

Arsenic (+ 3 Oxidation State) Methyltransferase and the Methylation of Arsenicals in the Invertebrate Chordate *Ciona intestinalis*

David J. Thomas,^{*,1} Gerardo M. Nava,[†] Shi-Ying Cai,^{‡,§} James L. Boyer,^{‡,§} Araceli Hernández-Zavala,[¶] and H. Rex Gaskins^{†,‡,1}

^{*}Pharmacokinetics Branch, Integrated Systems Toxicology Division, National Health and Environmental Effects Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711; [†]University of Illinois at Urbana-Champaign, Urbana, Illinois 61801; [‡]Mount Desert Island Biological Laboratory, Salisbury Cove, Maine 04672; [§]Liver Center, Yale University School of Medicine, New Haven, Connecticut 06519; and [¶]Center for Environmental Medicine, Asthma, and Lung Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

¹To whom correspondence should be addressed. Fax: 919-541-1937 or 217-333-8286. E-mail: thomas.david@epa.gov or hgaskins@illinois.edu.

Received June 8, 2009; accepted October 1, 2009

Biotransformation of inorganic arsenic (iAs) involves methylation catalyzed by arsenic (+ 3 oxidation state) methyltransferase (*As3mt*) yielding mono-, di-, and trimethylated arsenicals. To investigate the evolution of molecular mechanisms that mediate arsenic biotransformation, a comparative genomic approach focusing on the invertebrate chordate *Ciona intestinalis* was used. Bioinformatic analyses identified an *As3mt* gene in the *C. intestinalis* genome. Constitutive *As3mt* RNA expression was observed in heart, branchial sac, and gastrointestinal tract. Adult animals were exposed to 0 or 1 ppm of iAs for 1 or 5 days. Steady-state *As3mt* RNA expression in the gastrointestinal tract was not modulated significantly by 5 days of exposure to iAs. Tissue levels of iAs and its methylated metabolites were determined by hydride generation-cryotrapping-gas chromatography-atomic absorption spectrometry. At either time point, exposure to iAs significantly increased concentrations of iAs and its methylated metabolites in tissues. After 5 days of exposure, total speciated arsenic concentrations were highest in branchial sac (3705 ng/g), followed by heart (1019 ng/g) and gastrointestinal tract (835 ng/g). At this time point, the sum of the speciated arsenical concentrations in gastrointestinal tract and heart equaled or exceeded that of iAs; in branchial sac, iAs was the predominant species present. *Ciona intestinalis* metabolizes iAs to its methylated metabolites, which are retained in tissues. This metabolic pattern is consistent with the presence of an *As3mt* ortholog in its genome and constitutive expression of the gene in prominent organs, making this basal chordate a useful model to examine the evolution of arsenic detoxification.

Key Words: arsenic; *Ciona intestinalis*; comparative genomics; methyltransferases; toxicogenomics.

Inorganic arsenic (iAs) is a ubiquitous component of the biosphere, and organisms have evolved a variety of strategies to handle this potentially toxic metalloid. In bacteria, genes in the *ars* operon encode proteins that catalyze reduction of pentavalent arsenic (As) to trivalency and extrusion of trivalent arsenic from cells (Lin *et al.*, 2007; Rosen, 2002a,b). The bacterial *arsM* gene encodes an As methyltransferase that facilitates extrusion of As by catalyzing formation of volatile methylarsines (Qin *et al.*, 2006; Yaun *et al.*, 2008). In vertebrates, including humans, ingested or inhaled iAs is converted to methylated species (Hughes, 2006; Tseng, 2007). Therefore, inorganic, methylated, and dimethylated arsenicals are found in urine of humans exposed to iAs. Arsenic (+ 3 oxidation state) methyltransferase (*As3mt*), the prototype As methyltransferase in mammals, was initially identified in *Rattus norvegicus* (Lin *et al.*, 2002; Thomas *et al.*, 2007). Genomic analyses in organisms ranging in complexity from the sea urchin *Strongylocentrotus purpuratus* to *Homo sapiens* have identified candidate genes that could encode *As3mt*-like proteins. For vertebrates, occurrence of a predicted *As3mt* ortholog in the genome correlates with reports of the presence of methylated arsenicals in tissues or urine after exposure to iAs (Li *et al.*, 2005; Thomas *et al.*, 2007). Recombinant *As3mt* from human, rat, and mouse catalyzes the *S*-adenosylmethionine (AdoMet)-dependent conversion of arsenicals to methylated metabolites (Fomenko *et al.*, 2007; Walton *et al.*, 2003).

