

I Einleitung

I.1 Erhaltung der vollständigen genetischen Information in linearen Genomen

Organismen können als Art nur erhalten bleiben, wenn sie ihr Genom vollständig und fehlerfrei kopieren und weitergeben. Um diese *conditio sine qua non* zu erfüllen, hat die Evolution unterschiedliche Strategien hervorgebracht.

(i) ein zirkuläres Genom und einen Replikationsmechanismus, der sicherstellt, daß das gesamte Genom ohne den Verlust einer einzigen Base repliziert wird. Dabei spielt es keine Rolle, ob die DNA einzelsträngig oder doppelsträngig ist. In vielen Bakterien (z.B. *E. coli*, und *B. subtilis*) und Bakteriophagen (ϕ X174, P1, P4 und λ ; nur während des lytischen Zyklus) ist ein zirkuläres Genom und ein solcher Mechanismus zu finden. Der Vorteil eines zirkulären Genoms besteht darin, daß es keine DNA-Enden gibt, bei deren Replikation genetische Informationen verloren gehen könnte. Für die Replikation zirkulärer Genome existieren mehrere gut verstandene Modelle, die aus den Bakteriophagensystemen M13 oder ϕ X174 entwickelt wurden (Schekman *et al.*, 1972; Arai *et al.*, 1980; Meyer *et al.*, 1979). Die an Bakteriophagen gewonnenen Erkenntnisse wurden dann zur Entwicklung eines Modells der DNA-Replikation in *E. coli* verwendet (Kornberg & Baker, 1992; Messer & Weigel, 1996)

(ii) ein lineares Genom, kombiniert mit einem Mechanismus, der sicherstellt, daß alle Basen kopiert werden. Enden linearer DNA sind dabei problematisch, da sie von replikativer Polymerase nicht kopiert werden können (Watson, 1972). Die Bakteriophagen T4, T7, ϕ 29, PRD1, N15 und PY54 haben z.B lineare Genome während des lytischen Modus. Ihre DNA-Enden sind durch die Bildung konkatemerer Replikationsintermediate mit terminaler Redundanz (T4, Sechaud *et al.*, 1965; T7, White & Richardson, 1987a; White & Richardson, 1987b), terminale Proteine (ϕ 29, Blasco *et al.*, 1992; PRD1, Jung *et al.*, 1987; Hsieh *et al.*, 1990; Jung *et al.*, 1987; Hsieh *et al.*, 1990) oder kovalent geschlossene DNA-Enden (N15, Golub & Ravin, 1967; PY54, Popp *et al.*, 2000; Popp *et al.*, 2000) geschützt. Eine Reihe von Bakterien (z.B. *Streptomyces spp.*) haben ebenfalls lineare Genome, deren Enden durch terminale Proteine geschützt sind. Es gibt auch Mechanismen, die sicherstellen, daß ein Informationsverlust an den DNA-Enden kein Gen beschädigt - durch Telomersequenzen, die bei der Replikation für den Organismus schadlos verloren gehen können. Die meisten Eukaryonten haben Telomere. Telomere sind doppelsträngige DNA-Enden, die keine

genetische Information tragen und aus bis zu 1000 *tandem repeats* mit einer speziesabhängigen G-reichen Sequenz (z. B. TTGGGGTTG bei *Tetrahymena*) bestehen. Die Gesamtlänge der Telomere hängt von der Spezies und - bei Mehrzellern - vom Alter des Organismus ab. Diese Sequenz schützt die vor den DNA-Enden liegenden Gene vor replikativem Schwund bei der Zellteilung. Telomere werden unter bestimmten Bedingungen (z. B. embryonales Stammzellstadium beim Menschen) vom einem Enzym namens Telomerase wieder verlängert. Dieses Enzym besteht aus einer Protein- und einer RNA-Komponente. Die RNA dient der Telomerase als *primer/template* bei der Synthese neuer *repeats* (Greider & Blackburn, 1989).

Die Replikation linearer Genome mit terminalen Proteinen und terminalen Wiederholungen war - genau wie die Replikation zirkulärer Genome - ebenfalls Ziel zahlreicher Untersuchungen, und ist größtenteils aufgeklärt. Das Verständnis der Replikation von kovalent geschlossenen, linearen Genomen steht jedoch erst am Anfang (Casjens, 1999).

