Molecular Analysis of the UV Protection and Mutation Genes Carried by the I Incompatibility Group Plasmid TP110

JANICE A. GLAZEBROOK,† KANWAL K. GREWAL, AND PETER STRIKE*

Department of Genetics, University of Liverpool, Liverpool L69 3BX, United Kingdom

Received 17 March 1986/ Accepted 8 July 1986

The imp genes, responsible for the UV protection and mutation effects of the I incompatibility group plasmid TP110, have been cloned into vector plasmids, and their products have been analyzed. The genetic information required for expression of these properties was carried in a continuous DNA sequence of approximately 1.7 kilobases, encoding the production of two proteins with molecular weights of 11,000 and 51,000. The genetic arrangement of this system therefore appears similar but not identical to the functionally related umuDC and mucAB operons. A third protein with a molecular weight of 40,000 was produced from sequences downstream from imp and could be overproduced by high-level transcription through the imp genes. This protein was not required for the protection and mutation properties.

A variety of bacterial plasmids are able to increase the UV resistance and levels of UV-induced mutation in cells which harbor them, including the N incompatibility group (IncN) plasmid pKM101, the IncI plasmid TP110, and the IncB plasmid R16 (5, 13, 22, 23). Each of these plasmids appears to encode functions which act as analogs of the products of the chromosomal umuC and umuD genes, which are essential for UV-induced mutation in Escherichia coli and form an operon under the control of the rec-lex system (1, 10). The products of these genes are proteins of molecular weight (MW) 15,000 and 47,000, respectively (6), and are similar in size to those produced by the analogous genes (mucA and mucB) of pKM101 (11, 14, 15). The pKM101 muc genes are also under rec-lex control but show only limited nucleotide sequence homology to umuDC (14) and do not cross-complement them (23).

We have previously shown that TP110, which is quite unrelated to pKM101, also specifies products which can function as analogs of the umuDC proteins (4) and that the genes responsible (designated imp, for I group mutation and protection) are under rec-lex control (5). The work included in this paper describes the further characterization of the imp genes by subcloning into suitable plasmid vectors and the identification of two proteins encoded by the genes.

MATERIALS AND METHODS

Bacterial strains and plasmids used. The properties and sources of the bacterial strains and plasmids used are shown in Table 1.

Media. Luria broth contained 1% tryptone (Difco Laboratories), 0.5% yeast extract, and 1% NaCl, with 1.5% agar (Difco) for plates. M9 buffer contained (in grams per liter): NH4Cl, 1; Na2HPO4, 11; NaCl, 5; MgSO4·7H2O, 0.246; and CaCl2, 0.011.

DNA preparation and manipulation. Plasmid DNA was prepared by the polyethylene glycol (PEG) precipitation method of Humphries et al. (9), followed by overnight centrifugation of the redissolved PEG precipitate at 50,000 rpm in a cesium chloride-ethidium bromide equilibrium density gradient.

Rapid screening of recombinant clones was performed by the method of Birnboim and Doly (2).

For restriction analysis, DNA was digested in reaction volumes of 10 to 40 μl for 1 to 3 h in buffer conditions appropriate to the enzyme. Fragments were analyzed by agarose gel electrophoresis as described previously (5).

Restriction fragments were purified from 1% low-gelling-temperature agarose gels by phenol extraction of the molten agarose, followed by ether extraction of the aqueous phase and ethanol precipitation of the fragment.

DNA was transformed into calcium-treated cells as described previously (21).

UV survival and mutagenesis. UV survival was determined as described previously (5). UV mutagenesis was performed at lower doses than for UV survival, scoring for reversion of the his-4 marker to His+ Details have been given previously (5).

Analysis of plasmid-encoded proteins. [35S]Methionine-labeled proteins were synthesized from plasmid DNA by using cell extracts of Escherichia coli, as described by Pratt et al. (16). Protein products were analyzed by electrophoresis through 11 or 15% polyacrylamide–sodium dodecyl sulfate gels, as appropriate.

