Coordinate Expression of NADPH-dependent Flavin Reductase, Fre-1, and Hint-related 7meGMP-directed Hydrolase, DCS-1*[S]

Dorota A. Kwasnicka‡§, Agnieszka Krakowiak¶‖, Colin Thacker, Charles Brenner¶‡‡, and Steven R. Vincent‡§§**

*From the ‡Department of Psychiatry and Brain Research Centre and the **Biotechnology Laboratory, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada and the ¶Structural Biology and Bioinformatics Program, Kimmel Cancer Center, Philadelphia, Pennsylvania 19107*

A novel human cytosolic flavin reductase, Nr1, was recently described that contains FMN, FAD, and NADPH cofactors. Though the targets of the related NADPH-dependent flavoprotein reductases, cytochrome P450 reductase, methionine synthase reductase, and nitric oxide synthase, are known, the cellular function of Nr1 is not clear. To explore expression and regulation of Nr1, we cloned *fre-1*, the *Caenorhabditis elegans* ortholog of *Nr1*, and discovered that it is transcribed as a bicistronic pre-mRNA together with *dcs-1*, the ortholog of the recently described scavenger mRNA decapping enzyme. We used the novel substrate, 7meGpppBODIPY, to demonstrate that DCS-1 has low micromolar specificity for guanine ribonucleotides with the 7me modification, whereas trimethylated G substrates are poor competitors. Contrary to earlier classification, DCS-1 is not a pyrophosphatase but a distant member of the Hint branch of the histidine triad superfamily of nucleotide hydrolases and transferases. These observations are consistent with the hypothesis that DCS-1 homologs may function in the metabolism of capped oligonucleotides generated following exosome-dependent degradation of short-lived mRNA transcripts. We find that *fre-1* and *dcs-1* are coordinately expressed through worm development, are induced by heat shock, and have a nearly identical expression profile in human tissues. Furthermore, immunocytochemical analysis of the endogenous proteins in COS cells indicates that both are present in the nucleus and concentrated in a distinct perinuclear structure. Though no connection between these enzymes had been anticipated, our data and data from global expression and protein association studies suggest that the two enzymes jointly participate in responses to DNA damage, heat shock, and other stresses.

Determination of the cellular and biochemical specificities of thousands of newly described enzymes is a major problem of post-genomic biology. In humans, the flavin-containing NADPH cytochrome P450 reductases (CPR)[1] were described more than 25 years ago (1, 2), whereas related nitric oxide synthases (NOS) (3, 4) and methionine synthase reductases (MSR) (5) were described ten and five years ago, respectively. CPRs are endoplasmic reticulum-localized enzymes that serve to maintain cytochrome P450s and heme oxygenases in the reduced active state (6). NOS mediates the synthesis of NO from arginine in the immune, vascular, and nervous systems (7), whereas MSR is required to maintain the methylcobalamin-dependent enzyme methionine synthase in a reduced state for sulfur amino acid biosynthesis and are mutated in children with homocystinuria and other birth defects (5, 8). A novel cytoplasmic reductase named Nr1 was recently described that binds FMN, FAD, and NADPH cofactors and is highly expressed in human cancer cells (9). The substrates and biological pathways involving this enzyme are unknown.

Histidine triad (HIT) enzymes are a superfamily of nucleoside monophosphate hydrolases and nucleoside monophosphate transferases named for an active site sequence related to His-ϕ-His-ϕ-His-ϕ-ϕ (ϕ is a hydrophobic amino acid) in which the nucleoside monophosphate portion of substrates becomes covalently linked to the second His in the first step of the reaction (10). Branches 1 and 2 of the HIT superfamily are hydrolases. Branch 3 enzymes include transferases such as galactose-1-phosphate uridylyltransferase. The most well understood branch 1 enzymes are rabbit Hint and baker's yeast Hnt1, which hydrolyze the adenosine 5′-monophosphoramide substrate, AMP-lysine (11). Losses in Hnt1 lead to yeast cells that are hypersensitive to mutations in Kin28 and other components of general transcription factor TFIIH (11), which has lead to the hypothesis that a post-translationally modified form of Cdk7/Kin28 may be a lysyl-adenylylated protein substrate of Hint (10). In birds, Hint and an apparent Hint inhibitor, Asw, are encoded on the Z and W sex chromosomes, respectively, suggesting that Asw-mediated inhibition of Hint may play a role in female avian development (12). Branch 2 enzymes consist of the human tumor suppressor Fhit (13–15) and its fungal (16) and animal homologs (17), which hydrolyze dinucleoside polyphosphates such as ApppA and ApppppA to AMP plus the other mononucleotide (10).

Human Dcps (scavenger decapping enzyme) was recently purified and cloned as an enzyme that hydrolyzes compounds

such as 7meGpppG and small capped oligoribonucleotides and was proposed to function in the hydrolysis of short oligomers that remain after 3′ to 5′-exonucleolytic degradation of mRNA (18). Deletion of *DCS1* from baker's yeast showed that the enzyme has 7meGpppG-hydrolase activity *in vivo* (18). In fission yeast, the homologous enzyme (Nhm1) was co-purified with the mRNA cap-binding complex eukaryotic initiation factor-4F (19).

Here we identify *fre-1*, the *Caenorhabditis elegans* homolog of the human *Nr1* NADPH-dependent flavin reductase, as the second gene in a two-gene operon with *C. elegans dcs-1*. Biochemical analysis with a newly synthesized 7meGpppBODIPY substrate and structure-based sequence analysis of DCS-1 indicate that DCS-1 homologs are a new sub-branch of HIT hydrolases (Hint branch) with a 7meGMP-binding pocket in place of the AMP-binding pocket of Hint and Fhit. In *dcs-1:fre-1*-green fluorescent protein (GFP) transgenic worms, the *dcs-1:fre-1* promoter directs expression to neurons and pharyngeal muscle and is induced by heat shock. Reverse transcription-polymerase chain reaction (RT-PCR) analysis of the two genes in worms and humans indicates that they are co-regulated in time and tissue across phyla.

Immunocytochemistry indicates that these two proteins are found together in the nucleus and in a distinct perinuclear structure in COS cells. Expression data suggest that the two enzymes jointly participate in responses to DNA damage, heat shock, and other stresses.

EXPERIMENTAL PROCEDURES

Cloning and Characterization of Worm Transcripts—Cultures of *C. elegans* (Bristol N2) were grown and maintained as described (20). Total RNA, prepared from a mixed population of wild type hermaphrodites, was reverse-transcribed into cDNA with SuperScript reverse transcriptase (Invitrogen) using oligo(dT) primers followed by PCR using gene-specific oligonucleotides 1 and 2 for *fre-1* and 3 and 4 for *dcs-1*. Full-length cDNAs were cloned into pCR2.1 (Invitrogen) and sequenced. Trans-splicing specificity was assessed by performing RT-PCR with SL1 and SL2 as the 5′-primers and primers 2 and 4 as the 3′-primers as described (21). To examine developmental expression patterns, RNA was isolated from synchronized N2 nematodes cultured at 20 °C and RT-PCR was performed using the primers for *fre-1* and *dcs-1*, along with primers 5 and 6 for the *S*-adenosyl L-homocysteine hydrolase (AHH) gene, whose mRNA abundance is unchanged throughout development.[2] For expression analysis by RT-PCR, 1 μg of total RNA was subjected to 35 cycles of amplification with an annealing temperature of 59 °C. The same primers were used for RT-PCR analysis of worms subjected to a two-hour, 33 °C heat shock. Primer sequences are provided in Supplemental Table I.

