University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Vadim Gladyshev Publications | Biochemistry, Department of

5-25-2001

Chapter 13. The 15 kDa selenoprotein (Sepl5): functional studies and a role in cancer etiology

Vadim Gladyshev
University of Nebraska - Lincoln, vgladyshev1@unl.edu

Alan M. Diamond
University of Illinois at Chicago, Chicago, IL

Dolph L. Hatfield
National Institutes of Health, Bethesda, MD

Follow this and additional works at: http://digitalcommons.unl.edu/biochemgladyshev

Part of the Biochemistry, Biophysics, and Structural Biology Commons

Gladyshev, Vadim ; Diamond, Alan M.; and Hatfield, Dolph L., "Chapter 13. The 15 kDa selenoprotein (Sepl5): functional studies and a role in cancer etiology" (2001). *Vadim Gladyshev Publications.* Paper 32.
http://digitalcommons.unl.edu/biochemgladyshev/32

This Article is brought to you for free and open access by the Biochemistry, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Vadim Gladyshev Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in:

Selenium: Its Molecular Biology and Role in Human Health, edited by Dolph L. Hatfield, National Cancer Institute, U.S.A.

Published by Kluwer Academic Publishers, Boston/Dordrecht/London, 2001.

Portions of this work are copyright © 2001 Kluwer Academic Publishers. Portions authored or co-authored by U.S. government employees are not subject to copyright.

Chapter 13. The 15 kDa selenoprotein (Sep15): functional studies and a role in cancer etiology

Vadim N. Gladyshev

Department of Biochemistry, University of Nebraska, Lincoln, NE 68588, USA

Alan M. Diamond

Department of Human Nutrition, University of Illinois at Chicago, Chicago, IL 60612, USA

Dolph L. Hatfield

Section on the Molecular Biology of Selenium, Basic Research Laboratory, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

Summary: The 15 kDa selenoprotein (Sep15) is one of several recently identified selenoproteins. It contains a single selenocysteine residue in the middle of a 162-amino acid open reading frame and has no detectable homology to known proteins. The human Sep15 gene spans 51 kb, has 5 exons and is located on chromosome 1 at position p31. The gene contains two single nucleotide polymorphisms in the 3'-untranslated region (3'-UTR) including one in the SECIS element, that are distributed differently between Caucasians and African Americans. Sep15 localizes to the endoplasmic reticulum where it is tightly bound to UDP-glucose:glycoprotein glucosyltransferase, a protein involved in the quality control of protein folding. Sep15 may be involved in the chemopreventive effect of dietary selenium. This hypothesis is based on its differential expression in normal and malignant tissues, the distribution and functional consequences of natural polymorphisms within its gene, and the location of the Sep15 gene in a region that is often altered in a variety of cancers.

Introduction

The number of selenoproteins identified thus far in vertebrates stands at 22 (see Chapter 9). The function of only a few of these proteins is known. Since there is mounting evidence that at least some of the beneficial effects of selenium on health, including its anticarcinogenic properties, are mediated through selenoproteins, it is important to identify, characterize and determine functions for as many selenoproteins as possible.

One of these proteins for which the function is not known is the 15 kDa selenoprotein (Sep15). Sep15 was discovered about three years ago [1] when it was purified from human T cells as a protein that was strongly radiolabeled when cells were grown in the presence of ^{75}Se. The Sep15 cDNA sequence was deduced from the analysis of the EST database following sequencing of several tryptic peptides obtained from the isolated selenoprotein. The new open reading frame (ORF) contained an in-frame TGA codon (Figure 1) that was predicted to encode selenocysteine (Sec). This conclusion was also based on the presence of selenium in the protein and the fact that sequences downstream of the in-frame TGA matched those predicted to encode tryptic peptides in the isolated human T-cell Sep15.

Analysis of human Sep15 sequence using the EST database revealed that this selenoprotein was expressed in a variety of human cell types and identified its orthologs in other mammals. However, Sep15 sequences had no homology to known proteins making this protein a difficult target for functional characterization. Recent biochemical and genetic studies [2-4] significantly advanced our understanding of the genetics and biochemistry of Sep15. These studies identified a binding partner for Sep15 that implicated this selenoprotein in the quality control of protein folding [4] and revealed unexpected genetic factors that implicated Sep15 in the dietary effect of selenium in cancer prevention [2,3]. These studies are discussed in this chapter.