Lineare Genome mit kovalent geschlossenen Enden, also eine Untergruppe der linearen Genome, sind in verschiedenen Viren wie Poxviridae (z.B. das Vaccinia Virus, Baroudy *et al.*, 1982), Iridopoxviridae (z.B. das afrikanische *swine fever virus*, Gonzalez *et al.*, 1986), einem Parovirus (*minute virus* in Mäusen, Cotmore & Tattersall, 1987) und in den Genomen und Plasmiden von Mitochondrien und Plastiden (Morin & Cech, 1986, Cummings, 1992, Dinouel *et al.*, 1993) zu finden. In Eubakterien sind diese Strukturen z. B. in *Borrelia spp.*, aber auch in *Agrobacterium tumefaciens* bekannt, sowohl in Chromosomen als auch in Plasmiden (Casjens *et al.*, 1997) Jumas-Bilak *et al.*, 1998; Goodner *et al.*, 2001; Barbour & Garon, 1987.

Aus der Gruppe der Bakteriophagen sind zwei Vertreter bekannt, die diese DNA-Konfiguration besitzen: N15 ein *E. coli* Phage (Golub & Ravin, 1967) und PY54 (Popp *et al.*, 2000), ein Phage, dem *Yersinia enterocolitica* als Wirt dient und der im Zeitraum der Entstehung dieser Arbeit entdeckt wurde. Vorläufige Untersuchungen zeigen, daß PY54 große genetische und morphologische Ähnlichkeit mit N15 besitzt (S. Hertwig, & B. Appel, pers. Mitteilung).

1.2 Der Bakteriophage N15 als Modellsystem für das Studium linearer Genome

Der Bakteriophage N15 wurde Anfang der 60er Jahre im Abwasser der Kanalisation Moskaus entdeckt und 1967 von Viktor Ravin erstmalig beschrieben. Die Nukleotid-Sequenz ist

bekannt (46.375 bp, GenBank No. AF064539, Hendrix *et al.*, 1998). N15 gehört zur Gruppe der Siphophagen, zu der auch λ und P4 zählen, und benutzt *E. coli* als Wirt (Svarchevsky & Rybchin, 1984). Neueste DNA-sequenzbasierte, computergestützte, taxonomische Untersuchungen zeigen, daß die Verwandtschaft von N15 zu λ enger ist, als die der lambdoiden Phagen HK022 oder HK97 (Rohwer & Edwards, 2002), da etwa 50 % des Genoms von N15 für Proteine kodieren, die entsprechende Homologe des Phagen λ sind. Es handelt sich hauptsächlich um Morphogenesegene.

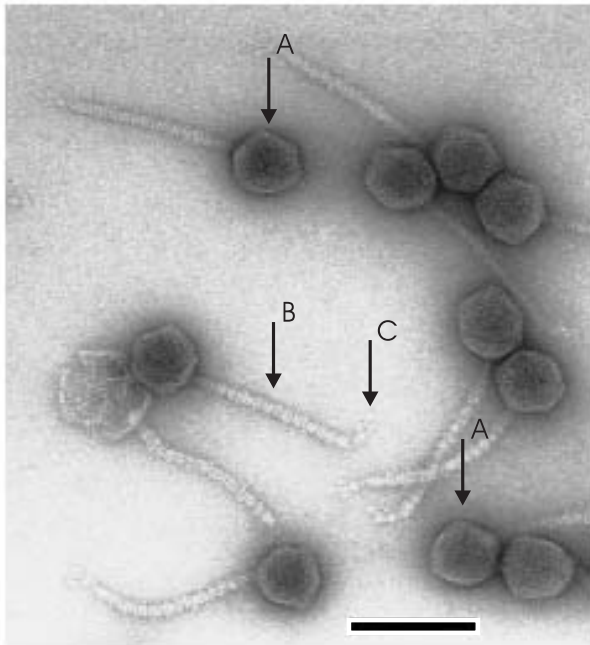


Abb. 1 Die Morphologie des lambdoiden Bakteriophage N15

In dieser elektronenmikroskopischen Aufnahme sind die Phagenköpfe (A), sowie der Phagenschwanz (B) zu sehen. C markiert die Schwanzfasern. Der Balken entspricht 100 nm (Bildquelle: Ravin *et al.*, 2000)

