The Chloroplast-Located Homolog of Bacterial DNA Recombinase

Jun Cao, Carolyn Combs and André T. Jagendorf

Plant Biology Section, Plant Science Building, Cornell University, Ithaca NY 14853, U.S.A.

The cDNA for the chloroplast-located homolog of bacterial RecA protein, designated recA-AT, was placed in a plasmid appropriate for in vitro transcription and translation. Translation with 35S-labeled Met permitted demonstration of uptake of the protein product into isolated pea chloroplasts, and processing to a mature size. Preliminary evidence for the first amino acid was estimated from results using both 35S-Met and 3H-Leu for in vitro transcription and translation, followed by uptake into chloroplasts and processing. The labeled protein was subject to sequential amino acid hydrolyses, and radioactivity was measured in each round. Induction of gene transcription in leaves infiltrated with the DNA-damaging agent, methyl methane-sulfonate was shown by Northern blot analysis. Further constructs were made for over-expression of the gene in E. coli; and one out of many tried permitted production of some soluble protein. Extracts from transformed bacteria were shown to have RecA activity using the "POM" assay [Bertrand et al. (1993) Nucl. Acids Res. 21: 3653] for DNA strand transfer. The protein was purified to close to homogeneity using methods developed for E. coli RecA isolation.

Key words: Arabidopsis thaliana — Chloroplast — DNA damage — DNA recombinase — Homologous recombination — RecA.

The plastomes (chloroplast genomes) of higher plants and algae undergo homologous recombination. The most direct evidence for this is found in the transformation of chloroplast DNA using bio-blastic (i.e., the "gene gun" using particles coated with DNA) technology (Svab et al. 1990, Staub and Maliga 1992, Svab and Maliga 1993, Golds et al. 1993, Carrer et al. 1993). The introduced sequences combine randomly in homologous regions of the plastome of higher plants. In a few cases the artificial production of "cybrids" or somatic hybrids has led to observable recombination between chloroplasts of different species

Abbreviations: MMS, methyl methane sulfonate; recA-AT, Arabidopsis thaliana gene homologous to the recA genes of bacteria; RecA-AT, Arabidopsis thaliana protein homologous to RecA proteins of bacteria; SSC, 0.15 M NaCl, 0.015 M sodium citrate.
1 Current address: Pharmacology Department, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, U.S.A.
2 Corresponding author.
3 Genbank Accession Numbers M93899 for the cDNA, and L15229 for the genomic sequence.

Materials and Methods

Plants—Seeds of Arabidopsis thaliana ecotype Columbia were obtained from Dr. Rob Last. Seeds of pea (Pisum sativum, cv. Progress No. 9) were from Agway, Ithaca, NY. Arabidopsis and pea plants were grown in a growth chamber under a 12 h pho-
Higher plant chloroplast RecA

Protein expression in vivo—For production of either the mature or the preprotein without a hexa-histidine handle, *E. coli* cells carrying pJC-5 were grown overnight in LB broth containing 50 μg ml\(^{-1}\) ampicillin at 37°C, then inoculated into fresh LB broth containing 50 μg ml\(^{-1}\) ampicillin. When the absorbance at 600 nm reached 0.4, isopropyl-β-thiogalactopyranoside (IPTG) was added to a concentration of 1 mM, and growth was continued for 3 h.

For expression of proteins with the hexa-histidine handle, the plasmid pJC-6 (for the mature protein) or pCON 3 (for the preprotein) were transfected into M9 salt and NZC broth containing 50 μg ml\(^{-1}\) ampicillin and 1 mM IPTG at 30°C (present from the start of growth) for 15 h.

Bacterial cells were harvested by centrifugation and resuspended in HEBES buffer (25 mM HEPES pH 7.0, 50 mM NaCl, 10% glycerol). Cells were lysed with 1.0 mg ml\(^{-1}\) lysozyme followed by sonication. Total protein concentration was measured using the Biorad Protein Assay, and RecA-AT protein concentration was estimated by immunoblot analysis.

