Submitted Thursday, July 31, 2014
Revised version submitted Tuesday, September 09, 2014

The pathway of human AS3MT arsenic methylation

Supplemental Information

Dharmendra S. Dheeman, Charles Packianathan, Jitesh K. Pillai and Barry P. Rosen[1]

Department of Cellular Biology and Pharmacology, Herbert Wertheim College of Medicine, Florida International University, Miami, FL 33199 USA

Running title: AS3MT mechanism

To whom correspondence should be addressed: Barry P. Rosen, Department of Cellular Biology and Pharmacology, Herbert Wertheim College of Medicine, Florida International University, Miami, FL 33199 USA, brosen@fiu.edu

Table S1: Oligonucleotide primers for mutagenesis of cysteine residues

Primer	Sequence
Oligonucleotides for mutagenesis of conserved cysteine residues	
C32S forward	5'-GCTGATCTGCAAACGAACGGTAGCGTGACCACG-3'
reverse	5'-CGTGGTCACGCTACCGTTCGTTTGCAGATCAGC-3'
C61S forward	5'-CTGCGTTATTACGGCAGCGGTCTGGTGATCC-3'
reverse	5'-GGATCACCAGACCGCTGCCGTAATAACGCAG-3'
C156S forward	5'-ATGATATCGTGGTTTCTAACAGCGTCATTAATCTGGTGC-3'
reverse	5'-GCACCAGATTAATGACGCTGTTAGAAACCACGATATCAT-3'
C206S forward	5'-GTTCTGTGGGGTGAAAGCCTGGGCGGTG-3'
reverse	5'-CACCGCCCAGGCTTTCACCCCACAGAAC-3'
Oligonucleotides for mutagenesis of non-conserved cysteine residues	
C72S forward	5'-CCCGGAACATCTGGAAAACAGTTGGATTCTGGATC-3'
reverse	5'-GATCCAGAATCCAACTGTTTTCCAGATGTTCCGGG-3'
C85S forward	5'-GGTTCTGGCCGCGACAGTTATGTTCTGAGCC-3'
reverse	5'-GGCTCAGAACATAACTGTCGCGGCCAGAACC-3'
C250S forward	5'-TGGAACGTGTTATTGGCGATAGCCGCTTCGTC-3'
reverse	5'-GACGAAGCGGCTATCGCCAATAACACGTTCCA-3'

Table S2: Oligonucleotide primers for mutagenesis of tryptophan residues

Primer	Sequence
W73R forward	5'-TCCCGGAACATCTGGAAAACTGTAGGATTCTGGATC-3'
reverse	5'-GATCCAGAATCCTACAGTTTTCCAGATGTTCCGGGA-3'
W73L forward	5'-AACATCTGGAAAACTGTTTGATTCTGGATCTGGGCAG-3'
reverse	5'-CTGCCCAGATCCAGAATCAAACAGTTTTCCAGATGTT-3'
W203L forward	5'-TCGCACGCACAAAGTTCTGTTGGGTGAATGCC-3'
reverse	5'-GGCATTCACCCAACAGAACTTTGTGCGTGCGA-3'
W213L forward	5'-GCGGTGCACTGTACTTGAAGGAACTGGCAGT-3'
reverse	5'-ACTGCCAGTTCCTTCAAGTACAGTGCACCGC-3'
W73R/W203Y forward	5′-TCGCACGCACAAAGTTCTGTATGGTGAATGCCTGGG-3′
reverse	5′-CCCAGGCATTCACCATACAGAACTTTGTGCGTGCGA-3′
W73R/W213Y forward	5′-GGCGGTGCACTGTACTATAAGGAACTGGCAGTGC-3′
reverse	5′-GCACTGCCAGTTCCTTATAGTACAGTGCACCGCC-3′
W203L/W213Y forward	5′-GGCGGTGCACTGTACTATAAGGAACTGGCAGTGC-3′
reverse	5′-GCACTGCCAGTTCCTTATAGTACAGTGCACCGCC-3′

Fig. S1. Nucleotide sequence of the hAS3MT cDNA clone and the synthetic gene. Conserved nucleotides are indicated (*).

