

## 4 Molekularbiologische Grundlagen des DNA-Computing

Die praktische Ausführung von Rechenvorgängen mittels technischer Geräte ist immer an die Anwendung von Naturgesetzen gebunden. Mechanische Rechenmaschinen realisieren Rechenvorgänge beispielsweise durch koordinierte Bewegungen von Zahnrädern, Walzen und ähnlichen Bauelementen, die auf den physikalischen Gesetzen der Mechanik basieren. Elektronische Computer nutzen die steuerbare elektrische Leitfähigkeit bestimmter Werkstoffe, indem sie Stromflüsse und elektrische Spannungen innerhalb eines Leitungsnetzwerkes gezielt beeinflussen, um auf diese Weise Rechenvorgänge praktisch umzusetzen. Ihrem Arbeitsprinzip liegen naturwissenschaftliche Gesetzmäßigkeiten der Elektrotechnik zugrunde. Abhängig davon, welche konkreten Naturgesetze zur Ausführung von Rechenvorgängen herangezogen werden, unterscheidet man verschiedene Computingkonzepte. Das DNA-Computing stützt sich im Gegensatz zu konventionellen Computingkonzepten auf Prinzipien der Biochemie sowie der Molekularbiologie. Dieses Kapitel vermittelt das für das laborpraktische DNA-Computing relevante Grundlagenwissen.

Das Erbmolekül DNA (Desoxyribonucleinsäure, deoxyribonucleic acid) ist neben der RNA (Ribonucleinsäure, ribonucleic acid) das wichtigste Medium zur Datenspeicherung in der Natur. Die Erbinformationen über Aufbau und Funktion aller derzeit bekannten zellulären Organismen werden durch spezifische DNA- oder RNA-Moleküle kodiert und bei der Fortpflanzung an die Individuen der nächsten Generation weitergegeben. DNA besitzt im Vergleich zu RNA favorisierte Eigenschaften wie beispielsweise eine größere Stabilität der Moleküle und eine höhere Spezifität vieler Reaktionen. Es liegt daher nahe, das von der Natur hervorgebrachte Speichermedium DNA als Datenträger im Sinne der Informatik einzusetzen.

DNA lässt sich durch eine Vielzahl molekularbiologischer Prozesse gezielt modifizieren und nach verschiedenen Kriterien analysieren. Als molekularbiologische Prozesse dienen biochemische Reaktionen und spezielle physikalische Abläufe, die bestimmte Eigenschaften der DNA ausnutzen. Die naturwissenschaftlichen Gesetzmäßigkeiten, auf denen diese molekularbiologischen Prozesse beruhen, bilden in ihrer Gesamtheit ein Computingkonzept, das zur Realisierung beliebiger Rechenvorgänge geeignet ist. Die Ausführung eines molekularbiologischen Prozesses auf einem DNA-Pool entspricht dabei einem Rechenschritt. Insofern kann jeder geeignete molekularbiologische Prozess auf DNA als eine Operation im Sinne der Mathematik und Informatik aufgefasst werden. Eine definierte Prozessabfolge stellt die Abarbeitung eines Algorithmus dar. Das Verständnis dieser Prozesse sowie die Kenntnis der damit einhergehenden Techniken und Methoden ist Voraussetzung für erfolgreiche Implementierungen DNA-basierter Algorithmen im molekularbiologischen Labor.

Viele molekularbiologische Prozesse vollziehen sich auf natürliche Weise in der lebenden Zelle (*in vivo*) und bewirken beispielsweise die Replikation (Duplizierung) von DNA bei der Zell-

teilung. Infolge umfangreicher Entdeckungen und Entwicklungen auf dem Gebiet der „Wissenschaften des Lebens“ gelingt es in zunehmendem Maße, molekularbiologische Prozesse auf DNA auch außerhalb lebender Zellen in Reagenzgläsern (*in vitro*) ablaufen zu lassen. Ergänzend wurden weitere molekularbiologische Prozesse sowie entsprechende Methoden zu ihrer Ausführung *in vitro* erdacht, die kein unmittelbares Vorbild in der lebenden Zelle besitzen, aber vor allem die Auswertung der in der DNA gespeicherten Informationen mit den menschlichen Sinnen ermöglichen. Bis auf sehr wenige Ausnahmen wird das praktische DNA-Computing auf dem heutigen Stand der Technik *in vitro* betrieben. Das verfügbare Repertoire laborpraktisch nutzbarer molekularbiologischer Prozesse hat sich in den vergangenen Jahren rasant erweitert. Die Wissenschaften des Lebens gelten als Schlüsseltechnologie des 21. Jahrhunderts. Es ist abzusehen, dass die Entdeckung und Synthese neuer Biokatalysatoren im Einklang mit Fortschritten auf dem Gebiet der Strukturbioologie zu weiteren, bisher nicht hinreichend verstandenen Möglichkeiten zur Verarbeitung von DNA-Informationen führen wird.

Das Kapitel gibt einen Überblick über den Stand der Technik. Die Gliederung orientiert sich an einer pragmatischen Herangehensweise, die der Methodik bei der laborpraktischen Implementierung von Algorithmen des DNA-Computing aus dem Blickwinkel der Informatik nachempfunden ist. Ausgehend von der Beschreibung der chemischen DNA-Struktur einschließlich ihrer Konformationen (räumlichen Anordnung der Atome) und der daraus ableitbaren Eigenschaften werden die molekularbiologischen Prozesse vorgestellt, die für das DNA-Computing *in vitro* relevant sind. Diese Prozesse lassen sich in fünf Gruppen einteilen. Die erste Gruppe enthält Methoden zur *Gewinnung von DNA*. Da Eingabedaten im laborpraktischen DNA-Computing kodiert in Form geeigneter DNA vorliegen müssen, werden Methoden benötigt, die spezifische DNA mit gewünschten Eigenschaften bereitstellen können. Betrachtet werden die Oligonucleotidsynthese sowie die Isolation von DNA aus Organismen. Methoden, die zur *Bereitstellung von DNA-Pools in wässriger Lösung* dienen, gehören zur zweiten Gruppe wie das Mischen, Aliquotieren und Verdünnen. Gemeinsames Merkmal dieser Prozesse ist, dass keine chemischen Bindungen innerhalb der DNA verändert werden und darüber hinaus auch keine gezielte DNA-Separation nach definierten Kriterien erfolgt. Die dritte Gruppe fasst diejenigen Prozesse zusammen, die ein *Knüpfen bzw. Aufbrechen von Wasserstoffbrückenbindungen* bewirken, wodurch einzel- und doppelsträngige DNA ineinander überführt werden kann (Hybridisierung, Denaturierung). Die vierte, bedeutendste und zugleich größte Gruppe bilden die *enzymatischen Reaktionen*, mit denen sich DNA auf vielfältige Weise rekombinieren lässt, indem die in der DNA gespeicherte Information gezielt modifiziert wird. Vorgestellt werden die Ligation, die Restriktionsspaltung, die Strangendenmodifikation, die Polymerisation sowie die Polymerase-Kettenreaktion. In die fünfte Gruppe werden Prozesse zur *Separation und Analyse von DNA* aufgenommen wie die Avidin-Biotin-Separation, die Gel-Elektrophorese und die im DNA-Computing häufig als Ausgabeoperation dienende Sequenzierung von DNA.

Neben der Beschreibung der chemischen bzw. physikalischen Grundlagen des jeweiligen Prozessablaufes wird auch dargelegt, wie sich die daraus resultierenden Labortechniken ausführen lassen. Einen besonderen Stellenwert nimmt die Analyse der *Seiteneffekte* ein, mit denen jeder molekularbiologische Prozess behaftet ist. Unter Seiteneffekten versteht man die nicht gezielt reproduzierbaren, sporadisch auftretenden und zumeist unerwünschten Wirkungen der Prozesse, die man beim DNA-Computing verhindern oder zumindest eindämmen bzw. kompensieren möchte. Eine abschließende Auswertung klassifiziert mögliche Seiteneffekte, nennt Methoden zu ihrem Nachweis und vergleicht die betrachteten Prozesse der einzelnen Gruppen hinsichtlich verschiedener Kriterien, die für die Konstruktion von DNA-Algorithmen wichtig sind.

## 4.1 DNA als Datenträger – Struktur und Eigenschaften

Die Nucleinsäuren gelten als Moleküle des Lebens, wobei der Begriff *Nuclein* davon abgeleitet ist, dass diese Substanzen erstmals in Zellkernen aufgrund ihrer Säureeigenschaften nachgewiesen wurden, jedoch ohne die gesamte chemische Struktur und die Funktion dieser Moleküle in der Zelle zu kennen. Außer DNA und RNA gehört auch PNA (Peptidnucleinsäure, peptide nucleic acid) zu dieser Stoffgruppe.

Im chemischen Sinne sind Nucleinsäuren strangartige Moleküle, die als *Polymere* bezeichnet werden. Darunter versteht man Stoffe, die sich aus einzelnen Bausteinen, den *Monomeren*, entsprechend eines definierten Aufbauprinzips zusammensetzen. Verschiedene Nucleinsäuren unterscheiden sich dadurch, welcher Satz an Monomeren jeweils zugrunde liegt und wie sich die einzelnen Bausteine chemisch zu einem Polymer verbinden können.

Prinzipiell eignen sich alle Polymere zur chemischen Kodierung von Informationen durch die räumliche Anordnung und Auswahl der Monomere. Damit die im Polymer gespeicherten Informationen gezielt gesetzt, verändert und ausgelesen werden können, bedarf es darauf abgestimmter, zuverlässig ausführbarer Methoden. Die Verwendung von DNA als Datenträger ist im Vergleich zu RNA, PNA und Mischformen derzeit am besten untersucht.

DNA kommt fast ausschließlich in Form von *DNA-Einzelsträngen* (Oligonucleotiden) und *DNA-Doppelsträngen* (Duplexen) vor. Unter speziellen Bedingungen können auch Triplexe und Quadruplexe auftreten, die in natürlicher DNA aber sehr selten sind und für das DNA-Computing keine Bedeutung haben.

### 4.1.1 DNA-Einzelstränge und ihre Primärstruktur

DNA-Einzelstränge bestehen aus einer linearen Abfolge (Kette) von *Nucleotiden*, die die monomeren Bausteine bilden. Jedes Nucleotid ist aus der Desoxyribose (einer Zuckerart), Phosphorsäure und einer Base aufgebaut. Man unterscheidet dabei vier mögliche Basen: Adenin, Cytosin, Guanin und Thymin. Die entsprechenden Nucleotide werden mit den Buchstaben A, C, G und T bezeichnet. Die Basen Adenin und Guanin werden durch den Oberbegriff Purin (Doppelringbasen) zusammengefasst, Cytosin und Thymin sind hingegen Pyrimidine (Einfachringbasen).

#### Definition 4.1 Nucleotid

Ein Nucleotid ist die kleinste informationstragende Einheit (monomerer Baustein) in DNA. Man unterscheidet die Nucleotide A, C, G und T.