Construction and Characterization of Transgenic Worms—A PCR product was obtained consisting of 1.8 kbp of DNA upstream of the *dcs-1* initiator codon, the entire *dcs-1* coding region, and the first 38 bp of the *fre-1* open reading frame using primers 7 and 8. This fragment, digested with *Xba*I and *Sma*I (sites engineered in the primers), was introduced into GFP expression vector pPD95.70 (23). The resulting plasmid was injected at 10 μg/μl along with p*lin-15* containing the wild type *lin-15* gene at 50 ng/μl and pBlueScript KS at 50 ng/μl into the gonad arms of *lin-15 (n765ts)* adult hermaphrodites as described (24). Transgenic animals were identified by rescue of the *lin-15 (n765ts)* multivulval phenotype at 20 °C, and stable lines were isolated. Cell identification was based on the characteristic morphology and position of GFP-positive nuclei viewed by simultaneous fluorescence and differential interference contrast microscopy.

Cloning and Characterization of Human Transcripts—Human *Nr1* (primers 9 and 10) and *Dcps* (primers 11 and 12) cDNAs were amplified from placental total RNA using *Pfu* polymerase (Stratagene) and cloned into pCR2.1. Using the same primers, PCR-based expression analysis was performed with a panel of human cDNAs (Clontech) according to the manufacturer's instructions.

Purification of Recombinant DCS-1—Worm *dcs-1* was cloned into pSGA04 (25) using primers 13 and 14 to express a His-tagged polypeptide after induction with isopropyl-1-thio-β-D-galactopyranoside

---

[2] C. Thacker, unpublished results.

(IPTG). The resulting plasmid pB344 was transformed into *Escherichia coli* strain BL21 and grown in Luria Bertani broth with 150 μg/ml ampicillin. A 2-liter culture was grown at 30 °C to an $A_{600 nm}$ of 0.4. After addition of IPTG to 0.4 mM, cells were shaken an additional 7 h. The washed, frozen, and thawed cell pellet was resuspended in 30 ml of 100 mM NaCl, 50 mM sodium phosphate, pH 8, 1 mM imidazole with two tablets of complete protease inhibitor mixture (Roche Applied Science) and lysed by sonication. Clarified lysate was loaded onto a 10-ml nickel-nitrilotriacetic acid column (Sigma) and washed with six column volumes of the same buffer with 40 mM imidazole and two column volumes with 50 mM imidazole. DCS-1 was eluted using a gradient of imidazole from 50 to 200 mM. Homogeneous enzyme was concentrated and dialyzed against 20 mM K·HEPES, pH 7.5, 50 mM KCl, 2% glycerol and frozen at −80 °C.

Synthesis of 7meGpppBODIPY—The critical precursor, 7meGTPγS, was synthesized enzymatically by a modification of procedures used to synthesize other thio-substituted nucleoside triphosphates (26–28). Briefly, 10 mM 7meGDP and 2 mM ATPγS were incubated in 25 mM Tris, pH 7.4, 2 mM MgCl$_2$, with 50 units of nucleoside diphosphate kinase at 30 °C for 2 h in a 1-ml reaction volume. The 7meGTPγS product was isolated by RP-HPLC, with a 5-μm ODS Hypersil column (Alltech) using 0.1 M triethylammonium bicarbonate, pH 7.4 (Buffer B, 0.1 M triethylammonium bicarbonate plus 40% CH$_3$CN). At a flow rate of 1 ml/min, a 30-min cycle consisted of 24 min from 0 to 20% Buffer B, 3 min for 20–100%, and 3 min 100–0%. Under these conditions, 7meGTPγS was eluted at 19.78 min. 7meGpppBODIPY was synthesized and purified as described for GpppBODIPY (29), *i.e.* by reacting 7meGTPγS with *N*-(4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacine-3-yl)methyl)iodoacetamide in dioxane.

Enzymatic Assays—To determine kinetic parameters, duplicate or triplicate assays were performed at 30 °C in 40 mM K·HEPES, pH 7.3, 1 mM MgCl$_2$ in total volumes of 20 μl with six or seven substrate concentrations ranging from 0.31 to 20 μM for 7mGpppBODIPY and 1.25–40 μM for GpppBODIPY and ApppBODIPY. In initial rate assays, the amount of enzyme varied from 0.1 to 0.55 pmol for 7meGpppBO-DIPY, 2.47–6.3 pmol for GpppBODIPY, and 4.39–11 pmol for Appp-BODIPY. Reactions samples were spotted on silica TLC plates (Merck) at 60–120 s intervals. Plates were developed in 2-propanol:NH$_4$OH:1,4-dioxane:H$_2$O (50:33:6:11), imaged with a Typhoon 9400 (Amersham Biosciences) (excitation 488 nm, emission 520 nm), and quantitated using ImageQuant 5.2 software. To determine the K_i of nonlabeled competitor nucleotides, these compounds were titrated into 20 μl of complete hydrolysis reactions containing 1.25 μM 7meGpppBODIPY and 0.27 pmol enzyme. Using methods established earlier with Fhit and GpppBODIPY (29), K_i values of the nonlabeled nucleotides were determined from the concentration-dependence of reduction of k_{cat}/K_m(apparent) for 7meGpppBODIPY.

Antibodies and Immunocytochemistry—Antibodies were raised in rabbits against synthetic peptides derived from human Nr1 (amino acids 197–207) and Dcps (amino acids 181–195), conjugated to KLH, and affinity-purified using the peptide antigens. In Western blots of protein from COS and HEK cells or mouse brain, the antibodies recognized bands of the predicted size for Nr1 (68 kDa) and Dcps (38 kDa). For immunocytochemistry, COS cells were grown on coverslips, fixed in 4% paraformaldehyde, and incubated with the antibodies at a final concentration of 2 μg/ml. Following washing in phosphate-buffered saline, the cells were incubated in Cy3-conjugated donkey anti-rabbit antisera (The Jackson Laboratories), counterstained with 4′,6-diamidino-2-phenylindole (DAPI), and examined by fluorescence microscopy.

RESULTS

Homologs of Human NADPH-dependent Flavin Reductase and Scavenger Decapping Enzyme Are Encoded on a Two-gene Operon in Worms—We examined the inventory of predicted flavoprotein reductases in the worm genome (30) and discovered homologs of CPR (accession number NP_498103), MSR (accession number NP_495978), and Nr1 (accession number 6425246), the latter of which we named *fre-1*. As shown in Fig. 1, the predicted 585-amino acid fre-1 polypeptide conserves sequences defined to bind FMN, FAD, and NADPH but lacks the N-terminal hydrophobic sequences that anchor CPR to the endoplasmic reticulum, the N-terminal heme-binding oxidase domain found in NOS, and the central sequences that are diagnostic for MSR.