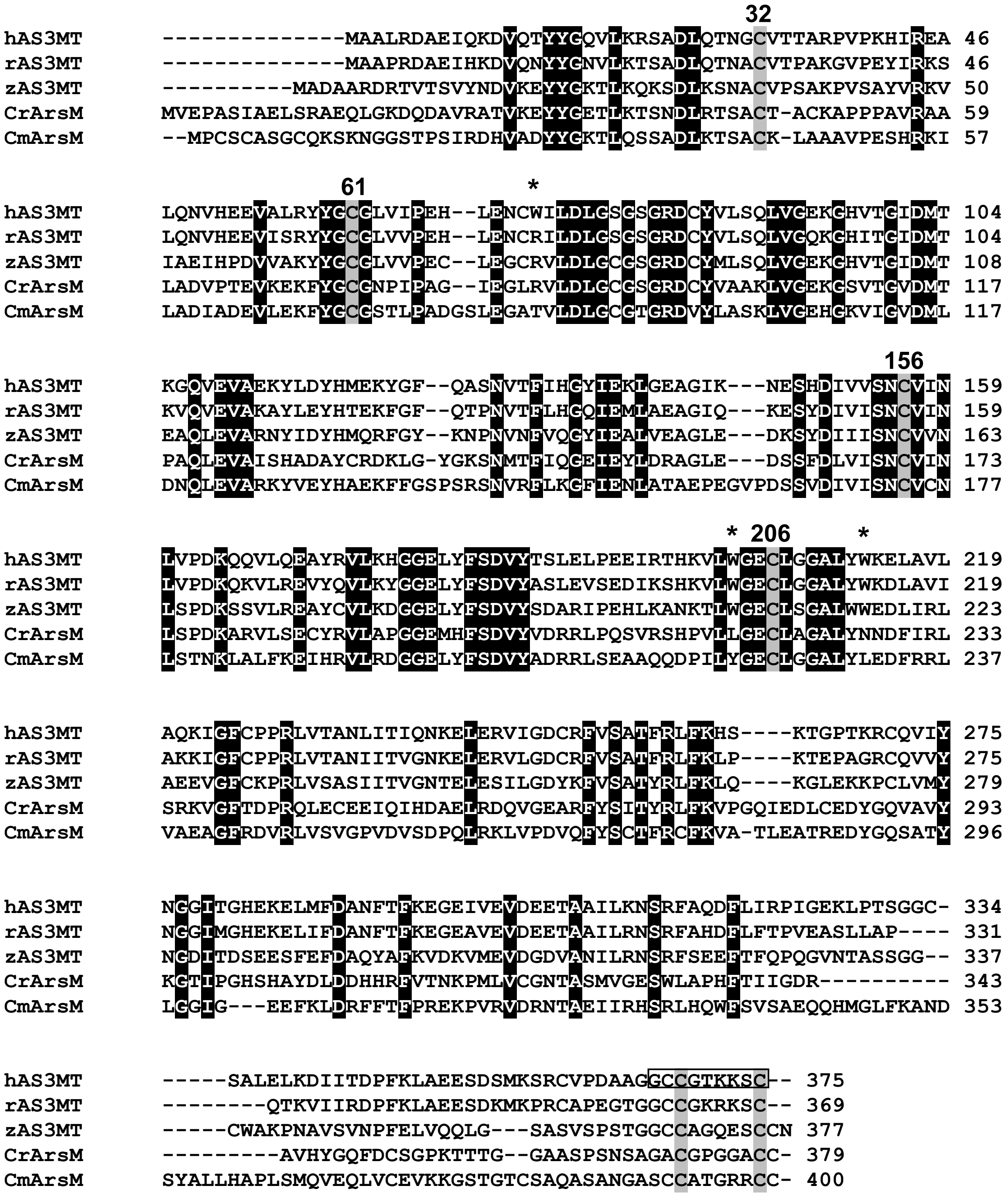
```
cDNA       ATGGCTGCACTTCGTGACGCTGAGATACAGAAGGACGTGCAGACCTACTACGGGCAGGTG  60
Synthetic  ATGGCTGCTCTGCGTGATGCCGAAATCCAAAAAGATGTCCAAACCTACTACGGTCAAGTG  60
           ******** ** ***** ** ** ** ** ** ** ** ** *********** ** ***

cDNA       CTGAAGAGATCGGCAGACCTCCAGACCAACGGCTGTGTCACCACAGCCAGGCCGGTCCCC  120
Synthetic  CTGAAACGCTCCGCTGATCTGCAAACGAACGGTTGCGTGACCACGGCACGTCCGGTTCCG  120
           *****  * ** ** ** ** ** ** ***** ** ** ***** **  * ***** **

cDNA       AAGCACATCCGGGAAGCCTTGCAAAATGTACACGAAGAAGTAGCCCTAAGATATTATGGC  180
Synthetic  AAACATATCCGCGAAGCGCTGCAGAATGTGCACGAAGAAGTTGCCCTGCGTTATTACGGC  180
           ** ** ***** *****  **** ***** *********** *****  * ***** ***

cDNA       TGTGGTCTGGTGATCCCTGAGCATCTAGAAAACTGCTGGATTTTGGATCTGGGTAGTGGA  240
Synthetic  TGCGGTCTGGTGATCCCGGAACATCTGGAAAACTGTTGGATTCTGGATCTGGGCAGCGGT  240
           ** ************** ** ***** ******** ****** ********** ** **

cDNA       AGTGGCAGAGATTGCTATGTACTTAGCCAGCTGGTTGGTGAAAAAGGACACGTGACTGGA  300
Synthetic  TCTGGCCGCGACTGTTATGTTCTGAGCCAGCTGGTCGGTGAAAAAGGCCACGTGACCGGT  300
             **** * ** ** ***** ** *********** *********** ******** **

cDNA       ATAGACATGACCAAAGGCCAGGTGGAAGTGGCTGAAAAGTATCTTGACTATCACATGGAA  360
Synthetic  ATTGATATGACGAAAGGCCAAGTCGAAGTGGCAGAAAAGTATCTGGACTACCATATGGAA  360
           ** ** ***** ******** ** ******** *********** ***** ** ******