Die einzelnen Nucleotide sind über Phosphodiesterbindungen zu einem Strang verknüpft. Jede Phosphodiesterbindung liegt zwischen zwei benachbarten Desoxyribosen. Die Desoxyribose ist ein Zucker, der eine ringförmige Struktur besitzt und fünf Kohlenstoffatome enthält. Die Positionen dieser Kohlenstoffatome innerhalb der Desoxyribose werden mit 1' bis 5' angegeben. Die Base ist an das 1'-Kohlenstoffatom kovalent gekoppelt. An das 3'- und das 5'-Kohlenstoffatom kann sich über Phosphodiester-Brücken jeweils eine weitere Desoxyribose anlagern, so dass eine Kette (ein Strang) entsteht. Die 3'- und die 5'-Bindungsstelle sind nicht nur

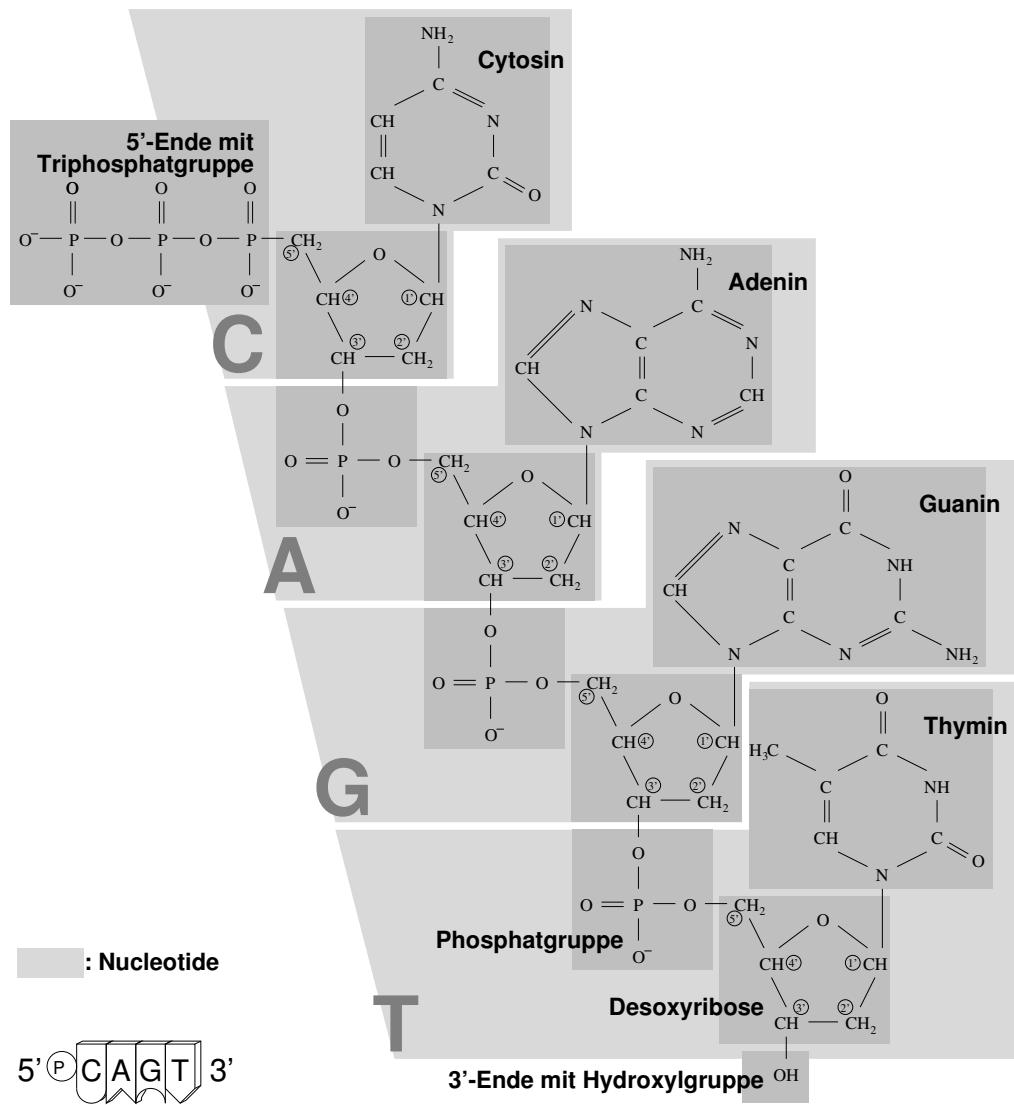


Abb. 4.1: Primärstruktur des DNA-Einzelstranges 5'P-CAGT-3'

räumlich voneinander getrennt, sie besitzen infolge der abweichenden Anordnung und Art der Atome in der lokalen Umgebung auch eine spezifische chemische Struktur. Man kann deshalb das *3'-Ende* und das *5'-Ende* eines DNA-Stranges voneinander unterscheiden. Dies bedeutet, dass DNA-Stränge gerichtet (orientiert) sind, wobei die *5'-3'*-Richtung als *sense* und die *3'-5'*-Richtung als *antisense* bezeichnet wird. Es ist üblich, Nucleotidsequenzen in *5'-3'*-Richtung zu notieren.

An *5'*- und *3'*-Enden von DNA-Strängen können spezifisch bestimmte chemische Gruppen oder Moleküle kovalent angelagert bzw. abgebaut werden, so dass jedes Strangende entsprechend chemisch markiert ist. Die Enden natürlich vorkommender DNA sind mit *Phosphatgruppen* ( $-\text{PO}_4$ ) oder *Hydroxylgruppen* ( $-\text{OH}$ ) versehen. Ein DNA-Strangende, an das eine Hydroxylgruppe gekoppelt ist, wird als *freies Ende* bezeichnet, weil anstelle der Hydroxylgruppe leicht eine andere Strangendenmarkierung angebracht werden kann. Insbesondere für Analyseverfahren stehen hierfür zahlreiche Moleküle zur Verfügung, wie beispielsweise *Biotin* und verschiedene *Fluoreszenzmarker*.

#### Definition 4.2 Strangendenmarkierung

Chemische Gruppen oder Moleküle, die spezifisch an *5'*- oder *3'*-Enden von DNA-Strängen kovalent angelagert oder abgebaut werden können, bezeichnet man als Strangendenmarkierung.

Analog zu den Nucleotiden werden auch Strangendenmarkierungen durch Symbole bezeichnet, die zur kompakten Notation von DNA-Strängen dienen. Phosphatgruppen erhalten unabhängig von der Anzahl die Bezeichnung P und Biotin die Bezeichnung B. Freie Enden werden nicht gesondert gekennzeichnet. Die Bezeichnungen der Strangendenmarkierungen dürfen nicht mit ihren chemischen Strukturformeln verwechselt werden.

Die Sequenz (Abfolge) der Basen in der Nucleotidkette bestimmen gemeinsam mit den Strangendenmarkierungen die Information, die der entsprechende DNA-Einzelstrang enthält (kodiert). Ein DNA-Einzelstrang kann folglich durch seine Nucleotidsequenz und seine Strangendenmarkierungen charakterisiert werden. Die Angabe der Nucleotidsequenz muss jedoch mit der Angabe der Leserichtung (*sense* oder *antisense*) einhergehen, um Verwechslungen mit der reversen Sequenz auszuschließen. Die durch die kovalenten Bindungen stabilisierte Nucleotidsequenz mit den Strangendenmarkierungen kennzeichnet die *Primärstruktur* jedes DNA-Einzelstranges.

#### Definition 4.3 DNA-Einzelstrang

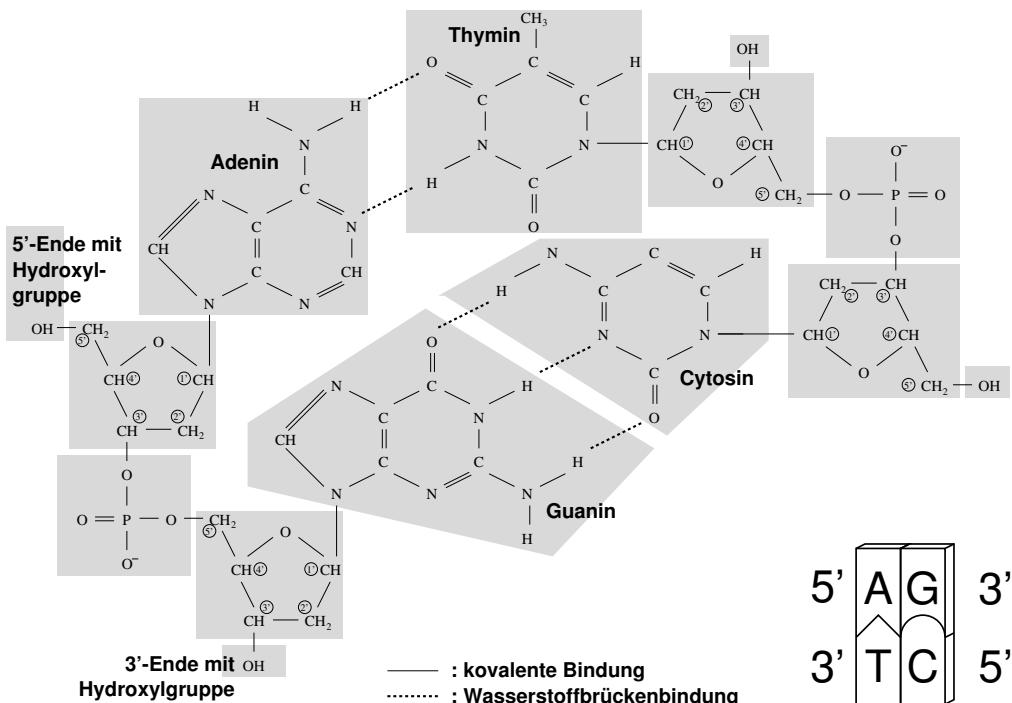
Ein DNA-Einzelstrang ist eine Sequenz der Nucleotide A, C, G und T. Die Enden der Nucleotidsequenz werden mit *5'* und *3'* angegeben. Jedes der beiden DNA-Einzelstrangenden trägt eine Strangendenmarkierung.

Die Abbildung 4.1 illustriert die Primärstruktur des DNA-Einzelstranges  $5'\text{P-CAGT-3}'$ . An das *5'-Ende* dieses Stranges ist eine Triphosphatgruppe angelagert, an das *3'-Ende* eine Hydroxylgruppe.

Jede im DNA-Strang befindliche Phosphatgruppe ist elektrisch negativ geladen, erkennbar durch die  $O^-$  in Abbildung 4.1. Dies bewirkt eine elektrisch negative Ladung des gesamten DNA-Stranges und somit seine hohe Polarität.

### 4.1.2 DNA-Doppelstränge und ihre Sekundärstruktur

DNA-Stränge besitzen unter bestimmten Voraussetzungen die Fähigkeit zur Basenpaarung. Basenpaarung kann zwischen benachbarter einzelsträngiger DNA auftreten. Dabei bilden sich zwischen je zwei gegenüberliegenden Basen Wasserstoffbrückenbindungen aus. Die Basen Adenin und Thymin können sich paaren, wobei zwei Wasserstoffbrückenbindungen entstehen. Die Basen Cytosin und Guanin können sich ebenfalls paaren. Hierbei entstehen drei Wasserstoffbrückenbindungen. Andere Paarungsmöglichkeiten der Basen gibt es nicht.



*Abb. 4.2: Basenpaarung im DNA-Doppelstrang  $^{5'}\text{AG}-^{3'}\text{TC}$  mit Kennzeichnung der Wasserstoffbrückenbindungen. An allen Stränden sind Hydroxylgruppen angelagert.*

Eine Wasserstoffbrückenbindung ist eine thermisch instabile chemische Bindung, bei der sich jeweils zwei Atome ein Proton ( $\text{H}^+$ ) teilen. Dieses Proton lässt sich als Brücke zwischen den beiden Atomen veranschaulichen. In DNA wird dieses Proton durch das an ein Sauerstoffatom (O) oder Stickstoffatom (N) kovalent gebundene Wasserstoffatom bereitgestellt. Auf der gegenüberliegenden Seite der Wasserstoffbrücke befindet sich entweder eine Amidgruppe ( $-\text{NH}$ ) oder eine Carbonylgruppe ( $-\text{CO}$ ). Eine Wasserstoffbrücke bildet sich aus, wenn sich das Proton und die gegenüberliegende Gruppe auf eine Entfernung von etwa 0,28nm nähern. Die

Bindung ist dann am stärksten, wenn die beteiligten Atome auf einer Linie angeordnet sind. Die Bindungsenergie liegt zwischen 12 und 29kJ/mol und beträgt somit nur etwa ein Zehntel der Energie kovalenter Bindungen innerhalb der DNA. Die Wasserstoffbrücken zwischen Adenin und Thymin sowie zwischen Cytosin und Guanin sind in Abbildung 4.2 durch gestrichelte Linien dargestellt. Cytosin-Guanin-Paare sind stärker gebunden als Adenin-Thymin-Paare, da sie eine Wasserstoffbrückenbindung mehr besitzen.

Infolge der Möglichkeit zur Basenpaarung bezeichnet man die Nucleotide A und T sowie die Nucleotide C und G als komplementär. Nur komplementäre Nucleotide können sich über Wasserstoffbrückenbindungen zusammenlagern. Der DNA-Strang und die in ihm enthaltenen Wasserstoffbrücken kennzeichnen seine *Sekundärstruktur*.

**Definition 4.4 Komplementarität von Nucleotiden**

Die Nucleotide A und T sind zueinander komplementär, ebenso die Nucleotide C und G.

Die Betrachtung der Basenpaarung lässt sich von einzelnen Nucleotiden auf Nucleotidsequenzen ausdehnen. Zwei gleichlange DNA-Einzelstränge können derart beschaffen sein, dass sich jede Base des einen Stranges mit der gegenüberliegenden Base des anderen Stranges paart. Dabei entsteht ein zweisträngiges Band komplementärer Nucleotide, ein DNA-Doppelstrang. Beide Stränge verlaufen in entgegengesetzter Richtung, der eine Strang von 5' nach 3', der andere Strang von 3' nach 5'.

Die entgegengesetzte Ausrichtung von einzelsträngiger DNA wird durch den Begriff *antiparallel* beschrieben. Er charakterisiert insbesondere eine räumliche Lage zweier DNA-Einzelstränge zueinander. Viele molekularbiologische Prozesse können nur dann in der gewünschten Weise ablaufen, wenn die als Reaktionspartner fungierenden DNA-Stränge eine bestimmte räumliche Anordnung haben.