Here, we have identified an orthologous *As3mt* gene in the sea squirt *Ciona intestinalis* and examined the metabolism of iAs in this primitive chordate. *Ciona intestinalis* is the closest living invertebrate relative of vertebrates, representing a lineage that diverged from the last common ancestor of all vertebrates at least 500 million years ago (Blair and Hedges, 2005; Dehal *et al.*, 2002; Delsuc *et al.*, 2006). Evolution of *C. intestinalis* likely occurred in an environment that contained iAs. Background levels of total iAs in seawater range from about 1 to 5 $\mu\text{g/l}$, although higher levels are often found near

Disclaimer: This manuscript has been reviewed in accordance with the policy of the National Health and Environmental Effects Research Laboratory, U.S. EPA, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

TABLE 1
Conserved Domains in As3mt Amino Acid Sequences Identified in Pfam Database

Species	Accession number	Pfam accession	Pfam description	E value
Human	NP_065733.1	PF08241	Methyltransferase domain	2.00E-14
Rat	NP_543166.1	PF08241	Methyltransferase domain	2.00E-12
Mouse	NP_065602.2	PF08241	Methyltransferase domain	2.00E-14
<i>Ciona intestinalis</i>	ENSCINP00000027827	PF08241	Methyltransferase domain	6.00E-11

point discharge sources (Sidiq, 1992). Adult specimens of *C. intestinalis* filter seawater at a rate of 2–5.5 l/g/day (Goodbody, 1974; Petersen and Svane, 2002). At background levels of total iAs in seawater, this would result in exposures ranging from 2 to 27.5 µg/g/day. If absorption of As from filtered water were 10%, then absorbed doses would range from about 0.2 to 2.8 µg/g/day. Although there are no data on rates of As absorption by *C. intestinalis*, this organism absorbs oxyanionic vanadate from seawater (Bielig *et al.*, 1961). This suggests that an oxyanionic As species (e.g., arsenate) could be absorbed from seawater. In this case, the steady-state concentration of As in tissues of *C. intestinalis* would be determined by fractional absorption of As from filtered seawater and the rate of clearance of As from the organism. Coordinating the formation of methylated arsenicals, some of which are more reactive than iAs, with clearance processes can minimize accumulation of toxic species and reduce the risk associated with exposure to this toxicant (Thomas, 2007; Thomas *et al.*, 2001). Here, we report that *C. intestinalis* rapidly accumulates iAs present in seawater and converts it to mono-, di-, and trimethylated arsenicals. The capacity of *C. intestinalis* to convert iAs is associated with the expression of a gene encoding a protein with considerable sequence identity to As3mt identified in other species.

MATERIALS AND METHODS

Experimental animals. *Ciona intestinalis* specimens were obtained from Marine Biology Laboratory (Woods Hole, MA) and maintained in circulating seawater tanks (average seawater temperature 15°C) at Mount Desert Island

TABLE 2
RBH Analysis of *Ciona intestinalis* As3mt Gene in GenBank

Species	Accession number	E value	Annotation
Human	NP_065733	2.00E-14	AS3MT
Rat	NP_543166	2.00E-12	As3mt
Mouse	NP_065602	2.00E-14	As3mt

Biological Laboratory (<http://www.mdibl.org/>). Animals were adjusted to the new environmental conditions for at least 5 days before use.

Preparation of As-containing seawater. Sodium arsenite (NaAsO₂; Sigma, St Louis, MO) was dissolved in water at 1M concentration as stock solution. It was used at 13.3µM (equivalent to 1 ppm) by diluting the stock solution into fresh seawater containing approximately 1 ppb of iAs.

Exposure conditions. Adult *C. intestinalis* specimens (average length = 5.6 ± 0.5 cm) were maintained in a 15°C seawater tank with the water constantly circulating through a standard aquarium pump. Individual animals were attached to a fiberglass mesh screen via suture of the basal tunic to enable continual submersion and equal distribution in the seawater tank. Animals were randomly divided into two groups ($n = 8-11$): (1) 1 ppm arsenite in seawater and (2) control unamended seawater. Two independent experiments of 1- or 5-day exposure to arsenite (or control seawater) were conducted. For the 5-day treatment, fresh seawater was replaced daily.

Tissue collection. At exposure's end, animals were anesthetized by immersion in a saturated magnesium chloride solution (Petersen and Svane, 2002). Animals were then placed in a petri dish and dissected under a low-magnification microscope. Heart, gastrointestinal tract, and branchial sac were collected separately, washed twice in PBS, placed in cryovials, and snap frozen in liquid N₂. Tissue samples were stored at -20°C until analysis.