Just 180 nucleotides upstream of the *fre-1* initiator codon, we

```
                                                      FMN phosphate
K10D2.6     1  ---------------MLAWIVSGLDTSDLVVLTLLAGGAIIFLFMKVFNQQPSSSRYSPTV
CPR         1  MGDSHVDTSSTVSEAVAEEVSLFSMTDMILFSLIVG---LLTYWFLFRKKKEEVPEFTKI
NR1         1  ------------------------------------------------------------
Y113G7A.8   1  ------------------------------------------------------------

                                                          FMN ring
K10D2.6    47  ASVTTSAAASKSNQSFIDRMKNENRQVLIMYGSQTGTAEEMSGRLAKDLTRYTKKAVVVD
CPR        58  QTLTSSVRES----SFVEKMKKTGRNIIVFYGSQTGTAEEFANRLSKDAHRYGMRGMSAD
NR1         1  --------------------MPSPQLLVLFGSQTGTAQDVSERLGREARRRRLGCRVQA
Y113G7A.8   1  --------------------MSIAILYGSETGTAQDIAESLRREAQQRHLQARVHE

K10D2.6   107  PEDIECEDLNRLSEVEDALLVLCIATYGEGDPTDNAVTLVEYLNAG---DCDLSGVRFAV
CPR       114  PEEYDLADLSSLPEIDNALVVFCMATYGEGDPTDNAQDFYDWLQET---DVDLSGVKFAV
NR1        40  LDSYPVVN---L--IINEPLVIFVCATTGQGDPPDNMKNFWRFIFRKNLPSTALCQMDFAV
Y113G7A.8  37  LDEYDVSQLP-----MEKVVLFVVSTTGQGEMPPNMRKTWKLLLRKSLGADFLKNVNIGV

                          FMN ring
K10D2.6   164  FGLGNKTYEHFNEIGIQMDKQLEKLGAKRIFHLGLGDDDANLEEDFMI--WREAFLPKVA
CPR       171  FGLGNKTYEHFNAMGKYVDKRLEQLGAQRIFELGLGDDDGNLEEDFIT--WREQFWPAVC
NR1        95  LGLGDSSYAKFNFVAKKLHRRLLQLGGSALLPVCLGDDQHELGPDAAVDPWLRDLWDRVL
Y113G7A.8  92  LGLGDSSYQKYNFAGKKLYRRLVQLGAKMMCGVHLADDQHEIGIDGAFIPWKTECWKKIK

K10D2.6   222  EEFG-WELNTEAETMRQYQLEPVE---EGKALFKGEFGRLGAYERPRPPFDVKNPYLATV
CPR       229  EHFG-VEATGEESSIRQYELVHTDI-DAAKVYMGEMGRLKSYENQKPPFDAKNPFLAAV
NR1       155  GLYPPPPGLTEIPPGVPLPSKFTLLF-LQEAPSTGSEGQRVAHPGSQEPPSESKPFLAPM
Y113G7A.8 152  EIAENLSIYREMTAEIDTSVQITTKYRFLELGGESGNDDDDSDEDLEPQIEIPDYFPLKV

K10D2.6   278  AINDELHTEHSDRSCRHIEFSVEG---SRIRYEAGDHLAVFPTNDPVLVDRLINMLQFDP
CPR       287  TTNRKLN-QGTERHLMHLELDISD---SKIRYESGDHVAVYPANDSALVNQLGKILGADL
NR1       214  ISNQRVTGPSHFQDVRLIEFDILG---SGISFAAGDVVLIQPSNSAAHVQRFCQVLGLDP
Y113G7A.8 212  LKNRRLTSAEHFQDTRLVDFEISDRNRSKMGFQPGDVLMVREYNPEETVKIAIEALGLTE

K10D2.6   335  DHAFRLVNVDEDASKRHPFP-----CPTTFRTALSHYVDICAPVKSHVLKAISEYCTDDT
CPR       343  DVVMSLNNLDEESNKKHPFP-----CPTSYRTALTYYLDITNPPRTNVLYELAQYASEPS
NR1       271  DQLFMLQPREPDVSSPTRLP-----QPCSMRHLVSHYLDIASVPRRSFFELLACLSLHEL
Y113G7A.8 272  EQLKKPLKIVKNDRFSKNPNFLVGTKTTLLTCLQRYFDLQQVPKRSFFEMLGYYSTNPP

K10D2.6   390  EKEFLNKLSTANEEGLKEYARYIVKERRSIVDVLTDQKT--CKPPIEYLLELLPRLQARY
CPR       398  EQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPS--LRPPIDHLCELLPRLQARY
NR1       326  EREKLLEFSSAQ--GQEELFEYCNRPRRTILEVLCDFPHTAAAIPPDYLLDLIEVIRPRA
Y113G7A.8 332  EKERLQELASPE--GLDDYLDYANRSRRTTAEALRDFVATSKNLKPDYLFEILTTIRPRA

                 FAD ring                  FAD adenine              FAD pyrophosphate
K10D2.6   448  YSIASSPRLNEEKIAICAVVTKYSIGDRDING-VCTRYLTTKDAGSKS-----P-VFVRK
CPR       456  YSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRK
NR1       384  FSIASSLLTHPSRLQILVAVVQFQTRLKEPRRGLCSSWLASLDPGQGP---VRVPLWWRP
Y113G7A.8 390  FSIASAP--SPFHLELLVAKVEYKSRMADKRRGLCSTFISRLKPGDEVF------CKIRP

                                    NADPH pyrophosphate
K10D2.6   501  STMRLPHRTTTQVINIGPGTGFAPFRGFLQDRQFHKNAGKEIGAMHLYYGCRHPDHDYIY
CPR       516  SQFRLPFKATTPVINVGPGTGVAPFIGFIQERAWLRQQGKEVGETLLYYGCRRSDEDYLY
NR1       441  GSLAFPETPDTPVINVGPGTGVAPFRAAIQERVAQGQTGN-----FLFFGCRWRDQDFYW
Y113G7A.8 442  GTFKFPS-PEAPVICIGPGTGVAPFRSLFGHRSLFSAHFPG----ILFFGCRSEHHDYYF

                                                              NADPH adenine
K10D2.6   561  KDELAKFQEDEVLTHLVCAFSRAQEHKIYVQDRLWETRDRIWDAINV-GAHVYICGDARN
CPR       576  REELAQFHRDGALTQLNVAFSREQSHKVYVQHLLKQDREHLWKLIEG-GAHIYVCGDARN
NR1       496  EAEWQELEKRDCLT-LIPAFSREQEQKVYVQHRLRELGSLVWELLDRQGAYFYLAGNAKS
Y113G7A.8 497  SDEWPELSGVAVIP---AFSRDTDGRKVYVQHKMGENAGKIKRMLDL-GAQVFIAGSSGD

               FAD ring
K10D2.6   620  MARDVQATLQKIFREIGGKSETEAVAYFKDMEKTKRYQADVWS
CPR       635  MARDVQNTFYDIVAELGAMEHAQAVDYIKKLMTKGRYSLDVWS
NR1       555  MPADVSEALMSIFQEEGGLCSPDAAAYLARLQQTRRFQTETWA
Y113G7A.8 553  MPKAVSAVLAQIQ---G-DE------WTRRAEETGKIQYETWS
```
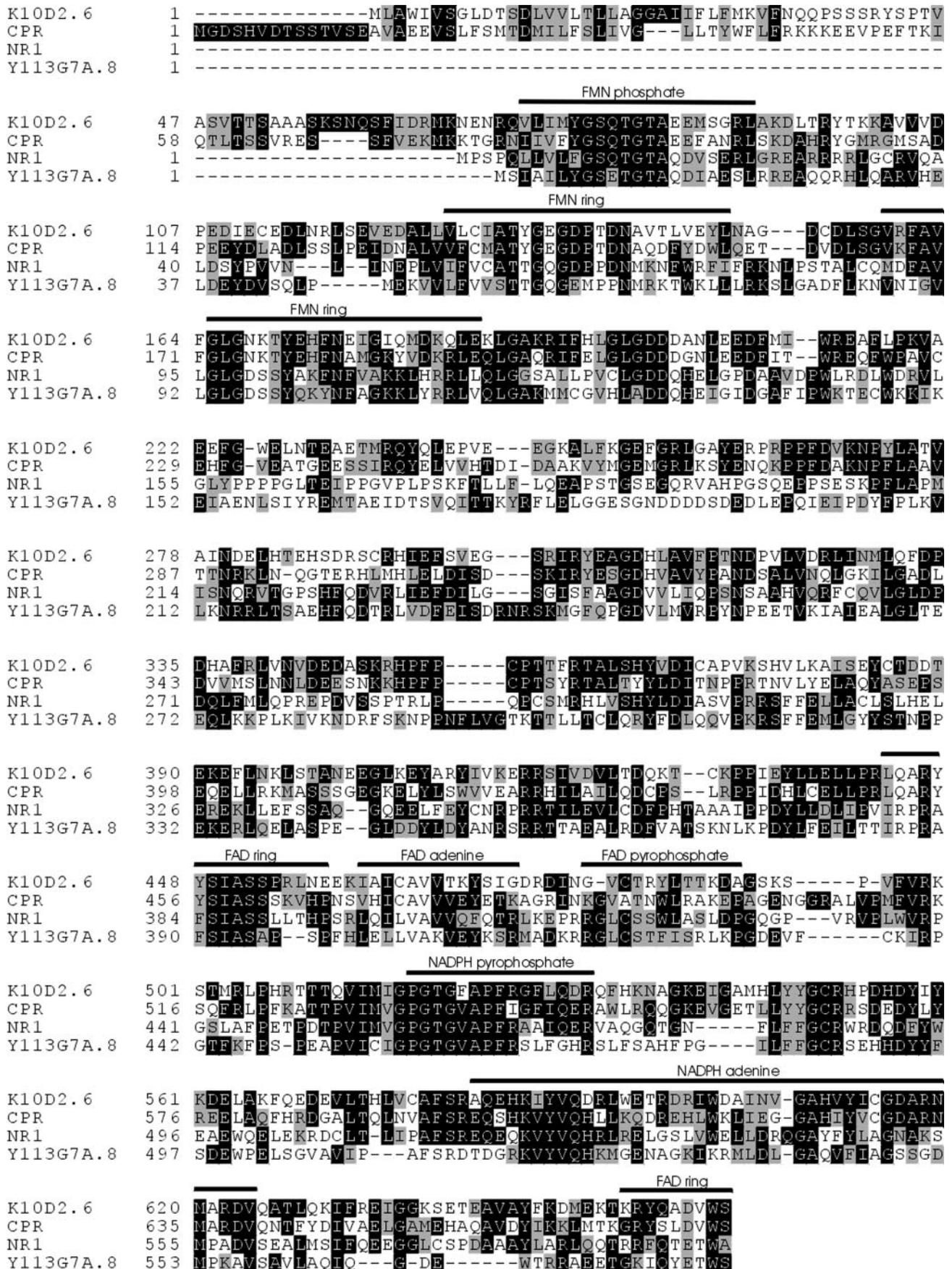
F<small>IG</small>. 1. **Alignment of the Fre-1 sequence with other members of the cytochrome P450 reductase family.** Sequences shown are: *K10D2.6*, *C. elegans* microsomal cytochrome P450 reductase (accession no. U21322); *CPR*, human microsomal cytochrome P450 reductase (accession no. A60557); *NR1*, human novel dual NADPH-dependent reductase (accession no. AF199509); *Y113G7A.8*, novel *C. elegans* reductase (Fre-1) (accession no. AY079165). Identical amino acids are *boxed in black* and conserved residues in *gray*. Cofactor-binding domains are indicated by *black lines*. Amino acids are numbered on the *left*.
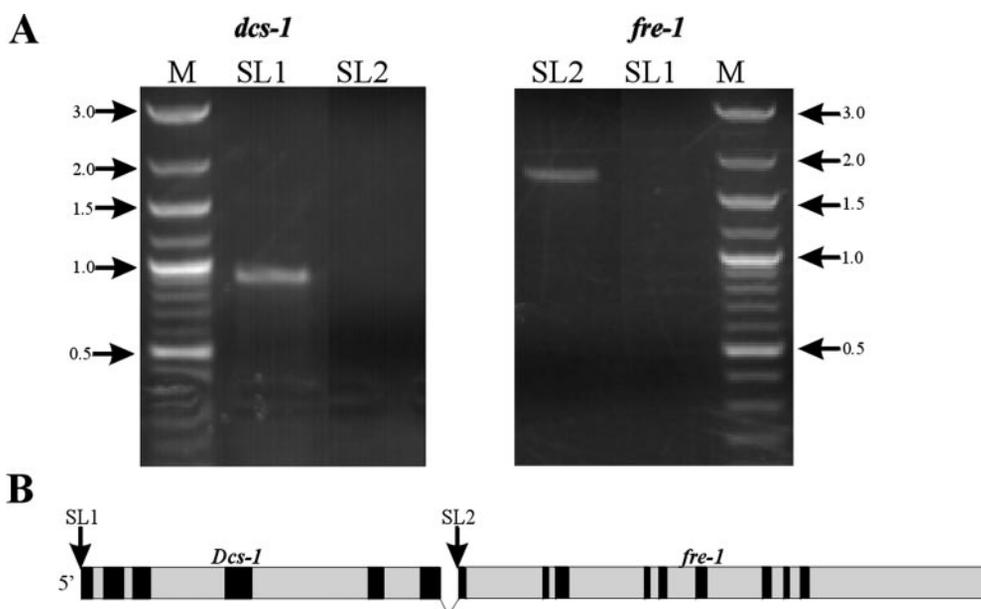
FIG. 2. **dcs-1 and fre-1 are in a two-gene operon.** *A*, RT-PCR was performed with gene-specific primers and SL2 and SL1 oligonucleotides. *dcs-1* (Y113G7A.9) is an upstream gene in the operon and is trans-spliced to SL1, whereas the *fre-1* (Y113G7A.8) is trans-spliced to SL2. The sizes of the PCR products were 936 bp (*dcs-1*) and 1758 bp (*fre-1*) as predicted. *B*, schematic diagram of the two-gene operon in *C. elegans* as predicted from the genome sequence. The insertions of the two splice leaders are indicated by *arrows*. Exons are shown in *black*. The 180-nt sequence between the two genes is also indicated.