For transcription of templates by T7 RNA polymerase, 1 μg of purified DNA was transcribed in a 50-μl reaction volume containing 40 mM Tris hydrochloride, pH 8.0, 15 mM MgCl2, 5 mM dithiothreitol, 1 mM each ATP, GTP, CTP, and UTP, 25 μg of bovine serum albumin (nuclease free), 12.5 U of T7 RNA polymerase, and 20 U of RNasin RNase inhibitor. The reactions were carried out in diethylpyrocarbonate-treated glassware for 60 min at 37°C and were then added directly to the cell-free preparations described above.

Isolation of Tn1000 insertion mutants. To isolate Tn1000 insertion mutants of the cloned genes, clones in pHSG415 or pLC28 were transformed into the F+ strain AB2414 and mobilized by conjugation into CA222 umuc36 nala. Transconjugants were obtained at a frequency of about 1 in 106 with respect to the number of recipient cells, and all proved to contain Tn1000 inserts when checked by gel electrophoresis. Transconjugants were screened for UV survival and mutation, and those which had lost the Imp+ phenotype were analyzed for the position of the insertion.

* Corresponding author.
† Present address: Department of Genetics, University of Leeds, Leeds LS2 9JT, United Kingdom.
plasmids isolated from the three selected colonies revealed a common region of DNA of approximately 7 kilobases (kb) present in each of the clones, which we presumed to contain the \textit{imp} genes. Transformation of the purified plasmids (designated pJAG253, pJAG256, and pJAG278) back into TK702 gave 100% coinheritance of the plasmid drug-resistance markers and the Imp\textsuperscript* phenotype. The conclusion that we had cloned the \textit{imp} genes was confirmed by the presence in the cloned \textit{PstI} fragments of \textit{EcoRI}, \textit{BglII}, \textit{HindIII}, and \textit{SalI} sites corresponding to the \textit{EcoRI} E, \textit{BglII} J, \textit{HindIII} H, and \textit{SalI} E fragments of TP110, which were known from previous work with \textit{Mu} d\textit{(Ap lac)} to be present in the \textit{imp} region (5). In addition, a previously unidentified \textit{HindIII} fragment was found, mapping adjacent to \textit{HindIII} fragment H of TP110, within \textit{BglII} J. This small 0.6-kb fragment had been missed during previous, less sensitive mapping studies (17), and we designated it TP110 \textit{HindIII} fragment K. The order of these sites in the \textit{PstI} partial fragment was a clear indication that we had cloned a single contiguous section of

\begin{table}
\centering
\caption{Strains and plasmids used}
\begin{tabular}{|l|c|c|}
\hline
Strain or plasmid & Relevant genotype or phenotype & Source or reference \\
\hline
\textbf{Strain} & & \\
AB1157 & $F^{-}$ thi-1 thr-1 leu-6 proA2 argE3 his-4 lacY1 galK2 xyl-5 ara-14 rps sup-37 & P. Howard-Flanders \\
TK701 & $F^{-}$ thi-1 proA2 his-4 lacY1 galK2 xyl-5 mit-l supE44 & T. Kato \\
TK702 & As TK701, but umaC36 & T. Kato \\
CA222 & As TK702, but gyrA & This laboratory \\
K-12 DH1 & $F^{-}$ lacZ bio-uvrB trpEA2 (λSam7 Num53 c1857 ΔH1) & W. Fiers \\
\textbf{Plasmid} & & \\
TP110 & Km\textsuperscript{r} cib\textsuperscript{+} imp\textsuperscript{+} & G. O. Humphries \\
pHSG415 & Ap\textsuperscript{r} Km\textsuperscript{r} Cm\textsuperscript{r} & Hashimotoh-Gotoh et al. (8) \\
pBR325 & Ap\textsuperscript{r} Te\textsuperscript{r} Cm\textsuperscript{r} & Bolivar (3) \\
PLC28 & Ap\textsuperscript{r} & Remault et al. (18) \\
pJAG253 & Km\textsuperscript{r} Cm\textsuperscript{r} imp\textsuperscript{+} & This paper \\
pJAG256 & Km\textsuperscript{r} Cm\textsuperscript{r} imp\textsuperscript{+} & This paper \\
pJAG278 & Km\textsuperscript{r} Cm\textsuperscript{r} imp\textsuperscript{+} & This paper \\
pJAG10 & Ap\textsuperscript{r} Km\textsuperscript{r} imp\textsuperscript{+} & This paper \\
pJAG11 & Ap\textsuperscript{r} Km\textsuperscript{r} imp\textsuperscript{+} & This paper \\
pJAG66 & Ap\textsuperscript{r} imp\textsuperscript{+} & This paper \\
pJAG68 & Ap\textsuperscript{r} imp & This paper \\
pJAG70 & Ap\textsuperscript{r} imp\textsuperscript{+} & This paper \\
pJAG72 & Ap\textsuperscript{r} imp & This paper \\
\hline
\end{tabular}
\footnote{Abbreviations: Ap\textsuperscript{r}, ampicillin resistance; Km\textsuperscript{r}, kanamycin resistance; Cm\textsuperscript{r}, chloramphenicol resistance; Te\textsuperscript{r}, tetracycline resistance.}
\end{table}