Genomic analyses. The sequence of a candidate As3mt gene (ENSC-ING00000015865) was retrieved from the *C. intestinalis* genome (www.ensembl.org/Ciona_intestinalis/index.html). A systematic approach was followed to infer gene homology. This included the use of methods based on hit clustering (Wapinski *et al.*, 2007) and using the results (hits) from sequence similarity searches among highly annotated proteins from human, rat, and mouse to derive an orthology assignment between genes. First, the widely used approach of reciprocal (bidirectional) best hit (RBH) (Wall *et al.*, 2003) was applied. BLASTP (Altschul *et al.*, 1997) analysis was then used to back search the full retrieved amino acid sequence against RefSeq (NCBI Reference Sequence project) protein sequences. Second, the Pfam algorithm (Bateman *et al.*, 2004) was used for the identification of conserved domains within the predicted *C. intestinalis* sequence. Third, the highly annotated and curated SWISS-PROT and TrEMBL protein databases (Bairoch *et al.*, 2004) were used to assess statistical significance of pairwise sequence similarity for the *C. intestinalis* As3mt sequence using probability of random shuffle (PRSS) and SSEARCH analyses (Pearson, 1991; Smith and Waterman, 1981). Finally, BLAST search analysis against the OrthoMCL database was used for the identification of orthologous groups across multiple taxa (Chen *et al.*, 2006). Sequences from the OrthoMCL As3mt orthologous group (accession number OG2_76202) including *Chlamydomonas reinhardtii* (195338), *C. intestinalis* (ENSCINP00000027827), *Cyanidioschyzon merolae_1* (CMT024C), *C. merolae_2* (CME010C), *Danio rerio_1* (ENSDARP00000085262), *D. rerio_2* (ENSDARP00000057878), *Debaryomyces hansenii* (DEHA0D15466g), *Dehalococcoides ethenogenes* 195 (YP_182128.1), *Filobasidiella neoformans* (185.m02734), *Gallus gallus* (ENSGALP00000013184), *Halobacterium* sp. NRC-1 (NP_046066.1), *H. sapiens* (ENSP00000358896), *Mus musculus* (ENSMUSP0000000365), *Neurospora crassa* (NCU05616), *Ostreococcus tauri_1* (Ot17g00760), *O. tauri_2* (Ot17g00590), *R. norvegicus* (ENSRNOP00000051584), *Rhodospirillum baltica* SH 1 (NP_867437.1), *Synechococcus* sp. WH 8102 9 (NP_896777.1), *Tetraodon nigroviridis_1* (GSTENP00022702001), *T. nigroviridis_2* (GSTENP00022701001), *Thalassiosira pseudonana_1* (140233), and *T. pseudonana_2* (21295) were retrieved from OrthoMCL database. In addition, reference protein sequences (NCBI database) from *Bos taurus* (NP_001030195.1), *Canis familiaris* (XP_543995.2), *Macaca mulatta* (XP_001113391.1), *Ornithorhynchus anatinus* (XP_001511844.1), *Pan troglodytes* (XP_508007.2), and *Xenopus tropicalis* (AAI33181.1) were retrieved from the GenBank as well as *Takifugu rubripes* (ENSXETP00000020854) from Ensembl database. These sequences were aligned using the ClustalW software (Larkin *et al.*, 2007), and phylogenetic trees were constructed with the maximum parsimony method (Eck and Dayhoff, 1966) using MEGA software version 4.0 (Tamura *et al.*, 2007).

