Research and Reports in Biochemistry

Dovepress
open access to scientific and medical research

Open Access Full Text Article

REVIEW

Comparative studies of vertebrate scavenger receptor class B type 1: a high-density lipoprotein binding protein

This article was published in the following Dove Press journal:
Research and Reports in Biochemistry
16 June 2012
Number of times this article has been viewed

Roger S Holmes[1,2]
Laura A Cox[1]

[1]Department of Genetics and Southwest National Primate Research Center, Texas Biomedical Research Institute, San Antonio, TX, USA; [2]School of Biomolecular and Physical Sciences, Griffith University, Nathan, Queensland, Australia

Abstract: Scavenger receptor class B type 1 protein (SCARB1) plays an essential role in cholesterol homeostasis and functions in binding high density lipoprotein cholesterol (HDL) in liver and other tissues of the body. SCARB1 also functions in lymphocyte homeostasis and in the uptake of hepatitis C virus (HCV) by the liver. A genetic deficiency of this protein results in autoimmune disorders and significant changes in blood cholesterol phenotype. Comparative SCARB1 amino acid sequences and structures and *SCARB1* gene locations were examined using data from several vertebrate genome projects. Vertebrate SCARB1 sequences shared 50%–99% identity as compared with 28%–31% sequence identities with other CD36-like superfamily members, ie, SCARB2 and SCARB3 (also called CD36). At least eight N-glycosylation sites were conserved among most of the vertebrate SCARB1 proteins examined. Sequence alignments, key amino acid residues, and conserved predicted secondary structures were also studied, including: cytoplasmic, transmembrane, and exoplasmic sequences; conserved N-terminal and C-terminal transmembrane glycines which participate in oligomer formation; conserved cystine disulfides and a free SH residue which participates in lipid transport; carboxyl terminal PDZ-binding domain sequences (Ala507-Arg/Lys508-Leu509); and 30 conserved proline and 18 conserved glycine residues, which may contribute to short loop formation within the exoplasmic HDL-binding sequence. Vertebrate *SCARB1* genes usually contained 12 coding exons. The human *SCARB1* gene contained CpG islands, micro RNA binding sites, and several transcription factor binding sites (including PPARG) which may contribute to the high level (13.7 times the average) of gene expression. Phylogenetic analyses examined the relationships and potential evolutionary origins of the vertebrate *SCARB1* gene with vertebrate *SCARB2* and vertebrate and invertebrate *SCARB3* (*CD36*) genes. These suggested that *SCARB1* originated in a vertebrate ancestral genome from a gene duplication event of an ancestral invertebrate *CD36* gene.

Keywords: vertebrates, amino acid sequence, SCARB1, evolution, high-density lipoprotein receptor

Introduction

Scavenger receptor class B type 1 protein (SCARB1, also called CLA1, SRB1, and CD36 L1) is one of at least three members of the type 1 thrombospondin receptor CD36 (cluster of differentiation 36) family that serves as an homo-oligomeric plasma membrane cell surface glycoprotein receptor for high-density (HDL) lipoprotein cholesterol, other phospholipid ligands, and chylomicron remnants.[1–10] SCARB1 has also been implicated in the uptake of hepatitis C virus (HCV) in the liver[11–14] and in platelet activation.[15] and may also participate in the phagocytosis of apoptotic cells.[16–18] SCARB2, also called lysosomal integral membrane protein 2, SRB2, and

Correspondence: Roger S Holmes
School of Biomolecular and Physical Sciences, Griffith University, Nathan, Queensland 4111, Australia
Tel +61 7 3735 7773
Email r.holmes@griffith.edu.au

CD36 L2, is a second member of the CD36 family, predominantly integrated within lysosomal and endosomal membranes which may contribute to membrane organization and transport functions.[19–23] A third member of the CD36 family, SCARB3, also called CD36, fatty acyl translocase, and glycoprotein 88, is an integral membrane protein of many tissues of the body which plays a role in fatty acyl translocation and as a multiple ligand cell surface receptor of oxidized low-density lipoproteins, and has been implicated in several diseases, including insulin resistance, diabetes, atherosclerosis, and malaria.[24–29]

The gene encoding SCARB1 (*SCARB1* in humans, *Scarb1* in mice) is expressed at very high levels in various cells and tissues of the body, including the adrenal cortex and medulla, liver, and lymphocytes.[1,30,31] Studies of *Scarb1*$^{-}$*/Scarb1*$^{-}$ knockout mice have shown that SCARB1 deficiency causes autoimmune disease and an accumulation of cholesterol-rich HDL in the circulation and a reduction in biliary cholesterol excretion, which supports a major role for SCARB1 in the reverse cholesterol pathway.[18,32,33] Human clinical studies have also examined *SCARB1* polymorphisms associated with increased plasma HDL cholesterol levels, altered platelet functions, and a reduction in cholesterol efflux from macrophages.[34–38] In addition, human *SCARB1* polymorphisms are associated with a number of major diseases, including type 2 diabetes, atherosclerosis, autoimmune disorders, and malarial *Plasmodium* liver infection.[33,39–42]

Structures of vertebrate *SCARB1* genes have been examined, including human,[43,44] mouse,[45] rat,[46] cow,[47] and salmon.[48] *SCARB1* genes usually contain 12 exons of DNA encoding SCARB1 sequences which undergo exon shuffling, generating several isoproteins in each case.[31,44,49–51]

This paper reports the predicted gene structures and amino acid sequences for several vertebrate *SCARB1* genes and proteins, the secondary structures for vertebrate SCARB1 proteins, several potential sites for regulating human *SCARB1* gene expression and the structural, phylogenetic, and evolutionary relationships for these genes and enzymes with those for vertebrate *SCARB2* and *SCARB3* (*CD36*) gene families.