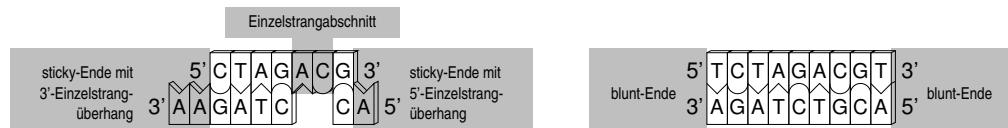
**Definition 4.5 antiparallel**

Entgegengesetzt ausgerichtete einzelsträngige DNA (ein Strang in 5'-3'-Richtung, der andre in 3'-5'-Richtung) wird als antiparallel bezeichnet.

**Definition 4.6 DNA-Doppelstrang**

DNA, bei der sich mehrere abschnittsweise antiparallel-komplementäre DNA-Einzelstränge (oder ein DNA-Einzelstrang mit sich selbst) durch Basenpaarung (Ausbildung von Wasserstoffbrücken) miteinander verbunden haben, wird als doppelsträngig bezeichnet, ein entsprechender DNA-Strang heißt DNA-Doppelstrang.

Jeder DNA-Doppelstrang weist mindestens ein durch Wasserstoffbrücken gebildetes Paar komplementärer Nucleotide auf. Die Beschaffenheit der äußeren Enden von DNA-Doppelsträngen



**Abb. 4.3:** Beispiele für sticky-Enden, blunt-Enden, Einzelstrangüberhänge und Einzelstrangabschnitte bei DNA-Doppelsträngen in linearisierter Darstellung. Alle 5'- und 3'-Enden sind frei.

beeinflussen den Ablauf zahlreicher molekularbiologischer Prozesse. Man unterscheidet äußere DNA-Doppelstrangenden mit und ohne *Einzelstrangüberhang*, wobei ein Einzelstrangüberhang durch eine zumeist sehr kurze ungepaarte Nucleotidsequenz an einem äußeren Doppelstrangende definiert ist, siehe Abbildung 4.3. DNA-Doppelstrangenden mit Einzelstrangüberhang werden als *sticky* (klebrig) bezeichnet, weil sie ein sehr effektives Zusammenfügen (Verkleben) von DNA-Doppelsträngen gestatten. DNA-Doppelstrangenden ohne Einzelstrangüberhang nennt man *blunt* (glatt).

#### Definition 4.7 blunt-Ende

Ein äußeres DNA-Doppelstrangende ist blunt, wenn es keinen Einzelstrangüberhang besitzt.

#### Definition 4.8 sticky-Ende

Ein äußeres DNA-Doppelstrangende ist sticky, wenn es einen Einzelstrangüberhang besitzt. Man unterscheidet hierbei 3'- und 5'-Einzelstrangüberhänge.

Neben Einzelstrangüberhängen können auch *Einzelstrangabschnitte* in DNA-Doppelsträngen vorkommen. Der in Abbildung 4.3 links dargestellte DNA-Doppelstrang setzt sich aus drei DNA-Einzelsträngen zusammen und hat beidseitig sticky-Enden mit 5'-Überhang sowie zusätzlich einen Einzelstrangabschnitt. Der rechts dargestellte DNA-Doppelstrang ist aus zwei über ihre gesamte Länge antiparallel-komplementären DNA-Einzelsträngen aufgebaut und endet beidseitig blunt.

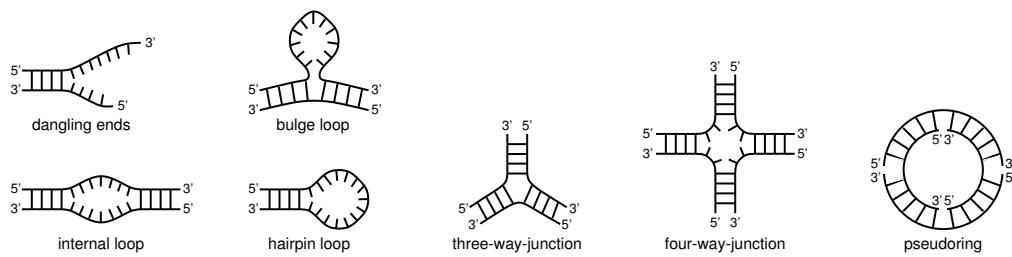
Viele im DNA-Computing vorteilhaft eingesetzte DNA-Doppelstränge zeichnen sich durch eine vergleichsweise einfache Sekundärstruktur aus, die es ermöglicht, diese DNA-Doppelstränge *linearisiert* zu notieren. Hierfür wird vereinbart, dass die in den zweizeiligen Darstellungen mittels passgeformter Bausteine jeweils oben abgebildeten Einzelstränge eines DNA-Doppelstranges in 5'-3'-Richtung und die unteren in 3'-5'-Richtung notiert werden.

#### Definition 4.9 linearer DNA-Strang

Ein DNA-Strang heißt linear, wenn er ausschließlich aus einer Sequenz (linearen Abfolge) von Nucleotidpaaren und/oder Nucleotiden besteht, genau zwei äußere Enden besitzt und an keiner Stelle nichtkomplementär-antiparallel ausgerichtete Nucleotide enthält.

Die in der Abbildung 4.3 dargestellten DNA-Doppelstränge sind linear, ebenso DNA-Einzelstränge, innerhalb derer keine Wasserstoffbrücken existieren. Lineare DNA-Stränge werden im DNA-Computing bevorzugt benutzt und lassen sich durch viele molekularbiologische Prozesse gut verarbeiten.

Nichtlineare DNA-Doppelstränge entstehen häufig im Ergebnis einer nichtoptimal verlaufenden Basenpaarung, die zu so genannten *Basenfehlpaarungen* (Mismatches) führen kann. Bei der Zusammenlagerung mehrerer DNA-Einzelstränge zu einem DNA-Doppelstrang wird hierbei nicht die maximal mögliche Anzahl von Wasserstoffbrücken ausgebildet. Beispiele für nichtlineare DNA-Doppelstränge zeigt Abbildung 4.4.



**Abb. 4.4:** Beispiele für mögliche Sekundärstrukturen nichtlinearer DNA-Doppelstränge

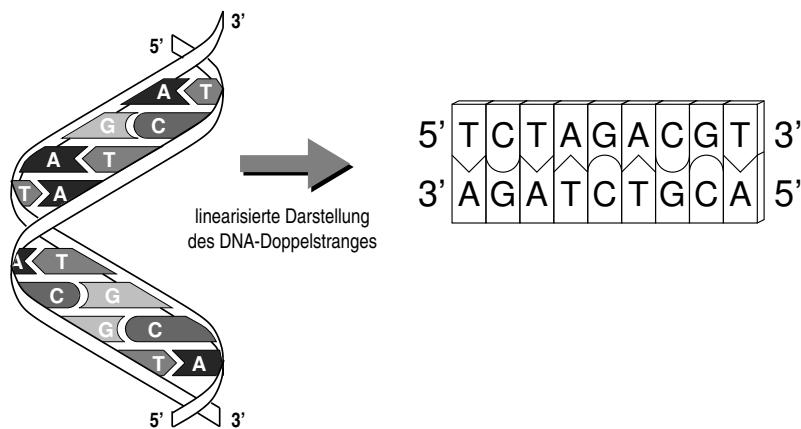
DNA-Einzelstränge, lineare DNA-Doppelstränge wie auch nichtlineare DNA-Doppelstränge werden mit dem Oberbegriff *DNA-Strang* bezeichnet.

#### 4.1.3 DNA-Konformationen und Tertiärstruktur

Aus Zellen von Organismen gewonnene DNA-Doppelstränge bestehen aus genau zwei, über ihre gesamte Länge antiparallel-komplementären DNA-Einzelsträngen, bei denen sich alle Nucleotide zu Nucleotidpaaren verbunden haben, siehe Abbildung 4.5.

Die räumliche Struktur (*Tertiärstruktur*) dieser DNA-Doppelstränge hat die Form einer Doppelhelix: Die beiden antiparallel-komplementären DNA-Einzelstränge (Molekülbänder) sind um eine gemeinsame helikale (zentrale) Achse gewunden, so dass die Gestalt einer Spirale aus gestapelten, fortlaufend gedrehten Sprossen entsteht („verdrillte Sprossenleiter“). Die durch Wasserstoffbrücken zusammengehaltenen Basen gepaarter Nucleotide bilden jeweils eine Sprosse. Alle Atome der Basen einer Sprosse befinden sich in einer gemeinsamen Ebene, so dass die Basen als planar bezeichnet werden. In der Doppelhelix sind die Sprossen aus den planaren Basenpaaren nach innen gerichtet, während die Phosphat-Desoxyribose-Einheiten außen liegen und das Bandrückgrat formen. Die Doppelhelix repräsentiert den Kern des DNA-Modells von Watson/Crick ([WaCr\_53]). Der linke Teil der Abbildung 4.5 zeigt ein Beispiel für eine DNA-Doppelhelix, deren räumliche Struktur vereinfacht dargestellt ist. Diese Struktur entsteht durch die Gesamtwirkung aller Bindungskräfte, wobei außer den kovalenten Bindungen und den Wasserstoffbrücken auch verschiedene Wechselwirkungen auftreten.

Die Doppelhelix beinhaltet eine Redundanz der in ihr gespeicherten Informationen, denn aus der Nucleotidsequenz des einen Einzelstranges kann die Nucleotidsequenz des antiparallel-



**Abb. 4.5:** Beispiel einer DNA-Doppelhelix und vereinfachte (linearisierte) Darstellung des entsprechenden DNA-Doppelstranges durch passgenaue Nucleotidbausteine.

komplementären Einzelstranges rekonstruiert werden. Dieser Umstand wird bei der DNA-Replikation ausgenutzt und ermöglicht eine nahezu fehlerfreie Weitergabe der Erbinformationen.

Die räumliche Anordnung der einzelnen Atome, aus denen ein Molekül zusammengesetzt ist, nennt man *Konformation*. Für die DNA-Doppelhelix sind drei mögliche Konformationen bekannt, die als *A-DNA*, *B-DNA* und *Z-DNA* bezeichnet werden. Abhängig von den physikalischen und chemischen Umgebungsbedingungen nimmt jeder DNA-Doppelstrang vollständig oder abschnittsweise eine dieser Konformationen ein und kann bei Änderung relevanter Umgebungsbedingungen in eine andere Konformation übergehen.

Die am häufigsten vorkommende Konformation von DNA in wässriger Lösung ist die *B-DNA*, auf der das DNA-Modell von Watson/Crick beruht. *B-DNA* stellt eine rechtsgängige Doppelhelix dar, deren mittlerer Durchmesser 2,37nm beträgt. Eine Windung enthält 10 Basenpaare, so dass benachbarte Basenpaare auf der helikalen Achse jeweils um 36° gedreht sind. Die Längenausdehnung (Ganghöhe) einer Windung misst 3,54nm. In der äußeren Gestalt der Doppelhelix wechseln sich so genannte kleine und große Furchen ab.

Die Konformation der *A-DNA* lässt sich beobachten, wenn die DNA dehydratisiert vorliegt, d.h. die relative Feuchtigkeit unter etwa 75% absinkt bzw. das Lösungsmittel eine geringere Polarität aufweist als die DNA. *A-DNA* ähnelt der *B-DNA*, sie ist ebenfalls rechtsgängig, erscheint aber gestauchter (Ganghöhe 2,53nm, 11 Basenpaare pro Windung) und besitzt einen größeren mittleren Helixdurchmesser (2,55nm).

Das Auftreten von *Z-DNA* setzt einen hohen Salzgehalt des Lösungsmittels oder Abschnitte von alternierenden Pyrimidinen und Purinen voraus. Die im Gegensatz zur *A-DNA* und *B-DNA* linksgängigen Windungen führen zu einer Ganghöhe von 4,56nm, einem mittleren Helixdurchmesser von 1,85nm und 12 Basenpaaren pro Windung. In *Z-DNA* lassen sich große und kleine Furchen nicht signifikant unterscheiden, die äußere Form kann durch einen regelmäßigen Zickzackverlauf versinnbildlicht werden.

