Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 344 (2005) 221–231
GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES
www.elsevier.com/locate/gene

Dimerization of bacteriophage P2 integrase is not required for binding to its DNA target but for its biological activity

Clara Frumerie, Jesper M. Eriksson[1], Marc Dugast[2], Elisabeth Haggård-Ljungquist*

Department of Genetics, Microbiology and Toxicology, Stockholm University, S-106 91 Stockholm, Sweden

Received 21 June 2004; received in revised form 30 August 2004; accepted 28 September 2004
Available online 25 November 2004
Received by D.L. Court

Abstract

Coliphage P2 integrates into the host chromosome upon lysogenization via site-specific recombination mediated by the phage integrase (Int). P2 integrase belongs to the tyrosine family of recombinases. In this work, it is shown that P2 integrase forms dimers but not oligomers in the absence of its DNA target. Furthermore, the C-terminal end of the protein and amino acid (aa) E197 have been found to be involved in dimerization. Amino acid E197 is located in a conserved region of the tyrosine recombinases that has not previously been implicated in dimerization. The dimerization deficient mutants were unaffected in binding to its phage attachment site (*attP*) substrate, but had a reduced ability to complement an *int*-defective prophage.

Keywords: *Escherichia coli*; Bacteriophage; Site-specific recombination; Integrase; Protein–protein interaction; DNA binding

1. Introduction

Bacteriophage P2 is a temperate phage that integrates into the host chromosome upon lysogenization. P2 integration occurs by conservative site-specific recombination between the phage attachment site (*attP*) and the bacterial attachment site (*attB*) generating the host/phage junctions *attL* and *attR*. The integration process requires the phage-encoded integrase (Int), and the host-encoded integration host factor (IHF) (Yu and Haggård-Ljungquist, 1993a). Int is responsible for cleavage and joining of the DNA backbone during recombination, while IHF is an architectural protein that binds and bends DNA. The excision event requires a third protein, Cox (control of excision) (Yu and Haggård-Ljungquist, 1993b).

P2 *attP* is about 210 nt long and it contains two Int arm-binding sites, P and P′, and the core-binding site. Each arm-binding site has two direct repeats, in contrast to the core-binding site that contains a poor inverted repeat. The DNA sequence of the repeats in the arms differs from those of the core, indicating that two different parts of the Int protein are involved in core and arm binding. The *attB* region is only about 30 nt long and contains only the core-binding site, which has 27 nt identity to the core sequence in *attP*. The IHF binding site is located between the P arm and the core sequence. The Cox-binding region, which contains an inverted repeat where each repeat contains three direct repeats, is located between the core and the P′ arm.

P2 integrase is a member of the Int, or “tyrosine”, family of recombinases and consists of 337 amino acid (aa) residues (37.9 kDa) (Yu et al., 1989). This family includes over 100 members that show very little sequence identity; only four amino acids are conserved in all

Abbreviations: Int, integrase; Cox, control of excision; IHF, integration host factor; PCR, polymerase chain reaction; PAA, polyacrylamide; SDS, sodium dodecyl sulphate; DTT, dithiothreitol; aa, amino acids; *attP*, phage attachment site; *attB*, bacterial attachment site; CFU, colony forming units; PFU, plaque forming units.

* Corresponding author. Tel.: +46 8 161270; fax: +46 8 164315.
E-mail addresses: Jesper.Eriksson@biovitrum.com (J.M. Eriksson), marc.dugast@curie.fr (M. Dugast), Elisabeth.Haggard@gmt.su.se (E. Haggård-Ljungquist).
[1] Present address. Biovitrum AB, S-112 76 Stockholm, Sweden.
[2] Present address. Inst. Curie, Inserm U520, F-75005 Paris, France.

0378-1119/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2004.09.026

proteins. However, two conserved boxes and three patches of charged amino acids have been identified (Nunes-Düby et al., 1998). Box I contains the invariant Arg residue and Box II contains the conserved His-Xxx-Xxx-Arg motif and the active site tyrosine. The DNA binding and catalytic properties of λ Int, the prototype of the Int family of recombinases, in conjunction with its accessory factors have been studied extensively (Nash, 1996; Landy, 1989; Azaro and Landy, 2002). The catalytic site is located at the C-terminal half of the protein, while the N-terminus contains motifs responsible for binding to the arm sites and to accessory proteins. The structures of the catalytic domains of λ integrase, and the integrase of the *Haemophilus influenzae* phage HP1 have been determined (Kwon et al., 1997; Hickman et al., 1997). In addition, the structures of full-length Cre and XerD recombinases have been determined (Guo et al., 1997; Subramanya et al., 1997), but these two members of the Int family differ from λ and HPI in that they require no accessory proteins and, therefore, do not have the same directionality. According to the study of the HP1 integrase, dimers are formed by mutual interactions of the C-terminal tails of each monomer, where the tail of one monomer reaches across to nestle into a hydrophobic cleft on the opposite monomer, which orients the active-site clefts antiparallel to each other (Hickman et al., 1997). However, the biological relevance of this dimerization has not yet been demonstrated. Secondary structure predictions and a 3D-prediction of the P2 integrase suggest high similarity of the C-terminal ends of P2 Int and HP1 Int.

