Development of shuttle vectors for halobacteria

We have developed systems for genetic analysis of halophilic archaebacteria. These include (i) transformation systems which we have shown to work efficiently in at least three genera of halobacteria, (ii) shuttle vectors which can be selected in either halobacteria or E. coli, and (iii) host systems including hundreds of mutant halobacterial strains. We have applied these methods of genetic analysis in characterizing a halobacterial HMGCoA reductase gene (which is also the basis of our selection for antibiotic resistance in halobacteria) and clusters of genes involved in the biosynthesis of tryptophan, histidine, arginine and leucine. These techniques have also permitted us to position dozens of genetic markers on a physical map of the Haloferax volcanii genome. Other applications include the construction of halobacterial suppressor tRNA genes by site directed mutagenesis and cloning of genes encoding halobacterial membrane proteins.
OBJECTIVE: Develop transformation systems for halophilic archaea, construct a variety of useful halobacterial plasmid vectors including halobacteria - E. coli shuttle vectors, and construct suitable halobacterial host strains. Demonstrate the utility of these genetic tools with several model applications.

ACCOMPLISHMENTS:

Transformation. During the course of this contract we have developed efficient and broadly applicable PEG-mediated spheroplast transformation systems for halophilic archaea including *Halobacterium halobium*, *Halofex volcanii*, and *Halococcus hispanica*. Using these methods we have demonstrated transfection with phage DNA and transformation with linear genomic DNA or plasmid DNA (including shotgun-cloning ligations). Transforming DNA can be either single or double stranded. Spheroplast regeneration under our conditions is also very efficient.

At this point we have performed thousands of transformations and the techniques can be considered completely routine. We have extended our basic protocols to include a microtiter plate-based mini transformation protocol for conveniently performing very large numbers of transformations and a quick transformation protocol which uses cells simply scraped from the surface of plates. In addition, we have developed convenient procedures for storing frozen cells for transformation.

Basically the transformation method can be summarized in six steps: (i) Resuspend cells in spheroplasting solution, (ii) Form spheroplasts by adding EDTA, (iii) Add DNA, (iv) Add PEG solution, (v) Dilute or wash cells with regeneration salts solution, (vi) Plate cells in regeneration top agar.

While our transformation methods work well for representatives of at least three halobacterial genera, we have found that some halobacterial species (*Halobacterium saccharovorum*, *Halobacterium trapanicum*, *Halococcus morrhuae*) are refractory to spheroplast formation by treatment with EDTA and therefore appear to be unsuitable for transformation by this method.

The bulk of our work has been performed using *Hf. volcanii* as a model organism. We have made extensive use of non-replicating homologous DNAs (genomic DNA, cosmid clones, restriction fragments from agarose gels, etc.) in transforming *Hf. volcanii*. Transformations with these DNAs require recombination with the genome and we have demonstrated that this is a reasonably efficient process in *Hf. volcanii*. Typically such experiments involve transformation with wild-type DNA and selection for correction of an auxotrophy, or the use of directly selectable markers such as drug resistance.

Generally, DNAs propagated in foreign hosts are restricted by a factor of about 10^4-fold by *Hf. volcanii*, but, thanks to the high efficiency of the transformation system, this has not hindered our work. We have recently learned that plasmid DNA which has been passaged through a dam- E. coli strain (such as JM110) evades restriction by *Hf. volcanii* (M.L. Dyall-Smith and S.N.)
Nuttall, University of Melbourne, personal communication). We have confirmed that the efficiency of transformation of *Hf. volcanii* by JM110-propagated plasmid DNA can exceed $10^8$ per microgram, compared to $10^4$ - $10^5$ transformants per microgram with plasmid DNA propagated in *dam* strains such as DH5α.