#### 4.1.4 Eigenschaften von DNA-Strängen

DNA-Stränge verfügen über eine Reihe von Eigenschaften, die ihre Nutzung als Datenträger für ein Computingkonzept favorisieren:

- DNA-Stränge sind im Vergleich zu herkömmlichen Speichermedien sehr klein und durch ihre gewundene Struktur sehr kompakt. DNA erlaubt eine *Speicherdichte* von bis zu  $10^{21}$  Basenpaaren pro Liter, das entspricht etwa 1 bit pro  $\text{nm}^3$  ([Kari\_97]).
- DNA-Stränge sind unter geeigneten Bedingungen beliebig lange konservierbar. Durch die *Langlebigkeit* eignet sich DNA somit auch als persistentes Speichermedium.
- DNA-Stränge können auch außerhalb von Zellen verarbeitet und aufbewahrt werden. Dies ermöglicht ein effizientes *in-vitro-Handling* im Labor.
- DNA-Stränge lassen sich leicht millionenfach duplizieren und auf mehrere Reagenzgläser verteilen. Auf diese Weise ist eine *redundante, dezentrale, verlustsichere Informationsspeicherung* möglich.
- DNA-Stränge können sehr *energieeffizient* verarbeitet werden ( $2 \cdot 10^{19}$  molekulare Operationen pro Joule, [Kari\_97]).
- DNA-Stränge sind *richtungsbehaftet*, was die Kodierung und Dekodierung von Daten in DNA-Sequenzen erleichtert und die Modellierung von DNA-Strängen durch Zeichenketten erlaubt.
- DNA-Stränge sind *elektrisch negativ geladen*, wodurch elektrophoretische Analysemethoden anwendbar sind.
- Die Enden von DNA-Strängen können *chemisch markiert* werden, indem dort geeignete chemische Gruppen oder Moleküle angelagert oder abgebaut werden. Dies erleichtert die Selektion und Separation von DNA-Strängen.
- Einfache Fehler in der DNA-Struktur (wie beispielsweise Strangbrüche) können unter geeigneten Bedingungen in bestimmtem Umfang effizient behoben werden. Hierdurch ist eine begrenzte *Reparaturfähigkeit* von DNA gegeben.
- DNA-Stränge lassen sich unter Verwendung einer Vielzahl spezifisch wirkender Enzyme rekombinieren (verarbeiten), wodurch ein *großes Spektrum an molekularbiologischen Prozessen* existiert.
- Sowohl die *Synthese* als auch die *Sequenzbestimmung* von DNA-Einzelsträngen sind möglich. Es existiert somit ein Instrumentarium zur Kodierung und Visualisierung von DNA-Daten.
- DNA-Stränge sind in hohem Maße *recyclingfähig* und *wiederverwendbar*. Dadurch wird eine *umweltfreundliche* und *verschleißfreie* Informationsverarbeitung und -speicherung erreicht.

## 4.2 Allgemeine Grundsätze zum laborpraktischen Umgang mit DNA

Die weiterführend als *DNA-Operationen* bezeichneten molekularbiologischen Prozesse des DNA-Computing werden laborpraktisch meist in einem Reagenzglas (synonym: Reaktionsgefäß, Test Tube) ausgeführt, wofür sich der Begriff *in vitro* etabliert hat. Im Gegensatz dazu steht der Begriff *in vivo*, der Abläufe in der lebenden Zelle charakterisiert. Ein Reagenzglas kann eine beliebig große, aber endliche Menge von DNA-Strängen enthalten, die dort im Allgemeinen in wässriger Lösung vorliegen. Übliche Reaktionsansätze beinhalten bis zu etwa 20nmol DNA, die in bis zu etwa 2ml Reinstwasser gelöst sind<sup>1</sup>. Jeder DNA-Strang wird durch eine Vielzahl identischer Moleküle repräsentiert, wobei deren Anzahl gewöhnlich in der Größenordnung von mehreren Millionen liegt. Dadurch ist eine hohe Redundanz der Informationsspeicherung gewährleistet. In den nachfolgenden Beispieldarstellungen, die die Wirkungen der DNA-Operationen illustrieren, wird jedoch zumeist nur *ein* Exemplar identischer DNA-Stränge angegeben.

Die Eingangsoperanden einer DNA-Operation liegen als Menge von DNA-Strängen in einem Reagenzglas vor, wobei jede DNA-Operation idealerweise gleichzeitig auf die Gesamtheit aller DNA-Stränge in diesem Reagenzglas wirkt. Im Ergebnis entsteht ein modifizierter Reagenzglasinhalt. DNA-Operationen sind somit *massiv datenparallel*.

DNA-Operationen sind parameterbehaftet. Über die Belegung von Parametern erfolgt die Prozesssteuerung. Typische Parameter sind Ausgangsstoffe und ihre Konzentrationen, Reihenfolge und Zeitpunkte ihrer Zusammenführung, der Temperatur-Zeit-Verlauf des Reaktionsansatzes, pH-Wert sowie von außen einwirkende Kräfte (z.B. Schütteln, Zentrifugieren, Durchmischen und Ruhelage). Neben den Parametern gibt es auch eine Reihe unkontrollierbarer Einflussgrößen, welche die Reproduzierbarkeit von DNA-Operationen beeinträchtigen und die Seiteneffektanfälligkeit bedingen. Durch die Konstruktion des Versuchsaufbaus als möglichst geschlossenes System, eine Parameteroptimierung sowie Kontrollexperimente ist man bestrebt, die Seiteneffektanfälligkeit zu minimieren. Dazu tragen insbesondere eine gründliche Vorabplanung aller Arbeitsschritte einschließlich ihres Zusammenwirkens sowie eine sterile und präzise Arbeitsweise bei, die Verunreinigungen der verwendeten Materialien weitestgehend vermeidet. Alle für die Ausführung von DNA-Operationen notwendigen laborpraktischen Arbeitsschritte werden unter detaillierter Angabe aller beeinflussbaren Parameter in einem Abarbeitungsprotokoll erfasst, um eine möglichst exakte Wiederholung des Ablaufs zu einem späteren Zeitpunkt oder an einem anderen Ort zu gewährleisten.

Eine hinreichend große Menge konzentrierter reiner DNA erscheint dem Betrachter als weißes Pulver. DNA, die als Eingangsoperand für DNA-Operationen dient oder das Ergebnis von Operationsausführungen repräsentiert, lässt sich in wässriger Lösung unter geeigneten Bedingungen über längere Zeiträume *in vitro* aufbewahren. Die DNA wird hierfür durch Zugabe von Reinstwasser gelöst, wobei sich die Menge des eingesetzten Reinstwassers nach der angestrebten Zielkonzentration richtet. Größere DNA-Vorräte konserviert man vorteilhaft als so genannte Stock-Solutions, die gewöhnlich eine DNA-Konzentration in verschiedenen Stufen zwischen

<sup>1</sup>Die Stoffmenge 1mol umfasst etwa  $6,0221367 \cdot 10^{23}$  DNA-Moleküle. Dieser Wert ergibt sich unmittelbar aus der Avogadroschen Konstante  $N_A = 6,0221367 \cdot 10^{23} \text{ mol}^{-1}$ , die als Proportionalitätsfaktor zwischen der Molzahl der Stoffmenge und der Molekülanzahl definiert ist. Eine Stoffmenge von 20nmol enthält folglich ungefähr  $1,2 \cdot 10^{16}$  DNA-Moleküle.

einigen  $\frac{\text{mmol}}{1}$  und etwa  $1 \frac{\mu\text{mol}}{1}$  aufweisen. Die wässrige DNA-Lösung wird in einem gut verschließbaren Reagenzglas, das gewöhnlich aus durchsichtigem Kunststoffmaterial mit Deckel besteht, bereitgestellt und zur längerfristigen Aufbewahrung in einem Gefrierschrank bei einer Temperatur von  $-20^\circ\text{C}$  oder darunter tiefgefroren. Durch diese Temperatur wird verhindert, dass eventuell als Verunreinigung in der wässrigen Lösung befindliche Stoffe (wie beispielsweise Enzyme) unerwünschte Reaktionen auslösen oder fördern können, die die DNA modifizieren. Zusätzlich ist die DNA im Gefrierschrank auch keinen Umwelteinflüssen ausgesetzt, die zu Veränderungen der DNA führen können, wie beispielsweise radioaktive Strahlung. Da die DNA außerhalb lebender Zellen nicht mehr durch entsprechende natürlich ablaufende Reparaturmechanismen geschützt ist, kommt der sorgfältigen Einhaltung DNA-konservierender Bedingungen eine hohe Bedeutung zu.

Bei der Ausführung von DNA-Operationen im molekularbiologischen Labor treten bestimmte elementare Arbeitsschritte häufig und wiederkehrend auf:

**Auftauen/Bereitstellen der wässrigen DNA-Lösung:** Im Vorfeld einer Verarbeitung von DNA in wässriger Lösung wird der tiefgefrorene Reagenzglasinhalt bei Raumtemperatur aufgetaut und das Reagenzglas anschließend in einer mit Eiswürfeln gefüllten Wanne bereitgehalten. Die Eiswanne stellt sicher, dass die Reagenzglasinhalte zwischen den einzelnen Arbeitsschritten bei einer Temperatur knapp über  $0^\circ\text{C}$  in flüssigem Zustand gehalten werden.

**Pipettieren:** Mit Hilfe des Pipettierens lassen sich Reagenzglasinhalte vollständig oder teilweise von einem Reagenzglas in ein anderes Reagenzglas überführen. Moderne Pipetten gestatten es, das aufzunehmende Volumen vor dem Ansaugen genau einzustellen. Nach jedem Pipettierungsvorgang auswechselbare Pipettenspitzen aus Einwegmaterial ermöglichen ein steriles Arbeiten. Die Pipettenspitzen sind so beschaffen, dass das angesaugte Volumen auch möglichst vollständig wieder abgegeben werden kann und kaum Reste in der Pipettenspitze verbleiben. Einwegspitzen werden nach Abschluss jedes Pipettierungsvorganges in einen Entsorgungsbehälter abgeworfen.

**Zentrifugieren:** In Flüssigkeiten gelöste Stoffe können durch Zentrifugieren nach ihrer Dichte und Teilchengröße räumlich getrennt werden. Ebenso lassen sich unlösliche Bestandteile aus Flüssigkeiten separieren. Die entsprechenden Geräte (Zentrifugen) besitzen einen kreisförmigen Rotor, in den eine bestimmte Anzahl von Reagenzgläsern radial mit ihren Böden nach außen gerichtet eingesteckt werden kann. Beim Zentrifugieren dreht sich der Rotor sehr schnell um den Kreismittelpunkt, wodurch infolge der Fliehkräfte die in den Reagenzgläsern enthaltenen Stoffe umso stärker in Richtung des Reagenzglasbodens gedrückt werden, je höher ihre Dichte und Teilchengröße ist. Bei der Bestückung des Rotors ist auf eine genaue Tarierung zu achten, damit keine Unwucht entsteht, die zu Beschädigungen an der Zentrifuge führen kann. Moderne Tischzentrifugen (Ultrazentrifugen) mit Einsätzen für 2ml-fassende Einwegtubes erlauben frei wählbare Rotationsgeschwindigkeiten bis etwa 12 000rpm (Umdrehungen pro Minute, rotations per minute). Das Zentrifugieren dauert bei der DNA-Separation von Fremdstoffen üblicherweise zwischen wenigen Sekunden und einigen Minuten. Wird durch Zentrifugieren eine Flüssigkeit in mehrere Schichten (Phasen) aufgeteilt, so nennt man die am Boden des Reagenzglases befindliche Schicht *Unterstand* und die am entgegengesetzten Ende der Flüssigkeitssäule liegende oberste Schicht *Überstand*. Angesammelte unlösliche Bestandteile

aus Flüssigkeiten heißen *Pellet*. Unterstände, Überstände wie auch Pellets können nach dem Zentrifugieren gesondert abpipettiert werden.

**Vortexen/Schütteln:** Um Reagenzien innerhalb der wässrigen Lösung im Reagenzglas möglichst gleichmäßig zu verteilen und eine gute Durchmischung zu erzielen, vortext bzw. schüttelt man die dabei gut verschlossenen Reagenzgläser. Das Vortexen geschieht mit Hilfe eines Gerätes, das über einen exzentrisch gelagerten Gummistutzen verfügt, der nach innen gewölbt ist und auf den man per Hand ein Reagenzglas gut aufpressen kann. Beim Vortexen wird der Gummistutzen in Rüttelbewegungen versetzt, die sich auf das aufgepresste Reagenzglas übertragen. Die Intensität der Rüttelbewegungen kann am Gerät variiert werden. Das Schütteln von Reagenzgläsern erfolgt in der Regel von Hand. Beide Vorgänge dauern nur wenige Sekunden. Wenn sich nach dem Vortexen bzw. Schütteln separate Flüssigkeitstropfen an den Innenwänden des Reagenzglases gebildet haben, empfiehlt sich ein sehr kurzes Zentrifugieren mit niedriger Rotationsgeschwindigkeit (short spin), um die einzelnen Tropfen am Reagenzglasboden wieder zusammenzuführen.

