Knockout Targeting of the Drosophila Nap1 Gene and Examination of DNA Repair Tracts in the Recombination Products

Susanne Lankenau,*† Thorsten Barnickel,* Joachim Marhold,‡ Frank Lyko,† Bernard M. Mechler‡ and Dirk-Henner Lankenau*†,†

*Department of Zoology, University of Heidelberg, D-69120 Heidelberg, Germany and †Research Group Epigenetics and ‡Department of Developmental Genetics, Deutsches Krebsforschungszentrum, D-69120 Heidelberg, Germany

ABSTRACT

We used ends-in gene targeting to generate knockout mutations of the nucleosome assembly protein 1 (Nap1) gene in Drosophila melanogaster. Three independent targeted null-knockout mutations were produced. No wild-type Nap1 protein could be detected in protein extracts. Homozygous Nap1Δ51 knockout flies were either embryonic lethal or poorly viable adult escapers. Three additional targeted recombination products were viable. To gain insight into the underlying molecular processes we examined conversion tracts in the recombination products. In nearly all cases the I-SceI endonuclease site of the donor vector was replaced by the wild-type Nap1 sequence. This indicated exonuclease processing at the site of the double-strand break (DSB), followed by replicative repair at donor-target junctions. The targeting products are best interpreted either by the classical DSB repair model or by the break-induced recombination (BIR) model. Synthesis-dependent strand annealing (SDSA), which is another important recombinational repair pathway in the germline, does not explain ends-in targeting products. We conclude that this example of gene targeting at the Nap1 locus provides added support for the efficiency of this method and its usefulness in targeting any arbitrary locus in the Drosophila genome.

The completion of the genome sequence provides unlimited access to all genes of Drosophila melanogaster (Adams et al. 2000). Nevertheless, despite nearly a century of Drosophila genetics, there are many Drosophila genes for which corresponding mutants are still unavailable. Means to overcome the drawback had been site-selected transposon mutagenesis (Ballinger and Benzer 1989; Kaiser and Goodwin 1990) and RNA-mediated interference (RNAi; Kennerdell and Carthew 1998). While transposon mutagenesis involves elaborate PCR screening, RNAi generates only genespecific phenocopies of loss-of-function mutations and does not always cause a true null phenotype. Therefore, methods of gene knockout targeting have been developed. Drosophila gene targeting is accomplished by two alternative techniques (Gloor et al. 1991; Rong and Golic 2000). Both take advantage of the fly’s endogenous homologous recombination machinery in the germline. One method utilizes a P-element-induced double-strand break in a target gene, which then is repaired from an ectopic donor construct by means of synthesis-dependent strand annealing (SDSA; Nassif et al. 1994). Pinduced gap repair was developed by Engels and colleagues (Gloor et al. 1991; for a review see Lankenau 1995; Lankenau and Gloor 1998). The drawback of Pinduced gap repair is the need for a suitable P-element tightly linked to the gene to be modified. Unfortunately, not all (i.e., 20%) of the Drosophila genes are available as P insertions (Spradling et al. 1999). The other technique resembles knockout targeting in mouse embryonic stem cells (Capelchi 1989a; Rong and Golic 2000). Rong and Golic’s approach now can target a mutation to any arbitrary locus in the Drosophila genome (Rong et al. 2002). The method involves four components: (1) a transgene that expresses a heat-shock-inducible site-specific recombinase (FLP); (2) a second transgene that expresses a heat-shock-inducible site-specific endonuclease (I-SceI); (3) a transgenic donor vector that contains recognition sites for both enzymes in addition to the white gene as a positive selection marker; and (4) the native wild-type target gene. Through heat shock, the FLP recombinase excises a circular episome containing the white marker gene and an in vitro modified donor gene. The extrachromosomal DNA molecule is linearized within the modified donor gene through the activity of the heatinduced I-SceI endonuclease.

Rong and Golic (2000) pioneered the new approach first at the yellow gene. Targeting additional genes at central chromosomal positions demonstrated that arbitrary loci can be modified (Rong and Golic 2001; Rong et al. 2002; Seum et al. 2002). As a further example of gene targeting in Drosophila, we set out to target an essential gene relevant for our future research. Because of its chromosomal location (1.4 Mb to the telomere),

Corresponding author: Department of Zoology, Im Neuenheimer Feld 230, D-69120 Heidelberg, Germany. E-mail: d.lankenau@uni-hd.de

Genetics 165: 611–623 (February 2003)
the expectation of obtaining a visible phenotype, and the lack of null mutations, we chose to target the nucleosome assembly protein 1 (Nap1) gene (Iro et al. 1996). Drosophila NAP1 has been shown to be required for chromatin assembly in vitro and was found to be associated with core histones H2A and H2B as a chaperone in embryonic extracts (Iro et al. 1996). The only in vivo data were obtained from the yeast Nap1 homolog, but these indicated a role in cell cycle regulation rather than in chromatin assembly (Kellogg and Murray 1995). The generation of a Nap1 knockout mutant would also be a first step toward determining its in vivo function in Drosophila. Six targeted knockout mutations (Nap1<sup>B</sup>) with three complete Nap1 gene knockouts were obtained. Three homozygous knockout mutations expressed no Nap1 protein and showed a semilethal phenotype.

