



Doctoral Thesis

"Linker insertion mutagenesis" auf der Basis der Insertionssequenz IS21

Author(s):

Seits, Thomas

Publication Date:

1996

Permanent Link:

<https://doi.org/10.3929/ethz-a-001754131> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

Diss. ETH Nr. 11807

**«Linker insertion mutagenesis» auf der Basis der
Insertionssequenz IS21**



Abhandlung
zur Erlangung des Titels
Doktor der Naturwissenschaften
der Eidgenössischen Technischen Hochschule
Zürich

vorgelegt von
THOMAS SEITZ
dipl. Natw. ETH
geboren am 29. Januar 1967
von Berneck (SG)

Angenommen auf Antrag von
Prof. Dr. Th. Koller, Referent
Prof. Dr. D. Haas, Korreferent

Lausanne 1996

1. KURZFASSUNG

R68 ist ein Plasmid aus *Pseudomonas aeruginosa* mit weitem Wirtsspektrum und trägt eine Insertionssequenz, IS21. Die spontane Bildung einer Tandemduplikation von IS21 auf R68 ergibt das Derivat R68.45. Dank dieser IS21-Tandemduplikation kann R48.45 in die Chromosomen vieler Gram-negativer Bakterien sehr effizient integriert werden, und kann dadurch chromosomale Gene mobilisieren.

IS21 ist 2131 bp gross und weist zwei offene Leseraster auf, *istA* und *istB*, die als Operon organisiert sind. An den Enden enthält IS21 "inverted repeats" von 11 bp. Das Gen *istB* codiert für das Genprodukt P30 (30 kDa). Das Gen *istA* spezifiziert im selben Leseraster zwei Genprodukte P45 (45 kDa) und P46 (46 kDa), die sich in ihren N-Termini um 8 Aminosäuren unterscheiden. Plasmide mit einer IS21-Tandemduplikation bilden mit einer hohen Frequenz Cointegrate mit anderen Replicons, wahrscheinlich nach einem "cut-and-paste"-Mechanismus. In den Cointegraten sind die beiden Replicons durch einzelne IS21-Elemente miteinander verknüpft, die in derselben Orientierung angeordnet sind. Einzelne IS21-Elemente können ebenfalls transponieren. Nach der Transposition bzw. Cointegration sind die IS21-Elemente von Zielduplikationen von 4 bp (selten 5 bp) flankiert.

Die IS21-Gene eines IS21-Elementes konnten nach einer Deletion der "inverted repeats" mit einem vorgeschalteten *tac*-Promoter in *Escherichia coli* überexprimiert werden. Die Transpositionsfunktionen konnten so für die Bestimmung der Cointegratbildungsfrequenz und der Transpositionsfrequenz *in trans* angeboten werden. Sie wirkten auf ein einzelnes IS21-Element, das die intakten IS21-Enden und ein eingesetztes Kanamycin-Resistenzgen trug (IS21-Km) oder auf ein Plasmid mit einer kurzen IS21-IS21-Verbindungsregion (VBR). Die VBR genügte in diesem Experiment für die Cointegratbildung und konnte die IS21-Tandemduplikation ersetzen.

Die separate Expression von P45 und P46 ermöglichte es, die unterschiedliche Substratspezifität von P45 und P46 zu untersuchen. Gleichzeitig wurde untersucht, ob die Länge der Verbindungssequenz (VBS) in IS21-Tandemduplikationen einen Einfluss auf die Cointegratbildungsfrequenz (CF) hatte. P46 katalysierte die Transposition eines einzelnen IS21-Elementes (IS21-Km) wie auch die Cointegratbildung mit IS21-IS21-VBR-Plasmiden mit Frequenzen von ca. $8 \cdot 10^{-4}$. Die VBS konnte dabei zwischen 2 bp und 1.4 kb variieren. P45 katalysierte die Cointegratbildung von IS21-IS21-VBR-Plasmiden mit kurzen (2-4 bp) VBS mit einer CF von ca. $5 \cdot 10^{-2}$. Mit längeren VBS (9 bp bis 1.4 kb) sank die CF auf ca. $3 \cdot 10^{-6}$. P45 wurde daher als Cointegrase und P46 als Transposase bezeichnet. Sowohl P45 als auch P46 benötigten IstB als Hilfsprotein. Waren die beiden IS21-Enden direkt verknüpft (0 bp-VBS), so war die CF weder mit P45 noch mit P46 messbar ($<10^{-7}$). Die VBS gingen bei der Cointegratbildung jeweils

verloren. Unter optimalen Bedingungen gehört die Cointegrase zu den aktivsten Enzymen in der Gruppe der Transpositions-Enzyme.