located the stop codon of the ortholog of human scavenger decapping enzyme Dcps, which we name *dcs-1*. Approximately 15% of worm genes occur in operons of two to eight genes (31). In these cases, pre-mRNAs are formed containing multiple cistrons that are cleaved and polyadenylylated. In all known cases in which there is an intercistronic length of greater than a few nucleotides, the downstream genes are trans-spliced with the downstream gene-specific SL2 RNA (32). To examine whether the potential downstream gene *fre-1* is, in fact, synthesized as part of an operon, we performed RT-PCR with an SL2 primer and *fre-1*-specific primer 2 (21) and demonstrated amplification of *fre-1* (Fig. 2). Moving upstream to *dcs-1*, we showed that *dcs-1* cDNA can be amplified with a primer to the SL1 trans-spliced leader but not the SL2 leader, demonstrating that *dcs-1* is the 5′-cistron within the *dcs-1:fre-1* operon. Because the next downstream and upstream genes are more than 10 kbp away from *dcs-1* and *fre-1*, exceeding the maximally observed intercistronic distance by more than 20-fold (31), we conclude that *dcs-1* and *fre-1* constitute a two-gene operon in *C. elegans*. A recent high-throughput study of SL2-transpliced transcripts predicted that Y113G7A.9 (*dcs-1*) and Y113G7A.8 (*fre-1*) are in a two-gene operon (31). It should be noted that because the mRNAs for both *dcs-1* and *fre-1* are trans-spliced to SL sequences, these transcripts are expected to contain the trimethylguanosine (TMG) cap from the nuclear SL RNA molecules, like those of other small nuclear RNAs (33).