Identification of proteins expressed from \textit{PstI}. Cells of strain K-12ΔH1 containing fragments cloned in PLC28 were grown to mid-log phase at 30°C and then switched to 42°C to initiate expression from \textit{PstI}. At 0, 1, 2, 3, and 4 h samples were taken, labeled with $[^{35}\text{S}]$methionine, and then analyzed by polyacrylamide gel electrophoresis (PAGE) as described above. The 3-h time point was found to be optimal for visualization of the \textit{imp} protein(s).

\section*{RESULTS}

Cloning of the \textit{imp} genes from TP110. Initial attempts to clone the UV protection and mutation genes of TP110 into either pBR322 or pBR325 proved unsuccessful despite the use of a variety of enzymes to generate restriction fragments, and it seemed likely that the genes could not be tolerated in high copy number. Partial digestion with \textit{PstI}, which is known to cut TP110 frequently (M. Ramsden, Ph.D. thesis, University of Liverpool, Liverpool, United Kingdom, 1981), was therefore used to generate fragments which were ligated into the \textit{amp} gene of pHSG415, a vector with a copy number of approximately four (8). The DNA so generated was transformed into TK702 umaC36, and the ampicillin-sensitive colonies derived from the transformation were patch-tested for UV survival and UV-induced reversion of the his-4 marker as described previously (5). From 100 such ampicillin-sensitive colonies, 3 were isolated which showed both higher survival and higher mutation rates than the TK702 control; none were obtained which displayed just one of these properties. Restriction mapping (Fig. 1) of the

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Cloning of the \textit{imp} genes from TP110 into the vector pHSG415. \textit{PstI} partial digestion products of TP110 were ligated into the unique \textit{PstI} site of the vector which lies within the ampicillin resistance gene (\textit{Amp}). Three independent \textit{Imp}\textsuperscript{+} plasmids were isolated by selection for the kanamycin and chloramphenicol resistance markers of the vector (\textit{Kan} and \textit{Cam}, respectively) and screening for restored mutability in the \textit{umaC36} recipients. The extent and orientation of the cloned fragment in the three plasmids (253, 256, and 278) are shown. The coordinates of the TP110 parent molecule and the relevant restriction sites relate to the complete map of this plasmid presented previously (17). The positions of eight independent \textit{Tn1000} insertions which abolished \textit{imp} activity are shown. Two fragments, \textit{EcoRI}-E and \textit{BglII}-J, are shown as hatched boxes. These fragments were respectively subcloned and deleted from the original clones in later work (see text).}
\end{figure}
the TP110 genome and not a composite of unrelated fragments.

The fragments inserted in clones pJAG256 and pJAG253 were identical in size but were present in opposite orientation. As both clones complemented the *umuC36* mutation of TK702 to the same extent (Fig. 2), we concluded that the *imp* genes were expressed from their own promoter on the cloned fragment and not from the *amp* promoter of the vector. The DNA cloned in pJAG278 did not contain the 1.3-kb *PstI* fragment which lies downstream from the *imp* genes in pJAG253 and pJAG256 but contained an extra 7.6-kb *PstI* fragment upstream. As these fragments were not common to all of the *imp* clones, we assumed them to be unnecessary for *imp* activity. This conclusion was supported by the ability of pJAG278 to enhance mutation and confer protection to exactly the same extent as the other two clones (data not shown), although in each case the restoration of mutability was never quite as high as that observed with TP110 itself. This small discrepancy is apparently a feature of these particular clones, as subclones derived from them were just as effective as TP110 in restoring UV protection and mutation (see below).