Three additional targeted Nap1<sup>B</sup> recombinants were viable. To understand the molecular differences between the six targeted Nap1<sup>B</sup> mutants we performed a detailed molecular analysis of these recombination products. We constructed the targeting vector such that the Nap1 donor gene included protein-function-destroying mutations, which simultaneously introduced five molecular repair-tract markers. Because end-in-targeting produces target-gene duplications, each Nap1 copy and the five corresponding tract markers were duplicated at each targeted event. Thus from the six targeting events we monitored 60 marker positions. The 18-bp I-SceI endonuclease site where the Nap1 donor was cleaved for double-strand break (DSB)-induced targeting was in all but one case replaced by wild-type Nap1 sequence. This indicated that exonuclease processing at the DSB and replicative repair at the donor-target biotinylated by random priming and used as a probe for I-SceI cleavage was in all but one case replaced by wild-type Nap1 sequence. This indicated that exonuclease processing at the DSB and replicative repair at the donor-target junctions had taken place. The results are discussed mechanistically with relation to major models of recombinational DNA repair. Either the classical DSB repair model involving double Holliday junctions (Szo</text>stak et al. 1983) or the break-induced recombination (BIR) model (Malkova et al. 1996) best explains the tract data. We conclude that this example of targeted gene modification at the Nap1 locus was efficient and confirmed the expectation that any arbitrary locus can be targeted. The described procedure is powerful and it clearly represents a general method for targeted mutagenesis.

**MATERIALS AND METHODS**

**Drosophila:** Genetic symbols are defined in standard reference works (Lindsley and Zimm 1992; FlyBase 1999). Genomic DNA sequences of the Nap1 gene and flanking sequences were accessed via http://www.ncbi.nlm.nih.gov/PMG/i</text>ngenomes/Genomes/7927.html and derived from the assembled and annotated genome sequence of D. melanogaster as available in GenBank (Myers et al. 2000). Fly stocks used for the targeting screen were a gift from Kent Golic (Rong and Golic 2001; Rong et al. 2002):

1. y<sup>w</sup> p(y<sup>w</sup>, 70 FLP/4 P[<i>v<sup>+</sup></i>], 70 I-SceI/2B Sco/S<sup>2</sup> Cyo
2. w<sup>1118</sup> p(y<sup>w</sup>, 70 FLP/10; + (strong constitutive expression of FLP).

All experimental fly stocks were tested for absence of endo</text>biotic, cytoplasmically inherited Wolbachia bacteria using published Wolbachia <i>pipientis</i>-specific 16S RNA PCR primers (O'Neill et al. 1992).

**Construction of donor plasmid and microinjection:** On the basis of the genome DNA sequence of Drosophila, recombinant PCR (Higuchi 1990) was used to generate a 4.275-kb Nap1-containing fragment from genomic DNA of Canton-S wild-type flies. The following oligonucleotides were used:

1. (Acc65I) 5’ CGCGGTATTCTAACGCTTAAAGCATTATGGTCAAGCGCTT 3’
2. (NotI) 5’ CGCGGTATTCTAACGCTTAAAGCATTATGGTCAAGCGCTT 3’
3. (I-SceI) 5’ TAGGAGATACAGCGTGTAATCTGGTCTGTCGATATTCC 3’
4. (I-SceI) 5’ ATTACCTCGTGGTATCCCTATgaggacccgtagaga 3’
5. (HindIII) 5’ GGTGCAAGAATCTCGGGGTCAGATGAAGCTTCGGCTGGGCGGTCATTG 3’
6. (HindIII) 5’ CAATGGACCCCGAGCGCCGAAGCGATTCAAGCTGACCGGATCTCGGGCTGGGCGGTCATTG 3’
7. (BclI) 5’ GCTTCTGGGATGCTGTCGATCACGATAGGGTGTTACG 3’
8. (BclI) 5’ GATGACACTTATGTCGATCAGACCATCGCCGAAGGAGCG 3’.

The isolated DNA fragment finally encompassed five molecular markers (Xbal–, HindIII+, I-SceI+, BclI+, and SalI–), which destroyed the open reading frame (ORF) structure to render the knockout mutant viable. To understand the molecular differences between each of the five targeted knockout duplicates, each Nap1 copy and the five corresponding tract markers were duplicated at each targeted event. Thus from the six targeting events we monitored 60 marker positions. The 18-bp I-SceI endonuclease site where the Nap1 donor was cleaved for double-strand-break (DSB)-induced targeting was in all but one case replaced by wild-type Nap1 sequence. This indicated that exonuclease processing at the DSB and replicative repair at the donor-target junctions had taken place. The results are discussed mechanistically with relation to major models of recombinational DNA repair. Either the classical DSB repair model involving double Holliday junctions (Szo</text>stak et al. 1983) or the break-induced recombination (BIR) model (Malkova et al. 1996) best explains the tract data. We conclude that this example of targeted gene modification at the Nap1 locus was efficient and confirmed the expectation that any arbitrary locus can be targeted. The described procedure is powerful and it clearly represents a general method for targeted mutagenesis.

**Analysis of recombinant flies:** Polytene in situ hybridization: The <i>w<sup>+</sup></i> gene from the pTV2 vector (Rong et al. 2002) was biotinylated by random priming and used as a probe for in situ hybridization to polytene chromosomes as described (Lin 1993; Figure 2A).

**Southern blot:** Southern blots were performed according to standard protocols (alkaline transfer onto positively charged nylon membranes; Ausubel et al. 1995) and hybridized to a biotinylated Nap1 probe (Figure 2C). Detection using streptavidin-alkaline phoshatase and CDP star as substrate was performed using the detector system (KPL, Gaithersburg, MD).