DCS-1 Is a Member of the Hint Branch of the HIT Hydrolase Superfamily with Specificity for 7meGMP in the Primary Nucleotide Binding Site—Human scavenger decapping enzyme was recently reported to hydrolyze 7meGpppG to 7meGMP plus GDP (18). Sequences of Dcps and Dcs orthologs (hereafter, these enzymes are termed Dcs to distinguish them from unrelated Dcp mRNA decapping enzymes (34)) were described as containing a His-φ-His-φ-His-φ-φ motif in which the second His is required for activity but with no other similarity to HIT hydrolases (18). The orthologous enzyme from fission yeast, named Nhm1p, was described as having low level similarity to Fhit, but only similarity with the orthologous enzymes was

shown (19). Based on our crystal structures of nucleotide-bound forms of Hint (35) and Fhit (15), we performed a structure-based alignment of three Hint orthologs, three Aprataxin orthologs, three Fhit orthologs, and three Dcs orthologs. As shown in Fig. 3, although the level of sequence identity between Dcs sequences and the three other enzymes is low, there is significant conservation that begins one residue before the first α-helix of the Hint structure through one residue after the fifth and final β-strand that is conserved in all HIT hydrolases (this strand is numbered 7 in Fhit because two short helices replace the first helix in Hint (15, 36)). There is no similarity outside this core structure, suggesting that Dcs homologs are unique outside of the dimeric 10-stranded GalT half-barrel that has been found in all HIT enzymes (10). Although human Dcs has a Thr in the seventh position of the His-φ-His-φ-His-φ-φ motif, the alignment shows that this β-branched and moderately hydrophobic residue does not violate the rule because Dcs homologs have β-branched hydrophobic residues Ile and Val, which along with Leu, are the consensus. More surprising than Thr in the seventh position of the HIT motif is the lack of a His that aligns with Hint residue 51 and Fhit residue 35. This residue is in a hydrogen-bonding network interacting with the final His of the HIT motif that has been proposed to be involved in leaving group leaving and water attack (10). Dcs orthologs have a polar residue at this position. Beyond conserved N- and C-terminal extensions that are diagnostic for DCS sequences, the DCS enzymes have a short, acidic loop insertion between β-strands 1 and 2 and a distinctive loop between β-strands 2 and 3 containing the sequence L-P-D-φ-K-W-D-G. On the basis of the crystal structure of Hint bound to GMP (35), this loop is positioned to make contact with 7me group of 7meG, which we predict accounts for the 7meGMP-specificity of DCS-1 and its homologs.

Specificity of DCS-1 Defined with 7meGpppBODIPY—We expressed His-tagged worm DCS-1 in *E. coli* and purified it to homogeneity. To characterize the specificity of this enzyme, we synthesized a novel fluorescent analog, 7meGpppBODIPY. Because the specificity of human Dcps was reported to be for analogs of capped RNAs such as 7meGpppG, 7meGpppA, or

FIG. 3. **Dcs-1 aligns with the structured entirety of Hint.** Mammalian Hint, Aprataxin, DcpS, and Fhit sequences were multiply aligned with two homologs of each enzyme and with each other. The beginning and the ending of the region of sequence similarity with Dcs-1 homologs is one amino acid N-terminal and two amino acids C-terminal to the first and last structured element in the crystal structure of rabbit Hint (35). The sequences of Dcs-1 orthologs are slightly closer to those of Hint orthologs than to other HIT hydrolases. As discussed under "Results," the L-P-D-ϕ-K-W-D-G loop between β-strands 2 and 3 is predicted to account for 7meG-specificity in the primary nucleoside monophosphate-binding site.

TABLE I
Biochemical characterization of DCS-1

Substrate	K_m	k_{cat}	k_{cat}/K_m	K_i
	μM	s^{-1}	$M^{-1} s^{-1}$	μM
7meGpppBODIPY	1.21 ± 0.05	0.174 ± 0.034	1.44×10^5	
GpppBODIPY	8.52 ± 0.40	0.0346 ± 0.0035	4.06×10^3	
ApppBODIPY	5.04 ± 0.91	0.0096 ± 0.0008	1.90×10^3	
7meGDP				2.23 ± 0.61
7meGTP				2.90 ± 0.71
7meGpppG				3.47 ± 0.84
2, 3, 7meGpppG				28.1 ± 2.5
AMPS-pNA-(R_p)				116 ± 34
GpppA				167 ± 18
AMPS-pNA-(S_p)				172 ± 36
AMP-pNA				247 ± 70
AMP-NH$_2$				$>10,000$
ApppA				$>10,000$
GTP				$>10,000$
GDP				$>10,000$

7meGppp(Np)$_n$, where $n <10$ (18), we reasoned that the lack of specificity for substrate features beyond 7meGTP would allow us to substitute a BODIPY reporter group for a nucleoside, nucleotide, or short oligonucleotide. In the past, GpppBODIPY was synthesized from GTPγS by conjugation of an iodoacetamide-linked BODIPY probe (29). Here, we used nucleoside diphosphate kinase to transfer the γ phosphorothio group of ATPγS to 7meGDP. The resulting 7meGTPγS was then converted to 7meGpppBODIPY by established methods (29). As shown in Table I, this turned out to be a nearly essential modification for DCS-1 characterization. Whereas Fhit substrates ApppBODIPY and GpppBODIPY were cleaved by DCS-1 with k_{cat} values of less than 0.01 s^{-1} and 0.035 s^{-1}, the 7me modification of GpppBODIPY increased k_{cat} to >0.17 s^{-1}. Taking k_{cat}/K_m as the specificity constant, DCS-1 prefers 7meGpppBODIPY to ApppBODIPY and GpppBODIPY by >35- and 75-fold, respectively. In the context of the BODIPY nucleotide substrates, the 7me modification to G accounts for a 7-fold discrimination in K_m and a 5-fold discrimination in the k_{cat} term. The 1.21-μM K_m measured for DCS-1 hydrolysis of 7meGpppBODIPY compares favorably with good substrates for other histidine triad enzymes: Fhit cleaves GpppBODIPY with a 1.4 μM K_m, though the higher k_{cat} gives an overall higher specificity constant (29).