Es wurde untersucht, inwiefern sich Punktmutationen in den "inverted repeats" auf die CF auswirken. Symmetrisch angeordnete Mutationen der 5'-CA-3'-Nucleotide unmittelbar an den 3'-Enden der IS21-Elemente in einer VBR hatten erniedrigte CF-Werte zur Folge: Die Mutation von C zu G bewirkte eine leichte Reduktion der CF, während die Mutation von A zu T die Cointegratbildung völlig unterband. Das letzte Nucleotid scheint also sehr wichtig zu sein für den korrekten Schnitt der Transposase und Cointegrase an den 3'-Enden der IS21-Elemente. Mutationen an einigen anderen Stellen der "inverted repeats" zeigten nur sehr schwache Auswirkungen auf die CF.

Auf der Basis von IS21 wurde eine "Linker insertion mutagenesis" *in vivo* entwickelt, die gegenüber bestehenden Methoden zahlreiche Vorteile aufweist. Die Methode dient der Untersuchung von Struktur-Funktionsbeziehungen von Proteinen. Die Methode besteht aus einem "Suicide"-Plasmid mit einem Replikationsursprung des Plasmides R6K (*oriR6K*) und einer IS21-IS21-VBR, die Restriktionsschnittstellen für *SalI* und *BglII* an den Enden der IS21-Elemente trägt, aus einem Zielplasmid mit einer hohen Kopienzahl, das ein zu mutierendes Zielgen trägt, und aus dem *E.coli*-Stamm ECOLIST mit chromosomal integrierten *istA*(P45)- und *istB*-Genen, welche unter der Kontrolle des induzierbaren *tac*-Promoters stehen. Der Stamm ECOLIST wurde nacheinander mit dem Zielplasmid und mit dem "Suicide"-Plasmid transformiert. Die resultierenden Cointegrate, welche das "Suicide"-Plasmid mit dem Zielplasmid bildeten, wurden *in vitro* mit *SalI* resp. *BglII* geschnitten. Bei anschliessender Ligation ging der grösste Teil des "Suicide"-Plasmids verloren und zurück blieb eine Insertion von 12 bp (für *SalI*) resp. 33 bp (für *BglII*). Die im Zielplasmid entstandenen Insertionen verändern den Leseraster nicht, führen jedoch zu Proteinen mit Insertionen von 4 bzw. 11 neuen Aminosäuren. Mit dieser Methode können in einem Zielgen schnell und einfach zahlreiche Linkerinsertionen produziert werden.

Als Zielgene wurden das *arcB*-Gen von *Pseudomonas aeruginosa*, das *rpoD*-Gen und die *hcn*-Gene von *Pseudomonas fluorescens* untersucht. Die Verteilung der Insertionen der "Suicide-Plasmide" war bei den *arcB*- den *rpoD*-Plasmiden fast zufällig. Es gab zwar Regionen mit gehäuften Insertionen, aber keine eigentlichen "hot spots". Möglicherweise wird die Insertionsspezifität durch DNA-Sekundärstrukturen beeinflusst. Die Resultate zeigen, dass transkribierte Regionen kein Hindernis für Insertionen sind.

Ausgehend von der "Linker insertion mutagenesis" auf der Basis von IS21 *in vivo* wurde dieselbe Methode ebenfalls *in vitro* entwickelt. Die Transpositionsproteine werden überproduziert in einem Rohextrakt von *E.coli* angeboten. Die *in vitro* Methode eignet sich nicht nur für die raschere Herstellung von Linkerinsertionen, sondern eröffnet auch die Möglichkeit, künftig die Transpositionsmechanismen mit gereinigten IS21-Transpositionsproteinen zu untersuchen.

1. ABSTRACT

R68 is a plasmid of *Pseudomonas aeruginosa* with a broad host-range. The spontaneous formation of a tandem duplication of *IS21* on plasmid R68 results in the derivative R68.45. This *IS21* tandem duplication enables R68.45 to integrate into the chromosomes and to mobilize these in many different species of Gram-negative bacteria.