**PCR analysis:** Genomic DNA of heterozygous flies was used to track the molecular markers by performing PCR with three different primer combinations and subsequent restriction digests (Figure 4, A and B). Primer pl primes specifically to the 5’ genomic region upstream of the distal Nap1 duplicate; the PCR product of primers pl and piII is therefore used to follow the markers on the distal (i.e., telomeric) side. Primer piIV is specific for the 3’ genomic region downstream of the proximal Nap1 duplicate, so that the PCR product of primers piII and piIV serves to characterize the proximal Nap1 duplicate. Primers II and III hybridize to all three different Nap1 copies (distal, proximal, and wild type). This PCR product was analyzed by a subsequent double digest with HindIII and BclI to prove the presence of the introduced restriction sites and to show the absence of the donor construct at the same time. Primers are as follows:

- pl: CGCGGACCTGCAACCATGATACTTTAGGA
- plII: CGCGGGCACACCATGATACTTTAGGA
- piII: CGCGGACCTGCAACCATGATACTTTAGGA
- piIV: CGCGGACCTGCAACCATGATACTTTAGGA

Genomic DNA of heterozygous flies was used for PCR analysis:

1. Southern blot:
2. PCR analysis: Genomic DNA of heterozygous flies was used to track the molecular markers by performing PCR with three different primer combinations and subsequent restriction digests (Figure 4, A and B). Primer pl primes specifically to the 5’ genomic region upstream of the distal Nap1 duplicate; the PCR product of primers pl and piII is therefore used to follow the markers on the distal (i.e., telomeric) side. Primer piIV is specific for the 3’ genomic region downstream of the proximal Nap1 duplicate, so that the PCR product of primers piII and piIV serves to characterize the proximal Nap1 duplicate. Primers II and III hybridize to all three different Nap1 copies (distal, proximal, and wild type). This PCR product was analyzed by a subsequent double digest with HindIII and BclI to prove the presence of the introduced restriction sites and to show the absence of the donor construct at the same time. Primers are as follows:

- pl: CGCGGACCTGCAACCATGATACTTTAGGA
- plII: CGCGGGCACACCATGATACTTTAGGA
- piII: CGCGGACCTGCAACCATGATACTTTAGGA
- piIV: CGCGGACCTGCAACCATGATACTTTAGGA
Sequencing: Distal- and proximal-specific PCR fragments (see PCR analysis) were gel purified and PCR fragments spanning the region of the Nap1 construct containing the I-Sce site were generated using primers pV and pVI (pV, CTCTGACCTCACTATGGGCCAGCAAATCTGA; pVI, CTCTCTAGACTACGCCTGACGACCTGAAATATCGA). The PCR fragments were directly sequenced using an ALF sequencer.

Western analysis: Protein extracts were prepared from Drosophila Oregon-R wild-type females and from homozygous mutant Nap1 females or dissected ovaries. Equal amounts of protein were separated on a 15% polyacrylamide gel and blotted using standard procedures. As primary antibody we used anti-NAP1 (Li et al. 1999) and anti-Rp-40 as a loading control (Torok et al. 1999). As secondary antibody we used anti-rabbit IgG coupled to HRP (Dianova). Secondary antibody was visualized by enhanced chemiluminescence (Perkin-Elmer, Norwalk, CT) and exposure to X-ray films (Figure 3, A and B).

Immunostaining of ovaries and confocal laser scanning microscopy: Ovaries were dissected from wild-type (Oregon-R) and homozygous mutant Nap1 females and the tissue was fixed for 40 min in 4% formaldehyde, 0.5% Tween 20, and 1× PBS followed by three washes in 1× PBS. After blocking in 1× PBS, 1% BSA, 0.1% Tween 20, ovaries were incubated with anti-NAP1 antibody. After three washes with blocking solution, incubation with the secondary antibody (anti-rabbit Cy5; Dianova) followed in combination with 10 μg/μl propidium iodide and 100 μg/μl RNase A to visualize DNA and phalloidin FITC to visualize F-actin. Finally, ovaries were washed in blocking solution, in 1× PBS, 1% BSA, and in Slow Fade Light Component C (Molecular Probes, Eugene, OR). Preparations were mounted in Slow Fade Light Component A and analyzed with a Zeiss LSM410 confocal microscope (Figure 3, C and D).

RESULTS

Design of the Nap1 knockout construct and the genetic screen: We used the D. melanogaster Nap1 gene mRNA sequence (GenBank accession no. U39553; Iro et al. 1996) to identify genomic DNA sequences flanking this gene. We identified a 4.5-kb fragment within a Drosophila scaffold section of the complete genomic sequence (accession no. AE003462). The intron/exon structure of Nap1 and its location within the 4.5-kb genomic fragment was roughly confirmed with the GENSCAN software. On the basis of this sequence we designed eight oligonucleotide primers, which were simultaneously used for three purposes:

1. The isolation of a 4.3-kb PCR fragment from genomic DNA of wild-type Drosophila flies containing the Nap1 gene in a central position.
2. The introduction of mutations into the wild-type Nap1-coding region, which destroys the function of its protein product. The mutations flank an introduced I-Sce endonuclease cutting site on both sides (Figure 1A).
3. The mutations further introduced HincIII and BclI as artificial restriction endonuclease cutting sites and a XhoI and a SalI site were destroyed. These sites were used to track the DNA repair activities responsible for targeted gene knockout events in vivo (Figure 1A).

With these eight primers we isolated and simultaneously mutagenized a 4.3-kb Nap1 fragment from genomic DNA by recombinant PCR (Higuchi 1990). The fragment was introduced into the pTV2 P-element vector (Figure 1A; Rong et al. 2002), and transgenic flies were established containing the Nap1/pTV2 donor construct integrated in a third balancer chromosome (Tm6, Ubx). These flies were used to screen for knockout events as shown in Figure 1B. Figure 1C shows the expected result of a knockout (“knock-in”) event at the Nap1 locus. The absence or presence of the I-Sce cutting site was not foreseeable. However, on the basis of published conversion frequency studies (Gloor et al. 1991; Preston and Engels 1996) we expected exonucleolytic processing of the I-Sce-induced DSB to result in wild-type sequence at the donor/target junctions.