Having demonstrated that 7meGpppBODIPY is a useful probe for DCS-1, we sought to further dissect substrate recognition by this enzyme. By titrating nonlabeled nucleotides into 7meGpppBODIPY assays and measuring the concentration-dependent reduction in k_{cat}/K_m(apparent) (29), we determined the K_i of 12 additional nucleotides (Table I). This analysis revealed that most but not all of the binding specificity is for the 7meG

mononucleotide group because we measured K_i values for 7meGDP, 7meGTP, and 7meGpppG in the 2.2–3.5 μM range. There is a hydrophobic binding site on the other side of the scissile bond, however, because GpppBODIPY competes for the DCS-1 active site better than GpppA, which, in turn, competes better than GTP. Similarly, binding specificity for a hydrophobic leaving group can be illustrated in the series of ApppBODIPY, AMP-pNA, ApppA, and AMP-NH$_2$, which exhibited increasingly poor binding. As a consequence of the lack of a 7me modification and the lack of a hydrophobic leaving group, the abundant guanine nucleotides such as GDP and GTP are not DCS-1 competitors even at 10 mM. The 2, 3, 7-trimethylated GpppG analog had a K_i value 8-fold higher than that of 7meGpppG. Thus, the *in vitro* specificity of the enzyme does not support the idea that small nuclear RNAs with 2, 2, 7-trimethylguanosine caps (37, 38) might be *in vivo* substrates for DCS-1 and its orthologs. Earlier, human Dcps was termed a pyrophosphatase (18). In contrast, our work indicates that DCS-1 enzymes are a distinct sub-branch of the Hint branch of HIT hydrolases (10) with specificity for 7meGMP in the primary nucleoside monophosphate binding site.

Evolutionary Origin of DCS-1 and Fre-1 Orthologs—Earlier we have argued that Hint, a prototypical AMP-lysine hydrolase, was the evolutionary precursor of Fhit homologs (10, 35). Although the specificity of DCS-1 is somewhat closer to Fhit (29) than Hint (11), *i.e.* DCS-1 prefers a phosphorylated leaving group such as ppBODIPY or GDP to an amine leaving group, structure-based sequence analysis shown in Fig. 2 indicates greater similarity to Hint than to Fhit. Moreover, sequence searches indicate that DCS-1 orthologs are found in animals including mammals, other vertebrates, invertebrates, fungi in-
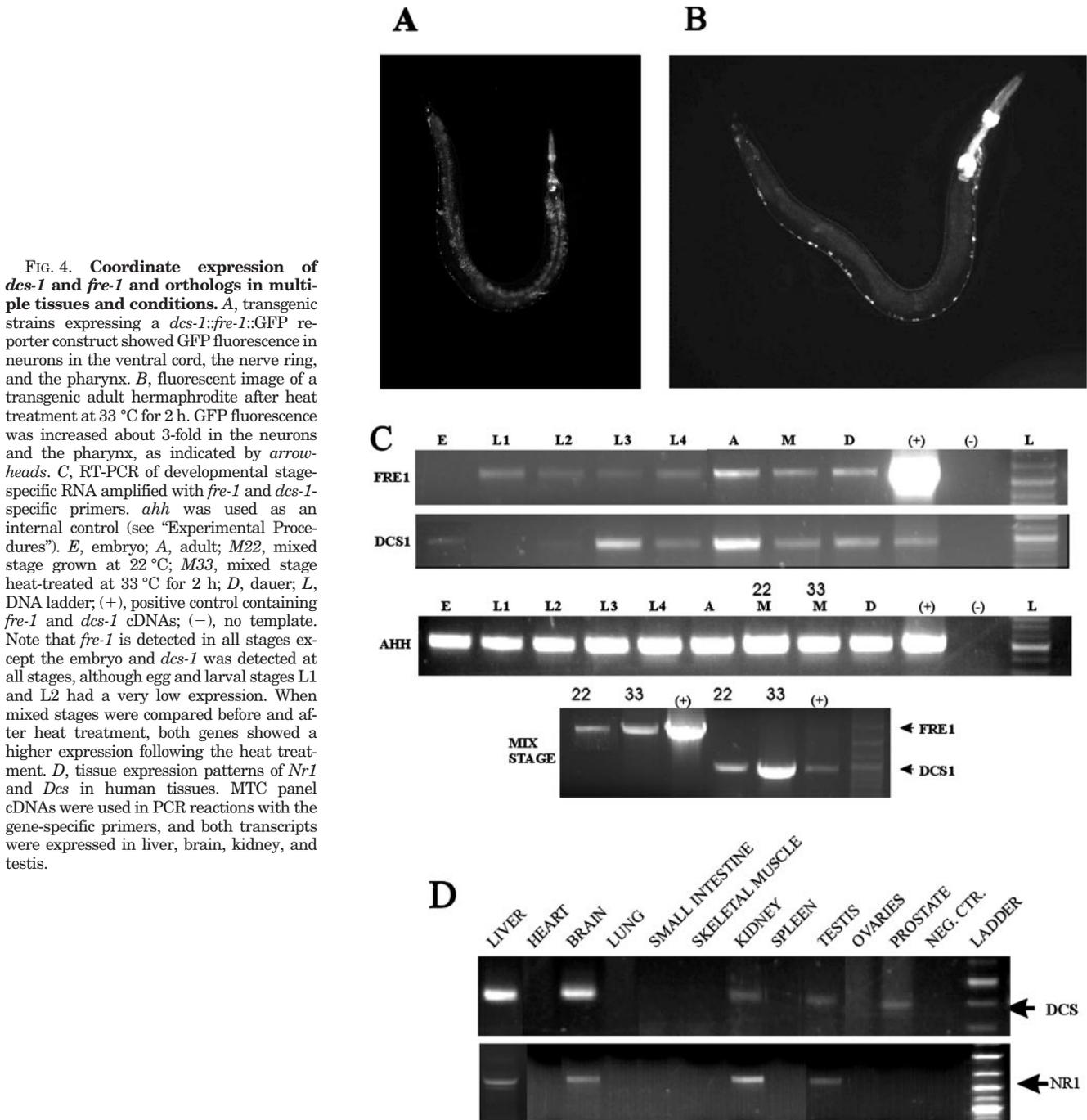
FIG. 4. **Coordinate expression of *dcs-1* and *fre-1* and orthologs in multiple tissues and conditions.** *A*, transgenic strains expressing a *dcs-1*::*fre-1*::GFP reporter construct showed GFP fluorescence in neurons in the ventral cord, the nerve ring, and the pharynx. *B*, fluorescent image of a transgenic adult hermaphrodite after heat treatment at 33 °C for 2 h. GFP fluorescence was increased about 3-fold in the neurons and the pharynx, as indicated by *arrowheads*. *C*, RT-PCR of developmental stage-specific RNA amplified with *fre-1* and *dcs-1*-specific primers. *ahh* was used as an internal control (see "Experimental Procedures"). *E*, embryo; *A*, adult; *M22*, mixed stage grown at 22 °C; *M33*, mixed stage heat-treated at 33 °C for 2 h; *D*, dauer; *L*, DNA ladder; (+), positive control containing *fre-1* and *dcs-1* cDNAs; (−), no template. Note that *fre-1* is detected in all stages except the embryo and *dcs-1* was detected at all stages, although egg and larval stages L1 and L2 had a very low expression. When mixed stages were compared before and after heat treatment, both genes showed a higher expression following the heat treatment. *D*, tissue expression patterns of *Nr1* and *Dcs* in human tissues. MTC panel cDNAs were used in PCR reactions with the gene-specific primers, and both transcripts were expressed in liver, brain, kidney, and testis.

cluding *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, and microsporidia including *Encephalitozoon cuniculi*.[3] As we have shown earlier, Fhit homologs are found additionally in green plants and mycetozoa such as *Dictyostelium discoidium* (10). Thus, to argue that DCS-1 orthologs evolved from Fhit orthologs would imply loss of DCS-1 from significant branches of the evolutionary tree. In contrast, independent descent of DCS-1 and Fhit from Hint requires only the assumption that DCS-1 is a more recently evolved hydrolase, *i.e.* after the rift of plants from animals and fungi (39), that specializes in 7meGMP hydrolysis from small nucleotide or oligonucleotide substrates. A similar analysis of Fre-1 orthologs indicates that they are a somewhat older group of enzymes, having homologs on either side of the plant/animal

rift.[3] Thus, in the first fungal-animal precursor cell in which a DCS-1 hydrolase appeared, a Fre-1-type reductase is expected to have already been part of the protein complement.