**Inkubieren:** Viele biochemische Reaktionen erfordern einen vorgegebenen Temperatur-Zeit-Verlauf, damit sie in der gewünschten Weise ablaufen. Insbesondere Enzyme (Biokatalysatoren) besitzen spezifische Temperaturen, bei denen sie optimal wirken. Die Zeitspanne, in der die zusammengeführten Stoffe miteinander reagieren und die entsprechenden chemischen Bindungen umgebaut werden, heißt Inkubation. Gerätetechnisch wird das Inkubieren durch geeignete Heiz- und Kühlaggregate ermöglicht, die Reagenzgläser aufnehmen können und diese einer zuvor eingestellten Temperatur bzw. einem weitgehend frei vorgebbaren Temperatur-Zeit-Verlauf aussetzen. Manche Geräte gestatten zusätzlich auch ein kontinuierliches Bewegen der Reagenzgläser, damit deren Inhalt ständig durchmischt wird. Die Dauer des Inkubierens ist sehr reaktionsspezifisch und liegt meist zwischen wenigen Minuten und einem Tag.

Weiterführend werden die für das DNA-Computing relevanten DNA-Operationen im Einzelnen beschrieben und anhand von Beispielen illustriert. Für viele DNA-Operationen existieren mehrere Begriffe, die aus der historischen Entwicklung der Methoden und ihrer Anwendungen entstanden sind. In den Definitionen der DNA-Operationen werden die synonymen Begriffe aufgezählt, die Operation selbst aber nach der international üblichen Bezeichnung benannt.

### 4.3 Gewinnen von DNA

Die Bereitstellung geeigneter DNA ist Voraussetzung, um DNA-Computing laborpraktisch betreiben zu können. Die zu verarbeitenden Daten müssen in Form von DNA-Strängen kodiert und diese DNA-Stränge anschließend in der erforderlichen Exemplaranzahl gewonnen werden. Prinzipiell gibt es zwei Strategien zur Gewinnung von DNA: Die *DNA-Einzelstrangsynthese* ermöglicht die Erzeugung von DNA-Einzelsträngen mit frei wählbarer Nucleotidsequenz und Exemplaranzahl bis zu einer Länge von etwa 100 Basen. Im Gegensatz dazu kann man alternativ auch DNA aus Trägerorganismen isolieren. Unterschieden wird hierbei, ob die DNA aus den Zellkernen entnommen wird – wie bei der DNA-Gewinnung aus Gewebe bzw. Blut – oder nicht aus dem Zellkern stammt, was zum Beispiel bei Plasmid-DNA der Fall ist. Bei der *DNA-Isolation aus Trägerorganismen* erhält man zumeist doppelsträngige DNA, deren Nucleotidpaarsequenz durch den Träger bestimmt ist.

### 4.3.1 DNA-Einzelstrangsynthese

**Definition 4.10** *Synthesis, DNA-Einzelstrangsynthese, Oligonucleotidsynthese*

Unter Synthesis versteht man das Erzeugen von DNA-Einzelsträngen mit frei wählbarer Nucleotidsequenz und in der Größenordnung vorgebbarer Exemplaranzahl. Die Enden der DNA-Einzelstränge sind mit Hydroxylgruppen markiert.



Abb. 4.6: Beispiel für Synthesis

Die ersten erfolgreichen Bestrebungen, DNA-Einzelstränge künstlich herzustellen, liegen mehr als 30 Jahre zurück. Heute gibt es eine Reihe unterschiedlicher Verfahren, von denen sich die im Jahr 1975 durch eine Forschergruppe um Letsinger veröffentlichte Methode durchsetzen konnte ([LFHL\_75]). Mit ihr lassen sich DNA-Einzelstränge bis zu etwa 100 Nucleotiden in großer Reinheit (weniger als 5% fehlerhafte Stränge) generieren. Die Synthese erfolgt in antisense-Richtung. Das Verfahren arbeitet nach dem Prinzip der wachsenden Kette: Fixiert an einem Glasträger wird jeder DNA-Einzelstrang beginnend am 3'-Ende zyklisch Nucleotid für Nucleotid aufgebaut, siehe Abbildung 4.7.

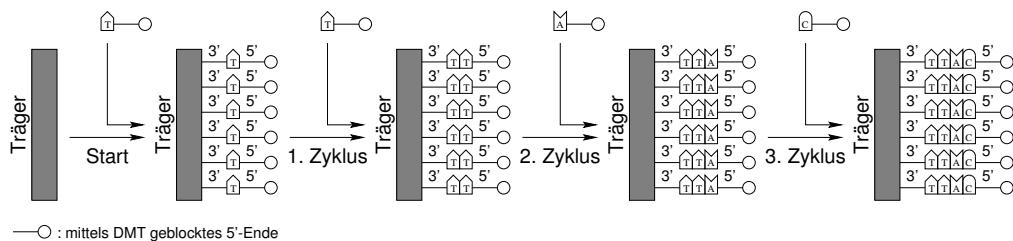


Abb. 4.7: DNA-Einzelstrangsynthese – Beispiel für das Prinzip der wachsenden Kette

Zum Anfügen eines neuen Nucleotids dient ein Zyklus, der aus vier Reaktionsschritten (Detritylation, Coupling, Capping und Oxidation) besteht. Die 5'-Enden der in die wachsenden Ketten integrierten Nucleotide sind chemisch mittels Dimethoxytrityl-Gruppen (DMT) geblockt, um Nucleotid-Mehrfrachanlagerungen an die gleiche Kette innerhalb eines Zyklus zu verhindern. Der erste Schritt, die *Detritylation*, hat die Aufgabe, das 5'-DMT abzuspalten und auszuschwemmen. Im zweiten Schritt, dem *Coupling*, wird jeweils ein neues, 5'-DMT-geblocktes Nucleotid an die 5'-Enden der wachsenden Ketten angefügt. Bei der Ausführung des Coupling kann der unerwünschte Seiteneffekt auftreten, dass sich an manche 5'-Enden kein neues Nucleotid anlagert. Die beim Coupling freibleibenden 5'-Enden werden im dritten Schritt, dem *Capping*, chemisch deaktiviert, damit sich dort in keinem nachfolgenden Zyklus noch neue Nucleotide anlagern können. Die durch das Capping deaktivierten DNA-Einzelstränge sind als

Syntheseprodukt verloren. Der vierte Schritt dient zum Stabilisieren der Bindung zwischen den Nucleotiden. Er wird durch Anlagern eines Sauerstoffatoms an die zwischen den Desoxyribozen befindlichen Phosphatgruppen realisiert und als *Oxidation* bezeichnet. Abbildung 4.8 illustriert die Schritte jedes Zyklus.

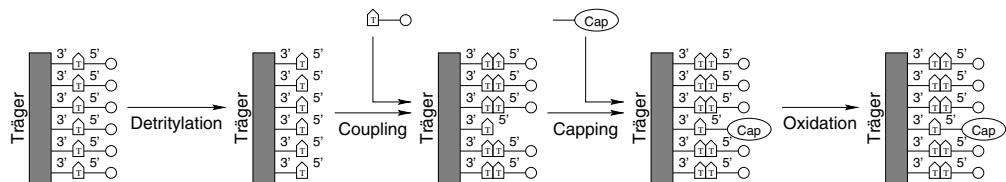


Abb. 4.8: DNA-Einzelstrangsynthese – Beispiel für die Schritte innerhalb jedes Zyklus

Nach der Abarbeitung aller Zyklen erfolgen drei abschließende Arbeitsschritte, die als Deprotection, Cleavage und Purification bezeichnet werden. Die *Deprotection* bewirkt das Entblocken der 5'-Enden, wobei das DMT entfernt und durch Hydroxylgruppen ersetzt wird. Mit Hilfe des *Cleavage* löst man die DNA-Einzelstränge vom Glasträger ab und markiert die 3'-Enden ebenfalls mit Hydroxylgruppen. Die *Purification* sondert die gewünschten DNA-Einzelstränge von Reaktionsresten und DNA-Einzelsträngen fehlerhafter Längen ab. Im Ergebnis liegen DNA-Einzelstränge der gewünschten Nucleotidsequenz und Exemplaranzahl in einem gemeinsamen Reagenzglas in wässriger Lösung vor, siehe Abbildung 4.9.

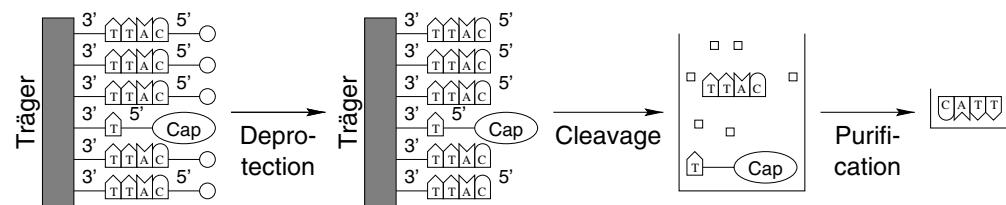


Abb. 4.9: DNA-Einzelstrangsynthese – Beispiel für die abschließenden Schritte nach Zyklenabarbeitung

Es sind Geräte verfügbar, die die DNA-Einzelstrangsynthese vollautomatisch übernehmen können. Die Ausbeute eines Synthesevorganges beträgt etwa 20nmol DNA. Das Hauptanwendungsgebiet liegt in der Bereitstellung von Primern (kurzen DNA-Einzelsträngen einer Länge von etwa 20 Basen) für Polymerase-Kettenreaktionen. Die DNA-Einzelstrangsynthese gilt allgemein als zuverlässig. Im Hinblick auf den Einsatzweck DNA-Computing sind folgende Seiteneffekte relevant:

- Neben den korrekt synthetisierten DNA-Einzelsträngen kann ein geringer Anteil zu kurzer DNA-Einzelstränge (hervorgerufen durch fehlende Nucleotide, so genannte Aussparungen (*Deletions*)) die Purification überdauern und im finalen Reagenzglas vorliegen.
- Ein geringer Anteil der synthetisierten DNA-Einzelstränge kann Abweichungen in der Nucleotidsequenz (bezüglich der Sequenzvorgabe falsch eingegebene Nucleotide, in Anlehnung an den vergleichbaren Effekt bei genomischer DNA als *Punktmutationen* bezeichnet) aufweisen.

- Verunreinigungen durch die verwendeten Chemikalien können in der DNA-Lösung verbleiben.
- Es können zu wenig Strangexemplare erzeugt werden (zu niedrige DNA-Konzentration), so dass nicht genügend DNA-Material für Folgeoperationen zur Verfügung steht.

### 4.3.2 DNA-Isolation aus Organismen

DNA, die aus Organismen gewonnen wird, kann in ihren für das DNA-Computing relevanten Eigenschaften wie Nucleotidpaarsequenz, Stranglänge und Struktur stark variieren. Entsprechend ihrer Herkunft unterscheidet man *genomische DNA*, *Plasmid-DNA* und *virale DNA*.

Genomische DNA stammt aus Zellkernen und wird hauptsächlich aus Gewebe, Blut, Zellkuluren, Pflanzen, Hefen und Bakterien isoliert. Im Ergebnis liegen lineare DNA-Doppelstränge ohne Einzelstrangabschnitte vor, deren Länge sehr groß ist und mehr als 200 000 Basenpaare betragen kann. Aus diesem Grund wird genomische DNA häufig als *hochmolekular* bezeichnet. Die Vorgehensweise bei der DNA-Isolation ist trotz der Vielfalt an Trägerorganismen sehr ähnlich. Im ersten Arbeitsschritt müssen die Zellwände aufgeschlossen (zersetzt) und die in der Zelle enthaltenen Proteine abgebaut werden. Man realisiert diesen Schritt meist durch enzymatisch katalysierte Reaktionen, bei denen auf den Trägerorganismus abgestimmte Enzyme (wie beispielsweise Proteininasen oder Lysozyme) zum Einsatz kommen. Anschließend erfolgt in einem zweiten Schritt die Reinigung der DNA von zusätzlich in der wässrigen Lösung befindlichen Fremdstoffen sowie das Abtrennen eventuell noch an der DNA anhaftender Verunreinigungen wie Proteine, aber auch RNA. Die während der Arbeitsschritte z.B. durch das Pipettieren und Zentrifugieren einwirkenden Scherkräfte können die hochmolekulare DNA unspezifisch in kürzere Fragmente einiger zehntausend Basenpaare zerreißen.