**Localization of the coding region.** To determine more precisely the region of the cloned fragment occupied by the *imp* genes, a series of *imp::Tn1000* insertion mutations were generated in pJAG256. All of the insertions causing loss of *imp* function mapped within TP110 *EcoRI* fragment E (Fig. 1), and we therefore adopted a variety of subcloning approaches to determine whether this fragment alone carried the whole of the *imp* functions. We had previously shown that a series of *imp::Mu d*(Ap lac) insertions also mapped within this fragment (5). Initially, we attempted to clone *EcoRI* fragment E into the *EcoRI* site of the multicopy plasmid pBR325 by purifying the fragment from pJAG256 and treating the vector with alkaline phosphatase. Such clones were obtained, in one orientation only, but proved to be particularly unstable and difficult to work with. To improve the stability of the construct, *EcoRI* fragment E was subcloned into the unique *EcoRI* site of pHSG415, which lies within the *cam* gene. In this vector, the fragment could be cloned in both possible orientations (pJAG10 and pJAG11), each of which was fully as effective as TP110 in complementing the *umuC36* allele for mutagenesis (data not shown). Strains carrying these plasmids were perfectly viable and showed no tendency to lose the plasmid. We concluded, therefore, that all of the functions necessary for the expression of the *imp* phenotype were carried on TP110 *EcoRI* fragment E and that overproduction of these products could not be tolerated by the cell, as evidenced by the behavior of the high-copy-number plasmid. Other factors also affected the stability of *imp* plasmids, as demonstrated by an attempt to delete *BglII* fragment J (Fig. 1) from pJAG256 by restriction digestion and religation. The resulting deletion plasmid appeared to retain its protection and mutation properties but was lost from host cells so rapidly that accurate studies were not possible. Similar problems of instability have been reported by Marsh and Walker (12) in cloning the functionally related *umu* genes, for which overproduction of the gene products appears to interfere with DNA replication.

**Analysis of gene products and their overproduction.** Attempts to analyze the proteins encoded by the *imp* genes of pJAG10 in minicells and maxicells were not successful. It seems that expression of *imp* proteins may occur at such a low level that only small amounts of protein are produced, making detection difficult.

To overcome this problem, we linked the *imp* gene(s) in the correct orientation to the lambda *pl* promoter by cloning *EcoRI* fragment E into the vector pLC28 (18), enabling us to switch expression from a low to a high level, as required, by thermal inactivation of the *cl* repressor. A restriction map of the plasmid used and the map positions of Tn1000 insertions which inactivated the *imp* genes are shown in Fig. 3. No attempt was made with this construct to remove the normal *imp* promoter, and, rather surprisingly, this high-copy-number plasmid proved to be perfectly stable. It may be that *imp* in this host-vector system was expressed less than when the fragment was cloned into pBR325, even though the *imp* promoter was still present, and that expression of *imp* was sufficient to give normal levels of protection and mutation even when *pl* was completely repressed. Both possible
orientations of EcoRI fragment E in this vector were obtained, and both complemented the umuC36 mutation in a lambda lysogen of TK702 to the same extent (data not shown). Only one orientation, of course, was of use in linking the imp gene(s) to plasmid.

With this expression vector, it was possible to identify two proteins produced by pJAG66 which were not produced when the imp genes of this plasmid were inactivated by Tn1000 insertion, these two proteins having apparent molecular weights of 40,000 and 51,000 (Fig. 4A).

To determine whether both of these proteins were required for protection and mutation functions, deletion derivatives of pJAG66 were constructed and analyzed for their protection and mutation properties and protein production (Fig. 3 and 4B). Deletion of the 0.75-kb PsI fragment (pJAG68) resulted in the loss of protection and mutation functions. Deletion of HindIII fragment H and sequences 3′ to it (pJAG72) had the same effect, while deletion of BglII fragment J and all sequences 3′ to it (pJAG70) left the mutation and protection properties intact. Analysis of the proteins produced by these deletion derivatives showed that the PsI deletion produced the 40,000-MW protein but not the 51,000-MW species; the HindIII deletion failed to produce either protein (data not shown), and the BglII deletion produced only the 51,000-MW protein. We concluded, therefore, that the 51,000-MW protein was essential for imp activity and was presumably encoded by an imp gene. The 40,000-MW protein was apparently made from a gene downstream from imp, and its transcription from plasmid was affected by the polarity of Tn1000 insertions in the imp genes. Whether this gene is normally transcribed from the imp promoter is not possible to determine from the data described, but it is clear that there cannot be a strong transcriptional terminator between imp and the 40,000-MW protein gene, which appears to lie in the region of DNA defined by BglII fragment J of TP110.