We carried out three targeting screens as shown in Figure 1B, each using slightly different heat-shock conditions. Table 1 shows the results of the screens. With the red eye-color phenotype as a positive marker we identified eight recombination events with six targeted events and three independent null-mutation Nap1 knockout products.

Verification of targeted knockout events: Two of the eight recombination events genetically segregated with the X chromosome and were not further analyzed. The three remaining recombination events segregated with the second chromosome where the Nap1 gene is located. None of them expressed a mosaic red/white eye-color phenotype when combined with a constitutively expressed FLP recombinase source (Rong and Golic 2001). Nonsyngamy confirmed that they were good candidates for targeted recombination events. Using the w51 gene as a probe for in situ hybridization to polytene chromosomes, we located the w51 gene at the Nap1 locus on the second chromosome (Figure 2A). Southern blot analysis confirmed the expected knockout duplications (Figure 2, B and C). Genomic DNA of heterozygous Nap1+/KO flies was digested either with BclI, diagnostic for the distal part (Figure 2B, bottom), or with HindIII, diagnostic for the proximal part (Figure 2B, top) of the predicted knockout duplication. In addition to the 12.9-kb BclI fragment diagnostic for the Nap1 wild-type gene, the recombinant flies Nap1KO1, Nap1KO2, Nap1KO3, and Nap1KO6 showed two bands (2.3 and 8.5) derived from the targeted Nap1 locus, identifying the incorporated BclI site. Next to the wild-type fragment (8.3 kb, Figure 2, B and C) the HindIII site of the targeted Nap1 duplication was found in Nap1KO1–Nap1KO4 but not in Nap1KO5 and Nap1KO8 (11.2- and 4.5-kb vs. 16.7-kb fragments, Figure 2, B and C). The DNA blot results indicated that the Nap1KO1 and Nap1KO5 alleles represented the anticipated knockout products. Because Nap1KO1–Nap1KO4 did not reveal identical fragment patterns in the DNA blot analysis, the recombination tracts of the six knockout events were studied to gain insight into the underlying DNA repair pathway (see below).

Absence of NAP1 56-kD protein in knockout mutants leads to lethality: We performed Western blot analyses
to further confirm successful targeting and to show if Nap1 knockout flies expressed NAP1 protein (Figure 3, A and B). The proximal construct of the targeted Nap1 duplication was designed by introduction of a HindIII site such that the ORF of the Nap1 knockout product would be destroyed close to the initial methionine, thus resulting in no protein product (Figure 1, A and C). The distal part of the knockout duplication introduced a BclI site to destroy the ORF in the middle of the Nap1 gene and to damage its nuclear localization signal (Figure 1, A and C). By Western blot analysis, we found that no 56-kD wild-type NAP1 protein can be detected either in whole fly protein extracts or in ovaries of homozygous Nap1KO and Nap1KO2 knockout flies (Figure 3, A and B). Nap1KO and Nap1KO2 were not further analyzed because too few homozygous offspring were obtained, which is possibly due to second-site mutations elsewhere in the genome. Homozygous Nap1KO and Nap1KO2, however, produced wild-type NAP1 protein (Figure 3B). Southern and repair tract analyses (see below) revealed that these alleles as well as Nap1KO were targeted recombination events in which one of the two Nap1 duplicates remained wild type (Figures 2, B and C, and Figure 4).

The Drosophila ortholog of Nap1 in humans was identified as a chaperone factor involved in the assembly of nucleosomes (Ishimi et al. 1984; Ito et al. 1996). Because nucleosome assembly is essential for eukaryotic cells, we observed absence of wild-type 56-kD NAP1 protein in either in whole fly protein extracts or in ovaries of homozygous Nap1 knockout flies expressed NAP1 protein (Figure 3, A and B). The viability of the Nap1KO mutant was investigated in Nap1KO/T;Cy;Roi Tb flies. The hatch rate of homozygous Nap1KO flies was by a factor of 5 lower than expected (6.3%), indicating a lethality analogous to perinatal lethal phenotypes of mice. No unusual pupal lethality was observed, suggesting that the limited hatch rates were due to larval or embryonic lethality. Low adult hatch rates were also observed for homozygous Nap1KO flies (7%). In the case of Nap1KO3, only a single fly hatched, and, similarly Nap1KO4 did not produce any homozygous adult flies, which is presumably due to epistatically interacting second-site mutations elsewhere in the genome. Consistent with their expression of wild-type NAPI protein (Figure 3B), Nap1KO and Nap1KO2 showed the expected percentage of homozygous flies (30%). Together with the Southern analysis this result indicated that one copy constitutive FLP expression, offspring flies that do not carry a targeted insert but still contain an unexcised donor on the TM6, Ubx balancer are white eyed with rare colored spots. (C) Expected Nap1 knockout targeting duplication (knock-in). Nap1 knockout sequences (shaded rectangle) are marked by BclI (BeI+) and HindIII (Hd+). Only one FRT site remains, stabilizing the w-hs gene under constitutive FLP expression.

tel., telomere; cent., centromere.

