Vol. 153, No. 1, 1988
May 31, 1988

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 359-364

AMINO ACID SEQUENCE OF S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE FROM DICTYOSTELIUM DISCOIDEUM AS DEDUCED FROM THE cDNA SEQUENCE

Judith Kasir, Robert R. Aksamit, Peter S. Backlund, Jr., and Giulio L. Cantoni

Laboratory of General and Comparative Biochemistry
National Institute of Mental Health, Bethesda, MD 20892

Received April 22, 1988

S-Adenosyl-L-homocysteine hydrolase has been cloned from a λgt11 cDNA library prepared from Dictyostelium discoideum that had been starved for 3 hours. The sequence of the cloned cDNA was determined and the deduced amino acid sequence was compared to the amino acid sequence of rat AdoHcy hydrolase. When the sequences from the two species were aligned, 74% of the amino acids were in identical positions. If conservative changes were taken into account the homology was 84%. Because differences have been reported in the binding characteristics of NAD^+ to the D. discoideum and rat AdoHcy hydrolases, changes in the amino acids of the putative NAD^+-binding site were of particular interest. Six changes were observed in this region but the changes appeared to be in regions that are not critical to the three dimensional folding of the NAD^+-binding site.

S-Adenosyl-L-homocysteine hydrolase catalyzes the reversible hydrolysis of AdoHcy, one of the products of methyl transfer reactions from S-adenosyl-L-methionine (1). The enzyme has been found in all cells with the exception of E. coli and certain other bacteria (2) where AdoHcy is hydrolyzed to adenine and ribosyl-homocysteine by a specific nucleosidase (3). AdoHcy hydrolases isolated from different sources exhibit significant structural similarities as well as some differences. The enzymes always consist of a number of identical subunits: the mammalian enzyme is a tetramer with a M_r of 190,000 and a subunit M_r of 47,000 (4-6), the plant enzymes may be either dimers or tetramers with a subunit of M_r of 55,000 (7,8), whereas the bacterial enzyme from Alcaligenes faecalis is composed of six subunits of M_r of 48,000 (2). Each subunit contains one mole of tightly bound NAD^+ which participates in the catalytic cycle, as first demonstrated by Palmer and Abeles (9). The amino acid sequence of the rat liver enzyme has been deduced from the cDNA nucleotide sequence (10), and contains a region of 31 amino acids (213-244) that has the characteristics of the nucleotide-binding domain discussed by Wierenga and Hol (11). This is the region of the AdoHcy hydrolase that presumably binds the ADP-moiety of NAD^+.

The slime mold, D. discoideum, is a primitive eukaryote that diverged from the mainstream of eukaryotic evolution at the earliest branch point yet

0006-291X/88 $1.50

characterized by molecular phylogeny (12). AdoHcy hydrolase from D. discoideum represents about 2% of the total soluble protein and has been purified to homogeneity by Hohman et al. (13). While its molecular and catalytic properties are similar to those of the rat liver enzyme it appears that the NAD^+ is bound to the D. discoideum enzyme considerably less tightly than to the mammalian enzyme (14). Therefore, it was of interest to examine whether the difference in NAD^+ binding affinity between the rat liver and the D. discoideum enzyme would be reflected in differences in the amino acid sequence of the nucleotide-binding domain.

We have isolated a cDNA clone coding for AdoHcy hydrolase from a D. discoideum cDNA library constructed in λgt11. Introduction of the cDNA into the E. coli expression vector pKK223-3 resulted in the production of AdoHcy hydrolase activity. We report here the amino acid sequence of D. discoideum AdoHcy hydrolase as deduced from nucleotide sequence of cloned cDNA.

MATERIALS AND METHODS

Cloning. The λgt11 cDNA library, prepared from D. discoideum mRNA that was obtained 3 hrs. after starvation, was kindly provided by P.N. Devreotes. The library was screened by in situ hybridization as described by Benton and Davis (15). The filters (Colony/Plaque Screen, New England Nuclear) were hybridized with a rat liver AdoHcy hydrolase cDNA (nucleotides 22 to 1970 as numbered by Ogawa et al. (10)) that had been labeled previously with ^{32}P by nick translation (16). Phage DNA from a positive plaque was prepared and the cDNA insert was subcloned into pUC13 to facilitate characterization.