N15 ist ein temperenter Phage (Abb. 1), dessen doppelstängige, linear verpackte DNA kohäsive Enden (*cos*-Enden) besitzt. Nach Infektion von *E. coli* Zellen wird die Phagen-DNA - wie beim Bakteriophagen λ - an den *cos*-Enden zirkularisiert (Abb. 2, oben). Die *cos*-sites von λ und N15 sind bis auf 3 bp identisch. Nach der Zirkularisierung fällt die Entscheidung, ob der lytische Zyklus beginnt, oder ob der Phage in den lysogenen Modus eintritt. Beim lytischen Zyklus werden neue Phagenpartikel produziert, deren Biogenese und Morphogenese nahezu identisch zu der beim Phagen λ beobachteten ablaufen sollte (Abb. 2, rechte Seite). Die Phagen-DNA wird repliziert, und die entstehenden Konkatemere werden an den *cos*-sites von der N15-codierten Terminase (N15 gpA und gpNu1) geschnitten und so auf Einheitslänge - jeweils ein komplettes Genom - gebracht. Die DNA wird anschließend in Phagenköpfe verpackt. Nach Komplettierung der Phagen erfolgt die Lyse der Wirtszelle. Dieses Modell wurde aufgrund der Analyse der Nukleotidsequenz des Genoms von N15 und dessen Vergleich mit dem Genom des Bakteriophagen λ vorgeschlagen (Ravin *et al.*, 2000).

Die Nukleotidsequenz läßt aber auch auf grundlegende Unterschiede zu λ im lysogenen Modus des Phagen N15 schließen. Anders als bei λ , ist der Prophage von N15 nicht in das Wirtsgenom integriert, sondern liegt als lineares Plasmid mit kovalent geschlossenen Enden vor. Beim Eintritt in den lysogenen Zyklus (Abb. 2, linke Seite), entsteht der lineare Prophage aus dem zirkulären Vorläufermolekül. Das neu gebildete, lineare Molekül wird über einen nur teilweise geklärten Mechanismus repliziert, dessen Modell in Abschnitt 1.3 ausführlicher diskutiert wird. Erfolgt im lysogenen Stadium die Induktion (z.B. durch Nahrungsstreß oder Einwirkung von UV-Licht), so entsteht - ebenfalls über einen