**Vector development.** We have developed a series of *E. coli*-halobacterial shuttle vectors. The workhorse among these has been the 10.5 kb pWL102 (see figure on highlight page). This vector carries (i) a ColE1 replicon and ampicillin resistance from pAT153, (ii) part of the sequenced *Hf. volcanii* plasmid pHV2 (iii) a number of useful cloning sites and (iv) a 3.5 kb fragment conferring resistance to mevinolin (a competitive inhibitor of HMGCoA reductase) in *Hf. volcanii*.

Early attempts at constructing shuttle vectors using a pUC-type vector as an *E. coli* replicon resulted in some instability in *E. coli*. Instability problems were completely overcome by relying on pAT153, with its lower copy number, as the basis for replication in *E. coli*.

Halobacterial replication functions of pWL102 are carried by 3747 bp representing about 60% of the parental pHV2 sequence. This DNA contains an ORF capable of encoding a polypeptide of 808 amino acid residues flanked by 1323 bp which include 532 bp of the 5-prime portion of a 600 bp ORF predicted in the original pHV2 sequence. Presumably this amount of DNA includes sufficient information for replication control and maintenance as well as the origin of replication.

Evidence suggests that pWL102 is not transmissible. Rosenshine et al. (Science 245:1387-1389, 1989) have relied on pHV2 as an immobile cytoplasmic marker in *Hf. volcanii* mating experiments. The shuttle vector cannot be mobilized in *E. coli* either since the ColE1 origin for conjugal transfer was deleted during the original construction of pAT153.

The mevinolin resistance (*Mev*R) locus of pWL102 was originally acquired from a spontaneous mutant of *Hf. volcanii*. We have sequenced the Mev*R* locus, as well as its wild-type counterpart, and found that it has a high degree of sequence similarity to eukaryotic genes for HMGCoA reductase but is very unlike a eubacterial example of this gene from *Pseudomonas*. The 1212 bp gene is about half the size of coding portions of typical eukaryotic homologs since its product, being a soluble enzyme in *Hf. volcanii*, lacks the eukaryotic membrane spanning domains. The point mutation which confers resistance to mevinolin has been identified as a change from GTIA to TTI'A 29 bases upstream of the initiating ATG in a region with similarity to a predicted consensus sequence for archaeabacterial promotors (see figure on highlight page). This appears to be a first *in vivo* example of an archaeabacterial promotor mutation (other than insertional inactivation) and, as such, should yield further insight into the structure of archaeabacterial promotors. Increase in HMGCoA reductase levels by a promotor "up" mutation is consistent with our finding that *Hf. volcanii* can also overcome the inhibition of this enzyme by mevinolin through a second mechanism: amplification of the HMGCoA reductase gene.

Sequence information is available for about 85% of pWL102 including all of the pHV2 and pAT153 portions and about 2 kb of the MevR insert. Unique restriction sites are present for Kpn I, Xba I, Eco RI, Cla I, Bam HI, Nco I and Sph I. Insertions can be made into any of these sites without disrupting plasmid maintenance or resistance functions.

Our shuttle vectors have been distributed to a number of other laboratories. Importantly, our *Hf. volcanii* mevinolin resistance locus has provided selection for a *Hb. halobium*- *E. coli* shuttle vector, pUBP2, constructed by F. Pfeifer and U. Blaseio (Proc. Natl. Acad. Sci. USA 87:6772-6776, 1990). Pfeifer's group has found that pUBP2 replicates in *Hf. volcanii* as well as *Hb. halobium* and we have employed this vector ourselves in experiments with the *Hb. halobium* bacteriorhodopsin gene (see below). Our early experiments indicated that *Hb. halobium* could not be stably transformed with pWL102. Pfeifer's group also found this to be the case, but, surprisingly, when they inserted a *Hb. halobium vac* gene into the pWL102 Kpn I site (between the MevR insert and the pHV2 sequence) stable replication in *Hb. halobium* resulted. We subsequently demonstrated that both pWL102 and pUBP2 replicate stably in *Ha. hispanica*, a halophile somewhat more distantly related to *Hb. halobium* and *Hf. volcanii* than either of these two species are to one another. Thus the halobacterial replicons of both of these shuttle
Vectors appear to have broad host ranges, transforming representatives of at least three halobacterial genera. This is roughly comparable to the phylogenetic depth spanned by broad host range vectors such as RSF1010 or RP4 within the γ subdivision of the purple bacteria.