Materials and methods

Vertebrate *SCARB1* gene and protein identification

BLAST (Basic Local Alignment Search Tool) studies were undertaken using web tools from the National Center for Biotechnology information site (http://blast.ncbi.nlm.nih.gov/Blast.cgi).[52] Protein BLAST analyses used vertebrate SCARB1 amino acid sequences previously described (Table 1). Non-redundant protein sequence databases for several mammalian genomes were examined using the BLASTP algorithm, including human (*Homo sapiens*),[53] chimpanzee (*Pan troglodytes*),[54] rhesus monkey (*Macaca mulatta*),[55] cow (*Bos taurus*),[56] mouse (*Mus musculus*),[57] rat (*Rattus norvegicus*),[58] opossum (*Monodelphis domestica*),[59] platypus (*Ornithorhynchus anatinus*),[60] chicken (*Gallus gallus*),[61] lizard (*Anolis carolinensis*),[62] frog (*Xenopus tropicalis*),[63] and zebrafish (*Danio rerio*).[64] This procedure produced multiple BLAST "hits" for each of the protein databases which were individually examined and retained in FASTA format, and a record was kept of the sequences for predicted mRNAs and encoded SCARB1-like proteins. These records were derived from annotated genomic sequences using the gene prediction method, ie, GNOMON and predicted sequences with high similarity scores for human SCARB1. Predicted SCARB1-like protein sequences were obtained in each case and subjected to analyses of predicted protein and gene structures.

BLAT (Blast Like Alignment Tool) analyses were subsequently undertaken for each of the predicted SCARB1 amino acid sequences using the UC Santa Cruz Genome Browser[65] with default settings to obtain the predicted locations for each of the vertebrate *SCARB1* genes, including predicted exon boundary locations and gene sizes. BLAT analyses were similarly undertaken for other vertebrate *SCARB1, SCARB2*, and *SCARB3* (*CD36*) genes using previously reported sequences in each case (see Table 1). Structures for human and mouse isoforms (splicing variants) were obtained using the AceView website to examine predicted gene and protein structures.[31]

Predicted structures and properties of vertebrate SCARB1

Predicted secondary structures for human and other vertebrate SCARB1 proteins were obtained using the PSIPRED v2.5 web site tools provided by Brunel University.[66] Molecular weights, N-glycosylation sites,[67] and predicted transmembrane, cytosolic, and extracellular sequences for vertebrate SCARB1 proteins were obtained using Expasy web tools (http://au.expasy.org/tools/pi_tool.html).

Comparative human and mouse *SCARB1* gene expression

The genome browser (http://genome.ucsc.edu)[65] was used to examine GNF Expression Atlas 2 data using various expression chips for human and mouse *SCARB1* genes (http://biogps.gnf.org).[68] Gene array expression "heat maps" were

examined for comparative gene expression levels among human and mouse tissues showing high (red), intermediate (black), and low (green) expression levels.

Phylogeny studies and sequence divergence

Alignments of vertebrate SCARB1, SCARB2, and SCARB3 sequences were assembled using BioEdit v.5.0.1 and the default settings.[69] Ambiguous alignment regions, including the amino and carboxyl termini, were excluded prior to phylogenetic analysis, yielding alignments of 431 residues for comparisons of vertebrate SCARB1 sequences with human, mouse, chicken, and zebrafish SCARB2 and SCARB3 sequences with the lancelet (*Branchiostoma floridae*) CD36 sequence (Table 1). Evolutionary distances were calculated using the Kimura option[70] in TREECON.[71] Phylogenetic trees were constructed from evolutionary distances using the neighbor-joining method[72] and rooted with the seasquirt CD36 sequence. Tree topology was reexamined by the boot-strap method (100 bootstraps were applied) of resampling, and only values that were highly significant (≥95) are shown.[73]

Results and discussion

Alignments of vertebrate SCARB1 amino acid sequences

The deduced amino acid sequences for pig (*Sus scrofa*), opossum (*M. domestica*), platypus (*O. anatinus*), and frog (*X. tropicalis*) SCARB1 are shown in Figure 1, together with previously reported sequences for human,[74] mouse,[75] and cow SCARB1 (Table 1).[47] Alignments of human with other vertebrate SCARB1 sequences examined were between 50% and 99% identical, suggesting that these are products of the same family of genes, whereas comparisons of sequence identities of vertebrate SCARB1 proteins with human SCARB2 and SCARB3 proteins exhibited lower levels of sequence identities of 30%–33% and 30%–32%, respectively, indicating that these are members of distinct *SCARB* or *CD36*-like gene families (Table 1).

The amino acid sequences for mammalian SCARB1 contained 509 residues whereas frog (*X. tropicalis*) SCARB1 sequences contained 505 amino acids (Figure 1). Previous studies have reported several key regions and residues for human and mouse SCARB1 proteins (human SCARB1 amino acid residues were identified in each case). These included: cytoplasmic N-terminal and C-terminal residues 2–11 and 461–509; N-terminal and C-terminal transmembrane helical regions, ie, residues 12–33 and 441–460;[8,46,74] key N-terminal glycine residues (Gly15/Gly18/Gly25) which form a dimerization motif in the N-terminal transmembrane domain and participate in forming SCARB1 oligomers;[76] a C-terminal PDZK1-interacting domain previously identified for the last three amino acids of SCARB1, ie, Ala-Lys-Leu;[75] exoplasmic Cys384, where a free SH group plays a major role in SCARB1-mediated lipid transport; four exoplasmic disulfide bond-forming residues, ie, Cys280, Cys321, Cys323, and Cys334;[77,78] and 10 exoplasmic N-glycosylation sites for human SCARB1, which have been identified or predicted for this protein (Table 2).[79] One of these sites (site 1) contained a proline residue at the second position and may not function as an N-glycosylation site due to proline-induced inaccessibility.[67] Eight of these sites were predominantly retained among the 19 vertebrate SCARB1 sequences examined (sites 2–3 and 6–12 in Table 2), whereas rat and mouse SCARB1 sequences contained an additional N-glycosylation site at 116 Asn-117 Arg-118Ser, and chicken, lizard, and salmon SCARB1 sequences contained a further site at 125 Asn-126Gly-127Thr (Figure 1, Table 2). Given the sequence conservation observed for these residues among the vertebrate SCARB1 sequences examined, it is apparent that they are essential for the structure and function of a glycoprotein. The multiple N-glycosylation sites observed for vertebrate SCARB1 sequences support a major role for glycan residues exposed on the external surface of plasma membranes in the performance of SCARB1 functions in binding various lipoproteins, including HDL, and in cholesterol transfer, as part of the reverse cholesterol transport role proposed for SCARB1.[79]