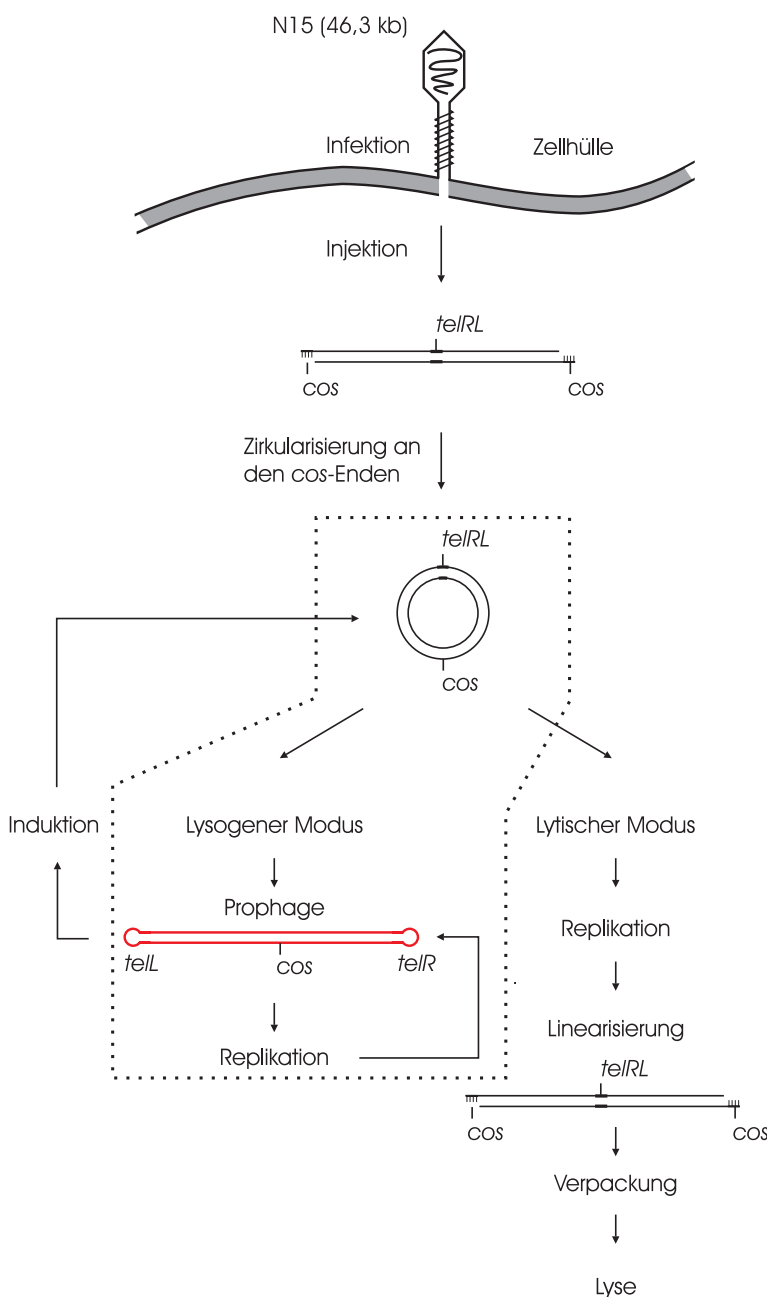


Abb. 2 Der lytische und lysogene Zyklus des Bakteriophagen N15

Schematische Darstellung der Infektion des Wirtes, Zirkularisierung des Genoms und der Entstehung des Prophagen (rot). Der lytische Weg der Replikation ist stark verkürzt dargestellt (Deneke *et al.*, 2002). Der mit Punkten umrandete Bereich hebt die Reaktion der *telomere resolution* hervor und ist in Abb. 6 nochmals gezeigt.

ungeklärten Mechanismus - wieder das zirkuläre N15-Phagengenom. Der Mechanismus der Induktion könnte aber auch in diesem Falle mit dem des Bakteriophagen λ verwandt sein, da die invertierten Repetitionen der *telomere resolution site* aus N15 den imperfekten Palindromen o_R und o_L des λ -*switch* ähneln (Ravin *et al.*, 2000). Der λ -*switch* reguliert den lysogenen Modus im Bakteriophagen λ (Johnson *et al.*, 1979). Analoge zu weiteren Regulationselementen des lysogenen Modus in λ sind ebenfalls bekannt und charakterisiert: CB aus N15 ist ähnlich dem λ -Prophagenrepressorprotein C1 und *Gen39* in N15 ist homolog zum Gen des Repressorproteins Cro aus λ .

Die Aufklärung des N15-Replikationsmechanismus während des Prophagenstadiums ist auch deshalb von besonderem Interesse, weil es tierpathogene (z. B. *swine fever virus*) und humanpathogene Organismen gibt, deren lineare Genomkonfiguration dem Prophagenstatus von N15 verwandt sind. Von besonderem Interesse ist dabei *Borrelia burgdorferi*, der Erreger der *Lyme-disease*, einer gefährlichen Hirnhautentzündung, die durch Zeckenbiß übertragen wird.

1.3 Die Replikation linearer, kovalent geschlossener Plasmide erfordert einen besonderen Reaktionsschritt

N15 bietet sich als Modellsystem für die Replikation von linearen Plasmiden mit kovalent geschlossenen Enden an, da von dem 46,4 kb großen Phagengenom ein 12,6 kb großes Minireplikon (Abb. 3) hergestellt wurde (Ravin & Ravin, 1998). Dieses repliziert autonom als lineares Plasmid, unabhängig von den chromosomal kodierten Replikationsinitiationsproteinen DnaA, DnaB, und DnaG (Svarchevsky, 1986; Vostrov *et al.*, 1992; Rybchin & Svarchevsky, 1999). Das 12,6 kb-Minireplikon kodiert für nur wenige

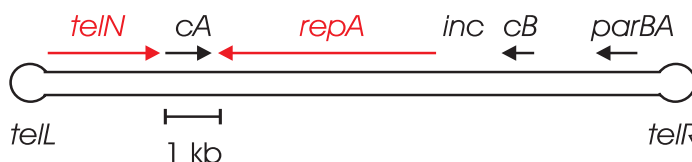


Abb. 3 Kleinstes noch linear replizierendes Miniplasmid, erzeugt aus N15

Genetische Elemente, die zur Replikation des linearen Plasmids essentiell sind, wurden über der schematischen Darstellung des Plasmids eingezeichnet (nach Rybchin & Svarchevsky, 1999). An der Replikation direkt beteiligte Proteine sind rot, regulatorische Elemente schwarz dargestellt. *telN* kodiert für die Protelomerase TelN, *repA* für das Replikationsprotein RepA. *cA* ist eine regulatorische RNA (analog zur P4 regulatorischen RNA CI). *inc* ist die Inkompatibilitätsregion. *CB* ist der Prophagenrepressor, *ParBA* sind *partitioning*-Proteine (beide analog zu P1). Der Balken entspricht 1 kb.

genetische Elemente (Abb. 3). Die beiden für die autonome Replikation des Miniplasmids zentralen Proteine sind RepA und TelN.