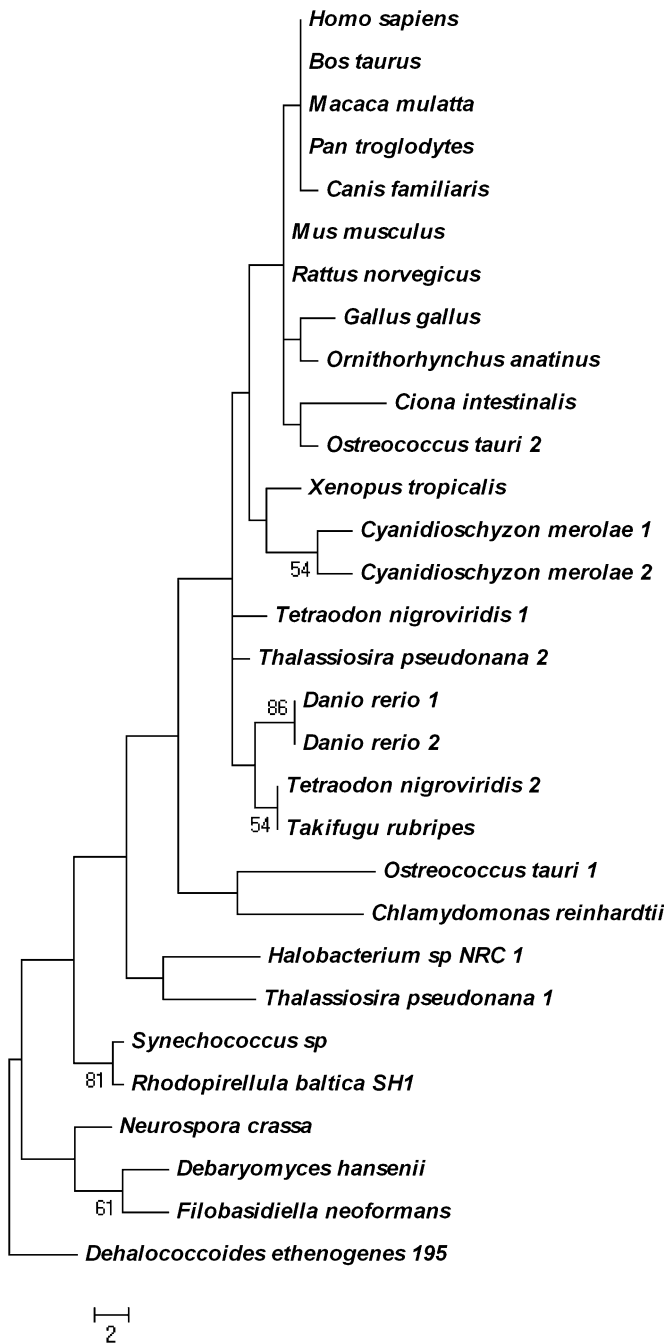


FIG. 2 Phylogeny of As3mt protein sequences in Eukarya, Archaea, and Bacteria. The phylogenetic tree was constructed with the maximum parsimony method using the PAUP* 4.0 b10 program. The statistical significance of branch order (numbers on branches) was estimated by the generation of 1000 replications of bootstrap resampling of the originally aligned amino acid sequences. The deduced amino acid sequence of arsenite *S*-adenosyl methyltransferase (YP_182128) from the bacterium *Dehalococcoides ethenogenes* was used as a root sequence. Scale bar indicates the number of amino acid changes among the aligned sequences.

single As3mt candidate gene (accession number ENSC-ING00000015865) encodes two putative proteins. The larger protein (ENSCINP00000027827, 293 amino acids) is encoded

TABLE 4
In silico Cloning of the As3mt Gene in *Ciona intestinalis*

Accession number	Description	E value
BW379540.1	cDNA library, adult digestive gland	0
BW418659.1	cDNA library, mature adult whole animal	0
BW462498.1	cDNA library, juvenile whole animal	0
BW063681.1	cDNA library, young adult	0
BW107407.1	cDNA library, cleaving embryo	0
AV863336.1	cDNA library, egg	0
BW056695.1	cDNA library, blood cells	0

by eight exons; the last three exons are missing for the shorter protein (ENSCINP00000027826, 199 residues). Both As3mt putative proteins possess the methyltransferase domain (Table 1). However, the ENSCINP00000027827 protein sequence harbors 94 additional amino acids at the C-terminus that are highly conserved among As3mt proteins. Thus, the larger As3mt candidate protein was used for RBH analysis (Wall *et al.*, 2003) against RefSeq databases, which revealed higher homology with well-annotated and biochemically characterized As3mt proteins from mammals (Table 2).

The predicted sequence of the *C. intestinalis* As3mt protein contains features consistent with its identity as an AdoMet-dependent As methyltransferase (Thomas *et al.*, 2007). As shown in Figure 1, it contains the four methyltransferase sequence motifs (I, I', II, and III) that are present in human, rat, and mouse As3mt (Kagan and Clarke, 1994). For each predicted protein, the invariant amino acid residues in the motifs are conserved. In addition, the five highly conserved Cys residues, which have been identified in all As3mt sequences, are also present in *C. intestinalis* As3mt (Fig. 1).

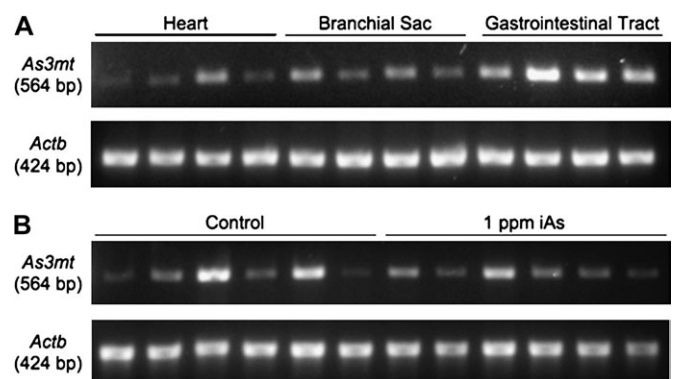


FIG. 3 End-point RT-PCR analysis of constitutive As3mt RNA expression in heart, branchial sac, and gastrointestinal tract of *Ciona Intestinalis* using four replicate animals (A). Steady-state As3mt RNA expression in the gastrointestinal tract was compared in six animals exposed to 1 ppm arsenite in seawater for 5 days to six animals maintained under similar conditions in control unamended seawater to examine the extent to which As3mt expression is modulated by exposure to iAs (B). The RNA isolation protocol, primer sequences, and reaction conditions are provided in the "Materials and Methods" section.