Conserved glycines in N-terminal oligomerization domain of the transmembrane sequence

The N-terminal region for vertebrate SCARB1 sequences (residues 1–42 for human SCARB1) contained cytoplasmic (residues 2–11) and transmembrane (residues 12–33) motifs which underwent changes in amino acid sequence but retained the predicted cytoplasmic and transmembrane properties in each case (Figure 1, Supplementary Figure 1). Vertebrate N-terminal transmembrane sequences in particular were predominantly conserved, especially for Gly15, Gly18, and Gly25 SCARB1 residues, which were fully conserved among all mammalian and frog SCARB1 sequences examined (Figure 1, Supplementary Figure 1). However, other lower vertebrate SCARB1 sequences, such as lizard and zebrafish SCARB1 sequences, contained a replacement at the Gly15 position (Ala15 and Leu12, respectively, data not shown). Additional glycine residues were observed in this region for some vertebrate SCARB1 sequences, including consecutive

Table 1 Vertebrate SCARB1, SCARB2, and SCARB3 genes and proteins

SCARB1 gene	Species	RefSeq ID 'Ensembl/NCBI	GenBank ID	UNIPROT ID	Amino acids	Chromosome location
Human	*Homo sapiens*	NM_00505	BC022087	Q8WVT0	509	12:125,267,232-125,348,266
Chimpanzee	*Pan troglodytes*	'XP_003314063.1	na	na	509	12:125,656,282-125,738,482
Gorilla	*Gorilla gorilla*	'ENSGGOP6367	na	G3QUG5	509	12:124,718,744-124,804,938
Gibbon	*Nomascus leucogenys*	'XP_003276298.1	na	G1QQ51	509	GL397391:1,704,805-1,787,403
Rhesus	*Macaca mulatta*	'XP_001101812.1	na	F7GMT2	509	11:126,096,068-126,176,851
Mouse	*Mus musculus*	NM_001205082.1	BC004656	Q61009	509	5:125,761,478-125,821,252
Rat	*Rattus norvegicus*	NM_031541	BC076504	P97943	509	12:32,390,385-32,453,972
Guinea Pig	*Cavia porcellus*	'XP_003461673.1	na	na	509	178:1,245,087-1,276,811
Cow	*Bos taurus*	NM_174597	BT19968	O18824	509	17:53,181,037-53,267,911
Dog	*Canis familaris*	'ENSCAFP10357	na	na	509	26:8,098,793-8,128,586
Pig	*Sus scrofa*	NM_213967.1	AF467889	Q8SQC1	509	14:28,301,496-28,388,350
Rabbit	*Oryctolagus cuniculus*	NM_001082788.1	AY283277.1	na	509	Un0166:481-18,522
Opossum	*Monodelphis domestica*	'XP_001379299	na	F7E0V5	509	3:475,428,162-475,518,702
Platypus	*Ornithorhynchus anatinus*	'XP_001508198	na	F7F3F3	509	196:2,257,972-2,313,254
Chicken	*Gallus gallus*	'XP_415106	na	na	503	15:4,543,054-4,558,954
Lizard	*Anolis carolinensis*	'ENSACAT14873	na	na	502	LGb:1,343,011-1,368,333
Frog	*Xenopus tropicalis*	'XP_002935333	na	na	505	GL172777:533,463-579,626
Zebrafish	*Danio rerio*	NM_198121	BC044516	E7FB50	496	11:21,526,513-21,572,478
Salmon	*Salmo salar*	NM_001204894.1	HQ403588	E9N3T5	486	na
SCARB2 gene						
Human	*Homo sapiens*	NM_005506	BT006939	Q53Y63	478	4:77,084,378-77,134,696
Mouse	*Mus musculus*	NM_007644	BC029073	O35114	478	5:92,875,330-92,934,334
Chicken	*Gallus gallus*	'XP_42093.1	BX931548	na	481	4:51,411,268-51,429,620
Zebrafish	*Danio rerio*	NM_173259.1	BC162407	Q8JQR8	531	5: 63,942,096-63,955,449
SCARB3 gene						
Human	*Homo sapiens*	NM_001001547	BC008406	P16671	472	7:80,275,645-80,303,732
Mouse	*Mus musculus*	NM_001159555.1	BC010262	Q08857	472	5:17,291,543-17,334,712
Chicken	*Gallus gallus*	'ENSGALG8439	AJ719746	F1NER9	471	1:12,077,308-12,107,415
Zebrafish	*Danio rerio*	NP_001002363.1	BC076048	Q6DHC7	465	4:21,594,449-21,606,961
CD36 gene						
Lancelet	*Branchiostoma floridae*	'XP_002609178.1	na	na	480	Un:534,334,234-534,343,082
Sea squirt	*Ciona intestinalis*	'XP_002127015.1	na	na	523	09p:2,872,362-2,873,903
Nematode	*Caenorhabditis elegans*	NM_067224	na	Q9XTT3	534	III:12,453,609-12,456,726

Notes: RefSeq: reference amino acid sequence; 'predicted Ensembl amino acid sequence; na, not available; GenBank IDs are derived from National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/genbank/); Ensembl ID was derived from the Ensembl genome database (http://www.ensembl.org); UNIPROT refers to UniprotKB/Swiss-Prot IDs for individual SCARB-like proteins (http://kr.expasy.org); *partial gene size only; Un refers to unknown chromosome; bps refers to base pairs of nucleotide sequences.

