Site-Specific Recombination in Human Cells Catalyzed by the Wild-Type Integrase Protein of Coliphage HK022

Mikhail Kolot, Adi Meroz, Ezra Yagil

Department of Biochemistry, Tel-Aviv University, Tel-Aviv 69978, Israel; telephone +972-3-640-9823; fax: +972-3-640-6834; e-mail: yagill@post.tau.ac.il

Received 16 January 2003; accepted 24 April 2003

Published online 23 June 2003 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/bit.10747

Abstract: The activity of the Integrase (Int) protein encoded by coliphage HK022 was tested in a human cell culture. Plasmids were constructed as substrates that carry the sites of the integration reaction (*attP* and *attB*) or the sites of excision (*attL* and *attR*). The site-specific recombination reactions were monitored in *cis* and in *trans* configurations by the expression of the green fluorescent protein (GFP) as a reporter. Cells cotransfected with the substrate plasmid(s) and with a plasmid that expresses the wild-type Int show efficient integration as well as excision in both configurations. The wild-type Int was active in the human cells without the need to supply the accessory proteins integration host factor (IHF) and excisionase (Xis) that are indispensable for the reaction in the bacterial host. © 2003 Wiley Periodicals, Inc. *Biotechnol Bioeng* **84**: 56–60, 2003.

Keywords: integrase; site-specific recombination; phage HK022; green fluorescent protein; human cells

INTRODUCTION

In its lysogenic cycle the mechanism of site-specific recombination of the coliphage HK022 is very similar to that of phage λ. The integration of the phage into the chromosome of its host (*Escherichia coli*) results from a site-specific recombination reaction between the *attP* site on the phage DNA and an *attB* site on the host DNA. The integrated prophage is flanked by the recombinant *attL* and *attR* sites, which are the sites for phage excision. Both integration and excision are catalyzed by the phage-encoded integrase (Int) protein and both reactions require the presence of the host-encoded accessory protein integration host factor (IHF). In addition, excision requires the phage-encoded accessory protein excisionase (Xis), whose activity can be partially replaced by the host-encoded FIS (factor for inversion stimulation) protein. The central 7 basepairs (bp) of the 21 bp-long *attB* (BOB′) are the site of DNA exchange and are known as the overlap (O). O is flanked by two 7 bp inverted repeats (B and B′) that serve as weak binding sites for Int. *attP* is composed of a similar core (COC′) flanked by two longer arms (P and P′) that carry tight binding sites for Int and for the accessory proteins (IHF, Xis, FIS). As a bivalent protein Int monomers bind tightly via the amino-terminal arm binding domain to the arms of *attP*, the bound accessory proteins bend the arms and thereby facilitate the binding of a carboxyl terminal domain of Int to the core, the site of the reaction. The resulting protein-DNA complex, known as the intasome, captures the *attB* site to execute the recombination. The recombinant *attL* (BOP′) and *attR* (POB′) sites of the integrated phage are each composed of a recombined core and one of the arms. The λ and HK022 systems are unidirectional; their integration and excision are not completely reversible reactions because their recombining *att* sites are not identical and because excision requires Xis as an additional accessory protein (reviewed by Azaro and Landy, 2001; Nash, 1996). Int of λ and HK022 are 70% identical, and each recognizes a different set of *att* sites. Xis is functionally identical in both phage (Weisberg et al., 1999).

Int-family site-specific recombinases of lower organisms have been recruited to operate in mammalian cells and in plants in order to facilitate the integration and excision of genes at specific sites. Among the best-developed systems are Cre-*lox* of bacteriophage P1 and FLP-*frt* of yeast. However, in these systems the recombination target sites (*lox* and *frt)* are identical, favoring excision over stable integration (Baer and Bode, 2001; Ow, 2002; Nagy, 2000; Farley et al., 2000; Sauer, 1998). Recently, the site-specific recombination system of the *Streptomyces* phage ϕC31 has been successfully developed to stably insert DNA at specific mammalian chromosomal sites. This system, which belongs to the resolvase family, is unidirectional because it requires its Xis for excision (Thyagarajan et al., 2001; Belteki et al., 2003).

Wild-type Int of λ expressed in mammalian cells is inactive. In contrast, IHF-independent mutants of Int-λ can promote site-specific integration and excision in mammalian cells (Christ et al., 2002; Christ and Dröge, 2002). We

Correspondence to: E. Yagil

Contract grant sponsors: Tel-Aviv University and by the Israel Science Foundation

Contract grant number: 637/02

© 2003 Wiley Periodicals, Inc.

have previously cloned and expressed the wild-type Int of HK022 and have shown, mainly by qualitative PCR reactions, that in the presence of plasmids that carry the proper *att* sites, integration and excision products can be detected (Kolot et al., 1999). In the present work, we developed the system for a quantitative assay for integration and for excision and show that, in contrast to the wild-type Int-λ, that of Int-HK022 is strongly active in mammalian cells.