Figure 1.—Strategy for Nap1 knockout targeting. (A) Structure of the donor targeting vector. The pTV2 plasmid contains a nonautonomous P transposable element. The P vector carries the white-hs (w-hs) marker gene, two FRT sites in direct orientation for FLP-mediated episomal excision, and the in vitro mutagenized Nap1 gene. The 4.3-kb genomic fragment containing the Nap1 gene was inserted into pTV2 at NotI and at Acc65I. The intron-exon structure is as indicated. Nap1 was mutagenized by recombinant PCR. The center of Nap1 contains an I-SceI site. On the left side of the I-SceI site, the inserted HinIII site (Hd+) introduces a reading frameshift into the open reading frame of Nap1 at the sixth most N-terminal amino acid. On the right side of the I-SceI site, the inserted BclI restriction site introduces a reading frameshift that truncates the Nap1 protein and destroys the nuclear localization signal. Two other restriction enzyme markers, XhoI and Sall, were deleted from the donor construct to support recombination tract analysis (XhoI–, Sall–). The intron/exon structure and its transcriptional orientation of Nap1 are indicated. (B) Cross to generate a targeted gene knockout of Nap1. The TM6, Ubx balancer chromosome (TM6) in a G1 female fly contains the transgenic Nap1P element donor construct. The wild-type Nap1 gene (Nap1+) is located on both second chromosome homologs (only one copy is shown, as a rectangle). One of these homologs contains two heat-shock inducible transgenes: FLP recombinase (shaded circle) and I-SceI endonuclease (solid circle). The X chromosomes are homozygous for a white (w) mutation. Upon heat shock during larval development, FLP and I-SceI produce the extrachromosomal targeting molecule. The female fly is crossed to a transgenic male with strong constitutive expression of FLP recombinase (70FLP, dark-shaded circle). Because of the efficiency of
The induction of mutations within genes is tightly coupled to our basic understanding of gene function. Precisely defined mutations are therefore a prerequisite to analyze the function of genes and their phenotypic impacts. Unfortunately, although the Drosophila genome has been mutated at very high density, no mutants...
**Figure 2.**—Mapping and verification of targeted recombination events. (A) Localization of \( w^{hs} \) at the \( Nap1 \) locus by polytene chromosome in situ hybridization. Chromosomes from flies homozygous for the targeted viable \( Nap^{P50S} \) allele were probed with labeled \( white-hs \) gene DNA. Two signals were detected: one at the 3C locus, which is the endogenous \( white \) locus, and the other at 60A, which is the \( white-hs \) insertion at the targeted \( Nap1 \) locus. (B) Map of genomic restriction fragments diagnostic for targeted knockout mutations KO1–KO6. Centered is a sequence-derived \( HindIII \) (H)- and \( BclI \) (B)-based restriction map of the wild-type \( Nap1 \) locus. The upper two restriction maps indicate \( HindIII \) digestion patterns with the 4.5-kb fragment diagnostic for successful targeting of the proximal part of the anticipated targeted knockout gene. The three bottom restriction maps indicate \( BclI \) digestion patterns with the 2.3-kb fragment diagnostic for successful targeting of the distal part of the anticipated targeted \( Nap1 \) gene. Arrows indicate positions of the I-SceI cutting site prior to the targeting event. (C) Southern blot analysis. Genomic DNA of wild-type flies and of heterozygous recombinants (\( Nap1^{KO1–KO6} \)) was digested with either \( BclI \) (left) or \( HindIII \) (right). A wild-type 4.3-kb genomic fragment containing the \( Nap1 \) locus was used as a biotinylated probe. The 2.3- and 8.5-kb \( BclI \) fragments of the recombinant flies \( Nap1^{KO1}, Nap1^{KO2}, Nap1^{KO5}, \) and \( Nap1^{KO6} \) identify the incorporated \( BclI \) site at the targeted \( Nap1 \) locus. In KO1–KO4 but not in KO5 and KO6, two \( HindIII \) fragments of 12.2 and 4.5 kb prove the introduction of the \( HindIII \) site in the proximal \( Nap1 \) duplicate.

are known for a significant fraction of genes. The \( Nap1 \) gene is one such example where a mutant null allele has not been available. While traditional mutagenesis procedures are based on phenotypic screening (with complex and time-consuming genetic crosses to look for recessive phenotypes), gene targeting requires no prediction of the mutant phenotype. There are convincing arguments that the technique of RNAi (which also does not require knowledge of a phenotype) is simpler on a practical level and therefore is better suited than targeted mutagenesis to overcome the lack of mutants (Carthew 2001). However, RNAi generates only gene-specific phenocopies of null mutations but does not always cause a true null phenotype (Adams and Sekelsky 2002). Therefore, altering specific endogenous genes within the metazoan germline represents a foundation for the highest possible level of experimental control over a particular locus of interest. The establishment of mouse embryonic stem (ES) cell lines, techniques transforming vector DNA into ES cells, and advanced methods to produce chimeras and completely ES-cell-derived fetuses trailed metazoan targeting and made the mouse a leading model organism (Thomas and Capecchi 1987; Capecchi 1989a,b; Joyner 1995).
null allele for the Drosophila Nap1 repair pathway. gene as well as upstream and downstream regulatory Nap1 of Nap1 (1) screening and verification of a targeted knockout ity of positively selected (red eye-color marker) flies basal concentration of NAP1 protein (green) observed in wild- map to the X chromosome. Neither of these integra-type cells is absent (arrowheads). tions were targeted events at the endogenous

Figure 3.—Analysis of NAP1 protein in targeted flies. (A) Western blot analysis of wild-type and homozygous Nap1KO knockout flies. Total protein extracts were obtained from adult female flies and anti-NAP1 antibody was used (Li et al. 1999). Anti-p40 antibody served to control for equal loading (bottom; Torok et al. 1999). (B) Western blot analysis of ovaries from wild-type and homozygous Nap1KO mutant flies. The genotypes were confirmed by PCR using genomic DNA from the car-
tases as substrate. (C) Comparative immunolocalization of NAP1 protein (green) in the follicle cell layer surrounding egg chambers of wild-type and Nap1 knockout mutant ovari-
les. DNA is stained with propidium iodide (red). Egg cham-
bers were stained with Alexa488-labeled phalloidin to reveal actin-rich structures (blue). (D) The same antibody applied to the follicle cell layer of egg chambers from a wild-type and a Nap1 knockout mutant fly. Actin (here, red) reveals the apical part of the follicle cells. In the knockout mutant the basal concentration of NAP1 protein (green) observed in wild-
type cells is absent (arrowheads).