The human Sep15 gene

The complete sequence of the human Sep15 gene has been recently determined [2]. The gene spans a region of 51 kilobase pairs on chromosome 1 and consists of five exons and four introns (see Figure 1). The 5'-untranslated region (5'-UTR) and the first 27 amino acid residues of a putative signal peptide are included in the first exon. This peptide is not present in Sep15 isolated from human T cells [1] consistent with posttranslational processing of the signal peptide. Such gene organization, when the coding sequence for an amino terminal signal peptide is located in a separate exon, has been observed for many other mammalian proteins. The largest exon in the human Sep15 gene is exon 5 which encodes C-terminal sequences and also includes the 3'-UTR. The TGA codon that dictates Sec incorporation into the resulting Sep15 gene product is located in exon 3. The site of initiation of transcription in the Sep15 gene was determined by primer extension and was found to lack a TATA box that is often found upstream of RNA polymerase II transcribed mammalian genes [2].

Two polymorphic sites, which are located in the 3'-UTR at positions 811 and 1125 in human cDNA, have been identified [1-3] and their locations in the Sep15 mRNA sequence are shown in Figure 1. In all DNA samples examined to date, C^{811} is found associated with G^{1125} and T^{811} is associated

with A^{1125} suggesting the presence of two alleles. Interestingly, the polymorphism at 1125 site is located in the apical loop of selenocysteine insertion sequence (SECIS) element.

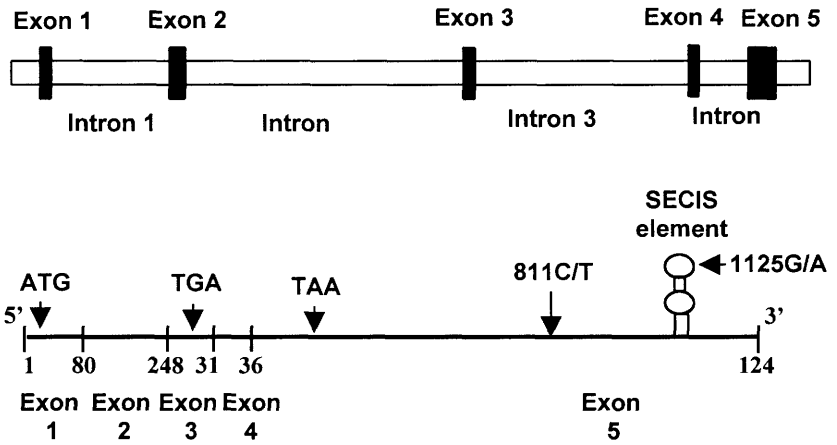

Figure 1. Structural organization of the human 15 kDa selenoprotein gene. In the upper panel, the exon-intron organization and relative sizes of exons and introns in the 51-kb human Sep15 gene are shown. Horizontal lines correspond to introns and flanking regions, and the closed squares to exons. In the lower panel, the organization of the human cDNA sequence obtained by splicing of the Sep15 gene is shown. The relative positions of the ATG initiation and the TGA Sec codons, the TAA termination signal and the detected polymorphisms (811C/T and 1125G/A) are shown. The location of 1125G/A polymorphism in the apical loop of the SECIS element is also indicated. The long horizontal line corresponds to the Sep15 cDNA, and short vertical lines correspond to exon-exon junctions. Numbers under junction sites correspond to last nucleotides in preceding exons.

