

# Über die Evolution der Struktur sich autokatalytisch vermehrender Moleküle

Dissertation  
zur Erlangung des Grades  
des Doktors der Naturwissenschaften  
der Naturwissenschaftlich-Technischen Fakultät  
der Universität des Saarlandes

von  
Dipl.-Phys. PHILIPP ZIMMER

Saarbrücken  
2017

Tag des Kolloquiums:

09.10.2017

Dekan:

Univ.-Prof. Dr. Guido KICKELBICK

Mitglieder des Prüfungsausschusses:

Univ.-Prof. Dr. Dr. Karsten KRUSE  
Univ.-Prof. Dr. Albrecht OTT  
Univ.-Prof. Dr. Karin JACOBS  
Dr. Andreas TSCHÖPE

# Zusammenfassung

Auf der Erde entwickeln sich immer neue Spezies und passen sich an das Ökosystem an. Diese ständige Weiterentwicklung der Spezies ist im Allgemeinen schlecht verstanden. Wir stellen in dieser Arbeit ein molekulares Modellsystem vor, welches spontan neue Moleküle erzeugen und bestehende autokatalytisch vermehren kann. Dieses System kann dazu genutzt werden, die grundlegenden Prinzipien der Evolution zu untersuchen.

Das System wird sowohl in einer Gillespiesimulation analysiert als auch experimentell durch eine Implementierung in DNA mit Hilfe der Taq-Ligase. In beiden Fällen kann eine Phase, in der alle möglichen Moleküle entstehen, und eine Phase, in der eine strikte Längenverdopplung stattfindet, festgestellt werden. Die Erweiterung des Phasenraums der vorkommenden Polymere durch eine zufällige Reaktion führt zu einer Längenverdopplungskaskade, welche für alle Reproduktionsarten (enzymatisch, autokatalytisch, sexuell) stabil ist. Dieser theoretische Prozess kann prinzipiell immer weiter fortgeführt werden und ist nur durch die zur Verfügung stehenden Ressourcen beschränkt.

## Abstract

On earth, ever new species evolve and adapt to the ecosystem. This unceasing progress of the species' evolution is generally poorly understood. In this work, we present a model system, which spontaneously generates new molecules and autocatalytically proliferates existing ones. This system is satisfactory to examine fundamental principles of evolution.

The system is analysed both with the help of a Gillespie simulation and a realisation using DNA and Taq DNA ligase. Both implementations show a phase in which all possible molecules emerge and a phase in which a strict length-doubling of polymers is observed. The expansion of the phase space of occurring polymers is induced by a spontaneous reaction and leads to a length-doubling cascade, which is stable for all types of reproduction (enzymatic, autocatalytic, sexual). In principle, this theoretical process may progress continuously and is only restricted by available resources.



# Inhaltsverzeichnis

## Zusammenfassung – Abstract

<b>Einleitung</b>	<b>1</b>
<b>1 Eigenschaften der biologischen Evolution</b>	<b>5</b>
1.1 Vorbemerkungen über Evolution . . . . .	5
1.2 Die Zeitskalen der Evolution, innere und äußere Einflüsse . . . . .	8
1.2.1 Äußere Einflüsse . . . . .	9
1.2.2 Innere Einflüsse . . . . .	14
1.3 Modellsysteme und Grundlagen für chemische Evolution . . . . .	17
1.3.1 Der Aufbau der DNA . . . . .	18
1.3.2 Antiparallele, kodierende, lineare Polymere . . . . .	19
1.3.3 Ein autokatalytischer Replikator . . . . .	22
1.3.4 Entwicklung von Zusammenhängen von Spezies . . . . .	24
1.3.5 Ein aufkonzentrierender Reaktor im thermischen Gefälle . . . . .	28
1.4 Analyse der Voraussetzungen für ein System für chemische Evolution und dessen Konstruktion . . . . .	33
1.4.1 Analysen von von Kiedrowski . . . . .	33
1.4.2 Vergleich von exponentiellem mit parabolischem Wachstum . . . . .	36
1.4.3 Hyperbolisches Wachstum . . . . .	38
1.4.4 Konstruktion eines Systems für Evolution . . . . .	42
1.5 Abläufe der Evolution . . . . .	46
<b>2 Experimentelle Realisierung mit DNA</b>	<b>55</b>
2.1 Biotechnischer Aufbau . . . . .	56
2.1.1 Grundsätzliche Verfahrensweise . . . . .	56
2.1.2 Taq-Ligase und Reaktionspuffer . . . . .	56
2.1.3 Mung-Bean-Nuclease . . . . .	58
2.1.4 AfeI und BstUI . . . . .	58
2.1.5 Gelelektrophorese und Auswertung . . . . .	58
2.2 Experimente: Sequenzen und Zyklenprotokoll . . . . .	59
2.2.1 Wahllose Wucherung (Taq-Ligase) . . . . .	59
2.2.2 Einzelstrangzerlegung (Taq-Ligase und Mung-Bean-Nuclease) . . . . .	60
2.2.3 Heterodimerzerlegung (Taq-Ligase, AfeI und BstUI) . . . . .	62
2.2.4 Längenselektion (Taq-Ligase und Linker) . . . . .	62
2.3 Ergebnisse . . . . .	69
2.3.1 Wahllose Wucherung . . . . .	69
2.3.2 Einzelstrangzerlegung . . . . .	71

2.3.3	Heterodimerzerlegung . . . . .	71
2.3.4	Längenselektion . . . . .	71
2.4	Exkurs: UNAFold und Hybridisierungsraum . . . . .	75
2.4.1	Funktionsweise von UNAFold . . . . .	76
2.4.2	Hybridisierungsraum . . . . .	76
<b>3</b>	<b>Theorie</b>	<b>81</b>
3.1	Systemdefinition . . . . .	81
3.2	Analyse . . . . .	83
3.2.1	Wahllose Wucherung: $\alpha$ klein . . . . .	83
3.2.2	Längenselektion: $\alpha$ groß . . . . .	83
3.2.3	Koexistenz . . . . .	87
3.2.4	Entropie . . . . .	90
3.3	Vergleich von Theorie, Experiment und Evolution . . . . .	91
3.4	Phasenübergang, Selbstähnlichkeit und Plateaulänge . . . . .	94
3.4.1	Selbstähnlichkeit: Plateaukonzentration . . . . .	94
3.4.2	Phasenübergang . . . . .	94
3.4.3	Gleichgewicht im Plateau und exponentielles Wachstum . . . . .	96
3.4.4	Plateaulänge . . . . .	96
3.4.5	Exponentielle Verteilung der Plateaulängen für $\alpha \gg \alpha_c$ . . . . .	98
3.4.6	Deterministische Plateaulängen für $\alpha \gtrsim \alpha_c$ . . . . .	101
3.5	Betrachtungen zur Allgemeingültigkeit der Analysen . . . . .	105
3.5.1	Evolution mit korrigierten Raten für DNA und Ligase . . . . .	105
3.5.2	Evolution mit parabolischem, exponentiellem und hyperbolischem Wachstum . . . . .	108
3.5.3	Vergleich der Maxima bei gleichzeitiger und getrennter Änderung von $\alpha$ und $[A_1]_0$ . . . . .	113
3.5.4	Zerfall . . . . .	113
3.5.5	Die unendliche Evolution . . . . .	117
3.5.6	Koexistenz . . . . .	118
3.6	Perkolation . . . . .	120
3.6.1	Methoden und Beobachtungsgrößen . . . . .	121
3.6.2	Beispielsysteme . . . . .	125
3.6.3	Analyse . . . . .	126
	<b>Diskussion, Zusammenfassung und Ausblick</b>	<b>133</b>
	<b>Anhang</b>	<b>135</b>
	Nomenklatur . . . . .	135
	<b>Eidesstattliche Versicherung</b>	<b>137</b>
	<b>Danksagung</b>	<b>139</b>
	<b>Literaturverzeichnis</b>	<b>141</b>

# Einleitung

Es ist von je her ein Anliegen des Menschen, mehr über seine Abstammung herauszufinden. Charles Darwin untersuchte den Vorgang der Evolution an Tieren und konnte die Prinzipien der Variation des Phänotyps in den Nachkommen und dessen Selektion durch die größeren Überlebenschancen des am besten angepassten Individuums ausmachen; vgl. Abschnitt 1.1. Dies legte den Grundstein, um die Entwicklung des Lebens verstehen zu wollen, angefangen von den ersten Zellvorläufern vor rund 3,7 Milliarden Jahren [1] hin zum modernen Menschen. Dieser Prozess ist durch einen Informationstransfer in Form von Erbmaterial entlang der Entwicklung der Spezies gekennzeichnet [1]. Dieses Erbmaterial, auch Genom genannt, beschreibt den Phänotyp des Individuums einer Spezies und variiert innerhalb einer Spezies nur geringfügig [1]. Neue Individuen entstehen durch Klonen, wie z. B. der Teilung von Bakterien oder geschlechtliche Fortpflanzung [1]. Letzterer Prozess beinhaltet eine Vermischung der Gene der Elternteile, die an die Nachkommen weitergegeben werden und stellt damit eine Variation dar. Diese kann auch bei der geschlechtslosen Fortpflanzung in Form eines horizontalen Gentransfers erfolgen [1]. Beim Kopieren des Genoms auf die Nachkommen oder durch äußere Einwirkungen z. B. von ionisierender Strahlung, kann es zu Veränderungen am Genom kommen, welche Mutation genannt wird [1]. Die Vererbungs-gesetze bei geschlechtlicher Fortpflanzung wurden erstmals von Mendel beschrieben [1]. Das Genom ist größer als der benötigte Informationsgehalt und insbesondere bei chromosomal\* vorliegender DNA in Eukaryoten kann es zu Änderungen der Genfunktion kommen, ohne deren Informationsgehalt zu ändern, in dem Bereiche (de)aktiviert werden. Beispiele sind das Angelman- oder das Prader-Willi-Syndrom. Beide sind auf eine Veränderung des 15. Chromosoms zurückzuführen [2]. Jedoch stammt die Änderung beim Prader-Willi-Syndrom vom Vater und führt u. a. zu Heißhunger, mandelförmigen Augen und einem hängenden Mund. Das Angelman-Syndrom wird von der Mutter vererbt und resultiert u. a. in einem flachen Gesicht, einem breiten, lachenden Mund und einem markanten Kinn. All diese Prozesse sind Gegenstand aktueller Forschung und recht gut untersucht [2].

Weniger gut untersucht ist das langfristige Wirken der (darwinschen) Evolution. Dabei wird davon ausgegangen, dass eine Spezies nur in einer Nische\*\* leben kann; vgl. Kapitel 1.1. Naheliegenderweise wurden Räuber-Beute-Systeme als Modellsysteme analysiert [3]. Diese sind in der Lage verschiedene Spezies und Nischen abzubilden. Der Konkurrenzkampf wird analysierbar. Beobachtungen in der Natur und von Züchtungsexperimenten könnten aber ebenfalls die Veränderungen der Individuen hin zu einer neuen Spezies verdeutlichen. So ändert sich beispielsweise die Flügellänge von Klippenschwalben in unmittelbarer Nähe zu Autobahnen [4]. Die Tiere bekommen kürzere Flügel, so dass sie den schnell heranfahrenden

---

\* In Eukaryoten liegt das **Genom** in Form von **DNA** vor, die ihrerseits in mehreren **Chromosomen** gebündelt ist [1].

\*\* Eine **Nische** bezeichnet die von einer Spezies genutzten Umweltfaktoren. Eine Nische ist also ein Unterraum der Umwelt und mehrere koexistierende Spezies teilen sich unterschiedliche Nischen untereinander auf.

Fahrzeugen besser ausweichen können. Dies ist möglich, da Variation und Mutation (zufällige) Veränderungen in den Individuen erzeugen, die einen Vorteil für die Nachkommen bedeuten können. Dieser Vorteil kann irgendwann zu einer neuen Nische führen und damit zu einer neuen Spezies, die besser an die Umweltgegebenheiten angepasst ist, weil die Nische diese besser abbildet [1].

Um das Wirken der Evolution besser analysieren zu können, bedient man sich auch molekularer Systeme. Viele dieser Systeme versuchen den Ursprung des Lebens zu modellieren. Es ist beispielsweise möglich, DNA in einer Umgebung (eine Pore einer heißen Quelle), die als guter Kandidat zur Entstehung des Lebens angesehen wird, der Länge nach aufzukonzentrieren; vgl. Abschnitt 1.3.5. Molekulare Systeme können aber auch als abstraktes Modellsystem für Lebewesen jeglicher Art verstanden werden; vgl. Abschnitt 1.3.4. Verschiedene Spezies replizieren sich autokatalytisch und bilden ein stabiles, reproduzierendes System. Das Ziel und gleichzeitig der Nachteil dieser Systeme ist, dass meist eine best-angepasste Spezies entsteht. Dies steht im Widerspruch zu den Erfahrungen der tatsächlich stattfindenden Evolution, die immer wieder neue Spezies hervorbringt; vgl. Abschnitt 1.5. Die Spezies ähnlicher Nischen, aber getrennter, nicht verbundener geografischer Regionen, weisen zudem besondere, den Nischen angepasste und vergleichbare Merkmale auf. So gibt es beispielsweise in Europa Hirsche, in Afrika Springböcke und in Australien Kängurus, die allesamt größere, pflanzenfressende Tiere darstellen und diese Nische ausfüllen [1]. In dieser Hinsicht ist die Evolution sogar zum Teil deterministisch.

Wir werden in dieser Arbeit jedoch einen anderen Ansatz wählen: Wir reduzieren das Modell auf das Wesentliche, wählen einen molekularen (autokatalytischen) Reproduktionsansatz und ergänzen diesen durch das spontane Entstehen neuer Individuen als Variationskomponente. Damit ist es möglich, losgelöst von speziellen Details der betrachteten Spezies, eine universelle Aussage über das Verhalten unter verschiedenen Bedingungen zu treffen. Dies führt zu einem molekularen Modell, welches den immerwährenden Prozess der Anpassung an immer neue Nischen abzubilden vermag.

In Kapitel 1 geben wir eine Übersicht über die verschiedenen Evolutionsperioden und charakterisieren typisches Verhalten der Evolution in verschiedenen biologischen und theoretischen Beispielen unter unterschiedlichen Einflüssen. Äußere Einflüsse eignen sich gut, um spezielle Reaktionen der Evolution zu studieren. Anhand innerer Einflüsse können wir die kollektive Entwicklung mehrerer Individuen, die anfänglich einer Spezies angehörten, nachvollziehen. Sie sollen einen Überblick über die möglichen Evolutionsszenarien geben, von denen wir auch einige im Modellsystem beobachten.

Anschließend erläutern wir die Grundlagen des Modells und dessen Voraussetzungen, wie DNA, abstrakte Eigenschaften der DNA und das Wachstumsverhalten im template-basierten Prozess. Wichtige Modellsysteme im Kontext von molekularer Evolution und DNA werden ebenfalls diskutiert. Im Folgenden leiten wir aus diesen Tatsachen eine mögliche Konstruktion des hier vorgestellten Modells her. Abschließend diskutieren wir wichtige Modellsysteme, die ebenfalls versuchen die Evolution zu erklären.

In Kapitel 2 stellen wir die biotechnischen Grundlagen der Umsetzung des Modellsystems sowie die technischen Parameter vor. Zur Umsetzung sind einige theoretische Berechnungen zur Schmelztemperatur der DNA notwendig, die wir ebenfalls nutzen, um ein phänomeno-

logisches Modell aufzustellen, welches in Simulationen weiter verwendet werden kann. Im Rahmen der technischen Umsetzungsmöglichkeit können wir das längstmögliche evolutionäre Verhalten beobachten.

Die theoretischen Analysen werden in Kapitel 3 vorgestellt und besprochen. Nach der Definition des Systems werden grundlegende Resultate diskutiert, die von einer Phasengrenze abhängen. Verschiedene Charakteristika und Erweiterungen des Modells werden ebenfalls analysiert. Unter gewissen Randbedingungen ist eine prinzipiell immer fortwährende Evolution möglich. Die grundlegenden Ergebnisse werden ebenfalls im Hinblick auf die Perkolationstheorie untersucht.



# 1 Eigenschaften der biologischen Evolution

In diesem Kapitel stellen wir die grundlegenden Eigenschaften der Evolution vor. Dies stellt sich als komplex dar, da die biologische Evolution auf verschiedenen Zeitskalen abläuft [5]. Es sind evolutionäre Antworten von relativ kurzen Zeiträumen [6, 7, 8], d. h. wenigen Generationen, bis hin zu relativ langen [9, 10, 11, 12, 13], d. h. vielen Generationen, zu beobachten. Es stellt sich die Frage, wie man einen solchen Ablauf erfassen, quantifizieren und eventuell vorhersagen kann. Dazu betrachten wir zunächst Antworten auf äußere und intrinsische Einflüsse, um Möglichkeiten der Entwicklung realer evolutionärer Systeme zu erfassen. Dann gehen wir über zu diversen Modellsystemen, um das bisher theoretisch und experimentell untersuchte Spektrum darzustellen. Bestehende Systeme geben sowohl darüber Auskunft, wie man ein evolutionäres System konstruieren, als auch, wie man bestehende (Modell-)Systeme charakterisieren kann.

Aus diesen Überlegungen werden wir dann eine Vorstellung entwickeln, welche grundlegenden Eigenschaften notwendig sind, um Evolution zu charakterisieren. Mit Hilfe der vorgestellten Systeme und deren Eigenschaften können wir dann ein evolutionäres System konstruieren, welches wir in dieser Arbeit untersuchen werden.

## 1.1 Vorbemerkungen über Evolution

Zunächst möchten wir eine Übersicht über das bisherige Verständnis von Evolution geben, angelehnt an [1] und [14]. Die Evolution lässt sich in drei Abschnitte einteilen: physikalische, chemische und biologische, je nachdem, welche Art der Interaktion zwischen den Spezies vorherrscht und zu deren Entwicklung beiträgt. Die physikalische Evolution befasst sich mit der Entstehung von Atomen aus den Elementarteilchen. Daraus entstanden dann anorganische und organische Moleküle, die chemische Evolution.

### Chemische Evolution

Das „Ergebnis“ der chemischen Evolution ist die Bildung von biologisch funktionalen Molekülen verschiedenster Art. Aus den ersten, einfachen organischen Molekülen bildeten sich Makromoleküle, u. a. Lipide. Lipide sind die Grundlage zur Bildung von Membranen, die ihrerseits zur Bildung von Organellen, wie sie in Zellen vorkommen, und auch Zellen selbst dienen. Makromoleküle sind auch Informationsträger, wie z. B. RNA. In der sogenannten RNA-Welt [15] übernimmt RNA sowohl die Funktion eines Erbinformationsträgers als auch eine Enzymfunktion. Daraus sollen sich dann gemeinsam die DNA und die Proteine entwickelt haben, die zusammen mit den Membranen die ersten biologischen Zellen darstellen, die in Sedimenten erstmals vor ca. 3,7 Milliarden Jahren festgestellt werden konnten. Einige

dieser Zellen konnten durch Photosynthese Sauerstoff produzieren, aber dieser wurde durch nicht vollständig oxidiertes Eisen in Eisenoxid gebunden. Die Zellen lebten daher in einer sauerstofffreien Umgebung.

In diesem Kapitel werden wir einen genaueren Überblick über die Abläufe der chemischen Evolution geben und ein theoretisches Modellsystem entwickeln, welches aus (abstrakten) Monomeren Polymere bildet. Wir werden zeigen, dass diese Polymere den Gesetzen der darwinschen Evolution unterliegen. Ebenfalls hat Emanuel Gregor Worst ein experimentelles System konstruiert, welches vergleichbare Eigenschaften aufweist und dessen Grundeinheiten aus kurzen DNA-Strängen bestehen. Prinzipiell können diese Einheiten biologisch relevante Information tragen [1]. Wenn es möglich ist, diese geordnet zusammenzubauen, dann könnten sich so ausreichend lange DNA-Stränge aufbauen, die für biologisches Leben relevant sein können.

### Biologische Evolution

Aus den ersten biologischen Zellen differenzierten und entwickelten sich dann vor ca. 2 Milliarden Jahren, als der Sauerstoffgehalt durch vollständige Oxidation von Eisen zu  $\text{Fe}^{3+}$  angestiegen war, die zellkernfreien Prokaryoten: Blaualgen, Bakterien und Archaea. Später, vor ca. 1,4 Milliarden Jahren, traten die ersten Eukaryoten auf, die das Erbgut in einem Zellkern vom Rest der Zelle separieren. Aus ihnen entwickelten sich dann vor ca. 1,2 Milliarden Jahren die ersten mehrzelligen Organismen. Vor ca. 0,4 Milliarden Jahren traten die ersten Hinweise auf Landpflanzen auf. Es entstanden weitere höhere Organismen, u. a. auch der Mensch. Darwin [16] postulierte Prinzipien, nach denen die Evolution abgelaufen sein soll.

### Darwinsche Evolution

Charles Darwin [16] hat zunächst die Vorgehensweise von Züchtern untersucht, die die gewünschten Merkmale durch Auswahl (Selektion) aussuchen und gegebenenfalls zur weiteren Zucht verwenden. Darwin ging davon aus, dass Lebewesen viel mehr Nachkommen erzeugen, als zur Erhaltung des Bestandes notwendig wären. Diese Nachkommen sind jedoch nicht identisch, sondern variieren in deren Merkmalen (durch Mutation, sexuelle Fortpflanzung, horizontalen Gentransfer). Lebewesen stehen ununterbrochen im Wettbewerb um Ressourcen, wie Nahrung und Lebensraum. Demnach folgerte er, dass nur die best-angepassten (fittesten) überleben und sich deshalb fortpflanzen können. Auch unterschiedliche Lebewesen stehen im Wettbewerb untereinander. Ist eine Ressource von Lebewesen nicht oder kaum beansprucht, so stellt diese eine Nische dar, die dann von einer neuen Art (Spezies) genutzt werden kann. Es wird davon ausgegangen, dass eine Nische nur von einer Art bevölkert werden kann, da die schlechter angepasste aussterben würde\* [1].

---

\* Wenn mehrere Arten die **exakt gleichen Nischen** bevölkern würden, würde die fitteste die anderen Arten letztlich vollständig verdrängen, da sie die Umweltbedingungen, um die die Arten konkurrieren besser ausnutzen kann. Dieses **Konkurrenzausschlussprinzip** konnte Gause [17, 18] an Hand der Pantoffeltierchenarten *Paramecium aurelia* und *Paramecium caudatum* nachweisen. Die Pantoffeltierchen werden in einer Bakterienkultur des Bakteriums *Bacillus pyocyaneus* als Nahrungsquelle gehalten, auf der die Pantoffeltierchen gleich gut wachsen. Die Bakterienkultur wird regelmäßig ausgetauscht und die Bakterien vermehren sich unter den gegebenen Bedingungen nicht weiter. Überschüssige Pantoffeltierchen werden (auch zur Biomassenbestimmung) entfernt. Anfänglich findet eine Konkurrenz um die noch vorhandene

Einflüsse, die zur Veränderung einer Nische oder deren Entstehung führen, nennt man Selektionsfaktoren. Abiotische Selektionsfaktoren sind äußere Einflüsse wie z. B. Feuchtigkeitsgehalt, Temperatur, Nährstoffgehalt der Umgebung oder Lichteinstrahlung. Biotische Selektionsfaktoren sind innere Einflüsse wie z. B. Feinde, Parasiten, Geschlechtspartner oder um Nahrung konkurrierende Individuen der gleichen Spezies. Damit kann eine Veränderung der äußeren Einflüsse, aber auch der Inneren zu einer Veränderung bzw. Entstehung von Nischen führen.

Verschiedene Spezies, die voneinander abhängig sind und sich zusammen in einer gemeinsamen oder getrennten Nische entwickeln, nennt man koevolvierende Spezies.

Spätestens seit der Entdeckung des genetischen Codes ist bekannt, dass der „Bauplan“ eines biologischen Lebewesens im genetischen Code, der meist in Form von DNA vorliegt, gespeichert ist. Dieser bestimmt dann auch die nicht erlernten Eigenschaften des Lebewesens, sog. Merkmale, wie z. B. den Phänotyp (die äußere Erscheinungsform). Änderungen an diesem Code werden Mutationen genannt. Bei geschlechtlicher Fortpflanzung diploider Lebewesen liegen fast<sup>\*</sup> alle Erbinformationen doppelt vor und stammen zur Hälfte vom Vater und zur Hälfte von der Mutter. Ist die Erbanlage rezessiv, so tritt das Merkmal nur auf, wenn beide Erbinformationen entsprechend mutiert sind. Bei einer Mutation einer dominanten Erbinformation ist das Merkmal in jedem Fall ausgeprägt.

## Unterscheidung der Evolutionsarten

In dieser Arbeit wollen wir die Bezeichnungen der verschiedenen Evolutionsarten wie folgt festlegen: Die physikalische Evolution befasst sich mit rein physikalischen, d. h. vornehmlich kerntechnischen Wechselwirkungen. Sobald chemische, d. h. auf Molekülebene wirkende Wechselwirkungen auftreten, sprechen wir von chemischer Evolution. Die biologische Evolution wirkt auf biologische Organismen. Organismen bestehen aus einer, mehreren oder vielen Zellen. Die natürliche Evolution bezeichnet alle diese Evolutionsabschnitte, insbesondere jedoch die biologische Evolution, falls keine äußeren, künstlichen Einflüsse auf die Gesamtheit der betrachteten Individuen wirken. Allen diesen Evolutionsarten ist gemein, dass eine kollektive Entwicklung der Spezies hin zu einem Optimum intrinsisch angestrebt wird. Dieses Optimum auf der Stufe der Spezies wird durch Optimierung einzelner Individuen erreicht.

Ausgehend von Beobachtungen der biologischen Evolution wäre es physikalisch sinnvoll, einen Satz grundsätzlicher Gesetze abzuleiten, deren Vorhersagen konsistent mit allen experimentellen Gegebenheiten sind, wie z. B. den ersten und zweiten Hauptsatz der Thermodynamik [19]. Die darwinsche Evolution ist hingegen abgeleitet aus grundlegenden Eigenschaften von Populationen biologischer Organismen [19]. Da für natürliche Systeme die Randbedingungen veränderlich und typischerweise unbekannt sind, kann dieses Prinzip jedoch meist Hinweise auf die tatsächlichen Gegebenheiten geben [19].

---

Nahrung statt, so dass es *P. caudatum* gelingt etwas besser als alleine zu wachsen. Nach dem Verbrauch der Nahrung findet eine Umverteilung der Nahrung zwischen den beiden Pantoffeltierchenarten statt und *P. aurelia* kann *P. caudatum* vollständig verdrängen.

\* Die **Erbinformationen**, die das Geschlecht bestimmen, sind für männliche und weibliche Lebewesen unterschiedlich und liegen demnach nicht notwendigerweise (z. B. im Fall von männlichen Säugetieren oder weiblichen Vögeln) doppelt vor.

Wir wollen uns aber von den genauen biologischen Hintergründen entfernen und die grundlegenden Eigenschaften von biologischer und chemischer Evolution an Hand von Beispielen und Modellsystemen vorstellen. Aus diesen Betrachtungen entwickeln wir dann in Abschnitt 1.5 eine Gesamtübersicht über die Eigenschaften der Evolution, die auch für die chemische Evolution maßgeblich sein können.

## 1.2 Die Zeitskalen der Evolution, innere und äußere Einflüsse

In diesem Abschnitt stellen wir zwei Möglichkeiten vor, die Antworten der Evolution auf Einflüsse zu erfassen: Man quantifiziert zum einen die Antworten auf äußere Einflüsse, Einflüsse die auch durch den Menschen generiert werden [6]\*. In einer kontrollierten Umgebung würde man z. B. Spezies in einem Ökosystem beobachten und dieses manipulieren, in dem man z. B. das Nahrungsangebot entsprechend verändert oder Individuen (selektiv) entfernt. Unkontrolliert kämen z. B. die wirtschaftliche Nutzung des Ökosystems, durch Jagd oder Abholzung in Frage, aber auch die Einleitung von produktionstechnischen Abfallstoffen wie Abwässer. Diese Vorgehensweise ermöglicht es insbesondere Reaktionen einzelner Spezies auf konkrete Änderungen zu erforschen. Zum anderen schafft man gleichbleibend konstante Bedingungen [9, 10, 11, 12], um mögliche intrinsische Faktoren zu untersuchen. Hier versucht man, ein Ökosystem von äußeren Einflüssen so weit wie möglich abzuschirmen und dessen Entwicklung zu beobachten. Damit ist es einfacher, das kollektive Verhalten eines Ökosystems zu untersuchen. Gestaltet man dieses Ökosystem minimalistisch, sind insbesondere auch Aussagen zu einzelnen Spezies ohne vielzählige andere Einflüsse machbar. Man kann beispielsweise Bakterien in einem minimalen Nährmedium züchten [9, 10, 11, 12].

In den vorgestellten Beispielen zeigt die Evolution Antworten, die auf verschiedenen Zeitskalen ablaufen können. Dabei ist die Ursache des Einflusses prinzipiell unerheblich für die resultierende Zeitskala. Jedoch weisen im Allgemeinen große Änderungen für gewöhnlich eine rasche bzw. starke Reaktion auf, während quasikonstante Bedingungen häufig zu einem Gleichgewichts- oder stationären Zustand führen [20, 21]. Es sind jedoch auch in der Nähe eines Phasenübergangs durch kleine Änderungen große Zustandsänderungen möglich [20, 21]. Für biologische Systeme besteht ein „stationärer Zustand“ aus einem Fließgleichgewicht, da sie sich inhärent im Nichtgleichgewicht befinden: Lebewesen benötigen (Freie-)Energie und damit negative Entropie in Form von Nahrung, um ihre Entropie auf einem niedrigeren Niveau als die Umgebung zu halten und um die abgegebene Wärme zu ersetzen [22]. Aber auch dieser „stationäre Zustand“ unterliegt Veränderungen, da es für gewöhnlich bei der Reproduktion von Individuen zu Genfehlern kommt [23, 24] und diese altern. Jedoch ist die Mutationsrate pro Genom relativ konstant zwischen  $10^{-2}$  und 1, obwohl sich die Genomgröße um mehr als 6 Größenordnungen unterscheidet [23, 24]. Letzterer Tatsache wird in dem hier vorgestellten Modellsystem ebenfalls Rechnung getragen, indem pro Reaktion nur an einer Bindestelle ligiert wird, während die beiden Edukte nicht verändert werden.

---

\* Strenggenommen sind auch diese Einflüsse intrinsisch, da der Mensch ein Teil der Biosphäre ist. Da dieser aber deutlich andere Möglichkeiten zu deren Gestaltung hat als andere Lebewesen, sind diese Eingriffe grundsätzlich anderer Natur, weswegen sie auch als extern angesehen werden.

### 1.2.1 Äußere Einflüsse

Seit jeher gab es Einflüsse von außerhalb auf das Ökosystem der Erde. Sei es durch Meteoriteneinschläge oder durch Erdbeben, Vulkanausbrüche und heiße Quellen. Die Folgen solcher Ereignisse sind regional begrenzt oder erdumfassend. So haben z. B. Vulkanausbrüche regional einen starken Einfluss durch die Ausbreitung von Lava und einen globalen durch die großflächige Verteilung von Flugasche und Schwefel, der zu Schwefelsäureaerosolen umgewandelt wird [25]. Asche und Schwefelsäure können sich, zunächst auf eine Halbkugel begrenzt, auf der ganzen Erde verteilen und zu einer signifikanten Reduzierung der Sonneneinstrahlung durch Abschattung und Wolkenbildung führen [25]. Heiße Quellen können z. B. über einen langen Zeitraum lokal warmes Wasser und Stoffe aus dem Erdinneren abgeben und so als Energie- und Nahrungsquelle für verschiedene Lebewesen dienen [26]. Wegen dieser recht konstanten Bedingungen werden heiße Quellen als guter Kandidat für den Ursprung des Lebens angesehen [26, 27]; vgl. Abschnitt 1.3.5.

Wir werden im Folgenden Beispiele vorstellen, bei denen verschiedene evolutionäre Reaktionen auf äußere, kontrollierte Einflüsse hervorgerufen werden. Die Änderungen äußern sich im Phänotyp, welchem ein oder mehrere Gene zu Grunde liegen [1]. Daher kann es auf einen Selektionsdruck eine genau angepasste Reaktion geben, wie in dem hier vorgestellten Beispiel der Diamantschildkröte. Es kann auch zu einer Verschiebung des Merkmals führen, wie z. B. bei den Klippenschwalben und der Jagd auf Rotwild und Elefanten. Die Jagd auf Füchse jedoch bewirkt das Aussterben einer Teilpopulation mit einem entsprechend gejagten Merkmal. Dies ist die sog. gerichtete Selektion [1]. Bei den Darwinfinken kam es vorher zu einer aufspaltenden Selektion, die mittlerweile in eine stabilisierende Selektion übergegangen ist [1]. Wenn mehrere Gene an der Ausprägung eines Merkmals beteiligt sind, treten verschiedene Phänotypen auf, wie z. B. beim Hausfuchs.

Einflüsse, die auf den Menschen zurückgehen, sind besser dokumentiert und kontrolliert als natürliche. Schon seit langem züchtet der Mensch Tiere und Pflanzen als Nahrungsquelle oder als Haustiere. Als Beispiele seien hier Getreide, insbesondere Weizen, der aus wildem Emmer gezüchtet wurde [28], sowie Hunde, die von wilden Wölfen abstammen, genannt [7]. Man ging bisher davon aus, dass ein solcher Zuchtprozess lange gedauert hat [7, 28]. Jedoch konnte Dmitri Konstantinowitsch Beljajew zusammen mit Ljudmila Trut zeigen, dass innerhalb weniger Generationen Silberfüchse, das sind Rotfüchse mit schwarzem Fell, typische Merkmale von Haustieren aufwiesen, wenn sie im Hinblick auf Zähmheit gezüchtet wurden [7]. Füchse scheinen ein geeigneter Modellorganismus zu sein, da bisher nicht bekannt ist, dass eine Domestizierung stattfand. Im Folgenden werden wir das Experiment zusammenfassen.

#### Hausfüchse

Beljajew wollte herausfinden, wie schnell die Anpassung von wilden Tieren an den Menschen erfolgt, ein Prozess der beim Hund tausende von Jahren gedauert haben soll. Dazu züchtete er nicht nach Aussehen oder physischen Eigenschaften, sondern ausschließlich nach Verhalten. Dabei werden Silberfüchse, die aus einer Fellzuchtfarm stammten, dahingehend ausgewählt, wie zutraulich sie sich Menschen gegenüber verhalten. Böartige Tiere werden ausgeschlossen. Damit wurden anfangs nur bis zu 5% der männlichen und 20% der weiblichen Population zur Zucht verwendet. Es entstanden relativ rasch Füchse, die sich aus der Hand füttern und streicheln ließen, die winselten und ab der sechsten Generation auch Füchse, die unbedingt



(a) silberner (schwarzer) Hausfuchs



(b) georgisch-weißer Hausfuchs

Abbildung 1.1: **Hausfüchse.** Ergebnis der Zucht nach Zähmheit aus kommerziell gezüchteten Silberfüchsen nach ca. 35 Generationen. Es treten typische Merkmale von Haustieren auf, wie z. B. scheckiges Fell, hängende Ohren, kürzere und gerollte Schwänze. Die Kopfform ist ebenfalls verändert. Hausfüchse sind gutmütig, treu wie Hunde und unabhängig wie Katzen [7]. Die Fellfarbe georgisch-weiß ist in der Natur überhaupt nicht zu finden [29, 30]. Bild (a) aus [31] und (b) aus [32].

menschlichen Kontakt suchten und sich ähnlich wie Hunde verhielten. Diese machten nach 30–35 Generationen bis zu 70–80% der untersuchten Population aus. Der Prozess der Domestizierung konnte also stark beschleunigt werden, weil strikt nach einem geeigneten Merkmal selektiert wurde.

Wesentlich erstaunlicher sind jedoch die einhergehenden Veränderungen der zahmen Hausfüchse gegenüber ihren kommerziell gezüchteten Verwandten. Viele dieser Änderungen sind auch bei anderen domestizierten Säugetieren bekannt, so Beljajew. Es gibt bei den Hausfüchsen Tiere mit kürzeren Beinen (bei allen domestizierten Säugetieren treten Zwerg- und Riesentiere auf). Auch tritt scheckiges Fell wie bei allen domestizierten Säugetieren auf. Des Weiteren können gerollte (wie bei Hunden und Schweinen) und verkürzte Schwänze mit weniger Wirbeln auftreten (wie bei Hunden, Katzen und Schafen). Es wurden auch hängende Ohren, die es in freier Wildbahn nur bei Elefanten gibt, beobachtet, wie sie z. B. bei Hunden, Katzen, Schweinen, Pferden, Schafen, Ziegen und Rindern zu finden sind. Diese Tatsache wurde bereits von Darwin erwähnt [16]. Alle domestizierten Säugetiere bis auf Schafe weisen veränderte Reproduktionszyklen auf. Diese Veränderung ist beim Hausfuchs selten und die Welpen von Tieren mit veränderten Reproduktionszyklen sind nicht überlebensfähig. Andere Zuchtmethoden bei Füchsen konnten jedoch keine Änderung des Reproduktionszyklus bewirken, obwohl dies zuchttechnisch von Vorteil wäre.

Diese Änderungen, so Beljajew, sollten auf eine Änderung des Gleichgewichts zwischen Neurotransmittern und Hormonen zurückzuführen sein. Änderungen an Genen, die diese steuern, könnten zu weitreichenden Folgen in der Entwicklung führen. Und tatsächlich werden die meisten Veränderungen durch Verschiebungen in der Entwicklungsgeschichte bedingt. Insbesondere wird die Sozialisationsphase, eingeleitet durch die Reaktion auf Geräusche, nun andauernd bis zum Stadium „Angst vor dem Unbekannten“ vergrößert und damit

vergleichbar zu der von Hunden. Dies ist auch auf Serotonin, einen Neurotransmitter, der die Entwicklung beeinflusst, zurückzuführen. Je zahmer der Hausfuchs, desto höher ist die Serotoninkonzentration im Gehirn. Des Weiteren produzieren die für die Stressanpassung zuständigen Nebennieren in Hausfüchsen weniger Hormone, so dass es zu einer geringeren Stressreaktion kommt. Außerdem wurde die Schädelform kleiner und ebenmäßiger zwischen Männchen und Weibchen sowie die Schnauze kürzer und breiter.

Das scheckige Fell, so Trut, ist auf eine Mutation des Gens *Star* zurückzuführen, welches die Migration der Vorläuferzellen, die später für die Farbgebung verantwortlich sind, in die später farbgebenden Gebiete um ca. zwei Tage verzögert. Man nimmt an, dass diese nicht mehr an ihrem Bestimmungsort ankommen und absterben.

Dies sind alles weitreichende Veränderungen, die nicht offensichtlich mit dem Zuchtkriterium Zahmheit zusammenhängen. Der Geschwindigkeit, mit der die Evolution voranschreitet, und der Universalität der Ergebnisse, hier der gemeinsamen Eigenschaften der Haustiere, werden wir auch in Kapitel 3 wieder begegnen. Den Einfluss von weiteren äußeren Faktoren auf die freie Evolution stellen wir anhand von mehreren Beispielen frei nach [6] dar und wenden uns zunächst den Auswirkungen der Jagd anhand von drei Beispielen zu [33].

### **Jagd: Füchse, Rotwild und Elefanten**

Seit Urzeiten betreibt der Mensch Jagd. Während es früher mit primitiveren Waffen sinnvoll war, kranke, schwächelnde und kleinere Tiere zu erlegen, tritt mit moderner Waffentechnik zunehmend auch der Jagdsport mit dem Ziel der Erlangung von Trophäen als signifikante Komponente der Jagd in Erscheinung. Auch andere moderne Fang- und Jagdmethoden, die einen großen Teil der Gesamtpopulation betreffen, führen zu einem starken Selektionsdruck, der von der Evolution auf die eine oder andere Art beantwortet wird. Besonders sei hier auch das Sammeln – z. B. für wissenschaftliche Zwecke – von sehr seltenen Tier- und Pflanzenarten, die dadurch stark dezimiert werden und deshalb aussterben können, genannt.

Rotfüchse in Kanada kommen in drei Varianten vor: als Rotfuchs, als dunkelroter Fuchs und als Silberfuchs. Um 1830 lag der Anteil an Silberfüchsen bei 16 % der Gesamtpopulation. Da die Felle der Silberfüchse stärker nachgefragt wurden und einen höheren Preis erzielten als die der Roten und Dunkelroten, wurden diese stärker gejagt. Dies führte bis 1930 zu einem Rückgang der Population von Silberfüchsen auf 5 %.

Ähnlich verhält es sich mit den Geweihen des Rotwilds und den Stoßzähnen von Elefanten. Die Sportjagd zielt hier auf große Geweihe und Stoßzähne ab und dementsprechend haben diese Tiere weniger gute Möglichkeiten sich fortzupflanzen, was die genetische Veranlagung für große Geweihe und Stoßzähne dezimiert. Dies kann auch die Partnerwahl beeinflussen, da eine Paarung mit weniger gejagten Tieren von Vorteil sein kann. Im Endeffekt wird sich der Anteil der (vormals) gejagten Tiere reduzieren. So ist der Anteil an Elefanten mit fehlenden Stoßzähnen von 10 % auf 38 % gestiegen.

Neben der selektiven Jagd hinterlässt auch eine Überjagd bzw. Überfischung in Gewässern genetische Spuren, die wie auch in den anderen Fällen nur schwer wieder rückgängig zu machen ist, da der Selektionsdruck in Richtung Relaxation geringer ist. Einen Überblick über die bisher beobachteten systematischen Zusammenhänge bietet [33] und dort insbesondere Abb. 1.

## Diamantschildkröte

Beobachtungen zu der Entwicklung der Diamantschildkröte finden sich in [34]. Diese ist in der Chesapeake Bay (mittlerweile nur noch teilweise) seit über 60 Jahren (10–20 Generationen) durch Krabbenfischerei mit bestimmten Krabbenfallen bedroht. Diese Krabbenfallen haben eine Fangöffnung, die es den Schildkröten ermöglicht als Beifang in die Fallen zu geraten. Da sich die Fallen unter Wasser befinden und die Schildkröten Luft zum Atmen benötigen, ertrinken sie. Diamantschildkröten weisen einen Sexualdimorphismus in der Körpergröße auf, so dass die kleineren Männchen stärker vom Beifang betroffen sind als die Weibchen. Dies führt zu einer nicht unerheblichen Reduktion der überlebenden Anzahl von Tieren. Im Vergleich zu Diamantschildkröten aus dem Long Island Sound konnte folgendes festgestellt werden: Die Weibchen der durch Krabbenfischerei bedrohten Schildkröten sind größer und wachsen schneller. Damit erreichen sie auch schneller eine Größe, die nicht mehr durch die Öffnung der Krabbenfalle passt und sind ab dem Alter von 5 Jahren im Mittel um 15 % größer. Die Männchen, die, weil sie kleiner sind, immer durch die Öffnung passen, sind in ihrer Größe vergleichbar. Jedoch wurden im Gegensatz zur Kontrollgruppe in der Chesapeake Bay keine männlichen Tiere gefunden, die älter als 8 Jahre waren. Eine Gompertzfunktion kann benutzt werden, um das Wachstumsverhalten von Tieren zu beschreiben:  $l(t) = l_{\min} + l_{\max - \min} \cdot e^{-e^{-\lambda(t - t_{\max})}}$ , wobei  $l(t)$  die Größe zur Zeit  $t$ ,  $l_{\min}$  die minimale,  $l_{\min} + l_{\max - \min}$  die maximale Größe,  $t_{\max}$  der Zeitpunkt des maximalen Wachstums und  $\lambda$  die Wachstumsrate ist. Damit finden Wolak *et al.* [34] aus den Messwerten der Chesapeake Bay ebenfalls eine Verschiebung des maximalen Wachstums gegenüber den Schildkröten des Long Island Sound von 6–8 zu 4–6 Jahren. Die weiblichen Schildkröten werden also größer und passen früher nicht mehr in die Fallen. Es ist erstaunlich, dass nur weibliche Tiere von der evolutionären Größenänderung betroffen sind, obwohl Schildkröten eine temperaturabhängige Geschlechtsbestimmung haben und sich folglich die Erbanlagen teilen, die die Größe bestimmen. Jedoch sind Männchen schon ab dem Alter von 3–4 Jahren geschlechtsreif und Weibchen erst ab 6–8 Jahren. Für Männchen ist es daher weniger sinnvoll, viel Energie in das Wachstum zu stecken.

In diesem Beispiel führt ein Selektionsdruck zu einer sehr konkreten evolutionären Reaktion.

## Klippenschwalben

Brown *et al.* untersuchen seit 30 Jahren Klippenschwalben, die an Straßenbrücken o. Ä. nisten [4, 35]. Dabei fiel auf, dass die Anzahl der durch den Straßenverkehr getöteten Schwalben, die sie täglich suchten und zählten, im Zeitraum von 1983–2012 signifikant von durchschnittlich 20 auf 4 zurückging. Im Gegensatz dazu ist die Zahl der Schwalbennester gestiegen. Da die Zahl der verschiedensten Aasfresser konstant oder rückläufig war, kommt eine dadurch begründete Abnahme der toten Schwalben nicht in Frage. Betrachtet man die Flügellänge der Schwalben, so fällt auf, dass diese unimodal verteilt ist und im Fall der lebenden Schwalben in diesem Zeitraum um durchschnittlich 3 mm kürzer geworden ist, während die Flügellänge der toten Schwalben um 2 mm größer wurde. Eine mögliche Erklärung wäre, dass die kürzere Flügellänge es erlaubt, den sich nähernden Fahrzeugen besser auszuweichen. Eine andere wäre, dass ein angelerntes Verhalten zur Meidung von Fahrzeugen führt. Dies würde jedoch bedeuten, dass vor allem jüngere Schwalben von Unfällen betroffen wären, was jedoch nicht der Fall ist.

Hamilton [36] schlussfolgert, weil geografische Änderungen der Flügel-, Schnabel- und Schwanzlänge im Allgemeinen nicht zufällig, sondern geordnet sind, dass Selektionskräfte für diese Änderungen verantwortlich sind, und gibt eine umfassende Übersicht über die möglichen und nachgewiesenen Faktoren, die diese bewirken können. Damit besteht die Möglichkeit, dass durch die Änderung der Flügellänge eine neue Spezies entstehen kann. Eine ähnliche Situation besteht in Bezug auf die Schnabellänge bei der Darwinfinkenart Mittel-Grundfink, welche von Hendry *et al.* [37, 38] untersucht wurde und der wir uns nun zuwenden wollen.

### Mittel-Grundfink, eine Darwinfinkenart

Darwinfinken sind Singvogelarten auf der Insel Galápagos. Häufig vertreten sind der Klein-, Mittel- und Groß-Grundfink. Klein-Grundfinken haben kleine Schnäbel und Groß-Grundfinken große. Beim Mittel-Grundfink ist die Schnabelgröße bimodal verteilt und größer bzw. kleiner als die des Klein-Grundfinks bzw. Groß-Grundfinks. Wie kommt es aber zu einer bimodalen Verteilung der Schnabelgröße? Huber *et al.* [38] sind sich sicher, dass die Aufspaltung ursprünglich durch unterschiedliche Nahrungstypen zustande kam. Mittlerweile ist auch das Paarungsverhalten der Weibchen verändert: Sie paaren sich nur mit Männchen, die einen vergleichbar großen Schnabel haben. Dies lässt sich auch in dem eingeschränkten Genfluss zwischen den klein- und großschnäbligen Exemplaren erkennen. Sie schlussfolgern, dass der Mittel-Grundfink deshalb am Anfang einer Artenentstehung steht.

Eine gänzlich andere Beobachtung macht Hendry *et al.* in [37]: Aus den Daten der Schnabellänge, die von 1964 bis 2005 erhoben wurden, fällt in Academy Bay, Insel Santa Cruz, Galápagos auf, dass sich zwischen 1968 und 1973 die Bimodalität der Schnabellängenverteilung sehr stark reduzierte. Da sich die Gesamtbreite der Verteilung nicht wesentlich geändert hat, gehen sie davon aus, dass sich die relative Häufigkeit von Vögeln mit mittlerer Schnabellänge vergrößert hat. Klimatische Ursachen können ausgeschlossen werden, da sich die anderen Orte der Vergleichsmessung nur ungefähr 10 km voneinander entfernt befinden. Es bleibt nur die Zunahme der Bevölkerung in der Academy Bay als Ursache übrig. Diese nahm in Puerto Ayora zwischen 1974 und 2001 von 900 auf 9.582 zu und die Gesamtpopulation auf den Galápagos nahm zwischen 1962 und 1974 von 2.391 auf 4.037 zu. Organisierte Touristentouren werden seit 1967 angeboten. Mit den Menschen wurden sicherlich auch nicht einheimische Pflanzen eingeführt. Zu dem gibt es schon seit langem sog. Finkenfutterröhren, in denen Reis als Nahrung angeboten wird, u. a. auch weniger als 500 m von der Beobachtungsstelle in der Academy Bay entfernt, die von mind. 100 Grundfinken gleichzeitig genutzt werden. Durch die Beifütterung fällt die Ursache der Speziation weg und die Artenspaltung wird unterdrückt.

Man sieht hier deutlich, dass die Evolutionsfolgen keinesfalls robust sind, sondern durchaus einfach gestört werden können, wenn die notwendigen Voraussetzungen aufgehoben werden.

### Zusammenfassung

Allen Situationen ist gemein, dass bei genügend großem Selektionsdruck eine relativ rasche Antwort auch auf genetischer Ebene erfolgt, eine Relaxation bei fehlendem Selektionsdruck in den Ursprungszustand nur langsam vorstättengeht [33] und die Auswirkungen vielfältiger Natur sind. Bei Vögeln wird eine Spezies z. B. u. a. durch die Flügelspannweite definiert, so dass man die Reduktion der Flügellänge als Entstehungsbeginn einer neuen Spezies auffassen kann [6], wobei diese Änderung zunächst nur einige Individuen betrifft. Eine ähnliche

Situation ergibt sich im Fall der Darwinfinken und Diamantschildkröten: Die Grenzen zwischen Spezies und Individuen verschwinden, so dass eine strikte Definition von beiden für einen beschränkten Artenumfang nur einen gegenwärtigen (zeitlich eingeschränkten) Zustand abbilden und nicht für alle Zeiten gültig sein kann.

### 1.2.2 Innere Einflüsse

Um diese Zusammenhänge besser zu studieren und eventuell eine Relaxation in einen dynamischen, stationären Zustand zu untersuchen, eignen sich Systeme, bei denen kein äußerer Selektionsdruck mehr vorherrscht. Ein solches System wird von Richard Lenski [9, 10, 11, 12] untersucht, indem das Darmbakterium *Escherichia coli* in einem speziellen Nährmedium gezüchtet wird [9, 10]. Es wurden zwei bis auf eine Mutation identische Bakterien benutzt. Diese Mutation ist neutral im benutzten Medium, erzeugt aber auf Tetrazolium/Arabinose-Platten unterschiedlich gefärbte Kolonien, da ein Stamm Arabinose verwerten kann. Die beiden Bakterienstämme wurden dann in je 6 Linien aufgeteilt, die sich unabhängig voneinander entwickeln. Das Nährmedium ist ein minimales und enthält als Nahrungsquelle 139  $\mu\text{M}$  Glucose. Des Weiteren sind in dem Nährmedium 1700  $\mu\text{M}$  Zitrat enthalten. Man kann ungefähr 6,6 Generationen pro Tag züchten, die jeweils 1:100 verdünnt werden. Das Experiment wird seit ca. 20 Jahren, d. h. seit ca. 40.000 Generationen durchgeführt. Es ist möglich, die Bakterien einzufrieren und später zur weiteren Kultivierung oder Analyse aufzutauen. Alle 500 Generationen werden auf diese Art Proben archiviert. Es konnten unter anderem sowohl eine Koevolution als auch der Erwerb einer neuen Eigenschaft in Form einer alternativen Nahrungsquelle nachgewiesen werden. Diese beiden Phänomene werden wir uns nun genauer anschauen.

#### Koevolution: Entstehung paralleler Linien

In [11] untersuchen Le Gac *et al.* das Auftreten einer Koevolution. Zwei verwandte Organismen können sich ähnliche Nischen teilen und sind daher einem starken Wettbewerb unterlegen. Das Konkurrenzausschlussprinzip besagt jedoch, dass sich zwei Spezies nicht eine gemeinsame Nische teilen können, so dass letztendlich die unterlegene Spezies ausstirbt.

Le Gac *et al.* konnten nach 6.500 Generationen eine Aufspaltung der Bakterienkultur in zwei Linien, die sie L und S nannten, feststellen. Die knappe Ressource ist Glucose. L kann schneller als S damit wachsen, aber L gibt Nebenerzeugnisse ab, die S verwerten kann. Betrachtet man die Gene, die exprimiert werden, d. h. welche genetische Information in der Zelle Verwendung findet, so stellt man fest, dass diese bei L als auch bei S teilweise verschieden sind; auch im Vergleich zum Ursprungsbakterium. In der S-Linie kommt ein alternativer Reaktionsweg für Glucose zum Tragen und in der L-Linie findet sich ein neues Glucosetransportprotein. Dies ermöglicht zum einen die Verwertung von Abfallprodukten des gewöhnlichen Glucosereaktionswegs und zum anderen die schnellere Aufnahme von Glucose. Das bedeutet, dass je mehr Bakterien der L-Linie vorhanden sind, desto mehr Abfallprodukte fallen an und desto mehr Nahrung für die S-Linie steht bereit. Die L-Linie erzeugt also durch ihr Vorhandensein eine Nische, die es der S-Linie ermöglicht zu wachsen.

Wie kam es zu einer solchen Entwicklung? Le Gac *et al.* betrachteten dazu die Evolution zwischen verschieden alten (6.500, 17.000 und 40.000 Generationen) L- und S-Linien in den Konzentrationen 9:1 und 1:9, um zu überprüfen wie sich die beiden Linien entwickelt

haben. Diese charakterisierten sie anhand häufigkeitsabhängiger Selektion. Es gibt positive und negative häufigkeitsabhängige Selektion [39, 11]. Bei einer positiven Kopplung nimmt die Fitness, d. h. die Fähigkeit seine Eigenschaften weiterzugeben, mit dem Vorkommen der Spezies zu. Bei einer negativen Kopplung nimmt die Fitness mit zunehmendem Vorkommen ab. Wenn die beiden Linien aus der selben Generation stammen, hat die jeweils niedriger konzentrierte einen Vorteil, also eine negative häufigkeitsabhängige Selektion. Geringer konzentrierte L-Linien können in S-Linien jeden Alters eindringen. Ältere S-Linien können in jüngere L-Linien eindringen, jedoch keine jüngeren S-Linien in ältere L-Linien. Damit ist gezeigt, dass die S-Linie aussterben würde, wenn sie sich nicht zusammen mit der L-Linie entwickelt hätte. Das bedeutet, dass die L-Linie in die Nische der S-Linie eingegriffen hat, da ansonsten die S-Linie nicht weiter von der Anpassung der L-Linie beeinflusst worden wäre.

Le Gac *et al.* können also feststellen, dass eine Linie ein Nebenprodukt produziert, welches die andere Linie verwerten kann, so dass sich eine Nische in zwei aufgespalten hat. Damit können sich die beiden Linien gemeinsam entwickeln, indem sie auf gegenseitige Entwicklungen reagieren.

### **Eine neue Eigenschaft: die Verwertung einer alternativen Nahrungsquelle**

Wir erinnern uns, dass im Nährmedium neben wenig Glucose Zitrat im Überfluss vorkommt. Es wäre also für die Bakterien sehr vorteilhaft, so Blount [10], auch Zitrat als Nahrungsquelle nutzen zu können. *Escherichia coli* kann allerdings Zitrat als Energiequelle in einer Sauerstoffatmosphäre nicht nutzen – eine Eigenschaft, die diese Spezies definiert. Jedoch kann *E. coli* mit Hilfe eines Cosubstrates wie Glucose oder Fructose [40] unter einer sauerstofffreien Umgebung Zitrat verwerten. Man geht davon aus, dass die fehlende Möglichkeit des Zitrattransports durch die Zellmembran unter Sauerstoff der einzige Grund für dessen Nichtverwertung ist. Atypische *E. coli* bewerkstelligen dies mit Hilfe eines Plasmids [41], das sind eigenständige DNA-Ringe, die sich autonom replizieren können [14]. Diese Plasmide wurden von anderen Bakterien erworben [41].

Dennoch hat sich nach weniger als 31.500 Generationen ein Zit<sup>+</sup> Phänotyp entwickelt, wie Blount *et al.* feststellen. Dieser Phänotyp kann Zitrat als einzige Energiequelle nutzen, da er in einem Medium und auf Agar-Platten, welche nur Zitrat als Energiequelle enthalten, gut wächst. Ein weiterer Test auf Zitratverwertung ist das Simmons-Agar. Wenn die Organismen das Zitrat als einzige vorhandene Energiequelle nutzen können, dann verbrauchen sie ebenfalls Stickstoff, der in Form von Ammoniumdihydrogenphosphat vorliegt. Der Verbrauch dieses Stoffes erzeugt einen Anstieg des pH-Werts, der mit Bromthymolblau angezeigt wird [42]. Dieser Test ist positiv beim Zit<sup>+</sup> Phänotyp. Jedoch erfolgt die Zitratverwertung langsamer als die Glucoseverwertung, da das Wachstum langsamer respektive schneller in entsprechenden Medien ist. Neben dem Zit<sup>+</sup> Phänotyp existiert ein Zit<sup>-</sup> Phänotyp, der Glucose besser verwerten kann und daher anfangs schneller wächst und letztlich, unabhängig vom anfänglichen Verhältnis, mit einem 1%-igen Konzentrationsverhältnis vorkommt; es handelt sich also um eine negative häufigkeitsabhängige Selektion.

Blount *et al.* folgern weiter, dass der Erwerb der Möglichkeit Zitrat zu verwerten in mehreren Schritten erfolgte. Dies legt zum einen die lange Zeitverzögerung bis zum Entstehen dieser Mutation nahe. Zum anderen wachsen Bakterien bis zur 33.000. Generation, wenn überhaupt, dann langsamer auf einem Medium, das nur Zitrat enthält, d. h. sie brauchen viele Stunden um zu wachsen; jüngere Bakterien wachsen sofort. Daher stellt sich die Frage, ab

welcher Generation die Bakterien überhaupt auf einem Zitrat-Agar wachsen können und wann notwendige Mutationen, um Zitrat verwerten zu können, stattgefunden haben. Durch den Glucosemangel herrscht dort ein stärkerer Selektionsdruck, so dass die darauf folgenden Anpassungsschritte möglicherweise schneller ablaufen. Von je 200 parallelen Experimenten traten bei der 20.000., der 27.000. und den neueren Generationen vereinzelt (d. h. bis zu 2) Zit<sup>+</sup> Phänotypen auf. Dabei kam es bei fast allen Experimenten zu einer Bildung mehrerer Kolonien, was darauf hindeutet, dass notwendige Mutationen schon früher stattgefunden haben müssen; die letztlich für diesen Phänotyp entscheidende Mutation aber später, da der Zit<sup>+</sup> Typ um mindestens 6 Tage verzögert, verglichen mit dem bereits ausgeprägten Zit<sup>+</sup> Typ, auftritt.

Blount *et al.* stellen zwei mögliche Szenarien der Evolution vor: jene von Gould [43, 10] und jene von Conway Morris [44, 10]. Gould behauptet, dass historische Zufälle die Evolution größtenteils unvorhersehbar machen. Obwohl jede lokale Änderung entlang der evolutionären Entwicklung eine kausale Beziehung zu den jeweiligen Umständen hätte, wäre das Ergebnis jedoch von den Details der jeweiligen Änderungen abhängig, so dass kleine Änderungen der Umstände weitreichende Auswirkungen hätten. Insbesondere bedeutet das, dass eine Evolution unter gleichen Bedingungen ausgehend von einem Punkt in der Vergangenheit eine gänzlich andere Welt als die heutige zur Folge hätte. Morris indes behauptet, dass die Wege der Evolution zwar verschieden seien, jedoch die Ziele begrenzt. Daher sei Evolution konvergent und könne daher allgemein wiederholt werden mit dem Ergebnis, dass bis auf unwesentliche Details die wesentlichen Eigenschaften reproduziert würden.

Wir meinen jedoch, dass beide Szenarien zugleich (un)zutreffend sind. Wesentlich ist der historische Verlauf der Anpassung einer Spezies. Hat eine Spezies im Laufe der Evolution eine Eigenschaft erworben, kann dies zu einer schnelleren oder auch langsameren Anpassung führen. Die Ausprägung erfolgt typischerweise schneller bei einem erneuten oder ähnlichen Selektionsdruck im Vergleich zum ursprünglichen Selektionsdruck oder kann andere Ausprägungen auch verhindern. Falls eine Ausprägung abhanden gekommen ist, ist auch die Geschichte des Verlusts wesentlich: Die Ausprägung kann noch latent vorhanden sein und gegebenenfalls schnell wieder zur Ausprägung gelangen oder während dieses Verlusts abhanden gekommen sein oder gar Ausprägungen verhindern.

### Stagnation des Artenreichtums im Meer für längere Zeiten

Auch die Entwicklung der Lebewesen im Meer und an Land läuft unterschiedlich ab und ist in weiten Teilen unabhängig voneinander. Daher ist es sinnvoll, diese getrennt zu betrachten. Stollmeier *et al.* [13] untersuchen in einem Modell die möglichen Ursachen von Stagnation und Wachstum der Biodiversität der Erde. Am Anfang des Phanerozoikums (vor 541 Ma) gab es nur wenige Spezies. Diese Anzahl hat sich bis heute auf einige Millionen vergrößert. Die Landlebewesen sind in dieser Zeit exponentiell gewachsen, wohingegen die Anzahl der Meereslebewesen auch Stagnationen unterlegen ist. Sie nehmen an, dass sich die Meereslebewesen dennoch exponentiell vermehren, jedoch dieses Wachstum durch vorübergehende Verlangsamungen und Unterbrechungen gekennzeichnet ist. Unterschiedliche Raten für Entstehung  $\mu$  und Aussterben  $\epsilon$  sollen für den qualitativen Unterschied zwischen dem kontinentalen und dem marinen Wachstum der Artenvielfalt verantwortlich sein. Spezies werden in Klassen nach der Anzahl der Abhängigkeiten eingeordnet, so dass ein Aussterben einer Spezies in Klasse  $l_0$  ebenfalls das Aussterben der abhängigen Spezies in den Klassen  $l > l_0$  zur Folge hat. Die

Aussterbensrate nehmen sie als konstant an (Van Valen-Gesetz) und jede entstehende Spezies ( $l \geq 2$ ) ist von einer anderen zufällig gewählten Spezies in der nächst niedrigeren Klasse abhängig. (Eine Abhängigkeit von mehreren Spezies führt zu einer zu großen Aussterbensrate.) Entstehende Spezies werden mit definierten Wahrscheinlichkeiten entweder in die Klasse der Ursprungsspezies  $l_1$  oder in die Klassen  $l_1 \pm 1$  eingeordnet.

Abhängig von dem Parameter  $\lambda := \frac{\epsilon}{\mu}$  treten im Zeitverlauf ausgewogene und unausgewogene Zustände auf. Unausgewogene Zustände kennzeichnen sich durch eine Akkumulierung hin zu größeren Klassen, so dass Diversitätsmaxima auftreten. Diese werden von größeren Aussterbenskaskaden gefolgt, welche sich als Diversitätsplateau zeigen. Am Ende eines Plateaus steht dann wieder ein ausgewogener Zustand und weiteres Wachstum ist möglich. Für die Anzahl der Meereslebewesen ist  $\lambda_{\text{mar}} \approx 0,68$  und für die kontinentale Artenvielfalt ist  $\lambda_{\text{kont}} \approx 0,57$ . Mit diesen Parametern treten Plateaus, die länger als eine Zeit  $d$  dauern, mehr als doppelt so häufig für das Wachstum der Anzahl der Meereslebewesen im Vergleich zu den Landlebewesen auf.

#### Zusammenfassung

Anhand dieser Beispiele kann man deutlich erkennen, dass es sinnvoll ist, auch weitestgehend abgeschlossene, keinem äußeren Selektionsdruck ausgesetzte Systeme zu untersuchen. Insbesondere können gänzlich neue Eigenschaften entstehen (wie die Möglichkeit, eine reichlich vorhandene Energiequelle wie Zitrat anzuzapfen), die wir in dem in dieser Arbeit untersuchten Modell als Restriktion auf eine der möglichen Nahrungsquellen wiederfinden werden; vgl. Abschnitt 3.2.2. Auch die Ausbildung neuer Linien in Form einer zusammenhängenden Entwicklung ist eine Eigenschaft, die wir später theoretisch darstellen werden; vgl. Abschnitt 3.2.3. Von besonderer Bedeutung ist im Weiteren auch die Herausbildung von Plateaus, wie eine Analyse zeigen wird; vgl. Abschnitt 3.2.2.

Im Folgenden werfen wir einen Blick auf die Möglichkeiten von Modellsystemen zur Erforschung der Anfänge und Grundlagen der Evolution und deren Fähigkeiten diese abzubilden.

## 1.3 Modellsysteme und Grundlagen für chemische Evolution

Wir wollen nun die Anfänge der chemischen Evolution beleuchten, d.h. den Übergang von chemischen Molekülen zu biologisch funktionalen Einheiten beschreiben. Dieser Übergang wird als chemische Evolution bezeichnet [1]. Grundlegende Experimente wurden von Miller [45] und Urey durchgeführt, in denen sie zeigen konnten, dass unter relativ einfachen und mit der Urerde kompatiblen Bedingungen verschiedene Aminosäuren entstehen. Aus den Proben von Originalexperimenten Millers, die er unter leicht abgewandelten Bedingungen durchgeführt hatte, konnten Johnson *et al.* [46] feststellen, dass sich gegebenenfalls auch größere Mengen von Aminosäuren herstellen lassen. Es liegt also nahe, dass eben solche Aminosäuren am Ursprung des Lebens vorhanden waren und in den ersten „Organismen“<sup>\*</sup> Verwendung fanden. Wollrab *et al.* [47] untersuchten Millers Experiment zusätzlich mit modernen Methoden und konnten verschiedene Klassen organischer Moleküle in Lösung nachweisen. Einen Überblick über weitere Reaktionswege und -methoden zur Entstehung biologisch relevanter Moleküle

---

\* **Organismus** meint in diesem Zusammenhang eine einfache, sich selbst reproduzierende und mutierende Entität.

bietet [48] in Kapitel 7. Als Beispiele seien hier die Fischer-Tropschsynthese, eine katalytische Synthese von Kohlenwasserstoffen aus Kohlenmonoxid, und die Streckersynthese, eine Synthese von Aminosäuren aus Aldehyden und Ammoniak sowie Cyanwasserstoff, welche alle als Produkte des Millerexperiments vorkommen, genannt. Die weitere Synthese und Stabilität von Bausteinen der RNA und DNA, wie z. B. Purinen (z. B. Adenin und Guanin), Pyrimidine (z. B. Cytosin, Uracil und Thymin) sowie Zucker (z. B. Ribose) sowie die Nukleosidsynthese, werden in [48] in den Kapiteln 8 und 9 vorgestellt.

Haben sich geeignete Moleküle für eine chemische Evolution gebildet, stellt sich die Frage, wie sich diese vermehren, entwickeln und selektieren können. Grundlegende Arbeiten wurden von von Kiedrowski [49, 50, 51], Orgel *et al.* [52], Eigen und Schuster [53, 19, 54, 55, 56], Kauffman [57, 56] und Braun *et al.* [58, 59, 60] veröffentlicht. Diese Arbeiten werden wir im Folgenden genauer vorstellen, aber zunächst wenden wir uns dem grundlegenden Aufbau der DNA zu.