Werden *telN*, *cA*, *inc*, *cB* und *parBA* entfernt, entsteht ein zirkuläres Miniplasmid der Größe 5,6 kb, bestehend aus *repA* und einem Selektionsmarker. Dieses ist für eine autonome Replikation hinreichend (Rybchin & Svarchevsky, 1999). Daraus wurde geschlossen, daß TelN für die Erhaltung das linear replizierenden Plasmids essentiell, und für die Replikation des zirkulären Plasmids überflüssig ist (Ravin *et al.*, 2001). Da in dem zirkulären Miniplasmides kein Replikations-*origin* gefunden wurde, muß es einen nicht definierten *origin* in *repA* geben, welcher der Initiation der Replikation durch RepA dient. Einen ähnlichen Aufbau zeigt ein Minireplikon, das aus dem P4 Genom entwickelt wurde, und welches nur das Gen für das Replikationsprotein α und einen Selektionsmarker trägt (Briani *et al.*, 2001). Interessanterweise gibt es zwischen RepA und P4 α signifikante Ähnlichkeiten (Rybchin & Svarchevsky, 1999).

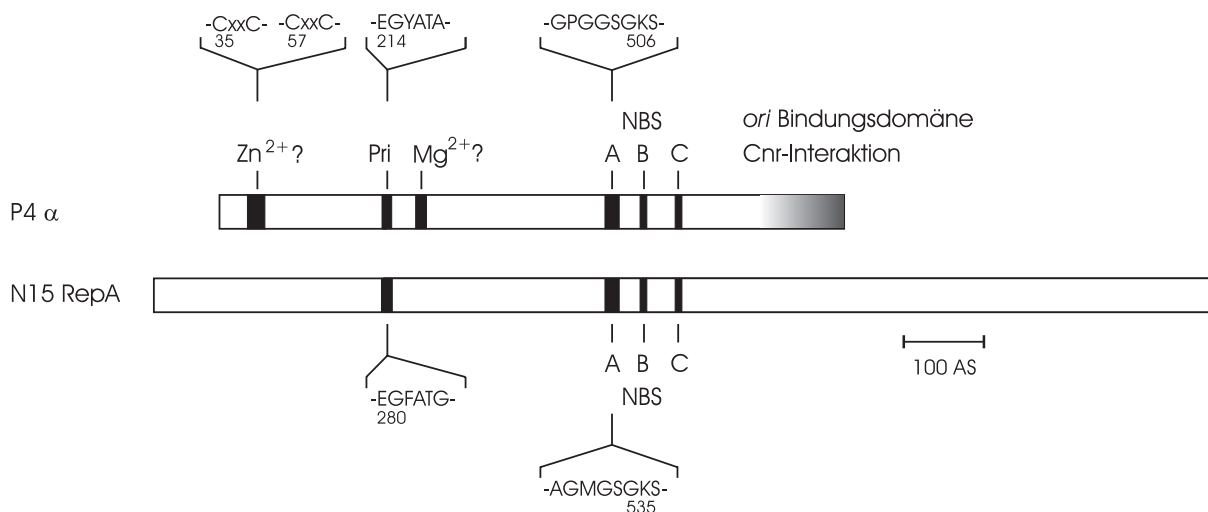


Abb. 4 Domänenstruktur von P4 α und N15 RepA

Schematische Balkendarstellung von Protein α des Bakteriophagen P4 und des postulierten RepA Proteins von N15. Motive sind als Boxen eingezeichnet und als Aminosäuresequenz über der Motivbezeichnung dargestellt. Zn^{2+} steht für ein Zinkfingermotiv, das in der T7 Primase/Helicase nachgewiesen wurde (Bernstein & Richardson, 1989), Pri für ein Primasemotiv (Strack *et al.*, 1992; Ziegelin *et al.*, 1993), Mg^{2+} für ein magnesiumionenbindendes DNA/RNA-Polymerasemotiv (Argos, 1988), und NBS für *nucleotide binding site* („Walker Box“ A und B, Walker *et al.*, 1982; Goralenya *et al.*, 1990). Die graue Box in P4 α markiert den Bereich der *origin*-Bindungsdomäne und der Interaktion mit dem Cnr-Protein (*copy number regulation* Protein). Der Balken entspricht 100 AS. Die Positionen der AS-Sequenzen im jeweiligen Protein sind angegeben.