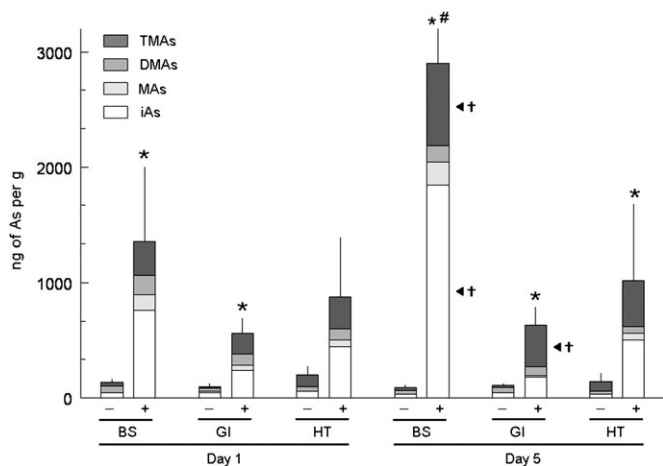


FIG. 4 Distribution of arsenicals in tissues of *Ciona intestinalis* following exposure to unamended seawater (–) or seawater with 1 ppm of arsenic (as arsenite) (+) for 1 day (day 1) or 5 days (day 5). Results shown for branchial sac (BS), gastrointestinal tract (GI), and heart (HT). Detected species are iAs, MAs, DMAs, and TMAAs. Means and SDs shown for summed concentrations of speciated arsenicals in tissues ($n = 3–8$). *A significant difference ($p \leq 0.008$) between treatment groups in the summed concentrations of arsenicals in a tissue on day 1 or day 5. #A significant difference ($p = 0.002$) in the summed concentrations of arsenicals in the branchial sac between days 1 and 5. †A significant difference ($p \leq 0.003$) in the concentration of a speciated arsenical in a tissue between days 1 and 5.

The function of As3mt as an As methyltransferase is dependent on the presence of Cys156 in rat and Cys157 in mouse. Replacement of these residues with Ser leads to a loss of catalytic function (Fomenko *et al.*, 2007; Li *et al.*, 2005); replacement of both Cys156 and Cys205 with Ser in human AS3MT nullifies its catalytic function (Thomas *et al.*, 2007). The persistence of these Cys residues in species that diverged over a period of about 500 million years suggests that their presence determines the protein's catalytic activity.

The statistical significance of pairwise sequence similarity for the *C. intestinalis* As3mt protein using PRSS and SSEARCH analyses (Pearson, 1991; Smith and Waterman, 1981) is consistent with its high homology to human, rat, and mouse proteins (Table 3). Both construction of orthologous groups across multiple eukaryotic taxa (OrthoMCL algorithm) and protein phylogenetic analysis (Fig. 2) reveal a clear evolutionary relation among homologous As3mt proteins. The phylogenetic tree indicates that As3mt might have diverged from the base of the chordate tree in another lower eukaryote. In sum, these analyses validate the accuracy of the manual annotation of the *As3mt* gene in *C. intestinalis*.

In silico cloning (Boguski *et al.*, 1994) of *As3mt* in *C. intestinalis* confirms the presence of its transcripts in tissues from juvenile and mature adult organisms and in the egg and dividing embryo (Table 4). Moreover, RT-PCR analysis from replicate animals demonstrated constitutive expression of *As3mt* in heart, branchial sac, and gastrointestinal tract (Fig. 3A). The level of steady-state *As3mt* RNA expression

was variable among replicate animals for all three tissues, and expression in the gastrointestinal tract did not appear to be modulated by 5 days of exposure to iAs (Fig. 3B). *As3mt* is also constitutively expressed in many tissues of the adult rat (Lin *et al.*, 2002). Although exposure to iAs does not appear to increase *As3mt* gene transcription or protein levels in mice (Drobna *et al.*, 2009; Xing, unpublished data), single nucleotide polymorphisms in the human *As3mt* gene, sequence variation in 5' flanking regions, and the number of tandem repeats in the 5' untranslated region affect protein expression and activity (Wood *et al.*, 2006). Hence, regulation of As3mt activity and expression is probably a complex interplay of transcriptional and translational controls.

In these experiments, organisms were exposed to iAs (added as arsenite) for up to 5 days. The background concentration of iAs in seawater used in these studies was approximately 0.001 ppm; in amended seawater, iAs concentration was approximately 1 ppm. Tissues from control animals maintained in unamended seawater contained iAs and its methylated metabolites (Fig. 4). The presence of methylated arsenicals in tissues under basal exposure conditions indicates that these organisms constitutively methylate iAs. At day 1 of exposure, the summed concentrations of speciated arsenicals in branchial sac or gastrointestinal tract from animals in amended seawater versus unamended seawater were significantly different ($p < 0.001$). However, the summed concentrations of speciated arsenicals in heart were not significantly different in these two treatment groups on day 1 of exposure. At day 5 of exposure, the summed concentrations of speciated arsenicals in all tissues from animals in amended seawater or unamended seawater were significantly different ($p \leq 0.008$). The summed concentrations of speciated arsenicals in tissues of control animals maintained in unamended seawater did not significantly change ($p > 0.05$) between days 1 and 5 of exposure. *Ciona intestinalis* exposed to iAs in amended seawater quickly accumulated iAs and its methylated arsenicals in tissues. After 1 day of exposure, iAs and methylated arsenicals were present in all tissues; branchial sac contained the highest summed concentrations of speciated arsenicals. Although the summed concentrations of speciated arsenicals increased in all tissues between days 1 and 5 of exposure, this increase was statistically significant ($p = 0.002$) only in branchial sac. In this tissue, the concentrations of iAs ($p = 0.001$) and TMAAs ($p = 0.001$) were significantly different on days 1 and 5 of exposure. In the gastrointestinal tract, the concentration of TMAAs was significantly different ($p = 0.003$) on days 1 and 5 of exposure.