**Immunoblot analysis**—Proteins were transferred from SDS gels (10% polyacrylamide) to nitrocellulose membranes electrophoretically (Towbin et al. 1979). The membranes were blocked with 5% (w/v) nonfat dry milk, incubated with 1:4,000 RecA antiserum, washed, then incubated with alkaline phosphatase-conjugated goat anti-rabbit IgG. After further washing, the RecA-AT protein was detected with the alkaline phosphatase substrate 5-bromo-4-chloro-3-indoly phosphate/nitro blue tetrazolium in alkaline phosphatase buffer (100 mM NaCl, 5 mM MgCl\(_2\), and 100 mM Tris-HCl pH 9.0).

**Enzyme assays**—The assay for activity of the RecA protein was conducted according to Bertrand et al. (1993). In this procedure, nitrocellulose membranes are coated with ssDNA (M13mp18, 0.5 μg cm\(^{-2}\)) in 2× SSC, then fixed on the membranes by heating in a vacuum oven for 2 h at 80°C. Proteins, separated by native PAGE, are electroblotted onto the ssDNA-coated nitrocellulose membranes in transfer buffer, then washed three times, for 5 min each, in buffer A (33 mM Tris-HCl pH 7.5; 1 mM MgCl\(_2\); 2 mM dithiothreitol, 0.5 mM ATP/γS). The membranes are then blocked for 1 h at 37°C with a boiled solution of 5% nonfat dry milk in buffer A and rinsed with the same buffer. The staining reaction was performed by incubating the membranes in a sealable plastic bag with 200 μl sample of labeled M13mp18 duplex DNA in buffer B (33 mM Tris-HCl pH 7.5; 12 mM MgCl\(_2\); 2 mM dithiothreitol, 0.1 mM ATP/γS). The duplex DNA had been labeled with \(^{32}\)P by end-labeling via a filling reaction in the EcoRI generated 5′ protruding ends, using polymerase I large fragment (Life Technologies). The probe was used at a concentration of 1 μg ml\(^{-1}\) with a specific activity of 10\(^{6}\) cpm μg\(^{-1}\). After the reaction, the membranes were treated with proteinase K (0.1 mg ml\(^{-1}\)) for 1 h at 37°C and then successively washed in 20× SSC; 10× SSC, 0.1% SDS; and 2× SSC, 1% SDS for 15 min at 65°C for each wash. The reaction products were visualized by autoradiography.

In vitro transcription and translation, and chloroplast import of the precursor—In vitro transcription and translation in the presence of \(^{35}\)S-Met were carried out in a TnT\(^{TM}\) coupled reticulocyte lysate system according to the manufacturer’s (Promega) protocol. \(^{35}\)S-Met was from the DuPont (NEN Life Sciences Products) Co. It had a specific activity of 37 TBq mmol\(^{-1}\), and 4 μl (1.48 MBq) was added to each reaction. The plasmid carrying the coding sequence of the *recA*-AT gene under the control of SP6 promoter was linearized with Xhol and used as the template. The reactions were
incubated with SP6 polymerase in a 50 μl volume for 2 h at 30°C.

Intact pea chloroplasts were isolated from 10-d-old pea shoots and purified on a Percoll gradient (Nivison et al. 1986). The import reaction contained 2 mg chlorophyll and 35S-Met labeled pre-RecA-AT precursor in 1.0 ml of import buffer (50 mM HEPES-KOH, pH 8.0, 330 mM sorbitol) together with 2 mM MgATP and 8 mM Met. After incubation at 23°C for 60 min, the chloroplasts were treated with 100 μg ml⁻¹ thymolysin and incubated for 30 min on ice before intact chloroplasts were re-isolated by centrifugation through 40% Percoll. The chloroplasts were washed in import buffer, and the radioactive translation products separated by 10% SDS-PAGE, then detected by autoradiography.