Other shuttle vectors which we have constructed include those bearing (instead of the mevinolin-resistance determinant) (i) the *H. volcanii* homolog of the *E. coli* trpB gene which is selectable in *H. volcanii* trpB mutants or (ii) *H. volcanii* tRNA<sup>Ser</sup>-derived amber, ochre or opal suppressors which are potentially selectable in *H. volcanii* nonsense mutants.

**Host development.** Initially, we cured wild-type *H. volcanii* of its endogenous pHV2 to create an appropriate host for transformation with pHV2-derived vectors. This strain, WFD11, has served as the parent for virtually all of our subsequent strain constructions. EMS mutagenesis of WFD11 was followed by screening and identification of 256 auxotrophs in the biosynthetic pathways for 16 amino acids as well as pathways for biotin, adenine, guanine, and pyrimidine biosynthesis. Double mutants have also been generated. We have an additional 126 auxotrophs which are defective in unidentified pathways. All of these and others, such as our Mev<sup>R</sup> mutants, spontaneous amino acid analog resistance mutants, and a number of *H. volcanii* strains from the collection of M. Mevarech (Tel Aviv University), are being maintained as frozen stocks in a well organized collection with a companion computerized data base.

Our vector and transformation systems have been integrated into ongoing work in this laboratory to create a physical map of the *H. volcanii* genome using bottom-up mapping by ordering a set of cosmid clones and top-down mapping of restriction fragments by pulsed field gel electrophoresis (mapping work supported by Medical Research Council). By transforming the EMS-generated auxotrophic mutants mentioned above with a minimal ordered set of about 160 *E. coli* cosmid clones of the *H. volcanii* genome, we have succeeded in genetic mapping of 139 alleles to 28 different clusters. These include alleles representing arginine, glutamine, histidine, isoleucine-valine, leucine, lysine, methionine, serine, threonine, tryptophan, tyrosine, early aromatic, adenine, guanine and pyrimidine biosynthetic pathways. We have also mapped the HMGCoA reductase gene genetically as well as several of those auxotrophic alleles defective in unidentified pathways.

We have characterized (by subcloning and sequencing) cosmid clones which complement auxotrophies in the tryptophan, arginine, histidine and leucine pathways. This has allowed us to identify homologs of the *E. coli* genes hisC, trpD, trpF, trpE, trpG, trpC, trpB, trpA, argF, argB, argC and pyrB, as well as oat (encoding ornithine acetyltransferase), a gene not found in *E. coli*.

The trpCBA and trpDFEG clusters were found to map to separate locations in the genome. The gene order for both clusters was unique compared with eubacterial examples and distinct from an example of methanogen organization as well. Genes in both trp clusters contain overlapping start and stop codons. The trpCBA cluster is preceded by sequence which can potentially be folded to form alternate stable stem-and-loop structures, and we're now taking a look at whether these may play some regulatory role.

We mapped the arginine mutations of 20 auxotrophs to one cosmid. Sequencing of DNA which complemented arginine auxotrophies revealed the following organization: -ORF->-argC->-argF->-argB->-oat->-URF->-pyrB->. Again, several of these genes have overlapping start and stop codons and the remainder have abutted stop/start codons. The presence of oat is diagnostic of a more energy efficient pathway compared to that found in some eubacteria such as *E. coli* and *B. subtilis*, which lack it. This enzyme has also been found to be present in methanogens and thermophiles.

The genetic defects in 18 of our histidine auxotrophs have been mapped to four chromosomal locations (i.e. mapped to non-overlapping cosmid clones). Sequencing of 2 kb of DNA from one of these complementing clones revealed a 1086 bp (361 aa) homolog to the *E. coli* hisC gene (encoding histidinol-phosphate aminotransferase). There do not appear to be additional genes near the hisC gene as judged by examination of upstream and downstream sequence.