### 1.3.1 Der Aufbau der DNA

Im Folgenden beschreiben wir den Aufbau der DNA, wie er in [1] und [14] beschrieben und in Abb. 1.2 dargestellt ist. Die DNA ist ein lineares, antiparalleles, gerichtetes Polymer aus Nukleotiden. Ein Nukleotid besteht aus einem Phosphatrest und einem Nukleosid, welches seinerseits aus einer Pentose und einer von 4 Nukleinbasen besteht. Bei der DNA ist als Pentose Desoxyribose verbaut. An dieser hängt dann eine der 4 Nukleinbasen Cytosin, Thymin, Adenin oder Guanin. Diese werden üblicherweise entsprechend mit C, T, A und G abgekürzt. Der Phosphatrest und die Pentose bilden das „Rückgrat“ der DNA, in dem sie als Polymer die lineare Struktur vorgeben, vgl. Abb. 1.2.a-b.A. Dabei gibt die Nummerierung der Pentose die Bezeichnung für die Enden der DNA an. Der Phosphatrest ist am 5. Kohlenstoffatom der Pentose und die OH-Gruppe am 3. gebunden. Folglich heißen die Enden 5' respektive 3'; vgl. Abb. 1.2.b.C. Als Nukleinbasen kommen die vier von Kossel identifizierten [61] Pyrimidine Cytosin und Thymin und die Purine Adenin und Guanin vor; vgl. Abb. 1.2.a-b.B. Dabei sind alle diese Bestandteile kovalent gebunden. Zwischen Purinen und Pyrimidinen kann sich jedoch auch zusätzlich eine schwache Bindung in Form von Wasserstoffbrücken ausbilden, welche Hybridisierung genannt wird. Dabei hybridisieren C und G (mit 3) sowie A und T (mit 2 Wasserstoffbrücken) jeweils untereinander, so dass 2 antiparallele DNA-Stränge mit den Nukleinbasen innen und einander gegenüber hybridisieren. Es entsteht eine Doppelhelix, deren Struktur Watson und Crick [62] richtig deuten konnten; vgl. Abb. 1.2.c.

DNA ist flexibel und kann auch stabile dreidimensionale Strukturen ausbilden [63].

RNA ist ein ähnlich aufgebautes, allerdings einsträngiges, lineares Polymer. Es hybridisiert auf Grund seiner veränderten Struktur hauptsächlich nur mit sich selbst und besitzt die Fähigkeit, sich falten zu können (d. h. eine dreidimensionale Struktur auszubilden), so dass diese eine enzymatische Funktionalität haben kann. Die Unterschiede zum Aufbau der DNA sind: Als Pentose ist bei der RNA Ribose verbaut und die Funktionalität der Nukleinbase Thymin übernimmt Uracil (U), dem im Vergleich die CH<sub>3</sub>-Gruppe fehlt.

Die Abfolge der einzelnen Nukleotide kodiert die genetische Information, welche im biologischen Sinn von den Eltern an deren Nachkommen weitergegeben wird. Im Hinblick auf die Konstruktion von evolutionären Systemen mit DNA- oder RNA-ähnlichen Molekülen ergeben sich interessante Konsequenzen, die wir im Folgenden analysieren werden.

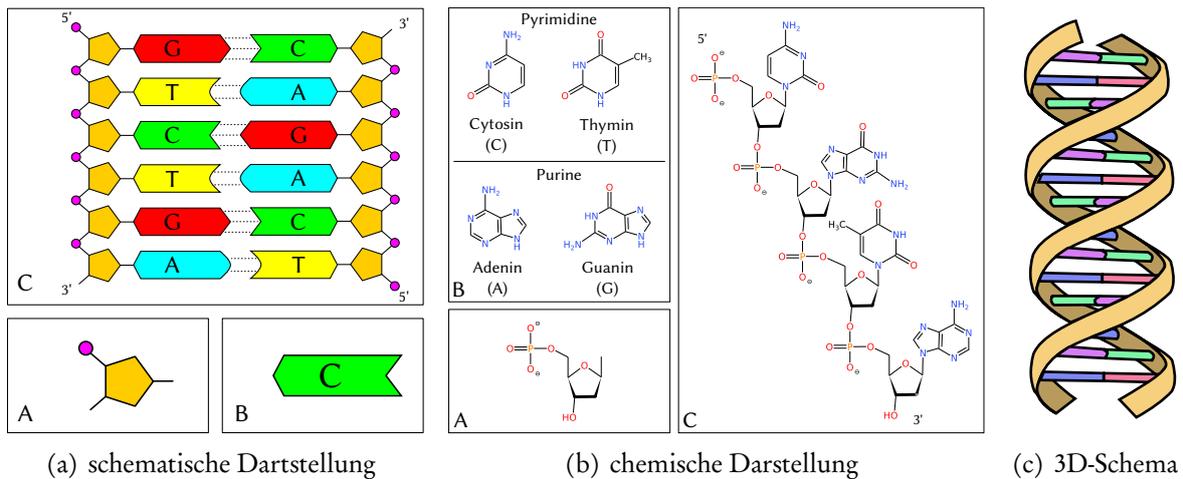


Abbildung 1.2: **Der Aufbau der DNA.** (A) Das „Rückgrat“ der DNA besteht aus einem linearen Polymer aus einem Phosphatrest am 5'-Ende einer Desoxyribose an deren 3'-Ende. (B, C) Am 1'-Ende ist eine der 4 Nukleinbasen Cytosin (C), Thymin (T) sowie Adenin (A) und Guanin (G) kovalent gebunden. (C) DNA hybridisiert antiparallel mit den konjugierten Basen A–T und C–G, so dass sich eine dreidimensionale Doppelhelix, wie in Schema (c) ergibt. Die Angabe von Sequenzen erfolgt normalerweise vom 5'-Ende zum 3'-Ende. Abb. (a-b) inspiriert von [1] und [14]; Abb. (c) aus [64] frei nach [62].

### 1.3.2 Antiparallele, kodierende, lineare Polymere

In diesem Abschnitt beschreiben wir die grundlegenden Möglichkeiten der Hybridisierung (d. h. der schwachen Bindung) von allgemeinen antiparallelen, kodierenden, linearen Polymeren und geben eine Übersicht über die möglichen Reaktionsformen, wie Faltung und Hybridisierung. Zwei vollständig komplementäre Polymere können vollständig hybridisieren; vgl. Abb. 1.3.a. Diese Situation bezeichnen wir im Folgenden als Vollüberlappung. Diese Eigenschaft lässt sich nutzen, um die vollständige Information, die in einem solchen Polymer kodiert ist, zu verarbeiten, d. h. zu erkennen, zu kopieren bzw. aneinander zu hängen. Werden zwei komplementäre Stränge kovalent gebunden, wie in Abb. 1.3.b, so hat der resultierende Strang die Eigenschaft, sich auf sich selbst hybridisieren zu können. Das X beschreibt hier einen möglichen Sequenzanteil, der benötigt wird, um die strukturelle Biegung zu vollziehen und welcher durchaus ein Teil der ehemaligen Sequenzen C und  $\bar{C}$  sein kann. Bei der DNA beträgt die Länge einer solchen Sequenz X minimal 3 Basen [65, 66] und wird Hairpin\* genannt.

\* Ein **Hairpin** ist eine Hybridisierung innerhalb eines Moleküls, die nur teilweise ausgebildet ist, so dass eine Schleife entsteht.

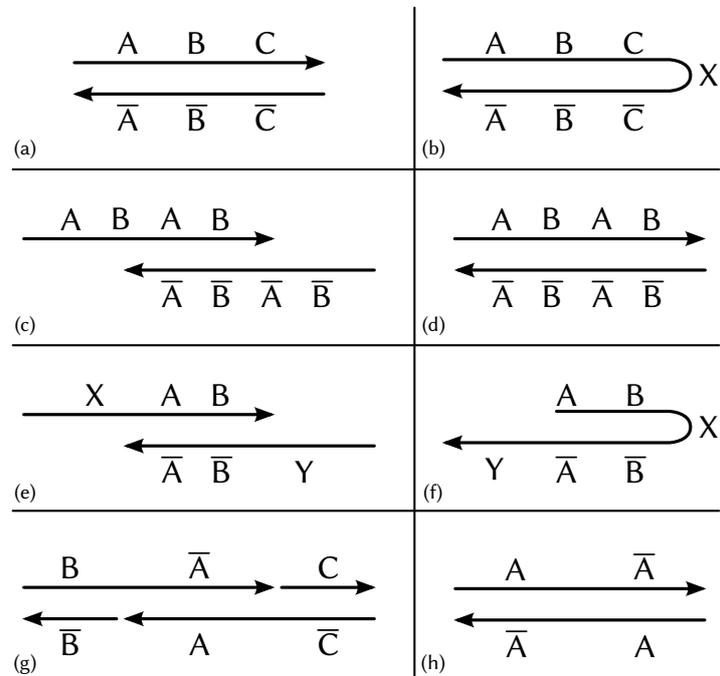


Abbildung 1.3: **Verhalten von kodierenden, antiparallelen, linearen Polymeren bei Hybridisierung.** Es werden die verschiedenen Möglichkeiten der Hybridisierung in Abhängigkeit von der Kodierung (Sequenzen) dargestellt. (a-b) Komplementäre, disjunkte Sequenzen hybridisieren. Im Fall der Selbsthybridisierung bleibt ein Bereich X möglicherweise unhybridisierbar aufgrund von strukturellen Einschränkungen. Dies ist bei sog. palindromischen Sequenzen, das sind selbstkomplementäre Sequenzen, der Fall. (c-d) Repetitive, komplementäre Sequenzen können zu einer Teilhybridisierung führen, die sich dann in eine Vollhybridisierung auflöst (d). (e-f) Überhänge aus beliebigen Strängen sind sowohl bei Hybridisierung als auch bei Selbsthybridisierung möglich. (g) Allgemeines Beispiel für Überhänge und den daraus resultierenden kovalenten Bindungsmechanismus. Zwei teilkomplementäre Stränge hybridisieren und es bilden sich Überhänge. Diese dienen als Templates für andere Stränge, die hybridisieren und so bei Anwesenheit entsprechender Enzyme kovalent gebunden werden können. (h) Überhänge treten bei rein selbstkomplementären Strängen nicht auf.

Tritt eine selbstkomplementäre\* Sequenz in einem Evolutionsexperiment als Template\*\* (Matrize) auf, so treten die Hybridisierung der Edukte auf dem Template und die Selbsthybridisierung des Templates in Konkurrenz, da die Hybridisierungsenergien vergleichbar sind, insbesondere je länger die Stränge sind. Erst wenn die Edukte kovalent gebunden sind, ist die Hybridisierungsenergie größer, sofern man von einer konstanten Bindungsenergie pro Bindungsstelle ausgeht, weil eine größere Anzahl an Bindungsstellen pro Molekül gebunden ist. Dies bedeutet, dass ein selbstkomplementäres Template sich selbst größtenteils inaktivieren kann; vgl. Abb. 2.2 und Kap 2.3.1. Für DNA kann die Bindungsenergie einfach berechnet werden. Diese Berechnung werden wir in Abschnitt 2.4.1 genauer erläutern.

Eine weitere mögliche Konstellation sind repetitive Sequenzen wie in Abb. 1.3.c-d. Diese können teilweise oder ganz überlappen; vgl. Abb. 1.3.c,d. Da die Vollüberlappung energetisch günstiger ist, wird im Allgemeinen diese eingenommen. Der Prozess zum Erreichen der Vollüberlappung ist Gegenstand aktueller Forschung [67]. Es werden zwei mögliche Kandidaten untersucht: Zum einen können versetzte Stränge aneinander ‚vorbeigleiten‘ (‚slithering‘ genannt). Zum anderen können mehrere, intern hybridisierte Regionen konkurrieren, um eine energetisch günstigere Überlappung (d. h. bei komplementären Sequenzen die vollständige Überlappung) zu erreichen. Dieser Prozess wird ‚interne Versetzung‘ (‚internal displacement‘) genannt. Pörschke [68, 67] folgert aus seinen Untersuchungen zu RNA, dass die Reaktionskinetik zweiter Ordnung sei und deshalb verschobene Sequenzen in der Lage sein müssen, sich so neu anzuordnen, dass eine voll ausgebildete Struktur entsteht, ohne sich gänzlich zu lösen.

Sind nur Teilstränge komplementär zueinander, so bilden sich Überhänge; vgl. Abb. 1.3.e-f. Dabei können sowohl zwei Stränge als auch nur ein Strang mit Selbsthybridisierung involviert sein. Überhänge treten insbesondere auch bei zwei selbstkomplementären Strängen mit invertierten Sequenzblöcken ( $\overline{AA} - \overline{AA}$ ) in einer Konstellation wie in Abb. 1.3.e ( $X = \overline{A}$ ,  $Y = A$  und  $B = \emptyset = \overline{B}$ ) auf. Auf solche Überhänge können dann andere, insbesondere auch kürzere Stränge hybridisieren und z. B. enzymatisch kovalent gebunden werden; vgl. Abb. 1.3.g. Dies führt im Strang zu wiederholtem Auftreten bereits vorhandener Sequenzen, falls die beiden Matrizen selbstkomplementär sind und aus zwei gleichen, komplementären Strängen bestehen, die in unterschiedlicher Reihenfolge angeordnet sind; vgl. Abb. 1.3.e mit  $X = \overline{A}$ ,  $Y = A$  und  $B = \emptyset = \overline{B}$  bzw. Abb. 1.3.g mit  $C = B$ . Bei nur einem selbstkomplementären Strang können keine Überhänge auftreten; vgl. Abb. 1.3.h.

### Folgen für ein Evolutionsexperiment

Auf ein Evolutionsexperiment, das auf antiparallelen, kodierenden, linearen Polymeren basiert, haben die vorhergehenden Betrachtungen weitreichende Folgen. Die Grundmechanismen solcher Polymere sind die Hybridisierung, kovalente Bindung und im Fall flexibler Polymere auch die Ausbildung von Strukturen. Wir wollen uns hier auf die ersten beiden Eigenschaften beschränken, da sie relevant für das in dieser Arbeit untersuchte System sind; vgl. Abschnitt 1.4.4

\* Eine Sequenz ist **selbstkomplementär**, wenn sie aus zwei gleich langen Teilen besteht, die komplementär zueinander sind. Eine solche Sequenz kann auf sich selbst vollständig hybridisieren bis auf die Stelle, an der die beiden Teilsequenzen aufeinander treffen; dort bildet sich ein Hairpin aus.

\*\* Als **Template** bzw. **Matrize** wird ein Einzelstrang bezeichnet, auf dem andere komplementäre (Teil-)Stränge hybridisieren können. Dadurch wird die Information über die Abfolge der hybridisierten Stränge vom Template auf diese übertragen [1]. Die kovalente Bindung der hybridisierten Stränge stellt eine (inverse) Kopie der Information des Templates dar.

und Kapitel 2 und 3. Diese können genutzt werden, um die kodierten Informationen weiterzugeben bzw. zu kopieren. Eine Kopie kann z. B. aus Monomeren erstellt werden, aber auch größere Polymereinheiten, die auf einem Template hybridisieren können, eignen sich dazu. Sie tragen zwar schon eine Teilinformation in sich, jedoch ist deren Anordnung auf einem Template im Allgemeinen nicht beliebig, so dass in der Anordnung ebenfalls Information gespeichert ist. Um diese Information zu kopieren, können diese Polymereinheiten kovalent gebunden werden. Einen solchen Fall möchten wir hier in Abhängigkeit von bestimmten Templates betrachten.

Ist das Template selbstkomplementär, so ist es wahrscheinlicher, dass das Template durch Faltung und Hybridisierung inaktiviert wird; vgl. Abb. 1.3.b,f. Wird dieses Template um ein zusätzliches Polymer erweitert, so ist im Allgemeinen die Selbstkomplementarität nur lokal und das Template kann sich durch Faltung in den selbstkomplementären Bereichen inaktivieren. Ersteres reduziert lediglich die Rate der Informationsweitergabe, letzteres kann jedoch Information vor dem Kopierprozess verschleiern. Eine unvollständige Kopie ist die Folge.

Im Fall von zwei unterschiedlichen, selbstkomplementären Templates, die aus den beiden gleichen komplementären Polymereinheiten aufgebaut sind, können sich auch Überhänge bilden; vgl. Abb. 1.3.e ( $X = \bar{A}$ ,  $Y = A$  und  $B = \emptyset = \bar{B}$ ). Diese Überhänge sind die Folge einer Indexverschiebung, so dass die kodierte Information nicht vollständig kopiert wird; vgl. Abb. 1.3.g ( $C = B$ ). Zudem kann über diesen Prozess auch neue Information hinzugefügt werden, wenn die Polymere C und  $\bar{C}$  länger sind und weitere Informationen tragen.

Diese beiden Szenarien sind für eine vollständige Informationsweitergabe hinderlich. Insbesondere in Zellen, bei der Zellteilung, werden die gesamten genetischen Informationen weitergegeben [1, 14]. Eine vollständige Informationsweitergabe wird erreicht, falls die Polymereinheiten vollständig und korrekt auf dem Template hybridisieren können und dann kovalent verbunden werden. Dazu müssen die benötigten Polymereinheiten vollständig vorhanden sein und das Template darf keine repetitiven oder selbstkomplementären Bereiche im Hinblick auf die Polymereinheiten aufweisen. Damit können die oben dargestellten hinderlichen Prozesse vermieden werden.

Im folgenden Abschnitt untersuchen wir die Möglichkeit, zwei Polymereinheiten auf einem Template kovalent zu binden. Die Experimente wurden durch von Kiedrowski beschrieben. Er benutzt zwar eine selbstkomplementäre Sequenz, die aber nur 6 Basen umfasst und daher praktisch keinen Hairpin ausbilden kann, da dieser nur 1–2 Basenpaare lang ist. Jedoch sind seine Überlegungen wichtig im Hinblick auf eine mögliche Implementierung einer autokatalytischen Reaktion, wie wir sie in Kapitel 2 und 3 verwenden werden.

### 1.3.3 Ein autokatalytischer Replikator

Von Kiedrowski hat das erste nicht-enzymatische, selbst-replizierende System, welches selbstkomplementäre Sequenzen nutzt und in dem sich Edukte und Produkte wieder voneinander lösen, in [49] vorgestellt. Damit ist es möglich, eine autokatalytische Reaktion in DNA durchzuführen und zu untersuchen. Die Sequenzen für das Template und die Edukte sind so gewählt, dass keine Verschiebungen auftreten können und die Reaktion nach der kovalenten Bindung der Edukte beendet ist.

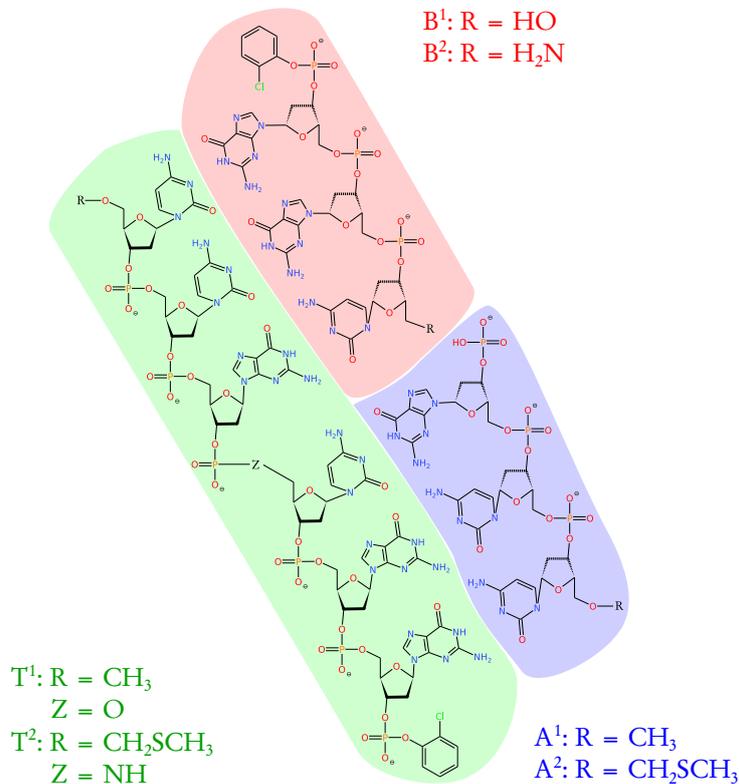


Abbildung 1.4: **Das selbst-replizierende Hexadeoxynucleotid von von Kiedrowski.** Das Hexadeoxynucleotid ist bis auf die Enden und Z identisch zum Aufbau der DNA und hat die Sequenz: GGC GCC (T, grün). Darauf können dann die Edukte A (blau, Sequenz: GCC) und B (rot, Sequenz: GGC) hybridisieren und schließlich kovalent binden. Diese Reaktion ist nur in Anwesenheit von EDC [N-(3-Dimethylaminopropyl)-N'-ethyl-carbodiimid] möglich. Die verschiedenen Versionen der Reste sind für unterschiedliche Experimente benutzt worden. In [50] wird eine Übersicht geboten und A<sup>2</sup>, B<sup>2</sup>, T<sup>2</sup> als Reaktanden sowie optional T<sup>1</sup> als Template genutzt. Abb. aus [50]; neu erstellt und angepasst an die Erklärung.

Von Kiedrowski hat zwei Varianten dieses Experiments konstruiert, z. B. [49, 50]. Die Abfolge der Basen ist bei den beiden Experimenten identisch, aber die Reaktanden sind über andere Moleküle verbunden und haben unterschiedliche, reaktive Reste; vgl. Abb. 1.4. Die Reaktion läuft in der Form:



ab. Dabei bezeichnen  $\alpha[T^*]^\zeta$  die autokatalytische Reaktionsrate mit der Konstanten  $\alpha$ , der Templatekonzentration  $[T^*]$  und der Ordnung  $\zeta$ , sowie  $\beta$  die spontane Bildungsrate für das Template  $T^*$ . Eine allgemeine Betrachtung dieses Reaktionstyps werden wir in Kapitel 1.4.1 durchführen.

Im Experiment [49] wurden die Edukte  $A^1$ ,  $B^1$  und  $T^1$  verwendet und es können die Produkte  $T^1$  und  ${}^5A^{3'} - {}^5A^{3'}$ , bei dem der A-Strang antiparallel gebunden ist, entstehen. Von Kiedrowski konnte nachweisen, dass das Produkt  $T^1$  autokatalytisch wächst, indem er die Templatekonzentration erhöhte. Dabei konnte er ebenfalls die Konstanten  $\alpha = 7,9 \cdot 10^{-4} \text{ M}^{-3/2} \text{ s}^{-1}$ ,  $\beta = 3,19 \cdot 10^{-5} \text{ M}^{-1} \text{ s}^{-1}$  und die Reaktionsordnung mit  $\zeta = 1/2$  bestimmen. Allerdings konnte die typische sigmoidale Kurve bei autokatalytischen Reaktionen wegen eines zu kleinen Verhältnisses der beiden Raten nicht reproduziert werden; vgl. Abschnitt 1.4.2.

In [50] konnte von Kiedrowski erstmals ein experimentelles System vorstellen, welches eine sigmoidale Wachstumskurve aufweist. Die Edukte sind  $A^2$ ,  $B^2$  und  $T^1$  als Template. Als Produkt entsteht  $T^2$ . Das Template  $T^1$  dient zur Aktivierung der autokatalytischen Reaktion; vgl. auch Abschnitt 2.2.4. Die Konstanten wurden zu  $\alpha = 7,43 \text{ M}^{-3/2} \text{ s}^{-1}$  und  $\beta = 1,77 \cdot 10^{-2} \text{ M}^{-1} \text{ s}^{-1}$  bestimmt, bei einer Reaktionsordnung von  $\zeta = 1/2$ . (Eine theoretische Analyse werden wir in Abschnitt 1.4.1 vorstellen.) Da das Verhältnis nun eine Größenordnung größer ist, kann das sigmoidale Verhalten gemessen werden.

Von Kiedrowski hat also herausgefunden, dass eine große autokatalytische Konstante für ein effizientes Wachstum zu dem gewünschten, sigmoidalen Verhalten führt und es bisweilen notwendig ist, eine autokatalytische Reaktion durch Zufügen des entsprechenden Templates zu aktivieren, um die Reaktionen schneller ablaufen zu lassen.

Welche Folgen autokatalytisches Wachstum im allgemeinen Fall haben kann, werden wir im Abschnitt 1.4.1 analysieren. Wir gehen nun von dem Wachstum eines einzelnen Individuums zu einem Modell einer kollektiven Dynamik über, das die Reaktionen zwischen solchen autokatalytisch reproduzierenden Einheiten beschreibt.

### 1.3.4 Entwicklung von Zusammenhängen von Spezies

Ein einzelnes Individuum ist nicht in der Lage, sich evolutionär zu verhalten; es wird eine kollektive Dynamik benötigt, die sicherstellt, dass sich mehrere Spezies stabil nebeneinander entwickeln können. Grundlegende und weitreichende Arbeiten wurden hierzu von M. Eigen und P. Schuster verfasst [19, 54, 55]. Dort werden mehrere mögliche Reaktionspfade im Hinblick auf eine evolutionäre Tauglichkeit untersucht. Als Lösung dieses Problems wird der Hyperzyklus vorgestellt. Ein Hyperzyklus ist ein (mehrfach) zyklisch verbundener Reaktionspfad aus autokatalytisch selbst-reproduzierenden Einheiten.

M. Eigen verfasste eine Abhandlung [53] über die Evolution biologischer Makromoleküle durch Selbstorganisation. In dem Artikel werden verschiedene theoretische und experimentelle Konzepte (u. a. der Hyperzyklus) vorgestellt, von denen wir hier ein für diese Arbeit

besonders relevantes hervorheben möchten: Den Entropieverlauf in einem Selektionsprozess. Ausgehend vom Prigogine-Glansdorff-Theorem [69] kann eine Verbindung zwischen Selektion und irreversiblen Prozessen hergestellt werden. Die Entropie in einem konstant durchflossenen System ist minimal und nimmt linear mit der Zeit zu. Wenn ein Mutant mit einem größeren Selektionswert als das vorher selektierte Exemplar durch eine stochastisch signifikante Fluktuation erzeugt wird, führt dies zu einem negativen Beitrag in der Entropie. Nach Prigogine und Glansdorff muss ein solches Ereignis zu einem Zusammenbruch des stationären Zustandes führen. Dies bedeutet, dass eine Spezies durch eine andere ausgetauscht wurde. Die linear anwachsende Entropie macht also einen „Knick“ und wächst letztlich parallel zur alten Kurve wieder genauso stark an. Der Versatz der beiden Kurven entspricht der Entropiedifferenz aus dem Zugewinn der internen Ordnung. Dies bedeutet auch, dass eine vormals selektierte Spezies im Zuge dieses Prozesses von einer anderen abgelöst worden ist. Die thermodynamische Beschreibung deckt aber im Gegenzug die internen, evolutionär relevanten Änderungen nur in der Form ab, dass sie stattgefunden haben; nicht jedoch, welche Änderungen mit welchen Auswirkungen sich vollzogen haben. Auch in dem hier vorgestellten System treten solche Umwälzungen auf, die sich in einer Verringerung der Entropie äußern; vgl. Abschnitt 3.2.4 und Abb. 3.6. Da sich das System nicht im stationären Zustand befindet, ist der Verlauf der Entropie nicht linear.

Wir gehen nun dazu über, die Hyperzyklen, wie sie von Eigen und Schuster [19, 54, 55] beschrieben wurden, darzustellen. Die Abhandlung ist in drei Teile gegliedert, von denen der erste die Entstehung behandelt, der zweite eine abstrakte Übersicht gibt und der dritte eine realistische Umsetzung beschreibt.

Der erste Teil [19] beschäftigt sich mit der Entstehung des Lebens, das sich, auch bis zum Auftreten der ersten Zelle, iterativ abgespielt haben muss, ähnlich der gegenwärtigen darwinischen Evolution. Der genetische Code ist demnach eine Einmal-und-für-immer-Selektion. Grundlegend zur Herstellung von Produkten ist ein Reaktionszyklus. Aus den Edukten und (mehreren) Zwischenschritten wird mit Hilfe eines Katalysators, z. B. eines Enzyms, die Reaktion vollzogen. Aus solchen Reaktionszyklen kann durch einen iterativen Prozess eine Hierarchie entstehen, die nacheinander die Produkte der verschiedenen Zyklen ineinander transformiert. Diese Hierarchie kann wiederum einen zyklischen Charakter aufweisen, wenn das finale Produkt wieder ein Edukt des ersten Zyklus darstellt. Dieser hat die Eigenschaft autokatalytisch zu sein. Jeder autokatalytische Reaktionszyklus ist ein Hyperzyklus der Ordnung 1, aber wird zur besseren Unterscheidung nicht so genannt und mit  $I_1$  bezeichnet. Ein zyklisches Reaktionsnetzwerk aus autokatalytischen Reaktionszyklen nennt man Hyperzyklus (der Ordnung 2); vgl. Abb. 1.5. Höhere Ordnungen werden entsprechend gebildet.

Grundvoraussetzung für ein darwinisches, selbst-selektives System ist, dass es gewisse Strukturen auf Kosten anderer dynamisch stabilisieren kann. Eigen und Schuster postulieren für ein molekulares, darwinisches System folgende Voraussetzungen an Metabolismus, Selbst-Replikation und Mutabilität: Im Gegensatz zu dem in dieser Arbeit vorgestellten System fordern sie, dass der Metabolismus, in Form von sowohl Bildung als auch Zersetzung, spontan und unabhängig voneinander sein muss. Diese Forderung ist in einem Gleichgewichtssystem notwendig, da sich ansonsten eine stabile Verteilung der bestehenden Konkurrenten einstellen würde. Selektion kann also nur wirken, wenn aus energiereichen Edukten Produkte und energieärmerer Abfall entstehen. Selbst-Replikation ist unerlässlich, so dass es zum einen zu keinem Informationsverlust kommt und zum anderen eine auftretende Mutation ebenfalls

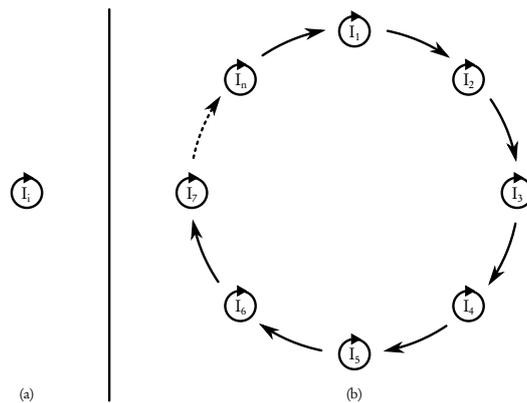


Abbildung 1.5: **Hyperzyklus**, abstrakte Darstellung. (a) Der Autokatalysator [Hyperzyklus 1. Ordnung] ist eine selbstreproduzierende Einheit. (b) Ein katalytischer Hyperzyklus (2. Ordnung) besteht aus einem Zyklus sich selbst antreibender Autokatalysatoren. Jeder Zyklus benötigt zur Reaktionsfortführung weitere Edukte und kann weitere Produkte abgeben (nicht eingezeichnet). Freie Darstellung nach [19].

die Möglichkeit hat, sich zu vermehren. Beim Prozess der Selbst-Replikation kommt es unweigerlich zu Fehlern, die sich in Mutationen äußern. Mutationen sind die Quelle für neue Information und können damit einen evolutionär wirksamen Vorteil aufweisen.

Eigen und Schuster betrachten ein durchflossenes System, in dem sich Individuen bilden und degradieren können. Die Bildung wird mit  $A_i Q_i x_i$  beschrieben, wobei  $Q_i$  die Qualität (Nicht-Mutation) der Kopie darstellt. Die Degradation wird mit  $D_i x_i$  beschrieben. Der Ein- bzw. Ausfluss wird mit  $\Phi_i$  bezeichnet. Es ergibt sich:

$$\dot{x}_i = (A_i Q_i - D_i) x_i + \sum_{k \neq i} w_{ik} x_k + \Phi_i. \quad (1.2)$$

Um die Massenerhaltung für die fehlerhaften Kopien einzuhalten gilt:

$$\sum_i (A_i Q_i - D_i) x_i = \sum_i \sum_{k \neq i} w_{ik} x_k. \quad (1.3)$$

Mit Hilfe der gemittelten Überschussproduktion  $E(t)$  ( $E_i \equiv A_i - D_i$ ) und Einführung einer Quasispezies  $y_i$  als Linearkombination der Individuen kann man schreiben<sup>\*</sup>:

$$\dot{y}_i = (\lambda_i - E(t)) y_i. \quad (1.4)$$

Jede Spezies mit  $\lambda_i < E(t)$  wird aussterben und, da  $E(t)$  mit der Auswahl von Spezies mit großen Eigenwerten zunimmt, wird  $E(t)$  ansteigen bis eine einzige Quasispezies ausgewählt ist; diese wiederum ist eine Verteilung von (unterschiedlichen) Individuen.

\* Die **Randbedingungen** für diese Zusammenfassung sind zum einen eine konstante Gesamtindividuumdichte, die durch einen anteilmäßigen Zu- bzw. Abfluss des Gesamtstroms geregelt wird. Zum anderen kompensiert der Gesamtstrom die Überschussproduktion. Wir definieren  $W_{ii} = (A_i Q_i - D_i)$ .  $\lambda_i$  sind dann die Eigenwerte des linearen Systems aus  $W_{ii}$  und  $w_{ik}$

Eigen und Schuster analysieren zunächst getrennt davon den Informationsgehalt der Quasispezies und stellen das daraus resultierende Stabilitätskriterium einer Spezies auf. Die Anzahl der Informationseinheiten (Nukleotide) ist beschränkt. Diese Grenze ist umgekehrt proportional zur mittleren Fehlerrate pro Informationseinheit.

In Teil II [54] beschreiben Eigen und Schuster verschiedene Szenarien gekoppelter, autokatalytischer Reaktionszyklen mit Hilfe einer Fixpunktanalyse. Dazu renormieren sie die entsprechenden Konzentrationen und bilden sie auf einen Simplex ab. Die hier vorgestellten Beispiele bestehen aus 3 unterschiedlichen, gleichartigen Individuen und werden daher auf einen 3-Simplex (Dreieck) abgebildet. Unverbundene Individuen aus Katalysatoren wachsen mit konstanter Rate und haben einen stabilen Fixpunkt im Inneren (Koexistenz). Autokatalytische Individuen nehmen mit einer linearen Rate zu und weisen, je nach Rate, einen instabilen, metastabilen oder stabilen (mit wachsender Rate) Fixpunkt bei den jeweiligen Einzelindividuen auf (reine Selektion). Drei katalytische Hyperzyklen prosperieren mit einer quadratischen Rate und sind bei jedem der Individuen stabil. Eine Quelle im Simplex treibt das System zu den Eckpunkten und folglich gibt es metastabile Fixpunkte, wenn ein katalytischer Hyperzyklus fehlt. Demnach wird, je nach Anfangsbedingungen, ein solcher Hyperzyklus selektiert. Es ist also nicht möglich, eine differenziertere Selektion darzustellen. Kopplungen selbstreplikativer Zyklen sind diesbezüglich ebenfalls nicht erfolgversprechend: Eine Kette von gekoppelten Autokatalysatoren ist nur unter geringen Konzentrationen und geordneten Raten, die eine Minimalkonzentration bestimmen, stabil<sup>\*</sup>. Verzweigungen sind ebenfalls nicht stabil und führen letztlich zu einem stationären Zweig. Ein Hyperzyklus jedoch weist neben den drei metastabilen Fixpunkten bei den jeweiligen Autokatalysatoren einen stabilen Fixpunkt im Simplex auf, welcher ein stabiler Strudel ist. Letztere stellt damit eine stabile Koexistenz mehrerer Spezies dar. Jedoch ist es für Hyperzyklen der Größe 5 oder größer möglich, einen stabilen Grenzzyklus zu erreichen.

Eigen und Schuster gehen dazu über, Netzwerke aus Hyperzyklen zu analysieren. Wenn einem Hyperzyklus ein externer Katalysator oder Autokatalysator als Parasit hinzugefügt wird, hängt es von den jeweiligen Raten ab, wer gewinnt. Die Rate zur nächsten Einheit des Hyperzyklus bezeichnen wir mit  $k_A$  und die zum Parasit mit  $k$ . Der externe Katalysator gewinnt, wenn das Verhältnis  $\frac{k}{k_A}$  größer als die Gesamtkonzentration ist. Andernfalls bleibt die Konzentration des Parasits gleich. Da dann der Hyperzyklus an Konzentration zunimmt, gewinnt dieser. Im Fall des externen Autokatalysators gewinnt derjenige Teil, der die größere Rate hat; der andere stirbt aus. Koppelt man zwei Hyperzyklen A und B über einen Wachstumsfaktor, dann treten sie bei einer linearen Ankopplung ( $\dot{c}_A = k_A c_A^2 c_B - \frac{c_A}{c_0} \Phi$ ) in Wettbewerb. Eine Kooperation ist mit einer quadratischen Ankopplung möglich ( $\dot{c}_A = k_A c_A^2 c_B^2 - \frac{c_A}{c_0} \Phi$ ).

Die Eigenschaften eines Hyperzyklus sind demnach die Koexistenz seiner einzelnen Teile und deren gemeinsames Wachstum, sowie eine adaptive Anpassungsfähigkeit im Falle von Verzweigungen im Zyklus. Spezies außerhalb des Hyperzyklus können verdrängt werden. Demnach ist der Hyperzyklus eine Einmal-und-für-immer-Selektion. Erweitert man die reaktiven Kopplungen innerhalb des Hyperzyklus um eine externe Katalyse, lassen sich stabile Oszillationen unter den beteiligten Spezies für größere ( $n > 5$ ) Hyperzyklen beobachten.

\* Das in dieser Arbeit vorgestellte System gehört zu dieser Klasse und ist stabil bei der größtmöglichen Spezies. Das Ziel der Analyse ist aber nicht der Endzustand, der auch für die natürliche Evolution unbekannt ist, sondern der Verlauf bis dieser erreicht ist.

In der Abhandlung [55] beschreiben Eigen und Schuster den realistischen Hyperzyklus. Die externe Katalyse der (Vorgänger der) DNA-Moleküle werden von (primitiven) Replikasen übernommen. Demnach entwickelten sich nach Eigen und Schuster aus den ersten Polynukleotiden GC-reiche Quasispezies, die die ersten Enzyme (Replikasen) codierten. Diese haben sich dann in Form von Hyperzyklen organisiert, die sich ebenfalls weiter entwickelten und auf einzelne Reaktionsräume aufgeteilt wurden. Daraus haben sich die ersten Vorläufer der heutigen Zellen entwickelt. Die Rolle des Hyperzyklus wird als einen (zur Organisation) notwendigen Schritt von den ersten biologisch aktiven Molekülen zur ersten Zelle dargestellt. Weitere Abläufe der Evolution beschreiben wir in Kapitel 1.5.

Wir wollen nun wieder einen Schritt von der abstrakten Organisation zurücktreten und betrachten einige zur Implementierung in DNA oder RNA notwendige Voraussetzungen für eine template-basierende Reaktion.

### 1.3.5 Ein aufkonzentrierender Reaktor im thermischen Gefälle

Heiße Quellen stellen eine Energie- und Nahrungsquelle im kalten Ozeanwasser dar, die z. B. von Archaea und Bakterien genutzt wird; vgl. Kap. 1.2.1 und [26]. Da heiße Quellen eine porenförmige Struktur aufweisen, gibt es Modelle, die die Entstehung des Lebens in Form einer RNA-Welt dort postulieren [27]. Die Poren könnten eine Funktionalisierung aufweisen, so dass bestimmte Reaktionsarten in bestimmten Poren ablaufen: Elementare Reaktionen finden auf der heißen Seite nahe des vulkanischen Felsens statt und auf deren Edukte aufbauende Reaktionen in kälteren Poren weiter im Wasser.

Der Aufbau von Strukturen, welcher notwendig für Leben ist, macht eine lokale Reduktion der Entropie notwendig [70, 58]. Diese kann aber nur im Nichtgleichgewicht auftreten; vgl. auch [22]. Zellen nutzen dazu einen komplexen Metabolismusapparat, um Nahrung durch die Zellmembran zielgerichtet in die Zelle zu transportieren. Auch werden unterschiedliche Funktionen in unterschiedlichen Regionen, z. B. Organellen, durchgeführt. So ist der Zellkern z. B. für die Aufbewahrung und Verarbeitung der Erbinformationen in Form von DNA zuständig und die Mitochondrien für den Stoffabbau und die Energiegewinnung [1], ähnlich zu der hier vorgestellten Separation in Form von Funktionalitäten und Energieflüssen.

Braun *et al.* [58, 59, 60] stellen eine solche vulkanische Pore im Labor nach und untersuchen die Reaktion von DNA in dieser unter einem thermischen Gradienten mit und ohne Durchfluss, um die möglichen Produkte zu analysieren bzw. eine Akkumulation der Produkte zu erreichen, so dass diese selektiert als Edukte einer potenziellen weiteren Reaktion in größerer Konzentration in beispielsweise einer anderen Pore zur Verfügung stehen. Dazu realisieren sie eine templatebasierte Reaktion mit DNA. Abgesehen von lebenden Zellen, in denen die Raten enzymatisch bei einer gegebenen Temperatur angepasst werden, erfolgt die Ablösung des Produkts vom Template und die anschließende Bindung der Edukte auf das Template typischerweise in einem thermisch oszillierenden Prozess. Dies ist ebenfalls für das Protokoll des Experiments in Kapitel 2 notwendig.

Wenn von einem anfänglichen Reservoir von Einheiten ausgegangen wird und sich diese im Laufe der Reaktionen zu größeren Einheiten zusammenlagern, wird folglich die Konzentration der Einheiten abnehmen. Dies kann man verhindern, indem man die größeren Einheiten

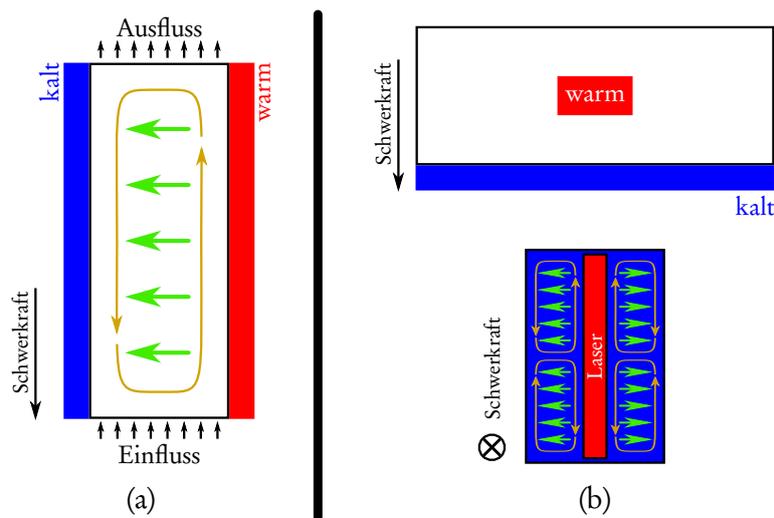


Abbildung 1.6: **Schematische Darstellung der Reaktoren.** Die Reaktoren bestehen aus einer Borsilikatkapillare, die mit Hilfe eines Peltierelements auf einer Seite gekühlt wird (blau dargestellt). Auf der gegenüberliegenden Seite wird mittels eines ohmschen Widerstandes (a) bzw. in der Mitte mit einem Laser (b) geheizt. Auf Grund des Temperaturunterschiedes bilden sich Konvektionsrollen (orange) aus und es kommt zu Thermoeffusion (grün). Damit werden längere DNA-Stränge, die sich im Reaktionsvolumen befinden, z. B. vornehmlich in der linken, unteren Ecke akkumuliert (a). Die Kapillare in (a) ist offen und durchflossen, während jene in (b) abgeschlossen ist. Abbildung (a) nach [59] und Abbildung (b) im Schnitt und der Draufsicht frei nach [58].

in einem kleineren Reaktionsvolumen anreichert. Eine vergleichbare Strategie werden wir im Abschnitt 3.5.5 anwenden, um die Limitierung der begrenzten Ausgangsmolekülzahl zu umgehen.

Die Anreicherung realisieren Braun *et al.* mittels zweier physikalischer Effekte: Konvektion und Thermoeffusion. Wie in Abbildung 1.6.a zu erkennen ist, wirkt die Thermoeffusion für DNA von einer heißen in Richtung einer kühleren Region [71]. Die Größe der Thermoeffusion im Gleichgewichtszustand wird durch den Soretkoeffizienten  $S_T$  beschrieben. Duhr und Braun finden, dass DNA schon von einer mit einem Laser um 1 K bis 2 K gegenüber Raumtemperatur erhitzten Region wegdiffundiert, was einen positiven Soreteneffekt darstellt. Dabei ist der Soretkoeffizient von der Länge der DNA abhängig:  $S_T \sim \sqrt{L}$ . D. h., dass längere DNA-Stränge einer größeren Thermoeffusion unterliegen. Demnach kann sich unter Thermoeffusion ein Längengradient entlang des Temperaturgradienten ausbilden und damit eine Längenselektion von DNA entlang des Temperaturgradienten erzielt werden.

Die Konvektion übernimmt in der Pore drei Aufgaben: Zum einen durchmischt sie das Reaktionsvolumen und zum anderen wirkt sie hauptsächlich senkrecht zum Temperaturgradienten und damit zur Thermoeffusion, so dass eine weitere Akkumulation entlang dieser Achse erfolgt, d. h. in Abb. 1.6.a werden sich die längeren DNA-Stränge in der linken unteren Ecke anreichern. Die Ausbildung von Konvektionsrollen ist stabil, wie Weggler *et al.* in [72] auch

für binäre Flüssigkeiten und große Rayleighzahlen feststellen konnten. Die Rayleighzahl ist proportional zum Temperaturgradienten und abhängig von physikalischen Fluideigenschaften und der Länge des Temperaturgradienten.

Zum dritten wird durch Konvektion die Suspension aus DNA umgewälzt und damit Temperaturänderungen ausgesetzt, die dafür sorgen, dass sich das Produkt vom Template lösen und neue Edukte darauf hybridisieren können. Damit wird das Temperaturprotokoll einer PCR\* nachgestellt.

Braun *et al.* realisieren die Pore durch eine Borsilikatkapillare, welche verschiedene Abmessungen für verschiedene Experimente hat: Der Querschnitt ist rechteckig mit 100 µm auf 50 µm beim abgeschlossenen Reaktor [58] bzw. 100 µm auf 2000 µm oder 70 µm auf 1400 µm beim durchflossenen Reaktor [59]. Der mit Paraffinöl abgeschlossene Reaktor wird auf der unteren Seite durch einen Siliziumwafer hindurch mit einem Peltierelement auf 24 °C gekühlt. Die Erhitzung erfolgt in der Mitte mit einem Laser auf 59 °C bis 86 °C [58]; vgl. Abb. 1.6.b. Der durchflossene Reaktor ist auf der heißen Seite mit Indiumzinnoxid beschichtet und zwischen zwei Saphirgläsern eingeschlossen, von denen die kalte Seite über einen Kupferblock mit einem Peltierelement auf 23 °C gekühlt wird. Die Erhitzung der heißen Seite auf 61 °C bis 94 °C erfolgt durch einen Strom durch das Indiumzinnoxid [59]; vgl. Abb. 1.6.a.

### Der abgeschlossene DNA-Reaktor

Die von Braun *et al.* eingesetzte DNA besteht zum einen aus einem 143 bp langen Template und zum anderen aus einem PCR-Reaktionsgemisch. Mit den oben genannten Temperaturen und Abmessungen bildet sich ein doppelter, zyklischer Konvektionsstrom mit einer mittleren Geschwindigkeit von 45 µm/s und damit mit einer Umlaufzeit von 80 s aus. Die längeren DNA-Stränge sammeln sich im unteren Bereich an den Rändern. Wegen der geringen Konzentration konnten nur bestimmte Längen (86 bp, 143 bp und 1530 bp) untersucht werden, die alle nachgewiesen werden konnten. Es entsteht also ein vermutlich vielfältiges Reaktionsgemisch, bei dem im Hinblick auf den Ausgangsstrang auch unvollständig replizierte Stränge gebildet werden.

### Der offene DNA-Reaktor

Kreysing *et al.* [59] untersuchen in einem vergleichbaren Experiment, ob es möglich ist, auch längere DNA-Stränge in einem durchflossenen Reaktor zu erzeugen; vgl. Abb. 1.6.a. Da sich kürzere Stränge schneller replizieren können, haben diese typischerweise einen Vorteil, der sich in deren Konzentrationszunahme äußert. Durch die Öffnung der Pore an deren beiden Enden und einem Zufluss eines DNA-Leitergemisches\*\* von 20 bp bis 200 bp in 20 bp Schritten, konnte gezeigt werden, dass es einen Bereich der Durchflussgeschwindigkeit gibt, der eine Trennung von kürzeren und längeren Strängen zulässt. Kürzere Stränge werden mit

---

\* Eine **Polymerasekettenreaktion (PCR)** ist ein Verfahren zur Vervielfältigung von beliebigen DNA-Strängen; vgl. [1]. Es werden sogenannte Primer, das sind kurze, zum Anfang des zu vervielfältigenden Stranges komplementäre DNA-Stränge, und einzelne Nukleotide zugefügt. Das Enzym Polymerase, üblicherweise aus dem Bakterium *Thermus aquaticus* und dann auch Taq-Polymerase genannt, verbindet ab dem Primer die Nukleotide zu einem komplementären DNA-Strang, so dass sich der Ausgangsstrang verdoppelt.

\*\* Eine **DNA-Leiter** ist ein Gemisch aus verschiedenen langen DNA-Strängen mit jeweils bekannten Konzentrationen. Diese wird üblicherweise dazu benutzt, um die Konzentration und die Länge unbekannter DNA durch Vergleich zu bestimmen [73].

dem Fluss aus dem System entfernt und längere sammeln sich auf der kalten Seite in der flussarmen Zone zwischen Rückfluss der Konvektion und dem Einlass, weil sie sich durch die Thermodiffusion in der kalten Region anreichern. Dies ermöglicht eine relativ genaue Trennung von akkumulierten und ausgewaschenen Strängen.

Der durch die Konvektion getriebene Temperaturwechsel für die akkumulierten DNA-Stränge ermöglicht auch hier eine PCR-ähnliche Reaktion. Es werden zwei unterschiedliche DNA-Stränge mit einer Länge von 36 bp und 75 bp im Zufluss zusammen mit dem PCR-Reaktionsgemisch zugefügt. Dies führt dazu, dass sich nur der längere Strang effizient replizieren kann und der kürzere ausgewaschen wird und schließlich nach ca. 6 h verschwindet.

#### Theorie eines RNA-Reaktors

Obermeyer *et al.* [60] untersuchen die oben genannte Versuchsanordnung theoretisch im Zusammenhang mit RNA. RNA speichert genetische Informationen und kann auch enzymatisch wirken, weil sie sich wie ein Enzym auf sich selbst faltet, existiert die Hypothese einer RNA-Welt [15]. Es ergibt sich die Frage, auf welchem Weg sich RNA zu RNA-Enzymen entwickelt hat. Obermeyer *et al.* analysieren deswegen die Strukturen der RNA, wie diese im RNA-Reaktor entstehen.

Im RNA-Reaktor-System können drei Arten von Reaktionen ablaufen. Zum einen können sich beliebige RNA-Moleküle ligieren, was einer spontanen Entstehung von neuen Molekülen entspricht und zum anderen können sich RNA-Moleküle falten und aufeinander hybridisieren, was zum Erhalten oder Kopieren von Information dient. Zum dritten können die Moleküle wieder zerkleinert werden, indem sie eher an einer Stelle, an der sie als Einzelstrang vorliegen, wieder gespalten werden. Insbesondere wird dadurch ein Selektionsdruck in Richtung hybridisierter Sequenzen erzeugt, was die aus informationstheoretischer Sicht relevante Eigenschaft der Informationsweitergabe eines RNA-Reaktors in diesem System implementiert und folglich kann auch die (Teil-)Sequenzweitergabe nachgewiesen werden.

Es werden ständig Monomere mit einer konstanten Rate hinzugefügt und Stränge mit der Rate  $\sim e^{-\sqrt{L/L_c}}$  entnommen, wobei  $L_c$  eine charakteristische Sequenzlänge ist, die durch die Randbedingungen\* bestimmt wird. Diese Rate bezieht die starke Akkumulation in einer Pore bzw. im Reaktor mit ein. Die Ligationsrate der Mono- bzw. Oligomere ist konstant und die Spaltungsrate an der Bindungsstelle K ist  $\beta_0 \cdot (1 - p_{L,K})$ . Dabei ist  $p_{L,K}$  die Basenpaarungswahrscheinlichkeit für zwei benachbarte Basen. Damit sind hybridisierte Basen stabiler. Es werden sowohl intra- als auch intermolekulare Basenpaarungen zugelassen, die mit Vienna [74] berechnet werden und temperaturabhängig für zwei verschiedene Temperaturen (10 °C und 60 °C) ausgewertet werden.

Obermeyer *et al.* untersuchen dann die Strukturen der entstehenden Moleküle, die jeweils ab einer bestimmten Länge für zwei verschiedene Längen (35 bp und 50 bp) auftreten. Dabei werden die Moleküle in fünf Klassen eingeteilt: Einmal in einen Hairpin, das ist ein größtenteils selbstkomplementärer Doppelstrang, der auf sich selbst hybridisiert und damit eine kurze Schleife bildet, und zum anderen in einen sog. Doppelhairpin, der aus zwei Hairpins besteht, die über einen kurzen Einzelstrang am Stamm (d. h. dem selbstkomplementären Doppelstrang) verbunden sind; zum dritten in einen sog. Hammerhead, was ein Doppelhairpin ist, der an

---

\* In die Randbedingungen gehen Parameter wie Soretkoeffizient, Temperaturgradient und Abmessungen [60] ein.

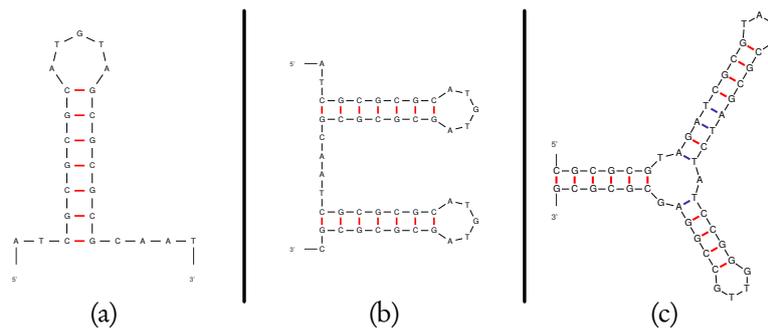


Abbildung 1.7: **Faltungen im stationären Zustand.** Obermeyer *et al.* können im stationären Zustand drei grundsätzliche Faltungen identifizieren; vgl. [60]. (a) Hairpin: Ein größtenteils selbstkomplementärer Doppelstrang, der auf sich selbst hybridisiert und eine Schleife bildet. (b) Doppelhairpin: Zwei über einen (kurzen) Einzelstrang an deren Enden verbundene Hairpins. (c) Hammerhead: Ein Doppelhairpin, der an den freien Enden wiederum auf sich selbst hybridisiert. Abbildung nach [60]. Strukturen berechnet mit [75, 76]. Die Sequenzen sind frei gewählt.

den freien Enden wiederum einen Doppelstrang bildet und damit eine Schleife. Diese drei Strukturen können ebenfalls wenige, nicht hybridisierende Basen aufweisen; vgl. Abb. 1.7. Des Weiteren wird zwischen anderen Faltungen und Einzelsträngen unterschieden.

Die Länge ist exponentiell abnehmend verteilt und temperaturabhängig. Bei niedrigeren Temperaturen können die Stränge besser hybridisieren und es treten längere Stränge auf. Dann ist auch die Faltungswahrscheinlichkeit, insbesondere für mittlere Längen, größer. Dementsprechend sieht auch die Verteilung der Strukturen aus: Bei großer Temperatur kommen insgesamt weniger Strukturen vor und es können nur Hairpins und wenige Doppelhairpins beobachtet werden. Bei niedriger Temperatur und kurzer Länge sind die Vorkommenswahrscheinlichkeiten in der Reihenfolge Hairpin, Doppelhairpin, Hammerhead und andere abnehmend. Insgesamt lässt sich in diesem Fall für diese Reihenfolge eine Zunahme für größere Längen beobachten, da diese längere Stränge zur Ausbildung benötigen.

Diese Beobachtungen gehen einher mit der beobachteten Informationsweitergabe; es werden nur auf relativ kurzen Sequenzabschnitten Korrelationen beobachtet. Schon ab Korrelationslängen größer als 3 kommen nicht mehr alle Sequenzmotive vollständig vor und es findet eine Selektion statt. Ausgehend von der Fließgleichgewichtslängenverteilung können Obermeyer *et al.* effektive Raten zur Erzeugung ( $k_+$ ) und Vernichtung ( $k_-$ ), welche Spaltung und Hybridisierung umfassen, bestimmen. Aus diesen konnten sie eine Differentialgleichung, die einem Replikationsprozess gleicht, für die mittlere Konzentration  $\langle n \rangle$  aufstellen:  $\partial_t \langle n \rangle = k_+ - k_- \langle n \rangle + r(\bar{n})$ . Dabei ist  $r$  die Replikationsrate und Motive werden mit der Rate  $k_+$  erzeugt und mit der Rate  $k_-$  zerstört.  $\bar{n}$  ist das Komplement von  $n$ . Die Replikationseffizienz  $r/k_-$  ist abhängig von der Spaltungsrate  $\beta_0$ . Sequenzen werden also spontan erzeugt, proportional zu ihrem Vorhandensein degradiert und proportional zu ihrem Komplement kopiert.

## 1.4 Analyse der Voraussetzungen für ein System für chemische Evolution und dessen Konstruktion

Braun *et al.* können experimentell und theoretisch die Situation in einer Pore eines vulkanisch aktiven Felsens nachstellen und dort die Bildung und Aufkonzentrierung längerer Stränge am Boden des Reaktionsvolumens nachweisen. Theoretisch konnte auch die Informationsselektionsdynamik untersucht werden und das Entstehen, Kopieren und Degradieren von Information festgestellt werden.

### Zusammenfassung

Wir haben nun einen Überblick über die Grundlagen der Materialien und Methoden zur Konstruktion eines chemischen evolutionären Systems gegeben. Der Informationstransport und die Informationsverarbeitung eines solchen Systems wurde am Beispiel der DNA vorgestellt. Es gibt aber auch konstruierte Informationsträger, die diese Aufgabe ebenfalls erfüllen können [77]. Betrachtet man lineare, kodierende, antiparallele Polymere, wie DNA, so gibt es mehrere Besonderheiten und Einschränkungen zu beachten, die allesamt auf deren Strukturbildung zurückzuführen sind. Damit kann man dann versuchen, einen Replikator wie den von von Kiedrowski [49, 50, 51] zu konstruieren, welcher auch die Grundlage der Betrachtungen des folgenden Abschnitts ist. Eigen und Schuster [19, 54, 55] konnten das kollektive Verhalten von autokatalytisch agierenden Reaktionen analysieren. Braun *et al.* [58, 59, 60] konnten theoretisch und experimentell die Aufkonzentrierung in einem Reaktionsvolumen mit thermischem Gefälle nachweisen. Dies ist für den Erhalt der Reaktion von entscheidender Bedeutung. Wir werden uns nun dem Wachstumsverhalten einer autokatalytischen Reaktion zuwenden und dies genauer analysieren.

## 1.4 Analyse der Voraussetzungen für ein System für chemische Evolution und dessen Konstruktion

In diesem Abschnitt untersuchen wir ein autokatalytisches Minimalsystem zur Demonstration der grundlegenden Wachstumsverhalten und deren Abhängigkeiten. Im Speziellen steht dabei das grundlegende Wachstumsverhalten von Dimeren bei einem anfänglichen Vorrat an Monomeren im Hinblick auf die Konstruktion eines chemischen evolutionären Systems im Fokus. Weitere Reaktionen laufen nicht ab.

### 1.4.1 Analysen von von Kiedrowski

Von Kiedrowski betrachtet in [78] ein minimales Replikationssystem, bei dem aus einem anfänglichen Vorrat an Monomeren autokatalytisch Dimere entstehen. Dies ist derzeit die Grundlage für alle künstlichen Selbstreplikationssysteme. Dabei wird ein lineares, selbstkomplementäres Molekül verwendet, welches als Template (Matrize) fungiert und damit autokatalytisch reagiert; vgl. Abb. 1.8. Dabei wird die Information des Moleküls  $\overline{AA}$ , d. h. die Abfolge der verbundenen Einzelmoleküle entlang des linearen Moleküls  $\overline{AA}$ , als Vorlage für die reversible schwache Bindung der Edukte ( $A$  und  $\overline{A}$ ) genutzt, die ihrerseits dann untereinander irreversibel kovalent binden ( $A|\overline{A}-\overline{AA} \rightarrow \overline{AA}-\overline{AA}$ ). Diese Reaktionsstruktur findet sich in der DNA wieder, welche demzufolge als Grundlage für künstliche selbstreproduzierende Systeme dient. Für solche Systeme werden häufig die Nukleinbasen Cytosin und Guanin

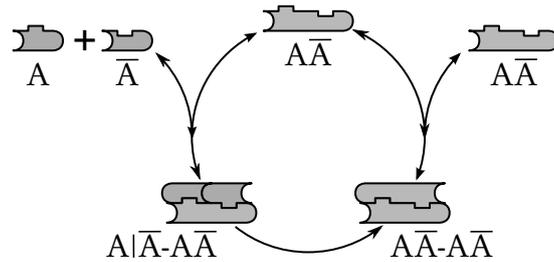


Abbildung 1.8: **Schema einer minimalen autokatalytischen Reaktion.** Die Reaktion wird im Fall der autokatalytischen Reproduktion links herum durchlaufen. Das Template (Matrize)  $\overline{AA}$  dient als Vorlage und die Edukte  $A$  und  $\overline{A}$  binden reversibel schwach darauf zum Produkt  $A|\overline{A-AA}$ . Dann reagiert  $A|\overline{A-AA}$  irreversibel kovalent zu  $\overline{AA-AA}$ , welches dann in zwei  $\overline{AA}$  dissoziiert. Nicht kovalente Reaktionen werden als reversibel angenommen. Frei nach [78].

verwendet, da sie eine größere Hybridisierungsenergie aufweisen; vgl. Abschnitt 1.3.1. Die DNA kann effizient allerdings nur enzymatisch kovalent gebunden werden\*, so dass eine von selbst ablaufende Reaktion durch von Kiedrowski dadurch realisiert wird, dass er die reaktiven Enden modifiziert (siehe Kapitel 1.3.3 und Abb. 1.4). Orgel [52] verwendet dagegen die mit der DNA chemisch verwandte RNA mit modifizierten Nukleotidbindungen, reaktiven Enden sowie einer Templatelänge von 4 Basen. In beiden Systemen sind zudem die linearen Moleküle selbstkomplementär, d. h. dass Produkt und Template strukturell identisch sind und damit das Produkt wieder als Template fungieren kann.

Um ein solches System theoretisch analysieren zu können, wird es auf zwei Reaktionsmöglichkeiten reduziert: die spontane und die autokatalytische Bildung des Produkts [80, 78] (und Kapitel 3.1). Da die Zwischenschritte durch reversible Reaktionen abgebildet werden und damit sich selbst equilibrieren, können die Teilraten zu einer Rate zusammengefasst werden. Damit lässt sich die Reaktion verkürzt schreiben als:



wobei  $A_1$  ein Monomer und  $A_2$  ein kovalent gebundenes Dimer bezeichnet. Die spontane Rate ist  $\beta$  und die autokatalytische  $\alpha[A_2]^\zeta$ , wobei  $\alpha$  die Stärke der Autokatalyse angibt. Von Kiedrowski findet in [49] (siehe auch Abschnitt 1.3.3) einen Wert von  $\zeta = \frac{1}{2}$ . Für ein autokatalytisches Wachstum liegt  $\zeta$  typischerweise im Bereich von  $\frac{1}{2}$  und 1. Man kann diese

\* Eine nicht enzymatische Bindung von Pyrimidin-DNA-Oligomeren ist durch Katalyse eines komplementären DNA-Doppelstrangs auch nicht enzymatisch möglich [79]. Zwei Pyrimidin-DNA-Oligomere können parallel an eine DNA-Doppelhelix binden und mit chemischen Wirkstoffen kovalent verbunden werden [79].

#### 1.4 Analyse der Voraussetzungen für ein System für chemische Evolution und dessen Konstruktion

Reaktion dann analytisch als Differentialgleichung darstellen:\*

$$\frac{dA_2(t)}{dt} = (A_1^0 - 2A_2(t))^2 \left[ \alpha (A_2^0 + A_2(t))^\zeta + \beta \right] \quad (1.6)$$

$A_2(t)$  bezeichnet die Konzentration von  $A_2$  im Zeitverlauf und  $A_x^0$  die Anfangskonzentration von  $A_x$ . Von Kiedrowski nutzt diese Gleichung, um das System zu analysieren. Da wir hier eine weitergehende Analyse vornehmen, stellen wir alle Graphen als direkte Simulation von Gl. 1.5 mit dem in Kapitel 3 diskutierten Programm dar.

Für den Reaktionsanfang kann man schreiben:  $\frac{dA_2(t)}{dt} = \alpha A_2(t)^\zeta$ . Nach Separation der Variablen und Integration erhalten wir für  $0 \leq \zeta < 1$ :

$$A_2(t) = \sqrt[1-\zeta]{(A_2^0)^{1-\zeta} + (1-\zeta)\alpha t} \underset{\zeta=\frac{1}{2}}{=} \left( \sqrt{(A_2^0)} + \frac{1}{2}\alpha t \right)^2 \quad (1.7)$$

und für  $\zeta = 1$  erhält man:

$$A_2(t) = A_2^0 e^{\alpha t}. \quad (1.8)$$

Dies ist ein exponentielles Wachstum. Gl. 1.7 beschreibt für den Fall  $\zeta = \frac{1}{2}$  ein parabolisches Wachstum. (Für  $\zeta = 0$  steigt  $A_2(t)$  linear mit Steigung  $\alpha$  an.)

Im Folgenden wollen wir die Unterschiede zwischen den beiden Wachstumsarten darstellen, so wie von Kiedrowski diese beobachtet. Betrachtet man in Abb. 1.9.a-b jeweils eine Kurve mit vergleichbarem Parameter, so stellt man fest, dass ein exponentielles Wachstum schneller abläuft als ein parabolisches. Da von Kiedrowski nicht den gesamten Parameterraum betrachtet, kann er eine Grenzkurve für  $\beta \rightarrow 0$  nur für parabolisches Wachstum feststellen, jedoch nicht für exponentielles. Eine genauere Analyse findet sich im folgenden Abschnitt.

Von Kiedrowski konnte eine Abschätzung bezüglich des Wachstumstyps\*\* für die in Abb. 1.8 gezeigte Reaktion herleiten. Demnach ist:

$$\zeta = \frac{4K_2cq^2}{8K_2cq^2 + (1+q)^2 - (1+q)\sqrt{8K_2cq^2 + (1+q)^2}}, \quad (1.9)$$

wobei  $K_1 = \frac{[A][\bar{A}-A\bar{A}]}{[A][\bar{A}][A\bar{A}]}$ ,  $K_2 = \frac{[A\bar{A}-A\bar{A}]}{[A\bar{A}]^2}$ ,  $q = \left[ K_1 ([A] + [A|\bar{A} - A\bar{A}]) ([\bar{A}] + [A|\bar{A} - A\bar{A}]) \right]^{-1}$  und  $c = [A\bar{A}] + [A|\bar{A} - A\bar{A}] + 2[A\bar{A} - A\bar{A}]$  ist. Diese Konstanten hängen nur von den Gleichgewichtskonstanten und den Gleichgewichtskonzentrationen  $[X]$  ab. In dieser Beschreibung wurden insbesondere eine schwache Bindung der Edukte mit sich selbst oder untereinander sowie strukturelle Bindungen wie Hairpinbildung vernachlässigt. Man erkennt sofort, dass der Grenzwert für  $K_2 \rightarrow \infty$  unabhängig von  $q$  bzw.  $K_1$  gegen  $\frac{1}{2}$  strebt. Ein großes  $K_2$  bedeutet,

\* Für ein reines **spontanes Wachstum** würde sich die Differentialgleichung  $\frac{dA_2(t)}{dt} = \beta \cdot (A_1^0 - 2A_2(t))^2$  ergeben. Diese hat die Lösung:  $A_2(t) = \frac{1}{2} \left[ A_1^0 - \frac{1}{2} \cdot \left( \beta t + \frac{1}{2A_1^0} \right)^{-1} \right]$ , welche eine Hyperbel darstellt, die in den Wert  $\lim_{t \rightarrow \infty} A_2(t) = \frac{1}{2} A_1^0$  sättigt.