DNA Sequence Determination. Restriction fragments of the hydrolase clone were subcloned into M13 phage vectors mp18 and mp19 and single stranded DNA was produced in the E. coli JM101 host. DNA sequence analysis was performed by the dideoxynucleotide technique of Sanger (17) with the Sequenase kit provided by U.S. Biochemicals.

RESULTS AND DISCUSSION

A D. discoideum cDNA library in λgt11 was screened with radiolabeled rat liver AdoHcy hydrolase cDNA. Of 250,000 plaques screened, one positive clone was obtained. Restriction enzyme analysis of the phage DNA showed that it contained a 1.3 kb EcoRI insert which was sequenced employing the strategy indicated in Fig. 1.

The nucleotide sequence and deduced amino acid sequence of the cDNA insert is shown in Fig. 2. The 1318 nucleotide sequence of the cDNA insert contains a single open reading frame of 1296 nucleotides beginning with a likely ATG initiation codon for methionine at nucleotides 1 to 3 and ending with the TAA stop codon at position 1294 to 1296. The 3' noncoding region contains the highly conserved poly(A) addition signal, AATAAA (18), that is located twelve nucleotides after the stop codon. Thus, the cDNA appears to code for a polypeptide of 430 amino acid residues.

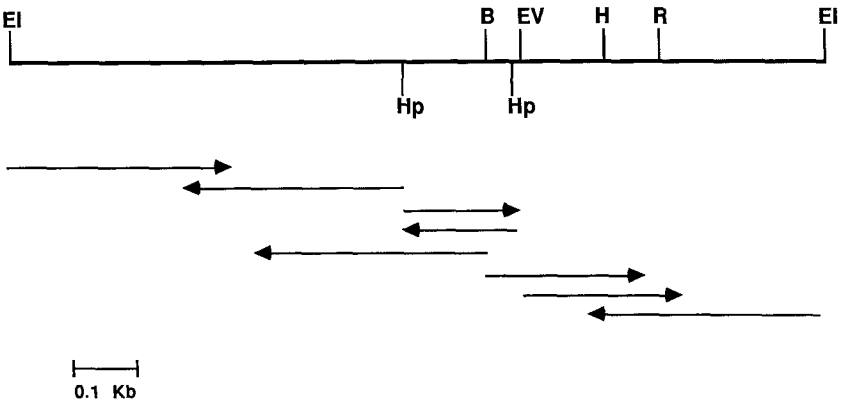

Figure 1. Restriction map and sequencing strategy for D. discoideum hydrolase cDNA. The arrows indicate the direction and extent of DNA sequenced. Restriction endonuclease sites indicated are: EI, EcoRI; Hp, HpaII; B, BstEII: EV, EcoRV; H, HincII; R, RsaI.

The amino acid sequence of the D. discoideum AdoHcy hydrolase was aligned with the sequence from the rat liver enzyme by the method of Needleman and Wunsch (19). As shown in Figure 2, the aligned sequences are very homologous with only one gap, which was inserted into the D. discoideum sequence. Seventy four percent of the amino acids are identical; and, if the conservative amino acid changes of Schwartz and Dayhoff (20) are considered, the homology between the two sequences is 84%. The three glycine and two cysteine residues which previously have been proposed to have an important role in either the binding of NAD^+ to the enzyme or in the catalytic activity (10) are conserved and indicated by asterisks in Figure 2. One stretch of 58 consecutive amino acids are identical in the two sequences. This region begins at amino acid 172 and extends into the putative NAD^+-binding domain indicated by the boxed amino acids in Figure 2. Within the last part of the NAD^+-binding domain six differences between aligned amino acids occur, and they occur at positions that are not believed to be critical in the three dimensional folding of this region (11). Other amino acid substitutions appear to occur randomly throughout the sequence with very few stretches where more than two consecutive amino acids are different between the rat and D. discoideum sequences. The longest stretch of consecutive amino acid substitutions is five, which occurs at amino acids 39 through 44. Thus, examination of the amino acid substitutions that have occurred in this region has not yet provided any insight that adequately explains differences in the binding of NAD^+ by the two different species of AdoHcy hydrolases. Additional work using site-directed mutagenesis may help illuminate the molecular basis for the functional difference between the two enzymes.