N-Terminal analysis of the mature RecA-AT protein—Transcription and translation were performed as above, with both 35S-Met and 3H-Leu in the reaction medium. Intact isolated chloroplasts were added after translation was finished, and the doubly labeled protein was imported as described above. After thymolysin treatment the chloroplasts were re-isolated, dissolved in SDS, electrophoresed on a 10% SDS-PAGE gel and blotted to Immobilon PVDF membranes. Radioactive protein at the predicted kDa values for the mature RecA-AT was eluted and sequenced by automated Edman degradation (Cornell DNA Technology Facility). The products of each sequence cycle were collected, and their radioactivity determined by liquid scintillation counting.

Total RNA isolation—Leaves of 3-4 week old Arabidopsis plants were collected and vacuum infiltrated with 7 mM MMS. Control leaves were infiltrated with water. Total RNA was isolated by a modification of the phenol/SDS method (Vervoord et al. 1989). MMS treated leaves were ground to a fine powder under liquid N2 with a pre-chilled mortar and pestle, suspended in grinding buffer and transferred to a 30 ml centrifugation tube. A half volume of chloroform was added, and the suspension centrifuged at 10,000 rpm for 20 min at 4°C. RNA was precipitated with an equal volume of 4 M LiCl, stored at −20°C overnight and collected by centrifugation at 10,000 rpm for 20 min at 4°C.

RNA gel blot hybridization—Twenty μg of total RNA were loaded on a formaldehyde-formamide (denaturing) 1.25% agarose gel and then transferred to nitrocellulose membrane as described by Sambrook et al. (1989). Hybridization conditions were as described (Cao et al. 1992). 32P-labeled DNA probes were prepared from the coding region of the recA-AT gene, and also of the cDNA for actin from Brassica anthers (Tobias et al. 1991) by random primer labeling. Filters were washed sequentially at 65°C in 2×, 1×, then 0.1× SSC, all containing 0.1% SDS; then exposed to X-ray film.

Purification of the over-expressed protein—E. coli cells containing vector pJC5 were grown in LB broth with 50 mg ml⁻¹ Ampicillin, treated with IPTG to induce the protein, harvested, and frozen and thawed 3 times. They were then treated with lysozyme at 2 mg ml⁻¹, shaken at 250 rpm and 37°C for 30 min, and centrifuged. The supernatant fluid contained considerable transformant protein, but more remained in the pellet, presumably in inclusion bodies. These were extracted by 5 successive sonication periods of 15 s each, centrifuged, and the successive supernatants contained more and more of the residual RecA protein (data not shown).

Polyethyleneimine (final concentration 0.5%) was added to the combined supernatants, which were then centrifuged (10,000 rpm for 10 min). The supernatant solution was discarded, and the pellet extracted with "R" buffer (20 mM Tris-HCl, pH 7.5, 10% glycerol, 1 mM DTT and 0.1 mM EDTA) containing 150 mM (NH₄)₂SO₄, and centrifuged. The final pellet was extracted with 300 mM (NH₄)₂SO₄ in R buffer, centrifuged, and RecA-AT in the extract was precipitated by 0.28 g (NH₄)₂SO₄ per ml. This precipitate was centrifuged, and the resulting pellet was redissolved in R buffer. Overnight dialysis against 20 mM MgCl₂ in R buffer caused precipitation of the RecA-AT protein, which was collected by centrifugation. After redissolving in 20 mM Na acetate, pH 6.5, it was precipitated by spermidine acetate at 7 mM, and collected by centrifuging. The pellet was redissolved, then dialyzed against R buffer.