\*\* Mit **Wachstumstypen** ist z. B. logarithmisches, lineares, parabolisches, exponentielles oder hyperbolisches Wachstum gemeint. Hier ist im Speziellen die Unterscheidung von parabolischem und exponentiellem Wachstum gemeint.

dass ein großer Teil des Produktes schwach auf dem Template gebunden bleibt und nicht dissoziiert. Damit stehen sowohl Produkt als auch Template für eine weitere Reaktion nicht zur Verfügung; dies nennt man Produktvergiftung. Für experimentell relevante Parameter gibt es für Moleküle, die mehrere Nukleotide umfassen, aber einen Bereich, der  $\zeta$  gegen 1 streben lässt. Dies ist insbesondere dann möglich, wenn Produktvergiftung vermieden wird. Für eine Abschätzung dieser Parameter sei auf [78] verwiesen.

Szathmary [81, 78] findet jedoch, dass exponentielles Wachstum unablässig für darwinische Selektion ist, falls die Gesamtkonzentration der konkurrierenden Templates konstant ist. Exponentielles Wachstum ist daher eine *condicio sine qua non*, um einen darwinischen Selektionsprozess in einem künstlichen System zu realisieren [78].

Szathmary [81] betrachtet zwei, sich exponentiell vermehrende Populationen  $P_1$  und  $P_2$ . Diese werden in einem Durchflussreaktor gezüchtet, in den genügend Nahrung zugeführt und Abfallstoffe sowie überschüssige Individuen mit einer Rate  $\Phi$  abgeführt werden. Die Wachstumsraten seien  $k_1$  respektive  $k_2$ , die Gesamtkonzentration  $P = P_1 + P_2$  und damit ergibt sich der Ausfluss zu  $\Phi = k_1 P_1 + k_2 P_2$ . Dann ergeben sich die Konzentrationsänderungen zu:

$$\dot{P}_i = k_i P_i - P_i \frac{\Phi}{P} \quad (1.10)$$

für  $i \in \{1;2\}$ . Aus Gleichgewichtszustand ( $\dot{P}_i \equiv 0$ ) lassen sich 3 Fixpunkte bestimmen:  $k_1 = k_2$ ,  $P_1 = 0$  und  $P_2 = 0$ . Im ersten Fall ist der Fixpunkt neutral, die beiden Populationen sind hinsichtlich ihres Wachstums identisch und die Konzentrationen konstant. Für  $k_1 > k_2$  ist  $P_2 = 0$  stabil und  $P_1 = 0$  instabil und umgekehrt. Für o. B. d. A.  $P_2 = 0$  können wir mit  $\Phi$  und  $P$  aus Gleichung 1.10 zwei entkoppelte Differentialgleichungen herleiten:

$$\dot{P}_1 = (k_1 - k_2)P_1 - (k_1 - k_2) \frac{P_1^2}{P} \quad (1.11)$$

$$\dot{P}_2 = (k_2 - k_1)P_2 - (k_2 - k_1) \frac{P_2^2}{P}. \quad (1.12)$$

Wenn  $P_1$  klein ist, dann überwiegt der lineare Teil der Gleichungen 1.11 und 1.12. Damit wächst  $P_1$  und  $P_2$  nimmt ab. Damit setzt sich die Population  $P_1$  durch, während  $P_2$  auf lange Sicht verschwindet. Die beiden Fixpunkte sind also Attraktoren für die beiden Fälle, dass eine Population ausstirbt und die andere überlebt.

Obwohl exponentielles Wachstum überlegen ist, wollen wir dennoch im folgenden Abschnitt exponentielles mit parabolischem Wachstum vergleichen.

## 1.4.2 Vergleich von exponentiellem mit parabolischem Wachstum

Um den Einfluss der verschiedenen Wachstumstypen zu untersuchen, vergleichen wir hier exponentielles mit parabolischem Wachstum, wie es teilweise zuvor schon von Kiedrowski durchgeführt hat [78]. Da hyperbolisches Wachstum ein intrinsisch anderes Verhalten zeigt, werden wir es im folgenden Abschnitt analysieren. Ziel dieser beiden Abschnitte ist es, die notwendigen Grundlagen zu schaffen, um ein System für Evolution zu konstruieren; vgl. Abschnitt 1.4.4, und die Gemeinsamkeiten der verschiedenen Wachstumstypen diesbezüglich zu erläutern.

#### 1.4 Analyse der Voraussetzungen für ein System für chemische Evolution und dessen Konstruktion

Exponentielles Wachstum zeichnet sich per Definition dadurch aus, dass die Wachstumsrate pro Einheitsdichte konstant ist; vgl. [81]:  $\frac{\dot{A}_2}{A_2} = \alpha$ . Damit verschwindet auch die 2. Ableitung  $\frac{d^2}{dt^2} [\ln(A_2)] = 0$ . Im Gegensatz dazu nimmt das parabolische Wachstum mit der Wurzel der Einheitsdichte ab; vgl. [81]:  $\frac{\dot{A}_2}{A_2} = \alpha \cdot \frac{1}{\sqrt{A_2}}$ . Damit ist klar, dass auf lange Zeiten exponentielles dem parabolischen Wachstum überlegen ist.

Parabolisches Wachstum findet man häufig bei templatebasierenden Prozessen, siehe [81]: Vereinfacht betrachten wir das Gleichgewicht zwischen hybridisiertem (H) und dissoziiertem (D) Template und den entsprechenden Reaktionskonstanten  $k_h$  und  $k_d$ , so dass sich die Reaktion zu:  $H \xrightleftharpoons[k_d]{k_h} 2D$  ergibt. Der Gleichgewichtszustand ist dann  $k_d H = k_h D^2$  und bei großen Konzentrationen kann die Gesamtkonzentration  $G$  als  $G = D + 2H \doteq 2H$  genähert werden, da die Hybridisierung schneller erfolgt als die Dissoziation ( $k_d \ll k_h$ ). Dann ist der Gleichgewichtszustand näherungsweise:  $D = \sqrt{\frac{k_d}{k_h}} \cdot \sqrt{H} \doteq \sqrt{\frac{k_d}{2k_h}} \cdot \sqrt{G}$ . Es ist ersichtlich, dass nur freie Stränge  $D$  als autokatalytisches Template dienen können und daher die Abhängigkeit von der Gesamtkonzentration in die Ratengleichung mit  $\zeta = 1/2$  eingehen muss.

Wir betrachten nun die Differentialgleichung  $\dot{A}_2 = \alpha \cdot A_2^\zeta + \beta$  wie auch von Kiedrowski in [78]. Die Wachstumstypen ergeben sich in Abhängigkeit von  $\zeta$ : exponentiell ( $\zeta = 1$ ), parabolisch ( $\zeta = 1/2$ ) und hyperbolisch ( $\zeta = 2$ ). Dazu betrachten wir wie von Kiedrowski Kurvenscharen ähnlich zu denen in Abb. 1.9.a-b. Wir werden ebenfalls den Reaktionsbeginn, d. h. die Bildung des ersten Moleküls, auf den Anfang, in unserem Fall willkürlich auf  $t = 10^{-22} > 0$ , verschieben, um den Reaktionsbeginn, der für alle Wachstumstypen identisch ist, gesondert zu untersuchen. Die Analyse des Wachstums für verschiedene Parameter ermöglicht es uns, Aussagen zu treffen, die spezifisch für diese Parameter sind. Des Weiteren sind die hier vorgestellten Analysen umfassender, da zum einen der untersuchte Parameterraum größer gewählt wurde und zum anderen die logarithmische Struktur der Abhängigkeit von den Wachstumsparametern dargestellt wird. Der Unterschied zur linearen Darstellung wird an entsprechenden Stellen erläutert.

Zunächst betrachten wir das Wachstum für einen festen autokatalytischen Parameter  $\alpha$  und verändern die spontane Bildungsrate  $\beta$ . Wir stellen fest, dass es sowohl für exponentielles als auch für parabolisches Wachstum für sehr kleine  $\beta$  eine Grenzkurve gibt, die im Wesentlichen ein freies, autokatalytisches Wachstum nach einem spontan gebildeten Dimer  $A_2$  beschreibt; vgl. Abb. 1.9.a-b. In linearer Zeit gibt es auch eine Grenzkurve in beiden Fällen für große  $\beta$ , die jedoch für logarithmische Zeiten verschwindet, da ein größeres  $\beta$  zu einer schnelleren spontanen Bildung von weiteren Molekülen führt. Kleine  $\beta$  zeigen die typischen sigmoidalen Wachstumskurven, die man bei autokatalytischem Wachstum bei begrenzten Ressourcen beobachten kann [78]. Für größere  $\beta$  wird die Konzentration anfänglich immer steiler, so dass eine Kurvenschar in Abhängigkeit von wachsendem  $\beta$  entsteht, die letztlich in einen sehr steilen Anstieg mündet. Wegen der schwächeren Autokatalyse im Fall des parabolischen Wachstums erfolgt dort die Aufspaltung schneller.

Da  $\beta$  daher im Wesentlichen die Zeitskala definiert, macht eine solche Betrachtung, zumal sie damit auch indirekt  $\alpha$  skaliert, im Hinblick auf die Charakterisierung von Wachstumsverhalten weniger Sinn. Aus Abb. 1.12.c geht hervor, wie sich die Zeit in Abhängigkeit von  $\beta$  bei korrekter Skalierung ändert:  $t \sim \beta^{-1}$ . Daher wenden wir uns nun den Abhängigkeiten von  $\alpha$  und der Anfangskonzentration zu und definieren die Zeitskala durch  $\beta = 1$ .

Die Dimerkonzentration in Abhängigkeit von einer logarithmischen Zeitskalierung zeigt einen sigmoidalen Verlauf, der für kleine  $\alpha$  zu äquidistanten Kurven führt. Für kleine  $\alpha$  gibt es eine Grenzkurve; vgl. Abb. 1.9.c-d. Die Grenzkurven sind für  $\zeta \leq 1$  und kleine  $\alpha$  identisch, da diese hauptsächlich durch die spontane Reaktion zu Stande kommen. Für große  $\alpha$  wird diese vernachlässigbar, so dass die Abhängigkeit der Halbwertszeit\* von  $\alpha$  zweigeteilt ist; vgl. Abb. 1.12.a. Für große  $\alpha$  (für  $\alpha \geq 10^1$  respektive näherungsweise  $\alpha \geq 10^{-2}$  bei  $[A_1]_0 = 2^{17}$ ) fällt die spontane Reaktion nicht mehr ins Gewicht und  $\alpha$  bestimmt im Wesentlichen die Zeit. Daher sind die beiden Exponenten mit  $b \approx 1$  vergleichbar; vgl. Abb. 1.12.a. Für kleine Zeiten (für  $\alpha \leq 10^{-6}$  respektive näherungsweise  $\alpha \leq 10^{-2}$  bei  $[A_1]_0 = 2^{17}$ ) ist die Halbwertszeit nahezu konstant. Die Sättigung kommt zu Stande, da die spontane Reaktion signifikant wird und damit die Reaktion unabhängig vom autokatalytischen Anteil ist. Zwischen den beiden Abschnitten ist der Übergang stetig. Im Allgemeinen ist die Halbwertszeit für kleinere  $\zeta$  größer.

Die Abhängigkeit der normierten Dimerzeitentwicklung von der Anfangskonzentration ist in beiden Fällen durch in der Zeit verschobene sigmoidale Kurven gekennzeichnet; vgl. Abb. 1.10.a-b. Der Abstand zwischen den Kurven ist kleiner für kleinere  $\zeta$ . Damit hängt die Halbwertszeit auch mit einem Potenzgesetz von der Anfangskonzentration ab und ist für kleinere  $\zeta$  ebenfalls größer und fällt langsamer ab; vgl. Abb. 1.12.b.

Die Zeit bis zum Reaktionsbeginn ist nur abhängig von  $\beta$  und der Anfangskonzentration, da der autokatalytische Term wegen des nicht vorhandenen Templates verschwindet. Dabei ist diese Zeit  $t_A$  proportional zur reziproken Anfangskonzentration  $[A_1]_0^{-1}$  und zur quadratisch reziproken spontanen Rate  $\beta^{-2}$ , also  $t_A \sim [A_1]_0^{-1} \cdot \beta^{-2}$ ; vgl. Abb. 1.11.

### 1.4.3 Hyperbolisches Wachstum

Im Gegensatz zum parabolischen oder exponentiellen Wachstum, nimmt das hyperbolische Wachstum ( $\zeta = 2$ ) eine Sonderstellung ein: Es hat ein Ende in endlicher Zeit [81]. Für  $\beta = 0$  ergibt sich die Lösung der Differentialgleichung zu (vgl. Gl. 1.7):

$$A_2(t) = \frac{1}{\frac{1}{A_2^0} - \alpha t} \quad (1.13)$$

Dieser Ausdruck ist eine Hyperbel und wächst, bei zur Genüge vorhandenen Edukten, ins Unendliche für  $t \rightarrow (\alpha \cdot A_2^0)^{-1}$ . Dieser Zeitpunkt tritt früher für große  $\alpha \cdot A_2^0$  ein [81]. Da die Wachstumsrate pro Einheitsdichte mit der Dichte wächst:  $\frac{\dot{A}_2}{A_2} = \alpha \cdot A_2$ , ist das Charakteristikum dieser Wachstumsart, dass das Häufigste überlebt, sei es, weil es zunächst häufiger vorkommt oder schneller wachsen kann. Selektion ist demnach eine Funktion der Anfangsbedingungen [81]. Folglich ist die Sigmoidalkurve bei Ressourcenbegrenzung sehr steil und näherungsweise (bis auf Anfang und Ende) eine senkrechte Gerade.

\* Die **Halbwertszeit** meint hier die Zeit, die benötigt wird, um die Hälfte der Monomere zu verbrauchen, womit auch die Hälfte der Dimere erzeugt worden ist.

## 1.4 Analyse der Voraussetzungen für ein System für chemische Evolution und dessen Konstruktion

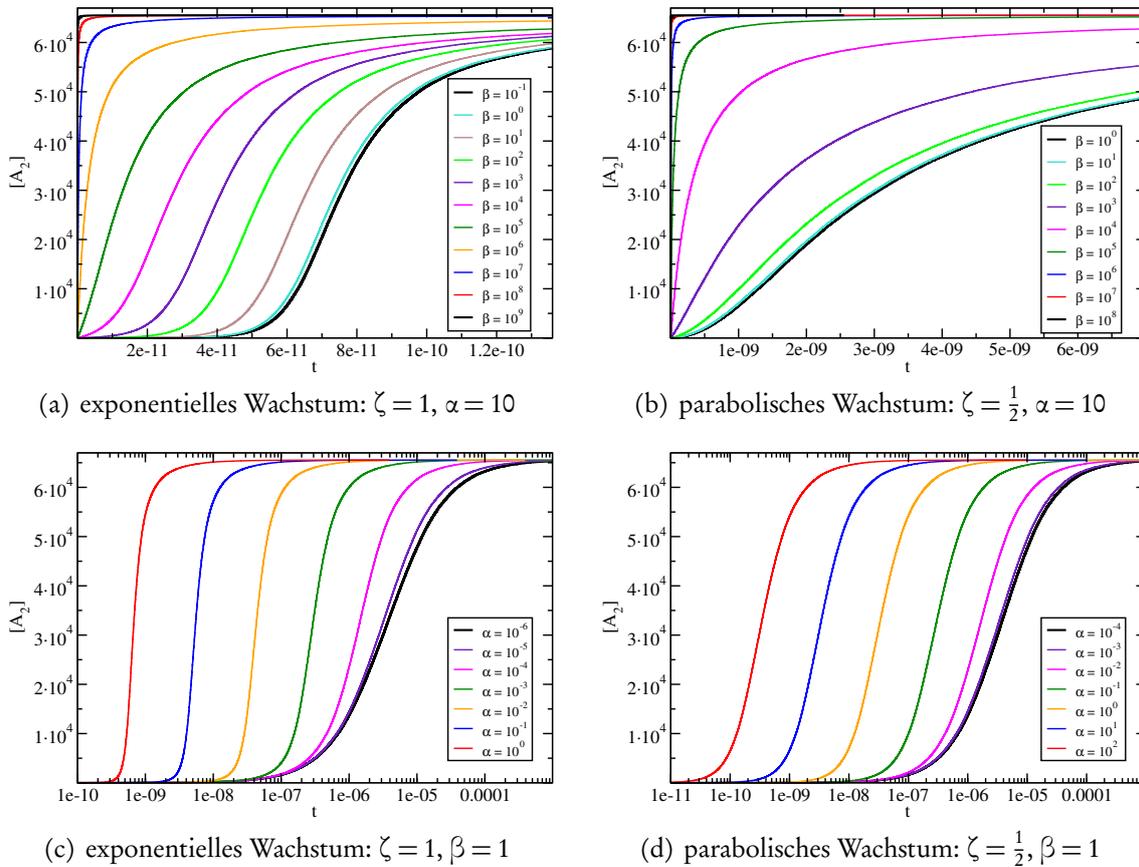


Abbildung 1.9: **Wachstumsverhalten von Dimeren.** Konzentration der Dimere bei der Reaktion nach Gl. 1.5 im Zeitverlauf für verschiedene Wachstumsarten. Abbildungen (a, c) zeigen exponentielles Wachstum ( $\zeta = 1$ ), welches einen schnelleren Anstieg zeigt als parabolisches Wachstum ( $\zeta = \frac{1}{2}$ ) in Abb. (b, d). Beide Wachstumsarten zeigen bei Veränderung von  $\beta$  eine nicht äquidistante Kurvenschar, die sowohl für große als auch kleine  $\beta$  eine Überlappung in den dickeren, schwarzen Kurven zur Folge hat. Von Kiedrowski [78] sah dies in Abb. (b) für kleine  $\beta$ . Eine Übersicht, wie  $\beta$  die Zeit reskaliert, findet sich im Text und in Abb. 1.12.c. Auch bei Variation von  $\alpha$  in Abb. (c, d) ergibt sich eine Überlappung für kleine  $\alpha$ . Bei logarithmischer Skalierung sind die sigmoidalen Kurven äquidistant für alle großen, betrachteten  $\alpha$ ; jedoch würden sich diese bei einer linearen Skalierung ebenfalls in einer Grenzkurve überlappen; vgl. Abb. 1.12.a. Anfangskonzentration  $[A_1] = 2^{17}$ . Alle Kurven wurden auf die Anfangszeit  $t = 10^{-22}$  verschoben. Das Verhalten der Zeit bis zum Reaktionsbeginn findet sich in Abb. 1.11.  $\alpha$  und  $\beta$  wurden im Bereich von  $10^k$ ,  $k = -9 \dots 9$  variiert. Abbildungen (a, b) teilweise nach [78].

## 1 Eigenschaften der biologischen Evolution

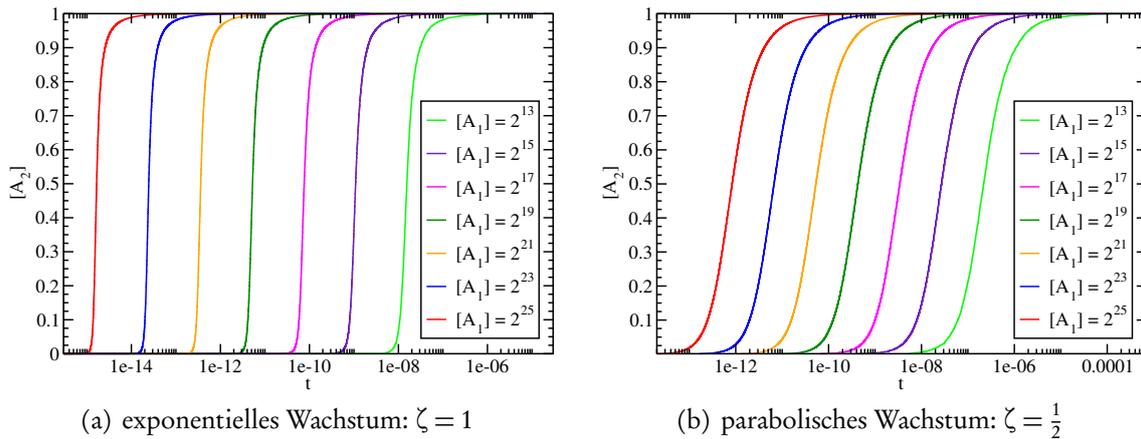


Abbildung 1.10: **Wachstumsverhalten von Dimeren für verschiedene Anfangskonzentrationen.** Zum besseren Vergleich wurde die Konzentration aller Kurven normiert. Alle Kurven sind in logarithmischer Zeit, beginnend mit der größten Anfangskonzentration, äquidistant angeordnet. Das exponentielle Wachstum erfolgt schneller als das parabolische; vgl. Abb. 1.9. Die Parameter sind  $\alpha = 10$  und  $\beta = 1$ , die Anfangsmonomerkonzentration wurde im Bereich von  $2^k$ ,  $k = 13, 15 \dots 25$  variiert. Alle Kurven wurden auf die Anfangszeit  $t = 10^{-22}$  verschoben.

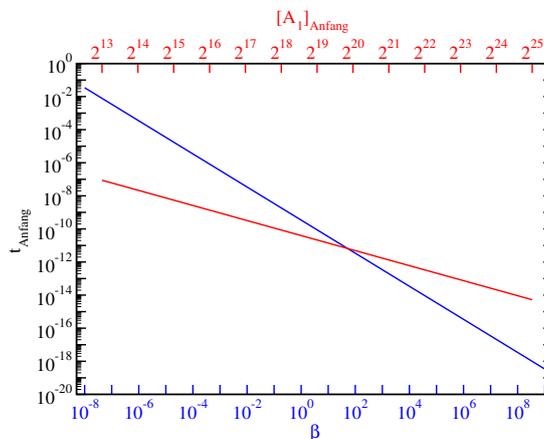


Abbildung 1.11: **Zeit bis zum Reaktionsbeginn für verschiedene Parameter** für eine spezielle Realisierung. Andere Realisierungen führen zu parallelen Geraden. Größere Anfangskonzentrationen führen zu einer kürzeren Zeit bis zum Reaktionsbeginn:  $t_A \sim [A_1]_0^{-1}$  und größere spontane Reaktionsraten ebenfalls zu einer kürzeren Zeit bis zum Reaktionsbeginn:  $t_A \sim \beta^{-2}$ . Die Proportionalitätskonstanten sind gerundet  $3,43 \cdot 10^{-10}$  für die Anfangskonzentration und 5,9 für  $\beta$ . Die Zeit bis zum Reaktionsbeginn ist unabhängig von  $\zeta$  und  $\alpha$  (nicht dargestellt). Vergleiche auch Abb. 1.9, 1.10 und 1.13.

## 1.4 Analyse der Voraussetzungen für ein System für chemische Evolution und dessen Konstruktion

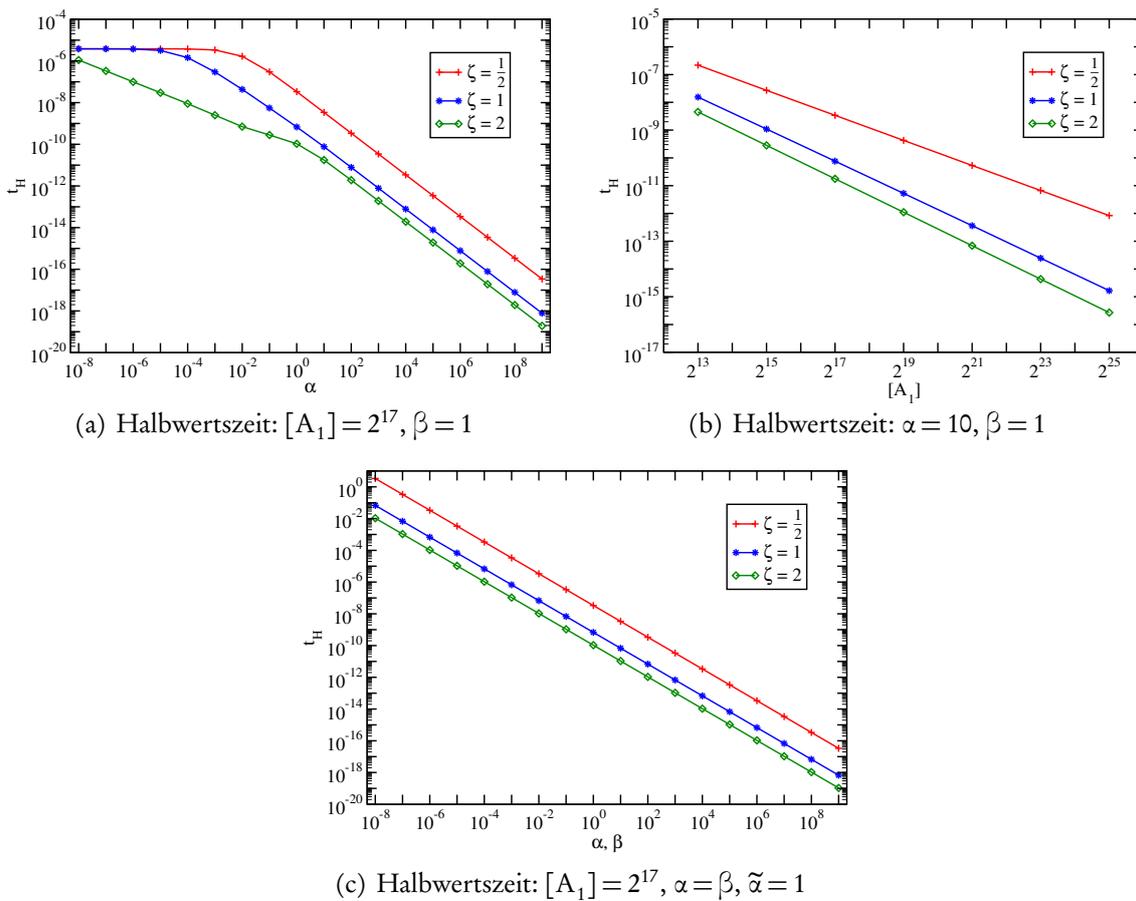


Abbildung 1.12: **Halbwertszeiten** des Wachstums in Abhängigkeit von  $\alpha$ ,  $\beta$  und der Anfangskonzentration. Die Kurven wurden mit einem Potenzgesetz angenähert:  $t_H = a \cdot x^{-b}$ , wobei  $x$  die Abszisse bezeichnet. Die Halbwertszeit bezeichnet in diesem Fall gerade die Zeit, die benötigt wird, um die Hälfte der Ausgangsmoleküle (Monomere) zu verbrauchen bzw. die Hälfte der Dimere zu erzeugen. (a) Die Abhängigkeit der Halbwertszeit von  $\alpha$  ist für alle Wachstumsarten zweigeteilt. Für kleine Zeiten ist die Abhängigkeit flacher als für große Zeiten und verläuft für  $\zeta \leq 1$  waagrecht  $|b| \leq 0,004(2)$  und mit  $b = 0,53(1)$  für  $\zeta = 2$  flacher als für große  $\alpha$ . Die Proportionalitätskonstanten sind (für kleine und große Zeiten) für  $\zeta = \frac{1}{2}$ :  $a \approx 3,8 \cdot 10^{-6}$ ,  $a \approx 3,4 \cdot 10^{-8}$ , für  $\zeta = 1$ :  $a = 3,5 \cdot 10^{-6}$ ,  $a = 7,4 \cdot 10^{-10}$  und für  $\zeta = 2$ :  $a = 6,4 \cdot 10^{-11}$ ,  $a = 1,9 \cdot 10^{-10}$ . Ab  $\alpha > 1$  verläuft die Halbwertszeit für alle betrachteten  $\zeta$  mit  $b = 0,99(1)$ . (b) Der Zusammenhang zwischen Anfangsmolekülzahl  $[A_1]$  und der Halbwertszeit ist für alle Wachstumsarten unterschiedlich und verläuft flacher je kleiner  $\zeta$  ist. Für  $\zeta = \frac{1}{2}$  ist  $a = 0,16(1)$  und  $b = 1,50(1)$ , für  $\zeta = 1$  ist  $a = 0,56(3)$  und  $b = 1,93(3)$  und für  $\zeta = 2$  ist  $a = 0,30(1)$  und  $b = 2,00(1)$ . (c) Ausgehend von Gl. 1.6 kann zur Reskalierung der Zeit  $\beta$  ausgeklammert werden, wenn man  $\alpha = \tilde{\alpha} \cdot \beta$  setzt. Dann ist  $b = 1,000(1)$  und  $a \approx 3,4 \cdot 10^{-8}$  für  $\zeta = 1/2$ ,  $a \approx 6,8 \cdot 10^{-10}$  für  $\zeta = 1$  und  $a \approx 1,1 \cdot 10^{-10}$  für  $\zeta = 2$ . Wie zu erwarten, ist der Exponent unabhängig von der Art der Autokatalyse. Alle Kurven wurden vor der Auswertung auf die Anfangszeit  $t = 10^{-22}$  verschoben.

Gl. 1.13 ist ein Modell für sexuelle Fortpflanzung, da (hier ohne Unterscheidung der Geschlechter) zwei Individuen benötigt werden, um einen zusätzlichen Nachkommen zu erzeugen [14, 81].

Für große  $\beta$  überwiegt dieser Reaktionsweg und es stellt sich eine vergleichbare Situation wie im vorherigen Abschnitt ein: eine Grenzkurve in linearer Zeit, die bei logarithmischer Zeitskala verschwindet und erkennbar sigmoidal verläuft. Bei kleinen  $\beta$  hat die spontane Bildung kaum Relevanz und es bildet sich im Grenzfall die steile Sigmoidalkurve für hyperbolisches Wachstum aus; vgl. Abb. 1.13.a.

Bei Veränderung von  $\alpha$  fällt im Gegensatz zum exponentiellen Wachstum auf, dass die Abstände hin zu großen autokatalytischen Raten einen Übergang von kleineren zu größeren Abständen erfahren. Dies ist bedingt durch den Übergang des Exponenten des Potenzgesetzes von  $b = 0,53(1)$  hin zu  $b = 0,99(1)$ ; vgl. Abb. 1.13.b und 1.12.a.

Da die Reaktionsraten nicht speziell von der Anfangskonzentration abhängen, ist das Verhalten vergleichbar zu den anderen Wachstumsarten, insbesondere zum exponentiellen Wachstum; vgl. Abb. 1.13.c und 1.12.b.

Die Halbwertszeit bei korrekter Reskalierung der Zeit ist qualitativ unabhängig von der Wachstumsart; vgl. Abb. 1.12.c.

Im folgenden Abschnitt wollen wir die hier gewonnenen Ergebnisse nutzen, um ein System, welches chemische Evolution in Form einer strikten Längenverdopplung zeigt, zu konstruieren.

#### 1.4.4 Konstruktion eines Systems für Evolution

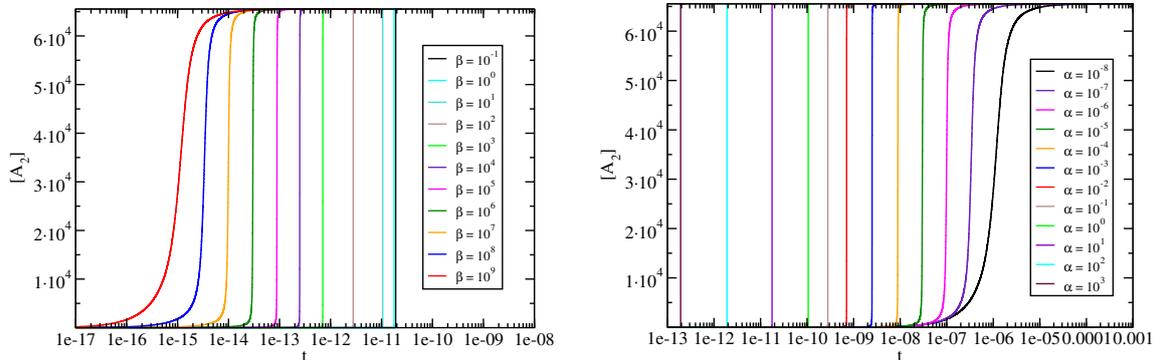
Ziel dieses Abschnitts ist es, ein System zu konstruieren, welches auf Grund seiner Selbstähnlichkeit in der Lage ist, eine Längenverdopplung aufzuweisen. Diese kann als dynamische Selektion von Polymeren bestimmter Länge im Zeitverlauf angesehen werden, so dass ein solches Modell als Grundlage für ein evolutionäres System dienen kann [80]. Insbesondere wollen wir überprüfen, ob es möglich ist, mit einem allgemeinen Parametersatz eine selbstständige Längenselektion zu konstruieren, d. h. ohne äußeres Eingreifen. Das schematische Vorgehen stellt sich wie folgt dar:

Man stelle sich eine Menge an chemisch identischen Ausgangsmolekülen vor, die in der Lage sind, die in Gl. 1.5 beschriebene Reaktion sowohl spontan als auch autokatalytisch durchzuführen. Diese schreiben wir mit  $n = 1$  um als:



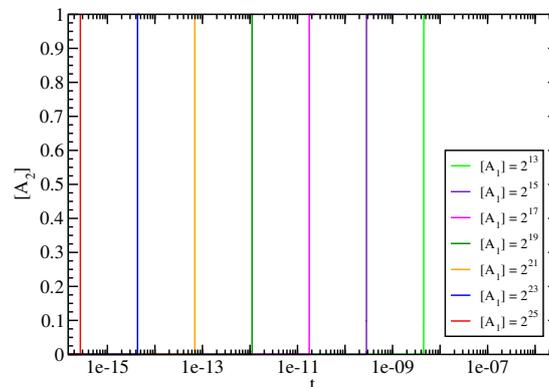
Dann gibt es eine endliche Zeit, in der fast alle Moleküle reagiert haben. Es sind doppelt so große Moleküle (hier: Dimere) entstanden. Aus diesen Molekülen ( $A_n$ ; hier:  $n = 2$ ) könnten sich dann in einem nächsten Schritt wiederum längenverdoppelte Moleküle ( $A_{2n}$ ) bilden. Dieser Schritt kann sich wiederholen, insbesondere, wenn sich die grundlegenden Reaktions-eigenschaften nicht ändern. Dafür darf sich die Zahl der reaktiven Zentren pro Molekül für verschiedene Molekülgrößen nicht ändern. Ein reaktives Zentrum ist eine Stelle in einem Molekül, an dem eine Reaktion stattfinden kann. Ohne Beschränkung der Allgemeinheit nehmen wir daher zur Illustrierung eine lineare Struktur an. Dies stellt sicher, dass nach jeder Reaktion die zwei Enden als Reaktionszentren in Betracht kommen. Des Weiteren ist so auch ein autokatalytisches Schema über eine Templatehybridisierung oder Vergleichbares denkbar.

## 1.4 Analyse der Voraussetzungen für ein System für chemische Evolution und dessen Konstruktion



(a) hyperbolisches Wachstum:  $\zeta = 2$ ,  $\alpha = 10$

(b) hyperbolisches Wachstum:  $\zeta = 2$ ,  $\beta = 1$



(c) hyperbolisches Wachstum:  $\zeta = 2$ ,  $\alpha = 10$ ,  $\beta = 1$

Abbildung 1.13: **Hyperbolisches Wachstum von Dimeren.** Konzentration der Dimere für hyperbolisches Wachstum ( $\zeta = 2$ ) nach Gl. 1.5. (a) Bei Veränderung von  $\beta$  gibt es einen Übergang von der für große  $\beta$  charakteristischen Sigmoidalkurve hin zu einer für hyperbolisches Wachstum charakteristischen sehr steilen Sigmoidalkurve bei kleinen  $\beta$ . Diese stellt eine Grenzkurve dar, da der Einfluss der spontanen Reaktion abnimmt. (b) Bei Veränderung der autokatalytischen Rate hin zu kleinen Werten deutet sich, wie zu erwarten, eine Überlappung an (nicht dargestellt). Wie beim (sub)exponentiellen Wachstum (vgl. Abb. 1.9) gibt es hier auch in der logarithmischen Darstellung keine Überlappung bei großen autokatalytischen Werten. Dort ist jedoch ein Übergang zu verschiedenen großen Abständen zu erkennen. Wie in Abb. 1.12.a zu erkennen ist, gibt es für  $\zeta = 2$  zwei verschiedene Bereiche unterschiedlich starker Autokatalyse. (c) Aus Abb. 1.10 und 1.12.b ist ersichtlich, dass die parallelen Abstände etwas größer als beim exponentiellen Wachstum sind. Falls nicht anders angegeben ist  $[A_1] = 2^{17}$ . Alle Kurven wurden auf die Anfangszeit  $t = 10^{-22}$  verschoben.

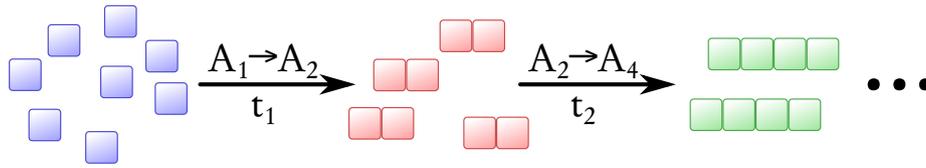


Abbildung 1.14: **Schema eines Reaktionssystems zur Längenverdopplung.** Aus Edukten, hier Monomeren, werden zunächst Dimere gebildet ( $2A_1 \rightarrow A_2$ ). Nachdem diese Reaktion (fast) vollständig abgelaufen ist, wird eine typengleiche Reaktion mit dem Produkt der vorhergehenden Reaktion gestartet:  $2A_{2^k} \rightarrow A_{2^{k+1}}$  (hier  $2A_2 \rightarrow A_4$ ). Nachdem diese beendet ist, kann repetitiv in gleicher Weise eine fortlaufende Längenverdopplung ablaufen, solange bis letztlich die Edukte aufgebraucht sind.

Um die Aussagen allgemeingültig zu halten, werden wir insbesondere in Kapitel 3 ein exponentielles Wachstum ( $\zeta = 1$ ) wählen, was auch eine Möglichkeit in einer templatebasierten Reaktion ist [78].

Ein Reaktionssystem zur Längenverdopplung können wir mit Hilfe folgender Beobachtungen konstruieren: Erstens die **Endlichkeit der Halbwertszeit**  $t_H$  und damit verbunden die endliche Zeitspanne bis fast alle Moleküle reagiert sind und sich ein Plateau ausbildet. Und zweitens die **Totzeit bis zum Reaktionsanfang**  $t_A$  der darauf folgenden Reaktion. Läge nun eine Vielfaches (größer als 1) der Halbwertszeit der Reaktion innerhalb der Totzeit zum Reaktionsstart der folgenden Reaktion ( $t_H > t_A$ ), so würde die erste Reaktion (fast) vollständig ablaufen und erst dann könnte eine weitere, auf den Produkten der ersten aufbauende Reaktion stattfinden. Die Reaktionen können mit Gleichung 1.14 beschrieben werden. Die erste Reaktion ist die Reaktion von Monomeren zu Dimeren ( $n = 1$ ), die zweite die Reaktion von Dimeren zu Tetrameren ( $n = 2$ ). Die Struktur aller (beider) Reaktionen ist identisch. Damit sind diese selbstähnlich und stellen eine Längenverdopplungskaskade dar, die – im Allgemeinen bei gleichen Reaktionsbedingungen – in gleicher Weise weitergehen kann, bis die nötigen Edukte nicht mehr in ausreichender Zahl vorhanden sind; vgl. Abb. 1.14, Abschnitt 3.2.2 und 3.4.1.

Die Halbwertszeit ist abhängig von der Plateaukonzentration und dem autokatalytischen Parameter. Aus Abb. 1.12.a und 1.12.b ergibt sich die Gesamtabhängigkeit zu:  $t_H \sim \alpha^{-1} \cdot [A_1]_0^{-\nu}$ . Insbesondere ist zu beachten, dass wir  $\alpha$  so wählen, dass die Halbwertszeit nicht sättigt; vgl. Abb. 1.12.a. Die entsprechende minimale Anfangskonzentration, die in der Nähe der Sättigung liegt, definiert den Phasenübergangspunkt. Des Weiteren wird hier der in Abhängigkeit von  $\zeta$  verschiedene Exponent mit  $\nu$  bezeichnet, welcher näherungsweise  $\nu = 2$  für  $\zeta \geq 1$  ist; vgl. Abb. 1.12.b.

Die Totzeit bis zum Reaktionsanfang ergibt sich aus Abb. 1.11 zu:  $t_A \sim [A_1]_0^{-1}$ . Da  $t_A \gtrsim t_H$  ist, lässt sich  $t_A$  und  $t_H$  gleichsetzen und man kann (für ein festes  $k$ ; vgl. Gl. 3.4) schreiben:

$$\alpha_c \sim [A_1]_0^{1-\nu} \rightarrow \alpha_{c,k} \sim \frac{1}{2^k} \cdot [A_1]_0^{1-\nu}. \quad (1.15)$$

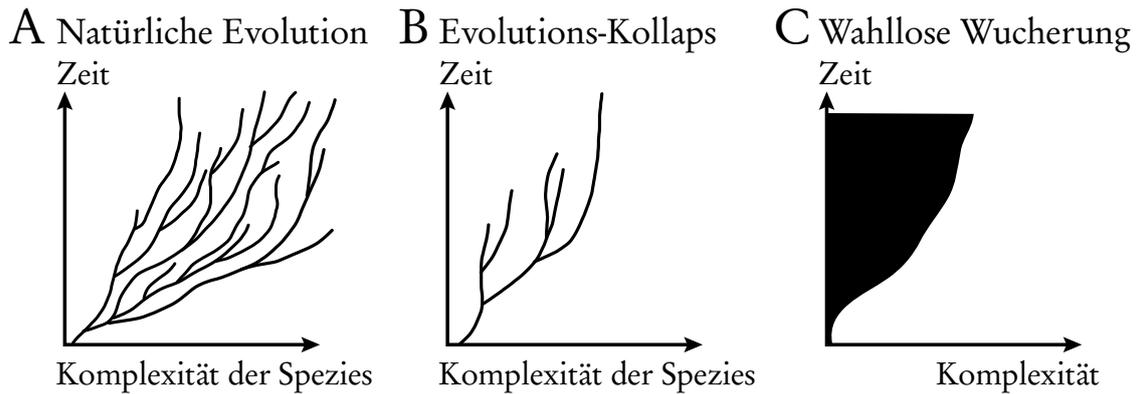


Abbildung 1.15: **Mögliche Verläufe der Evolution in einem Ökosystem.** Natürliche Evolution (A) zeichnet sich dadurch aus, dass im Zeitverlauf mehrere Spezies existieren, die im Mittel immer komplexer werden, da sie sich anpassen. Bei fixen, künstlichen Umweltbedingungen kommt es vor, dass sich eine am besten angepasste Spezies ausbildet und die weitere Evolution kollabiert (B). Fehlender Selektionsdruck kann zu einer wahllosen Wucherung (C) führen, da sich Lebewesen ausbilden, die sich kaum mehr unterscheiden, dennoch können diese komplexer werden. Veröffentlicht in [82].

Diese Gleichung hat die gleiche Struktur wie Gl. 3.5. Um eine Längenverdopplung für beliebige  $k$  bis zur Länge  $2^k$  zu erhalten, ist die Plateaukonzentration der jeweiligen Länge zu betrachten. Der Korrekturfaktor gibt die Konzentration der entsprechenden Länge an; vgl. Abschnitt 3.4.1 und Gl. 3.4. Der Exponent  $1 - \nu \doteq -1$  ist vergleichbar mit dem in Kapitel 3.4.2 in der vollen Rechnung gefundenen ( $\gamma = -1$ ); vgl. auch Abb. 3.8.

## Zusammenfassung

Wir konnten in diesem Abschnitt das allgemeine Verhalten von autokatalytischen Reaktionen für verschiedene Wachstumstypen untersuchen. Ein templatebasierender Prozess zeigt typischerweise ein parabolisches Wachstum. Durch Vermeidung von Produktinhibition ist es möglich, ein nahezu exponentielles Wachstum zu erreichen. Hyperbolisches Wachstum ist typisch für sexuelle Reproduktion.

Des Weiteren konnten wir Voraussetzungen für ein längenverdoppelndes System, welches auf einer autokatalytischen Reaktion beruht, herleiten. Diese unterscheiden sich qualitativ nicht von der zu Grunde gelegten Wachstumsart.

Inwieweit sich die Ergebnisse reproduzieren lassen und welche weiteren Schlussfolgerungen wir ziehen können, werden wir in Kapitel 3 behandeln.

Nun wenden wir uns abschließend allgemeinen Betrachtungen zur Evolution zu und welche Verläufe diese aufzeigen kann.

## 1.5 Abläufe der Evolution

In diesem Abschnitt fassen wir die vorhergehenden Ergebnisse zusammen und leiten daraus allgemeine Eigenschaften der Evolution ab. Wir postulieren, dass diese Eigenschaften allgemeingültig sind und sich auf alle Evolutionsstadien von chemischer zu weit fortgeschrittener biologischer Evolution übertragen lassen\*.

Zunächst werden wir einige Hypothesen zur Entstehung biologischen Lebens vorstellen. Die Progenote [83, 84, 85] ist eine Art Urzelle, aus der sich einzelne Zellarten, wie Bakterien, Eukaryoten und Blaualgen, differenziert haben. Die Grundlage zur Weitergabe der Information ist in einem aperiodischen Kristall, der DNA, gespeichert, dessen Eigenschaften Schrödinger [22] vor dessen Entdeckung untersucht hat. Aus diesen Hypothesen leiten wir dann grundlegende Eigenschaften evolutionärer Systeme ab, die wir zur Interpretation der Ergebnisse in den Kapiteln 2 und 3 benötigen.

Diese Eigenschaften können zu verschiedenem evolutionären Verhalten führen, welches wir anhand von Abbildung 1.15 charakterisieren und Beispiele dazu vorstellen.

### Die Progenote

Obwohl es prinzipiell in vielen Systemen möglich ist, dass beliebige Arten von Lebewesen entstehen, gibt es eine Begrenzung in deren Vielfalt. Dies macht es beispielsweise auch erst möglich, Arten zu klassifizieren. In der Anfangsphase der chemischen Evolution, in der sich erst sukzessive die nötigen, funktionalen Moleküle gebildet haben, hat anfangs ein Zustand existiert, in dem einzelne Exemplare von Urzellen nach und nach Vorläufer heutiger Zellbestandteile produziert haben, die sie auch untereinander ausgetauscht haben, so Woese [83, 84, 85]. Woese postuliert einen gemeinsamen Vorfahren von Prokaryoten und Eukaryoten, den er *Progenote* nennt. Dieser gemeinsame Vorfahre enthält die verschiedenen funktionalen Einheiten, die notwendig sind, eine lebende Zelle zu bilden. Jede dieser Einheiten beinhaltet sein eigenes informationsverarbeitendes System. Eine rudimentäre Zellteilungsfunktion ist vorhanden, jedoch stellt diese nicht sicher, dass in jeder Tochterzelle jedes dieser Systeme (in ausreichender Zahl) vorhanden ist. Daher können bei der Zellteilung durchaus sehr unterschiedliche Zellen entstehen, von denen auch viele nicht überlebensfähig wären. Da sich die Progenoten in einem einfachen Zustand befinden und daher kompatibel zueinander sind, ist ein horizontaler Gen- bzw. Systemtransfer zwischen verschiedenen Progenoten möglich, der eventuell fehlende Systeme ausgleichen kann. Die Systeme der Progenoten entwickeln sich ausgehend von der RNA-Welt von funktionalen RNA-Ketten hin zur Trennung vom Informationsspeicher (RNA und DNA) und der Funktion (Enzyme). Auf diesem Weg wird auch ein sicher arbeitendes Zellteilungssystem benötigt, um die verschiedenen Systeme vollständig auf die Tochterzellen aufzuteilen. Aus der Progenote entwickeln sich dann die drei

---

\* Die **Allgemeingültigkeit der Evolution** für alle biologischen Lebewesen – vom Virus über das Bakterium zu höheren Lebewesen wie Säugetieren – ist evident [1, 14]. Daher stellen wir hier die Hypothese auf, dass die gleichen Mechanismen ebenfalls ganz am Anfang bei der chemischen Evolution eine Rolle gespielt haben; eine vergleichbare Aussage treffen auch Eigen und Schuster [19]. Freilich kann dies erst als plausibel angesehen werden, wenn ein derartiger Übergang von selbstassemblierenden, chemischen Molekülen hin zu einfachsten biologischen Lebewesen im Labor vollzogen wurde. Im Gegensatz dazu geht Woese [83, 84, 85] jedoch davon aus, dass die darwinschen Eigenschaften der Evolution erst aufgetreten sind, als sich die ersten Zellen als Spezies ausbildeten, da erst dann eine Unterscheidung zwischen Geno- zu Phänotyp möglich ist.

Hauptzellarten Bakterien, Blaualgen und Eukaryoten auch durch horizontalen (d. h. artenübergreifenden) System- bzw. Gentransfer. Dies erklärt auch, dass verschiedene Zellorganellen der Eukaryoten aus verschiedenen Evolutionsstadien stammen: Mitochondrien besitzen z. B. ein eigenes Genom [1] während der Großteil des übrigen genetischen Codes, der zur Funktion einer Zelle notwendig ist, sich im Zellkern befindet. Der Übergang von der Progenote zu als Spezies ausgebildeten Zellen markiert dann auch erstmals die Unterscheidung Phänotyp und Genotyp, d. h. Aussehen und Informationsspeicher sind klar getrennt und bedingen einander.

Das bedeutet, dass der Übergang von einer Situation, wie sie sich am Anfang der chemischen Evolution darstellte, von einer wahllosen Wucherung, in der (fast) alle möglichen Konformationen der Lebewesen vorkommen und es daher keinen Sinn macht, sie zu unterscheiden (vgl. Abb. 1.15.c), hin zu der von Darwin charakterisierten natürlichen Evolution (vgl. Abb. 1.15.a) stattgefunden hat.

Diese Arbeit möchte in dieser Hinsicht auch Antworten auf die Fragen geben, die dadurch entstehen: Warum setzte plötzlich die darwinsche Evolution ein und war nicht vorher präsent? Wie konnten die notwendigen RNA-Sequenzen entstehen, die ein System letztlich ausmachen?\*

Neben der Arbeit von Woese möchten wir einen weiteren Zugang vorstellen, der nahelegt, dass der Informationsgehalt der ersten chemischen evolvierenden Moleküle eine zunächst untergeordnete Rolle gespielt hat.

### Der aperiodische Kristall

Schrödinger unterscheidet in seinem Buch „What is Life?“ [22] zwischen dem für das Leben interessanten, aperiodischen Kristall und dem unter physikalischen Aspekten interessanten, periodischen Kristall. Der periodische Kristall ist, so Schrödinger, ein sehr kompliziertes Objekt, welches aus der faszinierendsten und komplexesten Materialstruktur unbelebter Natur besteht. Einmal entschlüsselt, definiert die Kristallstruktur zusammen mit dem zu Grunde liegenden, erzeugenden Molekül den gesamten makroskopischen, aperiodischen Körper. Ebenso verhält es sich mit dem aperiodischen Kristall. Seine Kristallstruktur mag vordergründig unwichtig erscheinen\*\*, jedoch trägt seine lokale (Un-)Ordnung die Information, aus denen die schönsten und bemerkenswertesten Lebewesen entstanden sind.

Seit Watson und Crick [62], sowie Nirenberg und Matthaei [86] sind die Anfänge der Entschlüsselung des genetischen Codes bekannt, welcher dann auch alsbald grundlegend entschlüsselt wurde [1, 14]. Nirenberg und Matthei konnten nachweisen, dass ein RNA-Strang, der nur aus Uracil besteht, die Aminosäure Phenylalanin kodiert. Nirenberg *et al.* [87]

---

\* Zu Beginn der Evolution, also im oder vor dem Zustand der Progenote, kann keine sinnvolle Unterscheidung zwischen Geno- und Phänotyp getroffen werden, da die informations- bzw. funktionstragenden Moleküle nicht einer speziellen Spezies von Zelle zugeordnet sind. Daher ist es sinnvoll, nur den Phänotyp als die Spezies definierende Eigenschaft anzusehen, also auch im Fall der (zunächst funktionslosen) RNA- bzw. DNA-Stränge. Diese Arbeit legt zudem nahe, dass die Information zunächst eher eine untergeordnete Rolle gespielt hat. Die Selektion der Länge, wie sie hier vorgestellt wird, kann dazu dienen, genügend lange DNA- bzw. RNA-Stränge zu produzieren, die dann im weiteren Verlauf durch Selektion ein System im obigen Sinne darstellen könnten.

\*\* Die **Kristallstruktur** jedoch stellt sicher, dass die Information stabil in Form einer Doppelhelix aufbewahrt wird und dass diese räumlich stark komprimiert in Form von Chromosomen in der Zelle und besonders im Zellkern gespeichert werden kann [1].

fanden dann wenige Jahre später für fast alle Basen heraus, dass letztlich ein Basentriplett eine Aminosäure codiert. Damit ist der von Schrödinger postulierte Informationsträger – ein aperiodischer Kristall – identifiziert [22].

Jedoch wollen wir hier in dieser Arbeit diesen Ansatz nicht weiter verfolgen. Herauszufinden, wie sich dieser Kristall, die DNA, verändert und neue Arten entstehen, ist in vielfältiger Weise Gegenstand biologischer Forschung, der Genetik und Evolutionsbiologie. Vielmehr wählen wir aus mehreren Gründen einen anderen Ansatz. Zum einen kann es, so wie Schrödinger es postulierte, gar keinen kleinen aperiodischen Kristall in der chemischen Evolution gegeben haben, da chemische Moleküle schon die kleinsten biologisch und chemisch aktiven Einheiten darstellen. Damit wäre das Molekül schon selbst der Informationsträger und müsste zugleich eine biologische Funktion, nämlich sich zu evolvieren, erfüllen. Obwohl dies möglich erscheint, steht dies im Widerspruch zum Aufbau gegenwärtiger biologischer Zellen. Diese bestehen aus DNA und der dazu passenden biologischen „Maschinerie“, aus Enzymen, die eine entsprechende Interaktion und damit einen konstruktiven Ablauf im Sinne einer lebenden Zelle ermöglichen [1, 14]. Aber woher kamen die passenden Enzyme? Eine Antwort liefern hier Woese sowie Lehn im folgenden Abschnitt. Jedoch besteht das Problem, wie die erste RNA bzw. DNA sich zuverlässig gebildet hat.

### **Selbstorganisation und Evolution**

Lehn [88, 89] untersucht die Möglichkeiten supramolekularer Chemie und deren Erweiterungen. Supramolekulare Chemie bezeichnet die Interaktion von (bestimmten) Molekülen über geeignete Wechselwirkungen, wie z. B. intermolekulare Kräfte und Wasserstoffbrückenbildung. Die Supramoleküle, angelehnt an die Funktionalität von Biomolekülen, haben dadurch Interaktionsmöglichkeiten, die über die grundlegenden chemischen Eigenschaften hinausgehen. Durch Selbstassemblierung können so verschiedene, neue Strukturen entstehen, wie z. B. Flüssigkristalle, Liposome (Vesikel) und supramolekulare Polymere. Diese supramolekularen Einheiten können über äußere physikalische und chemische Einflüsse, wie z. B. Temperatur, Druck, elektrische und magnetische Felder, Redoxpotential, pH-Wert, (un)geladene Moleküle und Metallionen, verändert werden und so Information speichern, aber auch verarbeiten, z. B. durch molekulare Erkennungsprozesse. Damit ist es auch möglich, eine (durch äußere Einflüsse) programmierte Selbstassemblierung von Supramolekülen zu bewerkstelligen. Gibt es prinzipiell mehrere Möglichkeiten der Selbstassemblierung und können damit komplexere, also nicht direkt vorhersagbare Strukturen aus supramolekularen Einheiten entstehen, spricht man auch von Selbstorganisation. Hier tritt zunehmend die intrinsische, gegenseitige Beeinflussung unterschiedlicher supramolekularer Einheiten in den Fokus. Selbstassemblierung von Supramolekülen ist nur möglich, wenn die zu Grunde liegenden Moleküle sich dynamisch durch die Erkundung der Energie- bzw. Strukturhyperfläche an die Programmierung anpassen. Dann spricht man auch von der Chemie der dynamischen Strukturen. Auf einem supramolekularen Niveau ist die Interaktion nicht kovalent, die Dynamik physikalisch und die Supramoleküle sind durch komplementäre Wechselwirkungen gekennzeichnet. Aber auch für chemische Moleküle sind dynamische Strukturen möglich. Die Bindung erfolgt funktionell und kovalent, die Dynamik ist chemisch und die Wechselwirkung erfolgt durch komplementäre funktionelle Gruppen. Betrachtet man ein Ensemble solcher Moleküle, eine Bibliothek dynamischer Strukturen, so kann man einen Satz von Komponenten als Genotyp und das dynamische System und die Sätze der Bestandteile als Phänotyp bezeichnen. Die Chemie der

dynamischen Strukturen wurde bisher benutzt, um dynamische Rezeptoren bzw. Substrate für molekulare Erkennungsprozesse zu erzeugen, bioaktive Substanzen zu suchen und dynamische Materialien (z. B. dynamische Polymere) zu entwickeln. Eine Bibliothek dynamischer Strukturen kann z. B. durch Erzeugung einer riesigen Menge an molekularen Schlüsseln für ein biologisches Schloss durch thermodynamische bzw. kinetische Selektion, d. h. stärkste bzw. schnellste Interaktion, einen passenden Schlüssel finden. Es ist z. B. iminhalten Einheiten auch möglich, Informationen kurzlebig in der Form als auch langlebig als neues Molekül abzuspeichern. Die Chemie der dynamischen Strukturen gibt letztlich Zugang zu chemischen Netzwerken dynamischer Struktur. In diesen sind die einzelnen Einheiten über verstärkende oder gegensätzliche Wechselwirkungen verknüpft.

Ausgehend davon kann sich die Chemie der dynamischen Strukturen daher auch zur Beschreibung und Erforschung chemischer Evolution eignen.\* Chemische Netzwerke dynamischer Struktur können prinzipiell zusammenhängend evolvieren, indem eine Rückkopplung zwischen mehreren Spezies zu einer gleichzeitigen Optimierung der (oder einiger) Spezies führt, also einem Koevolutionsprozess, z. B. der Optimierung und Expressionssteigerung einer Schlüssel-Schloss-Kombination. Da die funktionale Interaktion der supramolekularen Einheiten beliebig abstrahiert werden kann, kann eine supramolekulare Interaktion auf verschiedenen Ebenen der Komplexität wirken und auch z. B. aus Symmetriegründen zu einer Selektion führen. So wird in [90] das symmetrischste Supramolekül gebildet und von den ansonsten stabilen Supramolekülen katalysiert. Lehn fasst zusammen, dass Selbstorganisation der fundamentale kosmische Prozess ist, welcher im Laufe der Evolution des Universums zur Bildung komplexer Materie geführt hat, ausgehend von Teilchen bis hin zum denkenden Organismus.

## Nischenbildung

Bei vielen Lebewesen tritt, so Laland und Odling-Smee [91, 92, 93], lokale Nischenbildung auf. So suchen viele Lebewesen ihre Lebensräume, Partner und Hilfsmittel und gestalten wesentliche Teile ihrer lokalen Umwelt, wie Nester, Baue und Wege. Beispielsweise züchten Blattschneiderameisen in ihren Nestern auf den gesammelten Blattstücken Pilze (*Leucocoprianea* und *Pterulaceae*), von denen sie sich ernähren [94]. Dabei achten sie auf Reinlichkeit, um Verunreinigungen mit anderen Pilzen zu vermeiden und züchten das Bakterium *Streptomyces* auf ihrem Integument\*\*, um dessen *Candididin* zur Bekämpfung anderer Pilzarten einzusetzen [94]. Auch enthalten Walnussbäume einen Stoff, der ähnlich zu *Juglon* ist, der mit den abgefallenen Blättern in den Boden eingetragen wird und damit das Pflanzenwachstum hemmt [96].

---

\* Lehn stellt Selbstorganisation von insbesondere abstrahierbaren chemisch-funktionalen Einheiten, die auch Informationen verarbeiten können, im Zeitverlauf in den Vordergrund. Es ist aber ebenfalls wichtig, die funktionalen, chemischen Reaktionsarten in einem örtlichen bzw. zeitlichen Kontext zu betrachten; d. h., dass es u. U. ebenso wichtig ist, die richtige chemische Reaktion zur rechten Zeit am rechten Ort auszuführen, um eine Selektion zu ermöglichen. In Eukaryoten ist z. B. die Genexpression durch mehrere räumlich und zeitlich getrennte Reaktionen gekennzeichnet [1]. Dabei wird die Information des Gens auf RNA kopiert (Translation), verarbeitet und schließlich in der Transkription in ein Protein übersetzt [1].

\*\* Das **Integument** ist die äußere Hülle eines Lebewesens [95]. Im Fall der Ameisen ist dies der Chitinpanzer.

Odling-Smee und Laland leiten aus Beobachtungen in der Natur verschiedene Mechanismen ab, wie Nischenbildung auftreten kann. Zum einen durch genetische Vererbung, so dass, falls in jeder Generation jeder Organismus seine ontogenetische Umgebung gleichermaßen ändert, die Vorfahren eine Quelle natürlicher Selektion ihrer Nachkommen durch wiederholte Nischenbildung darstellen können. Das heißt, dass eine Änderung der lokalen Umwelt durch kollektives Verhalten einzelner Individuen eine Nische darstellt. Diese Nische, welche durchaus unterschiedlich für unterschiedliche Individuen sein kann, kann zu einer evolutionären Reaktion führen, so dass die entsprechenden Nachkommen entsprechend selektiert werden. Beispielsweise wählen Kuckuckseltern immer wieder ein Wirtsnest der selben Vogelart für ihre Eier, die der selbigen angepasst sind und vererben damit veränderte Gene an ihre Nachkommen. Im Allgemeinen ist die gewählte Vogelart von Kuckuck zu Kuckuck unterschiedlich. Die dadurch veränderten Selektionsdrücke führen zu einer kürzeren Brutzeit gegenüber der Wirtsvogelart oder dem Herauswerfen der Wirtsvogeleier. Aber auch eine außergenetische Vererbung ist möglich. Regenwürmer verändern die Bodenbeschaffenheit so, dass sie nur optimal in dem von ihnen bearbeiteten Boden leben können. Auch indirekte Geninteraktionen sind möglich.

Eine Nische kann sich auch in Form einer scheinbaren Symbiose darstellen. Produziert ein Lebewesen ein (Abfall-) Produkt, welches ein anderes verwerten kann, kann dies einen Vorteil für andere Lebewesen darstellen. Dies ist nicht notwendigerweise eine Symbiose im klassischen Sinn, falls das produzierende Lebewesen das Abfallprodukt nicht weiter verwenden kann. Die verbrauchenden Lebewesen nutzen lediglich das ohnehin anfallende Abfallprodukt zur eigenen Verwertung. Der Ursprung der Nischenbildung kann aber auch rein erlerntes Verhalten sein. Britische Meisen haben gelernt, Milchflaschen zu öffnen und die Milch zu trinken. Daraus könnten Fähigkeiten entstehen, besser zu lernen oder die Enzyme zur Verwertung von Milch zu optimieren. Nischenbildung kann damit die Dynamik der Evolution verändern, indem sie nicht nur Gene, die an verschiedene Umweltsituationen angepasst sind, sondern auch Möglichkeiten entsprechende Gene zu verändern, weitergibt.

### Kontinuierliche Nischenbildung als Evolutionsmotor

Wie bei Lenski *et al.* können die vorhandenen Spezies z. B. durch Einfluss auf die (begrenzte) äußere Nahrung einen Evolutionsdruck verursachen und damit eine innere Nische darstellen: Ist genügend Nahrung vorhanden, können Arten, die diese Nahrung verwerten können und ansonsten keinen Einschränkungen unterliegen, frei wachsen. Wird die Nahrung durch den nun erhöhten Verbrauch knapper, wird auch das Wachstum gehemmt. Im Experiment von Lenski *et al.* führt dies nach vielen Generationen dazu, dass eine andere Nahrungsquelle gefunden wird, die ergiebiger ist, und damit eine neue Nische darstellt.

In den nachfolgend vorgestellten Systemen kommt es zu einer fortwährenden Bildung neuer Nischen, die es den jeweiligen Spezies erlauben, sich zu vermehren und so eine (prinzipiell) unendliche Kette an neuen Spezies zu bilden (vgl. Abb. 1.14)\*. Dabei stellt die jeweilige, in überwiegender Zahl vorhandene Spezies ( $A_n$ ), die Nische in Form von verwertbarer Nahrung für die im Folgenden wachsende Spezies ( $A_{2n}$ ) dar, die die doppelte Länge besitzt. Eine Nischenabfolge wurde in [91, 92, 93] beschrieben.

---

\* Durch die beschränkte Verfügbarkeit von Nahrung im System muss es natürlich letztlich zu einem Abbruch der Kette kommen.

## Gaiahypothese und natürliche Auslese

Die Gaiahypothese wurde zuerst von Lovelock und später auch von Margulis entwickelt [97]. Sie besagt, dass das Ökosystem der Erde einen Zustand einreguliert, der sicherstellt, dass Leben ermöglicht wird. Die Gaiahypothese betrachtet den Zusammenhang zwischen der Selektion einzelner Organismen und deren Einfluss auf eine solche Selbstregulierung. Damit wird die größte aller Nischen, die Erde, von ihren Bewohnern, den Tieren und Pflanzen, so verändert, dass ein vielfältiges Leben möglich ist. Dem Einfluss von Organismen auf deren Umwelt, speziell den Nischen, in denen sie leben, wird in der klassischen Evolutionstheorie wenig Beachtung geschenkt.

In den folgenden Abschnitten stellen wir das Modellsystem von Lenton [98] zur Beschreibung einer Gaia-Welt vor. Betrachtet man die Erdatmosphäre, so sieht man, dass sich diese in einem starken Ungleichgewicht befindet, da sie große Mengen an reaktiven Gasen wie Sauerstoff und Methan beinhaltet. Im Gegensatz dazu sind die Atmosphären von Mars und Venus nahezu im Gleichgewicht und enthalten viel Kohlenstoffdioxid. Die Ursache des Ungleichgewichtszustandes der Atmosphäre der Erde liegt in dem Vorhandensein von Pflanzen, die aus Sonnenlicht und Kohlenstoffdioxid Sauerstoff gewinnen und so den Treibhauseffekt des  $\text{CO}_2$  mindern. Das führt letztendlich zu günstigen Lebensbedingungen. Ein Anstieg der  $\text{CO}_2$ -Konzentration würde auch zu Schäden an den Pflanzen führen [99].

Generell gibt es abiotische Antworten auf Umweltveränderungen. Diese beeinflussen das Wachstum der besser angepassten Lebewesen, welche wiederum Einfluss auf die Population haben und damit Umweltveränderungen bewirken können. Aber auch klassische Evolution führt zur Selektion auf Grund von Umweltveränderungen, die auch aus den beiden zuvor genannten Mechanismen resultieren können.

Das Modellsystem von Lenton [98] beschreibt einen Planeten mit einer Art von Vegetation, den „Gänseblümchen“. Der Planet hat eine Albedo, d. h. einen Reflexionsgrad, von 0,4 und wird von einer Sonne beschienen. Die Einstrahlungsintensität nimmt im Laufe der Zeit, wie auf der Erde, zu. Auf dem Planeten leben zwei Arten von Gänseblümchen, die exponentiell wachsen können. Die Schwarzen haben eine Albedo von 0,15 und die Weißen eine von 0,65 und erzeugen wegen des unterschiedlichen Reflexionsgrads daher eine unterschiedliche Temperatur bei einer gegebenen Sonneneinstrahlung. Beide haben eine optimale Wachstumstemperatur von  $22,5^\circ\text{C}$  und können zwischen  $5^\circ\text{C}$  und  $40^\circ\text{C}$  überleben. Demnach wäre es von Vorteil, wenn bei einer geringen Sonneneinstrahlung schwarze Gänseblümchen wachsen, da sie die Temperatur des Planeten erhöhen. Bei großer Sonneneinstrahlung entsprechend weiße, die die ursprüngliche Temperatur absenken. Es stellt sich eine Temperatur zwischen ca.  $20^\circ\text{C}$  und  $30^\circ\text{C}$  ein. Dieses Modellsystem beschreibt lediglich den Zusammenhang eines äußeren Einflusses auf statische Pflanzen. Daher betrachten wir ein erweitertes Modell, bei dem sich ebenfalls die Gänseblümchen evolutionär verändern können.