Drosophila has suffered so far from the lack of an equally efficient gene-targeting method. Only recently a promising method in Drosophila and for its applicability to any arbitrary locus. The study had three subgoals: (1) screening and verification of a targeted knockout event at the Nap1 locus, (2) a partial functional analysis of Nap1, and (3) examination of recombination tracts for an initial understanding of the underlying DNA repair pathway.

Generation of targeted Nap1 mutant alleles: A mutant null allele for the Drosophila Nap1 gene has not been available so far. Starting from genomic DNA of Canton-S wild-type flies, we used recombinant PCR (Higuchi 1990) to generate a mutagenized 4.3-kb Nap1-containing fragment holding all knockout mutations and tract markers. The altered Nap1 fragment was subse-
sequently cloned into the pTV2 vector (Rong et al. 2002) and transformed into the germline of w^{11601} embryos. We found that starting from genomic DNA substantially reduces both the time and the cost associated with traditional library screening methods and it is at least as efficient as the construction of mouse-targeting vectors using methods such as recombination protein E (RecE) and RecT-mediated cloning (Zhang et al. 1998, 2002). We then used the ends-in method of gene targeting (Rong and Golic 2000) to disrupt the Nap1 locus. In our experiments, the targeting efficiency was dependent on the experimental conditions. The three screens used an identical donor insertion on a third chromosome balancer. Variations of the heat-shock conditions resulted in a significant improvement (Table 1). While the initial screen was not successful (no targeting event in 1100 single crosses with 165,000 scored chromo-
somes), the final screen was very efficient (five targeting events in 450 single crosses with 80,300 scored chromo-
somes; Table 1) This efficiency (∼1 in 16,000 gametes) matched the reported efficiencies (∼1 in 500 gametes to ∼1 in 30,000 gametes) of Rong et al. (2002). We are not sure about the reasons for the initial failure, but, as we used an uncommon brand of thick-walled glass bottles for culturing flies, the duration (1 hr) of the initially applied heat shock may not have been sufficient to produce enough donor-construct excisions. Extend-
ing the duration of the heat shock and repeated heat shocks increased eye-color mosaicism and led to the targeted products (Table 1).

In addition to the six targeted events captured in our screen, we found only two donor integrations that did not map to the Nap1 locus (at polytene-chromosome map position 60A of the second chromosome) but did map to the X chromosome. Neither of these integra-
tions were targeted events at the endogenous white locus caused by homologous recombination with donor-inte-
0 internal white sequences. Targeting in mouse ES cells is often accompanied by high ratios of nontargeted insertions. Actually, when positive-negative selection is not used in ES cell transformation, the bulk of positively selected ES cell clones contain nontargeted insertions outnumbering targeted events by orders of magnitude (Man-
sour et al. 1988; Bollag et al. 1989). For Drosophila, Rong et al. (2002) report that in their screens the major-
ity of positively selected (red eye-color marker) flies were targeted. Our results are consistent with this.

A drawback of the insertional targeting procedure (also called “knock-in targeting”) is that the mutated gene as well as upstream and downstream regulatory sequences are duplicated. This might affect gene func-
tion in an unpredictable manner. Phenotypic and func-
tional analysis of a targeted gene may be impaired fur-
Figure 4.—Recombination tract analysis. (A) PCR analysis of genomic DNA from heterozygous recombinant KO flies. The boxed area shows the predicted genomic structure of two alleles of the Nap1 locus. The top map designates the wild-type Nap1 allele, the bottom map designates a full knockout allele. w-hs represents the white-hs positive selection marker gene. H and B designate the restriction enzymes HindIII and BclI. Open circles indicate the presence and solid circles the absence of these sites. Diagnostic primer combinations (pI–pIV) are outside of the boxed area. The expected PCR products and their restriction digest products using BclI and HindIII are indicated. Italic letters (a–l) indicate restriction fragments corresponding to sizes in B. (B) Detection of all predicted fragments. Here, only results for Nap1KO3 are shown. Italic letters correspond to predicted fragment sizes in A. The internal structure of Nap1-targeting events was confirmed by a BclI/HindIII double digest (lane 1); lane 2 is a control from a fly containing the Nap1 donor construct on the third chromosome. The 0.9-kb fragment (fragment e in B) diagnostic for the donor construct is missing in all KO flies (except KO4 where it derives from the proximal Nap1 duplication). The distal region of the knockout rearrangement was diagnosed using BclI (lane 3, fragments h and i) and the proximal region was confirmed using HindIII (lane 4, fragments h and i). (C) Structure and recombination tracts of six knockout events at the Nap1 locus. Recombination tract analysis was performed as described in A and B. Restriction enzyme markers were introduced into the donor vector, which are reciprocal to the wild-type Nap1 target gene (X, XhoI; H, HindIII; I, I-SceI; B, BclI; S, SalI). Markers correspond to wild-type (solid circle) and mutagenized (open circle) Nap1 sequences. Presence (open star) or absence (solid star) of the I-SceI cutting site is indicated. Primer combinations pV and pVI were used to amplify each I-SceI site for subsequent sequencing. (a) Knockout alleles Nap1KO1 and Nap1KO2 revealed identical recombination tracts. These recombination products represent the predicted knockout structure. (b) Nap1KO3 represents the third complete knockout event. The proximal duplication is completely derived from the donor construct except for the I-SceI cutting sequence. (c) Incomplete knockout allele Nap1KO4. The distal duplication is wild type. The proximal duplication is identical to the donor construct. (d) Nap1KO5 and Nap1KO6 represent partial knockout alleles. The proximal duplication is entirely wild type.