Unter Plasmiden versteht man DNA, die in Mikroorganismen außerhalb der Zellkerne vorkommt und zumeist in Form von doppelsträngigen DNA-Ringen auftritt. Diese spezielle nicht-lineare Struktur wird auch als *zirkuläre DNA* bezeichnet. Ein Plasmid enthält im Allgemeinen zwischen 2 000 und mehreren zehntausend Basenpaaren. Aufgrund dieser im Vergleich zu genomischer DNA deutlich geringeren Länge nennt man Plasmid-DNA *niedermolekular*. Da lineare Abschnitte aus Plasmiden laborpraktisch leicht herausgetrennt sowie durch andere lineare DNA-Abschnitte ersetzt werden können und sich darüber hinaus Plasmide in bestimmten Bakterien gut vermehren lassen, haben Plasmide eine große Bedeutung bei der Klonierung. Die die Plasmide tragenden Bakterien werden in Bakterienkulturen aufgezogen und bis zur gewünschten Größe vermehrt. Zur Isolierung der Plasmide werden die Bakterien zunächst durch bestimmte anorganische Basen (z.B. Natriumhydroxid) aufgebrochen (alkalisch lysiert), wobei gleichzeitig auch die enthaltene doppelsträngige DNA einschließlich der Plasmid-DNA durch Abbau der Wasserstoffbrücken in ihre einzelsträngigen Bestandteile zerfällt. Die anschließende Zugabe von neutralisierendem Kaliumacetat bewirkt, dass sich die niedermolekulare DNA wieder zu Doppelstrang-Ringen zusammenlagert und sich in der Flüssigkeit löst, während die hochmolekulare DNA gemeinsam mit den weiteren verunreinigenden Bakterienbestandteilen ungelöst bleibt. Durch Zentrifugieren können die ungelösten Bestandteile abgesondert werden, so dass eine gereinigte wässrige Lösung mit den Plasmiden vorliegt.

Virale DNA gewinnt man zumeist aus Bakteriophagen, die entweder in Flüssigkultur oder auf Agarplatten in Bakterienkulturen vermehrt werden. Die Länge der linearen DNA-

Doppelstränge beträgt etwa 50 000 Basenpaare. Die Isolierung der viralen DNA erfordert mehrere Stufen. In der ersten Stufe erfolgt die Trennung der Bakteriophagen von den Bakterien, indem die Bakterien aufgebrochen (lysiert) werden und die bakterielle DNA enzymatisch abgebaut wird. Die darauffolgende Stufe separiert die Bakteriophagen durch Zentrifugieren und Aufnehmen des dabei entstehenden Pellets. In der dritten Stufe werden die Bakteriophagen aufgeschlossen (enzymatisch lysiert) und die wässrige Lösung in weiteren Stufen schrittweise gereinigt sowie verbliebene an der DNA haftende Verunreinigungen entfernt.

In Bezug auf die Weiterverwendung der aus Trägerorganismen isolierten DNA für das DNA-Computing sind folgende Seiteneffekte relevant:

- Verunreinigungen (insbesondere durch Fremdstoffe, verwendete Chemikalien, Restbestandteile aus den Trägerorganismen und eingesetzte Enzyme) können in der wässrigen DNA-Lösung verbleiben und Folgeoperationen beeinflussen.
- Die gewonnene DNA kann Unspezifitäten (Inhomogenitäten) aufweisen. Dies bedeutet, dass die in ein und demselben Reagenzglas vorliegenden DNA-Stränge unterschiedliche Struktur, Länge und Nucleotidpaarsequenzen besitzen können, die durch Fehler bei der DNA-Replikation in den Organismen im Vorfeld der DNA-Isolation noch zusätzlich differieren können. Fehler bei der DNA-Replikation führen zu *Punktmutationen*, *Deletions* und *Insertions*. Bei Punktmutationen ist an bestimmten Stellen im Strang ein Nucleotid oder Nucleotidpaar durch ein anderes ersetzt, das von dem Nucleotid oder Nucleotidpaar an der gleichen Position in anderen Exemplaren dieses Stranges abweicht. Eine Deletion ist eine Auslassung (Fehlen) eines Nucleotid- oder Nucleotidpaarsequenz-Abschnittes. Bei einer Insertion hingegen liegt ein zusätzlich eingefügter Abschnitt vor.

Basenfehlpaarungen sind bei doppelsträngiger DNA aus Organismen sehr selten.

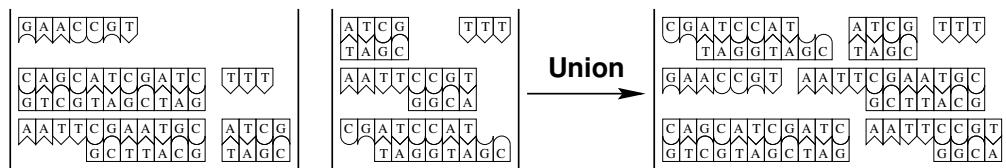
## 4.4 Mischen und Aufteilen von DNA in wässriger Lösung

Das Zusammenstellen von DNA-Pools in wässriger Lösung gehört zu den einfachsten, seiteneffektärmsten und laborpraktisch am schnellsten ausführbaren DNA-Operationen, weil keine chemischen Bindungen in den DNA-Strängen verändert werden und auch keine Separation der DNA-Stränge nach spezifischen Kriterien erfolgt. DNA-Operationen in dieser Klasse sind das *Mischen* von DNA-Pools in wässriger Lösung (Vereinigung, Union) sowie das *Aliquotieren* (Aufteilen, Split).

### 4.4.1 Vereinigung

#### **Definition 4.11** *Union, Merging, Mixing, Mischen, Vereinigung*

Unter Union versteht man das Zusammenführen mehrerer Reagenzglasinhalte in ein gemeinsames Reagenzglas.

**Abb. 4.10:** Beispiel für Union

Üblicherweise werden zur Ausführung der Operation die zu vereinigenden Reagenzglasinhalte nacheinander mittels einer Pipette aufgesaugt, in das Ergebnisreagenzglas überführt und anschließend der Ergebnisreagenzglasinhalt gründlich durchmischt, um einer DNA-Clusterung vorzubeugen.

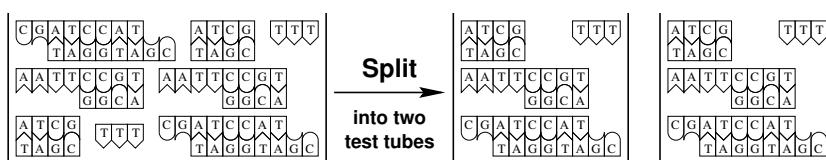
Folgende Seiteneffekte sind bei Operationsanwendung im DNA-Computing signifikant:

- DNA-Reste können in den Eingangsreagenzgläsern zurückbleiben und stehen im Ergebnisreagenzglas nicht zur Verfügung.
- DNA-Clusterungen im Ergebnisreagenzglas können dazu führen, dass nachfolgende Reaktionen unvollständig und ineffizient ablaufen, weil sich die Reaktionspartner nicht in räumlicher Nähe zueinander befinden.
- Verunreinigungen eines einzelnen Eingangsreagenzglases können das gesamte Ergebnisreagenzglas kontaminieren.

#### 4.4.2 Aliquotierung

**Definition 4.12** Split, Aufteilen, Aliquotierung

Unter Split versteht man das Aufteilen eines Reagenzglasinhaltes in mehrere Reagenzgläser.

**Abb. 4.11:** Beispiel für Split. Im Eingangsreagenzglas liegen von jedem enthaltenen DNA-Strang genau zwei Exemplare vor. Der Reagenzglasinhalt wird hier in idealer Gleichverteilung der DNA-Strangexemplare auf zwei Reagenzgläser aufgesplittet.

Zur Ausführung dieser DNA-Operation wird zunächst die wässrige DNA-Lösung im Reagenzglas durchmischt (durch mehrmaliges vorsichtiges Ansaugen und Zurückgeben des Reagenzglasinhaltes mit einer Pipette), danach durch Pipettieren der gewünschte Volumenanteil wässrige DNA-Lösung entnommen und in einem anderen, leeren Reagenzglas wieder abgelegt. Das

Durchmischen zu Beginn der Operationsausführung hat die Herstellung einer Gleichverteilung der DNA im Volumen der wässrigen Lösung zum Ziel und beugt DNA-Clusterungen vor.

Das Aliquotieren realisiert eine annähernde Gleichverteilung der DNA-Strangexemplare in die einzelnen Reagenzgläser entsprechend der überführten Volumina. Diese Volumina lassen sich auch so bemessen, dass der ursprüngliche Reagenzglasinhalt vollständig und gleichmäßig auf 2, 3 oder beliebig viele Reagenzgläser verteilt wird.

Folgende Seiteneffekte können beim Aliquotieren auftreten:

- Es ist ein Verlust von DNA-Strängen möglich, die in den Pipettenspitzen verbleiben und nicht in das jeweilige Zielreagenzglas gelangen.
- Abweichungen der Gleichverteilung bzw. eine nicht proportionale Aufteilung der Strangexemplare entsprechend der Volumenanteile in die einzelnen Reagenzgläser können vorkommen.
- Verunreinigungen des ursprünglichen Reagenzglasinhaltes können sich auf alle einbezogenen Reagenzgläser übertragen.

Sowohl Union als auch Split sind laborpraktisch physikalische Prozesse und keine molekularbiologischen Reaktionen, denn es werden in beiden Fällen keine chemischen Bindungen umgebaut.

## 4.5 Knüpfen und Aufbrechen von Wasserstoffbrückenbindungen

DNA kann in Form von DNA-Einzelsträngen und DNA-Doppelsträngen vorliegen. Beide Formen sind mittels temperaturgesteuerter Prozesse ineinander überführbar. Als maßgeblicher Prozessparameter dient folglich der Temperatur-Zeit-Verlauf, dem das Reagenzglas mit der DNA-Probe ausgesetzt wird. DNA besitzt eine spezifische Schmelztemperatur, unterhalb derer sich die thermisch instabilen Wasserstoffbrücken in der beschriebenen Weise ausbilden und oberhalb derer sie aufbrechen. Entsprechend formieren sich DNA-Doppelstränge aus den zugrunde liegenden DNA-Einzelsträngen, oder die DNA-Doppelstränge zerfallen in ihre Einzelstrangbestandteile. Zur Berechnung der Schmelztemperatur existieren mehrere, zumeist empirisch gewonnene näherungsweise Berechnungsvorschriften auf Basis messbarer bzw. bekannter Eigenschaften der DNA-Probe. Die Schmelztemperatur ist definiert als die Temperatur, bei der 50% der DNA-Probe einzelnsträngig vorliegt. In [LoZo\_98] ist die zugeschnittene Größengleichung 4.1 zur näherungsweisen Ermittlung der Schmelztemperatur  $T_m$  in °C eines linearen DNA-Stranges angegeben:

$$T_{m,\text{°C}} = \begin{cases} 81,5 + 16,6 \cdot \ln(K_{+,M}) + 0,41 \cdot \%GC - \frac{500}{L} & \text{für } L > 20 \\ 2 \cdot \#AT + 4 \cdot \#CG & \text{für } L \leq 20 \end{cases} \quad (4.1)$$

$L$  verkörpert die Stranglänge in Basen bei einzelsträngiger, in Basenpaaren bei doppelsträngiger DNA. Die obere Formel wird für DNA-Stränge einer Länge von mehr als 20 Basen bzw. Basenpaaren verwendet. Dabei bedeutet  $\%GC$  den prozentualen Guanin-Cytosin-Anteil in Prozent und  $K_+$  die Kationenkonzentration in  $M = \frac{\text{mol}}{\text{l}}$ . Kationen („Salz“) können der DNA-Lösung beigegeben werden, um die Schmelztemperatur strangstabilisierend zu erhöhen, so dass sie sich nicht seiteneffektbegünstigend auf Folgeoperationen auswirkt. Für sehr kurze DNA-Stränge liefert die untere Formel genauere Werte. Der Parameter  $\#AT$  bezeichnet die Gesamtanzahl von Nucleotiden A und T,  $\#CG$  die Gesamtanzahl von Nucleotiden C und G im DNA-Strang. Nichtlineare doppelsträngige DNA besitzt gewöhnlich eine niedrigere Schmelztemperatur als lineare doppelsträngige DNA mit gleichem Guanin-Cytosin-Anteil und bei gleicher Kationenkonzentration.

#### 4.5.1 Hybridisierung

**Definition 4.13 Annealing, Hybridisierung, Erstarren, Reassoziation, Renaturierung**

Unter Annealing versteht man das Zusammenlagern von mindestens zwei Molekülen einzelsträngiger DNA (oder eines DNA-Einzelstranges mit sich selbst) an ihren antiparallel-komplementären Stellen zu DNA-Doppelsträngen unter Bildung aller Anlagerungsmöglichkeiten. Beim Zusammenlagern werden zwischen jeweils zwei antiparallel-komplementär gegenüberliegenden Nucleotiden A und T je zwei, zwischen C und G je drei Wasserstoffbrücken ausgebildet, die eine entsprechende Basenpaarung bewirken. Zwei DNA-Stränge bleiben miteinander durch Wasserstoffbrücken verbunden, wenn sie über mindestens 50% der Länge mindestens eines beteiligten DNA-Stranges antiparallel komplementär sind.