Coordinate Regulation of dcs-1 and fre-1 in Worms and Human—Because *dcs-1* and *fre-1* mRNAs are spliced from the same operon, transcription of both genes is probably controlled by sequences within or upstream of *dcs-1*. Four sequences 5′ of the *dcs-1* coding sequence (−1800 to −1795, CGAAA; −1764 to −1759, GGAAT; −1716 to −1711, CGAAG; and −1634 to −1629, GGAAC) appear to be classical heat shock elements (40). To examine the expression pattern of *fre-1*, we amplified the 1.8 kbp 5′ to the *dcs-1* open reading frame, the entire *dcs-1* coding region, and the first 38 coding bp of *fre-1* and cloned this construct in-frame with GFP. Examination of transgenic animals carrying this construct showed GFP expression throughout development in the nuclei of neurons and the pharyngeal

---

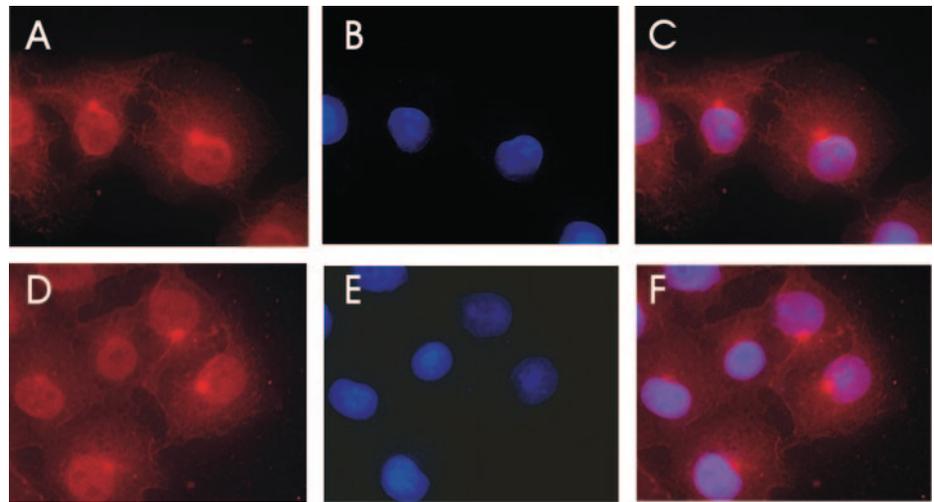[3] C. Brenner and S. Vincent, unpublished work.

FIG. 5. **Immunocytochemical localization of Nr1 and Dcs in COS cells.** Cells were incubated with affinity-purified rabbit antibodies to Nr1 (*A*) or Dcs (*D*) and detected with Cy3-labeled secondary antisera, and the nuclei were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) (*B* and *E*). In the overlay images (*C* and *F*), the accumulation of the two proteins in a distinct perinuclear structure is apparent.

muscle, with heat shock enhancing GFP expression (Fig. 4*A*). RT-PCR analysis of *dcs-1* and *fre-1* indicates that both mRNAs are expressed in all developmental stages, that L3 and adults have higher expression levels, and that both messages are present in dauer animals and similarly induced by heat shock (Fig. 4*B*). The RT-PCR assay was applied to human tissues and revealed that brain, liver, kidney, pancreas, and testes were positive for both messages. Heart, lung, small intestine, skeletal muscle, colon, leukocytes, spleen, ovaries, and thymus were negative for both messages. Only prostate was positive for *DCPS* and negative for *NR1*, and no characterized human tissue was positive for *Nr1* and negative for *DCPS*. Thus, though *DCPS* and *NR1* are not neighboring genes in humans, their expression profile is highly correlated (Fig. 4*C*).

Endogenous Nr1 and Dcps were detected in COS cells by immunocytochemistry (Fig. 5). Staining for both proteins was detected in the nucleus, but they were concentrated in a distinct perinuclear structure in these cells.

DISCUSSION

Consistent with our observations that *dcs-1* and *fre-1* are heat shock-induced genes in worm, the baker's yeast *DCS1* and *DCS2* genes are induced by heat shock and mutagens and are repressed by cAMP withdrawal. Dcs1 was reported to have a two-hybrid interaction with neutral trehalase (41) and to be found in a large complex with the stress-induced protein kinase Mpk1 by systematic mass spectrometry (42). These interactions were validated by showing that the genes are coordinately regulated at the level of mRNA accumulation and show similar heat shock-sensitive and caffeine-sensitive phenotypes (43). Though cell cycle initiation is highly sensitive to reduction of eukaryotic initiation factor-4E levels, translation of heat shock proteins such as polyubiquitin is insensitive (44), so it would appear unlikely that scavenger decapping activity is required to allow immediate-early stress-induced changes in gene expression. Resumption of the cell cycle after stress may entail an increased requirement for capping or cap-recognition, either of which could be inhibited by accumulation of short, capped RNAs.

The precise nature of the DCS-1 substrate remains an interesting question because it has little activity on long capped RNAs (18). It is interesting to note that the class of genes most frequently transcribed in operons in *C. elegans* encodes proteins involved in RNA degradation (31). mRNA decapping is mediated by the Dcp1 and Dcp2 Nudix-type hydrolases (45), which are located in distinct cytoplasmic foci (46–48). The human Hint-related Dcps enzyme was purified with the thought that its activity might be associated with a late stage

of mRNA degradation (18); however, the enzyme appears to be nuclear in fission yeast (19). Thus DCS-1 may function in exosome-mediated mRNA degradation in the cytoplasm and nucleus. We note that the stresses that induce *dcs-1* gene expression may result in production of aborted mRNA transcripts. Because capping and methylation of messages occurs early in the elongation phase of transcription (22), conditions that lead to the early dissociation of the polymerase II machinery would be expected to lead to production of short capped RNAs with no useful function. Given the lack of specificity for trimethylated nuclear caps, we suggest that capped, monomethylated, aborted transcripts may be relevant DCS-1 substrates. Discovery of *Dcs-1* and *fre-1* as coordinately expressed genes will allow testing of the hypothesis that scavenger decapping activity and a specific NADPH-dependent flavin reductase activity are required for responses to heat shock and genotoxic stress in animals.