MATERIALS AND METHODS

Cell Line, Culture Conditions, Transfection, and Immunoblot

Cells (~10^5) of the human embryonic kidney cell line 293T were transfected with 10 μg DNA of each plasmid using the calcium phosphate method as described previously (Kolot et al., 1999) in a 6-well culture plate. Immunoblots were performed using polyclonal Int antiserum as described (Kolot and Yagil, 2003).

Plasmid Construction

In all plasmids the "stop" site is a fragment that carries a transcription terminator that has been extracted from plasmid pBS302 (Sauer, 1993). *attP* originated from plasmid pKY110 (Kolot et al., 1996), *attB* from plasmid pNR185 (Nagaraja and Weisberg, 1990), *attR* from plasmid pMK24, and *attL* from plasmid pMK25 (Gottfried et al., 2000). In plasmids pMK218 and pMK189 the relevant fragments were cloned between the pCMV promoter and the GFP reading frame of plasmid pEGFP-N1 (ClonTech, Palo Alto, CA). In pMK218 they were cloned between the *EcoRI-BamHI* sites and in pMK189 between the *SmaI* and *EcoRI* sites. In plasmids pMK221 and pAM243 the *att* sites were cloned downstream of the pCMV promoter of plasmid pCDNA3 (Invitrogen, Carlsbad, CA). In pMK221 *attP* was cloned between the *EcoRI* and *HindIII* sites and in pAM243 between *KpnI* and *EcoRV*.

To construct plasmids pMK223 and pAM242, the CMV promoter was first removed from plasmid pEGFP-N1 by deleting its small *AseI-BglII* fragment, obtaining plasmid pMK219. In pMK223 the "stop" and *attB* fragments were cloned between the *HindIII* and *BamHI* sites of pMK219 and in pMK242 the "stop" and *attL* fragments were cloned between *HindIII* and *SalI* of pMK219.

To enable proper cloning, several of the cloned fragments were first subcloned into plasmids that carried multiple cloning sites and were cut out with the proper ends.

Fluorescent-Activated Cell Sorting (FACS) Analysis

Cells from one well were collected for FACS measurements that were done on a FACSort Becton Dickinson Instrument. Data analysis was performed using the WinMDI2.8 program.

RESULTS

Plasmids that carry the reporter gene encoding the GFP of the jellyfish *Aequorea victoria* (Prasher et al., 1992) were constructed such that the protein can express only as a result of an Int-catalyzed site-specific recombination reaction. Different constructs could monitor either the integration reaction (*attP* × *attB*) or the excision reaction (*attL* × *attR*). Each reaction was monitored in one of the following two configurations: In the *cis* configuration both recombining *att* sites are located on the same vector (Fig. 1A,C), and in the *trans* configuration each *att* site is located on a different vector (Fig. 1B,D). In the reactions in *cis* the human cytomegalovirus (CMV) promoter and the reading frame of the GFP were separated by a fragment carrying two *att* sites that flank a transcription terminator site. For the integration reaction in *cis* the two *att* sites were *attP* and *attB* (plasmid pMK218, Fig. 1A) and for the excision reaction in *cis* the sites were *attL* and *attR* (plasmid pMK189, Fig. 1C). Recombination between the two *att* sites evicts the transcription terminator and brings the CMV promoter proximal to the reading frame of the GFP gene allowing its expression. In the reactions in *trans* (Fig. 1B,D), one plasmid carried an *att* site downstream to the CMV promoter and the other plasmid carried an *att* site and the transcription terminator upstream to the promoterless GFP reading frame. In the integration reaction in *trans, attP* is on the first plasmid (pMK221) and *attB* is on the second one (pMK223). In the excision reaction in *trans, attR* is on the first plasmid (pAM243) and *attL* is on the second one (pAM242). Int-catalysis of these reactions recombines between the two *att* sites, resulting in a cointegrate that carries the CMV promoter upstream to the GFP gene, allowing its expression.