Das erweiterte Modellsystem beherbergt Gänseblümchen, die durch Mutation ihre Albedo im Bereich von 0,2 und 0,65 in 0,05 Schritten anpassen können. Anfangs wachsen nur Gänseblümchen mit einer Albedo von 0,4, die der des Planeten entspricht. Da aber zunächst die Sonneneinstrahlung niedrig und die optimale Temperatur noch nicht erreicht ist, entstehen solche Gänseblümchen mit einer Albedo von 0,35, wobei die Folgenden nach und nach mit

steigender Einstrahlung schrittweise immer heller werden. Diese wechseln sich in ihrem Vorkommen ab und die Temperatur wird zwischen ca. 23 °C und 30 °C mit dem Auftreten der ersten Mutation eingeregelt.

Wir können konstatieren, dass Selektion die Regulation der Temperatur nur leicht beeinflusst und keinesfalls zerstört, da der Einfluss der Umwelt überwiegt. Führt man zusätzlich noch Herbivoren und Carnivoren in das Modellsystem ein, so stellt man fest, dass der dominante Herbivor durch die Gänseblümchen-Umwelt-Rückkopplung bestimmt wird und nicht umgekehrt. Des Weiteren wird die Widerstandskraft durch ein verzweigteres Nahrungsnetz mit mehr Verknüpfungen gestärkt.

Das Kernresultat dieser Beobachtungen, welches wir weiter verfolgen möchten, ist, dass man die Gesamtheit aller Organismen und deren Umwelt vollständig betrachten muss, um zu verstehen, welche Merkmale fortbestehen und dominieren können.

### **Reaktionen der Evolution**

In Abschnitt 1.2.1 wurden mögliche evolutionäre Reaktionen auf äußere Einflüsse beschrieben. Diese Reaktionen möchten wir hier charakterisieren. Es sind zwei Reaktionstypen zu beobachten: Zum einen eine sehr konkrete, bisweilen auch vorhersehbare Reaktion und zum anderen eine vielfältige Reaktion.

Die konkrete Reaktion können wir z. B. bei der Jagd auf Silberfüchse, Rotwild und Elefanten beobachten. Die Jagd führt letztlich zur Reduzierung der gejagten Eigenschaft, sodass ein Überleben wahrscheinlicher wird. Im Einzelnen sind dies das praktische Aussterben der gejagten Silberfüchse zugunsten anderer Fellfarben, und die Reduzierung der Geweihgröße bzw. Stoßzahnlänge bei Rotwild und Elefanten.

Eine vielfältige Reaktion auf eine konkrete Selektion kann im Fall der Hausfüchse beobachtet werden. Nur nach Zahmheit selektiert, werden diese zum einen zahm in Form einer konkreten Reaktion. Aber es treten auch weitere Reaktionen auf: Die Hausfüchse weisen teilweise z. B. ein scheckiges Fell, Veränderungen an den Schwänzen und Ohren sowie verkürzte Beine auf.

Diese beginnenden Veränderungen markieren den Anfang der Differenzierung in eine neue Art. Aber auch das Gegenteil ist möglich. Wie beim Beispiel der Darwinfinken gezeigt wurde, vereinigen sich zwei Arten mit unterschiedlichen Schnabelgrößen auf Grund eines reichhaltigeren Nahrungsangebots wieder zurück in eine Art.

Diese Art von Dynamik und Artenbildung kann man in einem Evolutionsbaum, wie in Abb. 1.15.a darstellen. Aber wie können solch unterschiedliche Reaktionen auf einen einzelnen Einfluss folgen? Einen Hinweis gibt die Evolutionstheorie von Kimura [100]. Er geht davon aus, dass viele Gene inaktiv sind und die Mutationsrate konstant über das gesamte Genom ist. Daher können viele Mutationen still, d. h. ohne Effekt, auftreten und erst später „aktiviert“ werden, so dass dann die Mutation sich manifestiert. Eine Mutation, die viele stille Mutationen aktiviert, führt letztlich zu einer vielfältigen Reaktion.

Ein weiteres, in realen Systemen eher selten vorkommendes Szenario ist der Evolutions-Kollaps. Anstelle der kontinuierlichen Produktion von neuen Arten, führt diese Art der Entwicklung zu wenigen überlebenden Arten, von denen nach und nach weitere aussterben, bis letztlich auch die letzte stirbt. Ein Beispiel für eine solche Entwicklung ist ein umgekippter

See.\* Jedoch führt die Konstruktion vieler Beispielsysteme für (chemische) Evolution meist zu einer Art, die letztlich als am besten angepasste gesehen wird und als einzige überlebt [101]. Damit stellen diese beiden Beispiele einen Evolutions-Kollaps dar; vgl. Abb. 1.15.b.

### Darwinsche Eigenschaften des vorgestellten Systems

Um die verschiedenen möglichen Verläufe der Evolution zu untersuchen, stellen wir in dieser Arbeit ein molekulares System vor. Es ist reduziert auf die wesentlichen Eigenschaften, die für Evolution benötigt werden: Reproduktion und Variation. Wie in Kapitel 1.4.4 schon beschrieben, erfolgt die Reproduktion autokatalytisch und die Variation durch Entstehung von (neuen) Polymeren.

Die Polymere kodieren keine Information, so dass die Unterscheidung verschiedener Spezies über den Phänotyp erfolgt. Die Polymerlänge ist demnach die definierende Eigenschaft der Spezies.

Doch wie können in einem solchen System evolutionäre Eigenschaften auftreten? Es sind nicht die Moleküle selbst, sondern deren zeitliche Abfolge, die als Phänotyp die grundlegenden Eigenschaften der Evolution zeigen. Diese sind: Mutation, Selektion und (zumindest anfängliche) Koexistenz, so dass die tauglichsten selektiert werden, also ein „Überleben der Tauglichsten“ (am besten angepassten) [1].\*\* In Form einer Längenselektion, nämlich der Längenverdopplung, zeigen wir in den folgenden beiden Kapiteln, dass durch Mutation neue Arten (Längen) von Molekülen entstehen können und, dass dies durch einen Koexistenzkampf von mehreren Molekülen stattfindet und sich damit das Molekül mit der doppelten Länge als best-angepasstes am stärksten vermehrt. Aber auch die anderen Moleküle können (und werden) miteinander reagieren und es stellt sich eine Koexistenz der verschiedenen Moleküle ein. Ist die Wachstumsrate groß genug, können sich die längenverdoppelten Moleküle gegenüber den anderen durchsetzen, weil deren Wachstumsrate im entscheidenden Zeitraum größer ist als diejenige aller anderen. Bei zu geringer Wachstumsrate sind alle Reaktionen so stark miteinander verwoben, dass keine Vorzugsreaktionen auszumachen sind und praktisch alle möglichen Polymere entstehen. Diese Verwobenheit vergleicht Schrödinger [22] mit einem Meisterwerk von Stickarbeit, z. B. die Bildwirkereien (Teppichkartons) von Raphael.

---

\* Wenn nährstoffreiche (Ab-)Wässer, die typischerweise phosphatreich sind, in einen See gelangen, dann vermehren sich aufgrund dessen die Pflanzen (insbesondere Algen) sehr stark, so dass es zunächst tagsüber zu einer hohen und nachts zu einer niedrigen Sauerstoffkonzentration im See kommt. Durch diese **Eutrophierung** verbraucht der Abbau der abgestorbenen Pflanzen jedoch viel Sauerstoff, so dass es letztlich zu einem starken Sauerstoffmangel insbesondere in den unteren Wasserschichten kommt und anaerobe Bakterien die Zersetzung der organischen Stoffe übernehmen und der See umkippt. [1, 14] Dieser Prozess ist durch die schrittweise Einschränkung der Artenvielfalt gekennzeichnet, bis letztlich nur die anaeroben Bakterien sehr langsam die Biomasse zersetzen.

\*\* Die genaue Definition von **darwinscher Evolution** ist objektiv gesehen umstritten. Je nach Blickwinkel des Definierenden spielen Erkenntnisse, die erst später erlangt wurden, eine größere Rolle und es wird z. B. zwingend ein Genom vorausgesetzt, was bei biologischen Systemen jedoch nicht zu Widersprüchen führt. Hier werden wir aber auf die ursprünglichen Erkenntnisse Darwins zurückgreifen: Durch Mutation (Veränderung) werden (teilweise sichtbare, eventuell auch [orts-]abhängige) Eigenschaften [1] vererbt, die dann für die Nachkommen, die variieren, einen Vorteil darstellen, so dass sich eben diese selektiv, d. h. vorteilhaft, vermehren.

## 1 Eigenschaften der biologischen Evolution

Wir können also zum einen eine fortschreitende Selektion (mit Koexistenz) feststellen, die vergleichbar zu Abb. 1.15.a ist. Zum anderen, im Fall einer kleinen autokatalytischen Rate, entsteht eine Vielzahl an Polymerlängen, wie es in Abb. 1.15.c dargestellt ist.

### Zusammenfassung

In diesem Abschnitt haben wir eine Übersicht über mögliche Verläufe von Evolution gegeben. Keine Evolution findet bei der wahllosen Wucherung statt, bei der alles mögliche entstehen kann; sie könnte aber im Zustand *Progenote* den Ursprung der modernen Zelle darstellen [83, 84, 85]. Künstliche evolvierende Systeme zeigen meist einen Evolutionskollaps, bei dem letztlich nur eine Spezies überlebt [101]. Molekulare Modellsysteme, die ein vergleichbares Verhalten zur natürlichen Evolution zeigen, sind derzeit unbekannt.

Es wurden grundlegende Eigenschaften der darwinschen Evolution vorgestellt und diese in den Kontext des hier vorgestellten Systems übertragen. Ausgehend von der Gaia-Hypothese [98] und der Nischenbildung [91], ordnen wir den Selektions- und Variationsmechanismus des Systems ein.

In den folgenden beiden Kapiteln stellen wir die experimentelle Realisierung und die theoretische Beschreibung vor.

## 2 Experimentelle Realisierung mit DNA

Im Folgenden stellen wir die experimentelle Realisierung<sup>\*</sup> eines evolvierenden Systems vor. Teile dieses Kapitels wurden in [82] und [102] veröffentlicht. Anlehnend an Abschnitt 1.3.2 werden vier unterschiedliche Experimente konstruiert, die zur Überprüfung der dargestellten Limitierungen dienen. In einem Experiment kann eine selektive Längenverdopplung beobachtet werden.

Zuerst stellen wir den allgemeinen Aufbau der Experimente vor und dann die einzelnen Experimente und deren Unterschiede. Ein erster Ansatz, in dem ein Doppelstrang (Länge 20 bp) ligiert wird, führt zu einer „wahllosen Wucherung“ (vgl. Abb. 1.5.c), bei der alle Längen, die entstehen können, auftreten. Dies ist möglicherweise auf eine unvollständige Hybridisierung mit Überhängen (Indexverschiebungen) zurückzuführen. Um dies zu überprüfen, wird in einem zweiten Experiment das Enzym Mung-Bean-Nuclease dazu benutzt, freie Einzelstränge in Nukleotide zu zerlegen. Diese ist auch bei niedrigen Konzentrationen so effektiv, dass spontan gebildete Einzelstränge zerlegt werden, so dass keine Selektion von längeren Strängen stattfinden kann. In einem dritten Experiment werden mit den beiden Enzymen AfeI und BstUI an einem speziellen 10 bp-Doppelstrang die entstehenden, für Indexshifts typischen, heterogenen Stränge in ihre beiden ursprünglichen Teile zerlegt. Da heterogene Stränge während des gesamten Experimentverlaufs auftreten, stehen auch immer Monomerstränge für den weiteren Verlauf zur Verfügung. Es tritt keine Längenskalentrennung auf und es entstehen fast alle möglichen Stränge. In einem vierten Experiment werden neben den Monomeren (mehrere unterschiedliche 10 bp-Doppelstränge) und der Ligase spezielle Linker, das sind alle möglichen Produkte doppelter Länge, die den Informationsgehalt erhalten, hinzugefügt. Damit lässt sich die nötige Zeit- und Längenskalseparation erreichen und eine Längenverdopplung beobachten.

Wir beschreiben nun den allgemeinen Aufbau der Experimente, analysieren die Sequenzen bzw. konstruieren diese theoretisch. Ausgehend von diesen Berechnungen leiten wir eine Formel ab, die die Hybridisierung allgemein beschreibt und die wir im nächsten Kapitel benutzen werden. Am Ende präsentieren und diskutieren wir die Ergebnisse aus den entsprechenden Experimenten und beschreiben die Funktionsweise von UNAFold.

Die Ergebnisse des Abschnitts 2.2 wurden teilweise und die des Abschnitts 2.3 größtenteils in [82] veröffentlicht.

---

\* Die Experimente wurden von Emanuel Gregor Worst in der Arbeitsgruppe von Albrecht Ott durchgeführt [102]. Das Design der Experimente, Sequenzen und der verwendeten Enzyme stammen ebenfalls von ihm. Mein Beitrag bestand aus dem Designen der Linker und der damit verbundenen Sequenzen sowie der allgemeinen Analyse der Sequenzen, deren Hybridisierungseigenschaften und Schmelztemperaturen und des damit verbundenen allgemeinen Modells in Abschnitt 2.4.2. Die Experimente wurden von Emanuel Gregor Worst und mir ausgewertet.

## 2.1 Biotechnischer Aufbau

In diesem Abschnitt beschreiben wir zunächst den allgemeinen experimentellen Ablauf sowie die verwendeten Enzyme und den Reaktionspuffer. Danach stellen wir die Analyse mittels Gelelektrophorese und die Auswertung mit ImageJ [103] vor.

### 2.1.1 Grundsätzliche Verfahrensweise

Jedes Experiment wird in mehrfacher Ausführung parallel in PCR Soft Tubes in einem PCR-Thermozykler inkubiert. Dabei stellt jede PCR Soft Tube bei ihrer Entnahme eine Probe dar, die analysiert wird. Der Thermozykler ist ein Gerät zur gezielten Erhitzung und Abkühlung von PCR Soft Tubes, um z. B. eine PCR durchzuführen. Hier wird er genutzt, um eine LCR\*-Reaktion durchzuführen. Ein Template entsteht durch spontane Ligation oder wird hinzugegeben. Auf diesem können in der Sequenz passende Stränge hybridisieren und werden dann von der Ligase ligiert. Ligierte Stränge können ihrerseits wieder als Template fungieren. Um die entstandenen Doppelstränge wieder zu trennen und die Proben zu durchmischen, wird die Temperatur kurzzeitig erhöht und rasch wieder verringert. In einigen Experimenten werden zudem weitere Enzyme eingesetzt, die entweder auf Doppelstränge oder Einzelstränge einwirken. Das Temperaturprotokoll wird bei den jeweiligen Experimenten vorgestellt. Während die Enzyme hinzugegeben werden, werden die Proben eisgekühlt, schnell gevortexed und herunterzentrifugiert. Im Folgenden werden wir uns den verwendeten Enzymen zuwenden.

### 2.1.2 Taq-Ligase und Reaktionspuffer

Die Taq-Ligase [104] ist eine Ligase des Bakteriums *Thermophilus aquaticus*, welches aus einem Geysir des Yellowstone Nationalparks stammt. Sie ist hitzestabil bei einer Halbwertszeit von 40 min bei 95 °C bzw. von 5 min bei 100 °C. Für eine optimale Aktivität benötigt sie  $Mg^{2+}$ -Ionen und 80 °C [105]. Sie verbindet zwei auf einem Strang hybridisierte, aneinanderstoßende Einzelstränge. Auf Grund der im Folgenden durchgeführten Experimente konnte nachgewiesen werden, dass die Ligase ebenfalls in der Lage ist, zufällig zusammentreffende Einzelstränge mit einer geringeren Rate zu ligieren\*\* . Die Aktivität der Taq-Ligase ist temperaturabhängig; vgl. Abb. 2.1.

Der benötigte Reaktionspuffer ist in Tabelle 2.1 aufgelistet. Er stellt Energie in Form von NAD für die Ligase bereit und puffert eine reaktionsgünstige Umgebung. Damit keine Tropfen im Reaktionsgefäß entstehen, wird ein leichtes Benetzungsmittel verwendet. Die verwendeten Salze stellen eine bekannte Metallionenkonzentration ein, welche einen Einfluss auf die Bindungsenergie der Doppelhelix hat [106]. In diesem Puffer wurden alle Experimente durchgeführt.

---

\* Die LCR ist eine mit der PCR verwandte Reaktion, bei der auf einem Template keine Nukleotide, sondern längere DNA-Stränge mit Hilfe einer Ligase, wie z. B. der Taq-Ligase miteinander verknüpft, d. h. ligiert, werden [107].

\*\* Eine **zufällige Ligation** ist möglich, weil spontan neue Stränge in dem zufälligen Ligations-Experiment entstehen. Dieses enthält keine Templates unter den benutzten Sequenzen, so dass zwei von diesen darauf hybridisieren könnten, um auf diesem Weg ligiert zu werden.

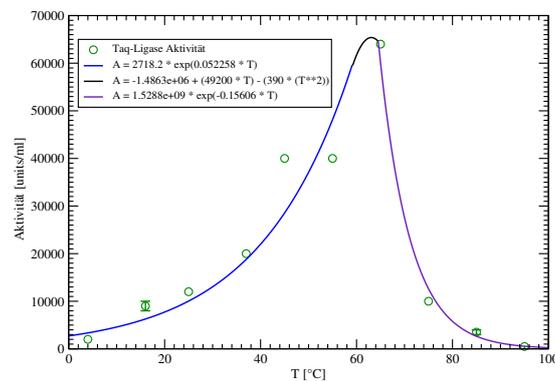


Abbildung 2.1: **Temperaturabhängigkeit der Taq-Ligase.** Die Taq-Ligase hat eine temperaturabhängige Aktivität, die im Temperaturverlauf langsam ansteigt und nach 65 °C rasch abfällt. Datenwerte aus [108]. Für die theoretischen Untersuchungen in Abschnitt 3.5.1 wurde die Aktivität durch eine stückweise definierte Funktion angefitet.

Konzentration	Stoff	Erläuterung
20 mM	Tris-HCl	biochemischer Puffer [109]
25 mM	Kaliumacetat	Salz, liefert K <sup>+</sup> -Ionen
10 mM	Magnesiumacetat	Salz, liefert Mg <sup>2+</sup> -Ionen
10 mM	Dithiothreitol	Proteinkonservator [110]
1 mM	NAD	Cofaktor für die Taq-Ligase [105, 111]
0,1 %	Triton X-100	Benetzungsmittel [112]

Tabelle 2.1: **Reaktionspuffer der Taq-Ligase.** Tris-HCL ist eine schwache Säure mit Pufferkapazität, um den pH-Wert (7,4 bei 25 °C) konstant zu halten [109]. Dithiothreitol ist eine Verbindung, die Proteine in ihrer funktionellen Form konserviert [110]. NAD ist, wie ATP, ein Cofaktor, der die benötigte Energie zur Reaktion (Ligation) liefert [105, 111]. Triton X-100 ist ein schwaches, nicht denaturierend wirkendes Benetzungsmittel, welches die Aktivität der meisten Enzyme bewahrt. Diese Stoffe dissoziieren keine Metallionen, die die Struktur der DNA verändern könnten [75]. Die einzigen Ionen sind die K<sup>+</sup>- und Mg<sup>2+</sup>-Ionen der beiden Acetate.

### 2.1.3 Mung-Bean-Nuclease

Die Mung-Bean-Nuclease [113] ist ein Enzym, welches Einzelstrang-DNA in ihre Nukleotide zerlegt und aus Mungbohnenkeimlingen stammt. Für hybridisierte DNA bedeutet das, dass Überhänge zerlegt und abgeschnitten sowie Hairpins gespalten werden. Der Nucleasepuffer enthält  $Zn^{2+}$ -Ionen, welche zur Reaktion notwendig sind und auch nach Verdünnung in einer ausreichenden Konzentration vorliegen, sowie  $Na^+$ -Ionen und hat einen pH-Wert von 5,0. Der Einfluss der Natriumionen ist vergleichbar mit den Kaliumionen und der pH-Wert stellt sich wegen der Pufferung auf den der von der Taq-Ligase benötigten ein. Damit arbeitet die Mung-Bean-Nuclease möglicherweise ineffizienter. Die Mung-Bean-Nuclease kann durch Erhitzen inaktiviert werden\* [114] und wird daher im Reaktionsverlauf hinzugefügt.

### 2.1.4 AfeI und BstUI

AfeI und BstUI sind Restriktionsenzyme [115, 116]. AfeI stammt ursprünglich aus dem Bakterium *Alcaligenes faecalis* und BstUI stammt aus dem Bakterium *Bacillus stearothermophilus U458*. Restriktionsenzyme schneiden doppelsträngige DNA an bestimmten Stellen. AfeI schneidet die Sequenz ...AGC $\nabla$ GCT... und BstUI die Sequenz ...CG $\nabla$ CG... an den angegebenen Stellen. Dies sind antiparallele Sequenzen. Bei geeignet gewählten Sequenzen können damit Hetero(di)mere geschnitten werden, während Homo(di)mere unbeeinflusst bleiben. Die Reaktionspuffer sind kompatibel und vergleichbar zum Taq-Ligase-Puffer. BstUI ist hitzestabil [117] aber AfeI nicht [118]; daher wird während des Reaktionsverlaufs mehr AfeI dazugegeben.

### 2.1.5 Gelelektrophorese und Auswertung

Nach der Probenentnahme werden diese mit einem Ladepuffer versehen und bei  $-20^\circ C$  eingefroren. Die Proben werden mit einer Polyacrylamid-Gelelektrophorese\*\* analysiert und unter UV-Licht abfotografiert. Anhand der Schwärzung und einer Kalibrierungskurve für verschiedene Längen und bekannte Konzentrationen kann durch Integration und Hintergrundsubtraktion auf die Konzentration rückgeschlossen werden. Diese Quantifizierung wird mit ImageJ [103] durchgeführt. Die Konzentration wird gegenüber der Zyklenzahl aufgetragen und analysiert.

---

\* Es sei darauf hingewiesen, dass Erhitzen keine geeignete Methode ist, die **Mung-Bean-Nuclease** zu **inaktivieren** [114]: Bevor die Nuclease thermisch inaktiviert wird, beginnt die DNA zu „atmen“, d. h. sie geht zumindest teilweise vom hybridisierten in den dissoziierten Zustand über. Es wird stattdessen eine Phenol-Chloroform-Extraktion und anschließende Ethanol-Fällung empfohlen.

\*\* Die **Polyacrylamid-Gelelektrophorese** [1] ist ein Verfahren zur Auftrennung von geladenen Molekülen. Die Proben werden mit einem Farbstoff (Bromphenolblau) versetzt und in sog. Taschen (Ausparungen im Gel) eingefüllt. Als Grundlage dient ein Polyacrylamid-Gel. Zwischen den Taschen und dem gegenüberliegenden Ende wird eine Spannung von 140 V für 60 bis 90 min angelegt. In dieser Zeit wandern kurze DNA-Stränge schneller durch das Gel und werden so von längeren separiert. Der Vorgang wird mit Eiswasser gekühlt, um weitere Reaktionen und eine Zersetzung zu verhindern. Der Farbstoff dient der Abschätzung, wie weit die Auftrennung schon erfolgt ist. Danach werden die DNA-Stränge mit Ethidiumbromid gefärbt und mit UV-Durchlicht sichtbar gemacht und abfotografiert. Wir verwenden eine Gelkonzentration von 15 bis 20 %, um eine Aufspaltung von 5 bis 150 bp zu erhalten [119].

Strang	Sequenz
Z	GGCAT CTTAG TACAT ATCAG
$\bar{Z}$	CTGAT ATGTA CTAAG ATGCC
Kombinationen	Sequenz
$\overline{ZZ}$	GGCAT CTTAG TACAT ATCAG GGCAT CTTAG TACAT ATCAG
$\overline{\overline{ZZ}}$	CTGAT ATGTA CTAAG ATGCC CTGAT ATGTA CTAAG ATGCC
$\overline{ZZ} = \overline{\overline{\overline{ZZ}}}$	CTGAT ATGTA CTAAG ATGCC GGCAT CTTAG TACAT ATCAG
$\overline{\overline{ZZ}} = \overline{\overline{\overline{\overline{ZZ}}}}$	GGCAT CTTAG TACAT ATCAG CTGAT ATGTA CTAAG ATGCC

Tabelle 2.2: **Sequenz des Experiments „wahllose Wucherung“**. Es wurde eine Sequenz zufällig ausgesucht, deren Schmelztemperatur im Bereich der größten Ligaseaktivität liegt und die keine Sekundärstrukturen bildet. Die verschiedenen Hybridisierungsmöglichkeiten dieser Sequenz und der daraus resultierenden Sequenzen sind in Abb. 2.2 dargestellt.

### Zusammenfassung

In diesem Abschnitt haben wir den grundlegenden biotechnologischen Aufbau der Experimente beschrieben. Es wird eine LCR-Reaktion mit der Taq-Ligase und deren Puffer durchgeführt. Um Überhänge zu schneiden oder die daraus entstehenden Produkte zu degradieren, werden zusätzlich die Mung-Bean-Nuclease bzw. AfeI und BstUI benutzt. Analysiert werden die Proben mittels Gelelektrophorese.

Nun gehen wir dazu über, die in den Experimenten verwendeten Sequenzen und die Temperaturprotokolle zu erläutern.

## 2.2 Experimente: Sequenzen und Zyklenprotokoll

Im Folgenden stellen wir die benutzten Sequenzen und das Zyklenprotokoll vor. Die zyklisch variierten Temperaturen sind so gewählt, dass sie in etwa mit den von UNAFold [75, 76] vorhergesagten übereinstimmen. Für die verschiedenen Experimente wurden zufällige Sequenzen gewählt, die gegebenenfalls wegen der eingesetzten Enzyme spezifische Besonderheiten in der Sequenz aufweisen oder entsprechend zufällig gewählt wurden, so dass sie nicht gegenseitig aufeinander hybridisieren.

Zunächst stellen wir das Experiment „wahllose Wucherung“ vor, in dem nur ein DNA-Strang und die Ligase benutzt werden.

### 2.2.1 Wahllose Wucherung (Taq-Ligase)

Das Experiment „wahllose Wucherung“ beginnt mit 20 bp DNA-Strängen mit der Sequenz Z; vgl. Tab. 2.2. Dies ist eine zufällig ausgewählte Sequenz, deren Schmelztemperatur im Bereich der größten Ligaseaktivität liegt und nicht in sich indexverschoben hybridisiert.

Die Proben werden für 15 min bei 60 °C inkubiert, dann für 1 min bei 66 °C. Nach einer weiteren 15 min Phase bei 60 °C wird die Temperatur für 1 min auf 84 °C erhöht. Angefangen mit einem Denaturierungsschritt für 1,5 min bei 84 °C, wird der oben beschriebene Zyklus wiederholt. Nach je 1 oder 2 Zyklen wird eine Probe entnommen.

Zunächst betrachten wir die möglichen Kombinationen des Strangs Z und dessen Komplement, wie sie im ersten Reaktionsschritt auftreten können, um Vorhersagen über den weiteren Verlauf des Experiments treffen zu können. In Tabelle 2.2 sind alle Kombinationsmöglichkeiten aufgelistet. Die Homodimere ZZ und  $\overline{ZZ}$  sind zueinander komplementär, während die beiden Heterodimere  $Z\overline{Z}$  und  $\overline{Z}Z$  jeweils selbstkomplementär sind. Wir erwarten, dass ein signifikanter Anteil der Heterodimere auf sich selbst hybridisiert, also in gefalteter Form vorliegt. Da hierzu nur ein Molekül vorliegt und diese Situation energetisch ähnlich zu einer Hybridisierung mit einer Teilüberlappung ist, wird erstere bevorzugt. Alle möglichen Typen von binären Reaktionen sind in Abbildung 2.2 aufgelistet. Der Ausgangsstrang und die Homodimere hybridisieren ohne einen signifikanten Anteil von Faltung (Abb. 2.2.e-f). Die Heterodimere können sich gut falten und es tritt dementsprechend auch ein großer Anteil an Faltungen unterhalb der Schmelztemperatur auf (Abb. 2.2.c). Die unterschiedlichen Heterodimere beeinflussen sich nicht und reagieren parallel nur mit sich selbst (Abb. 2.2.a). Heterodimere und Homodimere zeigen ein vergleichbares Verhalten (Abb. 2.2.b). Um den Einfluss der Überhangbildung zu verdeutlichen, wurde der Stoffmengenanteil der verschiedenen, überhangbildenden Hybridisierungstypen aus der binären Hybridisierung aus den Abb. 2.2.a-b in Abb. 2.2.d zusammengefasst. Für selbstkomplementäre Stränge sind die Wahrscheinlichkeiten untereinander zu hybridisieren geringer als für ein Homo- und ein Heterodimer und bei einer niedrigeren Temperatur. Diese Abbildung ist allenfalls als eine Abschätzung zu interpretieren, da kombinatorische und sterische Effekte eine volle Rechnung mit drei Strängen erfordern, welche von UNAFold nicht bereitgestellt wird [75]. Da sich möglicherweise auch andere Strukturen ausbilden könnten, die die Hybridisierungseigenschaften der DNA verändern, wie z. B. die Triple-Helix-Bildung [79], ist eine solche Berechnung nur in Ausnahmefällen einfach zu realisieren. Des Weiteren sind die Konformationen, die entstehen können, entsprechend eingefärbt dargestellt. Man sieht, dass die Homo-Heterodimerhybridisierung eine Überlappung bilden kann, die nicht zu ligierbaren Überhängen führt und daher nicht autokatalytisch verlängert werden kann. Welchen Einfluss die Überhangbildung auf die Bildung von Strängen hat, untersuchen wir im folgenden Abschnitt. Danach gehen wir über, den Einfluss von Hetero(di)meren zu untersuchen.

Die Ergebnisse der „wahllosen Wucherung“ werden in Abschnitt 2.3.1 vorgestellt.

### 2.2.2 Einzelstrangzerlegung (Taq-Ligase und Mung-Bean-Nuclease)

Um den Einfluss der Überhänge zu analysieren, werden in diesem Experiment mit Hilfe der Mung-Bean-Nuclease die überhängenden Einzelstränge in ihre Nukleotide zerlegt. Die Sequenz ist die gleiche wie beim vorherigen Experiment.

Der Zyklus ist ebenfalls derselbe mit einem weiteren Schritt bei 30 °C für 5 min. Ab dem sechsten Zyklus wird ein anderes Schema benutzt: 30 min bei 30 °C, 15 min bei 60 °C, 1 min bei 66 °C, 15 min bei 60 °C, 1 min bei 84 °C und 5 min bei 30 °C. Am Anfang eines solchen Zyklus wird eine Einheit Mung-Bean-Nuclease hinzugefügt.

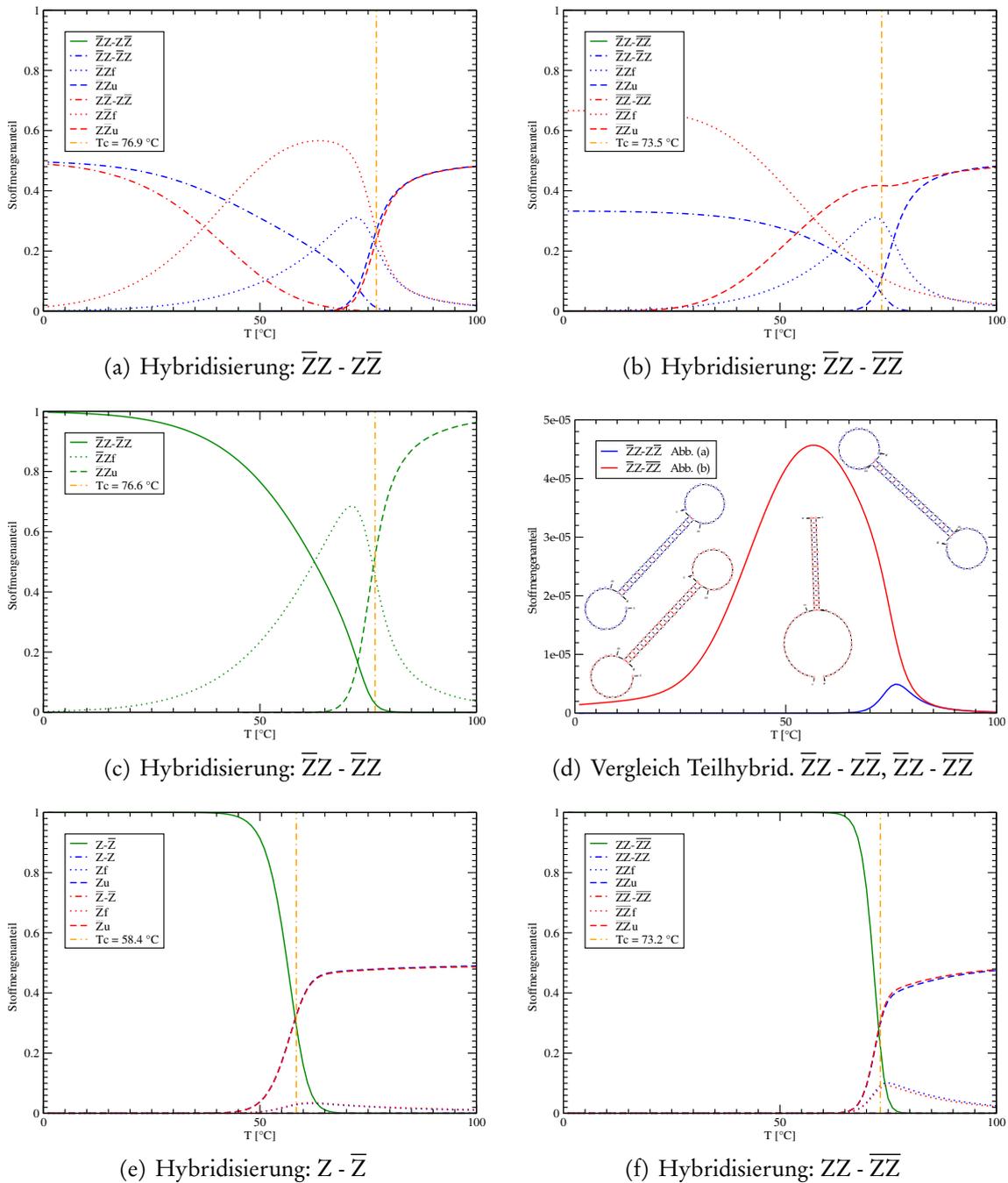


Abbildung 2.2: **Hybridisierungsverhalten des Experiments „wahllose Wucherung“** für (teil-) komplementäre Sequenzkombinationen. Abbildungen (a) und (b) zeigen, dass es bei teilüberlappenden, selbstkomplementären Sequenzen kaum zu einer teilweisen Crosshybridisierung kommt, sondern stattdessen die selbstkomplementären Stränge mit steigender Temperatur zunächst Dimere bilden, sich dann zu Monomeren falten und schließlich ungefaltet vorliegen. Der nicht selbstkomplementäre Strang  $\bar{Z}Z$  liegt nicht dimerisiert vor. Das Gemisch beeinflusst den Verlauf der Zustände nur marginal (c). Die Crosshybriden liegen, wie zu erwarten, genau teilhybridisiert vor (d). Die Farbe der Struktur entspricht der des Hybrids. Der Stoffmengenanteil ist jedoch sehr gering, so dass diese hier gesondert aus den Abb. (a) und (b) dargestellt werden. Für größere DNA-Konzentrationen steigt der relative Anteil (nicht dargestellt). Die Hybridisierung der komplementären Stränge (e) und (f) folgt dem typischen Verlauf. Struktur und Stoffmengenanteil wurde mit UNAFold [75, 76] berechnet.

## 2 Experimentelle Realisierung mit DNA

Da die Mung-Bean-Nuclease auch bei höheren Temperaturen, bei denen die Doppelstränge getrennt vorliegen, aktiv ist [114], besteht die Möglichkeit, dass sie in dieser Phase zerlegt werden können. Die Ergebnisse dieses Experiments werden in Abschnitt 2.3.2 vorgestellt.

### 2.2.3 Heterodimerzerlegung (Taq-Ligase, AfeI und BstUI)

Eine andere Art Überhänge zu unterdrücken, ist sie im bereits ligierten, hybridisierten Zustand wieder zu trennen und die dadurch entstandenen Hetero(di)mere wieder zu spalten. Dies ist die Aufgabe der Restriktionsenzyme AfeI und BstUI, falls geeignete Sequenzen gewählt wurden (vgl. Kap. 2.1.4). Daher suchen wir folgende Sequenz:  $Q = \text{CGTTA TTAGC}$  und deren Komplement  $\overline{Q} = \text{GCTAA TAACG}$  aus. Damit werden die Heterodimere  $Q\overline{Q}$  und  $\overline{Q}Q$  zerlegt, während die Homodimere unbeeinflusst bleiben.

Der Zyklus hat zwei Temperaturen und stellt sich wie folgt dar: 13,5 min bei 34 °C und 1 min bei 60 °C. Nach 20 Zyklen wurden die Temperaturen auf 60 °C und 76 °C erhöht, jeweils am Anfang mit einem Denaturierungsschritt bei 60 °C bzw. 76 °C für 1,5 min. BstUI wurde nach 8, 12, 20 und 32 Zyklen hinzugegeben und AfeI zusätzlich auch nach Zyklus 15 und 18, sowie alle 2 Zyklen ab dem 22. Zyklus.

Bei langsamer Reaktionsgeschwindigkeit der Zerlegung kann es dazu kommen, dass keine Zeitskalenseparation mehr eintritt und verschiedene Längen  $n$  von  $X_n$  vorliegen, die dann miteinander reagieren können. Die Folge wäre, dass keine Längenselektion möglich ist.

Im Abschnitt 2.3.3 stellen wir die Ergebnisse zur Heterodimerzerlegung vor.

### 2.2.4 Längenselektion (Taq-Ligase und Linker)

Eine Zeitskalenseparation lässt sich durch eine schnellere spontane Bildung von neuen Molekülen erreichen. Eine geringe Zugabe der sich spontan bildenden Moleküle erfüllt diesen Zweck. Diese Lösung bereitet aber auch den Weg, um Moleküle zu generieren, die den vollen Informationsgehalt erhalten und weitergeben. Damit sind auch Indexverschiebungen aller Art und Überhänge weitestgehend unterdrückt. Mit einem solchen Experiment ließe sich daher eine Längenverdopplung realisieren.

Das Zyklenprotokoll für dieses Experiment ist kürzer. Einmalig wird die DNA 1 min bei 80 °C denaturiert. Der Zyklus startet dann mit 2 min bei 28 °C und für 1 min bei 80 °C. Nach 24 Zyklen werden 80 fmol der 40 bp Linker hinzugegeben und die Temperatur auf 55 °C und 90 °C erhöht. Nach 29 Zyklen werden 40 fmol der 80 bp Linker zugefügt und die Temperatur auf 74 °C und 94 °C respektive erhöht, um den geänderten Schmelztemperaturen Rechnung zu tragen.

Im Folgenden werden wir ein vereinfachtes Experiment beschreiben, welches dazu dient, die Linkersequenzen zu finden und die Temperaturen auszurechnen. Danach berechnen wir die Sequenzen, Linker und Temperaturen für das tatsächlich durchgeführte Experiment.

Spezies	Sequenz
A	TGTGC CTTTG ACACC TGATC
$\overline{A}$	GATCA GGTGT CAAAG GCACA
B	GGAGT TGGTT ATTTG GACCG
$\overline{B}$	CGGTC CAAAT AACCA ACTCC
C	CTTAC ACGAG CTACA TGACG
$\overline{C}$	CGTCA TGTAG CTCGT GTAAG
D	GGGAA GTTGT GAAGA CTAGC
$\overline{D}$	GCTAG TCTTC ACAAC TTCCC
Linker	Sequenz
$\hat{\mathcal{L}}_{AB}$	AACCA ACTCC GATCA GGTGT
$\hat{\mathcal{L}}_{\overline{BC}}$	ATTTG GACCG CTTAC ACGAG
$\hat{\mathcal{L}}_{CD}$	ACAAC TTCCC CGTCA TGTAG
$\hat{\mathcal{L}}_{\overline{DA}}$	GAAGA CTAGC TGTGC CTTTG

Tabelle 2.3: **Sequenzselektion für das Experiment.** Es wurden aus zufällig generierten Sequenzen die am besten passenden herausgesucht (siehe Text). Die Linker  $\hat{\mathcal{L}}$  sind nicht identisch mit denen des Experiments; sie dienen lediglich dazu, die Temperaturen des Teilüberlapps einfacher zu quantifizieren.  $\hat{\mathcal{L}}_{XY}$  bedeutet hier, dass der Linker ein Template für die Ligation von Spezies X mit Spezies Y ist.

### Suche der Sequenzen und Temperaturoptimierung

Zur Optimierung der Sequenzen wurden 290.000 Sätze von Sequenzen mit festem CG zu AT-Verhältnis von 1 : 1 erzeugt, die dann weiter mit UNAFold analysiert wurden\*. Jeder Satz enthält 4 Sequenzen, deren Komplement und 4 Linker als Einzelstränge, die alle 20 bp lang sind. Die Linker verbinden je zwei unterschiedliche Einzelstränge, die dann als Template für ihr Komplement dienen. Aus den generierten Sätzen wurden diejenigen ausgeschlossen, deren Intersequenzschmelztemperaturen über 7 °C lagen. Aus diesen wurde eine Sequenz ausgewählt, die eine geringe Temperaturspanne sowohl zwischen den Linkern und den Sequenzen als auch unter den Sequenzen aufweist. Damit wird gewährleistet, dass die Reaktionsgeschwindigkeiten der verschiedenen, parallel ablaufenden Reaktionen vergleichbar sind und die Produkte daher in ähnlichen Konzentrationen vorliegen. Der ausgewählte Sequenzsatz und dessen Schmelztemperaturen sind in Tabelle 2.3 und 2.4 aufgelistet.

\* Dazu wurde zum einen das Programm pwgen mit den Optionen `-s -0` benutzt, die nicht benötigten, anderen Zeichen gelöscht, und Sequenzen mit dem passenden Verhältnis heraussortiert. Die Optionen stellen sicher, dass zufällige Zeichenketten generiert werden. Zum anderen wurde das Programm `me1t.pl` aus dem UNAFold-Paket [75, 76] mit den Optionen `-n DNA -N 0.025 -M 0.01 -C 0.000001` benutzt. Die Salz- und DNA-Konzentrationen sind vergleichbar zum Experiment:  $\text{Na}^+$ : 25 mM,  $\text{Mg}^{2+}$ : 10 mM und Gesamt-DNA-Konzentration: 1  $\mu\text{M}$ .

Sequenz	Template	Schmelztemperatur	Mittelwert Standardabweichung
zweite Reaktionsstufe – Linker			
A	$\hat{\mathcal{L}}_{AB}$	40,2°C	$\bar{T} = 40,3^\circ\text{C}$ $\sigma = 0,94^\circ\text{C}$
B	$\hat{\mathcal{L}}_{AB}$	39,5°C	
C	$\hat{\mathcal{L}}_{CD}$	39,7°C	
D	$\hat{\mathcal{L}}_{CD}$	40,5°C	
$\bar{B}$	$\hat{\mathcal{L}}_{\bar{B}\bar{C}}$	41,6°C	
$\bar{C}$	$\hat{\mathcal{L}}_{\bar{B}\bar{C}}$	40,2°C	
$\bar{A}$	$\hat{\mathcal{L}}_{\bar{D}\bar{A}}$	41,9°C	
$\bar{D}$	$\hat{\mathcal{L}}_{\bar{D}\bar{A}}$	39,0°C	
zweite Reaktionsstufe – Komplemente			
A	$\bar{A}$	65,2°C	$\bar{T} = 64,2^\circ\text{C}$ $\sigma = 0,66^\circ\text{C}$
B	$\bar{B}$	64,3°C	
C	$\bar{C}$	63,7°C	
D	$\bar{D}$	63,5°C	

Tabelle 2.4: **Schmelztemperaturen der Sequenzen**, wie sie zur Selektion der Sequenzen (siehe Tabelle 2.3) herangezogen wurden. Sowohl die Teilüberlappung (Hybridisierung mit den Linkern) als auch die Vollüberlappung sollten eine nicht allzu große Streuung aufweisen, damit die parallel ablaufenden Reaktionen mit vergleichbaren Geschwindigkeiten ablaufen (siehe Text).

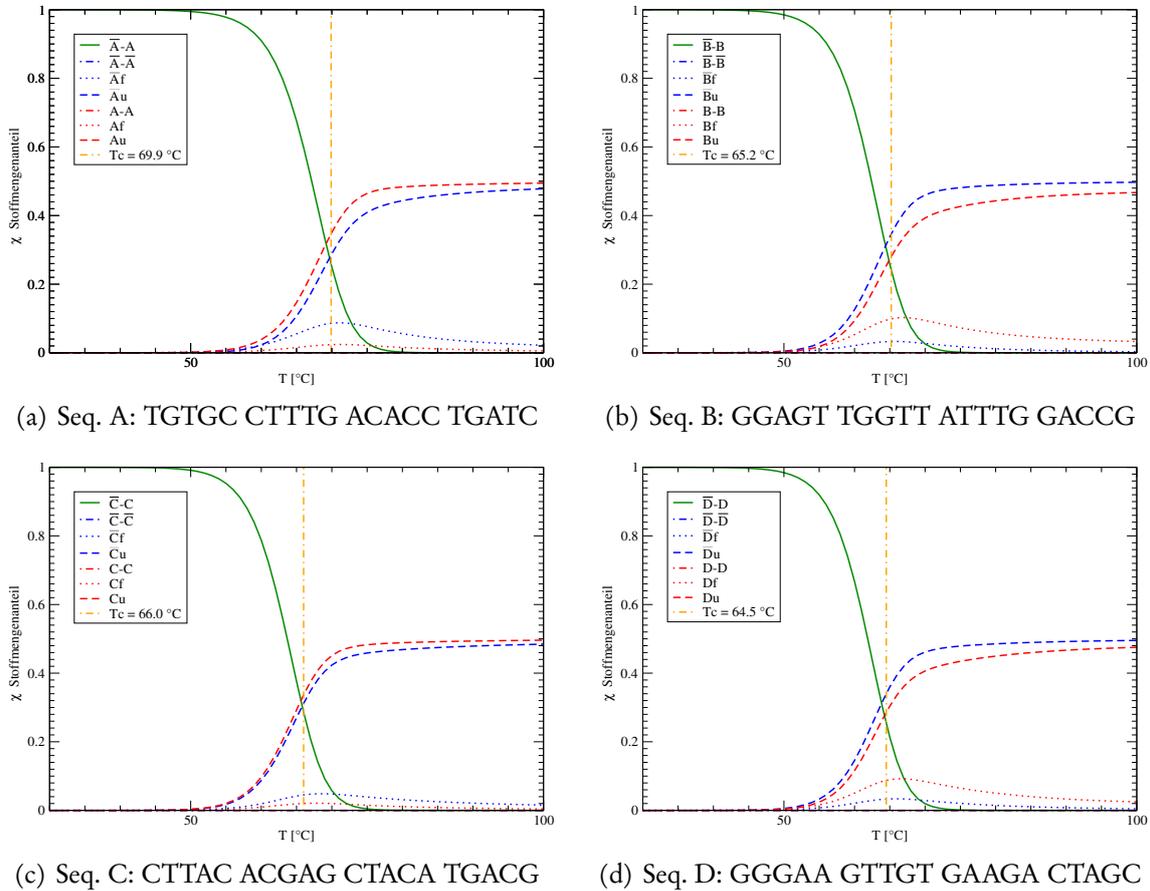


Abbildung 2.3: **Stoffmengenanteil der ausgewählten DNA-Sequenzen in Abhängigkeit von der Temperatur.** Das Schmelzverhalten der ausgewählten Sequenzen und deren Komplemente ist ähnlich. Bei niedrigen Temperaturen liegt eine vollständige Hybridisierung vor ( $\chi_{X_c, X} = 1$ ;  $X \in \{A, B, C, D\}$ ), während bei hohen Temperaturen die Stränge fast vollständig entfaltet sind ( $\chi_{X_u} \approx \frac{1}{2} \approx \chi_{X_u}$ ). Der Anteil der Stränge, die gefaltet sind ( $\chi_{X_f} \leq 0, 1$ ;  $\chi_{X_f} \leq 0, 1$ ), ist gering und Selbsthybridisierung ( $\chi_{X_c, X_c}$ ;  $\chi_{X, X}$ ) kommt praktisch nicht vor. Damit sind die ausgewählten Stränge gut geeignet, um das Reaktionsschema umzusetzen. Die jeweiligen Schmelztemperaturen ( $T_c$ ) sind ebenfalls vergleichbar: Mittelwert:  $66,4^\circ\text{C}$ , Standardabweichung:  $2,09$ .  $T_c$  ist diejenige Temperatur, bei der die Gesamtdimerkonzentration bei niedrigen Temperaturen halbiert ist.

## 2 Experimentelle Realisierung mit DNA

Eine weitere mögliche Einschränkung wäre eine ungleiche Verfügbarkeit der DNA-Stränge. Diese liegen, wie aus Abb. 2.3 hervorgeht, maximal nur zu einem Stoffmengenanteil von 10 % in der gefalteten Form vor, was keinen erheblichen Unterschied darstellt und einen parallelen Reaktionsverlauf nicht behindert.

Im Folgenden beschreiben wir die im Experiment tatsächlich verwendeten Sequenzen und Linker.

### Optimierte Linker und volle kombinatorische Analyse

Die im vorherigen Abschnitt besprochenen Sequenzen und Linker weisen im Hinblick auf die experimentelle Realisierung entscheidende Nachteile auf: Zum einen sind die Linker Moleküle, die im weiteren Reaktionsverlauf nicht mehr vorkommen und daher bei den autokatalytischen Reaktionen nicht verbraucht werden und im weiteren Reaktionsverlauf nur spontan ligieren können, was sich unter Umständen negativ auswirken könnte. Daher wählen wir hier Linker, die doppelsträngig vorliegen und dem zu ligierenden Produkt entsprechen. Damit ist sichergestellt, dass diese nicht als parasitäre Moleküle im Gemisch verbleiben und an der weiteren Reaktion teilnehmen können. Zum zweiten werden die anfänglichen Sequenzen auf eine praktikable, minimale Länge von 10 bp gekürzt\*. DNA hat eine Persistenzlänge von ca. 150 bp und ist auch ungefähr ab dieser Länge in der Lage in Ringen ligiert zu werden [121]. Damit stellt sie kein gutes Modell für ein starres, lineares Polymer dar, so dass wir davon absehen, Längen über 160 bp zu betrachten und daher Linker bis zu einer Länge von 80 bp verwenden.

In Tabelle 2.5 sind die im Linker-Experiment verwendeten Sequenzen samt den zugehörigen Linkern aufgelistet. Die Linker werden mit  $\mathcal{L}$  bezeichnet und dienen als Template für die jeweiligen komplementären Sequenzen und liegen als Doppelstrang vor. Die längst-mögliche Sequenz ist ABCD. Sie kann dann im nächsten Schritt nur noch rein spontan mit den noch vorhandenen Molekülen ligiert werden; im Idealfall mit sich selbst, so dass als längstes Molekül ABCDABCD entsteht mit den Nachteilen aus dem Experiment „wahllose Wucherung“ und der Möglichkeit, andere als lineare Strukturen zu bilden.

Tabelle 2.6 listet alle Hybridisierungstemperaturen auf, die im Experiment nach Konstruktion vorkommen. Die Temperaturen der exakten Komplemente liegen nahe beieinander, mit einer Standardabweichung von  $\sigma < 1,92^\circ\text{C}$ , die für längere Stränge kleiner wird. Die Temperaturen der Linker mit den jeweiligen Einzelsträngen ist geringfügig größer, da der Entropiegewinn geringfügig größer ist, was sich auf kleinere Temperaturen stärker auswirkt. Um die berechneten Temperaturen herum wurden dann die kalten und heißen Grenztemperaturen des Zyklus festgelegt. Durch entsprechende Wahl der niedrigen Temperatur, ist es möglich, oberhalb der Schmelztemperatur der kürzeren Stränge zu bleiben und diese damit nicht mehr templatebasiert, sondern allenfalls nur noch spontan reagieren zu lassen.

Die Ergebnisse der Längenselektion analysieren wir in Abschnitt 2.3.4.

---

\* Längen ab 5 bp sind auf einem Polyacrylamid-Gel noch zu erkennen [119], aber benötigen eine größere Konzentration als die letztlich im Experiment verwendete von weniger als 140 nM Gesamtkonzentration bei 10 bp, welche noch ausreichend gut zu quantifizieren ist. Des Weiteren brauchen Ligasen eine gewisse Erkennungslänge, um DNA zu ligieren, welche ebenfalls bei ca. 5 bp pro zu ligierendem Teilstrang liegt [120].

Sequenzen der ersten Reaktionsstufe		
Bezeichnung	Sequenz	
a <sub>1</sub>	TGTGC CTTTG	
a <sub>2</sub>	ACACC TGATC	
b <sub>1</sub>	GGAGT TGGTT	
b <sub>2</sub>	ATTTG GACCG	
c <sub>1</sub>	CTTAC ACGAG	
c <sub>2</sub>	CTACA TGACG	
d <sub>1</sub>	GGGAA GTTGT	
d <sub>2</sub>	GAAGA CTAGC	
Sequenzen der zweiten Reaktionsstufe		
Bez.	Linker	Sequenz
A = a <sub>1</sub> a <sub>2</sub>	$\mathcal{L}a_1a_2$	TGTGC CTTTG ACACC TGATC
B = b <sub>1</sub> b <sub>2</sub>	$\mathcal{L}b_1b_2$	GGAGT TGGTT ATTTG GACCG
C = c <sub>1</sub> c <sub>2</sub>	$\mathcal{L}c_1c_2$	CTTAC ACGAG CTACA TGACG
D = d <sub>1</sub> d <sub>2</sub>	$\mathcal{L}d_1d_2$	GGGAA GTTGT GAAGA CTAGC
Sequenzen der dritten Reaktionsstufe		
Bez.	Linker	Sequenz
AB	$\mathcal{L}AB$	TGTGC CTTTG ACACC TGATC GGAGT TGGTT ATTTG GACCG
CD	$\mathcal{L}CD$	CTTAC ACGAG CTACA TGACG GGGAA GTTGT GAAGA CTAGC
Sequenzen der vierten Reaktionsstufe		
Bez.	Linker	Sequenz
ABCD	$\mathcal{L}ABCD$	TGTGC CTTTG ACACC TGATC GGAGT TGGTT ATTTG GACCG CTTAC ACGAG CTACA TGACG GGGAA GTTGT GAAGA CTAGC

Tabelle 2.5: **Sequenzen des Linker-Experiments.** Alle Sequenzen, die im Experiment benutzt wurden ohne die zugehörigen Komplemente. Die mit Kleinbuchstaben bezeichneten Sequenzen sind die Ausgangssequenzen. Dabei entsteht nach Schema aus den Sequenzen  $x_1x_2$  ( $x \in \{a, b, c, d\}$ ) die Sequenz X. Großbuchstaben bezeichnen Sequenzen, die nach Schema entstehen. Linker werden mit  $\mathcal{L}$  bezeichnet und sind Moleküle, die im System entstehen können.  $\mathcal{L}XY$  besteht aus den Sequenzen X und Y und ist damit ein Template für  $\overline{Y}$  und  $\overline{X}$ .

2 Experimentelle Realisierung mit DNA

Template	Sequenz	Schmelztemperatur	Mittelwert Standardabweichung
erste Reaktionsstufe			
d <sub>2</sub>	$\overline{d_2}$	34,1 °C	$\overline{T} = 37,0^\circ\text{C}$ $\sigma = 1,92^\circ\text{C}$
c <sub>1</sub>	$\overline{c_1}$	35,1 °C	
c <sub>2</sub>	$\overline{c_2}$	35,3 °C	
a <sub>2</sub>	$\overline{a_2}$	36,7 °C	
b <sub>1</sub>	$\overline{b_1}$	37,7 °C	
d <sub>1</sub>	$\overline{d_1}$	37,7 °C	
b <sub>2</sub>	$\overline{b_2}$	39,1 °C	
a <sub>1</sub>	$\overline{a_1}$	40,0 °C	
LC	$\overline{c_1}$	35,8 °C	$\overline{T} = 39,2^\circ\text{C}$ $\sigma = 2,35^\circ\text{C}$
LD	$\overline{d_2}$	36,4 °C	
LC	$\overline{c_2}$	38,0 °C	
LA	$\overline{a_2}$	39,2 °C	
LB	$\overline{b_1}$	39,5 °C	
LD	$\overline{d_1}$	39,6 °C	
LB	$\overline{b_2}$	41,5 °C	
LA	$\overline{a_1}$	43,4 °C	
zweite Reaktionsstufe			
D	$\overline{D}$	63,5 °C	$\overline{T} = 64,2^\circ\text{C}$ $\sigma = 0,66^\circ\text{C}$
C	$\overline{C}$	63,7 °C	
B	$\overline{B}$	64,3 °C	
A	$\overline{A}$	65,2 °C	
LCD	$\overline{C}$	63,8 °C	$\overline{T} = 64,3^\circ\text{C}$ $\sigma = 0,54^\circ\text{C}$
LCD	$\overline{D}$	64,0 °C	
LAB	$\overline{B}$	64,2 °C	
LAB	$\overline{A}$	65,2 °C	
dritte Reaktionsstufe			
CD	$\overline{CD}$	78,3 °C	$\overline{T} = 79,0^\circ\text{C}$ $\sigma = 0,65^\circ\text{C}$
AB	$\overline{AB}$	79,6 °C	
LABCD	$\overline{CD}$	78,4 °C	$\overline{T} = 79,1^\circ\text{C}$ $\sigma = 0,65^\circ\text{C}$
LABCD	$\overline{AB}$	79,7 °C	
vierte Reaktionsstufe			
ABCD	$\overline{ABCD}$	86,6 °C	

Tabelle 2.6: **Schmelztemperaturen der Sequenzen** wie sie hauptsächlich im Experiment vorkommen. Aufgelistet sind die Schmelztemperaturen der Stränge mit ihren jeweiligen Komplementen und den Linkern. Für längere Stränge gibt es eine geringere Abweichung untereinander sowie zu den Linkern.

## Zusammenfassung

In diesem Abschnitt haben wir die Zyklenprotokolle der vier Experimente vorgestellt sowie die benutzten Sequenzen und deren Hybridisierungseigenschaften. Die Schmelztemperaturen der Sequenzen wurden mit UNAFold [75, 76] berechnet und es wurden Linker-Sequenzen zufällig generiert und die geeignetsten ausgewählt, welche eine geringe Streuung in der Schmelztemperatur aufweisen.

Nun gehen wir zu den Ergebnissen der jeweiligen Experimente über, um diese darzustellen und zu analysieren.

## 2.3 Ergebnisse

In diesem Abschnitt stellen wir die Ergebnisse der durchgeführten Experimente vor. Die „wahllose Wucherung“ führt, ausgehend von einem 20 bp-Strang, zum sukzessiven Auftreten aller möglicher Längen. Damit lässt sich keine Selektion erkennen. Die Einzelstrangzerlegung durch die Mung-Bean-Nuclease hat den Nachteil, dass alle Stränge bei Temperaturen, die größer als ihre Schmelztemperatur sind, als Einzelstrang vorliegen und daher ebenfalls zerlegt werden. Es lässt sich so also keine effektive Längenselektion bewerkstelligen. Die Heterodimerzerlegung führt zu einer größeren Verfügbarkeit der Ausgangsmoleküle einer Reaktion, die vormals aus einer Hetero(di)merreaktion stammen, so dass eine Zeitskalenseparation nicht eintritt und damit ebenfalls keine Längenselektion. Durch die Zugabe von Linkern kann die Zeitskalenseparation aufrechterhalten werden, so dass eine Längenselektion auftritt. Zunächst beschreiben wir die „wahllose Wucherung“.

### 2.3.1 Wahllose Wucherung

In Abbildung 2.4 ist das Ergebnis einer reinen Ligation mit der Taq-Ligase und des Doppelstrangs Z dargestellt. Es entstehen sukzessive jeweils um 20 bp längere Stränge mit abnehmender Konzentration. Der Konzentrationsverlauf für jede Länge nimmt zunächst zu, zeigt ein Maximum, welches für längere Stränge auch später und mit einer niedrigeren Konzentration auftritt, und nimmt schließlich ab. Wie wir bereits in Abschnitt 2.2.1 diskutiert haben, kann dies zum einen auf eine mögliche Überhangligation zurückgeführt werden: Es entstehen Stränge mit dem (End-)Motiv  $Z\bar{Z}$  bzw.  $\bar{Z}Z$ , wobei in diesem Fall Z bzw.  $\bar{Z}$  sowie das Motiv auch je als Repetition vorliegen können. Diese Motive ermöglichen zum einen eine Faltung aber auch eine Hybridisierung eines in der Länge nicht passenden Stranges, so dass beliebige Vielfache der Ausgangsstranglänge entstehen können; vgl. auch Abb. 2.2. Die Experimente Einzelstrangzerlegung und Heterodimerzerlegung dienen u. a. dem Nachweis dieser Hypothese. Zum anderen laufen die templatebasierenden Reaktionen im Vergleich zum Entstehungsprozess neuer Längen, z. B. spontan oder über den oben genannten Mechanismus der Überhangligation, zu langsam ab, um eine Zeitskalenseparation zu erreichen.

Ein vergleichbares Verhalten zeigt das theoretische Modell in Kapitel 3.2.1, welches wir dort analysieren.

Im Folgenden gehen wir der Frage nach, ob die Entfernung von Einzelsträngen, also insbesondere von Überhängen, ausreichend ist, um eine Längenverdopplung zu erreichen.

## 2 Experimentelle Realisierung mit DNA

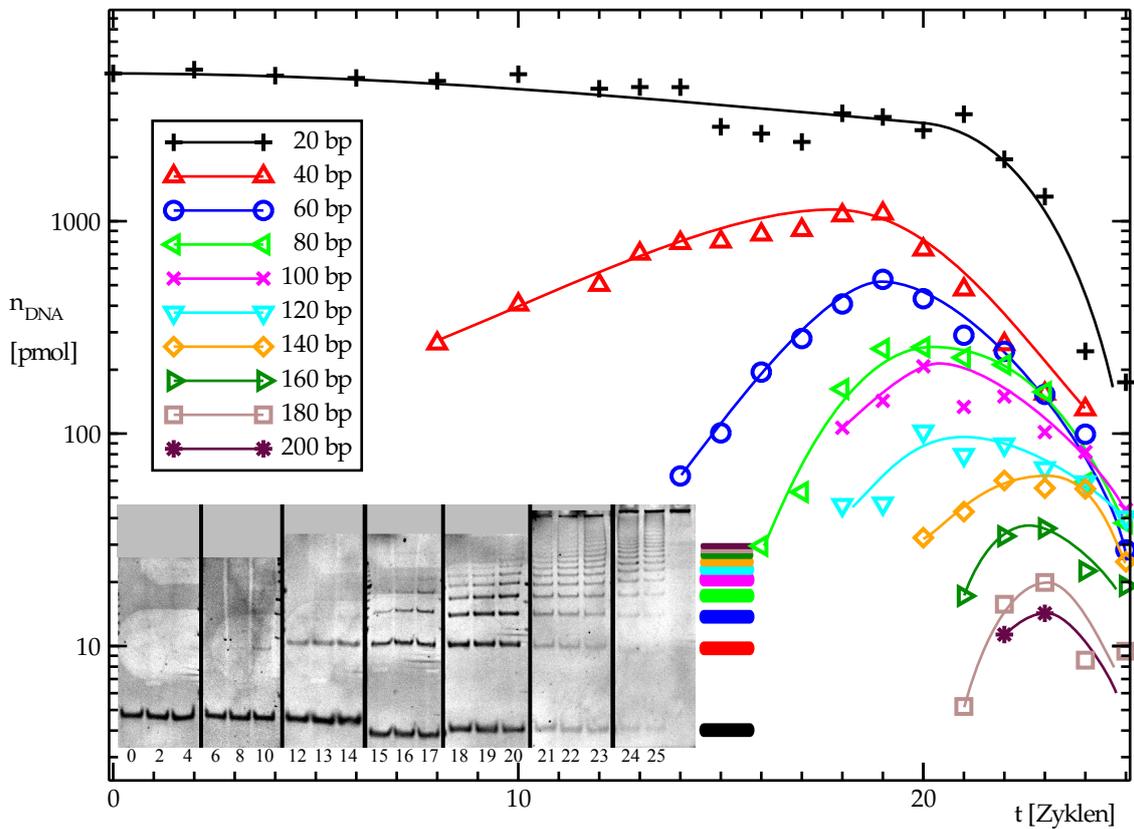


Abbildung 2.4: **Wahllose Wucherung.** Der ursprüngliche 20 bp DNA-Strang wurde zufällig gewählt, so dass keine Sekundärstrukturen entstehen, und dieser mit Taq-Ligase inkubiert. Die Zeit ist durch die Reaktionszyklen gegeben, deren Maximaltemperaturen bis über die Schmelztemperatur der längsten, jeweilig auftretenden Stränge reicht. Aus dem 20 bp DNA-Strang bilden sich sukzessive immer längere Stränge, deren Maximalkonzentration ebenso abnimmt. Alle messbaren Stränge haben eine Länge, die ein Vielfaches von 20 bp ist; vgl. Abb. 2.2. Die Kurven dienen der optischen Führung. Eingefügt sind die (zur besseren Darstellung bearbeiteten) Gele, die ausgewertet wurden. Veröffentlicht in [82].

### 2.3.2 Einzelstrangzerlegung

Wir untersuchen nun, welchen Einfluss die Entfernung von Einzelsträngen mit Hilfe der Mung-Bean-Nuclease auf das Experiment „wahllose Wucherung“ hat. Die Grundlagen wurden in Abschnitt 2.2.2 vorgestellt. Bei der Zugabe der Mung-Bean-Nuclease entstehen weniger neue Stränge und Längen über 40 bp sind nicht mehr detektierbar, wie aus Abbildung 2.5 zu erkennen ist. Damit wird ersichtlich, dass die Mung-Bean-Nuclease effektiv die Überhänge abbaut und auch die einzelsträngige DNA über der Schmelztemperatur zerlegt.

In [82, supplementary material] wird ein weiteres Experiment untersucht, in dem die Mung-Bean-Nuclease zu einem späteren Zeitpunkt und mit einer geringeren Konzentration zugefügt wird. In diesem Experiment bilden sich aber zeitnah Stränge der Länge 60 bp, so dass keine Längenselektion beobachtet werden kann. Wir können keine Konfiguration finden, in der sowohl Überhänge effizient entfernt werden und sich eine ausreichend große Reproduktionsrate vorhandener Stränge einstellt. Damit ist die Einzelstrangzerlegung ungeeignet um eine Längenselektion zu erreichen.

Wir wenden uns daher der anderen Möglichkeit der Verringerung der Folgen von Überhangbildung zu, nämlich dem Aufkommen von Hetero(di)meren, indem wir diese degradieren.

### 2.3.3 Heterodimerzerlegung

In Abschnitt 2.2.3 haben wir die Grundlagen der Heterodimerzerlegung mit Hilfe von AfeI und BstUI vorgestellt. Aus Abbildung 2.6 geht hervor, dass es möglich ist, bis zu einer Länge von 40 bp eine Längenselektion zu erreichen, da sich keine messbare Anzahl von 30 bp-Strängen bildet. Jedoch ab einer Länge von 40 bp bilden sich, wie im Experiment „wahllose Wucherung“, wieder alle möglichen Stränge aus.

Dies kann daran liegen, dass die Zeitskalen nicht vollständig separieren. Für kürzere Stränge läuft die Ligationsreaktion langsamer ab, so dass die Restriktionsenzyme mehr Zeit haben, die entstandenen Hetero(di)mere zu zerlegen. Bei höheren Temperaturen läuft die Ligation schneller ab und die Restriktionsenzyme arbeiten zu langsam, um alle Stränge zu zerlegen. Des Weiteren ist die Konzentration von 10 bp-Strängen noch nicht soweit abgefallen, so dass nicht nur längere Überhänge entstehen bzw. der Schmelztemperaturabstand der verschiedenen DNA-Längen abnimmt, so dass eine Hybridisierung zwischen verschiedenen Längen immer wahrscheinlicher wird und letztlich auch Homomere mit allen möglichen Längen entstehen können, die von den Restriktionsenzymen unbeeinflusst bleiben.