Amino acid sequences

Coding sequences within Sep15 genes do not show detectable homology to known proteins, nor do they have clear sequence motifs indicative of a protein structure, cofactor composition or function (Figure 2). However, Sep15 genes were found in a wide variety of animals, including rats, mice, zebrafish, fruit flies and nematodes (Figure 2) [1,2]. The position of Sec is conserved in the Sep15 vertebrate sequences, whereas in insects and nematodes, Sec is replaced with Cys (Figure 2). Homology analyses also indicate that vertebrate Sep15 sequences are highly conserved. For example, mature human and mouse Sep15 differ by only three amino acid residues.

```
Rat     MAAGQGGWLRPALGL RLLLATAFQAVSALG AEFSSEACRELGFSS 45
Mouse   MAAGQGGWLRPALGL RLLLATAFQAVSALG AEFASEACRELGFSS 45
Human   MAAGPSGCLVPAFGL RLLLATVLQAVSAFG AEFSSEACRELGFSS 45
Bovine  MAARRDGWLGPAFGL RLLLATVLQTVSALG AEFSSESCRELGFSS 45
Ciona   MVGPKRRKNGAIISK LLQLFLLGMIHSHAE ASLSAQECADLGFSS 45
Shrimp  -------MMVDPGGI GLFIAATLLSIVEAV QELSTEECFAVGLNK 38
Droso   --------------- --------------- ---------------  0
C.eleg  ---------MWVIFL LLAAVVSPMFGEVEE YKIDVEECKAAGFNP 36

Rat     N-LLCSSCDLLGQFN LLPLDPVCRGCCQEE AQFETKKLYAGAILE 89
Mouse   N-LLCSSCDLLGQFN LLPLDPVCRGCCQEE AQFETKKLYAGAILE 89
Human   N-LLCSSCDLLGQFN LLQLDPDCRGCCQEE AQFETKKLYAGAILE 89
Bovine  N-LLCSSCDLLGQFN LLQLDPDCRGCCQEE AQFETKKLYAGAILE 89
Ciona   E-LMCGSCSLLPKFN LTMLEDDCKKCCQSE VEEDTAKRFHSAILE 89
Shrimp  ANLLCSSCDTLKEFN LDVLEANCRGCCNVD DVNATPTKYPRAILE 83
Droso   ---MCSSCEKLDDFG LDTIKPQCKQCCTLD QQPAAQRTYAKAILE 42
C.eleg  ETLKCGLCERLSDYH LETLLTDCLQCCIKE -EEFKHEKYPTAILE 80
          ●
Rat     VCGUKLGRFPQVQAF VRSDKPKLFR-GLQI KYVRGSDPVLKLLDD 133
Mouse   VCGUKLGRFPQVQAF VRSDKPKLFR-GLQI KYVRGSDPVLKLLDD 133
Human   VCGUKLGRFPQVQAF VRSDKPKLFR-GLQI KYVRGSDPVLKLLDD 133
Bovine  VCGUKLGRFPQVQAF VRSDKPKLFK-GLQI KYVRGSDPVLKLLDD 133
Ciona   VCGUKIGRYPQVQAF VKGEKSRAFS-NLKI KYVRGADPVIKLLNE 133
Shrimp  VCGURLGAFPQVQAF VKSDRPAAFP-NLTI KYVRGADPIIKLMDE 127
Droso   VCTCKFRAYPQIQAF IQSGRPAKFP-NLQI KYVRGLDPVVKLLDA  86
C.eleg  VCECNLARFPQVQAF VHKDMARQFGGKVKV KHVRGVRPQVALKDA 125

Rat     NGNIAEELSILKWNT DSVEEFLSEKLERI- ----------------162
Mouse   NGNIAEELSILKWNT DSVEEFLSEKLERI- ----------------162
Human   NGNIAEELSILKWNT DSVEEFLSEKLERI- ----------------162
Bovine  SGNIAEELSILKWNT DSVEEFLSEKLERI- ----------------162
Ciona   DEQVQDTLSITKWNT DSVEEFLNEKLIRV- ----------------162
Shrimp  DGDVMETLAIDKWNT DSXEEFLNTYLILPG QRGRSR----------163
Droso   SGKVQETLSITKWNT DTVEEFFETHLAKDG AGKNSYSVVEDADGD 132
C.eleg  DFKTKEVLSVEKWDT DTLIDFFNQWLE--- ----------------153

Rat     -----------  162
Mouse   -----------  162
Human   -----------  162
Bovine  -----------  162
Ciona   -----------  162
Shrimp  -----------  163
Droso   DDEDYLRTNRI  143
C.eleg  -----------  153
```