Human cells (line 293T) were transfected with the relevant substrate plasmid(s) together or without an Int-expressing plasmid (pMK52) (Kolot et al., 1999) and 48 h following the transfection the cells underwent a FACS analysis. Dot plots that depict the fluorescence of the transfected cells in the presence and the absence of Int are shown in the inserts in Figure 1 and the quantitative results are plotted in Figure 2. In the untransformed control cells the fluorescence above the baseline was negligible (0.1%) and in cells transfected with a plasmid that expresses GFP (pEGFP-C3, ClonTech) 84% of them were fluorescent, indicating that this was the efficiency of transfection (not shown). Cells transfected with the substrate(s) alone, in the absence of Int, showed a fluorescence of 4.5 and 2.4% in the *cis* reactions and 2.1 and 1.2% in the *trans* reactions (Fig. 1, gray columns in Fig. 2). In contrast, in the presence of Int the percent fluorescence in the *cis* configuration were 30.6 in the integration reaction (Figs. 1A, 2) and 42.2% in the excision reaction (Figs. 1B, 2). The fluorescence in the *trans* configurations in the presence of Int was lower, 13.4% in the integration reaction (Figs. 1B, 2) and 15.6% in the excision reaction (Figs. 1D, 2). Because the increase in fluorescence is dependent on the presence of Int, its intensity reflects the rate of recombination. Thus, the results clearly

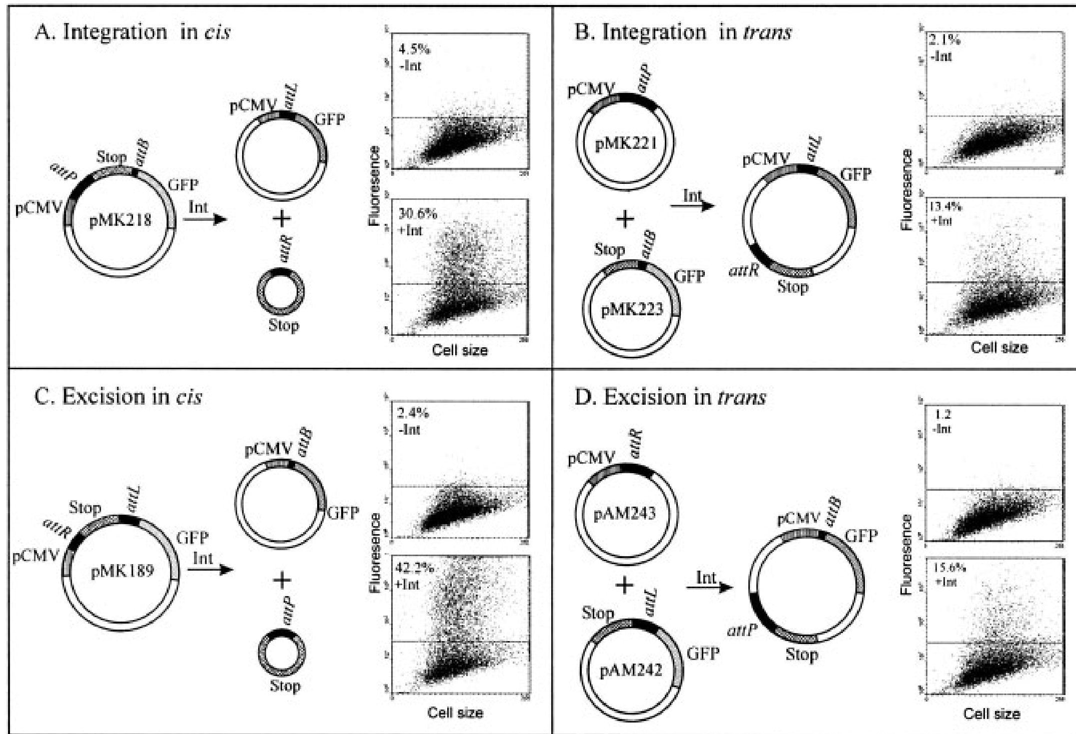

Figure 1. The structure of the plasmids used as substrates (numbered plasmids) and the expected products in the Int-catalyzed site-specific recombination reactions. Stop depicts the transcription termination site. The inserts show dot plots of the FACS analyses in each reaction in the absence (top) and presence (bottom) of the Int-expressing plasmid.

show that the wild-type Int protein of HK022 can catalyze the site-specific recombination reactions in the human cells. The basal-level fluorescence observed in the absence of Int could have resulted from some leakage, despite the presence of the transcription terminator upstream to the gene encoding the GFP. Substrates that were tested in the absence of the transcription terminator showed a much stronger fluorescence (data not shown), indicating that the presence of the terminator in these experiments was essential.