Results and Discussion

The Arabidopsis recA-AT gene codes for a preprotein targeted to the chloroplast—In view of the presumptive transit peptide in the primary sequence of the recA-AT gene, it was advisable to provide a more direct demonstration that the gene codes for a protein targeted for chloroplast entry. Accordingly the cDNA was linearized, then transcribed and translated in a cell-free reticulocyte-derived system (see Materials and Methods) with 35S-Met in the reaction mixture. After incubation with isolated intact pea chloroplasts, then digestion of non-imported proteins with thermolysin, the protected polypeptides were electrophoresed and subject to autoradiography. A significant proportion of the translation product (Fig. 2, lane 1) was

![Fig. 1 Restriction maps of constructs made for this project. For explanation see text. Unusual abbreviations include TP for the transit peptide of RecA-AT, and EK for Enterokinase.](image-url)
brought into a region inaccessible to thermolysin, and processed to a mature size of approximately 42 kDa, as predicted from the amino acid sequence of the presumed gene product (Fig. 2, lanes 2 and 3). In the absence of thermolysin treatment after the import reaction an additional higher molecular weight band was observed, indicating some of the in vitro translated precursor did not enter chloroplasts and remained unprocessed (data not shown).

The chloroplast RecA-AT preprotein processing site—The amino acid sequence, Val-Tyr-Ala-Ala in positions 50–53 of the amino acid sequence are consistent with 90% of the known chloroplast processing protease sites in chloroplast preproteins (Gavel and Van Heijne 1990). However, introns in many genes define different domains of the protein. A number of chloroplast proteins have the transit peptide in a separate domain from the mature protein; therefore the first intron-exon border could signal the start of the mature protein. In the case of recA-AT, the first exon ends at Arg83 and the second exon starts at Gln84 of the coding region. While this could also be the processing site, it would predict a mature protein of 355 amino acids, with a molecular mass of approximately 39 kDa. The observed apparent size of 42 kDa argues for the predicted processing sequence, rather than the exon/intron border.

An attempt was made to obtain more direct experimental evidence for the processing site. If the beginning of the second exon is the start of the mature RecA-AT protein, the first 10 amino acids would be: Gln-Lys-Ala-Leu-Glu-Ala-Ala-Met-Asn-Asp.

If the processing site lies within the predicted Val-Tyr-Ala-Ala sequence, the first 10 amino acids are: Ala-Lys-Lys-Leu-Ser-His-Lys-Ile-Ser-Ser.

While both sites would have a Leu residue as the 4th amino acid, only the exon border start site would have a Met in position 8. Accordingly, transcription and translation were performed with both 3H-Leu and 35S-Met. The translation products were incubated with intact chloroplasts, then any remaining external proteins were digested with thermolysin. Radioactive protein was isolated from the chloroplast stroma, and subject to micro-amino acid sequencing (see Materials and Methods). Significant radioactivity came off on round four of the sequencing cycle, as predicted for both possible start points. However there did not appear to be extra radioactivity from round 8, whether

![Samples kDa](image1)

**Fig. 2** Import and processing of pre-RecA by isolated intact pea chloroplasts. The recA-AT gene in plasmid pIM-3 was transcribed and translated in vitro with the Promega TNT system in the presence of 35S-Met for label. Import reactions were carried out by incubation of the pre-RecA-AT with intact pea chloroplasts isolated from a Percoll gradient; then thermolysin was added to degrade external proteins. EDTA was added to stop thermolysin action, then chloroplasts were sedimented and electrophoresed on SDS-PAGE. The gel was dried and autoradiographed. Lane 1: P indicates preRecA-AT protein. Lanes 2 and 3: M indicates mature RecA-AT after the import reactions for 30 (lane 2) or 60 min (lane 3).