Vom multifunktionellen Replikationsprotein α des Phagen P4 ist die Domänenstruktur bereits bekannt, so daß man aufgrund der an P4 α gewonnenen Daten über die Funktionen von RepA in N15 spekulieren kann. P4 α besitzt Primaseaktivität (N-terminaler Bereich, Strack *et al.*, 1992), Helikaseaktivität (gesamtes Protein, Ziegelin *et al.*, 1993) und *ori*-Bindungsaktivität (C-terminaler Bereich, Ziegelin *et al.*, 1997). Für RepA wird ebenfalls zumindest Helikase/Primaseaktivität postuliert (Ravin *et al.*, 2000; Pansegrau & Lanka, 1992) sowie eine *ori*-Bindungsdomäne, da es ein autonomes Replikon darstellt.

Die *origin*-Bindungsdomäne von P4 α liegt im C-terminalen Bereich des Proteins. Dieser Bereich interagiert auch mit dem Cnr-Protein und reguliert so die Kopiezahl von P4 im lysogenen Modus. Durch die Aufklärung der Kristallstruktur dieser Domäne (Yeo *et al.*, 2002) konnte das Motiv für die *origin*-Bindung in P4 α gefunden werden: es handelt sich um zwei α -Helices, die eine flügelartige Struktur (*winged helices*) haben. Die AS-Sequenz von N15 RepA zeigt im Bereich der *ori*-Bindungsdomäne von P4 α zwar keine signifikante Übereinstimmung, es ist aber dennoch möglich, daß *winged helices* auch in RepA vorhanden sind, da es sich um ein strukturelles Motiv handelt, für das nicht unbedingt eine AS-Sequenzübereinstimmung nötig ist.

Zur Replikation eines linearen Plasmids mit kovalent geschlossenen Enden existieren mehrere Modelle (Ravin & Lane, 1999; Casjens, 1999; Kobryn & Chaconas, 2001). Die beiden wahrscheinlichsten Wege wurden von Ravin diskutiert (Ravin *et al.*, 2001) und sind in Abb. 5 gezeigt. Bei beiden möglichen Wegen beginnt die Replikation an einem noch nicht definierten Replikationsorigin (*ori*) im Minireplikon von N15. Nach Beginn der Replikation sind zwei verschiedene Wege vorstellbar. A: Es entsteht intermediär ein zirkuläres Dimer, welches anschließend in zwei lineare Plasmide prozessiert wird. B: sobald die Replikationsblase ein Ende des Plasmids erreicht, wird das entstehende *tellL*' Sequenzstück prozessiert und die Replikation schreitet anschließend in einer klassischen Gabelform bis zum zweiten Ende des Plasmids fort, welches dann ebenfalls so prozessiert werden muß, so daß zwei lineare Plasmide entstehen. Bei der Replikation eines linearen Plasmides ist aber - unabhängig vom benutzten Mechanismus - ein besondere enzymatische Reaktion erforderlich. Diese Reaktion ist die Umwandlung linearer DNA in *hairpin*-Enden. Um diese Reaktion zu bewerkstelligen, wurde eine phagenkodierte, bisher nicht beschriebene Aktivität postuliert, die als *telomere resolution* bezeichnet wird und dem Genprodukt des Gens 29 (TelN) zugeschrieben wurde (Rybchin & Svarchevsky, 1999).