Abbreviation: SCARB, scavenger receptor class B.

glycines for opossum SCARB1 (residues 25 and 26, Figure 1) and for chicken (residues 23 and 24) and zebrafish (residues 15 and 16) SCARB1 sequences (sequences not shown). Site-directed mutagenesis studies have demonstrated key roles for the Gly15-$(X)_2$-Gly18-$(X)_5$-Gly25 motif within the N-terminus transmembrane sequence, by facilitating oligomerization and selective lipid uptake by SCARB1, particularly for the Gly18 and Gly25 residues.[76] A conserved glycine residue was also observed for the mammalian and frog C-terminal transmembrane sequences (human SCARB1 Gly451, Figure 1, Supplementary Figure 1), but it is not known whether or not this glycine has a similar role to that of the conserved N-terminal transmembrane glycine residues.

Conservation of cytoplasmic C-terminal PDZK1-interacting domain

The cytoplasmic C-terminal region of SCARB1, particularly the last three amino acids of mouse SCARB1, ie, Ala507-Lys508-Leu509, has been shown to interact with a PDZ domain-containing protein (PDZK1) which has been suggested to be responsible for the cell surface expression of this protein.[75,80,81] PDZ domains have been previously shown

Coding exons (strand)	Gene size bps	Subunit MW	Gene expression Level	% Identity with human SCARB1	% identity with human SCARB2	% identity with human SCARB3
12 (–ve)	81,035	56,973	13.7	100	29	31
12 (–ve)	82,201	56,973	na	100	29	29
12 (–ve)	86,195	57,034	na	99	29	31
12 (–ve)	82,599	57,029	na	99	29	31
12 (–ve)	80,784	57,046	na	97	29	30
12 (–ve)	63,985	56,754	5.1	79	29	29
12 (+ve)	63,588	56,973	0.5	78	28	28
12 (+ve)	31,725	56,969	na	75	29	30
12 (+ve)	86,875	57,610	na	83	29	29
11 (+ve)	29,794	57,149	na	79	28	27
12 (+ve)	86,855	57,514	na	87	29	28
12 (–ve)	18,042*	57,080	na	80	28	28
12 (+ve)	90,541	57,682	na	70	28	28
12 (–ve)	55,283	57,839	na	73	28	28
12 (+ve)	15,901	55,918	na	57	28	31
12 (+ve)	25,323	56,147	na	58	30	28
12 (–ve)	46,164	57,240	na	54	28	30
12 (–ve)	45,684	55,742	na	51	28	30
na	na	55,097	na	50	28	30
12 (–ve)	50,316	54,290	3.2	29	100	30
12 (–ve)	59,005	54,044	3.6	29	85	31
12 (+ve)	18,353	53,907	na	30	59	33
13 (+ve)	13,354	60,234	na	31	43	33
12 (+ve)	72,231	53,053	5.7	31	30	100
12 (–ve)	43,170	52,698	4.2	30	31	83
12 (–ve)	30,108	52,624	na	30	32	61
12 (–ve)	12,513	51,590	na	31	31	53
12 (+ve)	8,849	54,141	na	34	35	35
1 (–ve)	1,542	58009.00	na	26	33	31
8 (+ve)	3,118	60,182	4.6	21	26	24

to be abundant protein interaction motifs which function in regulating biological processes such as transport (in this case, lipid transport), ion channel signaling, and other signal transduction systems.[82,83] Comparisons of mammalian SCARB1 cytoplasmic C-terminal sequences (Supplementary Figure 1) support a high level of conservation for the last three amino acids, with a consensus sequence of Ala507-Lys/Arg508-Leu509 being maintained. This conservation extends further into the protein as follows: Gly501-Thr, Ser or Ala502-Val503-Leu504-Gln505-Glu506, which may also contribute to PDZK1 binding.

Conserved cysteine residues within the exoplasmic domain

Five cysteine residues of the mammalian and frog SCARB1 exoplasmic sequences (residues 34–440 for human SCARB1) were conserved among the proteins examined, of which four participate in disulfide bridge formation (cys281, cys321, cys323, and cys334), and the fifth (cys384) plays an essential role (via the free SH group) in lipid transport.[77] A sixth cysteine residue (cys251) is also conserved among the mammalian SCARB1 sequences examined and contains a free thiol group, although site-directed mutagenesis studies have

