA-Sequenz keine BLAST-Suche für die Aminosäuresequenz durchführen oder Aminosäuren extrahieren?
- A: Es muss ein Leseraster ausgewählt sein.