If the gene is small (e.g., 1 kb) and tightly flanked by neighboring genes. Because the donor sequence should be long (e.g., 4–5 kb) for homology requirements during recombinational DNA repair, it cannot always be avoided that the flanking genes are duplicated as well. This would make functional studies unreliable. These difficulties are now overcome by the use of a meganuclease cutting site (I-CreI) within the integrated pTV2-vector (Rong et al. 2002). A DSB can now be produced between the two copies of the duplicated target gene, and the DSB is repaired through single-strand annealing (SSA) repair such that a single copy of the targeted gene remains. The whole procedure therefore resembles a “hit-and-run” approach and satisfies the highest standards of gene targeting (Adair and Nairn 1995).

A partial functional analysis of NAPI: Because the Nap1 gene is large, the donor did not possess additional
genes whose altered expression pattern might affect a functional analysis of Nap1. Homozygous Nap1 flies did not express detectable amounts of NAP1 protein (Figure 3). We found that first-generation homozygous mutant Nap1 flies (derived from heterozygous parents) developed until the adult stage, albeit at sub-Mendelian...
frequencies. These flies showed reduced viability, but they were weakly fertile and gave rise to a second generation of homozygous flies. In these flies, the phenotype became much stronger and more penetrant. The few escaper flies that developed to the adult stage showed impaired development and died a few days after eclosion. A functionally strong maternal component of Nap1 expression at low concentrations (undetectable by Western blot) is probably sufficient to sustain relatively normal development in a significant fraction of homozygous mutant flies derived from heterozygous parents. Only after depletion of the maternally supplied components does the lethal phenotype become fully penetrant. The lethal phenotypes therefore were similar to the phenotype of other gene products thought to be important in nucleosome remodeling. For example, imitation switch (ISWI) homozygotes, where ISWI is the catalytic subunit of three essential chromatin-remodeling complexes NURF, ACF, and CHRAC, die as late larvae or integrations. The most frequent DSB repair mechanism (ISWI) homozygotes, where ISWI is the catalytic into a true targeted single-copy knockout mutation.

Switch phenotype of other gene products thought to be important in nucleosome remodeling. For example, imitation switch (ISWI) homozygotes, where ISWI is the catalytic subunit of three essential chromatin-remodeling complexes NURF, ACF, and CHRAC, die as late larvae or early pupae (Varga-Weisz and Becker 1998; Deuring et al. 2000). The Nap1 knockout mutants may therefore point toward related functions of NAP1.

Recombination tract analysis: In this study we engineered frameshift point mutations within the coding sequence of the Nap1 gene, which blocked protein expression in the three knockout alleles Nap1\(^{\text{P03}}\), Nap1\(^{\text{P02}}\), and Nap1\(^{\text{P01}}\). Simultaneously these mutations served as markers so that we could monitor conversion tracts in the recombination products over a sequence distance of 1.5 kb. Because ends-in targeting generates a duplication of the donor DNA at the target locus, all markers are represented twice in a given targeting product. With the aid of the markers we identified four different types of recombinant flies among the six targeted recombination events that needed explanation (Figure 4C): Ends-in targeting is triggered by a linearized donor episome, whose DNA ends are sensed by the cell as double-strand DNA damage. The artificial, I-Sce-induced break is thought to stimulate the DNA repair machinery of the cell. The donor construct contained the I-Sce cutting sequence located in the center of the Nap1 gene. Only one of the eight individual recombinant donor duplicates (Figure 4C) still contained an intact I-Sce site (Nap1\(^{\text{P03}}\)), but not a single nucleotide of the I-Sce recognition sequence was found in the other seven copies. This result indicated that cellular exonucleolytic activities enlarge the DSB made by I-Sce at least beyond the length of this sequence (18 bp). Both single-strand DNA as well as dsDNA from the staggered I-Sce site cut were removed. This necessitated some trimming activity of the 3′ ssDNA overhang and the 5′ double-strand end of the break. Resection of ends is known to occur by 5′-to-3′ exonucleases or by an endonuclease associated with a helicase producing long 3′-ended tails (Paques and Haber 1999). The tails then are thought to invade a homologous template. In the course of DSB repair the gap is finally restored by DNA replication initiated at the invaded 3′ ends (Szostak et al. 1983). Rong et al. (2002) describe the practical implications of this exonuclease activity. They successfully introduced mutations to the target genes 400–1300 bp from the I-Sce site. In agreement with this, the marker positions in our experiments were located 376 bp (HindIII), 552 bp (BglII), 613 bp (Xhol), and 856 bp (SodI) from the DSB.

How do the recombination tracts observed in the Nap1 recombination products match known DSB repair pathways? Four prominent DSB repair pathways are relevant for this study: (1) nonhomologous end joining (NHEJ; Roth and Wilson 1988), (2) the classical DSB repair model (Szostak et al. 1983), (3) SDSA (Nassif et al. 1994), and (4) BIR (Malkova et al. 1996). A fifth DSB repair pathway, SSA, is important when a targeted knock-in duplication (as produced here) is transformed into a true targeted single-copy knockout mutation.