The N-terminal of rat AdoHcy hydrolase is known to undergo at least two post-translational modification reactions in which the initial methionine is

```
 -1      G ATG ACT AAA TTA CAC TAC AAA GTT AAA GAT ATT TCA CTT GCC GCT TGG GGT CGT AAG   57
Dict.      Met Thr Lys Leu His Tyr Lys Val Lys Asp Ile Ser Leu Ala Ala Trp Gly Arg Lys
Rat     Met Ala Asp ___ ___ Pro ___ ___ ___ Ala ___ ___ Gly ___ ___ ___ ___ ___ ___   19

 58 GAA ATT GAA ATT GCC GCC AAT GAA ATG CCA GGT TTA ATG ACC TTA AGA AAG AAA TAT GGT  117
 20 Glu Ile Glu Ile Ala Ala Asn Glu Met Pro Gly Leu Met Thr Leu Arg Lys Lys Tyr Gly   39
    Ala Leu Asp ___ ___ Glu ___ ___ ___ ___ ___ ___ Arg Met ___ Glu Met ___ Ser

118 CCA GCT CAA ATC TTA AAA GGT GCT CGT ATT GCA GGT TGT TTA CAC ATG ACT ATC CAA ACC  177
 40 Pro Ala Gln Ile Leu Lys Gly Ala Arg Ile Ala Gly Cys Leu His Met Thr Ile Gln Thr   59
    Ala Ser Lys Pro ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ Val Glu ___ ___

178 GCC GTT TTA ATC GAA ACT TTA ACT GCT CTC GGT GCT CAA GTC CAA TGG TCA TCA TGT AAC  237
 60 Ala Val Leu Ile Glu Thr Leu Thr Ala Leu Gly Ala Gln Val Gln Trp Ser Ser Cys Asn   79
    ___ ___ ___ ___ ___ ___ ___ Val ___ ___ ___ ___ Glu ___ Arg ___ ___ ___ ___ ___

238 ATT TTC TCC ACT CAA GAT CAA GCC GCC GCT GCC ATC GCT GCC ACT GGT GTC CCA GTC TAT  297
 80 Ile Phe Ser Thr Gln Asp Gln Ala Ala Ala Ala Ile Ala Ala Thr Gly Val Pro Val Tyr   99
    ___ ___ ___ ___ ___ ___ His ___ ___ ___ ___ ___ Lys Ala ___ ___ Ile ___ ___ Phe

298 GCC TGG AAA GGT GAA ACC GAA GAA GAA TAC AAC TGG TGT GTC GAA CAA ACC ATT GTT TTC  357
100 Ala Trp Lys Gly Glu Thr Glu Glu Glu Tyr Asn Trp Cys Val Glu Gln Thr Ile Val Phe  119
    ___ ___ ___ ___ ___ ___ Asp ___ ___ ___ Leu ___ Ile ___ ___ ___ Leu His ___

358 CAA GAT GGT CCA TTA AAT ATG ATC TTA GAT GAT GGT GGT GAT TTA ACC ACC CTC GTC CAC  417
120 Gln Asp Gly Pro Leu Asn Met Ile Leu Asp Asp Gly Gly Asp Leu Thr Thr Leu Val His  139
    Lys ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ Asp ___ ___ Ile ___

418 GAG AAA TAC CCA CAA TTC TTA GCT GGT ATC AAA GGT ATC TCT GAA GAA ACC ACC CAT GGT  477
140 Glu Lys Tyr Pro Gln Phe Leu Ala Gly Ile Lys Gly Ile Ser Glu Glu Thr Thr His Gly  159
    Thr ___ His ___ ___ Leu ___ Ser ___ ___ Arg ___ ___ ___ ___ ___ ___ Thr ___

478 GTC CAC AAC CTC TAC AAA ATG TTC AAA GAA GGT AAA TTA AAG GTC CCA GCC ATC AAC GTC  537
160 Val His Asn Leu Tyr Lys Met Phe Lys Glu Gly Lys Leu Lys Val Pro Ala Ile Asn Val  179
    ___ ___ ___ ___ ___ ___ ___ ___ Ala Asn ___ Ile ___ ___ ___ ___ ___ ___ ___ ___

538 AAT GAC TCT GTC ACC AAA TCC AAA TTC GAT AAC TTA TAT GGT TGT CGT GAA TCT TTA ATC  597
180 Asn Asp Ser Val Thr Lys Ser Lys Phe Asp Asn Leu Tyr Gly Cys Arg Glu Ser Leu Ile  199
    ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

598 GAT GGT ATT AAA CGT GCC ACC GAT GTT ATG ATT GCC GGT AAA GTT GCC GTC GTC GCT GGT  657
200 Asp Gly Ile Lys Arg Ala Thr Asp Val Met Ile Ala Gly Lys Val Ala Val Val Ala Gly  219
    ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

658 TAC GGT GAT GTA GGT AAA GGT TGT GCT CAA TCA TTA TCA AAA ATG GGT GCT CGT GTT TTA  717
220 Tyr Gly Asp Val Gly Lys Gly Cys Ala Gln Ser Leu Ser Lys Met Gly Ala Arg Val Leu  239
    ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ Ala ___ Arg Gly Phe ___ ___ ___ ___ Ile

718 GTC ACT GAA ATC GAT CCA ATC AAT GCC CTC CAA GCC TGT ATG GAT GGT TAC CAA ATC GTC  777
240 Val Thr Glu Ile Asp Pro Ile Asn Ala Leu Gln Ala Cys Met Asp Gly Tyr Gln Ile Val  259
    Ile ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ Ala ___ Glu ___ ___ Glu Val Thr

778 ACC ATG GAA ACC GCC GCT CCA TTA TCA AAC ATT TTC GTC ACC ACC ACC GGT TGT CGT GAT  837
260 Thr Met Glu Thr Ala Ala Pro Leu Ser Asn Ile Phe Val Thr Thr Thr Gly Cys Arg Asp  279
    ___ ___ Asp Glu ___ Cys Lys Glu Gly ___ ___ ___ ___ ___ ___ ___ ___ ___ Val ___

838 ATC GTC AGA GGT GAA CAC TTT GCC GTC ATG AAA GAA GAT GCC ATC GTT TGT AAC ATT GGT  897
280 Ile Val Arg Gly Glu His Phe Ala Val Met Lys Glu Asp Ala Ile Val Cys Asn Ile Gly  299
    ___ Ile Leu ___ Arg ___ ___ Glu Gln ___ ___ ___ Asp ___ ___ ___ ___ ___ ___ ___

898 CAC TTT GAT TGT GAA ATC GAT GTC GCT TGG TTA AAC GCC AAC GCC    AAA AAA GAT ACC  957
300 His Phe Asp Cys Glu Ile Asp Val Ala Trp Leu Asn Ala Asn Ala Val Lys Lys Asp Thr  318
    ___ ___ ___ Val ___ ___ ___ Lys ___ ___ ___ Glu ___ ___ Glu ___ Val Asn

958 GTC AAA CCA CAA GTT GAC CGT TAC ACC CTT GCC AAC GGT GTC CAC ATC ATC CTC TTA GCT 1017
319 Val Lys Pro Gln Val Asp Arg Tyr Thr Leu Ala Asn Gly Val His Ile Ile Leu Leu Ala  338
    Ile ___ ___ ___ ___ ___ ___ Leu ___ Lys ___ ___ His Arg ___ ___ ___ ___ ___ ___

1018 GAA GGT CGT CTC GTC AAT TTA GGT TGT GGT ACT GGT CAT CCA TCT TTT GTT ATG TCA AAC 1077
339 Glu Gly Arg Leu Val Asn Leu Gly Cys Gly Thr Gly His Pro Ser Phe Val Met Ser Asn  358
    ___ ___ ___ ___ ___ ___ ___ ___ Ala Met ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

1078 TCT TTC TGT AAC CAA ACT TTA GCT CAA ATC GCC CTC TGG ACT AAA ACT GAA GAA TAC CCA 1137
359 Ser Phe Cys Asn Gln Thr Leu Ala Gln Ile Ala Leu Trp Thr Lys Thr Glu Glu Tyr Pro  378
    ___ ___ Thr ___ ___ Val Met ___ ___ ___ Glu ___ ___ ___ His Pro Asp Lys ___ ___

1138 TTA GGT GTC CAC TTA TTA CCA AAG ATT TTA GAT GAA GAA GTT GCT CGT TTA CAT TTA GAT 1197
379 Leu Gly Val His Leu Leu Pro Lys Ile Leu Asp Glu Glu Val Ala Arg Leu His Leu Asp  398
    Val ___ ___ ___ Phe ___ ___ ___ Lys ___ ___ ___ Ala ___ ___ Glu Ala ___ ___ Gly

1198 CAA TTA GGT GCT AAA TTA ACT ACC CTC ACT GAA AAA CAA TCC GAA TAT TTA TCA GTT CCA 1257
399 Gln Leu Gly Ala Lys Leu Thr Thr Leu Thr Glu Lys Gln Ser Glu Tyr Leu Ser Val Pro  418
    Lys ___ Asn Val ___ ___ ___ Lys ___ ___ ___ ___ Ala Gln ___ ___ Gly Met ___ ___

1258 GTC GCT GGT CCA TAC AAA GTT GAT CAC TAC AGA TAT TAA AACTTTTGT TTAATAAATC G      1317
419 Val Ala Gly Pro Tyr Lys Val Asp His Tyr Arg Tyr End                            430
    Ile Asn ___ ___ Phe ___ Pro ___ ___ ___ ___ ___
```