cDNA       AAATATGGCTTCCAGGCATCTAATGTGACTTTTATTCATGGCTACATTGAGAAGTTGGGA  420
Synthetic  AAATATGGTTTTCAGGCTTCTAACGTTACCTTCATTCACGGCTACATCGAAAAACTGGGT  420
           ******** ** ***** ***** ** ** ** ***** ******** ** **  ****

cDNA       GAGGCTGGAATCAAGAATGAGAGCCATGATATTGTTGTATCAAACTGTGTTATTAACCTT  480
Synthetic  GAAGCGGGCATTAAGAATGAAAGCCATGATATCGTGGTTTCTAACTGCGTCATTAATCTG  480
           ** ** ** ** ******** *********** ** ** ** ***** ** ***** **

cDNA       GTGCCTGATAAACAACAAGTGCTTCAGGAGGCATATCGGGTGCTGAAGCATGGTGGGGAG  540
Synthetic  GTGCCGGACAAGCAGCAAGTCCTGCAAGAAGCCTATCGTGTGCTGAAACATGGCGGTGAA  540
           ***** ** ** ** ***** ** ** ** ** ***** ******** ***** ** **

cDNA       TTATATTTCAGTGACGTCTATACGAGCCTTGAACTGCCAGAAGAAATCAGGACACACAAA  600
Synthetic  CTGTATTTTTCAGATGTCTACACCTCGCTGGAACTGCCGGAAGAAATTCGCACGCACAAA  600
            * *****    ** ***** **    ** ******** ********  * ** ******

cDNA       GTTTTATGGGGTGAGTGTCTGGGTGGTGCTTTATACTGGAAGGAACTTGCTGTCCTTGCT  660
Synthetic  GTTCTGTGGGGTGAATGCCTGGGCGGTGCACTGTACTGGAAGGAACTGGCAGTGCTGGCT  660
           *** * ******** ** ***** *****  * ************** ** ** ** ***

cDNA       CAAAAAATTGGGTTCTGCCCTCCACGTTTGGTCACTGCCAATCTCATTACAATTCAAAAC  720
Synthetic  CAGAAAATCGGCTTTTGTCCGCCGCGTCTGGTTACCGCAAACCTGATTACGATCCAAAAT  720
           ** ***** ** ** ** ** ** *** **** ** ** ** ** ***** ** *****

cDNA       AAGGAACTGGAAAGAGTTATCGGTGACTGTCGTTTTGTTTCTGCAACATTTCGCCTCTTC  780