The results show that 48 h following transfection the recombination reactions in the *cis* configuration were more efficient than those in *trans*. To test the time at which each type of reaction is at its maximum, we performed a kinetic analysis of the integration reactions in *cis* and in *trans* (Fig. 3). The results confirmed that the reaction in *cis* was faster and more efficient, reaching a peak of 34% fluorescence at about 48 h following transfection. In contrast, the reaction in *trans* was slower and reached a peak of over 20% fluorescence only at 72 h after transfection. In parallel, we also monitored the expression of Int in the transfected cells. Figure 4 shows that the protein already expressed at 24 h following transfection (column C) at which time the fluorescence was still negligible. Its expression reached a maximum at 72 h (column E) and declined thereafter.

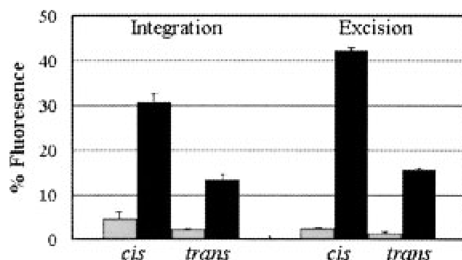

Figure 2. Fractions of fluorescent cells in the integration and excision reactions, each in the *cis* and the *trans* configurations. Gray columns depict the controls in the absence of Int. The data are an average of three experiments.

DISCUSSION

In a previous study we have shown that Int-HK022 can perform site-specific recombination in mammalian cells. However, in these studies we have only demonstrated the

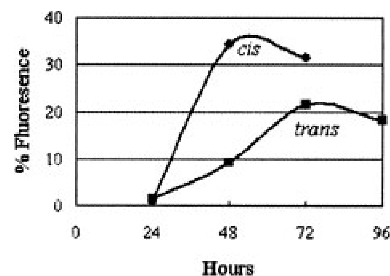

Figure 3. A kinetic analysis of the integration reaction in *cis* and in *trans*.

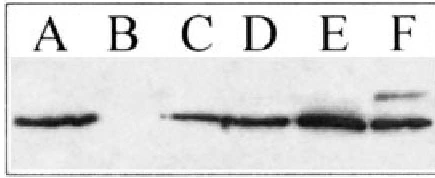

Figure 4. An immunoblot showing the expression of Int in cells transfected with the Int-expressing plasmid pMK52. Lane A, purified Int; Lane B, untransfected cells; Lanes C–F, cells 24, 48, 72, and 96 h, respectively, following transfection.

products by PCR analyses, which did not reveal the extent of the reactions. The only reaction, which was measured quantitatively, was an integration reaction in the *cis* configuration whose products showed only 3.5% recombination. In this experiment we used β-galactosidase as a reporter, which was assayed in a bacterial host (Kolot et al., 1999). In the experiments reported here we used the GFP reporter, which could be monitored directly in the mammalian cells. The results show that the wild-type Int of phage HK022 can promote site-specific recombination in human cells at least as efficient as the IHF-independent mutants of Int-λ (Christ et al., 2002).

The reactions in *trans*, which proved to be slower (Fig. 3) and less efficient (Fig. 2) than those in *cis*, reached a maximum at 72 h, at the time when the expression of Int was maximal (Fig. 4, lane E). In contrast, the *cis* reaction reached a maximum prior to the maximal Int expression. This could be for the following reasons: First, the *att* products of the *trans* reactions are arranged in *cis* (Fig. 1B,D), thereby enhancing the reverse reaction that seems to be thermodynamically favored. That is probably because in *trans* the recombining sites are distal apart, located on two different molecules, whereas in *cis* they reside on the same molecule. Second, the *trans* reactions require the presence of three cotransfected plasmids (two substrates and Int), whereas the *cis* reactions require only two plasmids. The efficiency of transfection of a single plasmid was about 84%, as judged from the transfection of the GFP-expressing plasmid. It is therefore reasonable to assume that the frequency of cells that absorbed two plasmids (as in the case of the reaction in *cis*) is higher than those that absorbed three plasmids (in the *trans* reactions).

In the mammalian milieu both the integration and excision reactions catalyzed by the λ and HK022 systems proved to be efficient without the need to supply the accessory proteins that are required in the bacterial and in the in vitro reactions. In vitro experiments have demonstrated that the eukaryotic chromatin-associated high mobility group DNA-bending proteins HMG1 and HMG2 can replace IHF in the excisive recombination reactions of λ (Segall et al., 1994). Perhaps the difference between the wild-type systems of λ and HK022 is because HMG or other eukaryal proteins that replace the bacterial DNA-bending proteins bind more effectively to the *att* sites of HK022. The fact that in both systems integration and excision occur rather frequently diminishes the advantage at that stage of the unidirectional nature of these lambdoid systems for stable gene integrations. In the phage ϕC31 system, which is also unidirectional, stable chromosomal *attP* × *attB* insertions have been accomplished (Thyagarajan et al., 2001; Belteki et al., 2003). However, to the best of our knowledge the ability of the ϕC31 system to perform the excisive *attL* × *attR* reaction in the mammalian milieu has not been reported.