Damit ist auch die Heterodimerzerlegung ungeeignet, insbesondere für größere Temperaturen und einer damit verbundenen größeren Ligationsrate, eine Längenselektion aufzuzeigen.

Wir wählen einen gänzlich anderen Ansatz und versuchen damit, die Bildung von Überhangligationen effizient zu unterdrücken und eine Längenselektion zu erhalten, indem wir sog. Linker zu der Taq-Ligase-Reaktion hinzufügen.

### 2.3.4 Längenselektion

Wie schon in Abschnitt 2.2.4 vorgestellt, diskutieren wir hier die Ergebnisse der Ligation mit Hilfe von Linkern  $\mathcal{L}$ . Wir können, wie in Abbildung 2.7 dargestellt, erkennen, dass eine Längenselektion in Form einer Längenverdopplung auftritt. Wesentlich ist, dass sich kurz nachdem sich das DNA-Molekül doppelter Länge gebildet hat das aktuelle DNA-Molekül

## 2 Experimentelle Realisierung mit DNA

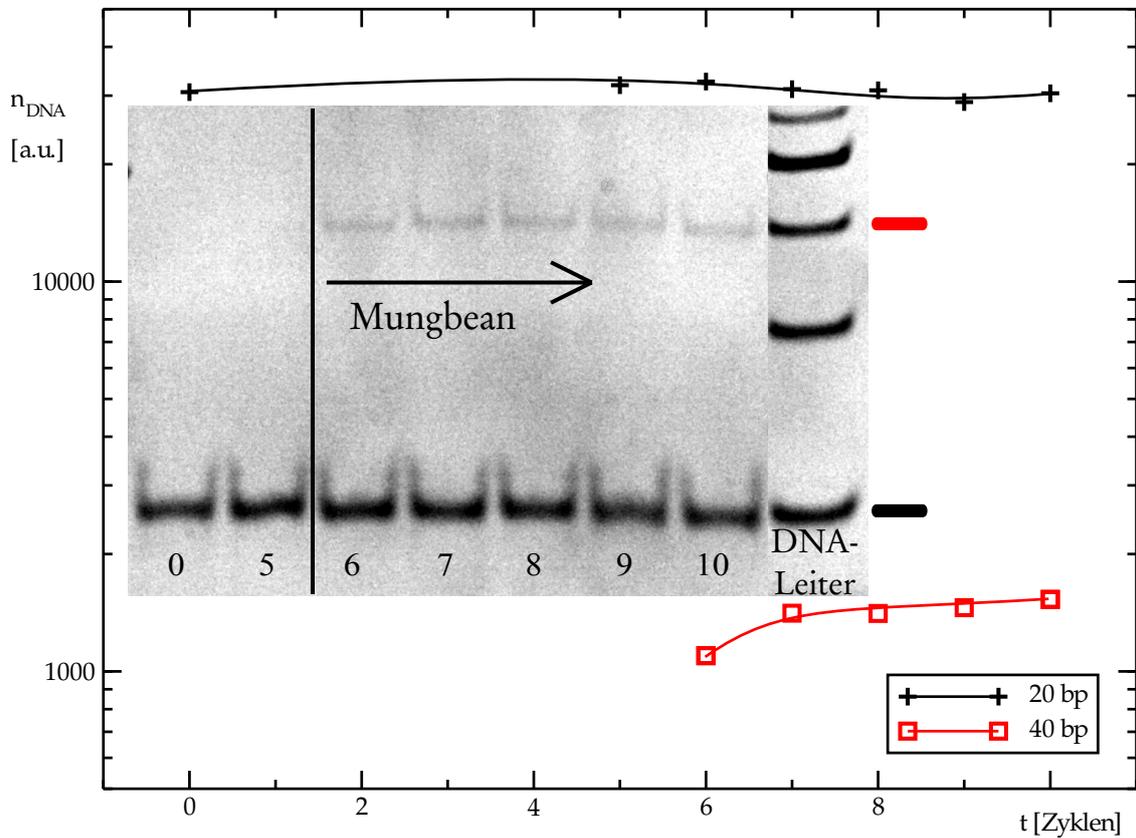


Abbildung 2.5: **Einzelstrangzerlegung** mit Hilfe der Mung-Bean-Nuclease. Zeitentwicklung des DNA-Doppelstrangs  $Z$  bei Ligation mit der Taq-Ligase und zugefügter Mung-Bean-Nuclease ab Zyklus 6. Ein Längenverdopplungsschritt ist möglich, aber längere Stränge entstehen nicht, da die Mung-Bean-Nucleaseaktivität so groß ist, dass sie im Bereich großer Temperaturen, bei denen die Stränge nicht hybridisiert sind, diese ebenfalls zerlegt, obwohl sie nicht temperaturstabil ist [114]. Damit stehen letztlich immer weniger Stränge zur Verfügung. Die Kurven dienen der optischen Führung. Eingefügt sind die (zur besseren Darstellung bearbeiteten) Gele, die ausgewertet wurden. Veröffentlicht in [82].

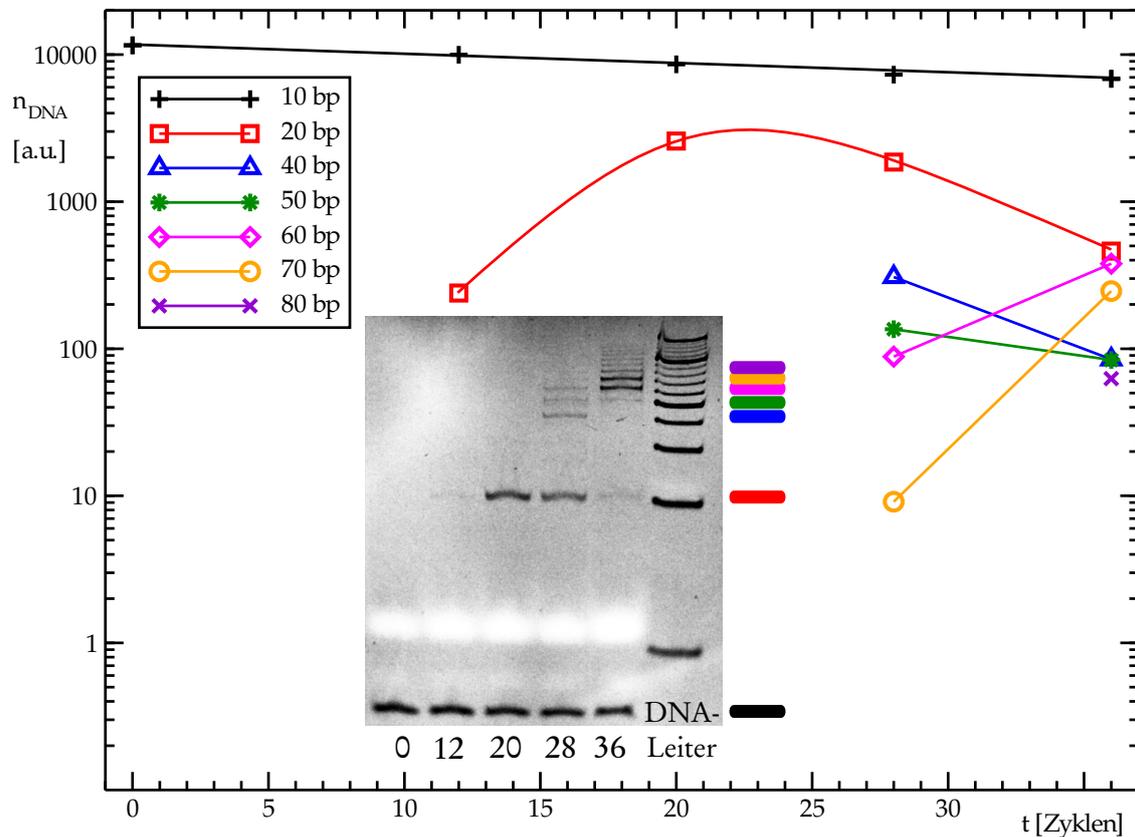


Abbildung 2.6: **Heterodimerzerlegung** mit Hilfe von AfeI und BstUI und Ligation mit der Taq-Ligase. Es wurde der Doppelstrang Q inkubiert. Es entstehen bis zur Auswertungsgrenze von 80 bp alle möglichen Längen bis auf 30 bp. Dies legt nahe, dass der Mechanismus der Hetero(di)merzerlegung am Anfang wirksam ist, da dort die Reaktionen langsamer ablaufen. Mit größeren Temperaturen und längeren Strängen entstehen auch wieder kleinere Überhänge, die eine Längenskalenseparation verhindern. Die Kurven dienen der optischen Führung. Eingefügt sind die (zur besseren Darstellung bearbeiteten) Gele, die ausgewertet wurden. Veröffentlicht in [82].

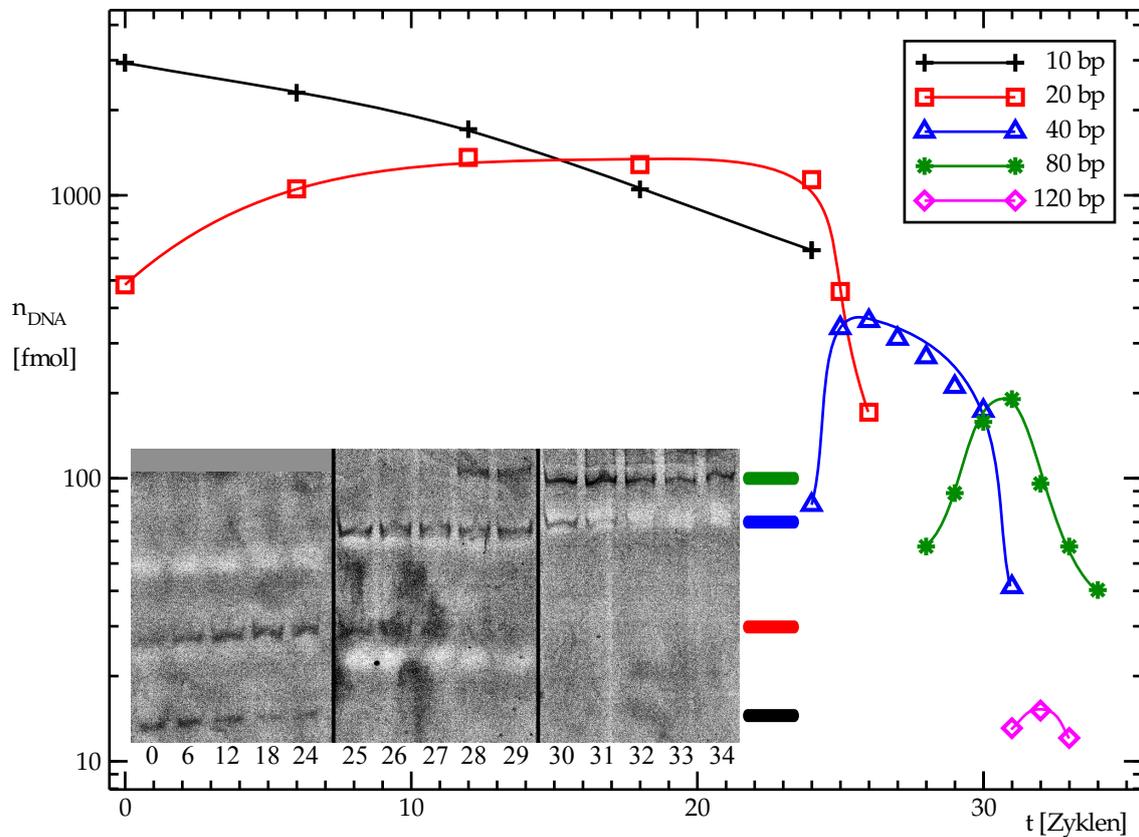


Abbildung 2.7: **Längenselektion.** Acht 10 bp DNA-Sequenzen, die nicht gegenseitig hybridisieren (siehe Tab. 2.5), bilden die Ausgangssequenzen und werden durch Linker, die doppelt so lange Templatessequenzen sind, verbunden. Linker stehen bis zur 80 bp Spezies zur Verfügung (siehe Text). Ab Zyklus 24 und 29 wurden die Maximaltemperaturen erhöht, um den Schmelztemperaturen der längeren DNA-Stränge Rechnung zu tragen. Die durch Linker angestoßene Ligation führt zu einer sukzessiven Längenverdopplung der jeweils überwiegenderen Moleküle. Moleküle anderer Längen sind in diesem Bereich nicht detektierbar. Die Kurven dienen der optischen Führung. Die Konzentration wurde aus den eingefügten (hier bearbeiteten) Elektrophoresegelel ermittelt. Veröffentlicht in [82].

(fast vollständig) abgebaut hat. Daher zeigen alle entstehenden Längen ein Maximum, welches sich schnell bei hohen Zyklustemperaturen und langsam bei niedrigen ausbildet. Der Abfall ist kongruent mit dem Wachstum des in der Länge verdoppelten, folgenden Moleküls. Die entstehenden Moleküle treten daher im Bereich des Maximums exklusiv und sukzessive auf.

Da für Längen über 80 bp keine Linker mehr konstruiert wurden, entsteht als nächst mögliches Molekül, vergleichbar dem Experiment „wahllose Wucherung“, ein DNA-Strang der Länge 120 bp.

Damit konnte eine Möglichkeit gefunden werden, eine Längenselektion durchzuführen. Wesentlich dafür ist eine Zeitskalenseparation zwischen dem Aufbau bestehender Moleküle und der Erzeugung neuer Moleküle (siehe Abschnitt 1.4.4). Dies konnte durch exponentielles Wachstum der nötigen Sequenzen (Tab. 2.5) durch die Hinzugabe von Linkern erreicht werden. Da exponentielles Wachstum schneller als jedes spontane Wachstum ist, entstehen kaum Moleküle mit nicht durch die Linker vorgegebenen Sequenzen; vgl. Kap. 1.4.

Den sukzessiven Aufbau von Molekülen mit doppelter Länge werden wir in einem theoretischen Modell in Kapitel 3.2.2 untersuchen.

### Zusammenfassung

Es wurden in diesem Abschnitt die Ergebnisse der vier vorgestellten Experimente diskutiert. Das Experiment „wahllose Wucherung“ führt zur Entstehung von allen möglichen Längen. Die Einzelstrangerlegung lässt nur einen Verdopplungsschritt zu und die Heterodimerzerlegung ist effektiv im Unterdrücken von 30 bp-Sequenzen; es entstehen jedoch alle möglichen längeren Sequenzen. Die Hinzugabe von Linkern führt im Experiment „Längenselektion“ zu einer Auswahl von sich verdoppelnden DNA-Längen.

Die Experimente „wahllose Wucherung“ und „Heterodimerzerlegung“ erinnern an das Evolutionsszenario „wahllose Wucherung“, da (fast) alle möglichen Längen entstehen (vgl. Abb. 1.15.c), während die Einzelstrangerlegung einen Evolutions-Kollaps zeigt, weil keine längeren Stränge mehr auftreten; vgl. Abb. 1.15.b. Lediglich das Experiment „Längenselektion“ hat Ähnlichkeiten mit der natürlichen Evolution, da sukzessive neue Längen (mit doppelter Länge) auftreten und andere Längen zumindest zeitweise während des Wachstums präsent sind; vgl. Abb. 1.15.a.

Im Folgenden erläutern wir die Funktionsweise des Programms UNAFold und konstruieren mit dessen Hilfe einen Satz an phänomenologischen Gleichungen, um die temperaturabhängige Hybridisierung von komplementären, unterschiedlich langen DNA-Strängen aufzustellen.

## 2.4 Exkurs: UNAFold und Hybridisierungsraum

Abschließend möchten wir die theoretische Funktionsweise der Software UNAFold erläutern. Ausgehend von einem einfachen Modell erklären wir die Grundlagen der Berechnungen über die Gibbs-Energie und der Zustandssumme.

Um im folgenden Kapitel auch Berechnungen mit den tatsächlich verfügbaren DNA-Strängen machen zu können, stellen wir hier ein phänomenologisches Modell auf. Dieses Modell gibt den Stoffmengenanteil der verfügbaren DNA-Stränge, die im Allgemeinen unterschiedliche Länge haben, unabhängig von ihrer Sequenz an.

### 2.4.1 Funktionsweise von UNAFold

Die Funktionsweise von UNAFold ist u. a. in den Veröffentlichungen [75, 76] beschrieben, an denen wir uns orientieren.

Der einfachste Ansatz zur DNA-Hybridisierung ist ein Zweizustandsmodell, bei dem die DNA entweder hybridisiert ist oder nicht. Dabei nimmt man üblicherweise an, dass die beiden betrachteten Stränge zueinander perfekt komplementär sind. Durch den Übergang sind Änderungen in der Gibbs-Energie ( $\Delta G$ ), der Enthalpie ( $\Delta H$ ) und der Entropie ( $\Delta S$ ) zu verzeichnen. Die Legendretransformation zwischen G und H ( $\Delta G = \Delta H - T\Delta S$ ) beschreibt den Übergang zwischen G und H. Durch die experimentell ermittelten Änderungen in G bzw. H [106] (dort in Tabelle 1) für paarweise Basen ( $\Delta G_i$ ), deren Enden ( $\Delta G_E$ ) und Selbstkomplementarität ( $\Delta G_S$ ), lässt sich  $\Delta G$  berechnen und ergibt sich zu<sup>\*</sup>:

$$\Delta G = \sum_i \Delta G_i + \Delta G_E + \Delta G_S \quad (2.1)$$

Da paarweise Basen betrachtet werden, gibt es für einen Strang der Länge n genau n-1 Beiträge. Für Änderungen in der Salzkonzentration werden ebenfalls Korrekturterme angegeben [106].

Aus dem van't Hoff'schen Gesetz  $\Delta G = -RT \cdot \ln\left(\frac{[A][B]}{[AB]}\right)$  lässt sich eine Beziehung zwischen  $\Delta G$  und der Gleichgewichtskonstante angeben, da beim Schmelzen gerade die Hälfte der Stränge dissoziiert sind:

$$-\frac{\Delta G}{R \cdot \ln\left(\frac{1}{2}[AB]_{\text{init.}}\right)} = T_m = -\frac{\Delta H}{\Delta S + R \cdot \ln(C_T/f)} \quad (2.2)$$

wobei  $T_m$  die Schmelztemperatur, R die Gaskonstante und A und B die Konzentration der beiden Stränge ist, bzw.  $C_T$  die Gesamtkonzentration an Strängen. Für selbstkomplementäre DNA-Stränge ist  $f = 1$  und für nicht selbstkomplementäre DNA-Stränge bei gleicher Konzentration ( $[A]_{\text{init.}} = [B]_{\text{init.}}$ ) ist  $f = 4$ . Für nicht selbstkomplementäre DNA-Stränge, die ungleich konzentriert sind, ergibt sich o. B. d. A.  $[A]_{\text{init.}} > [B]_{\text{init.}}$ :  $C_T = [A]_{\text{init.}} + \frac{1}{2}[B]_{\text{init.}}$  mit  $f = 1$  [106].

UNAFold berechnet daneben ebenfalls die gesamte Zustandssumme, so dass die beiden Stränge nicht komplementär sein müssen, mehrere Größenordnungen unterschiedlich konzentriert sein dürfen, und es wird sowohl Faltung als auch Hybridisierung entsprechend berücksichtigt. Es wird angenommen, dass das Volumen gut durchmischt ist und sich bei einer gegebenen Temperatur im thermodynamischen Gleichgewicht befindet.

Es können neben der Schmelztemperatur auch andere experimentell relevante Größen wie die UV-Absorptionsrate und die Wärmekapazität berechnet werden.

### 2.4.2 Hybridisierungsraum

Im Folgenden stellen wir ein Modell vor, um bei einer festen DNA-Sequenz Hybridisierungen von unterschiedlichen Längen schneller zu berechnen. Mit UNAFold werden die verschiedenen Kombinationen berechnet und mit einer Sigmoidfunktion angefitet. Die Sigmoidfunktion

\* Wir wollen  $\Delta G$  beispielhaft anhand der Hybridisierung der Sequenz CGTACG und deren Komplement berechnen:  $\Delta G = \Delta G_1(\text{CG/GC}) + \Delta G_2(\text{GT/CA}) + \Delta G_3(\text{TA/AT}) + \Delta G_4(\text{AC/TG}) + \Delta G_5(\text{CG/GC}) + \Delta G_E(\text{C/G}) + \Delta G_E(\text{G/C}) + \Delta G_S$ . Die Sequenz ist selbstkomplementär und mit den Zahlenwerten aus [106] berechnet sich  $\Delta G$  zu:  $\Delta G = -2,17 - 1,44 - 0,88 - 1,44 - 2,17 + 0,98 + 0,98 + 0,43 = -5,71 \text{ kcal/mol}$ .

ist eine gute Näherung für perfekt hybridisierende Stränge, wie man in Abb. 2.3 erkennen kann. Sicherlich gibt es sehr große Abweichungen von dieser Kurve bei nicht perfekt hybridisierenden Strängen wie z. B. bei sich faltenden Strängen, nicht hybridisierenden Strängen o.ä., wie aus Abb. 2.2.b ersichtlich ist. Daher ist der gewählte Ansatz, indem eine Sigmoidfunktion angefitet wird, in diesen Fällen nicht optimal, aber wir betrachten hier vornehmlich perfekt hybridisierende Stränge. Aus den so gewonnenen Parametern für die Sigmoidfunktion und deren Gültigkeitsbereich finden wir eine Funktion durch Anfiten, die diese Werte beschreibt, so dass es möglich ist, temperaturabhängige Näherungswerte für den Hybridisierungsanteil der DNA-Stränge schneller zu bestimmen, um diese in anderen Programmen einfacher verwenden zu können; vgl. Abschnitt 3.5.1. Alle Funktions- und Parameterbezeichnungen in diesem Abschnitt sind alphabetisch und unabhängig von denen des Rests dieser Arbeit gewählt.

Es wurde für eine feste Sequenz und dessen Komplement für verschiedene Längen mit dem Programm `hybrid2.pl`<sup>\*</sup> aus dem UNAFold-Paket die temperaturabhängige Hybridisierungskurve berechnet. Die Länge der DNA-Stränge wurde in 10 bp-Schritten vergrößert<sup>\*\*</sup>. An den Stoffmengenanteil wurde die Sigmoidfunktion:

$$\chi(T, n, m) = \left(1 + e^{[T - A(n, m)] \cdot B(n, m)}\right)^{-1} \quad (2.3)$$

mit den Parametern  $A(n, m)$  und  $B(n, m)$  angefitet, wobei  $A(n, m)$  die Schmelztemperatur angibt,  $B(n, m)$  die Breite der Funktion und  $n$  sowie  $m$  die Länge in 10 bp-Bausteinen. Aus den empirisch ermittelten Werten für  $A$  und  $B$  wurden entsprechende, passende Funktionen gewählt.  $A(n, m)$  beschreibt eine gekrümmte Ebene in  $n, m$ :

$$\begin{aligned} A(n, m) &= c(m) \cdot \left[1 - e^{-a(m) \cdot n + b(m)}\right], \text{ mit:} & (2.4) \\ a(m) &= d \cdot e^{o \cdot m + f} \\ b(m) &= g \cdot e^{h \cdot m + i} + j \\ c(m) &= k \cdot m + l \end{aligned}$$

Die Parameter ergeben sich zu:  $d = 9,13964$ ;  $o = -0,0326832$ ;  $f = -2,64004$ ;  $g = 2,90485$ ;  $h = -0,000941309$ ;  $i = 1,98539$ ;  $j = -21,1422$ ;  $k = 0,0459531$  und  $l = 85,8959$  und sind in Abb. 2.8.a-b zusammen mit den angefiteten Werten dargestellt. Die Breite der Sigmoidfunktion  $B(n, m)$  kann durch eine Ebene angenähert werden und ist in Abb. 2.8.c zusammen mit den Fitwerten dargestellt:

$$B(n, m) = p \cdot n + q \cdot m + r \quad (2.5)$$

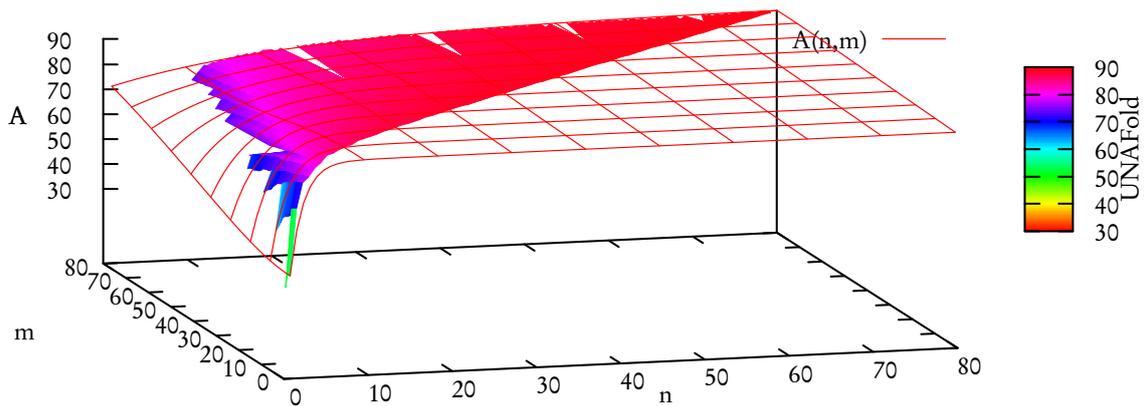
Dabei sind  $p = 0,207996$ ;  $q = 0,0141387$  und  $r = -0,248774$ .

Es werden nur Werte in Abb. 2.8 dargestellt, bei denen  $A(n, m)$  weniger als  $15^\circ\text{C}$  von der mit `melt.pl` berechneten Schmelztemperatur abweicht. Eine große Abweichung deutet auf Faltungen oder nicht hybridisierende Stränge hin. Für kleine Stränge, die auf langen

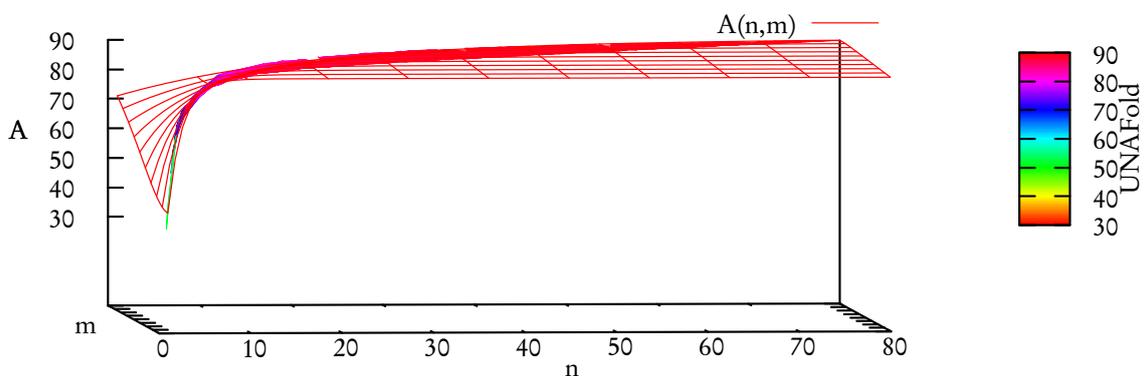
\* Das Programm `hybrid2.pl` wurde mit den Optionen `-n DNA -N 0.025 -M 0.010 -A 0.0000009 -B 0.0000009` und den beiden DNA-Stränge aufgerufen, um die Hybridisierungskurve zu berechnen.

\*\* Die zugrundeliegende Sequenz ist: TGTGC CTTTG ACACC TGATC GGAGT TGGTT ATTTG GACCG CTTAC ACGAG CTACA TGACG GGGAA GTTGT GAAGA CTAGC CCTGT TGAGT CTCGA AATGG CAGTT AAGCG CGCAA GTTAG GATAG TGGGT GTACA GATCG AACT GGTAG ACCAA GGTTC. Sie ist am Anfang identisch zur in Tab. 2.5 vorgestellten und nach dem gleichen Auswahlkriterium fortgeführt. Bei längeren Strängen liegt diese Sequenz repetitiv vor, bei kürzeren abgeschnitten.

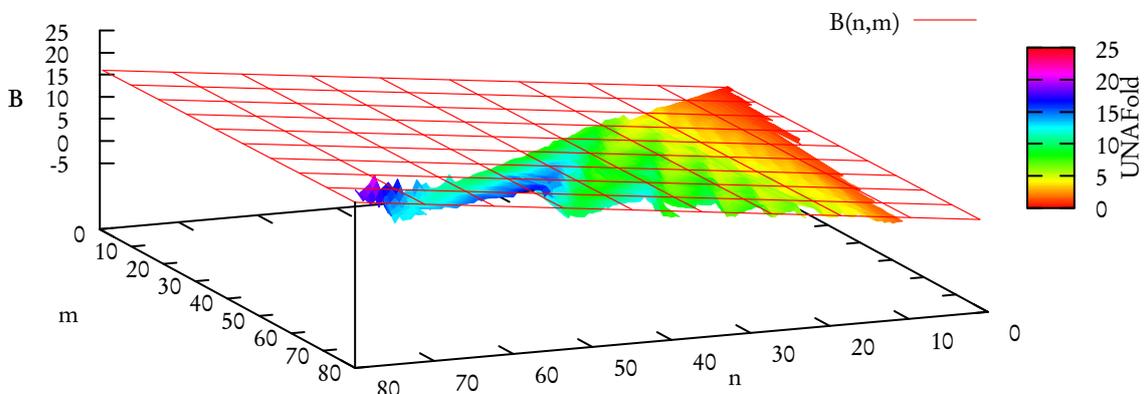
## 2 Experimentelle Realisierung mit DNA



(a)  $A(n,m)$  – Seitenansicht



(b)  $A(n,m)$  – Seitenansicht



(c)  $B(n,m)$  – Seitenansicht

Abbildung 2.8: **Hybridisierungsraum.**  $A(n,m)$  und  $B(n,m)$  sind Parameter zur Sigmoidfunktion (Gl. 2.3), die den Stoffmengenanteil der hybridisierten Stränge angibt.  $A(n,m)$  ist eine gekrümmte Funktion, deren Werte für kleine  $n$  nur teilweise definiert sind (siehe Text). Insbesondere die Krümmung in  $A(n,m)$  wird gut wiedergegeben (b).  $B(n,m)$  stellt eine oszillierende Ebene dar (siehe Text), die durch eine Fläche angenähert wurde. Da beide Funktionen symmetrisch in  $n$  und  $m$  sind, sind die Funktionswerte nur für  $m \leq n$  dargestellt bzw. maßgeblich.

hybridisieren sollen, liegt eher eine Faltung vor, weshalb eine Grenze  $m(n)$  für diese Fälle angefitzt wurde (nicht dargestellt), die sich ergibt zu:

$$m(n) \geq s \cdot n + t \quad (2.6)$$

Mit  $s = 7,90909$ ;  $t = -3,27273$  und für  $n \leq 10$ . Entsprechende  $m(n)$  hybridisieren nicht sigmoidal aufeinander, sondern können ein Maximum aufweisen. Dies ist insbesondere der Fall für die Parameter  $s = 12,57$ ;  $t = -6,14$  und  $n \leq 7$ ; der maximale Stoffmengenanteil an hybridisierten Strängen ist dann  $\chi_{\max.} \leq 0,02$ , so dass praktisch keine Hybridisierungen vorliegen. Zwischen den durch die beiden Parametersätze definierten Grenzen befindet sich der Übergang zur Ausbildung der Hybridisierungen, so dass dort keine allgemeingültige Aussage getroffen werden kann. Außerhalb der Grenze von Gl. 2.6 wird ein Stoffmengenanteil von mindestens  $\chi \geq 0,94$  erreicht.

Es sei angemerkt, dass die Hybridisierung von unterschiedlich langen DNA-Sequenzen weitestgehend (d. h. abhängig von der Sequenz) symmetrisch ist und daher ein Tupel von Längen  $(u, v)$  auf das Tupel  $(n, m) = (\min[u, v], \max[u, v])$  abgebildet wird, da  $m \geq n$ .

Gleichung 2.3 mit den angegebenen Parametern kann zur schnelleren, näherungsweisen Auswertung des hybridisierten Stoffmengenanteils in weitergehenden Anwendungen herangezogen werden. Für andere, nicht faltende Sequenzen müssen gegebenenfalls die Parameter angepasst werden. Dieser Ansatz ist für faltende Sequenzen u. U. nicht zielführend, da von einer sigmoidalen Temperaturabhängigkeit der Hybridisierung nicht mehr ausgegangen werden kann.

## Zusammenfassung

Wir haben die prinzipielle Funktionsweise von UNAFold erläutert. Ein einfaches Modell zur Bestimmung der Schmelztemperatur kann aus dem van't Hoff'schen Gesetz hergeleitet werden, in das die für paarweise Basen gemessene Gibbs-Energie eingeht. Die vollständige Zustandssumme ermöglicht UNAFold eine genauere Berechnung von  $T_m$ .

Wir konnten ein phänomenologisches Modell aufstellen, welches den sigmoidalen, temperaturabhängigen Hybridisierungsanteil zweier komplementärer DNA-Stränge näherungsweise beschreibt. Damit kann in einem weiteren Programm die hybridisierte Stoffmenge schneller berechnet werden. Der funktionale Zusammenhang der Hyperflächen für Schmelztemperatur und Breite der Sigmoidalkurve wurde gezeigt.

Wir gehen nun dazu über, das theoretische Modell mit spontaner und autokatalytischer Ligation vorzustellen und zu analysieren.



## 3 Theorie

In diesem Kapitel stellen wir die theoretische Modellierung und Analyse eines Systems vor, anhand dessen wir die Eigenschaften der Evolution untersuchen. Ausgehend von Abschnitt 1.4 stellen wir eine Reaktionsgleichung auf, die universell, je nach Parameterwahl, eine Längenselektion mit Koexistenz aufweist. Auszüge aus diesem Kapitel wurden in [82] und [122] veröffentlicht und die Grundlagen wurden in [80] gelegt.

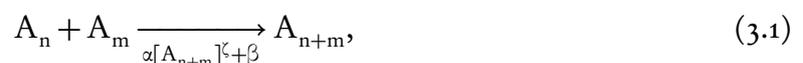
Zunächst analysieren wir die Ergebnisse in Abhängigkeit von den Parametern und im Hinblick auf Evolutionseigenschaften. Es werden Evolutionsbäume sowohl für die theoretischen als auch für die experimentellen Ergebnisse aufgestellt und verglichen. Danach wird näher auf den Aspekt der Selbstähnlichkeit eingegangen und ein Phasenübergang definiert. Das Wachstum und der stationäre Zustand während der auftretenden Plateaus werden ebenfalls beschrieben.

Weitergehend werden Modifikationen an den Gleichungen vollzogen, die es ermöglichen eventuelle Limitierungen der Beschreibung zu umgehen. Zum einen verwenden wir die korrigierten Raten für die Ligaseaktivität und die DNA-Hybridisierung aus Abschnitt 2.4.2 und Abb. 2.1. Zum anderen werden die Unterschiede der verschiedenen Wachstumstypen aus Abschnitt 1.4.2 und 1.4.3 mit dem exponentiellen Wachstum verglichen. Schließlich können wir mit Hilfe von Abschnitt 1.3.5 durch Aufkonzentrierung ein praktisch unendlich evolvierendes System vorstellen.

Abschließend analysieren wir das System in Bezug auf Perkolation. Perkolation tritt im Zusammenhang mit der Verknüpfung von gleichartigen Einheiten auf, falls diese sich nacheinander in verschiedenen Größenregionen verknüpfen.

### 3.1 Systemdefinition

Wir betrachten ein System aus linearen Polymeren  $A_n$ , die aus  $n$  Monomeren bestehen\*. Zwei Polymere  $A_n$  und  $A_m$  können sich zum Molekül  $A_{n+m}$  verketteten. Dies kann entweder spontan mit der Rate  $\beta$  stattfinden oder autokatalytisch mit der Rate  $\alpha[A_{n+m}]^\zeta$ :



die proportional zur Konzentration des Produkts  $A_{n+m}$  ist; vgl. Gl. 1.5. Dies ist die allgemeinste Reaktion zweier spontan und autokatalytisch reagierender Reaktanten. Im Folgenden setzen wir  $\zeta = 1$  und betrachten exponentielles Wachstum, da dies das Beispielsystem für ein

---

\* Allgemeiner betrachten wir **Entitäten** bzw. **Moleküle**, deren Struktur nebensächlich ist, solange sich die Reaktionsart nicht ändert, d. h. hat ein Molekül  $q$  Reaktionszentren, so kann es an  $q$  Stellen mit einem anderen reagieren, unabhängig von seiner tatsächlichen Größe (Länge). Bei linearen, codierenden, hybridisierenden Polymeren ist  $q = 2$  und bei linearen, gerichteten Polymeren ist  $q = 1$ . Wir werden im Folgenden von Polymeren aus einer Sorte ungerichteter Monomere sprechen und daher ist hier  $q = 1$ .

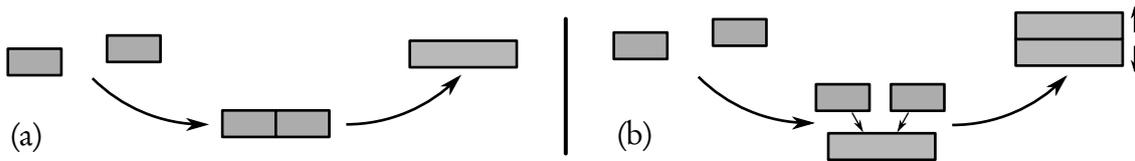


Abbildung 3.1: **Schematische Darstellung der Reaktion aus Gl. 3.1.** Zwei Moleküle können sich spontan treffen und verbinden (a). Zwei Moleküle können autokatalytisch miteinander reagieren, wie z. B. auf einem Target hybridisieren (d. h. nicht kovalent binden), welches als Template fungiert, dort untereinander kovalent reagieren und sich wieder vom Template ablösen (b). Dies ist die allgemeine Form der Reaktion aus Abb. 1.8.

evolutionäres Verhalten darstellt; vgl. Abschnitt 1.4.1. Der autokatalytische Reaktionsweg kann als eine Hybridisierung auf einem Template verstanden werden. Eine schematische Darstellung wird in Abb. 3.1 gezeigt.

Eine solche chemische Reaktion ist stochastisch, weswegen wir sie mit Hilfe eines modifizierten Gillespie-Algorithmus berechnen bzw. simulieren [123, 124, 125, 126]\*. Der Gillespie-Algorithmus setzt voraus, dass das Reaktionseinheitsvolumen gut durchmischt und homogen ist. Dabei gibt in den Simulationsergebnissen die Konzentration  $[A_n]$  die Anzahl der Moleküle  $A_n$  pro Einheitsvolumen an. Als Anfangsbedingung wird eine gegebene Konzentration an Monomeren betrachtet, die im Zeitverlauf reagieren können. Diese Anzahl ist eine Potenz von 2 ( $[A_1]_0 = 2^q$ ). Andere Moleküle sind anfangs nicht vorhanden ( $[A_n]_0 = 0$ , für  $n > 1$ ), wobei  $[A_n]_0$  die Anfangskonzentration von  $A_n$  zur Zeit  $t = 0$  darstellt. Da  $\beta$  die Zeit definiert, wählen wir  $\beta = 1$ ; vgl. Abschnitt 1.4.2.

Trotz der stochastischen Natur geben wir zum besseren Verständnis die chemischen Raten-gleichungen nach [127] an:

$$\begin{aligned} \frac{d}{dt}[A_{n+m}] &= (\alpha[A_{n+m}] + \beta)[A_n][A_m] \\ &\quad - \sum_k (1 + \delta_{n+m,k}) (\alpha[A_{n+m+k}] + \beta)[A_{n+m}][A_k] \end{aligned} \quad (3.2)$$

Dabei ist  $\delta_{i,j}$  das Kronecker-Delta mit  $\delta_{i,j} = \begin{cases} 1, & \text{falls } i = j \\ 0, & \text{falls } i \neq j \end{cases}$ . Es kann sich aus beliebigen Molekülen  $A_n$  und  $A_m$  ein neues Molekül  $A_{n+m}$  bilden. Diese Tatsache macht ein geschlossenes, analytisches Lösen der Gleichungen schwierig. Daher beschränken wir die analytischen Lösungen im Folgenden auf unter Voraussetzungen lösbare, aber dennoch wichtige Spezialfälle. Diese Voraussetzungen wurden aus den Simulationsdaten gewonnen.

\* Der entwickelte **Algorithmus** basiert auf dem Gillespie-Algorithmus [123, 124], der optimiert wurde. Die Auswahl erfolgt sukzessive für jeden Reaktanten und stellt daher sicher, dass alle Reaktionen ihr entsprechendes statistisches Gewicht bekommen. Dabei wird direkt gewählt [125]. Teilweise wurde das Speicherlayout stark minimiert und mit geeigneten Compileroptionen beschleunigt, was die große Anzahl an rechenintensiven Berechnungen erst ermöglichte. Viele Rechnungen wurden aber auch ohne letztere Optimierung durchgeführt. Um mehrere Berechnungen parallel durchzuführen, wurde GNU Parallel benutzt [126].

## 3.2 Analyse

Wir betrachten nun den Zeitverlauf der Konzentrationen für verschiedene autokatalytische Parameter  $\alpha$ . In Abhängigkeit von diesem gibt es zwei unterschiedliche Bereiche. Für kleine  $\alpha$ , d. h. eine schwache Autokatalyse, treten sukzessive immer größere Polymere auf. Für große  $\alpha$ , d. h. eine starke Autokatalyse, treten hauptsächlich Polymere auf, die sich sukzessive in der Länge verdoppeln. Dazwischen finden wir einen Phasenübergang, den wir näher in Abschnitt 3.4.2 untersuchen.

Im Folgenden analysieren wir die beiden Bereiche im Zeitverlauf.

### 3.2.1 Wahllose Wucherung: $\alpha$ klein

Der typische zeitliche Verlauf bei einem kleinen autokatalytischen Parameter ist in Abb. 3.2 dargestellt. Die einzelnen, um ein Monomer längeren Moleküle treten sukzessive auf und haben einen vergleichbaren Verlauf, der ein Maximum aufweist. Auch die Maxima treten nacheinander mit geringer werdender Konzentration auf. Die Maxima folgen näherungsweise einem Potenzgesetz mit  $[A_n]_{\max} \sim t^{-\nu}$ . Der Abfall der einzelnen Moleküle fällt auf eine geneigte, einhüllende Gerade zusammen, wobei kürzere Moleküle schneller abgebaut werden. Diese Gerade gehorcht dem gleichen Gesetz. Die ersten Moleküle, deren Wachstum streng sukzessive erfolgt, sind farbig markiert; das Wachstum der schwarz dargestellten Moleküle ist von Fluktuationen überlagert und daher nicht differenziert dargestellt.

Die verschiedenen Polymere erscheinen nach und nach in logarithmischer Zeit immer schneller. Dabei folgt der Zeitpunkt des Erscheinens der längeren Moleküle näherungsweise einem Potenzgesetz. Der Verlauf ist in Abb. 3.4 gezeigt.

### 3.2.2 Längenselektion: $\alpha$ groß

Eine starke Autokatalyse führt zu einer Längenverdopplung und damit zu einer Längenselektion, wie sie in Abb. 3.3 dargestellt ist. Es entstehen kontinuierlich und sukzessive neue Polymere  $A_n$  mit  $n = 2^k$ ,  $k \in \mathbb{N}$  also mit einer verdoppelten Länge: Bei einer großen autokatalytischen Verbindungsrate  $\alpha$  ist die Erzeugung von wachsenden Polymeren größer als alle anderen Reaktionen, so dass die Polymere, die die größte Konzentration aufweisen, ein Plateau zeigen. Da wir die Reaktion mit Monomeren beginnen, entstehen zuerst Dimere, die in einem Plateau sättigen. Danach besteht jedoch die Möglichkeit, dass Tri- oder Tetramere entstehen. Obwohl zunächst die Rate, mit der Trimere wachsen, größer ist und diese demnach zuerst entstehen, sinkt diese Rate, da ein Reaktant (Monomere) auch von einer vorausgehenden Reaktion verbraucht wird: der Erzeugung von Dimeren. Daher erwarten wir, dass diese Reaktionen langsamer werden und schließlich ausklingen. Die Tetramerreaktion verhält sich jedoch anders: Da die Edukte reichlich vorhanden sind und produziert werden, wächst damit die Reaktionsrate. Schließlich reagieren die Dimere zu Tetrameren und diese sättigen in einem Plateau.

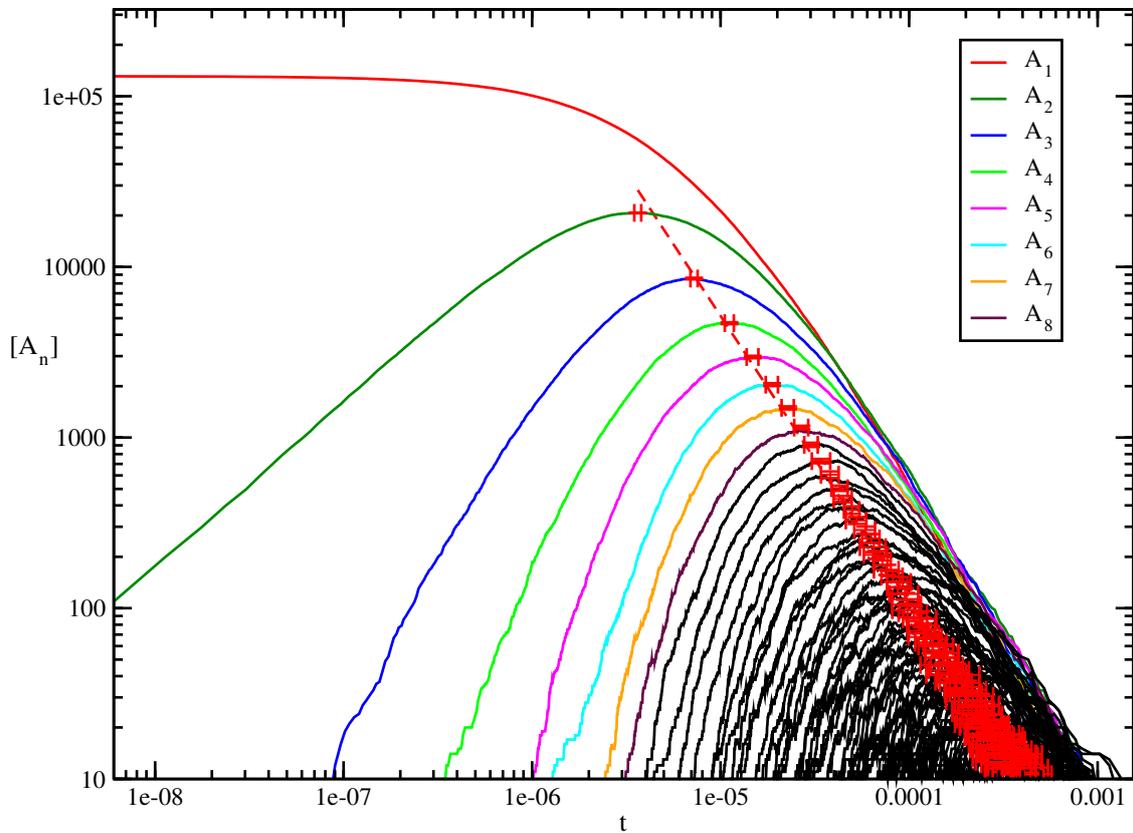


Abbildung 3.2: **Spontane Reaktion** linearer Moleküle bei kleinem autokatalytischen Parameter  $\alpha$ . Aus den am Anfang vorhandenen Monomeren bilden sich im Zeitverlauf nach und nach Moleküle, die immer um ein Monomer länger werden, da Monomere für einen langen Zeitraum vorhanden sind. Ab einer gewissen Länge ist diese Ordnung nicht mehr strikt und es fehlen Längen. Die streng geordneten Moleküle sind farbig hervorgehoben. Das Maximum der jeweiligen Moleküle ist ebenfalls geordnet. Die über verschiedene Realisationen gemittelten Maxima der jeweiligen Moleküle sind rot eingezeichnet. Sie folgen einem Potenzgesetz mit  $\nu = 1,67(1)$  (siehe Text). Die Fehlerbalken sind die Standardabweichungen über 20 Läufe. Die Verläufe der verschiedenen Polymere weisen einen ähnlichen stetigen Verlauf auf: Der Anstieg ist monoton, gefolgt von einem Maximum, der in einen monotonen Abfall übergeht, der allen Molekülen ähnlich ist. Parameter:  $[A_1]_0 = 2^{17}$ ,  $\alpha = 10^{-5}$ . Veröffentlicht in [82].

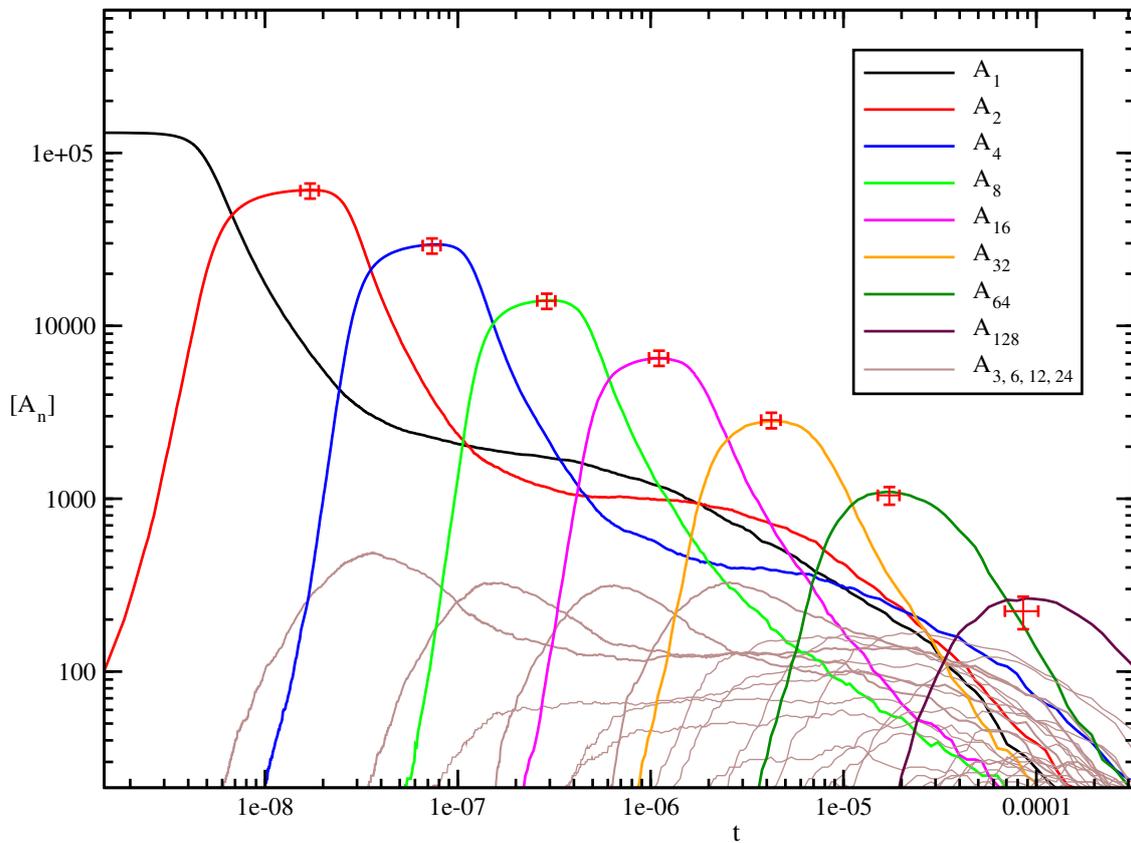


Abbildung 3.3: **Autokatalytische Reaktion** linearer Moleküle bei großem autokatalytischem Parameter  $\alpha$ . Es treten zwei Sorten von Molekülen auf: diejenigen, die aus einer Längenverdopplung resultieren, und andere. Erstere weisen eine um Größenordnungen größere Konzentration auf als die restlichen und sind farblich gekennzeichnet. Diese wachsen exponentiell, zeigen ein Plateau mit einem Maximum und einen Abfall, der eine Schulter aufweisen kann. Die Maxima fallen bei moderat überkritischem autokatalytischem Parameter  $\alpha$  für verschiedene Läufe zusammen und gehorchen für verschiedene Längen einem Potenzgesetz. Die Mittelwerte und Standardabweichungen sind rot eingetragen. Ebenfalls gut zu erkennen sind Moleküle, die aus zwei aufeinanderfolgenden, längenverdoppelten Polymeren resultieren ( $n = 3, 6, 12$  und  $24$ ). Diese sind jedoch um Größenordnungen geringer konzentriert. Eine noch geringere Konzentration haben sporadisch auftretende, weitere Moleküle. Parameter:  $[A_1]_0 = 2^{17}, \alpha = 10^{-1}$ . Veröffentlicht in [82].

### Selbstähnlichkeit

Es ist wesentlich festzustellen, dass dies ein allgemeingültiges Schema ist. Anstatt den Reaktionsweg von Tetrameren (Plateaukonzentration  $[A_4]_{\text{Plat.}}$ ) zu Oktameren zu untersuchen, könnte man auch die Reaktion von Monomeren zu Dimeren oder jede andere Verdopplungsreaktion von einem Polymer der Länge  $2^k$  betrachten. Demnach liefert die Betrachtung von Monomeren mit der Konzentration  $[A_4]_{\text{Plat.}}$ , die man als Tetramere identifiziert, vergleichbare Ergebnisse. Folglich wurde nur die Anfangskonzentration ungefähr geviertelt. Diese Selbstähnlichkeit hat weitreichende Folgen, auf welche wir näher im Abschnitt 3.4.1 eingehen.

### Längenverdoppelte Polymere

Ein längenverdoppeltes Polymer ( $A_{2m}$ ) wächst zunächst exponentiell (vgl. Abschnitt 3.4.3), da die Konzentration des vorausgehenden Polymers ( $A_m$ ), welches an dieser Stelle ein Plateau zeigt, relativ konstant ist. Danach geht die Konzentration über in eine Sättigung, während der das nächste Polymer mit doppelter Länge ( $A_{4m}$ ) wächst. Das Plateau weist ein Maximum auf. Darauf folgt ein stetiger Abfall, der eine Schulter zeigen kann. Diese tritt auf, falls einerseits noch ein Teil der Moleküle des Polymers mit halber Länge ( $A_m$ ) nicht reagiert hat und das Polymer mit doppelter Länge ( $A_{4m}$ ) in das Plateau sättigt. Zum anderen muss das System, in Abhängigkeit von  $\alpha$ , in geringer Konzentration andere Längen aufweisen, die nicht aus einer Längenverdopplung resultieren. Die Schulter kann ansteigen, insbesondere bei mittleren Werten von  $\alpha$ . In der Nähe des Phasenübergangs ist die Form des Plateaus leicht ansteigend, glatt und vergleichbar in der Länge. Höhere Ordnungen zeigen, wenn sie sich dem Phasenübergang weiter nähern, kein ausgeprägtes Plateau. Für größere  $\alpha$  wird das Plateau zunehmend waagrecht, dessen Länge\* statistisch verteilt und der Übergang ins Plateau immer kürzer.

Für kleinere  $\alpha$  sind die Plateaulängen und auch deren Maxima deterministisch. Die Maxima der längenverdoppelten Polymere folgen einem Potenzgesetz:

$$[A_{2^k}]_{\text{max}}(t) \sim t^\xi \quad (3.3)$$

mit einem Exponenten von  $\xi = -0,52(2)$ .

### Deterministische Plateaulänge: andere Längen und Reaktionsbeginn

Wenn die Plateaulänge deterministisch ist, bleibt eine signifikante Konzentration übrig, die nicht in der Längenverdopplungsreaktion verbraucht wird. Daher treten besonders Polymere der Länge  $m = 2^{k-1} + 2^k$  (z. B.  $m = 3, 6, 12$  und  $24$ ) in einer größeren Konzentration auf, wenn die entsprechenden Edukte in größerer Konzentration vorhanden sind. Diese und die Polymere aus der Längenverdopplungsreaktion sind gut zu erkennen und reagieren dann weiter zu beliebigen anderen Polymeren, die in geringerer Konzentration auftreten.

Folglich bricht in diesem Fall gegen Ende der Reaktion die Längenverdopplungskaskade zusammen, da die anderen Längen eine vergleichbare Konzentration erreicht haben. Dies tritt umso früher ein, je kleiner  $\alpha$  ist.

---

\* Mit **Plateaulänge** wird in dieser Arbeit die Dauer des Plateaus bezeichnet.

Da die Plateaulänge deterministisch ist, muss auch das Auftreten der neuen, längenverdoppelten Polymere deterministisch sein. Die Zeitspanne folgt einem Potenzgesetz und ist in Abb. 3.4 dargestellt.

### Statistische Plateaulänge

Für sehr große Werte von  $\alpha$  sind die Plateaulängen nicht mehr deterministisch, sondern statistisch verteilt; vgl. Abschnitt 3.4.4. Wegen der großen Autokatalyse erfolgt das Wachstum eines neuen längenverdoppelten Polymers sehr schnell, wohingegen die Entstehung eines neuen Polymers länger dauert. Daher treten typischerweise entweder keine anderen Längen auf oder nur in geringer Konzentration. Es kann aber passieren, dass das Plateau so kurz ist, dass der langsamere Abfall noch nicht abgeschlossen ist und eine signifikante Anzahl an Restmolekülen übrig bleibt. Diese können mit anderen Molekülen reagieren, typischerweise mit einem anderen Polymer aus der Längenverdopplungsskala oder mit sich selbst. Es tritt entweder eine gemeinsame, zunächst unabhängige Reaktion ein, die Moleküle von den Polymeren  $[A_{2^k}]$  verbraucht, oder eine davon unabhängige Längenverdopplung; vgl. Abb. 3.5.

Die statistische Verteilung der Plateaulänge resultiert aus dem Warten auf die Entstehung eines neuen Polymers. Daher haben die beiden Größen die gleiche Verteilung.

### 3.2.3 Koexistenz

Eine geringe Autokatalyse ( $\alpha$  klein) führt, da fast alle Polymerlängen entstehen, zu einer Koexistenz vieler Moleküle. Eine differenzierte Form der Koexistenz lässt sich im Bereich der Längenverdopplung erkennen; vgl. Abb. 3.3. Bei moderaten autokatalytischen Parametern, also einer deterministischen Plateaulänge, haben für größere Zeiten die Polymere  $A_{2^k}$  eine vergleichbare Konzentration und werden in einer Nische geduldet, d. h. sie tragen zur Längenverdopplungskaskade nicht mehr bei. Ebenfalls existieren Polymere der Länge  $m = 2^k + 2^{k+1}$ , die sich ebenfalls untereinander verdoppeln, aber auch aus der Längenverdopplungskaskade gespeist werden. In noch geringerer Konzentration sind andere Moleküle vorhanden. Wir stellen fest, dass weitere, teilweise auch geordnete Reaktionen neben der Längenverdopplungskaskade ablaufen und diese Linien damit koexistieren.

Ein anderes Bild zeichnet sich für sehr große autokatalytische Parameter  $\alpha$  ab. Dort kommt es bei einzelnen Realisierungen vor, dass andere (ursprünglich) nicht aus der Verdopplungskaskade stammende Polymere einen signifikanten Anteil erreichen und teilweise sowie zeitweise auch eine größere Konzentration erreichen als diejenigen, die aus einer Verdopplung resultieren; vgl. Abb. 3.5. Dieser Sachverhalt stellt eine Koexistenz zweier Entwicklungslinien dar, die aufeinander beruhen können, was wir als Koevolution bezeichnen. Andere Moleküle sind jedoch weitaus geringer konzentriert.

Ein koevolvierendes Polymer ( $A_m$ ) kann entstehen, wenn das Polymer doppelter Länge ( $A_{2m}$ ) ein kurzes Plateau auf Grund statistisch verteilter Plateaulängen aufweist und eine signifikante Restkonzentration von  $A_m$  im System verbleibt. Dieses Polymer kann dann mit einem anderen weiter reagieren, insbesondere falls dessen Plateau eine größere Länge aufweist, da dann eine Reaktion wahrscheinlicher ist, oder unabhängig davon mit sich selbst. Beide Reaktionstypen kommen in Abb. 3.5 vor:  $A_1 + A_{2^5} \rightarrow A_{33}$  und  $2A_{97} \rightarrow A_{194}$ . Die

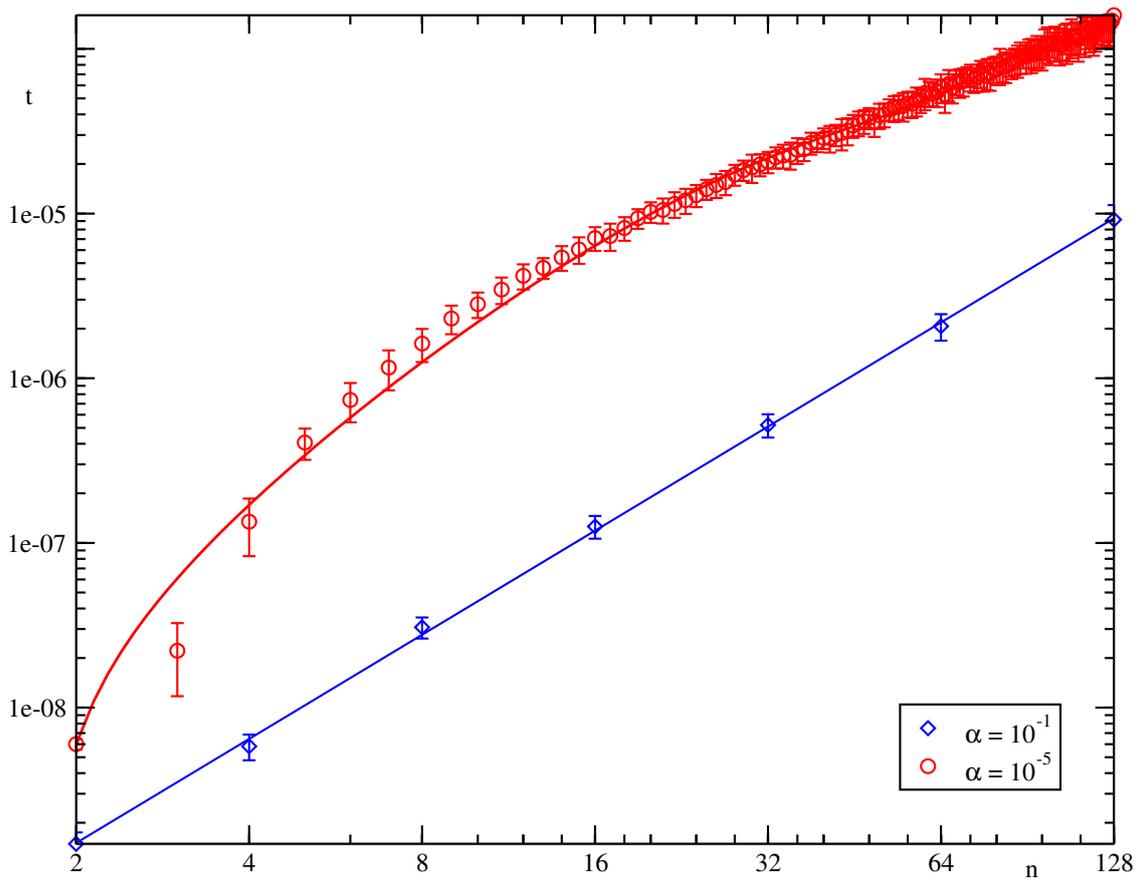


Abbildung 3.4: **Auftreten** aller Polymere für  $\alpha = 10^{-5}$  (rot) und für die längenverdoppelten Polymere  $\alpha = 10^{-1}$  (blau). Für  $\alpha = 10^{-1}$  folgt der Zeitpunkt des Auftretens einem Potenzgesetz:  $t_{\text{Auftr.}}(n) \sim n^{2.10(2)}$ . Die rote Kurve dient der optischen Führung. Veröffentlicht in [82].

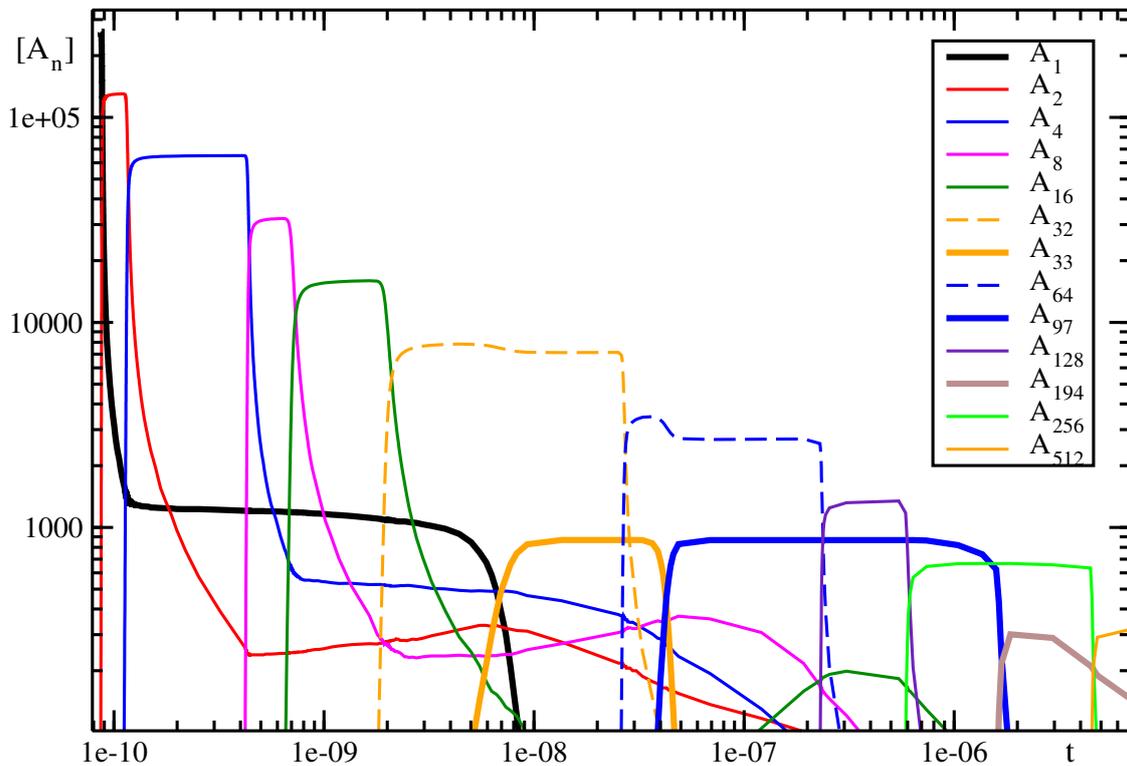


Abbildung 3.5: **Koevolution.** Längenverdopplung mit statistisch verteilten Plateaulängen. Es kann bei kurzen Plateaus möglich sein, dass noch eine signifikante Konzentration übrigbleibt ( $A_1$ ). Diese Moleküle können zu einem späteren Zeitpunkt weiter reagieren und Moleküle aus der Längenverdopplungskaskade verbrauchen (gestrichelt gezeichnet). Es entstehen neue Längen (fett markiert), die sich nach dem selben Schema entwickeln oder sich unabhängig davon verdoppeln können. Eine genaue Analyse der stattfindenden Reaktionen findet sich im Text. Andere Polymere haben eine geringe Konzentration und sind nicht eingezeichnet. Die Parameter sind  $\alpha = 10^2$  und  $[A_1]_0 = 2^{18}$ . Veröffentlicht in [82].