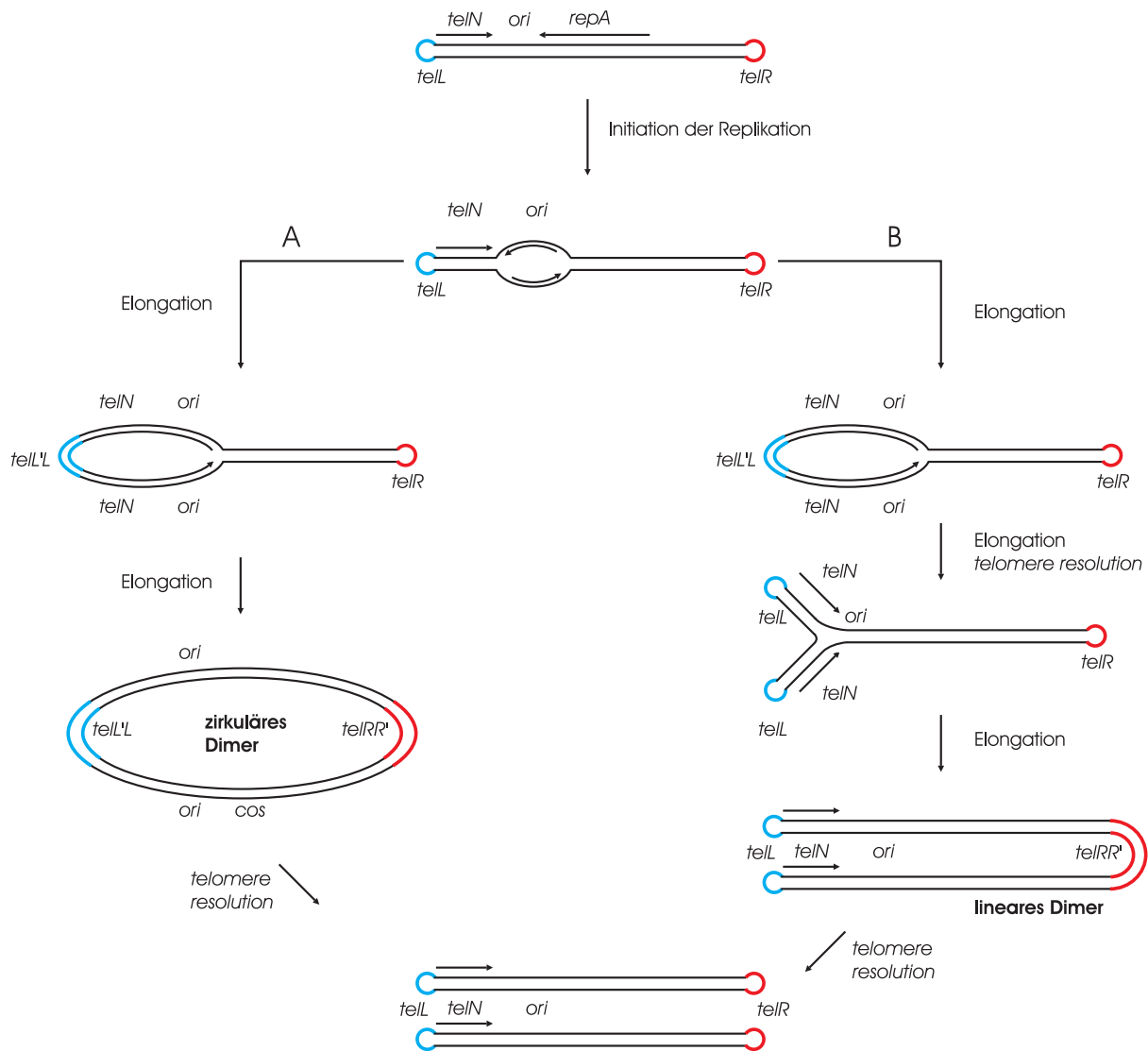


Abb. 5 Modell zur Replikation des linearen Plasmids des Bakteriophagen N15

Schematische Darstellung der Replikation des Prophagen von N15.

1.4 Was versteht man unter *telomere resolution*?

In N15 wurde *in vivo* beobachtet, daß beim Übergang vom lytischen Modus zum Prophagenstadium die vorher zirkuläre DNA sequenzspezifisch geschnitten und die DNA-Enden mit dem Gegenstrang zu *hairpins* verknüpft werden. Dabei gehen weder Basenpaare verloren, noch kommen welche hinzu. Die DNA-Bereiche, an denen dieser Vorgang stattfindet, heißen *telomere resolution sites*. Der Vorgang selbst wurde *telomere resolution* genannt (Abb. 2, Abb. 5 und Abb. 6, Chaconas *et al.*, 2001). Es wurde postuliert, daß *TelN* diese Reaktion katalysiert (Rybchin & Svarchevsky, 1999).

Die Aktivität von TelN wurde Protelomerase genannt. Die Analyse der Aminosäuresequenz legt aber nahe, daß TelN mit Telomerasen aus Eukarionten nichts zu tun hat. TelN funktioniert wahrscheinlich eher wie eine Tyrosinintegrase, da bei der *telomere resolution* die DNA-Sequenz nicht verändert wird (Rybchin & Svarchevsky, 1999).