Die Abbildung 4.12 veranschaulicht die Wirkung dieser Operation auf eine gegebene Menge von DNA-Strängen. Mehrere im Eingangsreagenzglas befindliche Exemplare eines DNA-Stranges werden durch Angabe eines Repräsentanten dargestellt. Alle DNA-Stränge des Eingangsreagenzglases sowie im Verlauf der Operationsabarbeitung entstehende Zwischenprodukte können miteinander interagieren.

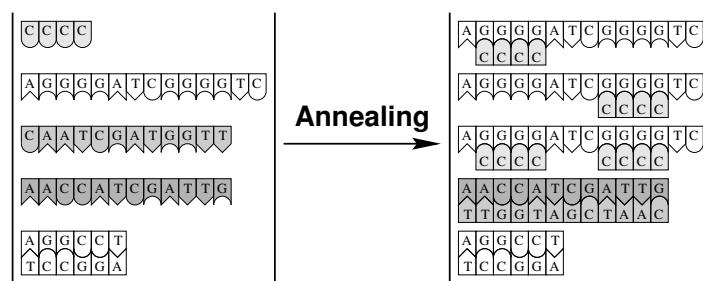


Abb. 4.12: Beispiel für Annealing

Annealing wird laborpraktisch gewöhnlich als ein langsam verlaufender Abkühlungsprozess realisiert. Hierzu erhitzt man die DNA-Probe auf eine Temperatur, die mit Sicherheit ober-

halb der Schmelztemperatur liegt (Richtwert: +94°C), um bereits vorhandene Wasserstoffbrücken abzubauen und hinreichend viel kinetische Energie zuzuführen, damit sich die DNA-Moleküle neu formieren können. Anschließend wird die DNA-Probe über mehrere Stunden auf Raumtemperatur abgekühlt. Je langsamer dieser Vorgang erfolgt, desto mehr Zeit bleibt zur Ausbildung einer optimalen Basenpaarung, bei der alle möglichen Wasserstoffbrücken auch tatsächlich entstehen.

Als Operation im DNA-Computing ist das Annealing mit mehreren signifikanten Seiteneffekten behaftet:

- Wenn die Exemplaranzahlen der zu annealenden DNA-Stränge nicht exakt aufeinander abgestimmt sind oder der Annealingprozess unvollständig abläuft, können Reste unbundener DNA zurückbleiben.
- Die Exemplaranzahlen der im Ergebnis der Hybridisierung entstandenen DNA-Stränge können stark voneinander abweichen.
- Die weitaus meisten Modelle des DNA-Computing setzen lineare DNA voraus, oder es werden speziell konstruierte nichtlineare DNA-Strukturen zugrunde gelegt. Unter dieser Annahme zählt auch als Seiteneffekt, dass sich einzelsträngige DNA nicht an *allen* antiparallel-komplementären Stellen verbinden kann oder Basenfehlpaarungen auftreten, so dass nichtlineare DNA entsteht, siehe Abbildung 4.4 auf Seite 105 und Kombinationen der dort dargestellten Artefakte.
- Die durch Annealing erzeugten DNA-Doppelstränge sind nur unterhalb ihrer Schmelztemperatur und in nicht stark alkalischem pH-Milieu stabil. Werden diese Randbedingungen – auch versehentlich – nicht eingehalten, können die Wasserstoffbrückenbindungen wieder aufbrechen.

Annealing ist eine sequenzspezifisch wirkende DNA-Operation, die auf Interaktionen zwischen verschiedenen DNA-Strängen basiert und deshalb ein hohes Strangkombinationspotenzial besitzt. DNA-Doppelstränge ohne Einzelstrangabschnitte, die beidseitig blunt enden, werden durch diese Operation nicht verändert.

#### 4.5.2 Denaturierung

##### **Definition 4.14** *Melting, Denaturierung, Schmelzen, Dissoziation*

Unter Melting versteht man das Aufspalten von DNA-Doppelsträngen in die zugrunde liegenden DNA-Einzelstränge.

Melting, häufig auch als Denaturieren bezeichnet, ist die Umkehrung des molekularbiologischen Ablaufs beim Annealing. Zum Denaturieren wird die DNA üblicherweise bis über ihre Schmelztemperatur erhitzt (Richtwert: +94°C). Dabei brechen die thermisch instabilen Wasserstoffbrückenbindungen auf. Die so gewonnenen DNA-Einzelstränge liegen jedoch nur oberhalb ihrer Schmelztemperatur vor. Diese kann bei DNA-Strängen, die für das DNA-Computing

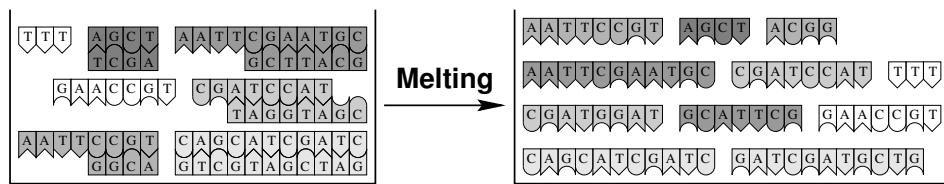


Abb. 4.13: Beispiel für Melting

typisch sind, durchaus mehr als +70°C betragen. Soll eine solch hohe Temperatur nach dem Melting nicht aufrechterhalten werden, besteht alternativ die Möglichkeit, DNA-Doppelstränge zu denaturieren, indem man sie stark alkalischen Bedingungen (hoher pH-Wert) aussetzt.

Auch das Denaturieren ist bezüglich einer Anwendung im DNA-Computing seiteneffektbehaftet:

- Die Reaktion kann unvollständig ablaufen, so dass nur ein Teil der Wasserstoffbrückenbindungen aufgebrochen wird. Im Ergebnis kann unerwünschte nichtlineare DNA entstehen oder die Operationsausführung überdauern. Bei zu niedrig gewählter Temperatur oder zu kurzer Wirkdauer können auch DNA-Doppelstränge im Ergebnisreagenzglas zurückbleiben.
- Bereits bei geringfügiger Abkühlung besteht die Gefahr einer Rehybridisierung.

Das Denaturieren wirkt nicht sequenzspezifisch. DNA-Einzelstränge werden durch diese Operation nicht verändert. Die dem Melting nachfolgende Operation sollte im unmittelbaren Anschluss ausgeführt werden, weil die einzelsträngige DNA bei einer Zwischenlagerung oberhalb ihrer Schmelztemperatur geschädigt werden kann.

## 4.6 Enzymatische Reaktionen

Enzyme gelten als „Katalysatoren des Lebens“. Sie besitzen eine große, wenn nicht gar die zentrale Bedeutung in der Molekularbiologie. Chemische Umwandlungen in lebenden Zellen werden durch Enzyme gefördert und beschleunigt. Dadurch haben sie die Fähigkeit, Lebensvorgänge aufrechtzuerhalten und zu steuern. Lebensvorgänge sind ein kompliziertes Zusammenspiel vieler, größtenteils voneinander abhängiger molekularbiologischer Reaktionen, die dem Stoffwechsel, der Energiegewinnung, der Reproduktion, dem Wachstum und dem Schutz des Lebewesens dienen. Enzyme sind in der Lage, alle Verbindungen zu bilden, die man in lebendem Gewebe findet.

Die Entdeckung der Enzyme ist eng mit der Untersuchung der alkoholischen Gärung verbunden, bei der Glucose in Ethanol und Kohlendioxid umgewandelt wird. Man erkannte im 19. Jahrhundert, dass für die alkoholische Gärung „lösliche Fermente“ notwendig sind, die sich in lebenden Hefezellen befinden. Der deutsche Biochemiker Wilhelm Friedrich Kühne nannte die

löslichen Fermente Enzyme. Dieser Name ist aus dem griechischen Wort für Sauerteig abgeleitet. Nachdem Enzyme als lösliche subzelluläre Substanz erkannt waren, dauerte es jedoch bis in die 60er Jahre des 20. Jahrhunderts, ehe die chemische Struktur und die Wirkungsweise von Enzymen vollständig aufgeklärt und nachgewiesen war.

Die weitaus meisten Enzyme sind Proteine. Als Ausnahme kennt man bisher Ribozyme, bei denen bestimmte RNA-Stränge selbst die Rolle eines Enzyms übernehmen und auf andere RNA-Stränge wirken können. Nachfolgend sollen ausschließlich Proteinenzyme betrachtet werden. Proteine sind Polypeptidketten. Sie bestehen aus einer Kette von Aminosäuren, deren Glieder über Amino- und Carboxylgruppen verbunden sind. Man unterscheidet 20 Aminosäuren, die als monomere Bausteine zur Verfügung stehen. Jede dieser Aminosäuren besitzt durch ihre Seitenketten spezifische chemische Eigenschaften. Jedes Protein ist bestimmt durch seine Aminosäuresequenz und durch die dreidimensionale Faltung der Kette, die hervorgerufen wird durch mannigfaltige Bindungen der Aminosäure-Seitenketten und des Kettenrückgrats untereinander. Es treten Wasserstoffbrückenbindungen, Ionenbindungen, kovalente Bindungen sowie verschiedene Wechselwirkungen auf, die die Kette zu einem Knäuel formen. Die dreidimensionale Faltung verleiht dem Protein eine charakteristische dreidimensionale Oberflächenstruktur aus einem Muster an Eindellungen, Rillen, Taschen, Höhlungen usw. Ein Protein kann mehr als 1000 Aminosäuren enthalten. Aus der Struktur eines Proteins ergibt sich seine Wirkung. Man unterscheidet eine Vielzahl von Proteinen bzw. Enzymen.

Enzyme verkörpern Biokatalysatoren. Das heißt, sie fördern bestimmte molekularbiologische Reaktionen. Während Hitze und die Zugabe von Säuren oder Basen die Reaktionsgeschwindigkeit bestenfalls um das Zehn- bis Hundertfache erhöhen, beschleunigt ein Enzym die Reaktion normalerweise um das Milliardenfache und mehr. Nur durch diese enorme Reaktionsbeschleunigung sind Lebensvorgänge in der Zelle überhaupt möglich. Die Reaktionen würden sonst so langsam ablaufen, dass sie keine Bedeutung für die Zelle haben und ihr Überleben nicht sichern. Jedes Enzym katalysiert eine bestimmte molekularbiologische Reaktion, geht jedoch selbst unverändert aus der Reaktion hervor.

Es erhebt sich die Frage, auf welche Weise ein Enzym spezifisch katalysierend wirken kann. Hierzu muss es das zu verändernde Molekül (Substratmolekül, hier DNA) und seine umzubauende chemische Bindung lokalisieren können. Enzyme bewerkstelligen dies aufgrund des Prinzips von „Schlüssel und Schloss“. Ein Teil der Enzymoberfläche ist komplementär zum Substratmolekül geformt. Dadurch kann das Enzym an das Substratmolekül andocken und es eng umschließen. Es bildet sich ein *Enzym-Substrat-Komplex*. Die molekularbiologische Reaktion besteht darin, innerhalb des Substratmoleküls ein Atom oder eine Atomgruppe hinzuzufügen oder zu entfernen. Hierzu ist es notwendig, mindestens eine chemische Bindung des Substratmoleküls aufzubrechen und schrittweise umzuordnen. Dies erfordert eine bestimmte Aktivierungsenergie.

Die Aktivierungsenergie ist die Energie, die Moleküle mindestens haben müssen, um in der gewünschten Weise miteinander reagieren zu können. Obwohl bei den auch unkatalysiert möglichen molekularbiologischen Reaktionen die Energie des Endzustandes niedriger ist als die des Ausgangszustands, muss zum Anstoß der Reaktion eine Energiebarriere überwunden werden. Diese erklärt sich daraus, dass auf dem Weg zum Reaktionsprodukt bestimmte energiereiche Übergangszustände durchlaufen werden müssen. Nur diejenigen Moleküle, die hinreichend viel Energie zum Einnehmen dieser Übergangszustände aufweisen, sind somit reaktionsfähig. Ohne Benutzung enzymatischer Katalysatoren kann man beispielsweise durch

Temperaturerhöhung erreichen, dass sich der Anteil der Moleküle mit Aktivierungsenergie vergrößert und die Reaktion schneller abläuft. Die Temperaturerhöhung eignet sich jedoch nicht zur Beschleunigung molekularbiologischer Reaktionen in der lebenden Zelle, da die Hitze von außen (fremdgesteuert) einwirken muss und die Gefahr der Zerstörung der Zelle bei zu hohen Temperaturen besteht. Eine andere Beschleunigungsmöglichkeit ist die Bereitstellung freier Protonen ( $H^+$ ) oder protonenaufnehmender Ionen ( $OH^-$ ) durch Zugabe von Säure oder Base (pH-Wert-Änderung), wodurch das Energieniveau des Übergangszustandes und damit die Aktivierungsenergie gesenkt wird. Enzyme bedienen sich dieses Mechanismus.