TABLE 1

Codon Usage for D. discoideum and Rat Liver AdoHcy Hydrolases

		Dict.	Rat			Dict.	Rat			Dict.	Rat
ALA	GCG	-	3	LEU	CTG	-	23	ARG	AGG	-	-
	GCA	1	7		CTA		1		AGA	3	-
	GCT	18	14		CTT	2	3		CGG	-	7
	GCC	22	17		CTC	8	8		CGA	9	2
					TTC	-	4		CGT	-	1
ASP	GAT	20	14		TTA	30	-		CGC	-	5
	GCA	2	11								
				LYS	AAG	4	25	SER	AGT	-	-
CYS	TGT	11	5		AAA	25	5		AGC	-	4
	TGC	-	4						TCG	-	-
				ASN	AAT	5	4		TCA	8	2
GLU	GAG	1	24		AAC	12	17		TCT	5	2
	GAA	25	3						TCC	3	5
				PRO	CCG	-	-				
PHE	TTT	3	4		CCA	14	5	THR	ACG	-	4
	TTC	7	8		CCT	-	5		ACA	-	2
					CCC	-	6		ACT	13	7
GLY	GGG	-	4						ACC	18	10
	GCA	-	5	GLN	CAG	-	10				
	GGT	35	9		CAC	17	1	VAL	GTG	-	20
	GGC	-	19						GTA	1	2
				ILE	ATA	-	3		GTT	12	2
HIS	CAT	3	1		ATC	18	14		GTC	22	8
	CAC	9	13		ATT	11	15				
								TYR	TAT	5	2
MET	ATG	12	17	TRP	TGG	6	6		TAC	10	10

Values are the number of times the codon occurs in either the D. discoideum (Dict.) or rat liver sequences.

removed and the N-terminal alanine is blocked (10,21). Whether or not the methionine is removed and the N-terminal is blocked in D. discoideum AdoHcy hydrolase in not yet known. Because the cDNA sequence extends only one nucleotide past the first ATG codon, some uncertainty exits as to whether or not this is the initiation codon. However, it should be noted that we have cloned the D. discoideum cDNA into expression vector pKK223-3, and AdoHcy hydrolase activity was detected in recombinant E. coli.