NHEJ is unlikely to play a role during targeted donor integrations. The most frequent DSB repair mechanism of metazoans, NHEJ was first recognized by Barbara McClintock in the early 1940s (McClintock 1987). It results in the ligation of broken DNA ends that share little or no homology with each other. This reaction is extremely efficient in eukaryotes, but the structures of the Nap1 knockout alleles (Figure 4C) cannot be explained by NHEJ. End joining left remnants of the I-Sce site behind, but after sequencing all target/donor joints we found only wild-type sequences at the Nap1 donor/target junctions, except at the proximal Nap1 copy of Nap1\(^{\text{P03}}\) where the I-Sce site remained intact (Figure 4C, c). NHEJ, however, may account for the two non-targeted Nap1 donor integrations on the X chromosome (Table 1).

Most likely, recombinational repair mechanisms are responsible for the targeted recombinants. Currently, the classical DSB repair model (Szostak et al. 1983) appears to be the most plausible explanation of the Nap1\(^{\text{P03}}\) recombinants (Figure 4). Figure 5 gives details on the DSB repair model for some of the Nap1\(^{\text{P03}}\) recombinants combined with the results of our tract analysis. The model readily explains the Nap1\(^{\text{P03}}\) and the Nap1\(^{\text{P02}}\) products by resolution of the double Holliday junctions before the tract markers have been copied by DNA synthesis or incorporated into heteroduplex. The tract patterns of Nap1\(^{\text{P05}}\), Nap1\(^{\text{P03}}\), and Nap1\(^{\text{P06}}\) can be explained readily as well if one assumes branch migration beyond one pair of diagnostic markers, followed by biased heteroduplex repair and subsequent branch resolution as indicated in Figure 5. The tract data encompass four diagnostic markers (excluding the nucleolytically processed I-Sce site) in each duplicate of the six targeted Nap1\(^{\text{P03}}\) products. Therefore, a total of 48 markers representing 24 tract marker pairs were analyzed. All pairs were continuous, which might not be a direct prediction of the “Szostak model.” Biased strand-specific mismatch repair might account for this result (Figure 5). For example, meiotic recombination is believed to occur via
Gene Targeting in Drosophila at Nap1

such events are not predicted by the conventional DSB repair model of Szostak (Szostak et al. 1983). According to the SDSA model (Nassif et al. 1994), the ends of the break independently undergo a genomewide homology search and it is thought that use of different, distantly located template sequences could be used for gene conversion (reviewed in Lankenau 1995; Lankenau and Gloor 1998). This prediction was subsequently demonstrated in yeast. It was shown that a broken plasmid can acquire genetic information from two different loci on two different chromosomes (Silberman and Kupiec 1994). A similar experiment involved templates and targets, where each end of a DSB on a plasmid was homologous to one of two overlapping truncated genes (LEU2) on two different chromosomes. Restoration of an intact LEU2 gene was made possible only by two separate strand invasion events and the subsequent annealing of DNA ends (Paques et al. 1998).

Interestingly, SDSA in Drosophila P-element-induced gap repair always resulted in nonreciprocal conversion between template DNA and target break (i.e., the template sequence was never altered; Gloor et al. 1991). This result indicates that SDSA is an improbable pathway for ends-in gene targeting because the invading strands of the linear donor episome would be finally unwound and returned to the broken strand (the linearized episome). Therefore, SDSA would not result in a targeted integration.

While SDSA is not the mechanism for targeted (knock-in) mutagenesis in Drosophila, the fact that the ends of a DSB undergo a genomewide homology search (Engels et al. 1994; Lankenau et al. 2000) may hold true for the ends of the broken pTV2 episome as well. Analogous to SDSA, invasion of the episomal ends would lead to replication. However, instead of unwinding and returning to the broken strand, further DNA synthesis might pause in a true replication fork. This might be captured and resolved by an endogenous replication fork during the following S phase of the cell cycle. The process resembles BIR (as shown in Figure 7), but it would not be restricted to the telomere as originally proposed for yeast (Malkova et al. 1996; Kraus et al. 2001). Engels (2000) proposed the mechanism for Drosophila partially to explain why targeting of the X-linked yellow gene was more efficient in females (with two X’s) than in hemizygous males. This, however, is not a relevant argument for BIR as it is now reported that targeting of autosomal genes is also more efficient in females (Rong et al. 2002). Further, as the models shown in Figure 7 include exonuclease activity and template switching, they are at least as complicated as those that explain the targeted Nap1 recombinants by classical DSB repair (Figures 5 and 6). However, the theoretical possibility of both broken ends invading different template strands may represent an option to test for the relevance
of a BIR-like process. One experiment might be to test whether targeted knock-in recombinants are frequently associated with crossing over between homologous chromosomes. An extended analysis of long recombination tracts in targeted genes may also give further evidence in favor or against one of the pathways.

We thank Kent Golic for the targeting kit fly stocks and the pTV2 vector DNA; David Kuttenkeuler, Julia Zander, Tobias Jursch, Kirsten Schubert, and Mingfa Li for discussions and technical help; and Carlos Flores for critically reading an early version of the manuscript. We also thank two anonymous referees for constructive suggestions on the classical DSB repair model. This work was supported by an EU grant (B.M.M., QLRI-CT-2000-00915), a DFG Emmy Noether fellowship (B.M.M., QLRI-CT-2000-00915), a DFG Emmy Noether fellowship (B.M.M., QLRI-CT-2000-00915), and a DFG grant (D.-H.L., LA711/5-1).

**LITERATURE CITED**


Ishimi, Y., J. Hiroseumi, S. Sato, K. Sugasawa, S. Yokota et al., 1984 Purification and initial characterization of a protein which