In site-specific recombination, the formation of ordered nucleoprotein complexes, intasomes, are of vital importance. A network of protein–protein and protein–DNA interactions stabilizes these structures. Because the structure of the C-terminus of λ and HP1 integrase differ, and the N-terminus of λ integrase is involved in dimerization (Jessop et al., 2000), we found it interesting to clarify the importance of the C-terminus of phage P2, the best characterized member of the large non-lambdoid family of phages infecting different enterobacteria to which also HP1 belongs. In this work, we show that P2 Int forms dimers and that C-terminal truncations of the protein and the Int mutation E197A, situated within the conserved Box I, abolishes dimer formation. The dimerization of the protein is not required for normal binding to *attP*, but for biological activity.

2. Materials and methods

2.1. Biological materials

Bacterial strains and plasmids used are listed in Table 1. The P2 *int* mutants were isolated after *N*-methyl-*N*-nitro-*N*-nitroso-guanidine treatment as previously described (Choe, 1969) and were a generous gift of Dr. L. E. Bertani.

2.2. Plasmid constructions

Plasmids were constructed using standard techniques (Sambrook et al., 1989). The genes were amplified by polymerase chain reaction (PCR) using Vent Polymerase (New England Biolabs), and primers from DNA-Technology, Denmark. The inserts and induced mutations were checked by automatic DNA sequencing on an ALFexpress II using the ThermoSequenase Kit (Pharmacia-Amersham).

2.2.1. pEE841

The *int* gene was amplified by PCR using primers-27intSalI (GGC GTC GAC GTA TGA AGT GGA CAT CCG) and intBamHI (GCG GAT CCG GGT GTC ACC AGT AGG GG), and, after digestion with *Sal*I and *Bam*HI, the gene was inserted between the *Sal*I and *Bam*HI sites of plasmid pJH391. The *int* gene lacks the first 30 bases, which is not expected to interfere with the putative dimerization domain (Yu and Haggård-Ljungquist, 1993a).

2.2.2. pEE849, pEE850, and pEE863

The plasmids were generated by site-directed mutagenesis using the Quickchange site-directed Mutagenesis Kit (Stratagene), with pEE841 as a template. Primers used were Int-ΔC17 top (CTC TCG GCC TCA GTA CCA CCT TAT TAC GGA TTA AGA GAA ATG GCG), Int-ΔC17 bot (CGC CAT TTC TCT TAA TCC GTA ATA AGG TGG TAC TGA GGC CGA GAG), Int-ΔC25 top (CGG ATT AAG AGA AAT GGC TTA TTA AAG GTA CTC TGG CGC AAA ATG), Int-ΔC25 bot (CAT TTT GCG CCA GAG TAC CTT TAA TAA GCC ATT TCT CTT AAT CCG), E197A top (GCT TTC AAA CGA GCT GCT GCT CCC CAA CGT GCT CC), and E197A bot (GGA GCA CGT TGG GGA GCA GCA GCT CGT TTG AAA GC).

2.2.3. pEE2008

The *int* gene was amplified by PCR using primers—9int (AAA AAA CTC GAT GAT GGT CGA TA) and 72.0R (TGG TCA ATG TGT GGA CTG GA), and inserted into the filled in *Nco*I site of plasmid pET8c. This generates an Int protein lacking the second and third amino acid.

2.2.4. pEE2016, pEE2017, and pEE2020

The plasmids were generated by site-directed mutagenesis, with pEE2008 as a template and the same primers as described for generation of pEE849, pEE850, and pEE863.

2.3. Dimerization/oligomerization assays

Dimerization was analyzed by measuring β-galactosidase activity in JH372 cells transformed with the respective plasmid. Strain JH372 contains the reporter λ O_R-P_R-*lacZ* fusion on the λ 202 prophage. Oligomeriza-