Figure 2. Amino acid alignment of animal Sep15 sequences. Residues identical in all sequences are highlighted. U represents selenocysteine and is shown by a closed circle above the sequence. Accession numbers for human and mouse Sep15 sequences are AF051894 and AF288740, respectively. Rat, bovine, *Ciona intestinals* (Ciona) and *Penaeus vannamei* (shrimp) sequences were assembled from EST sequences. The *Drosophila* (Droso) Sep15 sequence (accession number AE003523) was predicted by the *Drosophila* genome project. This sequence appears to lack the first exon containing the N-terminal signal peptide. The *C. elegans* (C.eleg) sequence (accession number AC024872) was predicted by the *C. elegans* Genome Project and corrected to remove an additional intron.

Pattern of expression

Following the initial EST analyses [1], expression of the Sep15 gene was examined in a number of human and mouse tissues by northern analysis and immunoblot assays [2]. Highest levels of gene expression were observed in prostate, liver, kidney, testes and brain, while lower levels were found in lung, spleen and skeletal muscle.

The high expression of Sep15 in the prostate may be relevant to the recent observation of an abundant ^{75}Se-labeled selenoprotein that migrated as ~15-16 kDa species on SDS PAGE gels [5]. In selenium deficiency, expression of this protein was conserved, in contrast to that of another selenoprotein, glutathione peroxidase 1. However, ~15-16 kDa protein has not been isolated or its gene sequenced, making it difficult to compare this protein and Sep15. Mammalian cells contain a number of selenoproteins with predicted molecular masses similar to that of mature Sep15 [6]. Thus, the rat prostate 15 kDa selenoprotein detected by Behne and colleagues [5] is either Sep15, one of the selenoproteins identified recently by searching mammalian genomes with various bioinformatics tools (Chapter 9), or a protein yet to be characterized.

Association of Sep15 and UGTR

Initial purification of Sep15 from human T cells employed reversed-phase chromatography as a final isolation step. Thus, Sep15 was isolated in a denatured state [1]. However, this protein migrated as ~160 kDa species on native gels. This observation suggested that Sep15 is either composed of multiple identical subunits or strongly binds another protein to form the 160 kDa complex. Taking advantage of the high expression level of Sep15 in the prostate, the protein was isolated in the native state from this organ by combination of conventional chromatography and HPLC [4].

Native Sep15 was found to occur in a complex with a 150 kDa protein. The latter protein was then identified by methods of protein microchemistry as UDP-glucose:glycoprotein glucosyltransferase (UGTR), an endoplasmic reticulum-resident protein [7]. The function of UGTR is to glucosylate misfolded proteins, thus retaining them in the endoplasmic reticulum until they are correctly folded or transferring them to degradation pathways [8]. The Sep15 gene was fused to a green florescent protein in order that the expressed fusion protein could be tracked intracellularly. Consistent with its association with UGTR, Sep15 was localized to the endoplasmic reticulum [4]. Furthermore, the N-terminal signal peptide, which was cleaved in the mature protein, was found to be essential for its translocation. The C-terminal sequence of Sep15 was not involved in retaining this protein in the endoplasmic reticulum, and most likely, Sep15 is retained in this compartment by its association with UGTR. Sep15 is the first known

endoplasmic reticulum-resident selenoprotein and its complex with UGTR suggests a role of selenium in the quality control of protein folding.

Influence of the A^{1125}/G^{1125} polymorphism on selenoprotein expression

The fact that the A^{1125}/G^{1125} polymorphic site occurs in the apical loop of the SECIS element suggested that it might differentially influence the level of Sep15 translation. This possibility prompted us to study the ability of polymorphisms in the Sep15 gene to support selenoprotein expression, as well as to study the allelic frequencies within the human population and their association with malignancy.