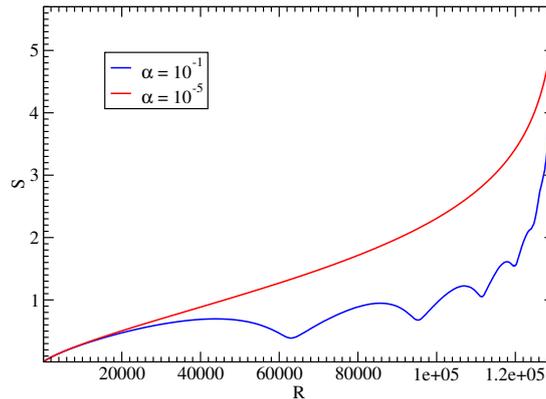


Abbildung 3.6: **Entropie**  $S = -\sum_n p_n \log(p_n)$ , wobei  $p_n$  die relative Häufigkeit des Polymers  $A_n$  ist.  $R$  ist die Anzahl der Bindungen, die proportional zur inneren Energie des Systems ist. Die Entropie der Längenselektion ist geringer als die der „wahllosen Wucherung“, so dass erstere eine größere Ordnung aufrechterhält. Veröffentlicht in [82].

Zwischenreaktion verbraucht ein weiteres Polymer ( $A_{26}$ ) aus der Längenverdopplungskaskade:  $A_{33} + A_{26} \rightarrow A_{97}$ . Dabei sei angemerkt, dass das Polymer  $A_{97}$  zu einem späteren Zeitpunkt eine größere Konzentration aufweist als  $A_{28}$ .

### 3.2.4 Entropie

Wie wir in den Vorüberlegungen zu Abschnitt 1.2 schon angedeutet haben, ist es notwendig, so Schrödinger, dass Lebewesen ihre Entropie auf einem niedrigeren Niveau halten als ihre Umgebung [22]. Dies wird mit Hilfe von Nahrung bewerkstelligt, die zugleich die abgegebene Wärme ersetzt. Dies ist notwendig, um dem irreversiblen Zerfall, der nach dem zweiten Hauptsatz der Thermodynamik für Gleichgewichtssysteme gilt, entgegenzuwirken und so eine Entropieabnahme zu erreichen, so dass spontane, selbst-organisierende Phänomene entstehen können [70, 128]; vgl. auch Abschnitt 1.3.4 und 1.3.5.

Aus Abb. 3.6 geht hervor, dass die Entropie in Abhängigkeit der Reaktionsanzahl, ein Maß für die innere Energie des Systems, für den Parameterbereich, der Längenselektion zeigt, geringer ist als im Fall der „wahllosen Wucherung“. Damit hält die Längenselektionsphase die Ordnung im System aufrecht, ermöglicht Kodierung von Information und zeigt Selbstorganisation.

Aus Abb. 3.6 geht hervor, dass die Entropie, ein Maß für die innere Energie des Systems, in Abhängigkeit von der Reaktionsanzahl für den Parameterbereich, der Längenselektion zeigt, geringer ist als im Fall der „wahllosen Wucherung“. Damit hält die Längenselektionsphase die Ordnung im System aufrecht, ermöglicht Kodierung von Information und zeigt Selbstorganisation.

## Zusammenfassung

In diesem Abschnitt konnten wir die Reaktionsgleichungen vorstellen, analysieren und einen Überblick über die möglichen Verläufe geben. Eine kleine Autokatalyse führt zu vielen, kaum differenzierten Polymeren und ähnelt einer „wahllosen Wucherung“. Große Werte für den autokatalytischen Parameter führen jedoch zu einer Längenselektion, der Längenverdopplungskaskade. Neben den Polymeren mit sich verdoppelnder Länge existiert eine Familie aus Polymeren, die aus zwei solcher aufeinanderfolgenden, längenverdoppelten Polymere besteht. Auch weitere Polymere treten auf. Die Längenverdopplungskaskade ist selbstähnlich. Für sehr große Autokatalyseraten werden die Plateaulängen der längenverdoppelten Polymere statistisch verteilt und es können davon unabhängige Kaskaden entstehen.

Die Entropie der „wahllosen Wucherung“ ist monoton steigend und größer als die der Längenselektion. Damit hält die Längenselektion eine größere Ordnung im System aufrecht.

Im Folgenden vergleichen wir die Ergebnisse von Experiment und Theorie.

## 3.3 Vergleich von Theorie, Experiment und Evolution

Der Reaktionsablauf ist in Theorie und Experiment verschieden und damit auch die Definition der Zeit. Das theoretische Modell geht von einer kontinuierlich ablaufenden Reaktion aus, d. h. in der templatebasierten Reaktion hybridisieren und reagieren die Edukte instantan und die Produkte trennen sich ebenfalls instantan. In einem Zeitschritt laufen diese Reaktionen also so schnell ab, so dass keine weitere Reaktion stattfindet; vgl. Abschnitt 3.1.

Im Experiment jedoch läuft die auf die autokatalytische Reaktion bezogene Zeit nicht kontinuierlich, sondern in Zyklen ab. Pro Zyklus kann nur ein Teil der Templates als solche fungieren; der Rest hybridisiert auf seinem Komplement oder faltet sich, falls möglich. Während des restlichen Zyklus kann dann mit Hilfe der Ligase ligiert werden; vgl. Abschnitt 2.1.1. Diese Reaktionen laufen während des Zyklus entsprechend schnell ab, aber sie werden von einem (lange dauernden) Zyklus gestoppt. Der Zyklus ist entsprechend gewählt, damit alle Reaktionen genügend Zeit haben, vollständig abzulaufen. Danach wird über die Schmelztemperatur erhitzt, alle Stränge schmelzen und können während des Abkühlens wieder entsprechend hybridisieren. Dabei stellt die Probenentnahme am Zyklusende eine Art Stop-Motion-Aufnahme des zugrunde liegenden Reaktionsverlaufs dar, wobei die Probenentnahme effektiv nichtlinear in der Zeit ist.

Damit erscheint ein Vergleich, insbesondere die Zeit betreffend, nicht sinnvoll. Dennoch wäre z. B. im Fall der Längenverdopplungskaskade die Abfolge der entstehenden Polymerlängen vergleichbar.

### Wahllose Wucherung

Abbildungen 2.4 und 3.2 haben einen vergleichbaren Verlauf. In beiden Fällen entstehen sukzessive aus den Monomeren<sup>\*</sup>, um die Monomerlänge längere, neue Polymere. Der Konzentrationsverlauf zeigt ein Maximum, welches für kürzere Polymere einen größeren Wert

---

\* Die Monomere sind im Experiment „wahllose Wucherung“ 20 bp-Stränge.

aufweist. Je länger die Polymere werden, desto geringer ist der Abstand der Maxima. Da es keinerlei Selektion bezüglich der entstehenden Längen gibt, ist der Verlauf, in Theorie und Experiment, ähnlich zur wahllosen Wucherung; vgl. Abb. 1.15.c und Abb. 3.7.a-b.

#### Längenselektion

In den Abbildungen 2.7 und 3.3 ist eine Längenselektion dargestellt. Ausgehend von Monomeren\* bilden sich sukzessive Polymere mit doppelter Länge. Die Konzentrationen haben einen an- und absteigenden Verlauf mit einem Maximum und kommen größtenteils nur exklusiv vor.

Anlehnend an Abschnitt 1.4.4 stellen wir fest, dass alle notwendigen Voraussetzungen zur Entstehung einer Längenselektion vorliegen: Die Reaktionen verlaufen genügend schnell, so dass ein Großteil der Verdopplungsreaktion abgelaufen ist, bevor sich ein neues Polymer bildet. In beiden Fällen wurde die autokatalytische Rate vergrößert: Im theoretischen Modell direkt und im Experiment durch die Zugabe von Linkern, die dazu führen, dass die entsprechende Längenverdopplungsreaktion anfangs schneller abläuft und somit der Totzeit bis zur Entstehung eines neuen Polymers zuvorkommt.

Das Ergebnis ist, mit Einschränkungen, vergleichbar mit der natürlichen Evolution; vgl. Abb. 1.15.a und Abb. 3.7.c-d. Die Einschränkungen sind zum einen das relativ schnelle Aussterben der Vorfahren und zum anderen, dass im dargestellten Fall der Entstehungszeitpunkt sowie die Länge der entstehenden Spezies deterministisch sind.

#### Gemeinsamkeiten mit der darwinschen Evolution

Um das Verhalten der Längenselektion auf die darwinsche Evolution hin zu untersuchen, werden im Folgenden die notwendigen Begrifflichkeiten erläutert bzw. identifiziert. Da die Polymere nur die Eigenschaft „Länge“ besitzen, ist diese das Unterscheidungsmerkmal für eine Spezies; vgl. Abschnitt 1.5. Da im biologischen Kontext eine Spezies von einer anderen deutlich zu unterscheiden ist, sehen wir davon ab, im Fall der „wahllosen Wucherung“ von Spezies zu sprechen. Im Fall der Längenverdopplungskaskade sind jedoch die einzeln auftretenden Polymere gut zu unterscheiden und können demnach als Spezies identifiziert werden. Im Folgenden dient eine solche Spezies ebenfalls als Nische; vgl. Kap. 1.1, 1.2.2 und 1.5. Im theoretischen Modell läuft die Entstehung einer neuen Spezies wie folgt ab: Aus den zwei vorherigen Spezies  $A_{2^k}$  und  $A_{2^{k+1}}$  bildet sich durch Variation zunächst ein neues Polymer:  $A_{2^k+2^{k+1}}$ . Dieses wächst jedoch nicht so stark, weil dessen Nische verschwindet (vgl. Abschnitt 3.2.2) und stattdessen in einer neuen, leicht veränderten Nische (nur das Polymer  $A_{2^{k+1}}$ ) das Polymer  $A_{2^{k+2}}$  durch Variation entsteht und weitaus stärker wächst. Dieser Vorgang kann als Selektion identifiziert werden. Damit treten alle typischen Eigenschaften der darwinschen Evolution auf.

Im Folgenden gehen wir dazu über, die physikalischen Eigenschaften des Systems genauer zu analysieren und beginnen mit dem Phasenübergang und dessen Konsequenzen.

---

\* Die Monomere sind im Experiment **Längenselektion** 10 bp-Stränge.

### 3.3 Vergleich von Theorie, Experiment und Evolution

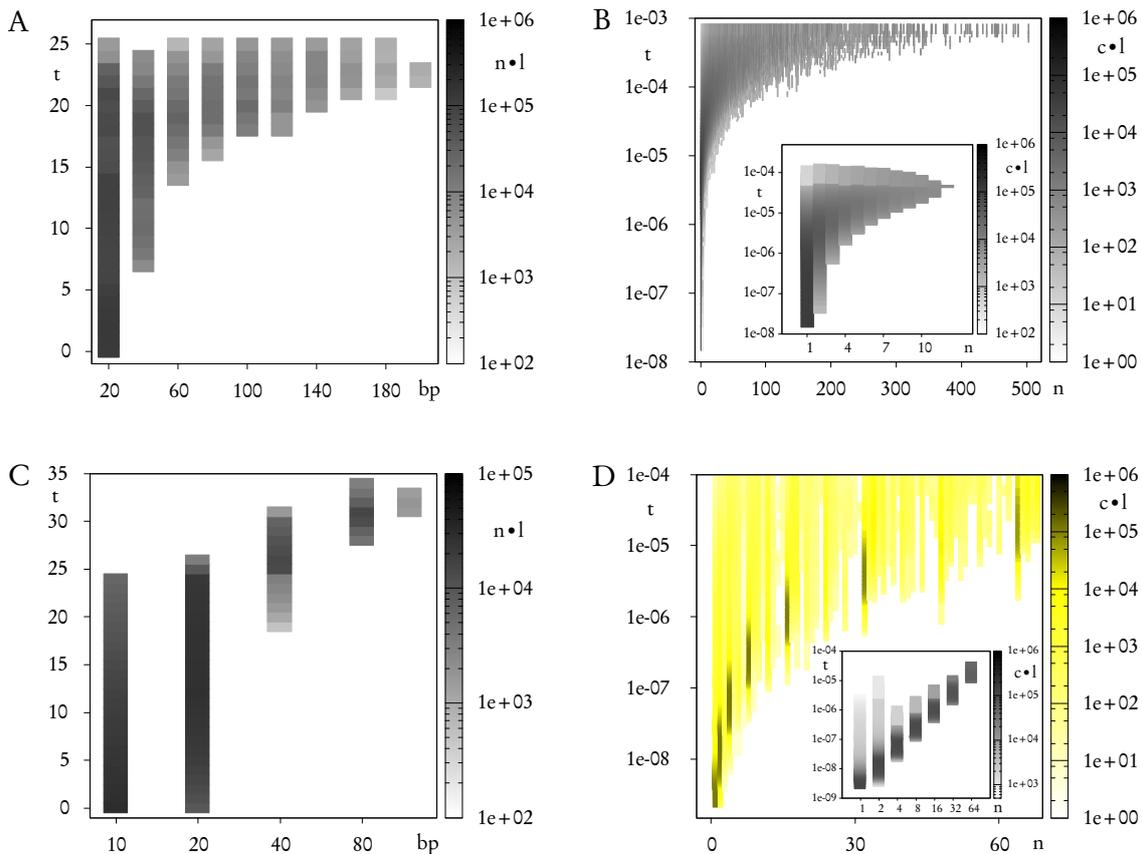


Abbildung 3.7: **Verläufe der Evolution als Evolutionsbaum.** Bild (A) und (B) zeigen ein zur „wahllosen Wucherung“ vergleichbares Verhalten (vgl. Bild 1.15.c) im unterkritischen Fall im Experiment und in der Simulation; Bild (C) und (D) entsprechend den überkritischen Fall, welcher Ähnlichkeiten zur Abb. 1.15.a aufweist (siehe Text). Es ist die Länge gegen die Zeit aufgetragen. Die Helligkeit bildet die Gesamtmasse (Konzentration  $\cdot$  Länge) ab. (A) Alle möglichen Längen erscheinen nacheinander und die Konzentration nimmt mit zunehmender Zeit ab. Ein vergleichbares Bild zeigt sich in (B). Im Inset ist ein Ausschnitt für kürzere Längen dargestellt. Bei allen entwickelt sich das Maximum der Gesamtmasse für größere Zeiten hin zu größeren Längen. (C) Aus den kürzeren Spezies wächst die nächste mit doppelter Länge sukzessive heraus. Die Spezies mit 120 bp Länge bildet sich nicht mit Hilfe von Linkern und resultiert daher nicht aus einer Längenverdopplung. (D) Betrachtet man nur die sich längenverdoppelnden Spezies (alle mit Konzentration größer als 500), so treten diese ebenfalls in der Simulation sukzessive auf (Inset). Die gelb dargestellten Moleküle sind in geringer Konzentration fast in allen Längen vorhanden, aber es gibt prominente Spezies, die in deutlich größeren Konzentrationen auftreten (hellgelb bzw. schwarz dargestellt). Logische, horizontale Verbindungslinien, die das Abstammungsverhältnis angeben, sind nicht eingezeichnet.

## 3.4 Phasenübergang, Selbstähnlichkeit und Plateaulänge

Nachdem wir einen Überblick über das Verhalten des Systems gegeben und den biologischen Zusammenhang erläutert haben, gehen wir nun dazu über, das System näher zu untersuchen.

Da zwei grundsätzlich unterschiedliche Ergebnisse („wahllose Wucherung“ und Längenselektion) im System auftreten, muss es einen Phasenübergang zwischen den beiden geben. Dessen Verlauf folgt aus der Selbstähnlichkeit, die wir nun näher charakterisieren.

### 3.4.1 Selbstähnlichkeit: Plateaukonzentration

Wenn Längenselektion auftritt, ist das System selbstähnlich, wie wir in einem Passus in Abschnitt 3.2.2 bereits erläutert haben: Jedes Polymer, welches ein Plateau zeigt und sich in einem solchen mit der Konzentration  $[A_n]_{\text{Plat.}}$  befindet, kann als Monomer mit entsprechender Konzentration angesehen werden und reagiert hauptsächlich nur mit sich selbst zu einem Dimer ( $A_{2n}$ ). Daraus resultiert die Längenverdopplungskaskade und wir können die zu erwartende Konzentration eines sich verdoppelnden Polymers in Abhängigkeit von der Anfangskonzentration angeben:

$$[A_{2^k}]_{\text{Plat.}}^+ \leq \frac{1}{2^k} \cdot [A_1]_0 \quad (3.4)$$

$[A_{2^k}]_{\text{Plat.}}^+$  ist das Supremum der maximal möglichen Plateaukonzentration. Wegen des Teilchenverlusts an andere Reaktionen wird die tatsächliche Plateaukonzentration entsprechend geringer sein.

Dies bedeutet, dass die Konzentrationen idealerweise äquidistant in logarithmischer Auftragsung sind. Damit ist alleine die Zeitverteilung maßgeblich für die Verteilung der Maxima; vgl. Gl. 3.3.

### 3.4.2 Phasenübergang

Der Phasenübergang, der in Abbildung 3.8 dargestellt ist, ist abhängig von der anfänglichen Konzentration der Monomere ( $[A_1]_0$ ) durch ein Potenzgesetz mit Exponent  $\gamma$ . Dieser Exponent ist negativ ( $\gamma \doteq -1$ ) und hat zusammen mit der Selbstähnlichkeit eine folgenreiche Konsequenz: Identifiziert man ein Polymer  $A_{2^k}$  mit einem Monomer entsprechend verringerter Konzentration, benötigt man einen größeren autokatalytischen Parameter, um immer noch Längenselektion zu beobachten, weil mit größerer Länge die Konzentration sinkt. Dies wiederum bedeutet, dass man einen größeren autokatalytischen Parameter benötigt, damit ein längeres Polymer ebenfalls selektiert wird. Daher ist die Phasengrenze, welche von der anfänglichen Monomerkonzentration abhängt, nur bis zu einer gegebenen Polymerlänge gültig<sup>\*</sup>:

$$\alpha_c \sim [A_1]_0^\gamma \rightarrow \alpha_{c,k} \sim \frac{1}{2^k} \cdot [A_1]_0^\gamma \quad (3.5)$$

\* Die Abhängigkeit von der maximalen Polymerlänge fällt möglicherweise für sehr große autokatalytische Parameter weg; vgl. Abschnitt 3.6.3

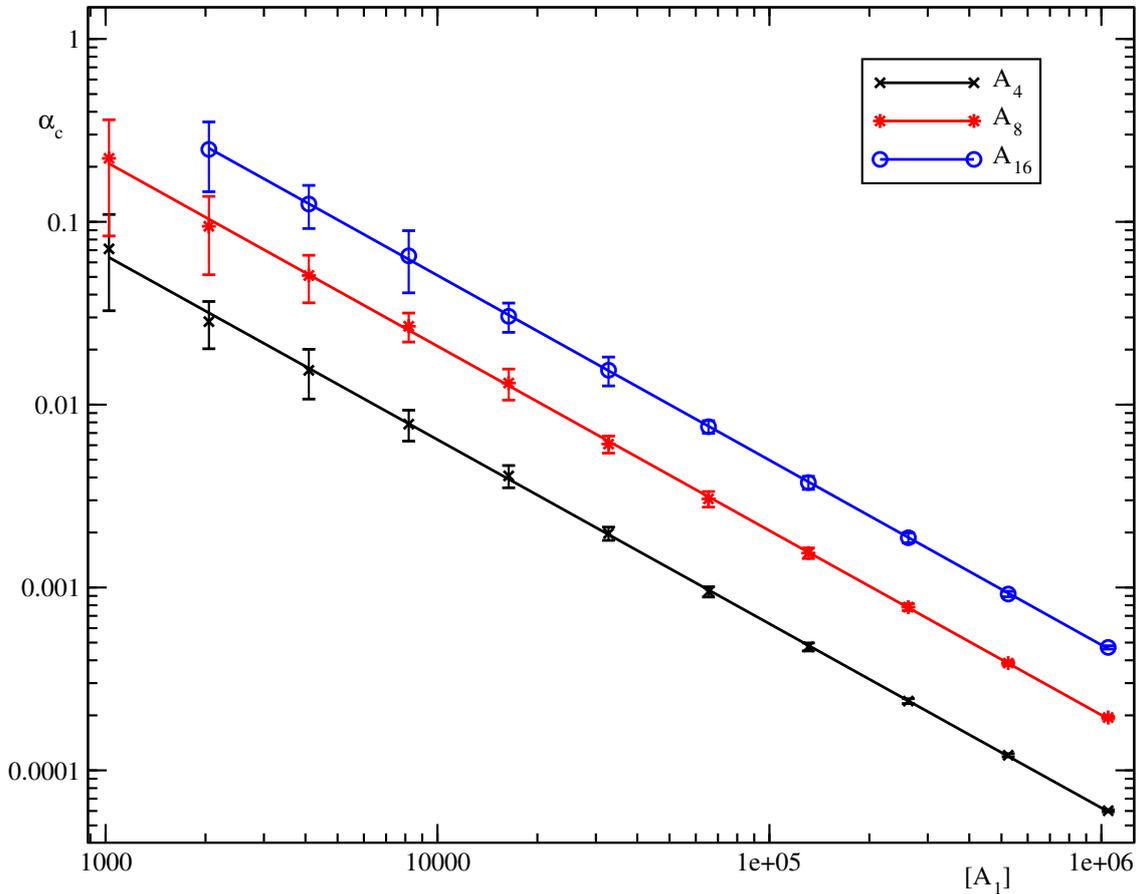


Abbildung 3.8: **Phasenübergang zur Längenverdopplung.** Der kritische autokatalytische Parameter  $\alpha_{c,k}$  ist skaleninvariant und hängt zum einen von der Anfangskonzentration ab:  $\alpha_{c,k} \sim [A_1]_0^\gamma$ ,  $\gamma = -1.0(1)$ . Das System ist selbstähnlich, wenn eine Längenverdopplung vorliegt für  $\alpha > \alpha_{c,k}$ , da, wenn (fast) alle Moleküle zu dem nächsten, doppelt so langen Molekül reagiert sind, man dieses wieder als Ausgangsmolekül mit halbiertes Anfangskonzentration ansehen kann; vgl. Gl. 3.4. Damit ist  $\alpha_{c,k}$  ebenfalls von der maximalen Länge, bis zu der eine Längenverdopplung stattfindet, abhängig und  $\gamma$  negativ. Also ergibt sich näherungsweise  $\alpha_{c,k}$  zu Gl. 3.5. Abbildung veröffentlicht in [82].

so dass wir mit Hilfe von Gl. 3.4 eine Abschätzung für eine Längenselektion bis mindestens zum Polymer  $A_{2^k}$  für  $\alpha > \alpha_{c,k}$  angeben können.

Unterhalb dieses Wertes und/oder für eine größere Polymerlänge entstehen Polymere vielzähliger, unterschiedlicher Längen; vgl. Abschnitt 3.2.1 und es ist nicht mehr möglich, hervorstechende Längen zu identifizieren.

Dabei definieren wir, dass eine Längenselektion bis zum Polymer  $A_{2^m}$  erfolgt, wenn es kein Polymer ( $A_{m+1} \dots A_{2^{m-1}}$ ) gibt, das eine größere Konzentration einnimmt als  $A_{2^m}$ , bis zu dem Zeitpunkt  $t$ , an dem dieses seine Maximalkonzentration ( $[A_{2^m}]_{\max}(t)$ ) erreicht.

Im Folgenden werden wir die Vorgänge während des Plateaus genauer analysieren.

### 3.4.3 Gleichgewicht im Plateau und exponentielles Wachstum

Das charakteristische Verhalten eines Moleküls, welches ein Plateau zeigt, beginnt mit einem exponentiellen Wachstum. Nach einem fließenden Übergang folgt ein Plateau, in dem die Molekülzahl (nahezu) konstant ist. Danach folgt ein exponentieller Abfall.

Im Bereich des Plateaus von  $A_{2^n}$  sind die Erzeugungs- und Verlustrate gleich. Wir nehmen an, dass  $\beta$  vernachlässigbar ist,  $\alpha > \alpha_{c,k}$  mit  $2^k \geq 4n$  und keine weiteren Reaktionen ablaufen als:  $2A_n \rightarrow A_{2n}$  und  $2A_{2n} \rightarrow A_{4n}$ . Dann ist  $A_{2n}$  stationär ( $[\dot{A}_{2n}] = 0$ ) und das Polymer  $A_{4n}$  wächst exponentiell:

$$\frac{d}{dt}[A_{4n}] = \alpha[A_{4n}][A_{2n}]_{\text{Plat.}}^2. \quad (3.6)$$

Aus der Massenerhaltung ( $4[\dot{A}_n] + 2[\dot{A}_{2n}] + [\dot{A}_{4n}] = 0$ ) folgt, dass der Verlust an  $A_n$  ebenfalls exponentiell ist. Die Lebensdauer des exponentiellen Wachstums bzw. Zerfalls von  $A_n$  ist demnach ( $[A_n](t) \sim e^{\pm \frac{t}{\tau_{ew}}}$ ):

$$\tau_{ew} = \alpha^{-1} \cdot [A_{\frac{1}{2}n}]_{\text{Plat.}}^{-2} \geq \alpha^{-1} \cdot 2^{2(k-1)} \cdot [A_1]_0^{-2}, \quad (3.7)$$

mit  $n \equiv 2^k$  und Gl. 3.4. Dieses Ergebnis stimmt mit den Simulationsdaten überein; vgl. Abb. 3.9.

### 3.4.4 Plateaulänge

Im überkritischen Fall ( $\alpha > \alpha_{c,k}$ ) entstehen Plateaus, die in ihrer Länge als auch Form unterschiedlich sind. Mit zunehmendem autokatalytischem Parameter  $\alpha$  wird die Form der Plateaus eckiger, da die Wachstumsreaktionen schneller ablaufen (Gl. 3.7). Dadurch ändert sich auch die Struktur der Plateaus und auch deren Entstehungsprozess: Bei schwach überkritischer Autokatalyse ( $\alpha \gtrsim \alpha_{c,k}$ ) ist die Plateauform runder und leicht ansteigend und hat eine deterministische Länge. Wenn die Autokatalyse größer gewählt wird ( $\alpha \gg \alpha_{c,k}$ ), ist die Plateaulänge statistisch verteilt und die Form eckiger; vgl. Nebengraphen Abb. 3.13.

Der Entstehungsprozess ist für ( $\alpha \gtrsim \alpha_{c,k}$ ) wie folgt: Wegen der schwachen Autokatalyse entsteht schon früh das Nachfolgemolekül  $A_{2n}$ , wenn  $A_n$  in der Plateauphase ist bzw. hineinkommt. Ersteres wächst aber nicht schnell an, weil die Reaktionen (Gl. 3.6), die zur Bildung der Plateauphase beitragen, auf Grund der größeren Konzentrationen der beteiligten Polymere ein größeres Gewicht haben, so dass die Aufbaureaktion von  $A_n$  fast bis zum Ende abläuft,

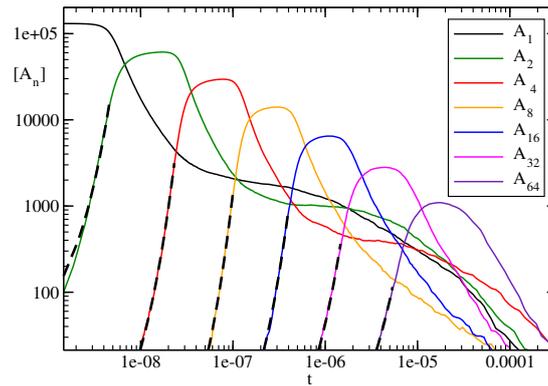


Abbildung 3.9: **Exponentielles Wachstum** der selektierten Polymere (gestrichelte Kurven). Die Lebensdauern sind  $6,67(5) \cdot 10^{-10}$  für  $A_2$ ;  $2,63(7) \cdot 10^{-9}$  für  $A_4$ ;  $1,1(6) \cdot 10^{-8}$  für  $A_8$ ;  $5,0(5) \cdot 10^{-8}$  für  $A_{16}$ ;  $2,1(4) \cdot 10^{-7}$  für  $A_{32}$  und  $1,1(4) \cdot 10^{-6}$  für  $A_{64}$ . Die Daten sind identisch mit Abb. 3.3. Veröffentlicht in [82].

bevor das Nachfolgemolekül  $A_{2n}$  stark wächst. Dies führt auch zu einer Bestimmung der Länge der Plateaus: Diese ist im Wesentlichen von der Stärke der Autokatalyse abhängig und daher für verschiedene Realisierungen konstant.

Im Fall einer sehr starken Autokatalyse ( $\alpha \gg \alpha_{c,k}$ ) ist der Entstehungsprozess ein anderer: Die starke Autokatalyse lässt das Polymer  $A_n$  so schnell und vollständig wachsen, dass währenddessen kaum eine andere Reaktion ablaufen kann. Da nun nur noch eine spontane Reaktion ( $2A_n \rightarrow A_{2n}$ ) ablaufen kann und diese selten ist, ist die (Warte-)Zeit bis zum Reaktionsanfang (und damit dem Ende des Plateaus des Polymers  $A_n$ ) exponentialverteilt [129]:

$$\text{expv}(t; \tau) = \frac{1}{\tau} e^{-\frac{t}{\tau}} \quad (3.8)$$

$$\text{EXPV}(t; \tau) = 1 - e^{-\frac{t}{\tau}} \quad (3.9)$$

mit  $t \geq 0$  und  $\tau > 0$ .

Im Folgenden analysieren wir die Länge zweier aufeinanderfolgender Plateaus unterschiedlicher Polymerlängen für verschiedene Parameter  $\alpha$ . Wir bezeichnen mit  $\tau_{A_n}$  ( $n \equiv 2^k$ ) den Erwartungswert der Zeit, in der die Konzentration des Polymers  $A_n$  größer als 70% der theoretischen Maximalkonzentration ist (vgl. Gl. 3.4). Der tatsächliche Prozentsatz hat keinen Einfluss auf das qualitative Ergebnis.

Aus den Daten kann der Parameter  $\tau$  aus Gl. 3.8 mittels eines Maximum-Likelihood-Schätzers bestimmt werden\*. Für die Exponentialverteilung ist der ML-Schätzer  $\bar{\tau}$  der Mittelwert der Stichprobe, welchen wir im Folgenden verwenden werden.

\* Der **Maximum-Likelihood-Schätzer** gibt den plausibelsten Parameterwert einer bekannten Verteilung an. Er wird durch Maximieren der Likelihood-Funktion bestimmt, die die Wahrscheinlichkeitsdichte der Zufallsvariablen darstellt. Ein großer Wert der Likelihood-Funktion stellt einen Satz wahrscheinlicher Ereignisse dar. Ziehen wir nun eine Stichprobe aus der Grundgesamtheit mit einer parameterabhängigen Wahrscheinlichkeitsdichte, so können wir durch Maximieren der Likelihood-Funktion den besten Parameter wählen. Der Maximum-Likelihood-Schätzer ist der maximierte Parameter. Siehe [130].

k	n	$\overline{\tau_{A_n}}$	$[A_1]^{-2} \cdot 2^{2k}$
2	4	$2,35 \cdot 10^{-10}$	$2,33 \cdot 10^{-10}$
3	8	$9,37 \cdot 10^{-10}$	$9,31 \cdot 10^{-10}$
4	16	$3,79 \cdot 10^{-9}$	$3,73 \cdot 10^{-9}$
5	32	$1,50 \cdot 10^{-8}$	$1,49 \cdot 10^{-8}$

Tabelle 3.1: **ML-Schätzer**  $\overline{\tau_{A_n}}$  der kumulierten Exponentialverteilungen (Gl. 3.9) aus Abb. 3.10 und dessen theoretische Abschätzung (Gl. 3.10). Mit sich verdoppelnder Moleküllänge vervierfacht sich die charakteristische Zeit. Gl. 3.10 ist eine untere Schranke für  $\tau_{A_n}$  und beide Werte stimmen bis auf die erste Nachkommastelle überein.

### 3.4.5 Exponentielle Verteilung der Plateaulängen für $\alpha \gg \alpha_c$

Die Plateaulängen sind bei einer großen autokatalytischen Rate ( $\alpha \gg \alpha_c$ ) exponentiell verteilt und die charakteristische Zeit ergibt sich mit Gl. 3.4 zu:

$$\tau_{A_n} = \frac{1}{\beta} \cdot [A_n]_{\text{Plat.}}^{-2} \geq \frac{1}{\beta} \cdot 2^{2k} \cdot [A_1]_0^{-2} \quad (3.10)$$

$\tau_{A_n}$  ist die Zeit bis zum nächsten Auftreten des Polymers  $A_n$ . Um die durch Simulationen ermittelte Plateaulängen gut mit einer Exponentialverteilung vergleichen zu können, betrachten wir die kumulative Verteilungsfunktion. Dafür werden die durch Simulationen ermittelten Plateaulängen sortiert und aufsummiert und ergeben die Zeitabhängigkeit. Die normierte Anzahl wird dann als Funktion der Zeit aufgetragen; vgl. Abb. 3.10. Die analytische Lösung (Gl. 3.9) hat nur die charakteristische Zeit als freien Parameter, welche durch den Maximum-Likelihood-Schätzer ermittelt wird.

Im Folgenden analysieren wir die Abhängigkeit der charakteristischen Zeit von den verschiedenen Systemparametern.

#### Vergleich von gemessener mit theoretischer charakteristischer Zeit

Für einen festen Parametersatz und verschieden lange Polymere ist ein Vergleich in Abb. 3.10 dargestellt. Bei lediglich 10.000 Realisationen des Parametersatzes stimmen die gemessenen und analytischen Verteilungen gut überein. Damit sind die Annahmen aus Abschnitt 3.4.4 als plausibel anzusehen: Das Wachstum der Polymere ist so schnell, dass keine andere Reaktion abläuft und die Zeit bis zum nächsten Wachstum nur durch einen Warteprozess, welcher exponentialverteilt ist, bestimmt wird.

Die Abschätzung der charakteristischen Zeit durch Gl. 3.10 ist in Tabelle 3.1 aufgelistet. Die Abschätzung ist eine untere Schranke und stimmt mit den gemessenen charakteristischen Zeiten bis auf die erste Nachkommastelle überein.

#### Vergleich von Verteilung verschiedener Anfangskonzentrationen

Die charakteristische Zeit ist nach Gl. 3.10 abhängig von der Anfangskonzentration und der Größe des längenverdoppelten Polymers  $A_{2^k}$ . Die charakteristische Zeit verschiedener Anfangskonzentrationen, die einer Zweierpotenz entsprechen, können demnach durch die

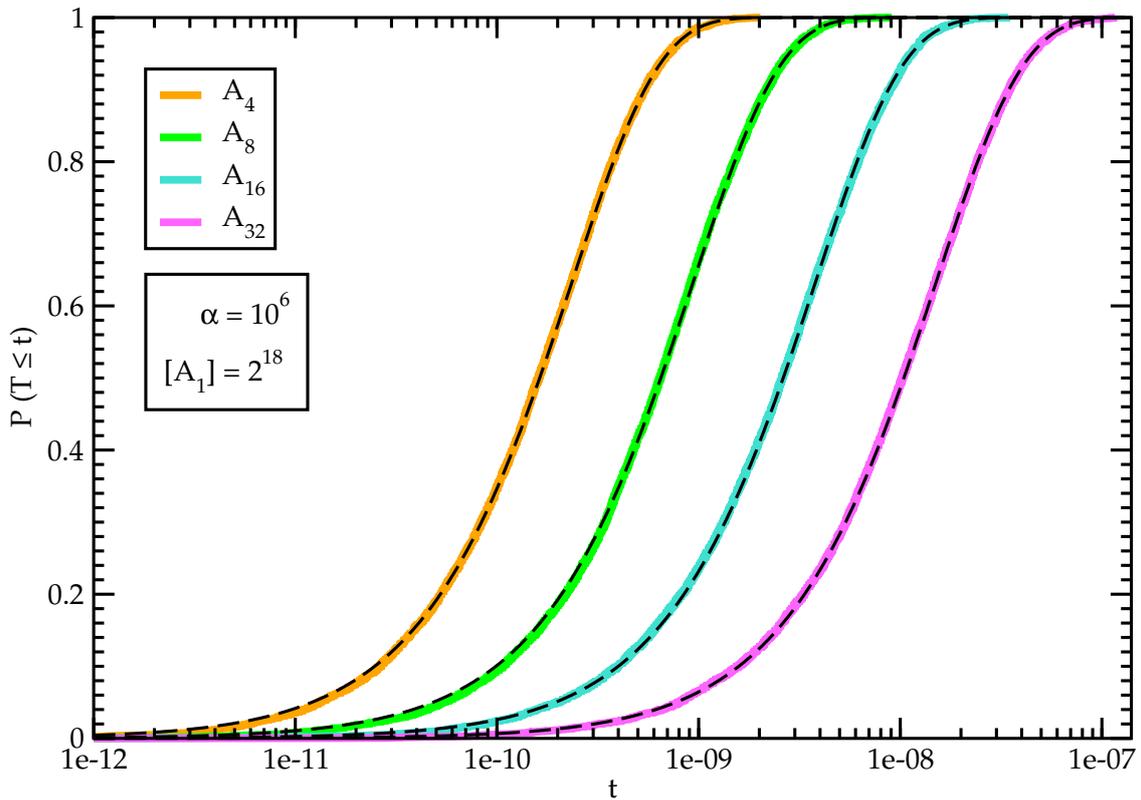


Abbildung 3.10: **Kumulierte Plateaulängenverteilung** für  $\alpha \gg \alpha_c$  und verschiedene Längen der Spezies  $n$ . Parameter sind:  $\alpha = 10^6$ ,  $[A_1]_0 = 2^{18}$ . Die gemessene Verteilung ist für die verschiedenen Moleküle in der Plateauphase farblich dargestellt. Eine Exponentialverteilung mit dem ML-Schätzer (Tab. 3.1) der gemessenen Verteilung ist zum Vergleich schwarz gestrichelt eingezeichnet. Die Verteilungen stimmen gut überein. Pro Verteilung wurden 10.000 Messwerte betrachtet. Veröffentlicht in [82].

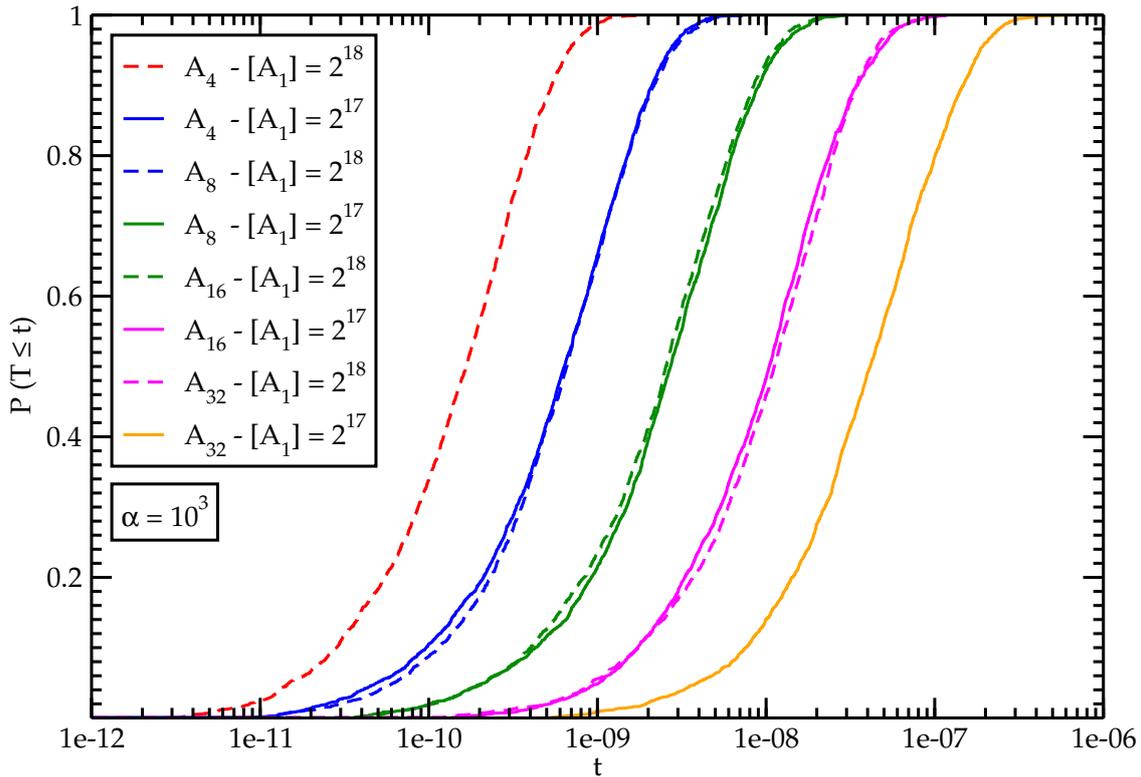


Abbildung 3.11: **Vergleich der kumulierten Plateaulängenverteilung** für verschiedene Anfangskonzentrationen  $[A_1]_0$  bei  $\alpha_c \ll \alpha = 10^3$ . Für  $[A_1]_0 = 2^{17}$  sind die Kurven durchgezogen und für  $[A_1]_0 = 2^{18}$  gestrichelt. Verteilungen mit vergleichbarer Molekülzahl in der Plateauphase haben eine vergleichbare Verteilung, siehe auch Tab. 3.2. Dies ist auf die Selbstähnlichkeit des Systems zurückzuführen. Die kumulierten Verteilungen beinhalten 2.500 Datenpunkte.

Betrachtung eines geeigneten längenverdoppelten Polymers zur Deckung gebracht werden, wenn sie entsprechend selbstähnlich gewählt wurden. Dies ist beispielhaft in Abb. 3.11 und Tab. 3.2 für eine verdoppelte Anfangskonzentration dargestellt. Folglich liegen diese Verteilungen idealerweise deckungsgleich auf der verdoppelten Polymerlänge: Die Verteilung von  $A_4$  bei  $[A_1]_0 = 2^{17}$  Anfangskonzentration wäre also deckungsgleich mit der von  $A_8$  bei  $[A_1]_0 = 2^{18}$ . Wie aus Abb. 3.11 liegen die beiden Verteilungen bis auf eine geringe Abweichung gut übereinander und auch die Maximum-Likelihood-Schätzer der Verteilungen in Tab. 3.2 stimmen auf der ersten Stelle überein. Damit sind die Betrachtungen zur Selbstähnlichkeit als plausibel anzusehen. Im Folgenden werden wir die Abhängigkeiten der charakteristischen Zeit genauer untersuchen.

$[A_1]_0 = 2^{17}$		$[A_1]_0 = 2^{18}$	
n	$\overline{\tau_{A_n}}$	n	$\overline{\tau_{A_n}}$
4	$9,17 \cdot 10^{-10}$	4	$2,34 \cdot 10^{-10}$
8	$3,96 \cdot 10^{-9}$	8	$9,48 \cdot 10^{-10}$
16	$1,51 \cdot 10^{-8}$	16	$3,71 \cdot 10^{-9}$
32	$6,23 \cdot 10^{-8}$	32	$1,56 \cdot 10^{-8}$

Tabelle 3.2: ML-Schätzer  $\overline{\tau_{A_n}}$  der kumulierten Exponentialverteilungen  $1 - e^{-t/\tau_{A_n}}$  aus Abb. 3.11 zum Vergleich.

### Abhängigkeit der charakteristischen Zeit von der Anfangskonzentration und dem autokatalytischen Parameter

Wir haben schon die Abhängigkeit von der Anfangskonzentration verdeutlicht und wollen jetzt zusätzlich auch die Abhängigkeit vom autokatalytischen Parameter analysieren. Dies ermöglicht es uns, die Abhängigkeiten in Gl. 3.10 zu überprüfen.

Solange die Plateaulänge statistisch verteilt ist, hat Gl. 3.10 Gültigkeit. Die Abhängigkeit der charakteristischen Zeit  $\overline{\tau_{[A_1]_0}}$  von der Größe der längenverdoppelten Polymere  $n = 2^k$  ist in Abb. 3.12.a dargestellt und gehorcht einem Potenzgesetz mit einem Exponenten von ungefähr 2, was mit Gl. 3.10 übereinstimmt. Dies gilt für alle autokatalytischen Parameter, die zu einer statistischen Verteilung der Plateaulängen führen. Der Übergang hin zu kleineren autokatalytischen Parametern  $\alpha$  ist in Abb. 3.13 wiedergegeben. Die Verteilungen überlappen im Fall statistisch verteilter Plateaus, deren Konzentrationsverlauf eckig aussieht. Bei kleineren  $\alpha$  werden die Plateaus runder und die Verteilung ist insbesondere am Anfang steiler und keine Exponentialverteilung mehr. Diesen Fall werden wir im folgenden Abschnitt 3.4.6 genauer untersuchen.

Der Zusammenhang zwischen der charakteristischen Zeit  $\overline{\tau_\alpha}$  und der Anfangskonzentration  $[A_1]_0$  ist in Abb. 3.12.b wiedergegeben. Er ist ebenfalls skaleninvariant mit einem Exponenten von nahezu  $-2$ , was sich ebenfalls mit dem analytischen Wert von Gl. 3.10 deckt. Dass bei einer gleichzeitigen Verdopplung von Anfangskonzentration und Polymerlänge die charakteristische Zeit gleich bleibt, haben wir bereits beispielhaft in Abb. 3.11 gezeigt. Dieser Zusammenhang ist auch in Abb. 3.12.b gut zu erkennen, da entsprechende Werte annähernd konstant bleiben (nicht eingezeichnet).

### 3.4.6 Deterministische Plateaulängen für $\alpha \gtrsim \alpha_c$

Für Werte des autokatalytischen Parameters  $\alpha$  in der Nähe von  $\alpha_c$  haben die Plateaus eine deterministische Länge und im Konzentrationsverlauf eine rundere Form. In Abb. 3.14 ist die kumulierte Verteilung zusammen mit der entsprechenden exponentiellen Verteilung dargestellt. Es wird deutlich, dass diese nicht übereinstimmen und die Plateaulänge scharf begrenzt ist. Weniger als 6 % weichen von diesem Mittelwert deutlich ab, wobei es auch zu einer geringfügig ( $<3\%$ ) fehlerhaften Erkennung der Plateaulänge, die sich prinzipbedingt

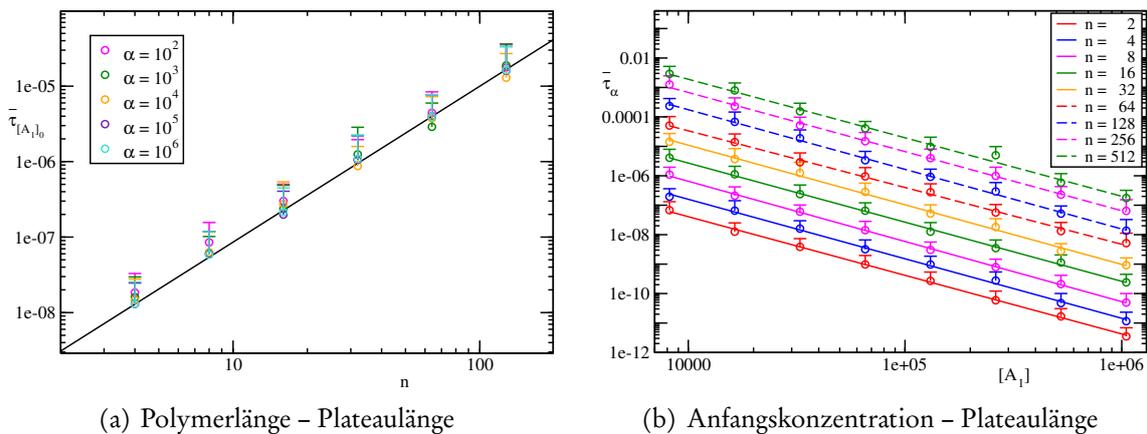


Abbildung 3.12: **Plateaulänge**  $\overline{\tau}_{[A_1]_0}$  und  $\overline{\tau}_\alpha$  in Abhängigkeit von der Anfangskonzentration und des autokatalytischen Parameters. (a) Verschiedene  $\alpha > \alpha_{c,k}$  in Abhängigkeit von der Polymerlänge sind skaleninvariant und nähern sich für große  $\alpha$  einer Kurve mit dem Exponenten 2,06(2) (für  $\alpha = 10^6$ ) an. Gezeigt ist  $[A_1]_0 = 2^{15}$ . (b) Die Plateaulänge ist in Abhängigkeit von der Anfangskonzentration auch für verschiedene Polymerlängen skaleninvariant mit dem Exponenten -2,00(3). Die Kurven sind parallel in doppeltlogarithmischer Darstellung wegen der Selbstähnlichkeit und weisen einen größeren Vorfaktor für größere Polymerlängen auf. Dargestellt sind die Kurven für  $\alpha = 10^3$ . In beiden Fällen wurde über 50 Realisierungen gemittelt. Veröffentlicht in [82].

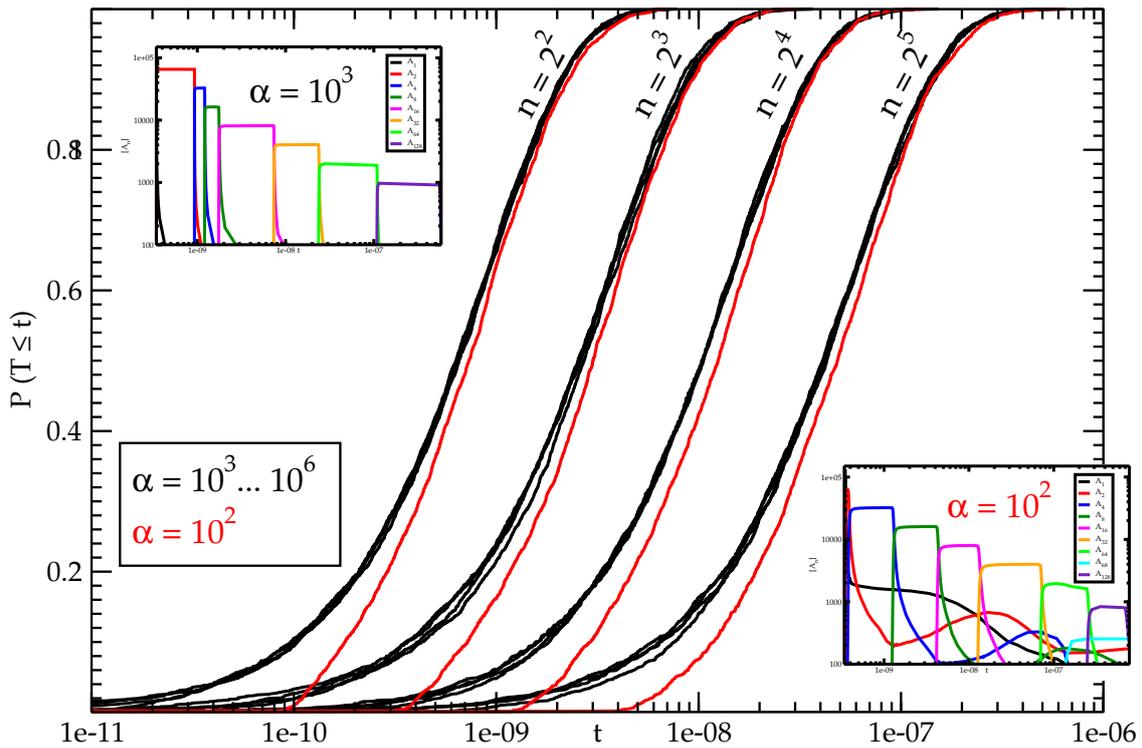


Abbildung 3.13: Vergleich der kumulierten Plateaulängenverteilung für verschiedene  $\alpha$  bei einer Anfangskonzentration von  $[A_1]_0 = 2^{17}$ . Für Werte von  $\alpha \geq 10^3$  fallen die Verteilungen aufeinander. Für kleinere Werte von  $\alpha$  (hier:  $\alpha = 10^2$ ) beginnt der Übergang zur fixen Plateaulänge, siehe Abb. 3.14. Zum Vergleich sind die entsprechenden Zeitentwicklungen eingefügt. Man sieht, dass der Konzentrationsanstieg in die Plateauphase sich für kleinere  $\alpha$  abrundet. Die Verteilungen bestehen jeweils aus 2.500 Datenpunkten.

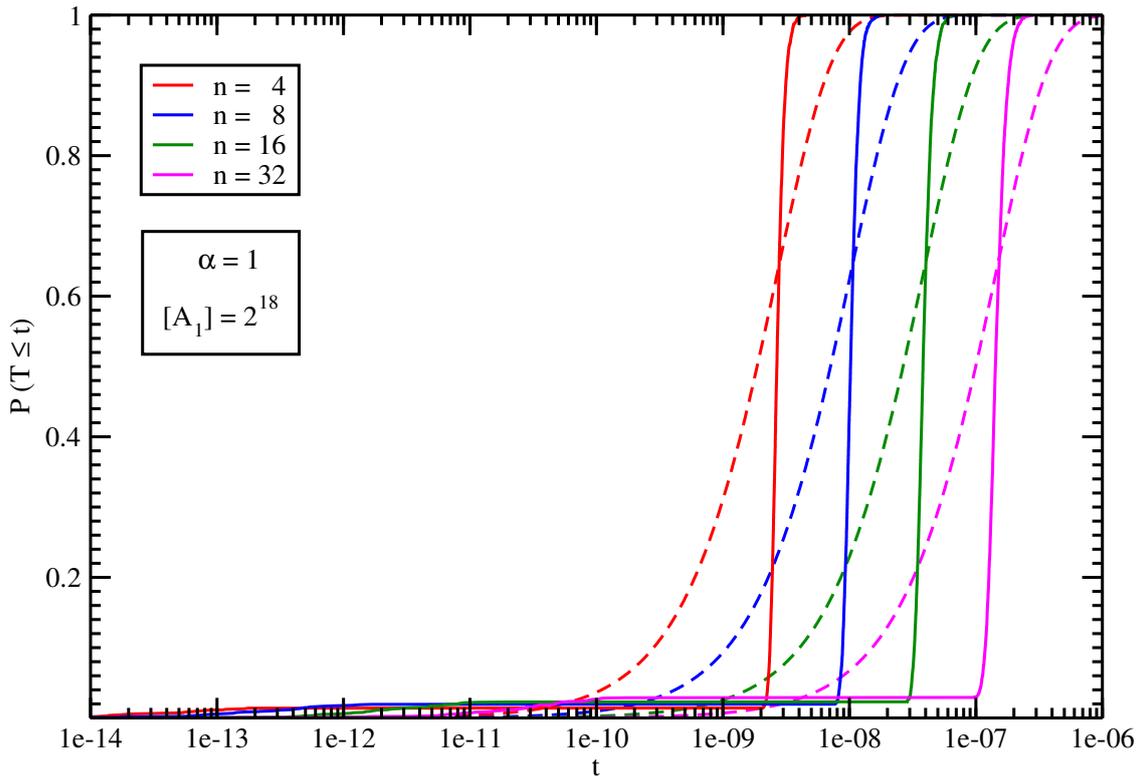


Abbildung 3.14: **Kumulierte Plateaulängenverteilung** für  $\alpha \gtrsim \alpha_c$  und verschiedene Längen der Spezies  $n$ . Parameter sind:  $\alpha = 1$ ,  $[A_1]_0 = 2^{18}$ . Die Verteilung ist nicht mehr exponentiell wie in Abb. 3.10 (gestrichelte Kurven sind exponentiell verteilt mit jeweiligem ML-Schätzer der dargestellten und gemessenen Verteilung). Vielmehr ist die Plateaulänge scharf begrenzt. Weniger als 4 % sind deutlich kürzer und weniger als 2 % größer. Für Mittelwerte und Standardabweichung siehe Tab. 3.3. Die kürzeren Plateaulängen sind schon zu sehr frühen Zeitpunkten vertreten. Das ist darauf zurückzuführen, dass 35 bis 74 Plateaus von den nominell 2.500 Realisierungen mehrfach erkannt wurden.

nicht beseitigen lässt, durch den auswertenden Algorithmus wegen statistischer Schwankungen kommt. Der Mittelwert der Plateaulänge ist in Tab. 3.3 aufgelistet. Bemerkenswert ist die geringe Standardabweichung.

Bisher ist kein analytischer Ausdruck für die Plateaulänge bekannt. Diese ist jedoch eine Funktion der Lebensdauer  $\tau_{ew}$  aus Gl. 3.7, da der Prozess der Plateaubildung auf dem sequentiellen Wachstum der längenverdoppelten Polymere beruht, wie wir in Abschnitt 3.4.4 gesehen haben. Wir konnten einen mit der Polymerlänge abnehmenden Proportionalitätsfaktor zwischen 39 und 47 bestimmen.

k	n	N	$\bar{t}$	$\bar{\sigma}$	$\tau_{ew}$	$\bar{t}/\tau_{ew}$
2	4	2383	$2,72 \cdot 10^{-9}$	$7,53 \cdot 10^{-20}$	$5,82 \cdot 10^{-11}$	$4,67 \cdot 10^1$
3	8	2397	$1,04 \cdot 10^{-8}$	$1,33 \cdot 10^{-18}$	$2,33 \cdot 10^{-10}$	$4,46 \cdot 10^1$
4	16	2404	$3,90 \cdot 10^{-8}$	$2,32 \cdot 10^{-17}$	$9,31 \cdot 10^{-10}$	$4,19 \cdot 10^1$
5	32	2420	$1,47 \cdot 10^{-7}$	$4,48 \cdot 10^{-16}$	$3,73 \cdot 10^{-9}$	$3,94 \cdot 10^1$

Tabelle 3.3: Mittelwerte  $\bar{t}$  und Standardabweichung  $\bar{\sigma}$  der konstanten Plateaulänge aus Abb. 3.14 für die verschiedenen Moleküle  $A_n$  in der Plateauphase. Die Standardabweichung ist sehr gering, was auf eine sehr konstante Plateaulänge hinweist. Diese ist von der Lebensdauer  $\tau_{ew}$  aus Gl. 3.7 abhängig (siehe Text). N der nominell 2.500 Realisierungen sind im Bereich  $0.04 < P < 0.98$ , in dem die Plateaulänge sich nur schwach ändert. Es wurden 35 bis 74 Plateaus mehrfach erkannt.  $\bar{t}/\tau_{ew}$  wurde auf 3 Stellen gerundet.

### Zusammenfassung

In diesem Abschnitt haben wir den Phasenübergang und die Plateaulänge des theoretischen Systems vorgestellt und analysiert. Im unterkritischen Fall ( $\alpha < \alpha_c$ ) entstehen der Reihe nach alle möglichen Polymerlängen und eine Vielzahl von unterschiedlichen Reaktionen findet gleichzeitig statt. Im überkritischen Fall ( $\alpha > \alpha_c$ ) werden bestimmte Polymere, die nach und nach eine doppelte Länge aufweisen, selektiert. Die Zahl der Reaktionen ist begrenzt und die längenverdoppelten Polymere zeigen eine Plateauphase. Die Länge der Plateauphase kann je nach  $\alpha$  deterministisch oder exponentiell verteilt sein. Die zu Grunde liegenden Prozesse wurden vorgestellt und hängen stark mit der Selbstähnlichkeit der längenverdoppelten Polymere zusammen, so dass wir analytische Abschätzungen insbesondere für die statistisch verteilte Plateaulänge angeben konnten.

Im Folgenden werden wir Erweiterungen dieses Systems vorstellen, die das Verständnis in Bezug auf die Evolution erweitern.

## 3.5 Betrachtungen zur Allgemeingültigkeit der Analysen

Bisher haben wir in diesem Kapitel hauptsächlich physikalisch relevante Eigenschaften des Systems beschrieben und analysiert sowie das grundlegende Verhalten im Hinblick auf die Evolution beschrieben. Im Folgenden werden wir das System sinnvoll erweitern, so dass der Einfluss bestimmter Modifikationen auf die Evolution verdeutlicht wird.

Zunächst betrachten wir bezüglich der Vereinbarkeit von Theorie und Experiment den Einfluss experimenteller Gegebenheiten.

### 3.5.1 Evolution mit korrigierten Raten für DNA und Ligase

In diesem Abschnitt wollen wir den Einfluss der Ligaseaktivität und des Anteils der hybridisierten DNA berücksichtigen, um einen besseren Vergleich zum Experiment in Kapitel 2 ziehen zu können. Wie wir bereits in Abb. 2.1 gesehen haben, hat die Ligase eine temperaturabhängige Aktivität. Ebenfalls konnten wir in Kapitel 2.4.2 die Temperaturabhängigkeit

### 3 Theorie

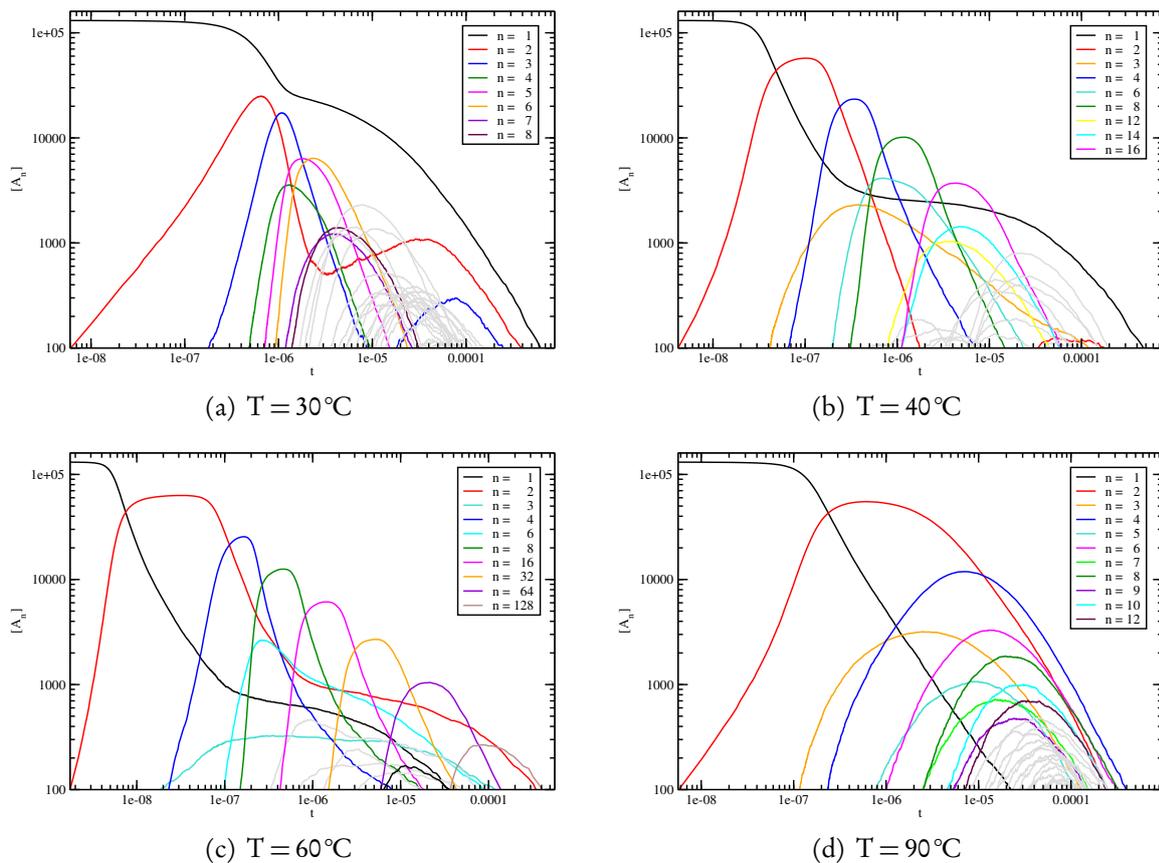


Abbildung 3.15: **Einfluss Ligaseaktivität und DNA-Hybridisierungsanteil.** Die Ligaseaktivität hat ein Maximum bei  $T \approx 62^\circ\text{C}$ ; vgl. Abb. 2.1. Während die Schmelztemperatur  $T_m$  für lange DNA-Stränge identisch ist, fällt diese für kürzere Stränge jedoch stark ab; vgl. Abschnitt 2.4.2. Damit ist es kürzeren Strängen bei höheren Temperaturen nicht gut möglich zu hybridisieren. (a) Ganz kurze DNA-Stränge ( $n = 2$ ; 20 bp) haben wegen ihrer sehr kleinen Schmelztemperatur ( $T_m \approx 30^\circ\text{C}$ ) eine kleinere effektive autokatalytische Rate als geringfügig längere ( $n = 3$ ; 30 bp;  $T_m \gtrsim 40^\circ\text{C}$ ) Stränge; vgl. Abb. 2.8.a. Daher wachsen die längeren ( $n \geq 4$ ; 40 bp) schneller. Wegen der geringeren Ligaseaktivität ist die effektive autokatalytische Rate jedoch so niedrig, dass keine Längenselektion stattfindet. (b) Praktisch unbeeinflusste Entwicklung bei ca.  $1/3$  der Ligaseaktivität. (c) Die größte Ligaseaktivität ermöglicht ein unbeeinflusstes Wachstum. Da die Stränge der Länge  $n \lesssim 4$  (40 bp) einen geringeren Hybridisierungsanteil aufweisen, wachsen sie langsamer. (d) Die Ligaseaktivität ist so niedrig, dass das Ergebnis für längere Stränge ( $n \gtrsim 4$ ; 40 bp) einer „wahllosen Wucherung“ ähnelt. Die kürzeren haben wegen ihrer geringen Schmelztemperatur eine noch niedrigere Reaktionsrate, so dass sie sich am Anfang von den längeren separieren. Die Anfangskonzentration ist  $[A_1]_0 = 2^{17}$  und der autokatalytische Parameter  $\alpha = 0,1$ .

der DNA vorstellen. Bereits in Abschnitt 2.4.1 haben wir die Abhängigkeit der Schmelztemperatur von den beiden Eduktkonzentrationen angegeben. Dies sind die bekannten bzw. berechenbaren Größen, um Korrekturen hinsichtlich der Hybridisierungseigenschaften der DNA zu berücksichtigen. Um eine vollständige Berechnung der Hybridisierung und Ligation mehrerer DNA-Stränge in einer Lösung realisieren zu können, fehlen insbesondere die freien DNA-Stränge (zur Anpassung der spontanen Reaktionsrate) und die gegenseitige Wechselwirkung mehrerer Edukte auf einem Template (Multistranginteraktion). Diese Prozesse müssen explizit implementiert werden, wie Berechnungen und Experimente in einem Microarray\* gezeigt haben, da sie nicht direkt in einen stationären Zustand relaxieren. Insbesondere kann ein nicht genau komplementärer Strang anfangs auf dessen Zielstrang stärker binden als das perfekte Komplement des Zielstranges, wenn dieser geringfügig (3-fach) geringere Binde- und Ablöseraten aufweist [131]. Eine Multistranginteraktion und die daraus resultierenden, freien DNA-Stränge ließen sich also nur mit einem erheblichen Aufwand korrekt berechnen [75], so dass wir in dieser Arbeit einen einfacher zu berechnenden Ansatz wählen.

Nach Gleichung 2.2 ist die Schmelztemperatur im Gleichgewicht auch von den jeweiligen Konzentrationen abhängig. Die benötigten Enthalpien ( $\Delta H$ ) und Entropien ( $\Delta S$ ) werden ebenfalls von UNAFold berechnet und könnten ähnlich zu Abschnitt 2.4.2 phänomenologisch dargestellt werden. Dieses Vorgehen würde jedoch den relaxierenden Charakter der Multistranginteraktion vernachlässigen und wir sehen daher davon ab, dies zu implementieren. Wir betrachten hier deshalb lediglich die Korrektur bezüglich der auf das Maximum normierten Ligaseaktivität  $L(T)$  (Abb. 2.1) und den temperaturabhängigen Hybridisierungsanteil der DNA aus Abschnitt 2.4.2. Gleichung 3.2 lässt sich dann mit Gl. 2.3 schreiben als:

$$\frac{d}{dt}[A_{n+m}] = (\iota\alpha[A_{n+m}] + \beta)[A_n][A_m] - \sum_k (1 + \delta_{n+m,k})(\iota\alpha[A_{n+m+k}] + \beta)[A_{n+m}][A_k], \quad (3.11)$$

mit  $\iota = L(T) \cdot \chi(T, n, n+m) \cdot \chi(T, m, n+m)^{**}$ . Es wird angenommen, dass sich das Produkt vom Template instantan nach der Ligation trennt und dass alle Edukte für eine spontane Ligation zur Verfügung stehen, ähnlich den Zyklen in Kapitel 2.