![Sequencing Cycle No.](image2)

**Fig. 3** NH2-terminal sequence analysis of the RecA-AT after the import reaction. In vitro synthesis of preRecA-AT was carried out with 35S-Met and 3H-Leu. The translation mixtures were then incubated with intact pea chloroplasts as described in Fig. 2 and in Material and Methods. Intact pea chloroplasts were then treated with thermolysin to degrade external proteins, resolved by SDS-PAGE. Proteins were blotted to Immobilon membrane, and RecA-AT identified by autoradiography, eluted and sequenced. The radioactivity released by each sequencing cycle was measured by scintillation counting. Background has been subtracted from the data presented.
the counter discrimination was set for $^3$H as shown, or for $^3$S (data not shown). Although a negative result cannot be conclusive, the experimental result is at least consistent with the conclusion from the molecular mass data. More critical data might have been obtained using radioactive Alanine as the label, but unfortunately attempts to introduce label from $^3$H-Alanine into the translation product were not successful.

Induction of Arabidopsis recA-AT mRNA—In an earlier study (Cerutti et al. 1993) the chloroplast protein responding to antibody against E. coli RecA was found to increase 2.5 to 3-fold when pea protoplasts were incubated with MMS or other DNA damaging agents. In the current work we detached Arabidopsis leaves, and vacuum infiltrated them with MMS. These leaves were incubated under aerobic conditions, in room light for varying times, then frozen. Northern blots showed that the abundance of the transcript of the recA-AT gene increased 5-fold or more (Fig. 4).

It is potentially interesting that the recA-AT mRNA is induced to a higher level than that of the protein in the previous study (Cerutti et al. 1993). However the conditions of induction were different in the two cases. In the earlier work pea leaf protoplasts were suspended in the DNA damaging agent; in the present study Arabidopsis leaves were vacuum infiltrated with the reagent. It remains to be seen whether the difference is due to use of different species, or leaves vs. protoplasts, or if it is consistent for protein level vs. mRNA level within the same species.

<table>
<thead>
<tr>
<th>Hours</th>
<th>Ctrl</th>
<th>MMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4 Increases in mRNA levels in response to treatment with MMS for varying times. Total RNA (20 μg), isolated from three to four-week-old Arabidopsis leaves vacuum infiltrated with MMS for indicated hours were electrophoresed on 1.25% agarose-formaldehyde gels, blotted onto membranes and analyzed by hybridization with the $^{32}$P-labeled recA-AT cDNA or actin cDNA as the probe, as indicated. Lane 1 is the control, with leaves treated with water for 9 h. In lanes 2 to 5 leaves were treated with MMS for 1, 3, 6, and 9 h respectively.

Northern dot blots were performed with extracts from roots, stems and flowers as well as leaves. These showed RecA-AT mRNA to be present in all of these organs (data not shown).

The protein product of recA-AT, over-expressed in E. coli, shows DNA strand transfer activity—The cDNA for recA-AT was introduced into E. coli in a number of

![Fig. 5](image)

Fig. 5 Western blot of over-expressed RecA-AT protein. E. coli cells transformed to produce either the preprotein, the mature protein, or the mature protein with a hexa-histidine handle, were induced with IPTG and extracted as described in Materials and Methods. Broken cells were fractionated into soluble vs insoluble proteins by centrifugation at 12,000 × g for 10 min. Soluble proteins were separated in 10% SDS-PAGE and transferred to nitrocellulose filter. The immobilized proteins were incubated with RecA antiserum. Lane 1, fusion protein of pre-RecA-AT with a hexa-histidine handle; lane 2, RecA-AT fusion protein produced from pJC-6; lane 3, RecA-AT protein without a hexa-histidine handle, produced from pJC-5; lane 4, pure E. coli RecA protein.