Figure 5 summarizes the deletion data and indicates the position of Tn1000 insertions which abolished imp activity. Taking the distance of the two extreme inserts as the limits of the imp operon, a minimal coding sequence of approximately 1.65 kb is indicated. Even allowing for a substantial nontranslated region between the promoter and the start of the structural gene, there is clearly an excess of coding capacity over that required for a single 51,000-MW protein. Previous studies with both the umu and muc operons have revealed that each is made up of two genes, overlapping by a small number of nucleotides (13 and 13, respectively [14]), and a similar arrangement might be expected to be found in the functionally related TP110 system. In an attempt to identify such a second gene product, DNA from pJAG70 was introduced into a Zobay cell extract system, and the proteins synthesized were analyzed by PAGE on a 15% polyacrylamide gel. Figure 6 shows tracks from such a gel, comparing vector DNA with pJAG70. Two extra proteins were clearly produced by the clone, one of 51,000 MW, the other of approximately 11,000 MW, the smaller of the two proteins being produced in much greater quantity than the larger. In previous maxicell and minicell preparations, this protein was not visible due to the background of low-MW proteins. The production of the 51,000-MW protein in the cell-free system clearly confirms the assignment of this protein as an imp product. To confirm that the smaller protein was also a product of these genes, the EcoRI-BglII imp-encoding fragment was cloned into the GeneScribe T7 promoter vector pT7-1 (U.S. Biochemical Corp.), the clone was linearized with either Sall (downstream from the imp genes) or HindIII (within the imp coding region), and extensive runoff trans-

**FIG. 3.** Coupling of the imp genes to the leftward promoter of lambda, pl, in the expression plasmid pJAG66, and the creation of deletion derivatives. The lines labeled pJAG68, pJAG70, and pJAG72 show the DNA deleted to yield these derivatives. pJAG68 was created by deleting the 0.75-kb PsI fragment from pJAG66. pJAG70 was created by joining the BglII site within EcoRI fragment E to the BamHI site of the vector multilinker, and pJAG72 was created by joining the HindIII site within imp to the HindIII site of the multilinker. Restriction sites: E, EcoRI; P, PsI; H, HindIII; B, BglII. The numbers 37, 65, 3, and 54 indicate the sites of four independent Tn1000 insertions which abolished imp activity.

**FIG. 4.** PAGE fluorogram of labeled proteins following pulse labeling with [35S]methionine of cells of strain K-12A1 containing pJAG66 and Tn1000 or deletion derivatives. Cells were incubated at the nonpermissive temperature (42°C) for 3 h before labeling. Previous experiments had shown this length of incubation to be optimal for reducing background protein production while allowing synthesis of plasmid-coded proteins to continue. (A) Lane 1, MW markers. Sizes (in kilodaltons) are given to the left of the panel. Lanes 2, 3, 4, and 5. Four independent pJAG66 imp::Tn1000 inserts (numbers 37, 3, 54, and 65, respectively; see Fig. 3). Lane 6, pJAG66 imp'. (B) Lane 1, pJAG68 imp. Lane 2, pJAG70 imp'.
scripts were created by transcription with T7 polymerase. The mixture of DNA and transcript (approximate ratio, 1:50) was again added to a Zubay in vitro transcription-translation system, and the protein products were identified by PAGE. The results (not shown) demonstrated extensive production of an 11,000-MW protein species in both transcribed preparations, but the 51,000-MW species was not produced in either.