Figure 2. Nucleotide sequence and deduced amino acid sequence of D. discoideum AdoHcy hydrolase. The A in the first ATG codon of the DNA sequence is nucleotide number 1. Numbering of the translated amino acid sequence begins with methionine, the first amino acid. The rat amino acid sequence is shown below the D. discoideum sequence. Only those amino acids that are different from the D. discoideum sequence are shown; identical amino acids are indicated by a line. The putative NAD^+-binding site from amino acid 213 to 244 and the poly(A) addition signal are boxed. Astericks indicate amino acids that are believed to have an important role in NAD^+-binding or catalysis.

The codon usage of rat and D. discoideum mRNA is very different as shown in Table I. As expected, there is a preferential use by D. discoideum of codons with A or T in the third position. This bias is characteristic of D. discoideum genes, which generally favor weak codon-anticodon interactions (22). Examination of the leucine codons used by the D. discoideum and rat sequences illustrates the conservation of the AdoHcy hydrolase primary structure. In the alignment of the amino acid sequences shown in Figure 2, leucine occurs in identical positions 33 times. Of these 33 conserved leucines, 22 of the codons were changed in two bases (TTA in D. discoideum to CTG, CTT or CTC in rat). In light of the very great homology between AdoHcy hydrolases from species that are separated by about one billion years in evolutionary history, it is tempting to postulate that the conservation of the amino acid structure serves an important function.

REFERENCES

1. de la Haba, G., and Cantoni, G.L. (1959) J. Biol. Chem. 234, 603-608.
2. Shimizu, S., Shiozaki, S., Ohshiro, T., and Yamada, H. (1984) Eur. J. Biochem. 141, 385-392.
3. Duerre, J. (1962) J. Biol. Chem. 237, 3737-3741.
4. Richards, H.H., Chiang, P.K., and Cantoni, G.L. (1978) J. Biol. Chem. 253, 4476-4480.
5. Døskeland, S.O., and Ueland, P.M. (1982) Biochim. Biophys. Acta 708, 185-193.
6. Fujioka, M., and Takata, Y. (1981) J. Biol. Chem. 256, 1631-1635.
7. Guranowski, A., and Pawelkiewicz, J. (1978) Planta (Berlin) 139, 245-247.
8. Sebestova, L., Votruba, I., and Holy, A. (1983) Collect. Czech. Chem. Commun. 49, 1543-1551.
9. Palmer, J.L., and Abeles, R.H. (1979) J. Biol. Chem. 254, 1217-1226.
10. Ogawa, H., Gomi, T., Mueckler, M.M., Fujioka, M., Backlund, P.S. Jr., Aksamit, R.R., Unson, C.G., and Cantoni, G. L. (1987) Proc. Natl. Acad. Sci. USA 84, 719-723.
11. Wierenga, R.K., and Hol, W.G.J. (1983) Nature (London) 302, 842-844.
12. McCarroll, R., Olsen, G.J., Stahl, Y.D., Woese, C.R., and Sogin, M.L. (1983) Biochemistry 22, 5858-5868.
13. Hohman, R.J., Guitton, M.C., and Veron, M. (1984) Arch. Biochem. Biophys. 233, 785-795.
14. de la Haba, Agostini, S., Bozzi, A., Merta, A., Unson, C., and Cantoni, G.L. (1986) Biochemistry 25, 8337-8342.
15. Benton, W.D., and Davis, R.W. (1977) Science 196, 180-182.
16. Rigby, P.W.J., Dieckmann, M., Rhodes, C., and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
17. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
18. Nevins, J.R. (1983) Annu. Rev. Biochem. 52, 441-466.
19. Needleman, S.D., and Wunsch, C.D. (1970) J. Mol. Biol. 48, 443-453.
20. Schwartz, R.M., and Dayhoff, M.O. (1979) In Atlas of Protein Sequence and Structure (M.O. Dayhoff, ed.), pp. 353-358, National Biomedical Research Foundation, Washington, D.C.
21. Gomi, T., Ishiguro, Y., and Fujioka, M. (1985) J. Biol. Chem. 260, 2789-2793.
22. Desrosiers, R., and Tanguay, R.M. (1988) J. Biol. Chem. 263, 4686-4692.