REFERENCES

1. Yasukochi, Y., and Masters, B. S. (1976) *J. Biol. Chem.* **251,** 5337–5344
2. Dignam, J. D., and Strobel, H. W. (1975) *Biochem. Biophys. Res. Commun.* **63,** 845–852
3. Bredt, D. S., Hwang, P. M., Glatt, C. E., Lowenstein, C., Reed, R. R., and Snyder, S. H. (1991) *Nature* **351,** 714–718
4. Schmidt, H. H., Pollock, J. S., Nakane, M., Gorsky, L. D., Forstermann, U., and Murad, F. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 365–369
5. Leclerc, D., Wilson, A., Dumas, R., Gafuik, C., Song, D., Watkins, D., Heng, H. H., Rommens, J. M., Scherer, S. W., Rosenblatt, D. S., and Gravel, R. A. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 3059–3064
6. Ding, S., Yao, D., Deeni, Y. Y., Burchell, B., Wolf, C. R., and Friedberg, T. (2001) *Biochem. J.* **356,** 613–619
7. Vincent, S. (1995) *Nitric Oxide in the Nervous System,* Academic Press, London
8. Wilson, A., Platt, R., Wu, Q., Leclerc, D., Christensen, B., Yang, H., Gravel, R. A., and Rozen, R. (1999) *Mol. Genet. Metab.* **67,** 317–323
9. Paine, M. J., Garner, A. P., Powell, D., Sibbald, J., Sales, M., Pratt, N., Smith, T., Tew, D. G., and Wolf, C. R. (2000) *J. Biol. Chem.* **275,** 1471–1478
10. Brenner, C. (2002) *Biochemistry* **41,** 9003–9014
11. Bieganowski, P., Garrison, P. N., Hodawadekar, S. C., Faye, G., Barnes, L. D., and Brenner, C. (2002) *J. Biol. Chem.* **277,** 10852–10860
12. Pace, H. C., and Brenner, C. (2003) *Genome Biol.* **4,** R18
13. Ohta, M., Inoue, H., Cotticelli, M. G., Kastury, K., Baffa, R., Palazzo, J., Siprashvili, Z., Mori, M., McCue, P., Druck, T., Croce, C. M., and Huebner, K. (1996) *Cell* **84,** 587–597
14. Barnes, L. D., Garrison, P. N., Siprashvili, Z., Guranowski, A., Robinson, A. K., Ingram, S. W., Croce, C. M., Ohta, M., and Huebner, K. (1996) *Biochemistry* **35,** 11529–11535
15. Pace, H. C., Garrison, P. N., Robinson, A. K., Barnes, L. D., Draganescu, A., Rosler, A., Blackburn, G. M., Siprashvili, Z., Croce, C. M., Huebner, K., and Brenner, C. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 5484–5489
16. Huang, Y., Garrison, P. N., and Barnes, L. D. (1995) *Biochem. J.* **312,** 925–932
17. Pace, H. C., Hodawadekar, S. C., Draganescu, A., Huang, J., Bieganowski, P., Pekarsky, Y., Croce, C. M., and Brenner, C. (2000) *Curr. Biol.* **10,** 907–917
18. Liu, H., Rodgers, N. D., Jiao, X., and Kiledjian, M. (2002) *EMBO J.* **21,** 4699–4708

19. Salehi, Z., Geffers, L., Vilela, C., Birkenhager, R., Ptushkina, M., Berthelot, K., Ferro, M., Gaskell, S., Hagan, I., Stapley, B., and McCarthy, J. E. (2002) *Mol. Microbiol.* **46,** 49–62
20. Brenner, S. (1974) *Genetics* **77,** 71–94
21. Zorio, D. A., Cheng, N. N., Blumenthal, T., and Spieth, J. (1994) *Nature* **372,** 270–272
22. Moteki, S., and Price, D. (2002) *Mol. Cell* **10,** 599–609
23. Miller, D. M., III, Desai, N. S., Hardin, D. C., Piston, D. W., Patterson, G. H., Fleenor, J., Xu, S., and Fire, A. (1999) *BioTechniques* **26,** 914–918
24. Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991) *EMBO J.* **10,** 3959–3970
25. Ghosh, S., and Lowenstein, J. M. (1997) *Gene* **176,** 249–255
26. Goody, R., Eckstein, F., and Schirmer, R. (1972) *Biochim. Biophys. Acta* **276,** 155–161
27. Buczynski, G., and Potter, R. (1990) *Biochim. Biophys. Acta* **1041,** 296–304
28. Lazarowski, E., Watt, W., Stutts, M., and Brown, H. (1996) *Brit. J. Pharmacol.* **117,** 203–209
29. Draganescu, A., Hodawadekar, S. C., Gee, K. R., and Brenner, C. (2000) *J. Biol. Chem.* **275,** 4555–4560
30. The *C. elegans* Sequencing Consortium (1998) *Science* **282,** 2012–2018
31. Blumenthal, T., and Gleason, K. S. (2003) *Nature Rev. Genet.* **4,** 110–118
32. Huang, T., Kuersten, S., Deshpande, A. M., Spieth, J., MacMorris, M., and Blumenthal, T. (2001) *Mol. Cell. Biol.* **21,** 1111–1120
33. Van Doren, K., and Hirsh, D. (1990) *Mol. Cell. Biol.* **10,** 1769–1772
34. Beelman, C. A., and Parker, R. (1995) *Cell* **81,** 179–183
35. Brenner, C., Garrison, P., Gilmour, J., Peisach, D., Ringe, D., Petsko, G. A., and Lowenstein, J. M. (1997) *Nature Struct. Biol.* **4,** 231–238
36. Lima, C. D., Damico, K. L., Naday, I., Rosenbaum, G., Westbrook, E. M., and Hendrickson, W. A. (1997) *Structure* **5,** 763–774
37. Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J., and Bordonne, R. (2002) *Mol. Cell* **9,** 891–901
38. Verheggen, C., Lafontaine, D. L., Samarsky, D., Mouaikel, J., Blanchard, J. M., Bordonne, R., and Bertrand, E. (2002) *EMBO J.* **21,** 2736–2745
39. Baldauf, S. L., Roger, A. J., Wenk-Siefert, I., and Doolittle, W. F. (2000) *Science* **290,** 972–977
40. Fernandes, M., Xiao, H., and Lis, J. T. (1994) *Nucleic Acids Res.* **22,** 167–173
41. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000) *Nature* **403,** 623–627
42. Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A. R., Sassi, H., Nielsen, P. A., Rasmussen, K. J., Andersen, J. R., Johansen, L. E., Hansen, L. H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sorensen, B. D., Matthiesen, J., Hendrickson, R. C., Gleeson, F., Pawson, T., Moran, M. F., Durocher, D., Mann, M., Hogue, C. W., Figeys, D., and Tyers, M. (2002) *Nature* **415,** 180–183
43. Kemmeren, P., van Berkum, N. L., Vilo, J., Bijma, T., Donders, R., Brazma, A., and Holstege, F. C. (2002) *Mol. Cell* **9,** 1133–1143
44. Brenner, C., Nakayama, N., Goebl, M., Tanaka, K., Toh, E. A., and Matsumoto, K. (1988) *Mol. Cell. Biol.* **8,** 3556–3559
45. Steiger, M., Carr-Schmid, A., Schwartz, D. C., Kiledjian, M., and Parker, R. (2003) *RNA* **9,** 231–238
46. Ingelfinger, D., Arndt-Jovin, D. J., Luhrmann, R., and Achsel, T. (2002) *RNA* **8,** 1489–1501
47. Van Dijk, E., Cougot, N., Meyer, S., Babajko, S., Wahle, E., and Seraphin, B. (2002) *EMBO J.* **21,** 6915–6924
48. Sheth, U., and Parker, R. (2003) *Science* **300,** 805–808