IS21 contains 2131 bp and two open reading frames designated *istA* and *istB*, which are organised as an operon. *IS21* is flanked by inverted repeats of 11 bp. The *istA* gene codes for two gene products P45 (45 kDa) and P46 (46 kDa), and the *istB* gene codes for the gene product P30 (30 kDa). P46 has 8 additional amino acids at the N-terminus, compared to P45. Plasmids containing an *IS21* tandem duplication form cointegrates at high frequencies with other replicons, probably by a "cut-and-paste" mechanism. In the cointegrates, both replicons are connected with each other by single *IS21* elements in the same orientation. In addition, *IS21* also transposes as a single copy. The integrations are flanked by target duplications of 4 bp (rarely 5 bp).

It was possible to overexpress the *IS21* genes with a *tac* promoter after deletion of the inverted repeats of a single *IS21* element in *Escherichia coli*. The transpositional functions were provided *in trans* for the determination of the cointegration frequency (CF) and the transposition frequency. The IstA and IstB proteins could act on a single *IS21*-element, which carried the intact *IS21*-ends and a kanamycin resistance cassette (*IS21*-Km), or on plasmids containing a short *IS21*-*IS21*-junction. In these experiments, the junction region could replace the *IS21* tandem.

The separate expression of P45 and P46 allowed us to investigate the different substrate specificities of these proteins. At the same time, it was investigated how the CF was influenced by the length of the junction sequence on an *IS21*-*IS21*-junction plasmid. P46 catalysed the transposition of a single *IS21*-element (*IS21*-Km) and the cointegrate formation of *IS21*-*IS21*-junction plasmids having 2 bp to 1.4 kb junctions with frequencies of about $8 \cdot 10^{-4}$. P45 catalysed the cointegrate formation of *IS21*-*IS21*-junction plasmids carrying short (2 to 4 bp) junctions with a frequency of about $5 \cdot 10^{-2}$. Longer junctions (9 bp to 1.4 kb) resulted in much lower CF of about $3 \cdot 10^{-6}$. P45 was therefore termed as cointegrase and P46 as transposase. Both P45 and P46 required the IstB helper protein. When the *IS21* ends were directly connected (0 bp junction) the frequency was not detectable ($<10^{-7}$) with P45 and P46. The junction nucleotides were lost during cointegrate formation. Under optimal conditions, cointegrase was one of the most active enzymes encoded by prokaryotic transposable elements.

It was investigated, if point mutations in the inverted repeats had an effect on the CF. Symmetrical point mutations placed at the terminal 5'-CA-3' nucleotides of the 3'-ends of *IS21* elements in a *IS21*-*IS21*-junction lowered the CF: Mutational C to G changes

slightly reduced the CF and A to T changes totally abolished cointegrate formation, confirming the essential function of the terminal A for the correct cutting at the 3'-ends of the IS21 by transposase or cointegrase. Mutations at some other positions in the inverted repeats changed the CF only slightly.

On the basis of IS21 a linker insertion mutagenesis *in vivo* was developed. This method has some advantages over existing methods and serves for the investigation of structure-function relationships in proteins. The method consists of a suicide plasmid carrying an *oriR6K* and an IS21-IS21-junction (with *SalI* und *BglIII* restriction sites in both IS21 ends), a high copy number target plasmid containing a gene to be mutated, and the *Escherichia coli* strain ECOLIST containing the chromosomally integrated *istA*(P45) and *istB* genes under the control of the inducible *tac* promoter. The strain ECOLIST was successively transformed with the target plasmid and the suicide plasmid. The cointegrates formed from the suicide plasmid and the target plasmid were cut *in vitro* with *SalI* or *BglIII*. After ligation most of the suicide plasmid was excised and insertions of 12 bp (for *SalI*) or 33 bp (*BglIII*) were created. The insertions do not change the reading frame but result in insertions of 4 or 11 amino acids, respectively. With this method it is possible to produce linker insertions within a target gene in a fast and simple way.

As target genes the *arcB* gene of *Pseudomonas aeruginosa*, the *rpoD* gene and the *hcn* genes of *Pseudomonas fluorescens* were investigated. The distribution of the inserted suicide plasmids was nearly random in the *arcB* and *rpoD* recombinant plasmids. There were a few regions with clustered insertions, but no real "hot spots". It is possible that DNA secondary structure may influence the insertion specificity. Transcribed regions did not prevent insertions.

Starting from IS21-dependent linker insertion mutagenesis *in vivo* a similar method also was developed *in vitro*. The transposition proteins were overproduced and provided in a crude extract from *E.coli*. The *in vitro* method was not only suitable for the production of linker insertions, but may also become useful for the investigation of the transposition mechanism with purified IS21 transposition proteins.