![Fig. 6](image)

Fig. 6 Analysis of RecA-AT activity in bacterial crude extracts separated on native PAGE. Bacterial crude extracts carrying vectors either with or without the coding sequence of the recA-AT gene were separated on native PAGE and blotted onto membrane pre-coated with ssDNA homologous to its duplex M13p18DNA as the probe. Lane 1, 5 μg of pure E. coli RecA protein; lane 2, 100 μg of protein from crude extract from the DF2566 cells carrying pJCS; lane 3, 100 μg of protein from crude extract DF2566 carrying the control plasmid, pTrc 99A.
plasmids. These included pJC-6 and pCON 3, versions with a hexa-histidine handle. In these, and other early attempts the protein yield was either very low; or else the RecA-AT was primarily located in inclusion bodies, and could not be renatured to give material with enzymatic activity. In the current procedure the mature protein, induced by IPTG in a recA-deficient E. coli strain, is found to be about 50% soluble (Fig. 5). The recA deleted E. coli strain FD2566 was transformed either with plasmid pJC-5 containing the coding sequence of the recA-AT gene or the same plasmid without the recA-AT gene. Expression was induced by growth in the presence of IPTG; cell extracts were obtained by use of lysozyme and sonic oscillation.

The extracts were tested for strand transfer activity using the "POM" procedure of Bertrand et al. (1993). This assay, which uses ssDNA bound to a nitrocellulose membrane, was asserted to be specific for RecA in bacterial extracts (Bertrand et al. 1993). It is apparent (Fig. 6) that the extract from cells containing the recA-AT insert, and which had been induced by use of IPTG, have activity in this assay but extracts from the control cells do not.

**Purification of over-expressed recA-AT**—Using the procedures appropriate for purification of RecA from *E. coli* (i.e., successive precipitations by polyethyleneimine, ammonium sulfate, MgCl₂, and spermidine acetate), the over-expressed protein was purified from the original *E. coli* extract. The appearance of the proteins on SDS-polyacrylamide gel electrophoresis is shown in Fig. 7, and approximate yields from each step are shown in Table 1. In the final step, our estimate is that 96% of the protein is immunoreactive. The success of these procedures emphasizes the relationship of the plant protein to the ones found in bacteria. The purified protein was used for antibody production in rabbits, and will be critical to the planned analysis of the biochemical function of this protein. It is of considerable comparative biochemical interest to see how closely the chloroplast enzyme resembles that of the bacteria from which it is most probably descended.

Supported in part by grant 93-01266 from the USDA/CSRS/NRICG Program. We wish to thank Drs. J. and M. Nasrallah for use of their Exomax machine, and Dr. J.B. Hays for a gift of the RecA-deficient strain of *E. coli*.

**References**


### Table 1 Purification of *Arabidopsis* RecA-AT protein over-expressed in *E. coli* FD 2566

<table>
<thead>
<tr>
<th>Fractions</th>
<th>Total protein (mg)</th>
<th>Total recA protein (mg)</th>
<th>recA protein (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insoluble extract</td>
<td>40</td>
<td>10.5</td>
<td>26</td>
</tr>
<tr>
<td>Soluble extract</td>
<td>180</td>
<td>12.0</td>
<td>7</td>
</tr>
<tr>
<td>Polyethyleneimine</td>
<td>80</td>
<td>9.0</td>
<td>11</td>
</tr>
<tr>
<td>(NH₄)₂SO₄ extract (150 mM)</td>
<td>40</td>
<td>8.0</td>
<td>20</td>
</tr>
<tr>
<td>(NH₄)₂SO₄ extract (300 mM)</td>
<td>15</td>
<td>6.8</td>
<td>45</td>
</tr>
<tr>
<td>(NH₄)₂SO₄ ppt.</td>
<td>10</td>
<td>5.3</td>
<td>53</td>
</tr>
<tr>
<td>MgCl₂ ppt.</td>
<td>4</td>
<td>3.0</td>
<td>75</td>
</tr>
<tr>
<td>Spermidine ppt.</td>
<td>2.5</td>
<td>2.4</td>
<td>96</td>
</tr>
</tbody>
</table>

*Estimated from Coomassie Blue stain.*

*Estimated from Western blot, using *E. coli* recA protein as standard.*
Higher plant chloroplast RecA


(Received June 10, 1997; Accepted September 29, 1997)