About two dozen leucine auxotrophs have been characterized by growth on pathway intermediates into at least three classes. All of these map to one cosmid (~40 kb of *H. volcanii*...
DNA). Three kilobases of DNA from this cosmid has been cloned into the shuttle vector pWL102 and shown to complement several auxotrophs. This DNA is being sequenced.

These detailed characterizations have lead to the ability to perform sophisticated host constructions in *H. volcanii* such as deletion of the genome's *trpB* region thus creating a host lacking sequence homology with the *trpB* shuttle vector referred to above.

**Special Applications. Suppressor tRNAs.** Using site-directed mutagenesis we have changed the anticodon of a cloned *H. volcanii* tRNA<sup>Ser</sup> gene to be complementary to the three nonsense codons, thus creating potential suppressor tRNAs. A 248 bp sequence including the tRNA gene has been cloned into pWL102 for each of the suppressor tRNAs plus the wild type. Thus each plasmid is isogenic except for the bases encoding the anticodon. Preliminary indications are that the UAG suppressor plasmid is lethal in *H. volcanii*. We are currently screening our collection of auxotrophs for nonsense mutants by transformation with the UAA and UGA suppressor tRNAs. Once we've identified nonsense mutants we will try using these as hosts for shuttle vectors which only rely on the cloned suppressor tRNAs for selection (as mentioned above).

**Membrane proteins.** In work being done in collaboration with J. Spudich (Albert Einstein College of Medicine, New York), and based on his observation that *H. volcanii* is motile, we have located probable *H. volcanii* sensory rhodopsin and flagellin genes by probing restriction digests of cosmid clones of the *H. volcanii* genome with the *Hb. halobium* sopI and flaB genes (sopI clone provided by D. Oesterhelt, Max-Planck-Institut für Biochemie, Martinsried; flaB clone provided by M. Sumper, Universität Regensburg). We are in the process of cloning these with the ultimate goal of functional analysis by construction of site specifically altered genes.

We have used pWL102 and the *Hb. halobium* shuttle vector pUBP2 to clone a 9 kb fragment carrying the *Hb. halobium* bacteriorhodopsin gene cluster (supplied by M. Betlach, University of California, San Francisco). These constructs transformed the Bop<sup>-</sup> *Hb. halobium* strain Pho81 to a Bop<sup>+</sup> phenotype. *H. volcanii* and *Ha. hispanica*, species which do not naturally produce purple membrane, were also transformed with these constructs, but, surprisingly, transformants did not appear to express bacteriorhodopsin even when grown in media containing retinal. Since the DNA cloned in these shuttle vectors includes all of the genes presently known to be involved in bacteriorhodopsin expression, our results suggest that there may remain further genetic elements to be recognized. *H. volcanii* and/or *Ha. hispanica* may constitute useful genetic backgrounds for determining what additional genetic elements are required for expression.

**PUBLICATIONS:**

We list only publications related to this contract. Those marked with asterisks were submitted since ONR support began, and acknowledge this support.


PhD THESES: Two PhD theses relate to work conducted under this contract.


MANUSCRIPTS IN PREPARATION:


**REPORTS:** Aspects of work performed under this contract have been presented at the following meetings.


UCLA Symposium on Molecular Evolution, Lake Tahoe, California, 1989.


Molecular Evolution Course, Wood's Hole, Massachusetts, 1989.


First Canadian Institute for Advanced Research Evolutionary Biology Meeting for Graduate Students, Montreal, Quebec, 1989.


Fourth International Conference on Retinal Proteins, Santa Cruz, California, 1990.

Fourth International Congress of Systematic and Evolutionary Biology, College Park, Maryland, 1990.

Molecular Evolution Course, Wood's Hole, Massachusetts, 1990.

American Society for Microbiology Symposium, Dallas, Texas, 1991.