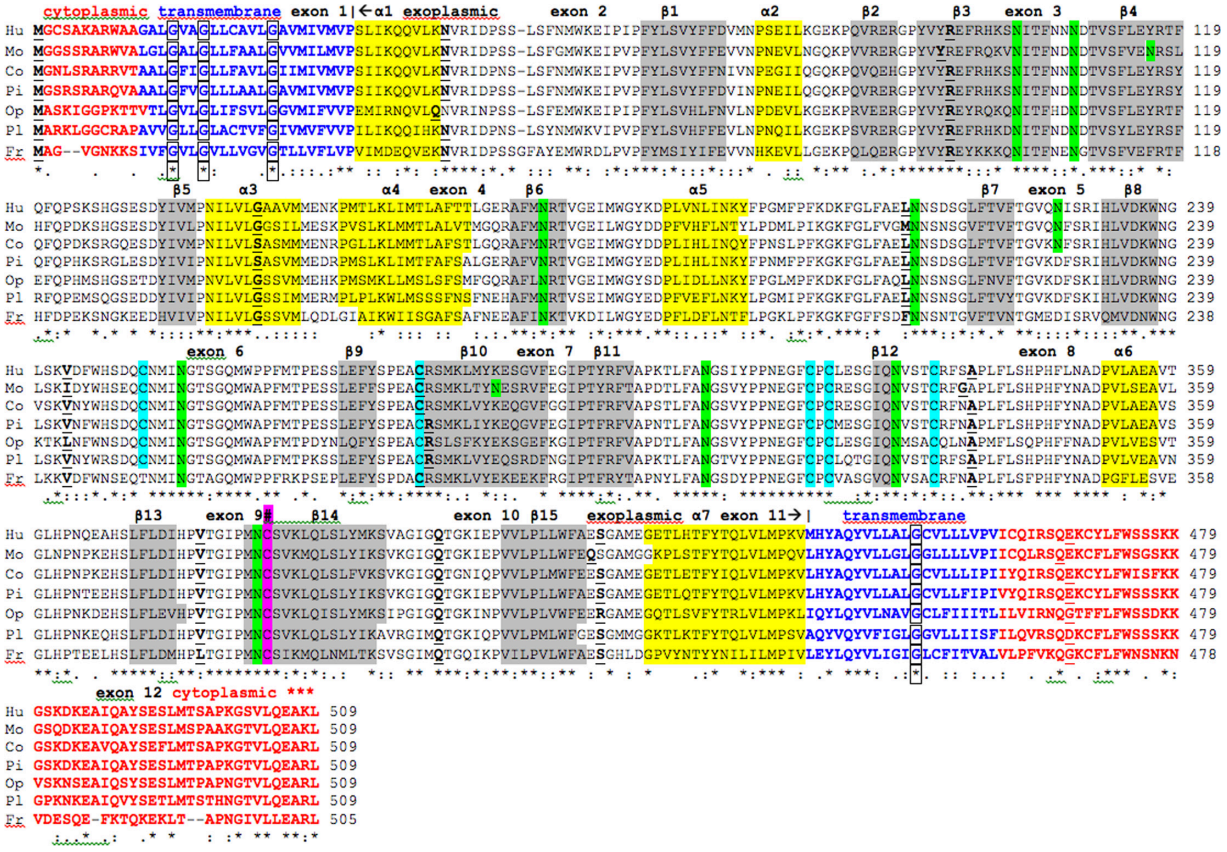

Figure 1 Amino acid sequence alignments for vertebrate SCARB1 sequences.
Notes: See Table 1 for sources of SCARB1 sequences. *Identical residues for SCARB1 subunits; similar alternate residues; dissimilar alternate residues; predicted cytoplasmic residues are shown in red; predicted transmembrane residues are shown in blue; N-glycosylated and potential N-glycosylated Asn sites are shown in green; free SH Cys involved in lipid transfer is shown in pink; predicted disulfide bond Cys residues are shown in blue; predicted α-helices for vertebrate SCARB1 are shown in yellow and numbered in sequence from the start of the predicted exoplasmic domain; predicted β-sheets are shown in gray and also numbered in sequence; bold underlined font shows residues corresponding to known or predicted exon start sites; exon numbers refer to human *SCARB1* gene exons. **Final three C-terminal residues which bind a PDZ domain-containing protein (PDZK1); G residues are conserved glycines in the N-terminal oligomerization domain of the transmembrane sequence.[76]
Abbreviations: Hu, human; Mo, mouse; Co, cow; Pi, pig; Op, opossum; Pl, platypus; Fr, frog; SCARB1, scavenger receptor class B type 1 protein.

demonstrated no significant involvement of this exoplasmic SH group in lipid transfer activity.[77] Moreover, the Cys251 residue has been substituted in the frog SCARB1 sequence by threonine, supporting the view that it is not an essential amino acid for vertebrate SCARB1 at this position (Figure 1).

Predicted secondary structures for vertebrate SCARB1

Predicted secondary structures for mammalian and frog SCARB1 sequences were examined (Figure 1), particularly for the exoplasmic sequences. α-helix and β-sheet structures were similar in each case, with an α-helix extending beyond the N-terminal and C-terminal transmembrane regions: α1 and α7. A consistent sequence of predicted secondary structure was observed for each of mammalian and frog SCARB1 sequences: N-terminal cytoplasmic sequence--N-terminal transmembrane sequence--α1--β1--α2--β2--β3--β4--β5--α3--α4--β6--α5--β7--β8--β9--β10--β11--β12--α6--β13--β14--β15--α7--C-terminal transmembrane sequence--C-terminal cytoplasmic sequence. Further description of the secondary and tertiary structures for SCARB1 must await the three-dimensional structure for this protein, particularly for the exoplasmic region which directly binds HDL and participates in reverse cholesterol transport.

Conserved proline and glycine residues within the SCARB1 exoplasmic domain

Figure 2 shows the alignment of 15 vertebrate SCARB1 amino acid sequences for the exoplasmic domain, with colors depicting the properties of individual amino acids and the strong conservation observed for these protein sequences. In addition to the key vertebrate SCARB1 amino acids detailed previously, others are also conserved, including 30 proline residues. Prolines play a major role in protein folding and protein-protein interactions involving the cyclic pyrrolidine