Wir betrachten nun die Zeitentwicklung bei bestimmten, konstanten Temperaturen; vgl. Abb. 3.15. Die effektiven Reaktionsraten hängen von der Ligase-Aktivität mit einem Maximum bei  $T = 62^\circ\text{C}$  ab, so dass größere ( $> 90^\circ\text{C}$ ) und kleinere ( $< 30^\circ\text{C}$ ) Temperaturen insbesondere für längere Stränge ( $n \geq 4$ ; 40 bp) ein Verhalten ähnlich zur „wahllosen Wucherung“ zeigen. Bei großen Temperaturen hybridisieren die kurzen Stränge kaum noch, so dass diese Reaktionen langsamer ablaufen und sich separieren; vgl. Abb. 3.15.d. Für niedrige Temperaturen gibt es zudem noch einen Unterschied im Hybridisierungsanteil der kürzeren DNA-Stränge, welcher bis  $n = 8$  (80 bp) zunimmt. Daher laufen die Reaktionen dieser Stränge mit deren Länge schneller ab, so dass diese nacheinander auftreten; vgl. Abb. 3.15.a.

\* Ein **Microarray** besteht aus fixierten DNA-Strängen in einer Lösung, an die sich freie, in Lösung befindliche DNA-Stränge binden können [131]. Zur Detektion können diese fluoreszierend markiert werden [131].

\*\* In den hier vorgestellten Beispielrechnungen wurde ein anderer Parametersatz für  $\chi(T, \mathbf{n}, \mathbf{m})$  verwendet, da der in Abschnitt 2.4.2 vorgestellte Parametersatz aus einer nochmaligen, händischen Optimierung der Startwerte der Fitparameter resultiert. Die hier verwendeten Parameter sind:  $d = 0,556801$ ,  $o = -0,0287587$ ,  $f = -0,0454737$ ,  $g = 2,04354$ ,  $h = -0,00350419$ ,  $i = 1,02461$ ,  $j = -5,74185$ ,  $k = 0,0376826$ ,  $l = 86,2402$ ,  $p = 0,220917$ ,  $q = 0,0119038$ ,  $r = -0,460872$ ,  $s = 9,10303$  und  $t = -4,46667$ .

Für mittlere Temperaturen gibt es kaum Abweichungen zur normalen Entwicklung, siehe Abb. 3.3. Bei erhöhten Temperaturen ( $\approx 60^\circ\text{C}$ ) sind kurze Stränge ( $n \leq 4$ ; 40 bp) nicht mehr so häufig hybridisiert, so dass deren Reaktionen entsprechend langsamer ablaufen. Diese Temperaturen sind experimentell relevant und können je nach Stranglänge angepasst werden, um eine gleichbleibende Wachstumsrate zu erzielen; vgl. Kap. 2.

#### Zusammenfassung

Wir konnten das Modell um den Beitrag der Ligaseaktivität und des paarweisen Hybridisierungsanteils der DNA zur Autokatalyse erweitern. Bei physikalisch relevanten Temperaturen ist der Einfluss auf die Längenselektion nur nahe der Phasengrenze relevant. Temperaturänderungen können aber dazu genutzt werden, die autokatalytische Rate für alle DNA-Längen gleich einzustellen. In den in Kap. 2 vorgestellten Experimenten ist der autokatalytische Parameter unterkritisch, so dass eine Temperaturanpassung diesen effektiv erhöhen kann.

Wir gehen nun dazu über, den autokatalytischen Term so anzupassen, dass auch parabolisches und hyperbolisches Wachstum möglich ist, und untersuchen das Verhalten auf Längenselektion. Diese Wachstumstypen treten z. B. bei templatebasierten Prozessen oder sexueller Fortpflanzung auf [81].

#### 3.5.2 Evolution mit parabolischem, exponentiellem und hyperbolischem Wachstum

In den folgenden Abschnitten stellen wir die verschiedenen Wachstumstypen in einer Reaktion nach Gl. 3.1 und deren Besonderheiten im Konzentrationsverlauf vor. Anschließend vergleichen wir die Ergebnisse untereinander.

##### Gemeinsamkeiten der Wachstumstypen

Falls der autokatalytische Parameter so klein gewählt wurde, dass das autokatalytische Wachstum keine Rolle spielt, sind die Ergebnisse identisch zum Abschnitt 3.2.1. Wir werden im Folgenden daher nur eine kurze Darstellung und Beschreibung in der Bildunterschrift geben.

Ebenfalls übertragbar sind die Aussagen zur Selbstähnlichkeit in den Abschnitten 3.2.2 und 3.4.1, falls sich eine Längenverdopplungskaskade ausbildet. Für genügend große Werte des autokatalytischen Parameters  $\alpha$  ist wegen des schnellen Wachstums in allen Fällen die Plateauverweildauer groß gegenüber der Wachstumsdauer, so dass eine statistische Plateaulänge auftritt. Demnach sind die Aussagen zur statistischen Verteilung der Plateaulänge aus den Abschnitten 3.2.2, 3.4.4 und 3.4.5 ebenfalls übertragbar.

##### Evolution mit parabolischem Wachstum

Das parabolische Wachstum zeigt im unterkritischen Bereich des autokatalytischen Parameters nahe der Phasengrenze eine charakteristische Besonderheit. Die Polymerkonzentrationsverläufe weisen einen parabelähnlichen Verlauf auf, der einen steileren Anstieg hat, und treten mit deren Maxima sortiert auf ( $n \leq 10$ ). Bemerkenswerterweise treten bestimmte Längen nicht auf. Die einzelnen Konzentrationsverläufe schließen einander nicht ein; vgl. Abb. 3.16.b.

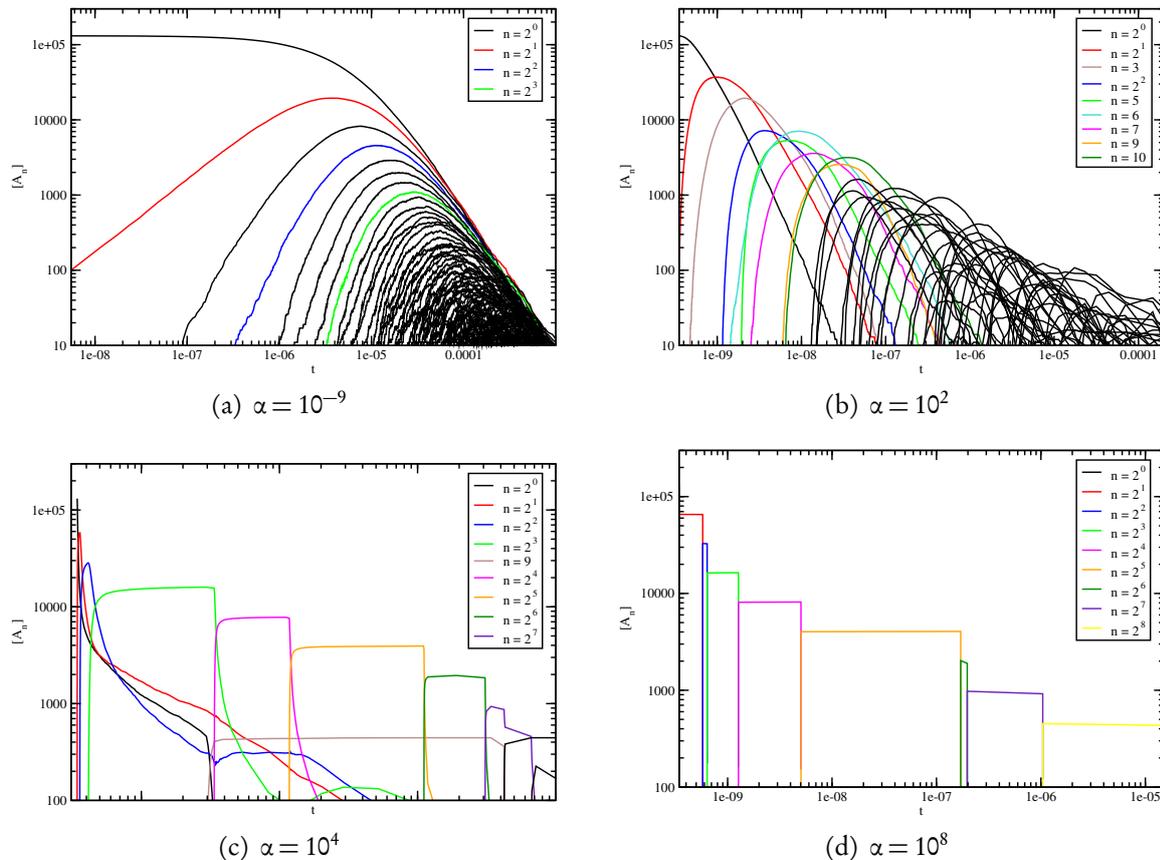


Abbildung 3.16: **Evolution mit parabolischem Wachstum** ( $\zeta = 0,5$ ). Konzentrationsverlauf verschieden langer Polymere. (a) Es stellt sich ein Verlauf, wie im Fall der „wahllosen Wucherung“ ein, bei dem anfangs (hier bis  $n = 8$ ) die Moleküle nach und nach auftreten und den charakteristischen Verlauf zeigen; vgl. Abb. 3.2. Die Ergebnisse sind für  $\alpha \lesssim 10^{-3}$  vergleichbar. Für größere  $\alpha$  tritt ein Übergang zur selektiven Phase ein. (b) Eine Besonderheit des Übergangs zur selektiven Phase ist das Auftreten einer chaotischen Selektion. In Abhängigkeit der konkreten Realisierung entstehen verschiedene Polymere nicht. Im Beispiel sind dies z. B. die Polymere (bis zu einer Länge von 30) 8, 13, 15, 16, 22, 23, 24, 25, 27 und 30. (c) Ab  $\alpha \gtrsim 10^3$  gibt es eine Längenselektion wie in Abb. 3.3. Nahe an dieser Grenze entstehen auch andere Polymere in Nebenreaktionen, wie hier die Länge  $n = 9$ . Die Plateaulängen sind im überkritischen Fall nicht vollständig deterministisch. (d) Für  $\alpha \gtrsim 10^5$  treten statistisch verteilte Plateaus auf; vgl. Abschnitt 3.4.5. Die Anfangskonzentration ( $[A_1]_0$ ) beträgt  $2^{17}$  Monomere und  $\beta = 1$ . Die Konzentration von Molekülen, die aus einer Längenverdopplung hervorgehen, und einige spezielle Moleküle sind farbig eingezeichnet.

### 3 Theorie

Die Plateaform nahe der Phasengrenze ist vergleichbar mit der des exponentiellen Wachstums, zeigt aber einen langsameren Anstieg. Das Plateau kann dort ebenfalls von seiner waagerechten Form abweichen und stärker ansteigen. Der Beginn des Abfalls erfolgt schneller als der Anstieg; vgl. Abb. 3.17.c.

Im Übergangsbereich, bis sich die Längenverdopplungskaskade vollständig ausgebildet hat, sind verschiedene Reaktionen außerhalb dieser Kaskade möglich. Es entstehen daher auch andere Längen, die miteinander und den längenverdoppelten koexistieren. Dies tritt insbesondere im Bereich zwischen Abb. 3.16.b und 3.16.c auf. In ersterem Fall koexistieren sehr viele Spezies in zeitlicher Abfolge miteinander bis auf wenige fehlende und werden zufällig selektiert. In letzterem Fall koexistiert z. B. das Polymer  $A_9$  mit den Längen  $2^4$  bis  $2^7$ .

#### Evolution mit exponentiellem Wachstum

Exponentielles Wachstum wurde bereits eingehend in den Kapiteln 3.2 bis 3.4 untersucht. Wir gehen daher hier nur auf die besonderen Unterschiede im Vergleich mit den anderen Wachstumstypen ein.

Der markanteste Unterschied ist die Entstehung von Schultern; vgl. Abb. 3.17.c. Diese entstehen, wenn die deterministische Plateaulänge zu kurz ist, um fast alle Moleküle in einer Längenverdopplungsreaktion reagieren zu lassen: Es bleibt ein Rest an Molekülen über, der dann z. B. für kleinere  $\alpha$  eine Reaktion der Form  $A_{2^n} + A_{2^{n+1}} \rightarrow A_{2^{2n+1}}$  vollzieht. Diese Reaktion baut die Moleküle ab und es bilden sich keine Schultern aus; vgl. Abb. 3.17.b. Für geringfügig größere  $\alpha$  findet diese Reaktion nicht mehr statt und die restlichen Moleküle reagieren unabhängig von den anderen in einer parallelen, verzögerten Verdopplungsreaktionskaskade. Die Maxima der aus dieser Reaktion resultierenden Moleküle zeigen sich als Schultern im Abfall, der auf die Plateauphase folgt; vgl. Abb. 3.17.c. Die Plateaulänge in diesem Bereich von  $\alpha$  ist nicht mehr rein deterministisch; vgl. auch Abb. 3.13.

#### Evolution mit hyperbolischem Wachstum

Die Selektion für hyperbolisches Wachstum tritt schon bei sehr kleinen Werten von  $\alpha$  auf und zeigt die charakteristische hyperbolische Wachstumsform. Andere Polymere treten in geringerer Konzentration ebenfalls auf; vgl. Abb. 3.18.b. Größere Werte von  $\alpha$  selektieren fast ausschließlich längenverdoppelte Polymere, deren Plateaulänge deterministisch ist; vgl. Abb. 3.18.c. In Abb. 3.18.d sind dann statistisch verteilte Plateaulängen zu verzeichnen und die charakteristische Wachstumskurve ist auf Grund des schnellen Wachstums nicht mehr zu erkennen.

#### Zusammenfassung

Bei allen betrachteten Wachstumsarten können wir einen Übergang finden, der eine Transition von der „wahllosen Wucherung“ hin zu einer Längenverdopplung beschreibt. Dieser ist, abhängig von der Geschwindigkeit der Wachstumsart, bei anderen Werten des autokatalytischen Parameters anzutreffen. Schneller wachsende Wachstumsarten können besser selektieren und weisen eine geringere Koexistenz auf.

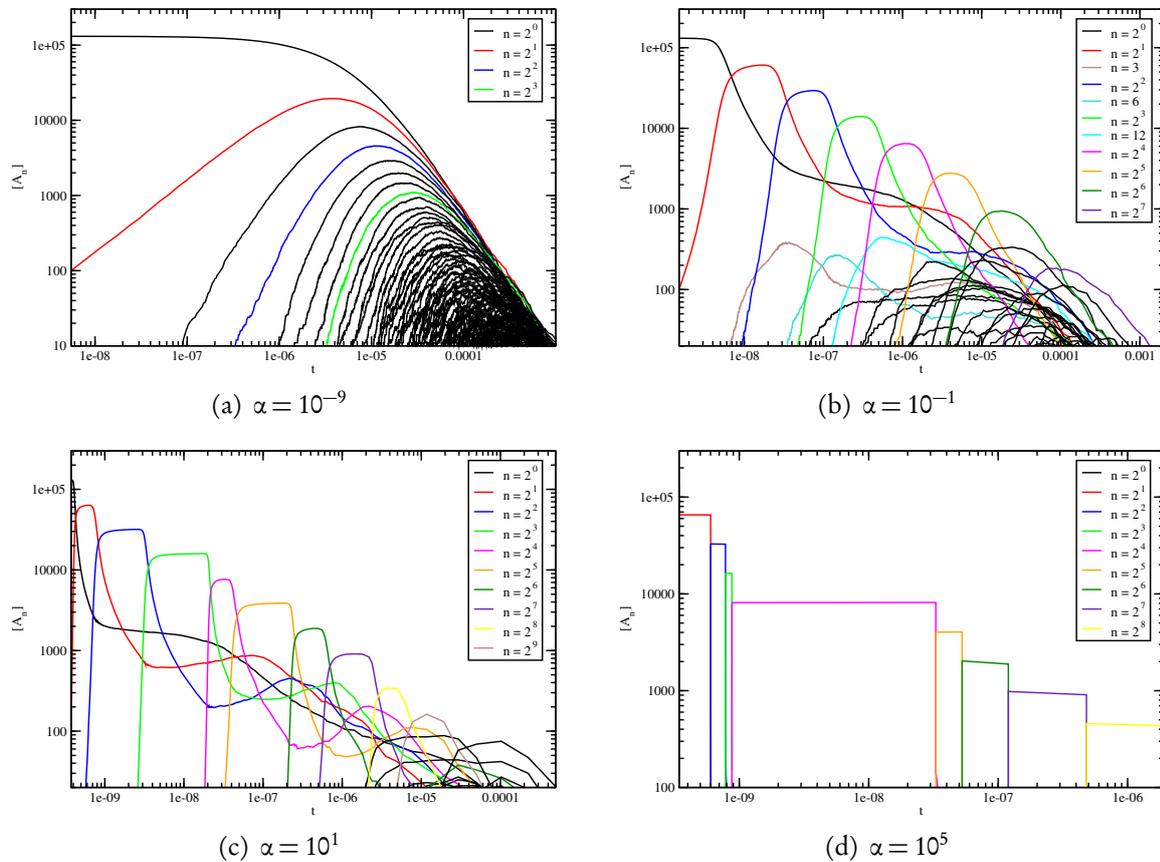


Abbildung 3.17: **Evolution mit exponentiellem Wachstum** ( $\zeta = 1,0$ ). Konzentrationsverlauf der entstehenden Polymere. (a) Für  $\alpha \lesssim 10^{-5}$  ist die Verteilung ähnlich dem Fall der „wahllosen Wucherung“. Anfangs (hier bis  $n = 15$ ) treten die Moleküle nach und nach auf und zeigen den charakteristischen Verlauf; vgl. Abb. 3.2. (b) Bis  $\alpha \lesssim 10^0$  entstehen Plateaus mit deterministischer Länge (vgl. Abschnitt 3.4.6), sowie Polymere aus Reaktionen der Form  $A_{2^n} + A_{2^{n+1}} \rightarrow A_{2^{2n+1}}$ , welche hier ebenfalls farbig markiert sind. (c) Für  $\alpha \lesssim 10^2$  sind die Plateaulängen abgerundet und können unterschiedliche Längen aufweisen. Da nicht alle Moleküle vollständig reagieren, entstehen an den Stellen Schultern, an denen sich keine Polymere der Form  $A_{2^{2n+1}}$  bilden. (d) Für größere  $\alpha$  treten statistisch verteilte Plateaus auf, vgl. Abschnitt 3.4.5. Moleküle, die nicht aus einer Längenverdopplung resultieren, sind rar. Die Anfangskonzentration ( $[A_1]_0$ ) beträgt  $2^{17}$  Monomere und  $\beta = 1$ . Die Konzentration von Molekülen, die aus einer Längenverdopplung hervorgehen, und einige spezielle Moleküle sind farbig eingezeichnet.

### 3 Theorie

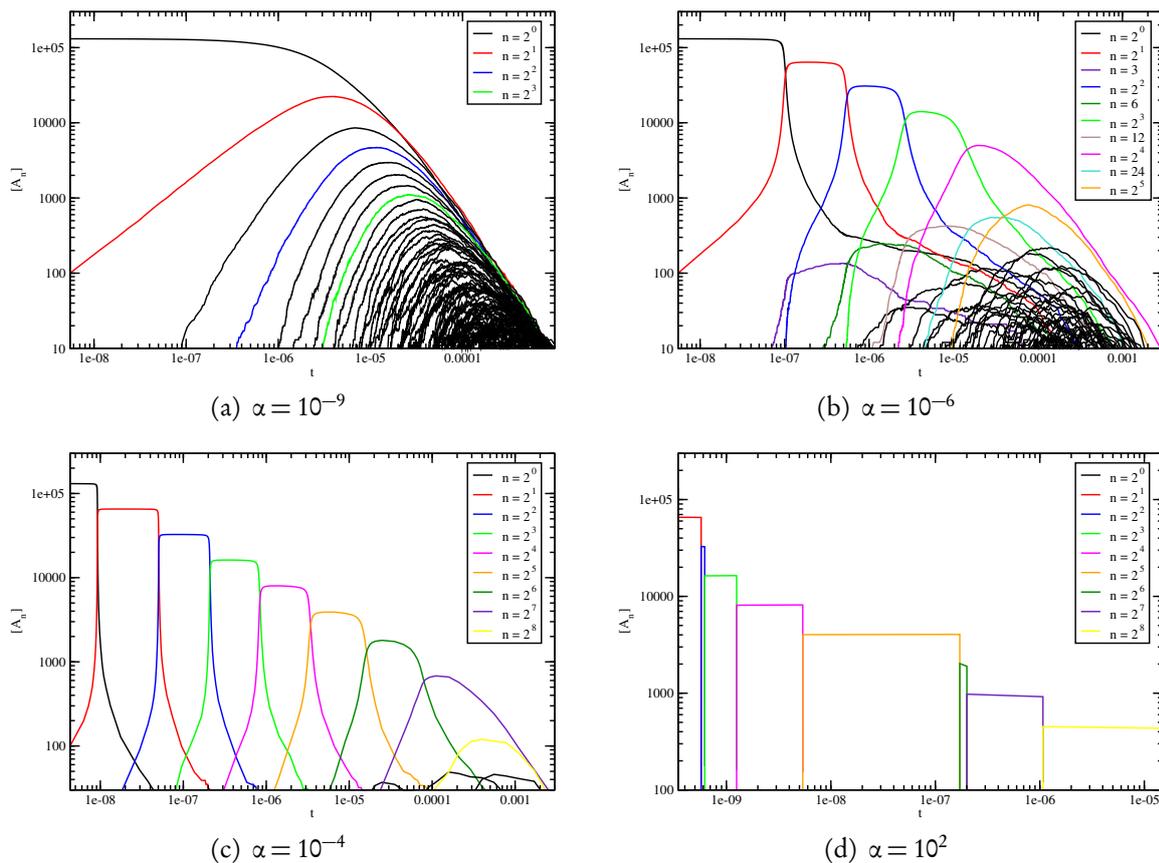


Abbildung 3.18: **Evolution mit hyperbolischem Wachstum** ( $\zeta = 2,0$ ). Konzentrationsverlauf der auftretenden Polymere. (a) Für  $\alpha \lesssim 10^{-9}$  hat das hyperbolische Wachstum keinen Einfluss und das Ergebnis gleicht dem Fall „wahllose Wucherung“. Am Beginn (hier bis  $n = 16$ ) entstehen die Polymere sukzessive und haben einen Verlauf wie in Abb. 3.2. (b) Nach und nach entstehen die für hyperbolisches Wachstum charakteristischen Wachstumskurven auch für längere Polymere, so dass ab einem  $\alpha \gtrsim 10^{-6}$  eine Längenselektion stattfindet und Plateaus deterministischer Länge entstehen; vgl. Abschnitt 3.4.6 und Abb. 1.13.b. Auch wenn die Moleküle der Form  $A_{2^{2n+1}}$  für größer werdende  $\alpha$  verschwinden, zeigt der Abfall keine Schulter, sondern allenfalls ein Plateau. (c) Für  $\alpha \gtrsim 10^{-3}$  verschwinden die Polymere, die nicht aus einer Längenverdopplung resultieren, fast vollständig, weil das Wachstum so stark ist, dass das Häufigste überlebt; vgl. Abschnitt 1.4.3. Die Plateaus haben größtenteils eine deterministische Länge und der charakteristische, hyperbolische Anstieg ist zu erkennen. (d) Falls  $\alpha \gtrsim 10^1$  ist, treten statistisch verteilte Plateaus auf (vgl. Abschnitt 3.4.5), und das Wachstum ist viel schneller als das Warten auf die Erzeugung eines neuen Polymers. Es treten praktisch keine Moleküle auf, die nicht aus einer Längenverdopplung stammen. Die Anfangskonzentration ( $[A_1]_0$ ) beträgt  $2^{17}$  Monomere und  $\beta = 1$ . Die Konzentration von Molekülen, die aus einer Längenverdopplung hervorgehen, und einige spezielle Moleküle sind farbig eingezeichnet.

### 3.5.3 Vergleich der Maxima bei gleichzeitiger und getrennter Änderung von $\alpha$ und $[A_1]_0$

Wir betrachten in diesem Abschnitt den Konzentrationsverlauf des jeweiligen am größten konzentrierten Polymers  $[\overline{A}]$ . Dadurch werden die Plateaus der (möglichen) Längenverdopplung zusammen in einem Graph dargestellt. Im Fall der „wahllosen Wucherung“ ist  $[\overline{A}]$  anfangs identisch mit der Monomerkonzentration  $[A_1]$  und später auch mit längeren Molekülen. Damit wird eine Unterscheidung der auftretenden Realisierungsmöglichkeiten von verschiedenen Parametern in einer Abbildung möglich. Wählt man die Anfangskonzentration  $[A_1]_0$  als freien Parameter, so überlappen sich die Graphen auf Grund der Selbstähnlichkeit, so dass wir davon absehen, diese darzustellen.

In Abb. 3.19.a wird für exponentielles Wachstum lediglich der autokatalytische Parameter  $\alpha$  variiert und in Abb. 3.19.b zusätzlich die Anfangskonzentration. Es ist in beiden Fällen die schrittweise Ausbildung von lokalen Maxima (d. h. von längenverdoppelten Polymeren) zu beobachten. Die Plateaus verlieren ihre rundliche Form und deterministische Länge und werden eckiger bei zunehmend statistisch verteilten Längen. Sowohl große  $\alpha$  (Abb. 3.19.a) als auch große Anfangskonzentrationen  $[A_1]_0$  (Abb. 3.19.b) führen zu einem schnelleren Reaktionsablauf. Dies ist auch für die beiden anderen Wachstumstypen der Fall.

Jedoch setzt für parabolisches Wachstum die Plateaubildung erst für größere Werte von  $\alpha$  und  $[A_1]_0$  ein; vgl. Abb. 3.19.c. Der Übergangsbereich verhält sich ebenfalls anders, da dort keine geordneten Plateaus, sondern unregelmäßige Maxima bzw. Plateaus auftreten; vgl. Abb. 3.16.b. Erst größere Parameter führen zu einem Graph, der aus einer strikten Längenverdopplungskaskade resultiert.

Im Gegensatz dazu ist bei hyperbolischem Wachstum ein direkter Übergang zu ausgezeichneten Maxima bzw. Plateaus schon für kleine Parameterwerte von  $\alpha$  zu verzeichnen. Die Plateauform nimmt schnell eine eckige Gestalt an; vgl. Abb. 3.19.d.

#### Zusammenfassung

In diesem Abschnitt konnten wir den Phasenübergang näher qualitativ untersuchen und haben für exponentielles und hyperbolisches Wachstum einen direkten Übergang zur Plateaubildung erkennen können. Bei parabolischem Wachstum erfolgt dieser Übergang über eine ungeordnete Zwischenstufe, bei der ebenfalls Maxima und auch Plateaus auftreten.

Wir gehen nun dazu über, das System so zu modifizieren, dass prinzipiell eine unendliche Evolution möglich ist. Zum einen führen wir einen Zerfall ein und zum anderen eine Aufkonzentrierung.

### 3.5.4 Zerfall

Um die Stabilität des Längenverdopplungsprozesses im Hinblick auf spontane Konzentrationsänderungen zu untersuchen, wurde Gl. 3.2 um einen Zerfallsprozess erweitert. Bei DNA tritt bei den in Kap. 2 benutzten Temperaturen ein solcher Effekt nicht auf. Wir haben einen Zerfall in Monomere ( $A_n \rightarrow nA$ ) untersucht, um einer weiteren Längenverdopplungskaskade die Möglichkeit zu geben, sich ausbilden zu können. Dieses Ergebnis konnten wir jedoch in

### 3 Theorie

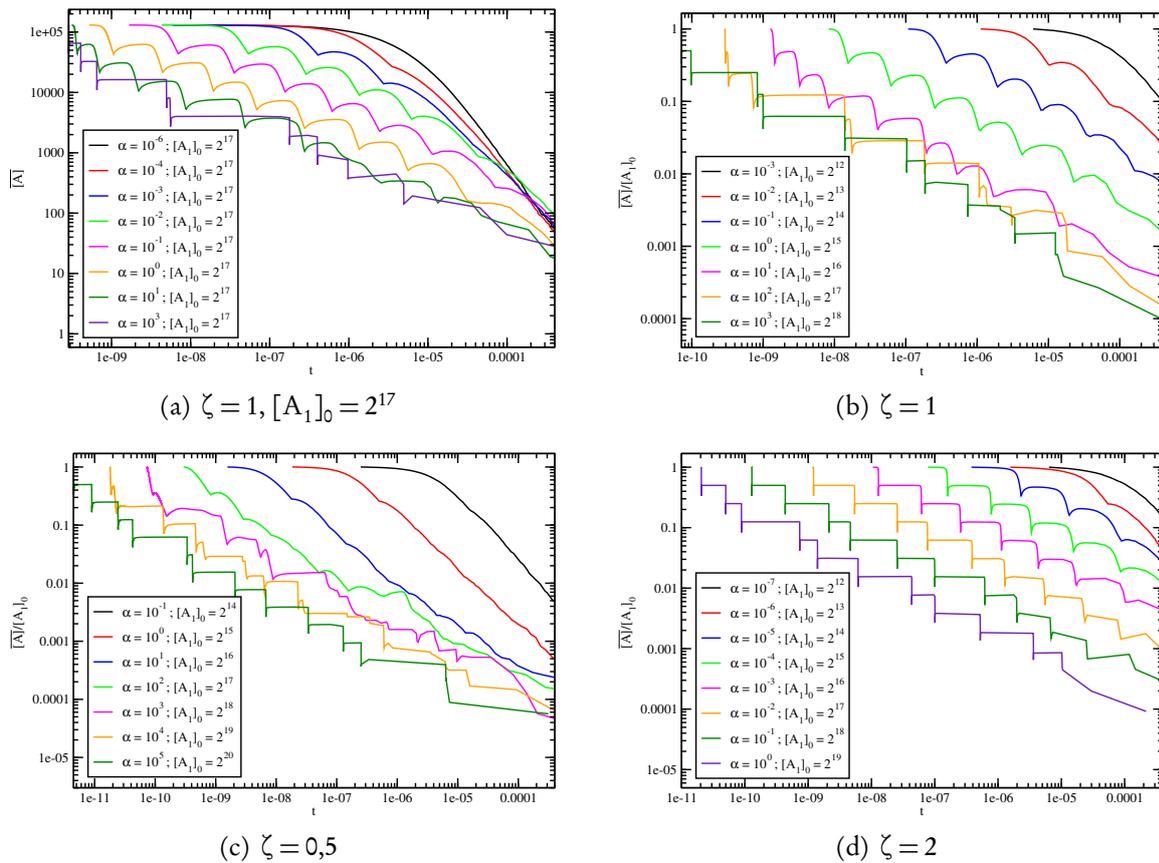


Abbildung 3.19: **Zeitverlauf der Konzentration des am größten konzentrierten Polymers**

$A_n$ , welches wir mit  $\overline{[A]}$  bezeichnen. Bild (a) zeigt die Änderung der Kurvenform für exponentielles Wachstum bei fester Anfangskonzentration. Falls  $\alpha$  fest gewählt wird und die Anfangskonzentration variiert, sind die Kurvenformen vergleichbar. Bilder (b) – (d) zeigen das Überqueren der Plateaugrenze bei gleichzeitiger Änderung von  $\alpha$  und  $[A_1]_0$  und damit alle vorkommenden maximalen Konzentrationsverläufe  $\overline{[A]}$ . Die verschiedenen Konzentrationen sind normiert. Für sehr kleine bzw. große autokatalytische Parameter überlappen sich die Kurven, weil lediglich die spontane Reaktion einen großen Beitrag hat. (a) Es ist ein fließender Übergang von sich ausbildenden, lokalen Maxima (der entsprechenden längenverdoppelten Polymere) gleicher Länge zu erkennen, der in eine statistische Länge übergeht. In (b) ist der Übergang vergleichbar zu (a). (c) Bei parabolischem Wachstum werden zunächst keine geordneten (d. h. aus einer Längenverdopplung resultierenden) Maxima ausgebildet, so dass eine Regelmäßigkeit nicht erkennbar ist; vgl. Abb. 3.16.b. Wenn die Plateaus statistisch verteilt werden, tritt auch hier eine Längenverdopplung ein. (d) Es bilden sich schon für kleine Werte von  $\alpha$  lokale Maxima aus, die den Plateauphasen der längenverdoppelten Moleküle entsprechen. Die Maxima weisen für größere  $\alpha$  zunehmend eine längere Plateauphase, einen steileren Anstieg und statistisch verteilte Längen auf.

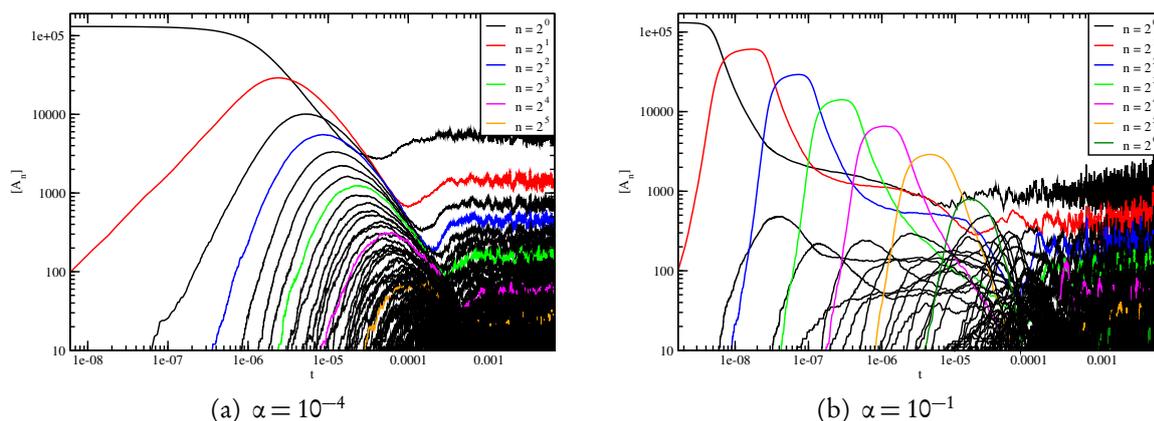


Abbildung 3.20: **Zerfall in Monomere bei exponentiellem Wachstum.** Simulationsergebnis von Gl. 3.12. Die Polymere zersetzen sich zusätzlich mit einer konstanten Rate  $\delta$  in Monomere. Diese Zersetzung beeinflusst das Ergebnis kaum; vgl. Abb. 3.2 und (a) sowie Abb. 3.3 und (b). Wenn die Konzentration aller Moleküle um ca. 2 Größenordnungen abgenommen hat, gewinnt der Zerfall an Gewicht, so dass die Monomerkonzentration und alle anderen Konzentrationen ( $[A_n]$ ,  $n \in \mathbb{N}$ ) ansteigen und um einen Wert fluktuieren. Diese Fluktuationen sind für größere  $\alpha$  größer. Die Parameter sind:  $[A_1]_0 = 2^{17}$ ,  $\delta = 1.000$ ,  $\zeta = 1$ . Abb. 3.20.b nach [82, supplementary material].

der Simulation nicht beobachten. Gl. 3.2 lässt sich dann schreiben als:

$$\begin{aligned}
 \frac{d}{dt}[A_{n+m}] &= (\alpha[A_{n+m}] + \beta)[A_n][A_m] \\
 &\quad - \sum_k (1 + \delta_{n+m,k}) (\alpha[A_{n+m+k}] + \beta)[A_{n+m}][A_k] \\
 &\quad - \delta + \delta_{n+m,1} \sum_k k \cdot \delta,
 \end{aligned} \tag{3.12}$$

wobei  $\delta$  eine konstante Zerfallsrate ist. Für kleine Werte von  $\delta$  hat der Zerfall keine erkennbaren Auswirkungen auf den Konzentrationsverlauf. Erst wenn die Konzentration aller Moleküle stark abgenommen hat, werden die Konzentrationen vom Zerfall beeinflusst und es entstehen größere Fluktuationen, je größer  $\alpha$  ist; vgl. Abb. 3.20. Die Verteilung ist dann über größere Zeiträume stationär.

Wird  $\delta$  um mehrere Größenordnungen größer gewählt, ist der Einfluss dominant. Von derart großen Zerfallsraten ist bei sinnvollen experimentellen Umsetzungen von Gl. 3.12 nicht auszugehen.

Wir erwarten daher, dass spontane Prozesse in einem sinnvollen Ausmaß den Reaktionsverlauf nicht nachhaltig beeinflussen.

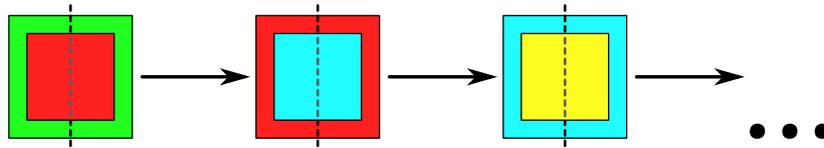


Abbildung 3.21: **Schematische Darstellung der Aufkonzentrationschritte.** Ein Reaktionsvolumen  $V_E$  (grün), in dem sich die Edukte befinden, wird auf das Reaktionsvolumen der Produkte  $V_P$  (rot) halbiert, in dem jede Reaktion, die in  $V_{E_1}$  abläuft, ebenfalls in  $V_{E_2}$  stattfindet. Die Produkte beider Reaktionen entstehen gleichzeitig im Reaktionsvolumen  $V_P$ . Wenn diese Reaktion abgelaufen ist, werden die Produkte zu Edukten und das Schema wiederholt sich mit  $V_E$  (rot) und  $V_P$  (cyan) sowie in einem weiteren Schritt mit  $V_E$  (cyan) und  $V_P$  (gelb) solange wie die Reaktionsverdopplung stattfindet.

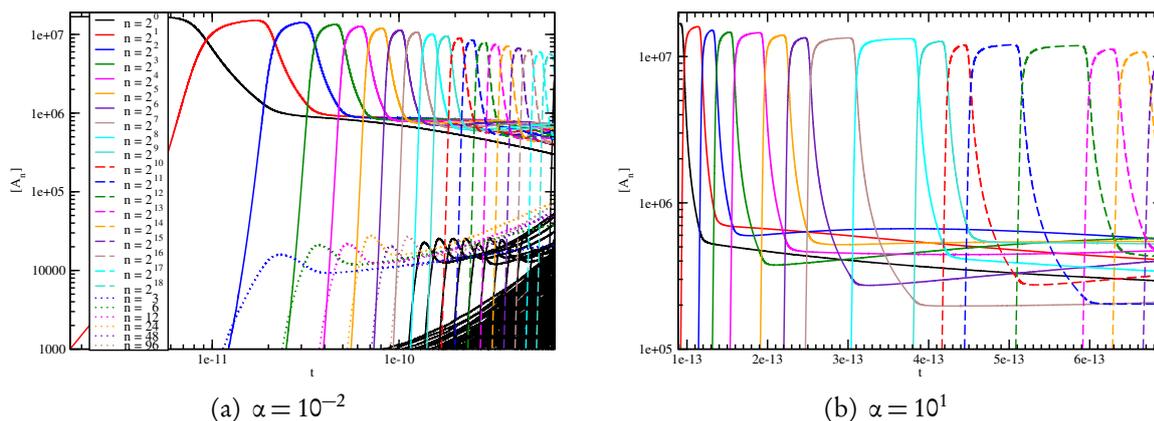


Abbildung 3.22: **Unendliche Evolution in aufkonzentrierenden Reaktionsvolumina.** Die Volumina der Produkte werden in jedem Längenverdopplungsschritt halbiert und aufkonzentriert. Die jeweils dargestellten Konzentrationen sind die des Einheitsreaktionsvolumens. Damit ergibt sich die effektive Verdopplungsreaktion zu  $A_n + A_n \rightarrow 2A_{2n}$  und die Rate zu  $\frac{1}{4}[\alpha[A_{2n}] + \beta]$ . Falls sich eine Längenverdopplungskaskade ausbildet, bleiben die Maximalkonzentrationen der  $A_{2^k}$  vergleichbar. (a) Für kleinere  $\alpha$  entstehen zusätzlich Polymere der Form  $A_{2^k+2^{k+1}}$  sowie andere, deren Konzentration für lange Zeiten monoton ansteigt. (b) Größere autokatalytische Parameter unterdrücken diese zusätzlichen Polymere fast vollständig und die einzelnen Längenverdopplungsschritte sind identisch; die Zeit verläuft daher linear. Anfangskonzentration  $[A_1]_0 = 2^{24}$  bei exponentiellem Wachstum ( $\zeta = 1$ ). Die Legenden von (a) und (b) sind identisch.

### 3.5.5 Die unendliche Evolution

Die in dieser Arbeit vorgestellten Ergebnisse des Modellsystems wurden mit einem endlichen Vorrat an Edukten (Monomeren) durchgeführt. Obwohl dies eine sinnvolle Ausgangssituation darstellt, da auch in natürlichen Systemen die Ressourcen begrenzt sind, führt dies unweigerlich dazu, dass letztlich bei immer größer werdenden Spezies deren Individuenzahl stark abnimmt, so dass die Rohstoffe zur Bildung neuer Individuen und damit größerer Spezies praktisch versiegen. – Die Längenverdopplungskaskade wird unterbrochen und es entsteht schließlich die größtmögliche Spezies  $A_n$ ,  $n = [A_1]_0$ . Ein solches Verhalten wird in der Natur nicht beobachtet, da die Individuen sterben und innerhalb des Ökosystems in ihre biologisch sinnvoll verwertbaren Einzelteile durch die Destruenten\* zerlegt werden. Dies entspricht weitestgehend dem Ansatz aus Abschnitt 3.5.4, in dem wir jedoch zeigen konnten, dass eine unaufhörliche Evolution nicht realisierbar ist. In biologischen Systemen sind Zerfall und Aufbau oft räumlich getrennt, was in dem hier vorgestellten Modellsystem mit einem gut durchmischten Reaktionsvolumen nicht der Fall ist.

Ausgehend von Abschnitt 1.3.5 konnten Braun *et al.* [58, 59, 60] jedoch einen anderen, für molekulare Systeme sinnvolleren Ansatz vorstellen: Die größenabhängige Aufkonzentrierung der vorhandenen Individuen. Angelehnt an diesen Gedanken führen wir in unser Modellsystem eine ähnliche Vorgehensweise ein. Anschaulich kann man sich eine thermische Falle, wie jene von Braun *et al.*, vorstellen. Wir setzen die Modellierung der Aufkonzentrierung für eine Längenverdopplungskaskade wie folgt um; vgl. auch Abb. 3.21: Die Edukte der Reaktion befinden sich in einem doppelt so großen Volumen ( $V_E$ ) wie die Produkte ( $V_P$ ),  $V_P \subset V_E$ , d. h. die Produkte können nur in ihrem Volumen ( $V_P$ ) entstehen und verbleiben dort, während die Edukte in ihrem Volumen ( $V_E$ ) stets gut durchmischt sind. Da die Reaktionen nur in  $V_P = \frac{1}{2}V_E$  stattfinden, ist die Gesamtrate  $\frac{1}{4}[\alpha[A_{n+m}] + \beta]$ , da je Edukte aus  $V_{E_1}$  und  $V_{E_2}$  benötigt werden. Weil die Längenverdopplungskaskade praktisch deterministisch abläuft, wird nur eine von zwei als identisch angenommenen Hälften von  $V_E$  simuliert: Findet eine Reaktion in einer der Hälften statt, so wird sie auch in der anderen genauso durchgeführt. Pro Reaktion entstehen daher 2 Produkte:  $A_n + A_m \rightarrow 2A_{n+m}$ .

Dieses Reaktionsdesign ermöglicht es uns, die Reaktion de facto in einem konstanten Reaktionsvolumen zu simulieren, so dass sich Gl. 3.2 ergibt zu:

$$\begin{aligned} \frac{d}{dt}[A_{n+m}] = & 2 \cdot \frac{1}{4} (\alpha[A_{n+m}] + \beta) [A_n][A_m] \\ & - \frac{1}{4} \sum_k (1 + \delta_{n+m,k}) (\alpha[A_{n+m+k}] + \beta) [A_{n+m}][A_k] \end{aligned} \quad (3.13)$$

Das kleinste Reaktionsvolumen, also dasjenige der letzten aus einer Längenverdopplung resultierenden Polymere, ist das für die Reaktion betrachtete Einheitsvolumen. Mit dieser Vorgehensweise lassen sich die Konzentrationen der anderen Polymere ( $A_u$ ,  $u \neq 2^k$ ) nicht genau berechnen, so dass diese als Abschätzung im Bereich  $\frac{1}{2}[A_u] \leq [\tilde{A}_u] \leq 2[A_u]$  gesehen werden können, wobei  $[\tilde{A}_u]$  die tatsächliche Konzentration ist. Diese fehlerbehaftete Konzentration fließt in die weitere Berechnung ein. Für genügend große  $\alpha$  treten solche Polymere jedoch mit einer nicht signifikanten Konzentration auf, so dass diese insgesamt vernachlässigbar sind.

\* **Destruenten** sind (Mikro-)Organismen, die abgestorbenes, organisches Material zersetzen und damit in organisch verwertbare Stoffe überführen [1].

Wie aus Abb. 3.22 ersichtlich ist, ist die Aufkonzentrierung ein geeignetes Mittel, die Konzentration der einzelnen nacheinander entstehenden Polymere auf einem nahezu konstanten Konzentrationsniveau zu halten. Da die jeweiligen Maximalkonzentrationen praktisch nicht abnehmen, sind auch die Reaktionszeiten der einzelnen Polymerlängen vergleichbar, so dass eine lineare Zeitskala gewählt werden kann; vgl. Abb. 3.22.b. Für kleinere  $\alpha$  entstehen zusätzlich, vergleichbar der bereits besprochenen Verdopplungskaskade in Abb. 3.3, Polymere der Länge  $A_{2^k+2^{k+1}}$  zusammen mit dem Polymer  $A_{2^{k+2}}$  und weitere in geringerer Konzentration; vgl. Abb. 3.22.a.

### Zusammenfassung

Durch schrittweises Aufkonzentrieren der Produkte in einem halbierten Reaktionsvolumen konnten wir die Maximalkonzentration der längenverdoppelten Polymere nahezu konstant halten. Damit bewegt sich die Evolution in einem linearen Zeitrahmen und kann prinzipiell beliebig lange andauern.

Wir gehen nun dazu über, das System auf eine andere Art zu modifizieren, um eine stabile Koexistenz zwischen den Polymeren einer Evolutionsstufe zu analysieren.

### 3.5.6 Koexistenz

Im Folgenden gehen wir der Frage nach, wie robust die Längenselektion im Hinblick auf eine nicht exakte Autokatalyse für  $\zeta = 1$  (exponentielles Wachstum) ist. Dazu modifizieren wir das Modell aus Gl. 3.2 derart, dass auch andere Längen von einem Template katalysiert werden können und verlassen demnach die strikt autokatalytische Reaktion. Die Reaktion  $A_n + A_m \rightarrow A_{n+m}$  kann jetzt auch von längeren Molekülen  $A_l$ ,  $l \geq n + m$  katalysiert werden\*, so dass sich Gl. 3.2 schreiben lässt als:

$$\begin{aligned} \frac{d}{dt}[A_{n+m}] = & \left( \alpha \sum_{l \geq n+m} [A_l] + \beta \right) [A_n][A_m] \\ & - \sum_k (1 + \delta_{n+m,k}) \left( \alpha \sum_{l \geq n+m} [A_{l+k}] + \beta \right) [A_{n+m}][A_k]. \end{aligned} \quad (3.14)$$

Anschaulich bedeutet das, dass längere Moleküle ebenfalls katalytisch tätig werden können und als Template fungieren. Im zweiten Verdopplungsschritt kann also das Polymer  $A_4$  sowohl  $A_4$ ,  $A_3$  als auch  $A_2$  katalysieren. Dementsprechend verändert sich das Verhalten in der längenselektierten Phase wie folgt: Neben dem längenverdoppelten Polymer  $A_4$  entsteht zusätzlich das Polymer  $A_3$ . Dieses weist zwar eine geringere Konzentration auf als  $A_4$  aber zeigt ebenfalls ein Plateau. Die Konzentration ist, obwohl  $A_4$  katalytisch wirkt, geringer, da ein Teil der Edukte ( $A_1$ ) stark an Konzentration abnimmt; vgl. Abschnitt 3.2.2. Demnach sind die längenverdoppelten Polymere immer größer konzentriert als die restlichen.

Es entstehen für jeden Längenverdopplungsschritt  $k$  Polymere der Länge  $n$ ,  $2^{k-1} < n < 2^k$ ; vgl. Abb. 3.23. Dabei verdoppelt sich in jedem Schritt auch das Polymer  $A_{2^{k-2}+2^{k-1}}$ , welches sich gleichzeitig zu  $A_{2^k}$  bildet und aus den beiden vorhergegangenen längenverdoppelten

\* Es ist auch der gegenteilige Fall denkbar, in dem kürzere Templates längere Produkte katalysieren können, also  $l \leq n + m$ . Dann ist jedoch keine Längenselektion mehr möglich (ähnlich der „wahllosen Wucherung“). Lediglich für  $n + m - 1 \leq l \leq n + m$  können noch Polymere selektiert werden.

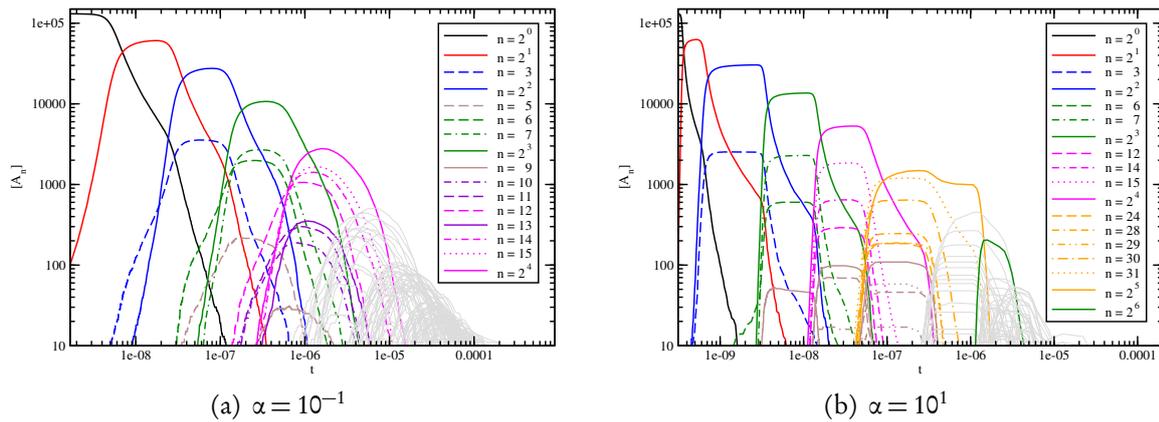


Abbildung 3.23: **Koevolution nach Gl. 3.14.** Es kommt im überkritischen Fall zu einer Längenselektion. Die längenverdoppelten Polymere haben jeweils die größte Konzentration. Es treten weitere, kürzere Polymere auf, die mit den längenverdoppelten koevolvieren. Polymere der Länge  $u(k) = 2^{k-1} + 2^{k-2}$  sind gestrichelt in der Farbe von  $A_v$ ,  $v(k) = 2^k$  eingezeichnet. Die sich aus  $u(k)$  und  $v(k)$ ,  $k = 2, 3, \dots$  entwickelnden entsprechend gestrichpunktet. Der Längenunterschied zwischen  $u(k)$  und  $v(k)$  wird mit jeder Längenverdopplung größer, so dass weitere Polymere zwischen  $u(k)$  und  $v(k)$  entstehen (in entsprechender Farbe gekennzeichnet). Dieses Schema lässt sich für kleine (a) und große (b)  $\alpha$  erkennen. In (b) kommt es für  $k = 5$  zu einem weiteren Evolutionsschritt für alle kürzeren Polymere als  $A_{2^5}$ . Die Anfangskonzentration beträgt  $2^{17}$  bei exponentiellem Wachstum. Die grauen Polymere sind nicht in der Legende aufgeführt, da sie nur eine niedrige Konzentration erreichen oder dicht beieinander liegen. Abb. 3.23.a in leicht modifizierter Form in [82] veröffentlicht.

### 3 Theorie

Polymeren besteht. Ebenfalls bildet sich gleichzeitig das Polymer  $A_{2^{k-1}}$ , welches sich aus  $A_{2^{k-1}}$  und  $A_{2^{k-1}-1}$  zusammensetzt. Diese beiden Moleküle ( $A_{2^{k-2}+2^{k-1}}$ ,  $A_{2^{k-1}}$ ) reagieren vornehmlich im darauffolgenden Schritt mit  $A_{2^k}$  oder auch untereinander; vgl. Abb. 3.23.

#### Koexistenz

Mit den im vorherigen Abschnitt entstehenden Molekülen stellt sich eine Koexistenz von diesen selektierten Molekülen und den längenverdoppelten ein. Diese sind zur katalytischen Vermehrung von den längenverdoppelten abhängig und ernähren sich größtenteils auch von diesen. Damit koevolvieren diese Moleküle zusammen. Speziell ist auch ein weiterer Evolutionsschritt ohne die längenverdoppelten Moleküle möglich, wie z. B. in Abb. 3.23.b. Dort ist das Plateau der anderen Moleküle kürzer und neue Moleküle entstehen, während  $A_{2^5}$  maximal ist.

Diese Situation unterscheidet sich grundlegend von der bisher beobachteten Koevolution, die eher spontan in einer Realisierung auftritt; vgl. Abb. 3.5. Hier ist der Prozess gekoppelt an die längenverdoppelten Moleküle und daher zu einem gewissen Grad (es können auch andere Moleküle auftreten) deterministisch.

#### Zusammenfassung

In diesem Abschnitt konnten wir durch eine Modifikation des katalytischen Terms die katalytische Wirkung eines (längenverdoppelten) Moleküls erweitern, so dass andere Moleküle in Abhängigkeit und zeitgleich von den längenverdoppelten entstehen. Diese anderen Moleküle können wiederum selbst miteinander reagieren und sich unabhängig von den längenverdoppelten Molekülen entwickeln.

## 3.6 Perkolation

Wir haben in dieser Arbeit bislang hauptsächlich die biologisch relevanten Grundlagen und Aspekte des Systems behandelt sowie die daraus resultierende physikalische Perspektive. Jedoch ist das System ebenfalls in einem anderen, nicht-biologischen Kontext von Relevanz, nämlich der Netzwerkforschung. Diese beschäftigt sich u. a. mit dem Verhalten von Graphen, die aus Knoten (*vertices*) und Kanten (*edges*) bestehen. Ein Knoten ist die kleinste Einheit in einem Graph und eine Kante eine Verbindung zwischen zwei Knoten. Mit Kanten verbundene Knoten werden Cluster genannt. Damit stellt das Polymer  $A_n$  ein Cluster der Größe  $n$  dar, welches aus  $n$  Knoten und  $n - 1$  Kanten besteht. Um die Notation und Ausdrucksform konsistent zu halten, werden wir bevorzugt die bisherige Nomenklatur beibehalten, obwohl die Resultate möglicherweise auch allgemeingültiger sein können.

Ein allgegenwärtiger Prozess, der mit Hilfe der Netzwerkforschung untersucht werden kann, ist das Fragmentieren [132]. Dieser Prozess beschreibt das In-Stücke-Brechen homogener Teile, wobei in vielen Fällen die daraus entstehenden Stücke eine vergleichbare Größe haben [132]. Der zeitinverse Prozess wird Perkolation genannt und beschreibt das Zusammenlagern ähnlich großer Einheiten; vgl. Abb. 3.24; der allgemeine Prozess der Zusammenlagerung heißt im Gegensatz dazu Aggregation [132]. Beide Prozesse können Phasenübergänge zeigen.

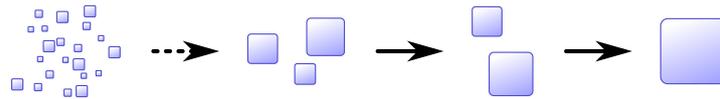


Abbildung 3.24: **Perkolation** ist der rückwärts ablaufende Prozess des Zerfalls. In jedem Schritt werden zwei Einheiten zu einer größeren, neuen Einheit verbunden bis letztlich eine einzelne große Einheit entsteht. Dabei werden im Gegensatz zur normalen Aggregation systematisch nur Einheiten vergleichbarer Größe vereinigt. Frei übernommen aus [132].

Schröder *et al.* [132] geben mehrere Beispielsysteme, die, wenn sie sich langsam ändern, knistern. Dazu gehören das Knittern von Papier oder das Verbrennen von Holz. Letzteres kommt durch verdampfende Wasserreste oder Harzeinschlüsse zu Stande. Es sind also Prozesse, die Einheiten verbinden oder vernichten, die ein solches Verhalten hervorrufen. Ein Beispiel für Perkolation ist der Barkhauseneffekt, bei dem die Weiss-Bezirke im Zuge der Magnetisierung eines ferromagnetischen Materials ummagnetisiert werden und sich damit zusammenlagern [132]. Bei der Entmagnetisierung werden die entsprechenden weisschen Bezirke wieder fragmentiert, um eine verschwindende Gesamtmagnetisierung zu erhalten. Ein anderes Beispiel ist eine Epidemie, bei der aus einzelnen mit einer Krankheit infizierten Personen, die gut mit anderen Personen vernetzt sind, immer größere, infizierte Personengruppen entstehen, so dass letztlich ein Großteil infiziert ist [133]. In diesem Fall wäre es beispielsweise wünschenswert, den Beginn der Perkolation zu verzögern [133]. Perkolation lässt sich also in einem Netzwerk [132] als auch auf einem Gitter beobachten [134].

Ein ähnliches Verhalten zeigt auch das in dieser Arbeit vorgestellte System im Fall der Längenverdopplung. Systeminhärent werden aus gleich (ähnlich) großen Polymeren solche mit doppelter Länge. Dies deutet auf einen perkolutiven Phasenübergang<sup>\*</sup> hin. Daher werden wir im Folgenden das Verhalten des Systems in Abhängigkeit vom autokatalytischen Parameter untersuchen. Eine Veröffentlichung darüber ist in Vorbereitung [122]. Zunächst geben wir aber einen kurzen Überblick über die verwendeten Methoden und häufig studierten Systeme.

### 3.6.1 Methoden und Beobachtungsgrößen

Es gibt mehrere Fragestellungen, die im Kontext der Perkolation beantwortet werden können. Diese sind im Einzelnen die Eigenschaft der (Nicht-)Selbstmittelung (*(non-)self-averaging*) [132], der Verlauf des Ordnungsparameters [133] und die (Nicht-)Existenz der diskreten Skaleninvarianz (*discrete-scale-invariance*) [135]. Anhand dieser Eigenschaften können die Systeme klassifiziert werden.

Als Ordnungsparameter wählt man für gewöhnlich die Größe des größten Clusters. Hier charakterisiert dieser die Länge (aber nicht die Häufigkeit des Vorkommens) des längsten Polymers. Wir bezeichnen den Ordnungsparameter mit  $\theta$ . Die Struktur des Problems ist

\* Der hier betrachtete **perkulative Phasenübergang** ist im Hinblick auf eine andere Fragestellung, nämlich der Herausbildung einer einzelnen großen Komponente, definiert. Damit unterscheidet er sich fundamental von dem in Kapitel 3.4.2 vorgestellten, der auf die Entstehung sich immerwährend reproduzierender Einheiten abzielt. Die Phasenübergänge sind demnach nicht direkt vergleichbar.

beispielsweise im zeitlichen Verlauf des Ordnungsparameters zu erkennen. D'Souza *et al.* [133] geben eine Übersicht: Es sind kontinuierliche Phasenübergänge möglich, die jedoch Sprünge im Ordnungsparameter für eine endliche Realisierung zeigen können. Zwei weitere Klassen bilden die Phasenübergänge, die einen einzelnen Sprung entweder lange vor oder am Ende des Prozesses aufweisen. Es gibt aber auch nicht konvergente, nicht selbstmittelnde Prozesse mit echten unstetigen Sprüngen. Cho *et al.* [136] konnten ein System finden, das durch Änderung zweier Parameter alle diese Eigenschaften zeigt und zudem noch eine ultralangsame Konvergenz aufweist. Solche echten unstetigen Sprünge zeigt beispielsweise das Bohman-Frieze-Wormald-Modell [137, 138, 139]. In einem gewissen Parameterbereich tritt lediglich ein einzelner Sprung auf, während in einem anderen sich mehrere nacheinander bilden [138].

Für selbstmittelnde Systeme verschwindet die relative Varianz des Ordnungsparameters [132, 140, 136]:

$$R_v = \frac{\langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2}{\langle \mathcal{O} \rangle^2} \xrightarrow{N \rightarrow \infty} a, \quad \begin{cases} a = 0, & \text{selbstmittelnd} \\ a > 0, & \text{nicht selbstmittelnd,} \end{cases} \quad (3.15)$$

wobei  $\langle \cdot \rangle$  eine Ensemblemittelung bedeutet und  $N$  die Systemgröße darstellt. Nicht-selbstmittelnde Systeme weisen im Grenzwert für unendlich große Systeme jedoch einen verteilten Ordnungsparameter auf, so dass die relative Varianz nicht verschwindet [132, 135, 140].

Die relative Varianz kann Oszillationen vor [141] und nach [135] dem Phasenübergang aufweisen, die auf einen (vergangenen) Phasenübergang hinweisen können [135]. Diese Oszillationen können bei einer strikten Längenverdopplung, wie sie generisch in vielen perkolativen Systemen vorkommt, aus einer strikten Längenverdopplung resultieren [141]. Diesem Prozess kann eine diskrete Skaleninvarianz zu Grunde liegen, die sich aus einer diskreten Translationsinvarianz ergeben kann, welche wiederum aus der Diskretheit des zu Grunde liegenden Gitters oder Netzwerks resultiert [141]. Daraus ist es möglich, den Perkolationsphasenübergang aus den ensemblemittelten, mikroskopischen Zustandsänderungen des Ordnungsparameters vorherzusagen [141].

Diskrete Skaleninvarianz ist eine besondere Form der Skaleninvarianz. Letztere geht durch die Skalierung des Ordnungsparameters um  $\mu(\lambda)$  hervor, wenn man den willkürlichen Parameter  $x$  mit  $\lambda$  skaliert [140]:

$$\mathcal{O}(x) = \mu \mathcal{O}(\lambda x). \quad (3.16)$$

Die Lösung ist ein Potenzgesetz [140]:

$$\mathcal{O}(x) \sim x^{\nu}, \quad \text{mit} \quad \nu = -\frac{\ln \mu}{\ln \lambda}. \quad (3.17)$$

Die Skaleninvarianz heißt diskret, wenn die obige Relation nur für bestimmte, abzählbare  $\lambda_1, \lambda_2, \lambda_3, \dots$  gilt, die als  $\lambda_n = \lambda^n$  geschrieben werden können [140].  $\lambda$  wird fundamentales Skalierungsverhältnis genannt [140]. Damit ist es möglich, die diskrete Skaleninvarianz in einem logarithmisch aufgetragenen Graph an Hand der Oszillationen in der relativen Varianz und, doppelt logarithmisch aufgetragen, an den linearen Einhüllenden der einzelnen Realisierungen des größten Polymers zu erkennen [135].

Schröder *et al.* [135] stellen fest, dass eine diskrete Skaleninvarianz möglicherweise bei der Ensemblemittelung verloren gehen kann, und schlagen als Abhilfe eine Reskalierung der Zeit nach Gl. 3.16 vor, die vor der Ensemblemittelung angewendet wird, so dass die einzelnen

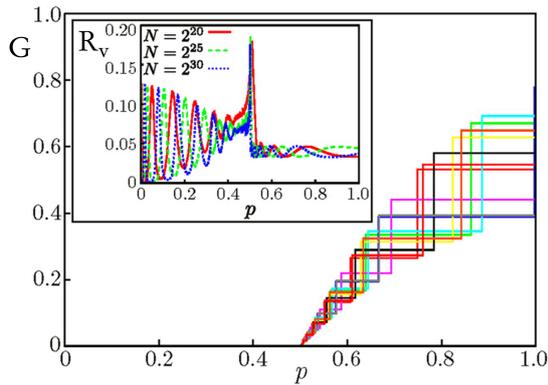
Realisierungen wieder gegeneinander ausgerichtet sind. Diese Reskalierung ändert bei einem selbstmittelnden System die Ensembleeigenschaften nicht, weil dann die beiden Reskalierungsfaktoren für alle Realisierungen (fast) identisch sind. Dazu verwenden Schröder *et al.* die relative Größe des größten Clusters und die Verknüpfungsdichte und normieren diese auf den jeweiligen Phasenübergangspunkt. Wir werden die relative Länge des längsten Polymers benutzen, die dann schon für verschiedene Systemgrößen im Endzustand den Wert 1 aufweist als Abhängigkeit von der Anzahl der Reaktionen. Die Anzahl der Reaktionen hat, beispielsweise im streng längenverdoppelnden Fall bei rein statistisch verteilten Plateaus, die Eigenschaft, die verschiedenen Realisierungen (fast) identisch werden zu lassen, da essentiell die gleichen Reaktionen ablaufen, die aber in der Realzeit wegen der statistischen Verteilung der jeweiligen Plateauanfänge nicht koinzidieren. Das (ideale) System ist also eigentlich selbstmittelnd, da die verschiedenen Realisierungen (fast) identisch sind, aber in Realzeit nicht-selbstmittelnd, da die identischen Realisierungen zu (sehr) verschiedenen Zeiten stattfinden.

In der Realzeitdarstellung sind die einzelnen Realisierungen von der Zeit abhängig und daher beeinflusst diese Zeitabhängigkeit das Ensemblemittel über die verschiedenen Realisierungen. Auch wenn prinzipiell die gleiche Abfolge der Zustände (Polymere) für jedes System generiert wird, können diese zu unterschiedlichen Zeiten auftreten. Wir wollen daher hier von einem zeitlichen Ensemblemittel sprechen. Im Gegensatz dazu wird bei einer Ensemblemittelung über die verschiedenen Realisierungen in Abhängigkeit der Reaktionszahl die Zeitfluktuation entfernt und das Ensemblemittel ist zeitunabhängig und hat den Charakter eines statistischen Mittels von Konfigurationen entsprechenden statistischen Gewichts. Wir nennen dieses Mittel daher hier ein statistisches Ensemblemittel.

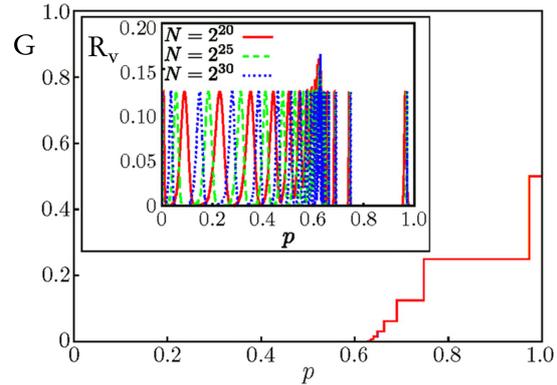
Diese Darstellung erinnert an physikalische Systeme, von denen sowohl ein Zeitmittel als auch ein Ensemblemittel gebildet werden kann. Treten alle Konfigurationen im Phasenraum im Vergleich der beiden Mittel gleich häufig auf, so spricht man von einem ergodischen System; ansonsten von einem nicht-ergodischen [125]. Ein Beispiel für ein ergodisches System [142] ist Sinais [143] zwei-Scheiben-Problem [125]. Es werden zwei Scheiben in eine Kiste mit periodischen Randbedingungen betrachtet, die aneinander stoßen können, wobei diese so groß sind, dass sie nicht aneinander vorbei können. Im stationären Bezugssystem einer Scheibe beschreibt der für die zweite Scheibe zugängliche Raum eine entsprechend „eingedellte“, ausgefüllte Raute [125]. Diese wird sowohl durch eine Simulation von zwei Scheiben beschrieben als auch durch das wiederholte Platzieren beider Scheiben bei korrektem statistischen Gewicht. Ein nicht-ergodisches System ist ein Sonnensystem mit wenigen, kleinen Planeten mit geringer Masse und einer schweren Sonne [125]. Vernachlässigt man die Planeten-Planeten-Interaktion und nimmt eine unendlich schwere Sonne an, so beschreiben diese Keplerbahnen [125]. Aber auch die Planeten-Planeten-Interaktion verändert diese Bahnen nur geringfügig [125, 144] und es entstehen damit stabile Unterräume im Phasenraum [144], in denen sich die Planeten bewegen. Weil diese Unterräume begrenzt sind, liegt ein nicht-ergodisches System vor.

Wir gehen nun dazu über, für das hier vorgestellte System wichtige Grenzfälle und oft behandelte Systeme zu diskutieren.