Notably, the pattern of metabolites in *C. intestinalis* exposed to iAs resembles that found in chronically exposed rats. In rats exposed to arsenate in drinking water for 2 weeks, the As-containing methylated metabolites DMA and TMA were found in tissues and excreta (Adair *et al.*, 2007). Recombinant rat As3mt has been shown to efficiently catalyze the mono-, di-, and trimethylation reactions (Waters *et al.*,

2004a). Catalysis of As methylation reactions by rat As3mt is dependent on physiological reductants (thioredoxin, glutaredoxin, and dihydrolipoic acid), suggesting that these or similar dithiol reductants must be available in *C. intestinalis*. In addition, glutathione (GSH) is an important modulator of the activity of rat As3mt (Waters *et al.*, 2004b). Addition of GSH to *in vitro* reaction mixtures containing rat As3mt, AdoMet, and a coupled thioredoxin/thioredoxin reductase/NADPH system suppressed formation of TMAs from either arsenite or dimethylarsinous acid. A recent analysis of the genome of *C. intestinalis* annotated a number of genes involved in the biosynthesis and utilization of GSH (Nava *et al.*, 2009). These investigators identified genes encoding the two enzymes required for GSH synthesis (catalytic and regulatory subunits of glutamate cysteine ligase and glutathione synthetase), which were expressed in branchial sac, heart, and gastrointestinal tract of adult *C. intestinalis*. Further investigation is required to determine whether GSH may modulate As methylation in *C. intestinalis*.

Other factors could also affect the pattern of As-containing metabolites found in tissues. Translocation of iAs and its methylated metabolites among tissues during the course of exposure could obscure differences among tissues in the capacity to form metabolites. Trivalent inorganic and methylated arsenicals enter eukaryotic cells via aquaporins and hexose permease transporters and exit cells via ATP-binding cassette (ABC) transporters (Thomas, 2007). Hence, intracellular concentrations of individual methylated arsenicals will be determined by the balance of influx and efflux processes. Variation among tissues in expression of these membrane-spanning proteins may determine levels of individual metabolites in different tissues. Genes encoding aquaporin-like proteins and ABC transporters have been identified in *C. intestinalis* (Annilo *et al.*, 2006; Okamura *et al.*, 2005). Notably, the branchial sac, which consistently showed the highest levels of As accumulation, also avidly accumulates vanadium (V) added as oxyanionic vanadate (Cheney *et al.*, 1997; Michibata, 1993). The relation, if any, between accumulation of V and of As in tissues is unclear. For V, there is evidence of specific binding proteins in tissues that sequester this element (Trivedi *et al.*, 2003). Analogous proteins in *C. intestinalis* could be involved in the binding of iAs or its methylated metabolites. For example, nicatransferrin, a monolobal transferrin from *C. intestinalis* (Tinoco *et al.*, 2008), could bind arsenicals in a manner analogous to As binding by mammalian transferrin (Zhang *et al.*, 1998).

Functional similarities between the metabolism of As in *C. intestinalis* and in more complex organisms support the hypothesis that enzymatically catalyzed methylation of As is a fundamental component of the organismic response to this metalloid, which arose early in the evolution of the chordate lineage. Similarities among developmental and physiological processes in this primitive chordate and higher organisms may make *C. intestinalis* a useful model organism for studies of

molecular processes involved in the accumulation, methylation, binding, and excretion of As and for studies of modes of action of iAs and its methylated metabolites as toxins or carcinogens.

FUNDING

National Institute for Environmental Health Sciences Center for Comparative Toxicology (P30 ES03828) at Mount Desert Island Biological Laboratory; DK 25636 to J.L.B.; Yale Liver Center (P30 34989).

ACKNOWLEDGMENTS

The authors thank Ann Benefiel and Debbie Piper for editorial assistance. The authors declare no conflict of interest.

REFERENCES

- Adair, B. M., Moore, T., Conklin, S. D., Creed, J. T., Wolf, D. C., and Thomas, D. J. (2007). Tissue distribution and urinary excretion of dimethylated arsenic and its metabolites in dimethylarsinic acid- or arsenate-treated rats. *Toxicol. Appl. Pharmacol.* **222**, 235–242.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.
- Annilo, T., Chen, Z. Q., Shulenin, S., Costantino, J., Thomas, L., Lou, H., Stefanov, S., and Dean, M. (2006). Evolution of the vertebrate ABC gene family: analysis of gene birth and death. *Genomics* **88**, 1–11.
- Bairoch, A., Boeckmann, B., Ferro, S., and Gasteiger, E. (2004). Swiss-Prot: juggling between evolution and stability. *Brief Bioinform.* **5**, 39–55.
- Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L., *et al.* (2004). The Pfam protein families database. *Nucleic Acids Res.* **32**, D138–D141.
- Bielig, H. J., Pflieger, K., Rummel, W., and Seifen, E. (1961). Uptake and distribution of vanadium by the tunicate *Ciona intestinalis* L. Hoppe Seylers *Z. Physiol. Chem.* **326**, 249–258.
- Blair, J. E., and Hedges, S. B. (2005). Molecular phylogeny and divergence times of deuterostome animals. *Mol. Biol. Evol.* **22**, 2275–2284.
- Boguski, M. S., Lowe, T. M., and Tolstoshev, C. M. (1993). dbEST—database for “expressed sequence tags.” *Nat. Genet.* **4**, 332–333.
- Boguski, M. S., Tolstoshev, C. M., and Bassett, D. E., Jr. (1994). Gene discovery in dbEST. *Science* **265**, 1993–1994.
- Chen, F., Mackey, A. J., Stoeckert, C. J., Jr, and Roos, D. S. (2006). OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res.* **34**, D363–D368.
- Cheney, M. A., Berg, J. R., and Swinehart, J. H. (1997). The uptake of vanadium (V) and other metals by the isolated branchial sacs of the ascidians *Ascidia ceratodes*, *Ciona intestinalis*, and *Styela montereyensis*. *Comp. Biochem. Physiol.* **116C**, 49–153.
- Dehal, P., Satou, Y., Campbell, R. K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D. M., *et al.* (2002). The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* **298**, 2157–2167.
- Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* **439**, 965–968.