Table 2 Predicted N-glycosylation sites for vertebrate SCARB1 sequences

Vertebrate SCARB1	Species	Site 1*	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8	Site 9	Site 10	Site 11	Site 12	No of sites
Human	*Homo sapiens*	74NPSE	102NITF	108NDTV			173NRTV	212NNSD	227NISR	255NGTS	310NGSI	330NVST	383NCSV	9
Chimp	*Pan troglodytes*	74NPSE	102NITF	108NDTV			173NRTV	212NNSD	227NISR	255NGTS	310NGSI	330NVST	383NCSV	9
Gorilla	*Gorilla gorilla*	74NPSE	102NITF	108NDTV			173NRTV	212NNSD	227NISR	255NGTS	310NGSI	330NVST	383NCSV	9
Gibbon	*Nomascus leucogenys*		102NITF	108NDTV			173NRTV	212NNSD	227NISR	255NGTS	310NGSI	330NVST	383NCSV	9
Rhesus	*Macaca mulatta*		102NITF	108NDTV			173NRTV	212NNSD	227NISR	255NGTS	310NGSI	330NVST	383NCSV	9
Mouse	*Mus musculus*		102NITF	108NDTV	116NRSL		173NRTV	212NNSN	227NFSR	255NGTS	310NGSV	330NVST	383NCSV	10
Rat	*Rattus norvegicus*	74NPSE	102NITF	108NDTV	116NRSL		173NRTV	212NDSS	227NFSK	255NGTA	310NGSV	330NVST	383NCSV	10
Guinea Pig	*Cavia porcellus*		102NISF	108NDTV			173NRTV	212NNSD	227NLSR	255NGTS	310NGSV	330NVST	383NCSV	9
Cow	*Bos taurus*		102NITF	108NDTV			173NRTV	212NNSD	227NFSR	255NGTS	310NGSV	330NVST	383NCSV	9
Dog	*Canis familaris*		102NITF	108NDTV			173NRTV	212NNSD		255NGTS	310NGSV	330NVST	383NCSV	8
Pig	*Sus scrofa*		102NITF	108NDTV			173NRTV	212NNSD		255NGTS	310NGSV	330NVST	383NCSV	8
Rabbit	*Oryctolagus cuniculus*		102NITF	108NDTV			173NRTV	212NDSD		255NGTS	310NGSM	330NVST	383NCSV	8
Opossum	*Monodelphis domestica*	74NPSS	102NISF	108NDTV			173NRTV	212NNSN		255NGTS	310NGSV	330NMSA	383NCSV	8
Platypus	*Ornithorhynchus anatinus*		102NITF	108NDTV			173NRTV	212NNSN		255NGTS	310NGSV	330NVST	383NCSV	8
Chicken	*Gallus gallus*		100NITF	106NDTV		125NGTE	171NRTV	210NNTN	225NISQ	253NGTA	307NGTD	327NVSS	380NCSI	10
Lizard	*Anolis carolensis*		102NITF	108NDTV		127NGSE	173NRTV	212NNSN		255NGTS	310NGTD	330NVSS		8
Frog	*Xenopus tropicalis*		101NITF	107NGTV			172NKTV	211NNSN		254NGTA	309NGSD	329NVSA	382NCSI	8
Zebrafish	*Danio rerio*		100NITF	106NHTV				210NNSN		253NGTA		328NVSS	381NVSN	6
Salmon	*Salmo salar*	71NPTE	99NITF	105NDTV		126NESD		209NNSN		252NGTA	307NGSD	327NVSS	380NVSI	8

Notes: Numbers refer to amino acids in the acid sequences, including N-asparagine; K-lysine; I-isoleucine; H-histidine; S-serine; T-threonine; Q-glutamine; D-aspartate; Y-tyrosine; and V-valine. Note that there are 12 potential sites identified, including nine sites for human SCARB1.[77] N-glycosylation sites were identified using the NetNGlyc 1.0 web server (http://www.cbs.dtu.dk/services/NetNGlyc/). *Site 1 contains a proline residue at position 2 which may restrict N-glycosylation at this site.

Abbreviation: SCARB, scavenger receptor class B.

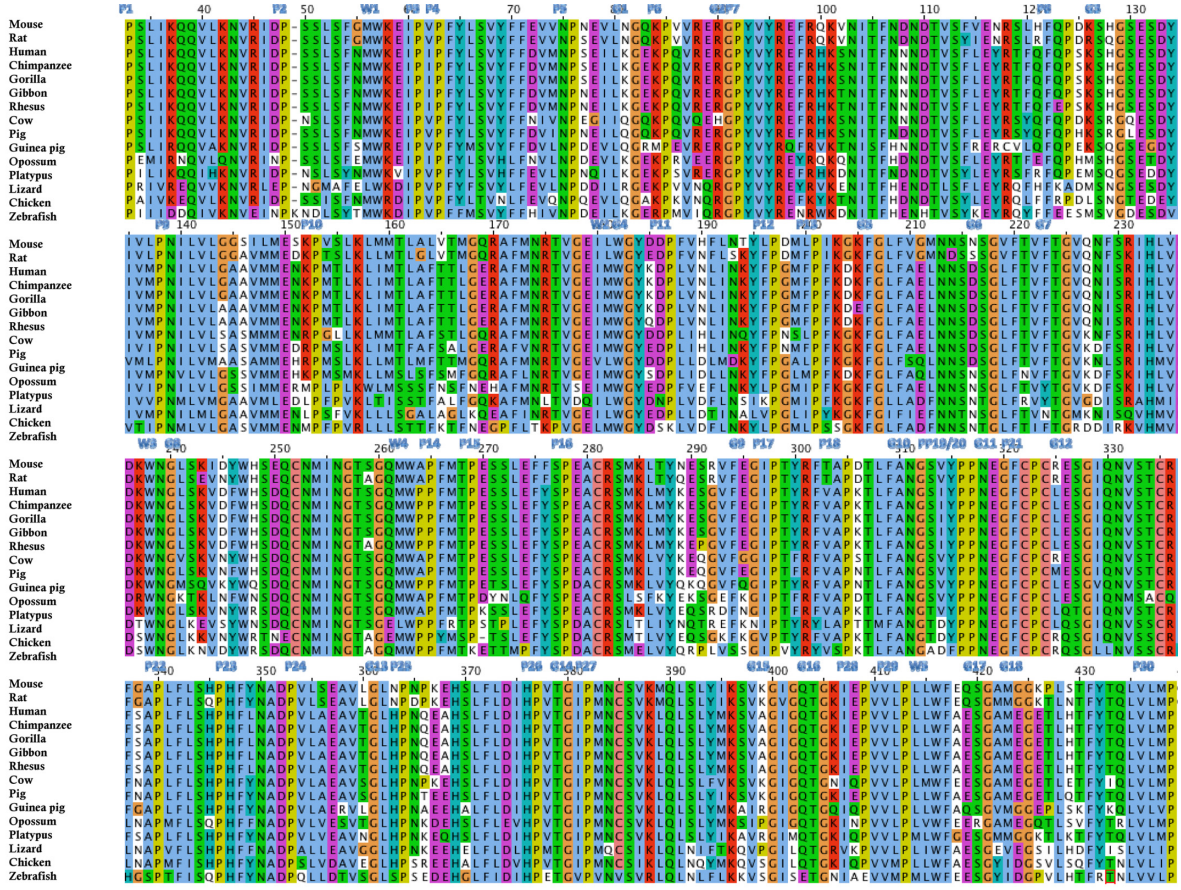

Figure 2 Amino acid sequence alignments for vertebrate SCARB1 exoplasmic sequences.