**DISCUSSION**

The *imp* genes encoding the UV protection and mutation properties of TP110 have been cloned into both low- and high-copy-number vectors, and two proteins apparently encoded by the *imp* genes have been identified. Of these, the 51,000-MW protein is essential for the protection and mutation properties, while the 11,000-MW species was shown to be encoded by DNA upstream from the gene for the larger protein. Deletion of this DNA caused a loss of the Imp + phenotype, and although it remains possible that a third, as yet unidentified protein is encoded by this region, the sizes of the proteins and the length of the coding region suggest very strongly that the 11,000-MW protein is indeed an *imp* product. The observation that these two proteins are not the same size as those encoded by the *umuDC* and *mucAB* operons demonstrates the variation which exists among these functionally related operons, as might be anticipated from the fact that the *umu* operon is chromosomally located, while the plasmid-borne operons are found on a wide variety of unrelated plasmid species. In the case of the *umu* and *muc* operons, it is clear from cloning and sequencing data that the two genes are all that are required for the UV protection and mutation phenotypes (11, 14). In the case of the *imp* operon, preliminary complementation data (to be presented elsewhere) also indicate that just two cistrons are involved, although a more extensive analysis needs to be completed before this statement can be made definitively. With respect to the size of the proteins, it is intriguing that the sum of the sizes of the gene products so far identified appears to be constant; 15,000 and 47,000 MW in the case of *umuDC*, 16,000 and 45,000 MW in the case of *mucAB*, and 11,000 and 51,000 MW in the case of *imp*. This may be an indication that all of these functionally related gene products have evolved from a common ancestral sequence; DNA and amino acid sequence data suggest that this is true for *umu* and *muc* (14), and we are currently sequencing the *imp* region to determine whether these genes could share the same ancestry; preliminary results indicate some areas of homology (data to be published elsewhere). The similarity in operon arrangement and the common control by LexA are also indications that *imp* may be related structurally to *umu* and *muc*.

If these genes do all prove to be structurally related, a clear implication would be that they have at one time been carried on a mobile genetic element which has allowed them to spread through a variety of unrelated replicons. Consistent with this hypothesis, Sedgwick and Goodwin (19) have provided evidence for DNA rearrangements associated with the *umuC* region of *E. coli* K-12 and demonstrated that many naturally occurring *E. coli* strains do not in fact carry these genes. We have attempted to obtain evidence for the trans-

---

FIG. 5. Summary of deletion data and collation of Tn1000 to indicate minimal *imp* coding region. The positions of Tn1000 insertions include those isolated in pJAG256, pJAG10, pJAG66, and pJAG70. The direction of *imp* transcription is shown with the broken arrow, and restriction site abbreviations are defined in the legend to Fig. 3.

FIG. 6. PAGE fluorogram of labeled proteins synthesized in a Zubay cell extract system with (A) pLC28 and (B) pJAG66 purified DNA as the template. The positions of marker proteins are shown (in kilodaltons), as are the positions of the 11,000- and 51,000-MW putative *imp* proteins.
position of the TP110 imp genes without success. The genes do not appear to be flanked by inverted repeats, as judged by electron microscopy, nor, from the limited studies we have carried out, do they appear to be capable of promoting cointegrate formation during conjugation. Whether or not they were at one time carried on a mobile element which has now lost the ability to transpose remains an open question.

The widespread occurrence of protection and mutation genes in a large number of unrelated replicons may have some implication for the nature of the action of the gene products. It seems likely, given the diversity of the genetic systems involved, that the plasmid-coded functions do not interact specifically with host replication proteins, but rather carry out some discrete function which permits the host to complete the fixation of the mutation. A specific interaction with polymerases would seem unlikely, for example, with the muc operon, as this plasmid system is reported to show its effects not only in a wide variety of bacterial hosts but also when transformed into cultured mouse cells (7).

A consistent observation with both the umu and muc operons is that the 15,000- to 16,000-MW protein is synthesized in much greater quantity than the 47,000- to 46,000-MW protein (6, 11, 20). Our observations that a similar effect is seen with the imp genes in the in vitro Zabub system and that only the 11,000-MW protein is translated when transcription is greatly increased suggest that the production of the second protein may require some processing which is achieved only inefficiently, if at all, in vitro. If the smaller of the two proteins were to act as a repressor of the synthesis of the larger, then its overproduction following translation of the T7 transcripts might prevent production of the larger protein. This hypothesis is currently being tested.

ACKNOWLEDGMENTS

The work in this paper was supported by a Science and Engineering Research Council (SERC) project grant to P.S. J.A.G. was the recipient of an SERC postgraduate training award.

The excellent technical assistance of Annette C. Ritson is gratefully acknowledged.

LITERATURE CITED