- Drobna, Z., Naranmandura, H., Kubachka, K. M., Edwards, B. M., Herbin-Davis, K., Styblo, M., Le, X. C., Creed, J. T., Maeda, N., Hughes, M. F., et al. (2009). Disruption of the arsenic (+3 oxidation state) methyltransferase gene in the mouse alters the phenotype for methylation of arsenic and affects distribution and retention of orally administered arsenate. *Chem. Res. Toxicol.* **22**, 1713–1720.
- Eck, R. V., and Dayhoff, M. O. (1966). In *Atlas of Protein Sequence and Structure*. National Biomedical Research Foundation, Silver Spring, MD.
- Fomenko, D. E., Xing, W., Adair, B. M., Thomas, D. J., and Gladyshev, V. N. (2007). High-throughput identification of catalytic redox-active cysteine residues. *Science* **315**, 387–389.
- Goodbody, I. (1974). The physiology of ascidians. In *Advances in Marine Biology* (F. S. Russell and M. Yonge, Eds), Vol. 12, pp. 1–149. Academic Press, London, UK.
- Hernández-Zavala, A., Matousek, T., Drobna, Z., Paul, D. S., Walton, F., Adair, B. M., Jiri, D., Thomas, D. J., and Stýblo, M. (2008). Speciation analysis of arsenic in biological matrices by automated hydride generation-cryotrapping-atomic absorption spectrometry with multiple microflame quartz tube atomizer (multiatomizer). *J. Anal. At. Spectrom.* **23**, 342–351.
- Hughes, M. F. (2006). Biomarkers of exposure: a case study with inorganic arsenic. *Environ. Health Perspect.* **114**, 1790–1796.
- Hughes, M. F., Devesa, V., Adair, B. M., Stýblo, M., Kenyon, E. M., and Thomas, D. J. (2005). Tissue dosimetry, metabolism and excretion of pentavalent and trivalent monomethylated arsenic in mice after oral administration. *Toxicol. Appl. Pharmacol.* **208**, 186–197.
- Kagan, R. M., and Clarke, S. (1994). Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. *Arch. Biochem. Biophys.* **310**, 417–427.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., et al. (2007). ClustalW and ClustalX version 2. *Bioinformatics* **23**, 2947–2948.
- Li, J., Waters, S. B., Drobna, Z., Devesa, V., Stýblo, M., and Thomas, D. J. (2005). Arsenic (+3 oxidation state) methyltransferase and the inorganic arsenic methylation phenotype. *Toxicol. Appl. Pharmacol.* **204**, 164–169.
- Lin, S., Shi, Q., Nix, F. B., Stýblo, M., Beck, M. A., Herbin-Davis, K. M., Hall, L. L., Simeonsson, J. B., and Thomas, D. J. (2002). A novel S-adenosyl-L-methionine:arsenic(III) methyltransferase from rat liver cytosol. *J. Biol. Chem.* **277**, 10795–10803.
- Lin, Y. F., Yang, J., and Rosen, B. P. (2007). ArsD: an As(III) metallochaperone for the ArsAB As(III)-translocating ATPase. *J. Bioenerg. Biomembr.* **39**, 453–458.
- Michibata, H. (1993). The mechanism of accumulation of high levels of vanadium by ascidians from seawater: biophysical approaches to a remarkable phenomenon. *Adv. Biophys.* **29**, 105–133.
- Nava, G. M., Lee, D. Y., Ospina, J. H., Cai, S.-Y., and Gaskins, H. R. (2009). Genomic analyses reveal a conserved glutathione homeostasis pathway in the invertebrate chordate *Ciona intestinalis*. *Physiol. Genomics* **39**, 183–194.
- Okamura, Y., Nishino, A., Murata, Y., Nakajo, K., Iwasaki, H., Ohtsuka, Y., Tanaka-Kunishima, M., Takahashi, N., Hara, Y., Yoshida, T., et al. (2005). Comprehensive analysis of the ascidian genome reveals novel insights into the molecular evolution of ion channel genes. *Physiol. Genomics* **11**, 269–282.
- Pearson, W. R. (1991). Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. *Genomics* **11**, 635–650.
- Petersen, J. K., and Svane, I. (2002). Filtration rate in seven Scandinavian ascidians: implications of the morphology of the gill sac. *Mar. Biol.* **140**, 397–402.
- Qin, J., Rosen, B. P., Zhang, Y., Wang, G., Franke, S., and Rensing, C. (2006). Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 2075–2080.
- Rosen, B. P. (2002a). Biochemistry of arsenic detoxification. *FEBS Lett.* **529**, 86–92.