Nachdem sich der Enzym-Substrat-Komplex gebildet hat, befinden sich auch bestimmte Aminosäuren des Enzyms in unmittelbarer Nähe der umzubauenden chemischen Bindung des Substrats. Diese werden als *aktives Zentrum* des Enzyms bezeichnet, an dem die *Katalyse* stattfindet. Die Seitenketten der Aminosäuren im aktiven Zentrum sind derart beschaffen, dass sie im schnellen Wechsel sowohl als Säure als auch als Base agieren können. Anders ausgedrückt bedeutet dies, dass das aktive Zentrum räumlich komplementär zum Übergangszustand ist und die Aktivierungsenergie herabgesetzt. Darüber hinaus führt das Enzym infolge des Schlüssel-Schloss-Prinzips die Reaktionspartner zielgerichtet zusammen, so dass die Reaktion nicht auf zufällige Zusammenstöße der Reaktionspartner angewiesen ist. Die selektive katalytische Wirkung von Enzymen beruht folglich auf Anpassung der Enzymstruktur an die Beschaffenheit des Substratmoleküls und der umzubauenden chemischen Bindung sowie auf der Eigenschaft bestimmter Aminosäuren, als Säure- und Basenkatalysator zu dienen.

Durch die Ausführung der eigentlichen molekularbiologischen Reaktion (Inkubation) verändert sich der Enzym-Substrat-Komplex zum *Enzym-Produkt-Komplex*. Das Enzym löst sich schließlich wieder vom Substrat ab (Dissoziation) und steht für die Katalyse einer gleichartigen molekularbiologischen Reaktion an einem anderen Substratmolekül zur Verfügung.

Enzyme können nur dann optimal arbeiten, wenn ihre Oberfläche die zum Substrat passende Struktur hat. Die Oberflächenstruktur von Enzymen ist angepasst an bestimmte Umgebungsbedingungen und insbesondere sehr empfindlich gegenüber Temperatur- und pH-Wert-Abweichungen. pH-Veränderungen in der Umgebung eines Enzyms können dazu führen, dass zusätzliche Protonen gebunden oder vorhandene Protonen abgegeben werden, wodurch die Faltung des Proteins und damit die Oberflächenstruktur verändert wird. Temperaturabweichungen können eine Entfaltung der Polypeptidkette hervorrufen, was ebenfalls mit einer Änderung der Oberflächenstruktur einhergeht. In der Zelle, dem ursprünglichen Wirkungsort von Enzymen, sind Temperatur und pH-Wert im Allgemeinen nahezu konstant. Bei Verwendung von Enzymen im Reagenzglas müssen diese Bedingungen geschaffen werden. Jedem Enzym ordnet man deshalb eine spezifische Temperatur und einen pH-Wert zu, bei dem es seine maximale Aktivität entfaltet. Um im Reagenzglas einen bestimmten pH-Wert aufrechtzuerhalten, bedient man sich einer *Puffer-Substanz*, die hinzugegeben wird. Eine bestimmte Temperatur des Reaktionsraumes kann man im Labor durch Einsatz geeigneter Heiz- bzw. Kühlgeräte erreichen.

Bei hohen Temperaturen werden Enzyme zerstört und verlieren ihre Wirkung (thermische Instabilität). Diesen Umstand kann man ausnutzen, um Enzyme im Reagenzglas nach Beendigung der molekularbiologischen Reaktion zu *deaktivieren*. Dadurch soll sichergestellt werden, dass das Enzym nachgeschaltete Reaktionen und zusätzliche Substrate nicht mehr beeinflusst.

Die einer enzymatischen Reaktion unterzogenen DNA-Doppelstränge sollten derart beschaffen sein, dass sie während der thermischen Enzymdeaktivierung nicht denaturieren, denn bei kurzen DNA-Doppelsträngen ist ein Denaturieren möglich. Alle Reagenzien sollten im Re-

genzglas gut durchmischt (gleichmäßig verteilt) vorliegen, um sicherzustellen, dass sich alle Reaktionspartner in räumlicher Nähe zueinander befinden.

Für das DNA-Computing sind besonders die Enzyme interessant, die auf DNA-Stränge als Substrate einwirken und diese zielgerichtet verändern können. Beispiele hierfür sind Ligasen, Restriktionsendonukleasen, Polymerasen, Kinasen und Phosphatasen.

#### 4.6.1 Ligation

Die Ligation im Sinne des DNA-Computing ist eine molekularbiologische Reaktion, die durch eine DNA-Ligase katalysiert wird und das fortgesetzte Verketten von DNA-Doppelsträngen unter bestimmten Voraussetzungen bewirkt. Jeweils zwei an ihren Enden *kompatible* DNA-Doppelstränge mit *5'-Phosphat* werden durch Herstellung einer Phosphodiesterbindung kovalent miteinander verbunden und bilden damit einen neuen DNA-Doppelstrang. In der lebenden Zelle hat die Ligation Bedeutung bei der DNA-Replikation und bei der Reparatur von Brüchen innerhalb der Nucleotidkette.

**Definition 4.15 Kompatibilität von DNA-Doppelstrangenden**

Zwei sticky-Enden von DNA-Doppelsträngen sind kompatibel, wenn die Einzelstrangüberhänge beider Stränge über ihre gesamte Länge antiparallel-komplementär sind. Blunt-Enden von DNA-Doppelsträngen sind generell kompatibel.

**Definition 4.16 5'-Phosphat**

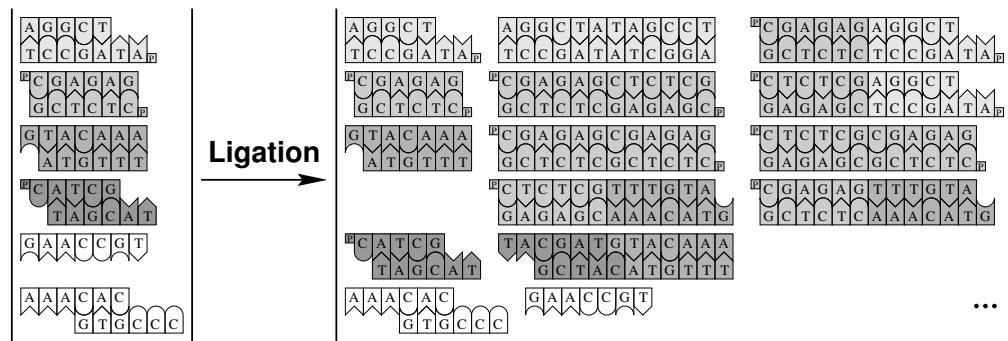
Die Strangendenmarkierung 5'-Phosphat bezeichnet Phosphatgruppen, die an freie 5'-DNA-Strangenden angelagert sind.

**Definition 4.17 Ligation**

Unter Ligation versteht man eine molekularbiologische Reaktion, bei der DNA-Doppelstränge mit kompatiblen Enden und mindestens einem 5'-Phosphat an diesen kompatiblen Enden fortgesetzt miteinander verkettet werden. Es werden alle möglichen Endenpaare, auch von Zwischenprodukten, in die Reaktion einbezogen. Die Reaktion wird durch eine DNA-Ligase katalysiert.

Die Abbildung 4.14 zeigt ein Beispiel für Strangverkettungen, die mittels Ligation gebildet werden können. Dargestellt sind DNA-Doppelstränge einer Länge von bis zu 15 Basenpaaren, ebenfalls rekombinierbare Stränge größerer Längen sind durch drei Punkte angedeutet.

Im biochemischen Sinne sind DNA-Ligasen Enzyme, die die Ausbildung von Phosphodiesterbindungen zwischen benachbarten 3'-Hydroxylgruppen und 5'-Phosphatgruppen an DNA-Doppelstrangenden katalysieren. DNA-Ligase ermöglicht auch das Reparieren von Brüchen



**Abb. 4.14:** Beispiel für Ligation

im Desoxyribose-Phosphat-Rückgrat doppelsträngiger DNA. Der Katalysemechanismus ist bekannt ([DrPo\_92]). Die Knüpfung von Phosphodiesterbindungen mittels Katalyse durch eine DNA-Ligase geschieht unter Nutzung des Cosubstrats *Adenosintriphosphat* (ATP) oder *Nicotinamid-Adenin-Dinucleotid* (NAD) als Energielieferant.

Gut untersuchte DNA-Ligasen entstammen dem Bakterium *Escherichia coli* und dem Bakteriophagen T4 ([WJLF\_68]). Aus dieser Herkunft werden auch die meisten kommerziell angebotenen und laborpraktisch verwendeten DNA-Ligasen gewonnen. In *Escherichia coli* und T4 liegt die DNA linear doppelsträngig vor, und die Enzymoberfläche ist an die Form von DNA-Doppelsträngen angepasst. Deshalb wird durch die genannten Enzyme die Ligation doppelsträngiger DNA, nicht jedoch von DNA-Einzelsträngen, gefördert.

T4-DNA-Ligase entfaltet ihre maximale Aktivität in einem pH-Bereich von etwa 7,5 bis 8,0. Zur Gewährleistung dieses pH-Milieus dient eine Puffersubstanz (T4-DNA-Ligase-Puffer), die dem Reaktionsansatz in der vom Hersteller spezifizierten Endkonzentration beigegeben wird. Eine in-vitro-Ligation führt man gewöhnlich bei einer Temperatur zwischen 16°C und 20°C durch, damit die kompatiblen sticky-Enden vor dem Verketten hybridisieren können. Diese Wirktemperatur stellt eine Balance zwischen der Schmelztemperatur der zu annealenden Einzelstrangüberhänge und einer hinreichenden Enzymaktivität dar. Die Inkubationszeit beträgt bis zu 16 Stunden. Die thermische Deaktivierung von T4-DNA-Ligase erfolgt bei 65°C für etwa 10 Minuten.

Die Ligation von DNA-Doppelsträngen mit sticky-Enden läuft wesentlich effizienter ab als die Ligation blunt endender DNA. Man benötigt etwa die 50-fache Enzymmenge, um eine vergleichbare Ausbeute an Reaktionsprodukt zu erzielen. Die Zugabe von *Polyethylenglycol* (PEG, etwa 1 Volumenprozent des Reaktionsansatzes als Richtwert) kann blunt-end-Ligationen unterstützen.

Im Hinblick auf die Anwendung der Ligation als Operation im DNA-Computing sind folgende Seiteneffekte signifikant:

- Infolge intramolekularer Ligation können sich DNA-Ringe bilden. DNA-Ringe weisen in Elektrophorese-Gelen gewöhnlich ein anderes Laufverhalten auf als lineare DNA ([Mart\_96]).

- Die Exemplaranzahlen der durch die Ligation entstandenen DNA-Fragmentketten-Kombinationen können stark voneinander abweichen. Während von einer bestimmten Verkettung sehr viele identische DNA-Moleküle gebildet wurden, kann eine andere Verkettung nur durch sehr wenige DNA-Moleküle oder gar nicht repräsentiert sein.
- Das deaktivierte Enzym verbleibt im Reagenzglas. Obwohl es selbst keine enzymatische Aktivität mehr besitzt, kann es als Verunreinigung die Auswertung von Reagenzglasinhalten erschweren.
- Durch aktivitätsgeminderte Enzyme, hervorgerufen durch Alterung oder unsachgemäße Lagerung, kann der durch die Ligation bezweckte Effekt nicht oder nur sehr schwach auftreten, und das Ligationsprodukt unterscheidet sich trotz des Vorhandenseins ligierbarer DNA-Doppelstränge kaum von der anfänglich vorliegenden DNA.

Die Operation Ligation wirkt auf DNA-Doppelstränge mit sticky-Enden sequenzspezifisch, auf DNA-Doppelstränge mit blunt-Enden sequenzunabhängig und generell abhängig von einer chemischen Strangendenmarkierung. DNA-Einzelstränge werden durch diese Operation nicht verändert.

Bei der Ligation können schrittweise DNA-Doppelstränge wachsender Längen entstehen, wenn sich die zu Beginn der Reaktion vorliegenden DNA-Doppelstränge im Verlauf der Reaktion zunehmend aneinander lagern. Alle potentiellen Verkettungsmöglichkeiten können dabei gebildet werden. Die Ligation bewirkt Interaktionen zwischen verschiedenen DNA-Doppelsträngen und besitzt deshalb ein hohes Strangkombinationspotenzial.