Synthetic  AAGGAACTGGAACGTGTTATTGGCGATTGCCGCTTCGTCAGTGCTACCTTTCGTCTGTTC  780
           ************ * ***** ** ** ** ** ** **   *** ** ***** ** ***

cDNA       AAACACTCTAAGACAGGACCAACCAAGAGATGCCAAGTTATTTACAATGGAGGAATTACA  840
Synthetic  AAACATTCCAAGACCGGTCCGACGAAACGCTGTCAAGTGATTTATAACGGCGGTATCACC  840
           ***** ** ***** ** ** ** **  * ** ***** ***** ** ** ** ** **
```

```
cDNA       GGACATGAAAAAGAACTAATGTTTGATGCCAATTTTACATTTAAGGAAGGTGAAATTGTT  900
Synthetic  GGCCACGAAAAGGAACTGATGTTCGATGCGAATTTTACGTTCAAAGAAGGTGAAATCGTG  900
            ** ** ***** ***** ***** ***** ******** ** ** *********** **

cDNA       GAAGTGGATGAAGAAACAGCAGCTATCTTGAAGAATTCAAGATTTGCTCAAGATTTTCTG  960
Synthetic  GAAGTGGATGAAGAAACCGCGGCCATTCTGAAAAATAGCCGTTTTGCGCAGGATTTCCTG  960
           ***************** ** ** **  **** ***    * ***** ** ***** ***

cDNA       ATCAGACCAATTGGAGAGAAGTTGCCAACATCTGGAGGCTGTTCTGCTTTGGAGTTAAAG  1020
Synthetic  ATTCGCCCGATCGGCGAAAAGCTGCCGACGAGTGGCGGTTGTTCCGCCCTGGAACTGAAA  1020
           **  * ** ** ** ** *** **** **   *** ** ***** **  ****  * **

cDNA       GATATAATCACAGATCCATTTAAGCTTGCAGAAGAGTCTGACAGTATGAAGTCCAGATGT  1080
Synthetic  GATATTATCACCGACCCGTTCAAACTGGCGGAAGAATCAGACTCAATGAAATCACGCTGT  1080
           ***** ***** ** ** ** ** ** ** ***** ** ***   ***** **  * ***

cDNA       GTCCCTGATGCTGCTGGA  1098
Synthetic  GTGCCGGATGCTGCGGGT  1098
           ** ** ******** **
```