Notes: Amino acids are color-coded: yellow for proline (P); S (serine); green for hydrophilic amino acids, S (serine), Q (glutamine), N (asparagine), and T (threonine); brown for glycine (G); light blue for hydrophobic amino acids, L (leucine), I (isoleucine), V (valine), M (methionine), and W (tryptophan); dark blue for amino acids, T (tyrosine) and H (histidine); purple for acidic amino acids, E (glutamate) and D (aspartate); and red for basic amino acids, K (lysine) and R (arginine); conserved prolines (P) numbered P1–P30; conserved glycines numbered G1–G18; and conserved tryptophans numbered from W1–W5.

Abbreviation: SCARB1, scavenger receptor class B type I protein.

amino acid side chain, which may introduce turns (or kinks) in the polypeptide chain as well as having destabilizing effects on α-helix and β-strand conformations.[84] In addition, the presence of sequential prolines within a protein sequence may confer further restriction in folding conformation and create a distinctive structure, such as that reported for the mammalian Na^+/H^+ exchanger, which plays a major role in cation transport.[85] Sequential prolines (Pro317-Pro318) were conserved for all vertebrate SCARB1 sequences examined and may confer a distinctive conformation in this region, supporting the lipid transfer functions for this protein (Figure 2). Moreover, regions of water-exposed proteins with high levels of proline residues are often sites for protein-protein interactions[86] and these residues may play essential roles in the binding of HDL and other lipoproteins by the exoplasmic region of SCARB1.

Figure 2 also shows conservation of 18 glycine residues for these vertebrate SCARB1 exoplasmic domains which, due to their small size, may be essential for static turns, bends, or close packing in the domain, or required for conformational dynamics during HDL receptor on-off switching, as in the case of the aspartate receptor protein.[87] Both proline and glycine residues are frequently found in turn and loop structures of proteins, and usually influence short loop formation within proteins containing 2–10 amino acids.[88] Evidence for these short loop structures within SCARB1 exoplasmic sequences was evident from the predicted secondary structures for vertebrate SCARB1 (see Figure 1) with proline and/or glycine residues found at the start of the following structures: α1 (Pro33), β1 (Pro62), α2 (Pro75), β2 (Pro84), β3 (Gly90-Pro91), α3 (Pro136), α4 (Pro152), α5 (Pro186), β7 (Gly206), β10 (Pro277), β11 (296Gly, 298Pro), β12 (Gly327), α6 (Pro352), β14 (Gly379, Pro381), β15 (Pro408), and α7 (Gly424). Moreover, conserved SCARB1 sequential proline residues (Pro315-Pro316) are located in a region with no predicted secondary structure (between β11 and β12) but

with disulfide bonds present,[77] which suggests that this is a region of conformational significance for SCARB1.

Alignments of mouse SCARB1, SCARB2, and SCARB3 (CD36) amino acid sequences

The amino acid sequences for mouse SCARB1, SCARB2, and SCARB3 (see Table 1) are aligned in Figure 3. The sequences were 30%–33% identical and showed similarities in several key features and residues, including: cytoplasmic N-terminal and C-terminal residues; N-terminal and C-terminal transmembrane helical regions; exoplasmic disulfide bond forming residues, previously identified for mouse SCARB3 (CD36): Cys243, Cys272, Cys311, Cys313, Cys322, and Cys333;[77] several predicted N-glycosylation sites for mouse SCARB1 (11 sites), SCARB2 (11 sites), and SCARB3 (7 sites), of which two are shared between these sequences (N-glycosylation sites 7 and 9 (Table 2)); and similar predicted secondary structures previously identified for SCARB1 (Figure 1). The free SH group Cys384 residue, which plays a major role in SCARB1-mediated lipid transport,[77] was unique for SCARB1, being replaced by other residues for the corresponding SCARB2 and SCARB3 proteins. N-terminal transmembrane glycine residues, which play a role in the formation of SCARB1 oligomers,[76] were also observed for the corresponding mouse SCARB3 (CD36) sequence, although only one of these glycines (Gly10) was retained for the mouse SCARB2 sequence. These results suggest that mouse SCARB1, SCARB2, and SCARB3 proteins share several important properties, features, and conserved residues, including being membrane-bound with cytoplasmic and transmembrane regions, and have similar secondary structures, but are sufficiently different to serve distinct functions.