Initially, the ability of the nucleotide positions comprising the A^{1125}/G^{1125} polymorphism to influence selenoprotein expression was examined [2,3]. The Sep15 3'-UTR, encoding one or the other polymorphism, was attached to the deiodinase 1 (D1) gene in an expression vector. The levels of D1 in the chimeric constructs were compared to that of wild type D1 [2]. The level obtained with the A^{1125} construct was approximately 75% of that observed with wild type and the level with the G^{1125} construct was approximately 50% of that of wild type. These data suggested that the identity of the base at position 1125 ultimately may influence Sep15 levels.

In contrast to the 1125 site, the base at position 811 did not influence the level of Sep15 expression. That is, a T^{811}/G^{1125} construct yielded approximately the same level of expression as C^{811}/G^{1125} and a C^{811}/A^{1125} construct yielded approximately the same level of expression as T^{811}/A^{1125} [2,3].

To further examine the consequences of the polymorphic nucleotide positions with regard to their ability to influence the translation of mRNAs containing an in-frame UGA codon, a construct was employed that permitted us to test SECIS function by monitoring expression levels of full size and truncated proteins. Again, the G form resulted in lower expression of the reported construct. Thus, the data examining SECIS element efficiency as a function of the two polymorphic positions in Sep15 clearly demonstrated that the identity of the base at position 1125 influences the level of UGA readthrough. However, in addition to influencing selenoprotein expression at concentrations of selenium provided by 10% fetal bovine serum in cell culture media, these polymorphisms differentially responded to additional selenium supplementation of culture media [2,3]. Although the G^{1125} allelic form resulted in lower SECIS efficiency compared to the A^{1125} form in the absence of selenium supplementation, it was more responsive to increasing selenium levels than the A^{1125} form. The possibility that these properties may have a role in cancer prevention is discussed below.

Possible role of Sep15 in cancer etiology

The finding that the Sep15 gene occurs in humans as two allelic forms allowed us to determine the frequencies at which these alleles are present in the human population. Over 700 samples were analyzed by the PCR amplification of the polymorphic region followed by diagnostic restriction enzyme digestion [3]. This analysis revealed surprising differences in allele frequencies between Caucasians and African Americans. In addition, this study revealed differences in allele frequencies obtained from tumors of breast or head and neck origin when compared to cancer-free individuals among African Americans [3]. Furthermore, loss of heterozygosity (LOH) at the Sep15 gene locus was detected when DNA from circulating lymphocytes of an African American patient with a supraglotis tumor was compared to the DNA obtained from that tumor [3]. Development of cancer is often accompanied by the LOH at the sites that encode protective genes that are also called tumor suppressor genes. Thus, genetic changes at the Sep15 locus suggest that the loss of the Sep15 gene (or possibly a gene in the vicinity of the Sep15 gene) contributes to the development of certain cancers.

The human Sep15 gene is located on chromosome 1 at position p31. This region is commonly mutated or deleted in human cancers and the presence of a tumor suppressor gene on 1p31 has been suggested, but its identity has not been established [9,10]. It is possible that the loss of one copy of Sep15 gene may result in a decrease in the biosynthesis of this selenoprotein gene product in malignant tissues relative to the corresponding normal tissues.

The observation that natural polymorphisms found in the 3'-UTR of the Sep15 gene may influence its translational efficiency in response to selenium levels may also be relevant to the putative role of the protein in cancer. It is possible that individuals representing different combinations of these two haplotypes may express different amounts of Sep15 and, in addition, they may differentially respond to changes in dietary selenium (i.e., differentially control translational levels of Sep15 in response to changes in selenium levels).

The fact that prostate in both humans and mice expresses high levels of Sep15 may also be relevant to the link between Sep15 and cancer. Human selenium supplementation trials revealed that dietary selenium can reduce the incidence of prostate cancer [11] and it is likely that the greatest protection is provided to those individuals with lower selenium intake. In addition, epidemiological data have indicated a statistically significant inverse correlation between selenium in the diet and prostate cancer [12]. However, the mechanism of cancer prevention by selenium is poorly characterized [13] and no selenoprotein has been implicated in such protection. The recent data on Sep15 raise the possibility that this protein may function in the prevention of cancer and possibly serve as an agent by which selenium supplementation exerts its chemopreventive effect.