Fig. S2. Amino acid sequence of synthetic hAS3MT. Nucleotide and amino acid sequence of synthetic hAS3MT. Amino acid residues 1-218 are GST; residues 37-42 are a thrombin site; residues 73-38 are an enterokinase site; residues 280-645 are hAS3MT (underlined), and residues 655-662 are the eight histidine tag.

```
atgtcccctatactaggttattggaaaattaagggccttgtgcaacccactcgacttcttttggaatatcttgaa
 M  S  P  I  L  G  Y  W  K  I  K  G  L  V  Q  P  T  R  L  L  L  E  Y  L  E   25
gaaaaatatgaagagcatttgtatgagcgcgatgaaggtgataaatggcgaaacaaaaagtttgaattgggtttg
 E  K  Y  E  E  H  L  Y  E  R  D  E  G  D  K  W  R  N  K  K  F  E  L  G  L   50
gagtttcccaatcttccttattatattgatggtgatgttaaattaacacagtctatggccatcatacgttatata
 E  F  P  N  L  P  Y  Y  I  D  G  D  V  K  L  T  Q  S  M  A  I  I  R  Y  I   75
gctgacaagcacaacatgttgggtggttgtccaaaagagcgtgcagagatttcaatgcttgaaggagcggttttg
 A  D  K  H  N  M  L  G  G  C  P  K  E  R  A  E  I  S  M  L  E  G  A  V  L   100
gatattagatacggtgtttcgagaattgcatatagtaaagactttgaaactctcaaagttgattttcttagcaAG
 D  I  R  Y  G  V  S  R  I  A  Y  S  K  D  F  E  T  L  K  V  D  F  L  S  K   125
ctacctgaaatgctgaaaatgttcgaagatcgtttatgtcataaaacatatttaaatggtgatcatgtaacccat
 L  P  E  M  L  K  M  F  E  D  R  L  C  H  K  T  Y  L  N  G  D  H  V  T  H   150
cctgacttcatgttgtatgacgctcttgatgttgttttatacatggacccaatgtgcctggatgcgttcccaaaa
 P  D  F  M  L  Y  D  A  L  D  V  V  L  Y  M  D  P  M  C  L  D  A  F  P  K   175
ttagtttgttttaaaaaacgtattgaagctatcccacaaattgataagtacttgaaatccagcaagtatatagca
 L  V  C  F  K  K  R  I  E  A  I  P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A   200
tggcctttgcagggctggcaagccacgtttggtggtggcgaccatcctccaaaatcggatggttcaactagtggT
 W  P  L  Q  G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D  G  S  T  S  G   225
tctggtcatcaccatcaccatcactccgcgggtctggtgccacgcggtagtactgcaattggtatgaaagaaacc
 S  G  H  H  H  H  H  H  S  A  G  L  V  P  R  G  S  T  A  I  G  M  K  E  T   250
gctgctgctaaattcgaacgccagcacatggacagcccagatctgggtaccggtggtggctccggtgatgacgac
 A  A  A  K  F  E  R  Q  H  M  D  S  P  D  L  G  T  G  G  G  S  G  D  D  D   275
gacaagagtcccatggctgctctgcgtgatgccgaaatccaaaaagatgtccaaacctactacggtcaagtgctg
 D  K  S  P  M  A  A  L  R  D  A  E  I  Q  K  D  V  Q  T  Y  Y  G  Q  V  L   300