Finally, we recently demonstrated that Int-HK022 accumulates in the mammalian nucleus (Kolot and Yagil, 2003). Work is in progress to test its potential for stable site-specific chromosomal manipulations, as are the IHF-independent mutants of int-λ (Christ and Dröge, 2002).

Brian Sauer kindly sent us plasmid pBS302, Itzchak Oschri helped us with the FACS analyses, and Pnina Gottfried and anonymous reviewers helped improve the manuscript.

References

Azaro MA, Landy A. 2001. Integrase and the λ int family. In: Craig NL, Craigie R, Gellert M, Lambowitz A, editors. Mobile DNAII. Washington DC: ASM Press. p 118–148.

Baer A, Bode J. 2001. Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. Curr Opin Biotechnol 12:473–480.

Belteki G, Gertsenstein M, Ow DW, Nagy A. 2003. Site-specific cassette exchange and germline transmission with mouse ES cells expressing phiC31 integrase. Nat Biotechnol 21:321–324.

Christ N, Dröge P. 2002. Genetic manipulation of mouse embryonic stem cells by mutant lambda integrase. Genesis 32:203–208.

Christ N, Corona T, Dröge P. 2002. Site-specific recombination in eukaryotic cells mediated by mutant lambda integrases: implications for synaptic complex formation and the reactivity of episomal DNA segments. J Mol Biol 319:305–314.

Farley FW, Soriano P, Steffen LS, Dymecki SM. 2000. Widespread recombinase expression using FLPeR (flipper) mice. Genesis 28: 106–110.

Gottfried P, Yagil E, Kolot M. 2000. Core-binding specificity of bacteriophage integrase. Mol Gen Genet 263:619–624.

Kolot M, Yagil E. 2003. Determinants that target the integrase of phage HK022 into the mammalian nucleus. J Mol Biol 325:629–635.

Kolot M, Gottfried P, Yagil E. 1996. A second site-specific recombination event in the lambdoid bacteriophage HK022. Mol Gen Genet 253: 362–369.

Kolot M, Silberstein N, Yagil E. 1999. Site-specific recombination in mammalian cells expressing the Int recombinase of bacteriophage HK022. Mol Biol Rep 26:207–213.

Nagaraja R, Weisberg RA. 1990. Specificity determinants in the attachment sites of bacteriophages HK022 and lambda. J Bacteriol 172: 6540–6550.

Nagy A. 2000. Cre recombinase: the universal reagent for genome tailoring. genesis 26:99–109.

Nash HA. 1996. Site-specific recombination: integration, excision, resolution, and inversion of defined DNA segments. In: Neidhardt FC, Curtiss R III, Ingraham JL, Lin ECC, Low KB, Magasanik B, Resnikoff WS, Riley M, Schaechter M, Umbarger HE, editors. *Escherichia coli* and *Salmonella*: cellular and molecular biology. Washington DC: ASM Press. p 2363–2376.

Ow DW. 2002. Recombinase-directed plant transformation for the post-genomic era. Plant Mol Biol 48:183–200.

Prasher DC, Eckenrode VK, Ward WW, Prendergast FG, Cormier MJ. 1992. Primary structure of the *Aequorea victoria* green-fluorescent protein. Gene 111:229–233.

Sauer B. 1993. Manipulation of transgenes by site-specific recombination: use of Cre recombinase. Methods Enzymol 225:890–900.

Sauer B. 1998. Inducible gene targeting in mice using the Cre/*lox* system. Methods: Comp Methods Enzymol 14:381–392.

Segall AM, Goodman SD, Nash HA. 1994. Architectural elements in nucleoprotein complexes: interchangeability of specific and non-specific DNA binding proteins. EMBO J 13:4536–4548.

Thyagarajan B, Olivares EC, Hollis RP, Ginsburg DS, Calos MP. 2001. Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. Mol Cell Biol 21:3926–3934.

Weisberg RA, Gottesmann ME, Hendrix RW, Little JW. 1999. Family values in the age of genomics: comparative analyses of temperate bacteriophage HK022. Annu Rev Genet 33:565–602.