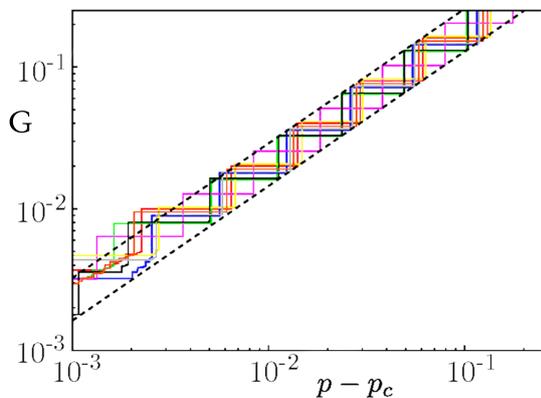
### 3 Theorie



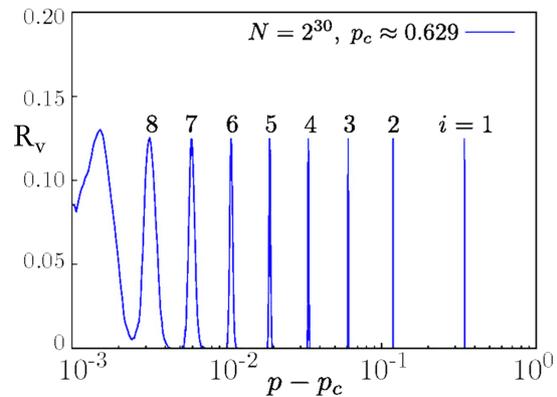
(a) modified Erdős-Rényi – größtes Cluster und relative Varianz



(b) globalhomophile Perkolations – größtes Cluster und relative Varianz



(c) modified Erdős-Rényi – reskaliertes größtes Cluster



(d) globalhomophile Perkolations – reskalierte relative Varianz

Abbildung 3.25: **Globalhomophile Perkolations und modified Erdős-Rényi.** Das modified Erdős-Rényi Modell zeigt Realisierungen, die auch nach Reskalierung nicht koinzidieren (vgl. Abb. (a,c)), und die relative Varianz ist endlich; vgl. Abb. (a). Damit ist das modified Erdős-Rényi Modell nicht selbstmittelnd. (a,c) Das größte Cluster wächst sprunghaft und die einzelnen Realisierungen können durch eine Einhüllende beschrieben werden. (a) Die relative Varianz zeigt vor und nach dem Phasenübergang Oszillationen, die für verschiedene Systemgrößen nicht koinzidieren. Der Phasenübergang ist für die betrachteten Systemgrößen gleich lokalisiert. (b) Die globalhomophile Perkolations weist deterministische Realisierungen auf: Die Größe des größten Clusters wird strikt verdoppelt. (b,d) Die Oszillationen der verschiedenen Systemgrößen in der relativen Varianz sind für verschiedene Systemgrößen vor dem Phasenübergang unterschiedlich. Überkritisch koinzidieren sie jedoch und haben eine logarithmische Periodizität. Diese wird für geringe Clusterzahlen leicht durchbrochen. Entnommen aus [145].

### 3.6.2 Beispielsysteme

Bei allen Beispielsystemen beginnen wir mit einem leeren Graph, d. h. es sind nur nicht verbundene Knoten vorhanden. Es wird dann zufällig eine gewisse Anzahl an Knoten gewählt, die dann nach den angegebenen Regeln verbunden werden. Es sei angemerkt, dass dieses Vorgehen sich grundlegend von dem in dieser Arbeit betrachteten System unterscheidet. Hier hat jedes Polymer das gleiche Gewicht, aber dort hat ein Cluster der Größe  $n$  auch  $n$  Knoten und wird damit auch  $n$ -mal so häufig ausgewählt.

Ein Beispiel für ein nicht selbstmittelndes Modellsystem ist das modified Erdős-Rényi Modell [133]. Es werden drei Knoten zufällig ausgewählt und die beiden größten Komponenten  $G_1$  und  $G_2$ ,  $G_1 \geq G_2$  des gesamten Netzwerks betrachtet. Falls  $G_1 = G_2$ , werden zufällig zwei der gewählten Knoten  $e_i$  und  $e_j$  verbunden. Falls  $G_1 > G_2$  und mindestens zwei der gewählten Knoten in Clustern sind, welche gleichgroß zu  $G_1$  sind, werden diese zwei Knoten verbunden, ansonsten zwei Knoten aus Clustern, die kleiner als  $G_1$  sind [146]. Der zeitliche Verlauf der Größe der größten Komponente und die relative Varianz ist in Abb. 3.25.a,c dargestellt. Wir fassen kurz die Erkenntnisse aus [135] zusammen: Der Verlauf der größten Komponente ist für verschiedene Realisationen unterschiedlich, was auf ein nicht selbstmittelndes System hinweist. Die relative Varianz ist positiv und daher ist das System nicht selbstmittelnd. Es sind jedoch Oszillationen zu erkennen, so dass bestimmte Clustergrößen bevorzugt werden. Eine Reskalierung bringt den linearen Zusammenhang im log-log-Plot der diskreten Skaleninvarianz zum Vorschein; vgl. Abb. 3.25.c.

Die globalhomophile Perkolation wurde von Schröder *et al.* [135] vorgestellt und wir stellen kurz die Analysen dar. Es wird eine Verbindung aus allen möglichen Verbindungen, die Cluster der gleichen Größe (und möglicherweise auch dasselbe Cluster) verbinden, gleichverteilt zufällig ausgewählt. Dies führt zu einer strikten Größenverdopplung und, da das zweitgrößte Cluster zunächst wachsen muss, um mit dem größten verbunden werden zu können, zu einer makroskopischen Verzögerung dieses Schritts. Dieses Modell ist im Gegensatz zum modified Erdős-Rényi Modell selbstmittelnd, obwohl es eine Treppe im Verlauf des größten Clusters zeigt; vgl. Abb. 3.25.b. Weil die Übergangspunkte insbesondere für größere Systeme zu gleichen, unendlich kleinen, deterministischen Zeitpunkten stattfinden, weist die relative Varianz lokalisierte Peaks auf, so dass die relative Varianz fast überall verschwindet; vgl. Abb. 3.25.b,d. Damit ist die globalhomophile Perkolation selbstmittelnd. Des Weiteren kann aus Abb. 3.25.d anhand der in der logarithmischen Auftragung periodischen Peaks auf eine diskrete Skaleninvarianz für überkritische Systeme geschlossen werden.

Cho *et al.* [147] beschreiben ein Modell mit einem Verbindungskern<sup>\*</sup>  $K_{ij} \sim (ij)^\omega$  mit einem variablen Gewicht  $\omega$ , wobei  $i$  und  $j$  die jeweiligen Clustergrößen sind. Für  $\omega = 1$  entspricht das dem Erdős-Rényi Modell und für  $\omega = 0$  einer linearen Polymerisation. Der letzte Fall ist deshalb hier interessant und die Ratengleichung ergibt sich dann zu:

$$\dot{n}_s(t) = \sum_{s=i+j} \frac{1}{c^2} n_i(t)n_j(t) - \frac{2}{c} n_s(t), \quad (3.18)$$

\* Der **Verbindungskern** ist die Zusammenfassung aller Vorfaktoren der beiden Clusterdichten  $n_i(t)$  und  $n_j(t)$  in Gleichung 3.18 [147].

wobei  $n_s(t)$  die Dichte der Cluster mit Größe  $s$  sind und  $c(t)$  die Gesamtdichte. Diese Gleichung ist äquivalent zu Gl. 3.2 mit  $\alpha \equiv 0$  und  $\beta = 1/c^2$ . Damit ist diese Gleichung zeitabhängig, wenn  $\omega \neq 1$ . Cho *et al.* finden, dass  $c(t)$  linear mit der Zeit abfällt und dass die Clustergrößenverteilung exponentiell abfällt, wenn  $s$  groß wird. Dazu betrachten sie die Erzeugendenfunktion von  $n_s(t)$ :  $f(z, t) = \sum_s n_s(t) z^s$  mit der Fugazität\*  $z$  im Bereich  $0 < z < 1$ . Damit und aus Gl. 3.18 erhalten sie eine Differentialgleichung für  $f(z, t)$ . Die Lösung geben sie an mit:  $f(z, t) = \frac{(1-t)^2 z}{1-zt}$  für  $t < 1$  und  $f(z, t) = 0$  für  $t > 1$ . Sie finden, dass für  $t < 1$  eine Reihenentwicklung in  $z$ :  $n_s(t) = (1-t)^2 t^{s-1}$  liefert. Cho *et al.* schließen daraus, dass die Clustergröße exponentiell abnimmt, wenn  $s$  groß wird. Insbesondere wählen sie  $\delta = 1-t$  so, dass sich  $n_s(t) \approx \delta^2 e^{-s/\delta^*}$  ergibt mit  $s^* \approx 1/\delta$ . Sie finden, dass die charakteristische Größe  $s^*$  für  $\delta \rightarrow 0$  divergiert und es zu einem Phasenübergang 1. Ordnung kommt, da die Clustergröße einen plötzlichen Sprung in der Größe des größten Clusters vollzieht. Für endliche Clustergrößen ergibt sich nach Cho *et al.* die Größe des größten Clusters bei einem Phasenübergang bei  $t_c = 1$  zu:

$$G(t) = \begin{cases} 0 & \text{falls } 0 < t < 1 \\ 1 & \text{falls } t > 1. \end{cases} \quad (3.19)$$

Dieses Ergebnis weicht vom Resultat ab, welches Cho *et al.* aus der Smoluchowski Koagulationsgleichung\*\* erhalten:  $t_c = \infty$ .

### 3.6.3 Analyse

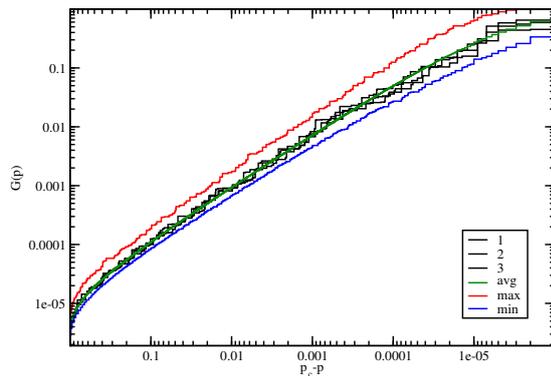
Im Folgenden analysieren wir das in dieser Arbeit vorgestellte System im Hinblick auf (Nicht-)Selbstmittelung, die diskrete Skaleninvarianz, den Verlauf des Ordnungsparameters und der relativen Varianz in Abhängigkeit verschiedener autokatalytischer Parameter  $\alpha$ . Diese wählen wir im Einzelnen zu  $\alpha = 10^{-6}$  („wahllose Wucherung“),  $\alpha = 10^1$  (runde Plateaus variabler Länge) sowie  $\alpha = 10^5$  und  $\alpha = 10^7$  (statistisch verteilte Plateaus), wobei diese Werte exemplarisch herausgegriffen wurden und typische Beispiele des jeweiligen Verhaltens darstellen\*\*\*.

Das Minimum und Maximum des größten Polymers in Abhängigkeit von der Reaktionszahl und der Realzeit liegen für  $\alpha = 10^{-6}$  dicht beieinander und weisen wie der Durchschnitt kaum Sprünge auf; vgl. Abb. 3.26.a,c. Diese resultieren aus der Endlichkeit des Systems. Die relative Varianz in Abhängigkeit von der Reaktionszahl verläuft stetig und nimmt für größere Systeme ab außer dem Fall geringer Teilchenzahlen; vgl. Abb. 3.26.b,d. Dies ist ein starker Hinweis auf eine Selbstmittelung. In Abhängigkeit von der Realzeit kommt es bei der relativen Varianz zu aperiodischen Oszillationen für kleine Zeiten, die vom Entstehen der ersten Polymere ( $A_1, A_2, \dots, A_5$ ) herrühren. Für große Zeiten zeigt die relative Varianz einen Peak, der für größere Systemgrößen schmaler wird. Dazwischen, im Bereich von Abb. 3.17.a, ist ihr Zahlenwert

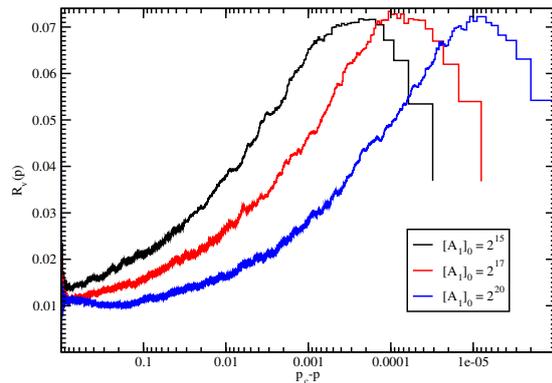
\* Die **Fugazität** ist eine Korrekturgröße für unterkritische, reale Systeme, um die Gleichungen des idealen Systems anwenden zu können [148].

\*\* Die **Smoluchowski Koagulationsgleichung** beschreibt das Koagulieren (Zusammenklumpen) von Teilchen in einer Lösung auf Grund von Diffusion [149].

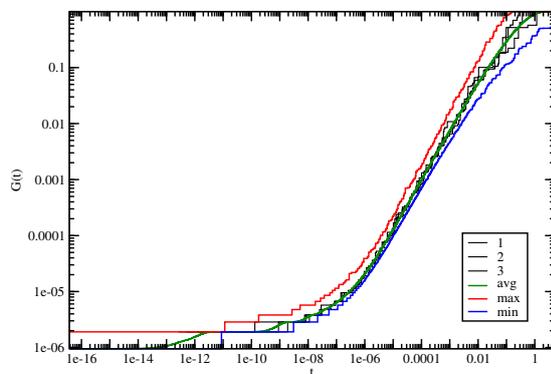
\*\*\* Für  $\alpha \leq 10^{-4}$  gibt es kaum Oszillationen in  $R_v(p)$ . Für  $\alpha \geq 10^1$  werden diese Oszillationen streng periodisch und gehen ab  $\alpha \geq 10^3$  in Peaks über. Bis  $\alpha \leq 10^2$  verläuft der Durchschnitt von  $G(p)$  glatt und wird bis  $\alpha \leq 10^4$  wellig, um dann nahe am Minimum zu liegen. ( $[A_1]_0 = 2^{20}$ ).



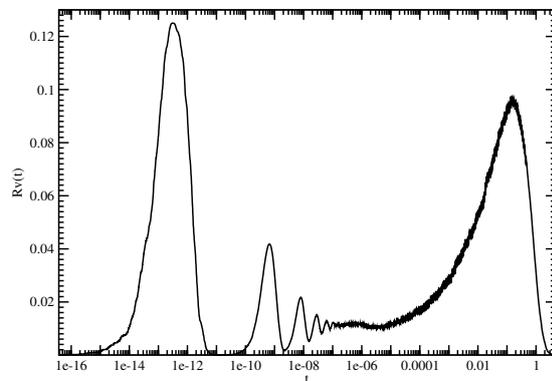
(a) größtes Polymer in Abhängigkeit von der Reaktionszahl



(b) relative Varianz in Abhängigkeit von der Reaktionszahl



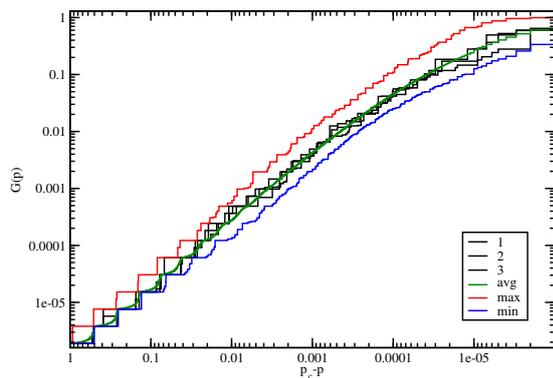
(c) größtes Polymer in Abhängigkeit von der Zeit



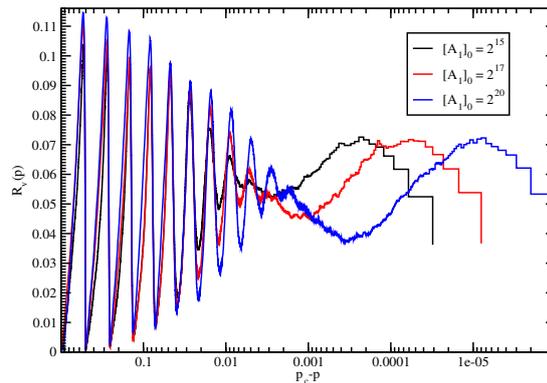
(d) relative Varianz in Abhängigkeit von der Zeit

Abbildung 3.26: **Größtes Polymer und relative Varianz** in Abhängigkeit von der Anzahl der Reaktionen und der Zeit für  $\alpha = 10^{-6}$ . (a) Die Realisierungen des größten Polymers werden von stetigen minimal (min) und maximal (max) möglichen größten Polymeren eingehüllt und der Durchschnitt (avg) der Realisierungen verläuft glatt. Das deutet darauf hin, dass das System selbstmittelnd ist. (b) Die relative Varianz nimmt für größere Systeme kleinere Werte an. Das weist darauf hin, dass das System selbstmittelnd ist. Auf Grund der endlichen Systemgröße sind am Anfang und am Ende größere Werte möglich. (c) Die Minima und Maxima des größten Polymers in Echtzeit sind stetig und liegen dicht beieinander. Dies ist ein starkes Indiz für Selbstmittelung. Schwankungen am Anfang und Ende, insbesondere des Durchschnitts des größten Polymers, sind auf die endliche Größe zurückzuführen. (d) Die relative Varianz in Realzeit weist bis auf die Randbereiche einen niedrigen Wert auf. Die Oszillationen sind auf die endliche Systemgröße zurückzuführen und resultieren aus dem sukzessiven Aufbau der ersten Polymere ( $A_1, A_2, \dots, A_5$ ). (a,c,d) für eine Anfangskonzentration von  $[A_1]_0 = 2^{20}$ . Der Konzentrationsverlauf ist vergleichbar mit Abb. 3.17.a.

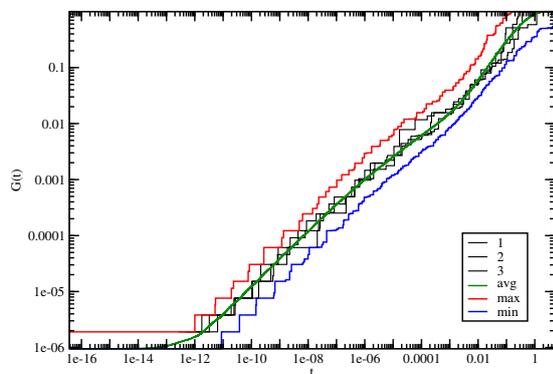
### 3 Theorie



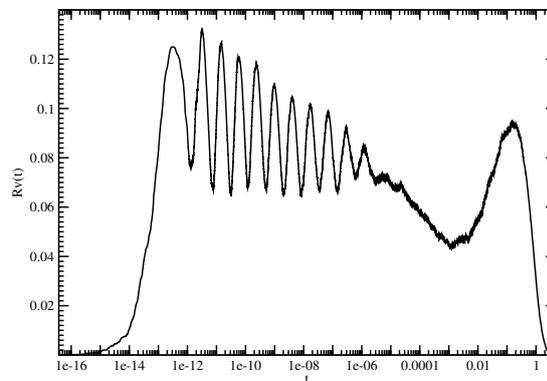
(a) größtes Polymer in Abhängigkeit von der Reaktionszahl



(b) relative Varianz in Abhängigkeit von der Reaktionszahl



(c) größtes Polymer in Abhängigkeit von der Zeit



(d) relative Varianz in Abhängigkeit von der Zeit

Abbildung 3.27: **Größtes Polymer und relative Varianz** in Abhängigkeit von der Anzahl der Reaktionen und der Zeit für  $\alpha = 10^1$ . (a) Die minimale und maximale Größe des größten Polymers weisen am Anfang Sprünge auf und sind klar getrennt, was auf eine Nicht-Selbstmittelung hindeutet. Die Sprünge sind auch im Durchschnitt erkennbar, der ansonsten glatt verläuft. (b) Wenn das größte Polymer Sprünge zeigt, oszilliert die relative Varianz. Die Oszillationen weisen für größere Systeme eine stärkere Amplitude und eine Zunahme der Anzahl der Schwingungen auf. Die größer werdende Amplitude verschwindet im Minimum für große Systeme und weist daher auf eine diskrete Skaleninvarianz und ein nicht-selbstmittelndes System hin. Für große Zeiten sind die Verläufe vergleichbar und durch eine geringe Teilchenzahl bedingt. (c) Verlauf wie in (a). Der Durchschnitt zeichnet die Treppenstruktur in Realzeit jedoch nicht nach. (d) Im Zeitverlauf weist die relative Varianz ebenfalls Oszillationen auf, die jedoch einen konstanten Wert für die Minima zeigen. Für große Zeiten (und damit geringen Teilchenzahlen) tritt ein Maximum am Ende auf. (a,c,d) für eine Anfangskonzentration von  $[A_1]_0 = 2^{20}$ . Der Zeitverlauf der einzelnen Konzentrationen ist in Abb. 3.17.c dargestellt.

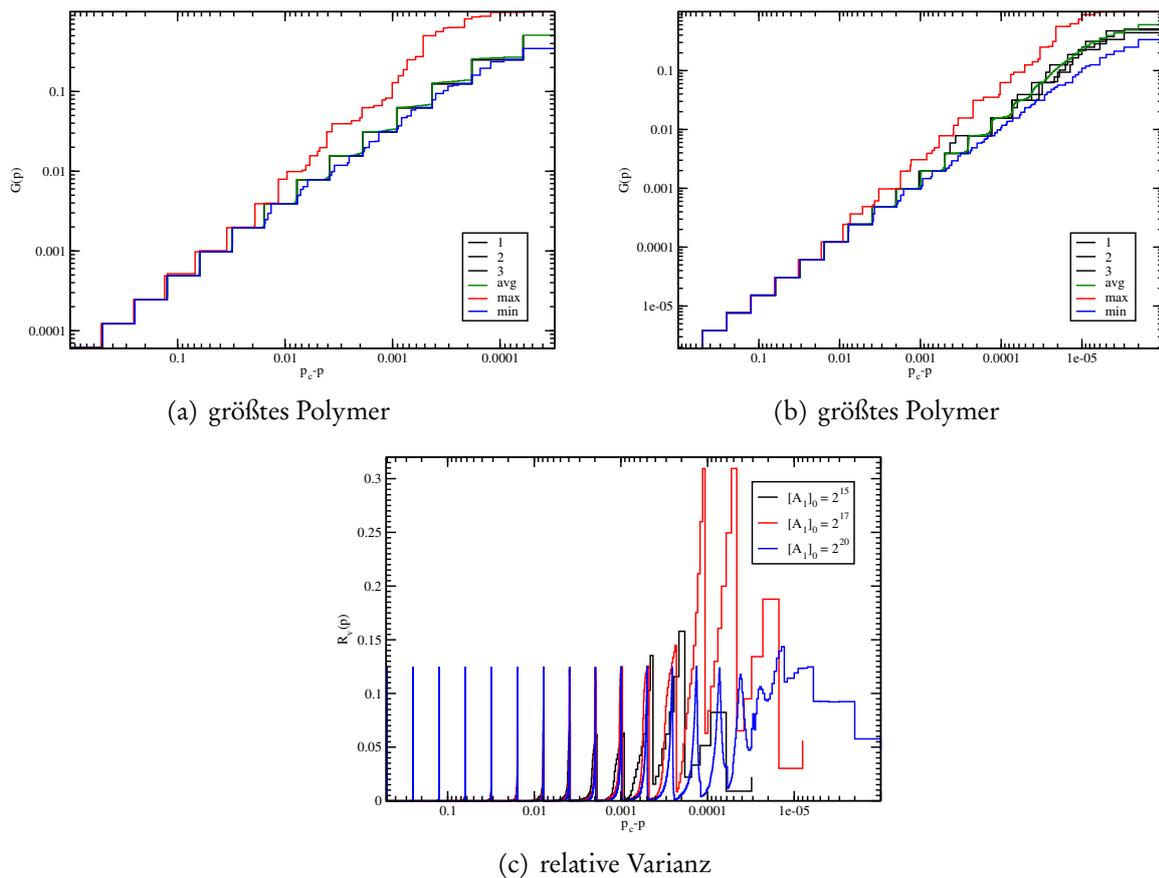
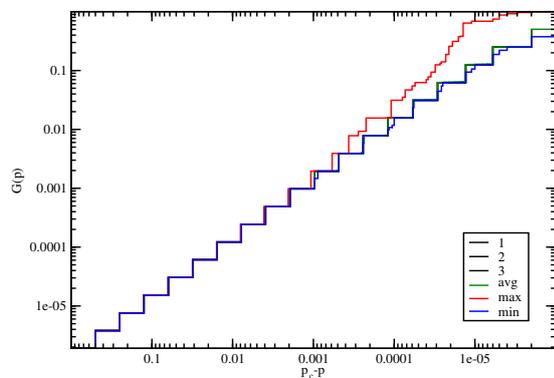
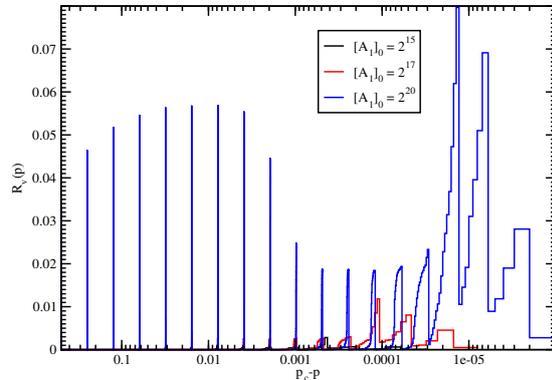


Abbildung 3.28: **Größtes Polymer und relative Varianz** in Abhängigkeit von der Anzahl der Reaktionen für  $\alpha = 10^5$ . (a) Für eine Anfangskonzentration von  $[A_1]_0 = 2^{15}$  fallen Minimum, Maximum und demnach auch der Durchschnitt der Größe des größten Polymers für kleine Reaktionszahlen nahezu aufeinander. Dies ist auf die strikte, ungestörte Längenverdopplung zurückzuführen. Für größere Reaktionszahlen ist der Durchschnitt nahe am Minimum, während das Maximum größere Werte annimmt. Hier treten auch andere Reaktionen auf, die zum einen die Reaktionsanzahl leicht verringern können, wenn vorher Zwischenreaktionen, d. h. Reaktionen, die keine Längenverdopplung zur Folge haben, aufgetreten sind und diese jetzt zu kürzeren als den längenverdoppelten Polymeren reagieren. Das Maximum kann jedoch leicht zu größeren Werten hin verschoben werden; es genügt eine vorzeitige Reaktion der größten Polymere. (b) Größere Anfangskonzentrationen ( $[A_1]_0 = 2^{20}$ ) führen zu einem ausgedehnteren Bereich, in dem Minimum und Maximum zusammenfallen. Es entstehen jedoch für größere Reaktionszahlen eine stärkere Abweichung von Durchschnitt und Minimum. Wegen der größeren Anfangskonzentration sind Reaktionen, die nicht zu einer Längenverdopplung führen, wahrscheinlicher. (c) Für den Bereich der strikten Längenverdopplung verschwindet die relative Varianz bis auf die Verdopplungsschritte fast überall: Ein Zeichen für ein selbstmittelndes, diskret skaleninvariantes System. Größere Reaktionszahlen zeigen eine endliche relative Varianz auf Grund der geringen Teilchenzahlen. Der Konzentrationsverlauf findet sich in Abb. 3.17.d.

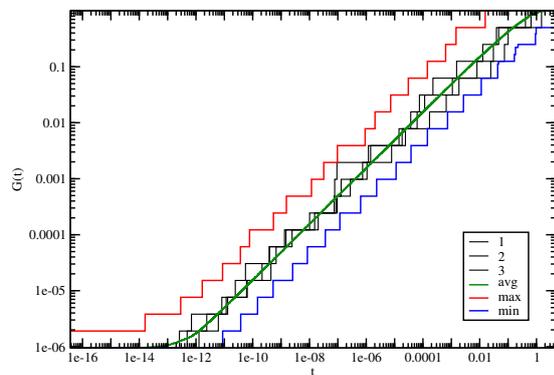
### 3 Theorie



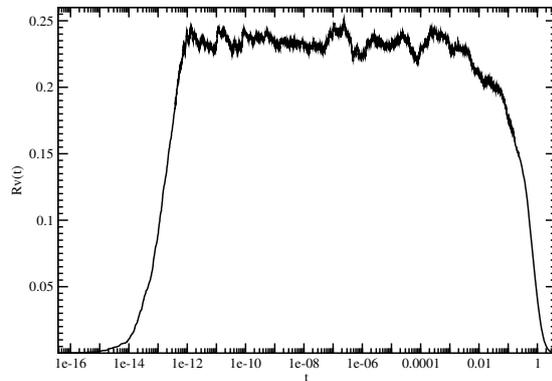
(a) größtes Polymer in Abhängigkeit von der Reaktionszahl



(b) relative Varianz in Abhängigkeit von der Reaktionszahl



(c) größtes Polymer in Abhängigkeit von der Zeit



(d) relative Varianz in Abhängigkeit von der Zeit

Abbildung 3.29: **Größtes Polymer und relative Varianz** in Abhängigkeit von der Anzahl der Reaktionen und der Zeit für  $\alpha = 10^7$ . (a) Der Bereich, in dem Maximum und Minimum aufeinander fallen, wird mit wachsendem  $\alpha$  größer; vgl. Abb. 3.28.b. Für kleine Reaktandenzahlen ist der Durchschnitt nahe beim Minimum und das Maximum erreicht größere Werte. (b) Die relative Varianz zeigt diskrete, periodische Peaks, die mit der Systemgröße zunehmen. Die relative Varianz verschwindet also fast überall, so dass ein selbstmittelndes System vorliegt. Die Periodizität weist auf eine diskrete Skaleninvarianz hin. (c) Minima und Maxima sind in Realzeit getrennt voneinander und der Durchschnitt mittelt sich zu einer glatten Kurve zwischen Minimum und Maximum. (d) Die relative Varianz zeigt in der Zeitabhängigkeit einen endlichen Wert. (a,c,d) für eine Anfangskonzentration von  $[A_1]_0 = 2^{20}$ . Der zeitliche Verlauf der Konzentrationen ist vergleichbar mit Abb. 3.17.d.

gering. Für  $\alpha = 10^{-6}$  zeigt das System ein vergleichbares Verhalten in Abhängigkeit von der Realzeit und der Reaktionszahl und die relative Varianz deutet auf ein selbstmittelndes System hin.

Im Fall der Längenverdopplung mit runden Plateaus unterschiedlicher Länge ( $\alpha = 10^1$ , Abb. 3.17.c) zeigt das größte Polymer für beide Abhängigkeiten anfangs im Minimum und Maximum eine Treppenstruktur, die sich nach dem Aufbau der längenverdoppelten Polymere auflöst und in einen stetigen Verlauf übergeht; vgl. Abb. 3.27.a,c. Der Durchschnitt in Abhängigkeit von der Reaktionszahl lässt durch seine Oszillationen nahe des Minimums den Aufbau der längenverdoppelten Polymere erkennen; vgl. Abb. 3.27.a. Im selben Bereich sind auch in der relativen Varianz periodische Oszillationen zu erkennen, deren Amplitude mit steigender Systemgröße wächst; vgl. Abb. 3.27.b,d. Daran schließt sich ein Minimum mit einem folgenden Peak an. Ersteres nimmt mit der Systemgröße ab und letzteres ist eine Folge der endlichen Systemgröße. Das System ist also anfangs nicht selbstmittelnd und zeigt eine diskrete Skaleninvarianz.

Größere autokatalytische Parameter ( $\alpha = 10^5$ ) führen zu statistisch verteilten Plateaus (Abb. 3.17.d) und das Minimum und das Maximum des größten Polymers fällt in Abhängigkeit von der Reaktionsanzahl im relevanten Bereich praktisch zusammen; vgl. Abb. 3.28.a,b. Die Längenverdopplungsreaktion läuft hier soweit ungestört ab, so dass kaum Abweichungen zu erkennen sind. Danach sind die Molekülzahlen gering und das Maximum nimmt größere Werte an, da es zur spontanen Bildung größerer Polymere kommt. Entsprechend wird auch das Minimum glatter, weil andere kleinste Polymere übrigbleiben. Für niedrige Anfangskonzentrationen ist der Durchschnitt davon kaum beeinflusst, während er für größere im Zeitverlauf zunächst welliger wird und dann glatt zwischen Maximum und Minimum liegt. Ein solcher Übergang ist demnach insbesondere auch für große Systeme zu erwarten und wird für große  $\alpha$  hin zu großen Systemgrößen verschoben; vgl. Abb. 3.29.a. Die relative Varianz weist im relevanten Bereich sättigende Peaks auf, die einen Maximalwert von 0,124 erreichen; vgl. Abb. 3.28.c. Diese Peaks sind periodisch und sonst verschwindet die relative Varianz fast überall. Demnach liegt ein diskret skaleninvariantes, selbstmittelndes System vor. Für größere Reaktionszahlen zeigt die relative Varianz wegen der geringen Polymeranzahlen und der endlichen Autokatalyse ein anderes Verhalten.

Größere autokatalytische Parameter ( $\alpha = 10^7$ ) führen zu vergleichbaren Ergebnissen, wobei der relevante Bereich ausgedehnter ist und die relative Varianz einen geringeren Maximalwert erreicht, der für größere Systemgrößen als noch zunehmend extrapoliert werden kann; vgl. Abb. 3.29.a,b. In Abhängigkeit von der Realzeit zeigen sowohl Minimum als auch Maximum des größten Polymers eine Treppe und liegen getrennt voneinander. Der Durchschnitt liegt glatt und gerade in der Mitte; vgl. Abb. 3.29.c. Entsprechend weist auch die relative Varianz ein entsprechend des geraden Durchschnitts ausgedehntes, fluktuierendes Plateau auf; vgl. Abb. 3.29.d. Dies deutet auf ein nicht selbstmittelndes System hin und ist die Konsequenz der statistischen Plateaulängenverteilung. Da deren Aufbau, aus Reaktionssicht gesehen, deterministisch ist und strikt erfolgt, ist auch die Abhängigkeit von der Reaktionszahl deterministisch und für alle Läufe bis auf Ausreißer identisch. Jedoch ist die Entstehung der nächsten Stufe aus Realzeitsicht rein statistisch und nicht vorhersagbar, so dass dies zu völlig unterschiedlichen Realisierungen führt, die daher nicht mit dem Mittel übereinstimmen.

### Zusammenfassung

Wir konnten das System im Hinblick auf Perkolation untersuchen und haben unterschiedliches Verhalten für Reaktionszahl und Realzeit feststellen können. Es obliegt dem Betrachter, wo sein Interesse liegt: Die Reaktionszahl betrachtet den reinen Verdopplungsprozess und die auftretenden Reaktionen, während die Realzeit auf die tatsächlich stattgefundenen Realisierungen eingeht.

Die relative Varianz unterkritischer Systeme weist darauf hin, dass diese selbstmittelnd sind; vgl. Abb. 3.26.b. Für größere  $\alpha$  folgt ein Bereich, der in der relativen Varianz Oszillationen mit wachsender Amplitude für größere Systeme zeigt. Dies deutet darauf hin, dass das System in diesem Bereich nicht-selbstmittelnd ist und eine diskrete Skaleninvarianz aufweist; vgl. Abb. 3.27.b. Im Bereich statistisch verteilter Plateaus sind in der relativen Varianz Peaks zu sehen, die auf ein selbstmittelndes, diskret skaleninvariantes System hinweisen; vgl. Abb. 3.28.c und Abb. 3.29.b.

Um eine unbeeinflusste Längenverdopplung zu erreichen, ist es notwendig, dass neben einer Vergrößerung der Anfangskonzentration auch der autokatalytische Parameter entsprechend Gl. 3.5 vergrößert wird. Jedoch legt die Extrapolation von Abb. 3.28.b und Abb. 3.29.a nahe, dass es einen Grenzwert für  $\alpha$  gibt, weil die Grenze bis zu der praktisch eine strikte Längenverdopplung stattfindet sich hin zu größeren Reaktionszahlen verschieben lässt. Falls sich diese Grenze bis zum Ende der Reaktion verlegen ließe, gäbe es keine anderen als die längenverdoppelten Polymere. Die Systemgröße könnte dann beliebig anwachsen, ohne dass der Längenverdopplungsmechanismus gestört wäre.

Wir haben ein System gefunden, das an Hand eines Parameters sowohl selbstmittelnd als auch nicht selbstmittelnd bzw. selbstmittelnd und diskret skaleninvariant sein kann. Damit deckt das hier vorgestellte System eine Auswahl der möglichen Systeme in Abhängigkeit eines einzelnen Parameters ab.

# Diskussion, Zusammenfassung und Ausblick

Das evolutionäre Verhalten auf der Erde ist gekennzeichnet durch ein immerwährendes Fortschreiten der Entwicklung. Es bilden sich aus den bestehenden Arten auf Grund neuer Umweltbedingungen und Interaktionen der Spezies neue Arten heraus, die jeweils temporär dominant in ihrer Nische sind [13]. Dieser Prozess ist unvereinbar mit der Annahme, es entstünde eine best-angepasste Spezies [19]. Wir konnten ein Modellsystem entwickeln, welches die Grundlagen molekularer, autokatalytischer Reproduktion einbezieht und temporär dominante Arten (Längen) hervorbringt. Gleichgültig, welche genaue Wachstumsart gewählt wird, z. B. eine sexuelle Reproduktion, ist eine solche Phase stabil in Abhängigkeit von der Stärke des Wachstums. Wenn man dieses Ergebnis auf größere Objekte übertragen möchte, wären die am schnellsten wachsenden Spezies solange stabil, bis aus ihren Ressourcen\* eine neue, kaum mit ihnen wechselwirkende Spezies entstünde. Wir haben hier vorwiegend ein exponentielles, autokatalytisches Wachstum untersucht.

In der unselektiven Phase, „wahllose Wucherung“ genannt, treten, angefangen von Monomeren, alle möglichen Längen an Polymeren auf. Diese sind anfangs in ihrer Reihenfolge des Auftretens auch der Länge nach sortiert und größere Polymere erreichen nicht die Konzentration von kürzeren. In diesem Fall kann keine selektive Evolution stattfinden.

Die selektive Phase ist gekennzeichnet durch nacheinander proliferierende Polymere, die in ihrer Länge jeweils verdoppelt sind. Diese können je nach Größe des autokatalytischen Parameters ein mehr oder weniger ausgeprägtes Plateau in ihrem Maximum aufweisen. Ausgeprägte Plateaus resultieren aus einem schnellen Wachstum, welches, nach einem vollständigen Ablauf der Reaktion, auf die Entstehung des nächsten Polymers wartet. Daher ist die Länge der Plateaus statistisch verteilt. Wenig ausgeprägte Plateaus haben eine deterministische Länge, da das folgende Polymer schon während der Wachstumsphase entsteht.

Das System ist in der Lage, Koexistenz zu zeigen, so dass neben den längenverdoppelten Polymeren auch andere auftreten. Diese entstehen zufällig bzw., wenn auch längere Polymere katalytisch wirken, deterministisch. Damit ist eine wesentliche Eigenschaft der Evolution, die Koexistenz mehrerer, unterschiedlicher Nischen, reproduzierbar.

Mit diesem System wurde die Grundlage für eine Modellsystemklasse zur Analyse evolutionärer Systeme gelegt. Diese beinhalten eine Möglichkeit, spontan neue, nicht vorhandene Individuen zu generieren und vorhandene Individuen autokatalytisch zu reproduzieren. Damit sich eine stabile Population und damit eine (temporär) dominante Spezies einstellt, dürfen die Variations- bzw. Mutationsraten nicht zu groß sein [19].

---

\* Die **Ressourcen** könnten in diesem Fall entweder direkt von der neuen Spezies verwertet werden oder in einem für andere Spezies nicht oder kaum zugänglichen Verwertungsprozess zur Herstellung von Nahrung. Es sei angemerkt, dass der Mensch sich durch die Landwirtschaft eine solche, von anderen Lebewesen wenig beeinflusste Nahrungsproduktion aufgebaut hat, die von deren Ressourcen Gebrauch macht.

Es besteht einerseits die Erweiterbarkeit um einen genetischen Code, der den Polymeren zusätzlich innewohnt. Hiermit ließe sich die Propagation und Stabilität genetischer Information untersuchen und somit weitere, realistischere Aussagen über das Fortschreiten der Evolution treffen. Andererseits könnte die Einführung von abgeschlossenen Einheiten, wie sie als Zellen in höheren Organismen vorkommen oder als Organellen, wie sie in Zellen vorkommen können, eine Entmischung der verschiedenen funktionalen Reaktionspfade ermöglichen. Diese könnten zu getrennt entstehenden Produkten führen, die zusammen interagieren und eine möglicherweise neuartige (funktionelle) Eigenschaft aufweisen. Demgegenüber steht die reine Interpretation des Modells als ein molekularbiologisches. In diesem Zusammenhang wäre eine weitergehende Funktionalität der Spezies durch beispielsweise eine mehrdimensionale Struktur ein zu implementierendes Merkmal. Man denke hier an RNA, die sich falten und in diesem Zustand eine enzymatische Funktion erfüllen kann [19].

Diese Erweiterungen kommen der speziell untersuchten Situation entgegen und können einen weiteren, genaueren Einblick in die Funktionsweise und Bedingungen der Evolution gewähren.

# Anhang

## Nomenklatur

A Molekül

$\bar{A}$  Komplement des Moleküls A.

A- $\bar{A}$  Hybridisierung des Moleküls A auf dem Molekül  $\bar{A}$ .

A| $\bar{A}$ -A $\bar{A}$  Hybridisierung der Moleküle A und  $\bar{A}$  auf dem Molekül A $\bar{A}$ .

[A] Konzentration des Moleküls A.

[A]<sub>0</sub> Anfangskonzentration des Moleküls A.

[A]<sub>Plat.</sub> Plateaukonzentration des Moleküls A.

[A]<sub>max</sub> Maximalkonzentration des Moleküls A.

$\overline{[A]}$  Konzentration des Moleküls, welches im Ensemble die größte Konzentration zeigt.

n, m beliebige Längen

k, l  $k, l \in \mathbb{N}$ : Längen der Form  $n = 2^k$  und  $m = 2^l$ .

**Vollüberlappung** Vollständige Hybridisierung von zwei komplementären Strängen (ohne Überhänge).

**ML-Schätzer** Maximum-Likelihood-Schätzer



# Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form in einem Verfahren zur Erlangung eines akademischen Grades vorgelegt.

---

Ort, Datum

---

Philipp Zimmer



# Danksagung

Für die vielen Anregungen, Ideen, Diskussionen und seine Unterstützung möchte ich mich bei meinem Betreuer Prof. Dr. Dr. Karsten Kruse bedanken. Mein Dank geht auch, für die vielen hilfreichen Diskussionen, an Prof. Dr. Albrecht Ott. Emanuel Worst möchte ich für die experimentelle Umsetzung und die vielen nützlichen Gespräche danken. Meinen Dank möchte ich auch Dr. Jan Nagler für die Anregungen und den Gedankenaustausch zum Thema Perkolation bekunden.

Mein Dank geht auch an alle Mitglieder der Arbeitsgruppe Kruse, insbesondere an Alexander Dreher, Denis Johann, Dominic Jourdain und Christopher Zapp für die angenehme Zusammenarbeit. Lisa Veemees und Florian Jäckel möchte ich auch für das sorgfältige Korrekturlesen danken. Und mein Dank gilt allen, die mir ihr Vertrauen in die Fertigstellung dieser Arbeit ermutigend entgegengebracht haben.



# Literaturverzeichnis

- [1] BAYRHUBER, KULL, Linder Biologie, (Schroedel Verlag, Hannover, DE), 1998
- [2] CASSIDY, SCHWARTZ, Prader-Willi and Angelman Syndromes, *Reviews in Molecular Medicine* 77: 140-151, 1998
- [3] VOLTERRA, Variations and Fluctuations of the Number of Individuals in Animal Species living together, *J. Cons. int. Explor. Mer* 3: 3-51, 1928
- [4] BROWN, BOMBERGER BROWN, Where has all the road kill gone? *Current Biology* 23: R233-R234, 2013
- [5] CARROLL, HENDRY, REZNIK, FOX, Evolution on ecological time-scales, *Functional Ecology* 21: 387-393, 2007
- [6] Evolution aus Menschenhand, Regie: Steve Nicholls, Sarah Whally, AUT, Arte FRA 2013
- [7] TRUT, Early Canid Domestication: The Farm-Fox Experiment, *American Scientist* 87: 160-169, 1999
- [8] HENDRY, NOSIL, RIESEBERG, The speed of ecological speciation, *Functional Ecology* 21: 455-464, 2007
- [9] LENSKI, ROSE, SIMPSON, TADLER, Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence During 2000 Generations, *The American Naturalist* 138: 1315-1341, 1991
- [10] BLOUNT, BORLAND, LENSKI, Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*, *Proc Natl Acad Sci USA* 105: 7899-7906, 2008
- [11] LE GAC, PLUCAIN, HINDRÉ, LENSKI, SCHNEIDER, Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with *Escherichia coli*, *Proc Natl Acad Sci USA* 109: 9487-9492, 2012
- [12] ELENA, LENSKI, Evolution experiments with micororganisms: the dynamics and genetic bases of adaptation, *Nat Rev Gen* 4: 457-469, 2003
- [13] STOLLMEIER, GEISEL, NAGLER, Possible Origin of Stagnation and Variability of Earth's Biodiversity, *Phys Rev Lett* 112: 228101 (2014)
- [14] CAMPBELL, Biologie, (Spektrum Verlag, Heidelberg, DE), 1997

- [15] GILBERT, Origin of life: The RNA world, *Nature* 319: 618, 1986
- [16] DARWIN, On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life, (John Murray, London, UK), 1859
- [17] GAUSE, Experimental Analysis of Vito Volterra's Mathematical Theory of the Struggle for Existence, *Science* 79: 16-17, 1934
- [18] GAUSE, Experimentelle Untersuchungen über die Konkurrenz zwischen Paramecium caudatum und Paramecium aurelia, *Archiv für Protistenkunde* 84: 207-224, 1935
- [19] EIGEN, SCHUSTER, The Hypercycle. A principle of natural selforganization. Part A: Emergence of the hypercycle, *Die Naturwissenschaften* 64: 541-565, 1977
- [20] NOLTING, Thermodynamik, (Springer, Wiesbaden, DE), 2010
- [21] NOLTING, Statistische Physik, (Springer, Wiesbaden, DE), 2007
- [22] SCHRÖDINGER, What is Life, (Cambridge University Press, Cambridge, UK), 1944
- [23] DRAKE, CHARLESWORTH, CHARLESWORTH, CROW, Rates of Spontaneous Mutation, *Genetics* 148: 1667-168, 1998
- [24] GAGO, ELENA, FLORES, SANJOUAN, Extremely high mutation rate of a hammerhead viroid, *Science* 323: 1308, 2009
- [25] SELF, The effects and consequences of very large explosive volcanic eruptions, *Phil. Trans. R. Soc. A* 364: 2073-2097, 2006
- [26] MARTIN, BAROSS, KELLEY, RUSSEL, Hydrothermal vents and the origin of life, *Nature Reviews* 6: 805-814, 2008
- [27] MARTIN, RUSSEL, On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells, *Phil. Trans. R. Soc. Lond. B* 358: 59-85, 2002
- [28] FELDMANN, KISLEV, Domestication of emmer wheat and evolution of free-threshing tetraploid wheat, *Israel Journal of Plant Sciences* 55: 207-221, 2007
- [29] BELYAEV, TRUT, RUVINSKY, Genetics of the W Locus in Foxes and Expression of its Lethal Effects, *Journal of Heredity* 66: 331-338, 1975
- [30] KOLDAEVA, KOLDAEV, The effects of domestication on the fur animal production, *Vestnik VOGiS* 11: 62-75, 2007

- [31] SILBERFUCHS: [https://commons.wikimedia.org/wiki/File:Silberfuchs\\_06.jpg](https://commons.wikimedia.org/wiki/File:Silberfuchs_06.jpg), Lizenz: CC-BY-2.0-DE, Autor: Zefram, 04.01.2016
- [32] GEORGIAN WHITE RUSSIAN DOMESTICATED RED FOX: [https://en.wikipedia.org/wiki/File:Georgian\\_white\\_Russian\\_domesticated\\_Red\\_Fox.jpg](https://en.wikipedia.org/wiki/File:Georgian_white_Russian_domesticated_Red_Fox.jpg), Lizenz: CC-BY-3.0, Autor: Kayfedewa, 04.01.2016
- [33] ALLENDORF, HARD, Human-induced evolution caused by unnatural selection through harvest of wild animals, *Proc Natl Acad Sci USA* 106: 9987-9994, 2009
- [34] WOLAK, GILCHRIST, RUZICKA, NALLY, RANDOLPH, CHAMBERS, A Contemporary, Sex-Limited Change in Body Size of an Estuarine Turtle in Response to Commercial Fishing, *Conservation Biology* 24: 1268-1277, 2009
- [35] BOMBERGER BROWN, BROWN, Intense Natural Selection on Morphology of Cliff Swallows (*Petrochelidon Pyrrhonota*) a Decade Later: Did the Population Move between Adaptive Peaks? *The Auk* 128, 96-77, 2011
- [36] HAMILTON, The Adaptive Significances of Intraspecific Trends of Variation in Wing Length and Body Size Among Bird Species, *Evolution* 15: 180-195, 1961
- [37] HENDRY, GRANT, GRANT, FORD, BREWER, PODOS, Possible human impacts on adaptive radiation: beak size bimodality in Darwin's finches, *Proc. R. Soc. B* 273: 1887-1894, 2006
- [38] HUBER, LEÓN, HENDRY, BERMINGHAM, PODOS, Reproductive isolation of sympatric morphs in a population of Darwin's finches, *Proc. R. Soc. B* 274: 1709-1714, 2007
- [39] MORRIS, WATSON, Positively versus Negatively Frequency-Dependent Selection, *Adv. Artif. Life. Lect. Notes Artif. Intell.* 5778: 77-84, 2009
- [40] LÜTGENS, GOTTSCHALK, Why a Co-substrate is Required for Anaerobic Growth of *Escherichia coli* on Citrate, *Journal of General Microbiology* 119: 63-70, 1980
- [41] SMITH, PARSELL, GREEN, Thermosensitive H1 plasmids determining citrate utilization, *Journal of General Microbiology* 109: 305-311, 1978
- [42] ZIMBRO, POWER, MILLER, WILSON, JOHNSON, Difco & BBL Manual, (Becton, Dickinson and Company, Sparks, Maryland, USA), 2009
- [43] GOULD, Wonderful Life, (W. W. Norton & Company, New York, USA), 1989
- [44] CONWAY MORRIS, Life's Solution: Inevitable humans in a Lonely Universe, (Cambridge University Press, Cambridge, UK), 2003

- [45] MILLER, Production of Some Organic Compounds under Possible Primitive Earth Conditions, *Journal of the American Chemical Society* 77: 2351-2361, 1955
- [46] JOHNSON, CLEAVES, DWORKIN, GLAVIN, LAZCANO, BADA, The Miller Volcanic Spark Discharge Experiment, *Science* 322: 404, 2008
- [47] WOLLRAB, SCHERER, AUBRIET, CARRÉ, CARLOMAGNO, CODUTTI, OTT, Chemical Analysis of a Miller-Type Complex Prebiotic Broth, *Orig Life Evol Biosph* 46(2-3): 149-169, 2016
- [48] MILLER, ORGEL, The Origins of Life in the Earth, (Prentice-Hall Inc., New Jersey, USA), 1974
- [49] VON KIEDROWSKI, A Self-Replicating Hexadeoxynucleotide, *Angew. Chem. Int. Ed.* 10: 932-935, 1986
- [50] VON KIEDROWSKI, WLOTZKA, HELBING, MATZEN, JORDAN, Parabolic Growth of a Self-Replicating Hexadeoxynucleotide Bearing a 3'-5'-Phosphoamidate Linkage, *Angew. Chem. Int. Ed.* 30: 423-426, 1991
- [51] ACHILLES, VON KIEDROWSKI, A Self-Replicating System from Three Starting Materials, *Angew. Chem. Int. Ed.* 32: 1198-1201, 1993
- [52] ZIELINSKI, ORGEL, Autocatalytic synthesis of a tetranucleotide analogue, *Nature* 327: 346-347, 1987
- [53] EIGEN, Selforganization of matter and the evolution of biological macromolecules, *Die Naturwissenschaften* 58: 465-523, 1971
- [54] EIGEN, SCHUSTER, The Hypercycle. A principle of natural selforganization. Part B: the abstract hypercycle, *Die Naturwissenschaften* 65: 7-41, 1978
- [55] EIGEN, SCHUSTER, The Hypercycle. A principle of natural selforganization. Part C: the realistic hypercycle, *Die Naturwissenschaften* 65: 341-369, 1978
- [56] SCHUSTER, Complex Adaptive Systems, (Scator-Verlag, Saarbrücken, DE), 2001
- [57] KAUFFMAN, Metabolic Stability and Epigenesis in Randomly Constructed Genetic Nets, *J. Theoret. Biol.* 22: 437-467, 1969
- [58] MAST, BRAUN, Thermal Trap for DNA Replication, *Phys. Rev. Lett.* 104: 188102, 2010
- [59] KREYSING, KEIL, LANZMICH, BRAUN, Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length, *Nature Chemistry* 7: 203-208, 2015
- [60] OBERMAYER, KRAMMER, BRAUN, GERLAND, Emergence of Information Transmission in a Prebiotic RNA Reactor, *Phys. Rev. Lett.* 107: 018101, 2011

- [61] FELIX, Albrecht Kossel. Leben und Werk., *Die Naturwissenschaften* 17: 473-478, 1955
- [62] WATSON, CRICK, A Structure for Deoxyribose Nucleic Acid, *Nature* 171: 737-738, 1953
- [63] DIMITROV, ZUKER, Prediction of Hybridization and Melting for Double-Stranded Nucleic Acids, *Biophysical Journal* 87: 215-226, 2004
- [64] DNA-DOPPELHELIX: [https://de.wikipedia.org/wiki/Datei:DNA\\_simple.svg](https://de.wikipedia.org/wiki/Datei:DNA_simple.svg), 11.02.2016
- [65] SANTA LUCIA, HICKS, The Thermodynamics of DNA structural motifs, *Annu. Rev. Biophys. Biomol. Struct.* 33: 415-440, 2004
- [66] XODO, BOOM, Oligodeoxynucleotide Folding in Solution: Loop Size and Stability of B-Hairpins, *Biochemistry* 27: 6321-6326, 1998
- [67] OULDRIDGE, ŠULC, ROMANO, DOYE, LOUIS, DNA hybridization kinetics: zippering, internal displacement and sequence dependence, *Nucleic Acids Research* 41: 8886-8895, 2013
- [68] PÖRSCHKE, Model calculations on the kinetics of oligonucleotide double helix coil transitions. Evidence for a fast chain sliding reaction, *Biophysical Chemistry* 2,2: 83-96, 1974
- [69] PRIGOGINE, GLANSDORFF, Thermodynamic Theory of Structure, Stability and Fluctuations, (Wiley Interscience, London, GB), 1971
- [70] WICKEN, Information transformations in molecular evolution, *J. theor. Biol.* 72: 191-204, 1978
- [71] DUHR, BRAUN, Why molecules move along a temperature gradient, *Proc Natl Acad Sci USA* 103: 19678-1968, 2006
- [72] WEGGLER, HUKU, LÜCKE, Roll and square convection in binary liquids: A few-mode Galerkin model, *Phys. Rev. E* 81: 016309, 2010
- [73] WORST, Enzymbasierte Selbstreplikation von DNA Fragmenten, Diplomarbeit, Physik und Mechatronik, Universität des Saarlandes, 2012
- [74] HOFACKER, FONTANA, STADLER, BONHOEFFER, TACKER, SCHUSTER, Fast folding and comparison of RNA secondary structures, *Monatshefte für Chemie* 125: 167-188, 1994
- [75] MARKHAM, ZUKER, UNAFold: software for nucleic acid folding and hybridization, *Bioinformatics, Volume II. Structure, Function and Applications, Humana Press* 453: 3-31, 2008
- [76] MARKHAM, ZUKER, DINAMelt web server for nucleic acid melting prediction, *Nucleic Acids Res* 33: W577-W581, 2005

- [77] SADOWNIK, MATTIA, NOWAK, OTTO, Diversification of self-replicating molecules, *Nat. Chem.* 8: 264-269, 2016
- [78] VON KIEDROWSKI, Minimal Replicator Theory I: Parabolic Versus Exponential Growth, *Bioorganic Chemistry Frontiers* 3: 113-146, 1993
- [79] LEUBKE, DERVAN, Nonenzymatic Ligation of Oligodeoxyribonucleotides on a Duplex DNA Template by Triple-Helix Formation, *J. Am. chem. Soc.* 111: 8733-8735, 1989
- [80] ZIMMER, Theoretische Untersuchung der Entstehung und Evolution von selbstreproduzierenden, chemischen Reaktionszyklen, Diplomarbeit, Physik und Mechatronik, Universität des Saarlandes, 2010
- [81] SZATHMÁRY, Simple Growth Laws and Selection Consequences, *Trends in Ecology & Evolution* 6: 366-370, 1991
- [82] WORST, ZIMMER, WOLLRAB, KRUSE, OTT, Unbounded growth patterns of reproducing, competing polymers – similarities to biological evolution, *New J. Phys.* 18: 103003, 2016
- [83] WOESE, FOX, The concept of cellular evolution, *J. Mol. Evol.* 10: 1, 1977
- [84] WOESE, The universal ancestor, *Proc. Natl. Acad. Sci. USA* 95: 6854, 1998
- [85] WOESE, On the evolution of cells, *Proc. Natl. Acad. Sci. USA* 99: 8742, 2002
- [86] NIRENBERG, MATTHAEI, The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides, *PNAS* 47: 1588-1602, 1961
- [87] NIRENBERG, LEDER, BERNFIELD, BRIMACOMBE, TRUPIN, ROTTMAN, O'NEAL, RNA codewords and protein synthesis, VII. On the general nature of the RNA code, *PNAS* 53: 1161-1168, 1965
- [88] LEHN, From supramolecular chemistry towards constitutional dynamic chemistry and adaptive chemistry, *Chem. Soc. Rev.* 36: 151-160, 2007
- [89] LEHN, Constitutional Dynamic Chemistry: Bridge from Supramolecular Chemistry to Adaptive Chemistry, *Top. Curr. Chem.* 322: 1-32, 2011
- [90] MÜLLER, DAS, BÖGGE, SCHMIDTMANN, BOTAR, PATRUT, Generation of cluster capsules ( $I_h$ ) from decomposition products of a smaller cluster (Keggin- $T_d$ ) while surviving ones get encapsulated: species with core-shell topology formed by a fundamental symmetry-driven reaction, *Chem. Commun.* 7: 657-658, 2001
- [91] ODLING-SMEE, LALAND, FELDMAN, Niche Construction, *The American Naturalist* 147: 641-648, 1996

- [92] LALAND, ODLING-SMEE, FELDMAN, Evolutionary consequences of niche construction and their implications for ecology, *PNAS* 96: 10242-10247, 1999
- [93] LALAND, ODLING-SMEE, FELDMAN, Niche construction, biological evolution, and cultural change, *Behavioral an brain sciences* 23: 131-175, 2000
- [94] MATTOSO, MOREIRA, SAMUELS, Symbiotic bacteria on the cuticle of the leaf-cutting ant *Acromyrmex subterraneus subterraneus* protect workers from attack by entomopathogenic fungi, *Biology Letters* doi:10.1098/rsbl.2011.0963, 2011
- [95] FREUDIG, Lexikon der Biologie, (Spektrum Akademischer Verlag, Heidelberg, D), 2006
- [96] ALISKAN, TERZI, Allelopathic effects of walnut leaf extracts and juglone on seed germination and seedling growth, *J. Horticultural Science and Biotechnology* 76: 436-440, 2011
- [97] LOVELOCK, MARGULIS, Atmospheric homeostasis by and for the biosphere: the gaia hypothesis, *Tellus A* 26: 1-2, 1973
- [98] LENTON, Gaia and natural selection, *Nature* 394: 439-447, 1998
- [99] SCHWARZ, CARBON TOXICITY IN PLANTS, *Acta Hortic.* 481: 685-688, 1999
- [100] KIMURA, Evolutionary Rate at the Molecular Level, *Nature* 217: 624-626, 1968
- [101] LINCOLN, JOYCE, Self-Sustained Replication of an RNA Enzyme, *Science* 323: 1229-1232, 2009
- [102] WORST, Dissertation, Naturwissenschaftlich-Technische Fakultät, Universität des Saarlandes, in Vorbereitung, voraussichtlich 2017
- [103] ABRÀMOFF, MAGALHÃES, RAM, Image Processing with ImageJ, *Biophotonics International* 11: 36-42, 2004
- [104] NEB: <https://www.neb.com/~media/Catalog/All-Products/58CE8774C4BC4BA1B2783E1418D93ED4/Datacards%20or%20Manuals/M0208Ddatasheet-Lot0101402.pdf>, 04.06.2015
- [105] RITTIÉ, PERBAL, Enzymes used in molecular biology: a useful guide, *J. Cell Commun. Signal.* 2: 25-45, 2008
- [106] SANTALUCIA, A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics, *PNAS* 95: 1460-1465, 1998
- [107] BARANY, The ligase chain reaction in a PCR world, *PCR Methods Appl.* 1: 5-16, 1991

- [108] NEB: <https://www.neb.com/faqs/1/01/01/what-is-the-activity-of-taq-dna-ligase-at-various-temperatures>, 27.06.2016
- [109] SIGMA-ALDRICH: <https://www.sigmaaldrich.com/Graphics/Supelco/objects/4800/4709.pdf>, 06.06.2015
- [110] DOCHECK: <http://flexikon.doccheck.com/de/Dithiothreitol>, 06.06.2015
- [111] BELENKY, BOGAN, BRENNER, NAD<sup>+</sup> metabolism in health and disease, *TRENDS in Biochemical Sciences* 32: 12-19, 2006
- [112] SIGMA-ALDRICH: <http://www.snowpure.com/docs/triton-x-100-sigma.pdf>, 06.06.2015
- [113] NEB: <https://www.neb.com/~media/Catalog/All-Products/96BB95515E8F4FF79AC5A97AD83116C2/Datacards%20or%20Manuals/M0250Datashet-Lot0251604.pdf>, 08.06.2016
- [114] NEB: <https://www.neb.com/faqs/2011/09/12/can-mung-bean-nuclease-be-heat-inactivated>, 08.06.2016
- [115] NEB: <https://www.neb.com/~media/Catalog/All-Products/F93132F30E20426AA2184D370130E815/Datacards%20or%20Manuals/R0652Datashet-Lot0041302.pdf>, 08.06.2016
- [116] NEB: <https://www.neb.com/~media/Catalog/All-Products/B0607458AF3047239E36E53E6966FE13/Datacards%20or%20Manuals/R0518Datashet-Lot0111212.pdf>, 08.06.2016
- [117] NEB: <https://www.neb.com/products/r0518-bstui>, 09.06.2016
- [118] NEB: <https://www.neb.com/products/r0652-afei>, 09.06.2016
- [119] LONZA: [http://bio.lonza.com/uploads/tx\\_mwaxmarketingmaterial/Lonza\\_BenchGuides\\_SourceBook\\_Section\\_VII\\_-\\_Separation\\_of\\_DNA\\_in\\_Polyacrylamide\\_Gels.pdf](http://bio.lonza.com/uploads/tx_mwaxmarketingmaterial/Lonza_BenchGuides_SourceBook_Section_VII_-_Separation_of_DNA_in_Polyacrylamide_Gels.pdf), 09.06.2016
- [120] HORSPOOL, COOPE, HOLT, Efficient assembly of very short oligonucleotides using T<sub>4</sub> DNA Ligase, *BMC Research Notes* 3: 291, 2010
- [121] SHORE, LANGOWSKI, BALDWIN, DNA flexibility studied by covalent closure of short fragments into circles, *PNAS* 78: 4833-4837, 1981
- [122] ZIMMER, KRUSE, NAGLER, Anomalous percolation in molecular evolution, (*in Vorbereitung*), 2017
- [123] GILLESPIE, A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions, *J. Comp. Phys.* 22: 403-434, 1976

- [124] GILLESPIE, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.* 81: 2340-2361, 1977
- [125] KRAUTH, Statistical mechanics: algorithms and computations, (Oxford University Press, Oxford, UK), 2006
- [126] TANGE, GNU Parallel - The Command-Line Power Tool, *The USENIX Magazine* 36: 42-47, 2011
- [127] BARROW, Physikalische Chemie, (Bohmann-Verlag, Heidelberg, DE), 1984
- [128] MACKLEM, Emergent phenomena and the secrets of life, *J Appl Physiol* 104: 1844-1846, 2008
- [129] WALZ, Lexikon der Mathematik, (Spektrum, Heidelberg, DE), 2001
- [130] PESTMAN, Mathematical Statistics, (Walter de Gruyter, Berlin, DE), 1998
- [131] ZHANG, HAMMER, GRAVES, Competitive hybridization kinetics reveals unexpected behavior patterns, *Biophys. J.* 89: 2950-2959, 2005
- [132] SCHRÖDER, EBRAHIMNAZHAD RAHBARI, NAGLER, Crackling noise in fractional percolation, *Nature Communications* 4: 2222, 2013
- [133] D'SOUZA, NAGLER, Anomalous critical and supercritical phenomena in explosive percolation, *Nature Physics* 11: 531-538, 2015
- [134] SABERI, RAHBARI, DASHTI-NASERABADI, ABBASI, CHO, NAGLER, Universality in boundary domain growth by sudden bridging, *Scientific reports* 6: 21110-7, 2016
- [135] SCHRÖDER, CHEN, NAGLER, Discrete scale invariance in supercritical percolation, *New J. Phys.* 18: 013042, 2016
- [136] CHO, MAZZA, KAHNG, NAGLER, Genuine non-self-averaging and ultraslow convergence in gelation, *Phys Rev E* 94: 022602-8, 2016
- [137] BOHMAN, FRIEZE, WORMALD, Avoidance of a Giant Component in Half the Edge Set of a Random Graph, *Random Structures & Algorithms* 25: 432-449, 2004
- [138] CHEN, CHENG, ZHENG, CHUNG, D'SOUZA, NAGLER, Unstable supercritical discontinuous percolation transitions, *Phys Rev E* 88: 042152-9, 2013
- [139] CHEN, NAGLER, CHENG, JIN, SHEN, ZHENG, D'SOUZA, Phase transitions in supercritical explosive percolation, *Phys Rev E* 87: 052130-9, 2013
- [140] SORNETTE, Critical Phenomena in Natural Sciences, (Springer-Verlag, Heidelberg, DE), 2006

- [141] CHEN, SCHRÖDER, D'SOUZA, SORNETTE, NAGLER, Microtransition Cascades to Percolation, *PRL* 113: 155701, 2014
- [142] SIMÁNYI, Proof of the Boltzmann-Sinai ergodic hypothesis for typical hard disk systems, *Invent. math.* 154: 123-178, 2003
- [143] SINAI, Dynamical systems with elastic reflections, *Russ. Math. Surv.* 25: 137-189, 1970
- [144] THIRRING, Lehrbuch der Mathematischen Physik. 1. Klassische Dynamische Systeme, (Springer-Verlag, Wien, AUT), 1977
- [145] Aus [135] unter Lizenz: CC-BY-3.0; Autoren: Malte Schröder, Wei Chen, Jan Nagler; Titel: Discrete scale invariance in supercritical percolation; Journal: New Journal of Physics 18(2016)013042; DOI: 10.1088/1367-2630/18/1/013042; Änderungen: Einzelabbildungen umsortiert, Abbildungsbezeichnung entfernt bzw. ersetzt und Achsenbeschriftung angepasst.
- [146] RIORDAN, WARNKE, Achlioptas processes are not always self-averaging, *Phys Rev E* 86: 011129-4, 2012
- [147] CHO, KAHNG, KIM, Cluster aggregation model for discontinuous percolation transitions, *Phys Rev E* 81: 030103-4, 2010
- [148] GMEHLING, KOLBE, KLEIBER, RAREY, Chemical Thermodynamics for Process Simulation, (Wiley-VCH, Weinheim, DE), 2012
- [149] SMOLUCHOWSKI, Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen, *Pisma Mariana Smoluchowskiego* 1: 530-594, 1927