aaacgctccgctgatctgcaaacgaacggttgcgtgaccacggcacgtccggttccgaaacatatccgcgaagcg
 K  R  S  A  D  L  Q  T  N  G  C  V  T  T  A  R  P  V  P  K  H  I  R  E  A   325
ctgcagaatgtgcacgaagaagttgccctgcgttattacggctgcggtctggtgatcccggaacatctggaaaac
 L  Q  N  V  H  E  E  V  A  L  R  Y  Y  G  C  G  L  V  I  P  E  H  L  E  N   350
tgttggattctggatctgggcagcggttctggccgcgactgttatgttctgagccagctggtcggtgaaaaaggc
 C  W  I  L  D  L  G  S  G  S  G  R  D  C  Y  V  L  S  Q  L  V  G  E  K  G   375
cacgtgaccggtattgatatgacgaaaggccaagtcgaagtggcagaaaagtatctggactaccatatggaaaaa
 H  V  T  G  I  D  M  T  K  G  Q  V  E  V  A  E  K  Y  L  D  Y  H  M  E  K   400
tatggttttcaggcttctaacgttaccttcattcacggctacatcgaaaaactgggtgaagcgggcattaagaat
 Y  G  F  Q  A  S  N  V  T  F  I  H  G  Y  I  E  K  L  G  E  A  G  I  K  N   425
gaaagccatgatatcgtggtttctaactgcgtcattaatctggtgccggacaagcagcaagtcctgcaagaagcc
 E  S  H  D  I  V  V  S  N  C  V  I  N  L  V  P  D  K  Q  Q  V  L  Q  E  A   450
tatcgtgtgctgaaacatggcggtgaactgtatttttcagatgtctacacctcgctggaactgccggaagaaatt
 Y  R  V  L  K  H  G  G  E  L  Y  F  S  D  V  Y  T  S  L  E  L  P  E  E  I   475
cgcacgcacaaagttctgtggggtgaatgcctgggcggtgcactgtactggaaggaactggcagtgctggctcag
 R  T  H  K  V  L  W  G  E  C  L  G  G  A  L  Y  W  K  E  L  A  V  L  A  Q   500
aaaatcggcttttgtccgccgcgtctggttaccgcaaacctgattacgatccaaaataaggaactggaacgtgtt
 K  I  G  F  C  P  P  R  L  V  T  A  N  L  I  T  I  Q  N  K  E  L  E  R  V   525
attggcgattgccgcttcgtcagtgctacctttcgtctgttcaaacattccaagaccggtccgacgaaacgctgt
 I  G  D  C  R  F  V  S  A  T  F  R  L  F  K  H  S  K  T  G  P  T  K  R  C   550
caagtgatttataacggcggtatcaccggccacgaaaaggaactgatgttcgatgcgaattttacgttcaaagaa
 Q  V  I  Y  N  G  G  I  T  G  H  E  K  E  L  M  F  D  A  N  F  T  F  K  E   575
ggtgaaatcgtggaagtggatgaagaaaccgcggccattctgaaaaatagccgttttgcgcaggatttcctgatt
 G  E  I  V  E  V  D  E  E  T  A  A  I  L  K  N  S  R  F  A  Q  D  F  L  I   600
cgcccgatcggcgaaaagctgccgacgagtggcggttgttccgccctggaactgaaagatattatcaccgacccg
 R  P  I  G  E  K  L  P  T  S  G  G  C  S  A  L  E  L  K  D  I  I  T  D  P   625
ttcaaactggcggaagaatcagactcaatgaaatcacgctgtgtgccggatgctgcgggtgtcgacaagcttgcg
 F  K  L  A  E  E  S  D  S  M  K  S  R  C  V  P  D  A  A  G  V  D  K  L  A   650
gccgcactcgagcaccaccaccaccaccaccaccactaa
 A  A  L  E  H  H  H  H  H  H  H  H  *
```