- Rosen, B. P. (2002b). Transport and detoxification systems for transition metals, heavy metals and metalloids in eukaryotic and prokaryotic microbes. *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.* **133**, 689–693.
- Sidiq, M. (1992). Arsenic in marine environments. In *Toxic Metal Chemistry in Marine Environments*, pp. 61–105. Dekker, New York.
- Smith, T. F., and Waterman, M. S. (1981). Identification of common molecular subsequences. *J. Mol. Biol.* **147**, 195–197.
- Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. 2007. *Mol. Biol. Evol.* **24**, 1596–1599.
- Thomas, D. J. (2007). Molecular processes in cellular arsenic metabolism. *Toxicol. Appl. Pharmacol.* **222**, 365–373.
- Thomas, D. J., Li, J., Waters, S. B., Xing, W., Adair, B. M., Drobna, Z., Devesa, V., and Stýblo, M. (2007). Arsenic (+3 oxidation state) methyltransferase and the methylation of arsenicals. *Exp. Biol. Med.* **232**, 3–13.
- Thomas, D. J., Styblo, M., and Lin, S. (2001). The cellular metabolism and systemic toxicity of arsenic. *Toxicol. Appl. Pharmacol.* **176**, 127–144.
- Tinoco, A. D., Peterson, C. W., Lucchese, B., Doyle, R. P., and Valentine, A. M. (2008). On the evolutionary significance and metal-binding characteristics of a monolobal transferrin from *Ciona intestinalis*. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 3268–3273.
- Trivedi, S., Ueki, T., Yamaguchi, N., and Michibata, H. (2003). Novel vanadium-binding proteins (vanabins) identified in cDNA libraries and the genome of the ascidian *Ciona intestinalis*. *Biochim. Biophys. Acta* **1630**(2–3), 64–70.
- Tseng, C. H. (2007). Arsenic methylation, urinary arsenic metabolites and human diseases: current perspective. *J. Environ. Sci. Health C. Environ. Carcinog. Ecotoxicol. Rev.* **25**, 1–22.
- Wall, D. P., Fraser, H. B., and Hirsh, A. E. (2003). Detecting putative orthologs. *Bioinformatics* **19**, 1710–1711.
- Wallace, I. M., O'Sullivan, O., Higgins, D. G., and Notredame, C. (2006). M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucl. Acids Res.* **34**, 1692–1699.
- Walton, F. S., Waters, S. B., Jolley, S. L., LeCluyse, E. L., Thomas, D. J., and Stýblo, M. (2003). Selenium compounds modulate the activity of recombinant rat AsIII-methyltransferase and the methylation of arsenite by rat and human hepatocytes. *Chem. Res. Toxicol.* **16**, 261–265.
- Wapinski, I., Pfeffer, A., Friedman, N., and Regev, A. (2007). Automatic genome-wide reconstruction of phylogenetic gene trees. *Bioinformatics* **23**, 549–558.
- Waters, S. B., Devesa, V., Del Razo, L. M., Stýblo, M., and Thomas, D. J. (2004a). Endogenous reductants support the catalytic function of recombinant rat cyt19, an arsenic methyltransferase. *Chem. Res. Toxicol.* **17**, 404–409.
- Waters, S. B., Devesa, V., Fricke, M. W., Creed, J. T., Stýblo, M., and Thomas, D. J. (2004b). Glutathione modulates recombinant rat arsenic (+3 oxidation state) methyltransferase-catalyzed formation of trimethylarsine oxide and trimethylarsine. *Chem. Res. Toxicol.* **17**, 1621–1629.
- Wood, T. C., Salavagionne, O. E., Mukherjee, B., Wang, L., Klumpp, A. F., Thomae, B. A., Eckloff, B. W., Schaid, D. J., Wieben, E. D., and Weinshilboum, R. M. (2006). Human arsenic methyltransferase (AS3MT) pharmacogenetics: gene resequencing and functional genomics studies. *J. Biol. Chem.* **281**, 7364–7373.
- Yaun, C., Lu, X., Qin, J., Rosen, B. P., and Le, X. C. (2008). Volatile arsenic species released from *E. coli* expressing AsIII S-adenosylmethionine methyltransferase gene. *Environ. Sci. Technol.* **42**, 3201–3206.
- Zhang, X., Cornelis, R., DeKimpe, J., Mees, L., and Lamiere, N. (1998). Study of arsenic-protein binding in serum of patients on continuous ambulatory peritoneal dialysis. *Clin. Chem.* **44**, 141–147.
- Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *J. Comp. Biol.* **7**, 203–214.