Although maximal protection against carcinogen-induced cancers in rodents has been achieved by providing dietary selenium in amounts exceeding levels that are necessary for maximal expression of GPx1 and GPx3 [13; and see Chapters 8, 14 and 23], these studies should not necessarily be viewed as proof that selenoproteins are not involved in cancer prevention. It is possible that distinct selenoproteins exhibit different expression patterns at high concentrations of selenium or when individuals are under environmental or genetic stress, such as those that are predisposed to the risk of developing certain cancers. Consistent with this notion was the observation that liver tumors of TGFα/*c-myc* double transgenic mice (a transgenic model of hepatocarcinoma) that were maintained on a selenium-sufficient diet had reduced levels of Sep15 compared to adjacent hepatic tissues [2]. In addition, Sep15 was essentially undetectable in a mouse prostate cancer cell line, while this protein was abundant in normal mouse prostate [2]. If lower levels of the Sep15 predispose an individual to malignancy, then the observation of differences in SECIS element function between the two naturally occurring alleles in the human population may indicate a segment of the population who are either at greater risk of cancer or whom might benefit from selenium supplementation. These observations, together with the fact that alterations in the region on chromosome 1 where the Sep15 gene is located is often associated with cancer progression suggest a possible role of Sep15 in cancer etiology.

Concluding remarks

Sep15 was discovered by virtue of the presence of selenium in the protein and the relative abundance of Sep15 compared to other selenoproteins. However, characterization of its gene sequences failed to detect significant homology to known proteins. Recent studies on the selenoprotein have suggested insights into its function and a role in human health. Its identification as a binding partner to the UGTR protein implicates Sep15 in the control of protein folding in the endoplasmic reticulum. In addition, several biochemical and genetic observations have suggested that Sep15 may be involved in cancer development and may mediate, at least in part, the chemoprotective effect of selenium. While these data hold great promise, further studies are necessary to determine the relevance of Sep15 to both protein folding and cancer prevention.

Acknowledgements: We thank members of our laboratories for contributions they made in characterization of Sep15 and its gene. This study was supported by a grant from the Cancer Research Foundation of America (to VNG).

References

1. VN Gladyshev, KT Jeang, JC Wootton, DL Hatfield 1998 *J Biol Chem* 273:8910

2. E Kumaraswamy, A Malykh, KV Korotkov, S Kozyavkin, Y Hu, SY Kwon, ME Moustafa, BA Carlson, MJ Berry, BJ Lee, DL Hatfield, AM Diamond, VN Gladyshev 2000 *J Biol Chem* 275:35540
3. Y Hu, KV Korotkov, R Mehta, DL Hatfield, C Rotimi, A Luke, TE Prewitt, RS Cooper, W Stock, EE Vokes, ME Dolan, VN Gladyshev, AM Diamond 2001 *Cancer Res* in press
4. KV Korotkov, E Kumaraswamy, Y Zhou, DL Hatfield, VN Gladyshev, submitted
5. M Kalcklosch, A Kyriakopoulos, C Hammel, D Behne 1995 *Biochem Biophys Res Commun* 217:162
6. GV Kryukov, VM Kryukov, VN Gladyshev 1999 *J Biol Chem* 274:33888
7. C Ritter, A Helenius 2000 *Nat Struct Biol* 7:278
8. AJ Parodi 2000 *Annu Rev Biochem* 69:69
9. S Avigad, H Benyaminy, Y Tamir, D Luria, I Yaniv, J Stein, B Stark, R Zaizov 1997 *Eur J Cancer* 33:1983
10. S Mathew, VV Murty, GJ Bosl, RS Chaganti 1994 *Cancer Res* 54:6265
11. LC Clark, GF Combs, BW Turnbull, EH Slate, DK Chalker, J Chow, LS Davis, RA Glover, GF Graham, EG Gross, A Krongrad, JL Lesher, HK Park BB Sanders, CL Smith, JR Taylor 1996 *J Am Med Assoc* 276:1957
12. K Yoshizawa, WC Willett, SJ Morris, MJ Stampfer, D Spiegelman, EB Rimm, E Giovannucci 1998 *J Natl Cancer Inst* 90:1219
13. HE Ganther 1999 *Carcinogenesis* 20:1657