Fig. S3. Multiple alignment of hAS3MT with orthologues. *Homo sapiens* As(III) *S*-adenosylmethionine methyltranferase (AAI19639) was aligned with eukaryotic orthologues from *Rattus norvegicus* (rAS3MT, NP_543166), *Danio rerio* (zAS3MT, NP_001034928), *Chlamydomonas reinhardtii* (CrArsM, XP_001703619) and *Cyanidioschyzon merolae* (CmArsM, FJ476310). The shading illustrates conserved residues, with conserved cysteine residues highlighted in gray. The asterisks denotes position of tryptophan residues. Boxed residues denote residues not included in the cDNA clone or synthetic hAS3MT.



```
                                                                 32
hAS3MT      --------------MAALRDAEIQKDVQTYYGQVLKRSADLQTNGCVTTARPVPKHIREA  46
rAS3MT      --------------MAAPRDAEIHKDVQNYYGNVLKTSADLQTNACVTPAKGVPEYIRKS  46
zAS3MT      ----------MADAARDRTVTSVYNDVKEYYGKTLKQKSDLKSNACVPSAKPVSAYVRKV  50
CrArsM      MVEPASIAELSRAEQLGKDQDAVRATVKEYYGETLKTSNDLRTSACT-ACKAPPPAVRAA  59
CmArsM      --MPCSCASGCQKSKNGGSTPSIRDHVADYYGKTLQSSADLKTSACK-LAAAVPESHRKI  57

                          61         *
hAS3MT      LQNVHEEVALRYYGCGLVIPEH--LENCWILDLGSGSGRDCYVLSQLVGEKGHVTGIDMT 104
rAS3MT      LQNVHEEVISRYYGCGLVVPEH--LENCRILDLGSGSGRDCYVLSQLVGQKGHITGIDMT 104
zAS3MT      IAEIHPDVVAKYYGCGLVVPEC--LEGCRVLDLGCGSGRDCYMLSQLVGEKGHVTGIDMT 108
CrArsM      LADVPTEVKEKFYGCGNPIPAG--IEGLRVLDLGCGSGRDCYVAAKLVGEKGSVTGVDMT 117
CmArsM      LADIADEVLEKFYGCGSTLPADGSLEGATVLDLGCGTGRDVYLASKLVGEHGKVIGVDML 117

                                                                  156
hAS3MT      KGQVEVAEKYLDYHMEKYGF--QASNVTFIHGYIEKLGEAGIK---NESHDIVVSNCVIN 159
rAS3MT      KVQVEVAKAYLEYHTEKFGF--QTPNVTFLHGQIEMLAEAGIQ---KESYDIVISNCVIN 159
zAS3MT      EAQLEVARNYIDYHMQRFGY--KNPNVNFVQGYIEALVEAGLE---DKSYDIIISNCVVN 163
CrArsM      PAQLEVAISHADAYCRDKLG-YGKSNMTFIQGEIEYLDRAGLE---DSSFDLVISNCVIN 173
CmArsM      DNQLEVARKYVEYHAEKFFGSPSRSNVRFLKGFIENLATAEPEGVPDSSVDIVISNCVCN 177

                                                      * 206    *
hAS3MT      LVPDKQQVLQEAYRVLKHGGELYFSDVYTSLELPEEIRTHKVLWGECLGGALYWKELAVL 219
rAS3MT      LVPDKQKVLREVYQVLKYGGELYFSDVYASLEVSEDIKSHKVLWGECLGGALYWKDLAVI 219
zAS3MT      LSPDKSSVLREAYCVLKDGGELYFSDVYSDARIPEHLKANKTLWGECLSGALWWEDLIRL 223
CrArsM      LSPDKARVLSECYRVLAPGGEMHFSDVYVDRRLPQSVRSHPVLLGECLAGALYNNDFIRL 233
CmArsM      LSTNKLALFKEIHRVLRDGGELYFSDVYADRRLSEAAQQDPILYGECLGGALYLEDFRRL 237

hAS3MT      AQKIGFCPPRLVTANLITIQNKELERVIGDCRFVSATFRLFKHS----KTGPTKRCQVIY 275
rAS3MT      AKKIGFCPPRLVTANIITVGNKELERVLGDCRFVSATFRLFKLP----KTEPAGRCQVVY 275
zAS3MT      AEEVGFCKPRLVSASIITVGNTELESILGDYKFVSATYRLFKLQ----KGLEKKPCLVMY 279
CrArsM      SRKVGFTDPRQLECEEIQIHDAELRDQVGEARFYSITYRLFKVPGQIEDLCEDYGQVAVY 293
CmArsM      VAEAGFRDVRLVSVGPVDVSDPQLRKLVPDVQFYSCTFRCFKVA-TLEATREDYGQSATY 296

hAS3MT      NGGITGHEKELMFDANFTFKEGEIVEVDEETAAILKNSRFAQDFLIRPIGEKLPTSGGC- 334
rAS3MT      NGGIMGHEKELIFDANFTFKEGEAVEVDEETAAILRNSRFAHDFLFTPVEASLLAP---- 331
zAS3MT      NGDITDSEESFEFDAQYAFKVDKVMEVDGDVANILRNSRFSEEFTFQPQGVNTASSGG-- 337
CrArsM      KGTIPGHSHAYDLDDHHRFVTNKPMLVCGNTASMVGESWLAPHFTIIGDR---------- 343
CmArsM      LGGIG---EEFKLDRFFTFPREKPVRVDRNTAEIIRHSRLHQWFSVSAEQQHMGLFKAND 353

hAS3MT      -----SALELKDIITDPFKLAEESDSMKSRCVPDAAGGCCGTKKSC-- 375
rAS3MT      --------QTKVIIRDPFKLAEESDKMKPRCAPEGTGGCCGKRKSC-- 369
zAS3MT      -----CWAKPNAVSVNPFELVQQLG---SASVSPSTGGCCAGQESCCN 377
CrArsM      ---------AVHYGQFDCSGPKTTTG--GAASPSNSAGACGPGGACC- 379
CmArsM      SYALLHAPLSMQVEQLVCEVKKGSTGTCSAQASANGASCCATGRRCC- 400
```