Abb. 6 Schematische Darstellung der *telomere resolution*

In N15 hat die DNA-Sequenz, aus der die Telomerenden entstehen und die auch das Zielsubstrat für TelN bilden, den Namen *tos* (*telomerase occupancy site*) erhalten (Abb. 7). Diese DNA-Sequenz stellt eine Intergenregion von 310 bp Länge dar und enthält drei invertierte Repetitionen (R1/L1, R2/L2 und R3/L3, in Abb. 7 durch hellgraue Boxen markiert). Es wird von den Genen *telN* (rechte Seite) und *Gen28* (linke Seite) flankiert. *Gen28* ist ein *partitioning gene*, Analog zum *sopA*-Gen des F-Plasmids von *E. coli* (Ravin & Lane, 1999; Lobočka *et al.*, 1996) und sorgt dafür, daß jede *E. coli* Zelle eine Kopie des in geringer Kopienzahl vorliegenden Prophagen hat. Seine Transkriptionsrichtung ist der von *telN* entgegengesetzt.

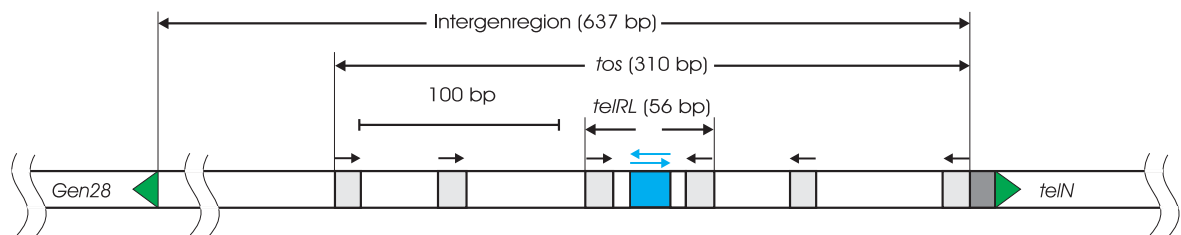


Abb. 7 Die *telomerase occupancy site* (*tos*) und *telN* bilden eine genetische Einheit

Die hellgrauen Boxen stellen *inverted repeats* dar. Die blaue Box markiert das innere Palindrom der *telomere resolution site*, die dunkelgraue Box bezeichnet den Promotor und die Shine-Dalgarno-Sequenz für *telN*. Der Balken in der *tos*-Region entspricht 100 bp. Das *telN* Gen ist nicht vollständig gezeigt, seine Länge beträgt 1,9 kb (Deneke *et al.*, 2000; Ravin *et al.*, 2000).

Es fällt auf, daß *telN* und *tos* auf der DNA von N15 direkt nebeneinander liegen. Sie scheinen eine funktionelle Einheit zu bilden. Zwischen *Gen28* und *tos* (in Abb. 7 links) befindet sich eine 327 bp lange Region ohne erkennbare Funktion.

Es ist von Vorteil, daß sich die genetische Einheit *tos/telRL* in dem sehr gut verstandenen Organismus *E. coli* untersuchen läßt. Aus diesem Grund bietet sich das System zur experimentellen Überprüfung der Hypothesen über die *telomere resolution* an. Der enzymatische Mechanismus der *telomere resolution* ist Thema dieser Doktorarbeit.

I.5 Aufgabenstellung

Ein Ziel dieser Doktorarbeit war es, zum Verständnis der Replikation linearer Genome mit kovalent geschlossenen Enden beizutragen. Es sollte ein Teilaspekt der Replikation, nämlich die *telomere resolution* Reaktion, aufgeklärt werden. Dazu war wegen seiner problemlosen Handhabbarkeit in *E. coli* der Bakteriophage N15 als Modellsystem vorgesehen. Für die *telomere resolution* Reaktion wurde die Aktivität eines bisher nicht bekannten Enzyms postuliert (Rybchin & Svarchevsky, 1999). Um dieses Enzym und seine Reaktionen zu identifizieren, war ein *in vitro*-System aus gereinigten Komponenten zu entwickeln. Im einzelnen war dazu (i) die Expressionsklonierung des *telN* Gens erforderlich, (ii) die Reinigung des Proteins und (iii) die Herstellung geeigneter Substrate. Die Aufgabe bestand darin, durch Mutagenesestudien und biochemische Analysen die Basis zu schaffen, die Reaktion in Teilschritte zu zerlegen. DNA-Bindungsstudien - unter anderem mittels Elektronenmikroskopie - sollten zur Eingrenzung der *telomere resolution site* durchgeführt werden. Die Bewertung der Reaktionsstöchiometrie sollte durch Einsatz der *Surface Plasmon Resonance* Methode erfolgen. Die *nearest neighbor* Analyse sollte zur Bestimmung der von TelN-spezifischen „Schnittstelle“ dienen. Der Vergleich von Reaktionen im N15-System mit denen verwandter Systeme wird diskutiert.