Canadian Institute for Advanced Research Genome Evolution Meeting, Quebec City, Quebec, 1991.
Development of Shuttle Vectors for Halobacteria

W.F. Doolittle, Dalhousie University; 1991

Objectives

- Develop transformation systems for halophilic archaeabacteria
- Construct *E. coli*-halobacterial shuttle vectors
- Construct useful halobacterial host strains
- Demonstrate utility of these genetic tools

Accomplishments

- Developed efficient, simple and reliable transformation system for *Halofexx volcanii*, *Halobacterium halobium* and *Haloarcula hispanica* which permits uptake and expression of circular or linear, double-stranded or single stranded, self-replicating or integrating DNAs.
- Generated and characterized hundreds of *Hf. volcanii* auxotrophs, genome deletion strains and other useful mutants.
- Constructed *E. coli*-halobacterial shuttle vectors selectable by drug resistance or by complementation of amino acid auxotrophies. Demonstrated that these shuttle vectors have a broad host range within the halobacteria.

Significance

These transformation methods and shuttle vectors provide powerful tools for investigating the molecular biology of the halophilic archaeabacteria by the methods of modern surrogate genetics.

- Mevinolin resistance conferred by the shuttle vector pWL102 is due to an up-promotor mutation of the *Hf. volcanii* HMGC0A reductase gene.
- Constructed a set of shuttle vectors bearing UAG, UAA, and UGA suppressors derived from *Hf. volcanii* tRNA<sup>Ser</sup>.
- Cloned *H. halobium* bacteriorhodopsin gene cluster into *Hf. volcanii* and *Ha. hispanica* using shuttle vector.
- Mapped and characterized genes of the *Hf. volcanii* arg, trp and his biosynthetic pathways using shuttle vectors and genomic DNA transformations.
Principal Investigator: W. Ford Doolittle

Institute: Department of Biochemistry
Faculty of Medicine
Dalhousie University
Halifax, Nova Scotia B3H 4H7

Grant Title: Development of Shuttle Vectors for Halobacteria

Number of publications acknowledging ONR support during contract period: 8

Number of patents/inventions: none

Total number of students/trainees: 7
  How many are female? 2
  How many are minority students? 2
  How many are not US citizens? 6

Awards/Honors to PI and/or to members of PI's research group (please describe): none

Equipment purchased (# and description of items >$1500): none

Your Email address: FORD@AC.DAL.CA
Distribution List for Final Reports

Attach a copy of the REPORT DOCUMENTATION PAGE (DD FORM 1473) to your final report as the first page and mail two copies (including the postcard labelled DTIC FORM 50) to:

Defense Technical Information Center
Building 5, Cameron Station
Alexandria, VA 22314

This allows other investigators to obtain copies of your report directly from DTIC. DTIC will fill out the postcard DTIC ACCESSION NOTICE (DTIC FORM 50) and return it to you with their number for your report. When you refer people to DTIC to get a copy of your report, give this number to expedite the request.

Mail one copy to each of the following and attach this very page to the back of your report – otherwise the folks below will think they have mistakenly received a copy meant for the Molecular Biology Program:

(a) Dr. Michael Marron
   ONR Code 1141
   Molecular Biology Program
   800 N. Quincy Street
   Arlington, VA 22217-5000

(b) Administrative Grants Officer
    ONR Resident Representative
    (address varies - see copy of your grant/contract)

(c) Director,
    Applied Research Directorate
    ONR Code 12
    800 N. Quincy Street
    Arlington, VA 22217-5000

(d) Director
    Office of Naval Technology
    Code 22
    800 N. Quincy Street
    Arlington, VA 22217-5000

(e) Director
    Chemical and Biological Sci Div
    Army Research Office
    P. O. Box 12211
    Research Triangle Park, NC 27709

(f) Life Sciences Directorate
    Air Force Office of Scientific Res
    Bolling Air Force Base
    Washington, DC 20332

(g) Director
    Naval Research Laboratory
    Technical Information Div
    Code 2627
    Washington, DC 20375

Encl (1) 3/12/91