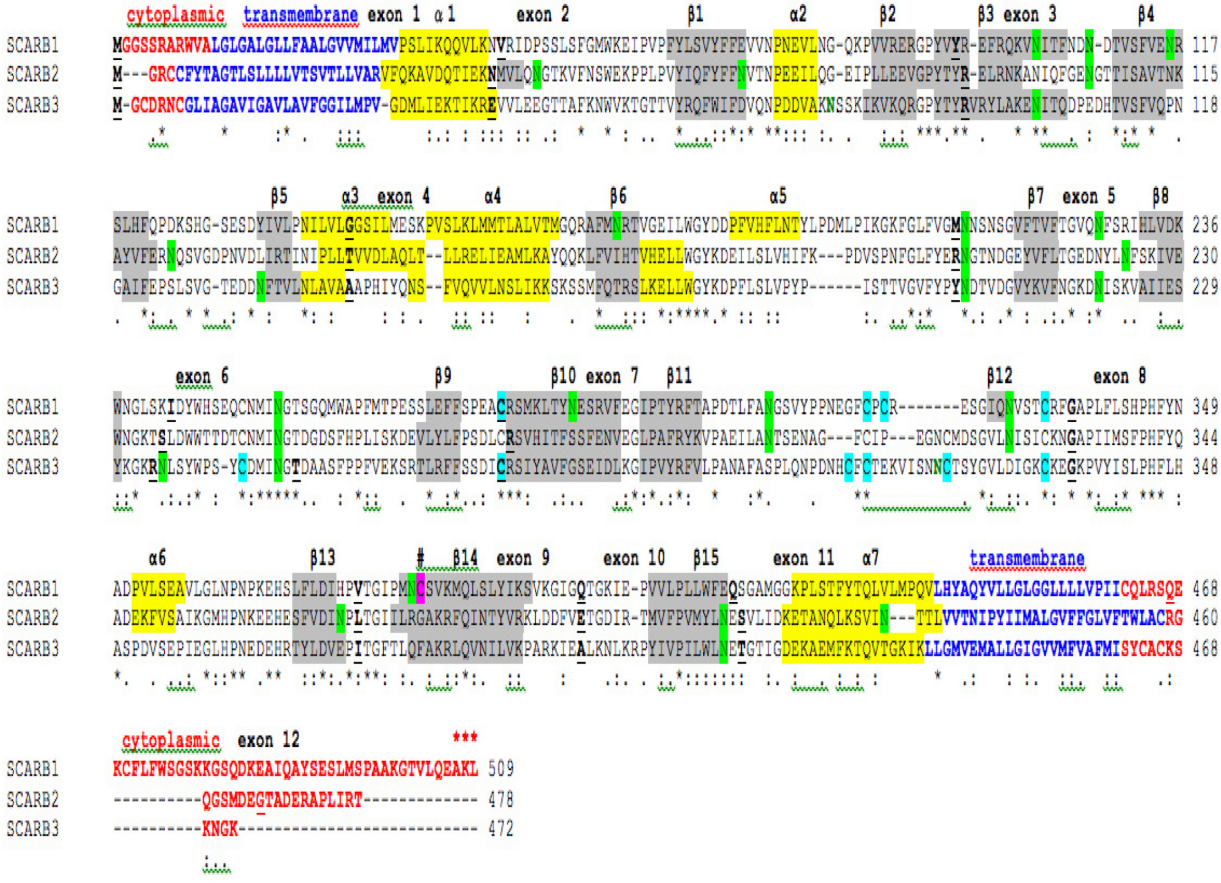

Figure 3 Amino acid sequence alignments for mouse SCARB1, SCARB2, and SCARB3 (CD36) sequences.

Notes: See Table 1 for sources of mouse SCARB1, SCARB2, and SCARB3 (CD36) sequences. *Identical residues for SCARB1 subunits; similar alternate residues; dissimilar alternate residues; predicted cytoplasmic residues are shown in red; predicted transmembrane residues are shown in blue; N-glycosylated and potential N-glycosylated Asn sites are in green; free SH Cys involved in lipid transfer is shown in pink; predicted disulfide bond Cys residues are shown in blue; predicted α-helices for vertebrate SCARB1 are in shaded yellow and numbered in sequence from the start of the predicted exoplasmic domain; predicted β-sheets are in shaded gray and also numbered in sequence; bold underlined font shows residues corresponding to known or predicted exon start sites; exon numbers are shown; ***Final three C-terminal residues which bind a PDZ domain-containing protein (PDZK1) for mouse SCARB1.

Abbreviation: SCARB, scavenger receptor class B.

Gene locations and exonic structures for vertebrate *SCARB1* genes

Table 1 summarizes the predicted locations for vertebrate *SCARB1* genes based upon BLAT interrogations of several vertebrate genomes using the reported sequences for human,[43,44] mouse,[45] rat,[46] cow,[47] and salmon[48] and the predicted sequences for other vertebrate *SCARB1* genes and the UC Santa Cruz genome browser.[65] The predicted vertebrate *SCARB1* genes were transcribed on either the negative strand (primate, mouse, rabbit, platypus, frog, and zebrafish genomes) or the positive strand (rat, guinea pig, cow, dog, pig, opossum, chicken, and lizard genomes). Figure 1 summarizes the predicted exonic start sites for human, mouse, cow, opossum, platypus, and frog *SCARB1* genes, with each having 12 coding exons in identical or similar positions to those predicted for the human *SCARB1* gene.[43,44]

Figure 4 shows the predicted structures of mRNAs for the three major human *SCARB1* transcripts and the major *Scarb1* transcript.[31] The human transcripts were 2.6–2.7 kilobases in length, with 12 (isoforms a and c) or 13 (isoform b) introns present for these *SCARB1* mRNA transcripts, and in each case, an extended 3′-untranslated region was observed. The human *SCARB1* genome sequence contained several predicted transcription factor binding sites, three microRNA binding sites (miR-125/351, miR-223, and miR-145) located in the 3′-untranslated region and three CpG islands (CpG78, CpG24, and CpG27), of which CpG78 is located in the 5′-untranslated promoter region of human *SCARB1* on chromosome 12. Mouse *Scarb1* also contains a CpG island within the gene promoter (CpG33) as well as two of the human predicted microRNA target sites (miR145 and miR125/351) within the mouse *Scarb1* 3′-untranslated region. The *SCARB1* CpG islands residing proximate to the promoter may contribute to the very high level of gene expression (13.7 times for human *SCARB1* and 5.1 times the average gene expression for mouse *Scarb1*)[31] and may be

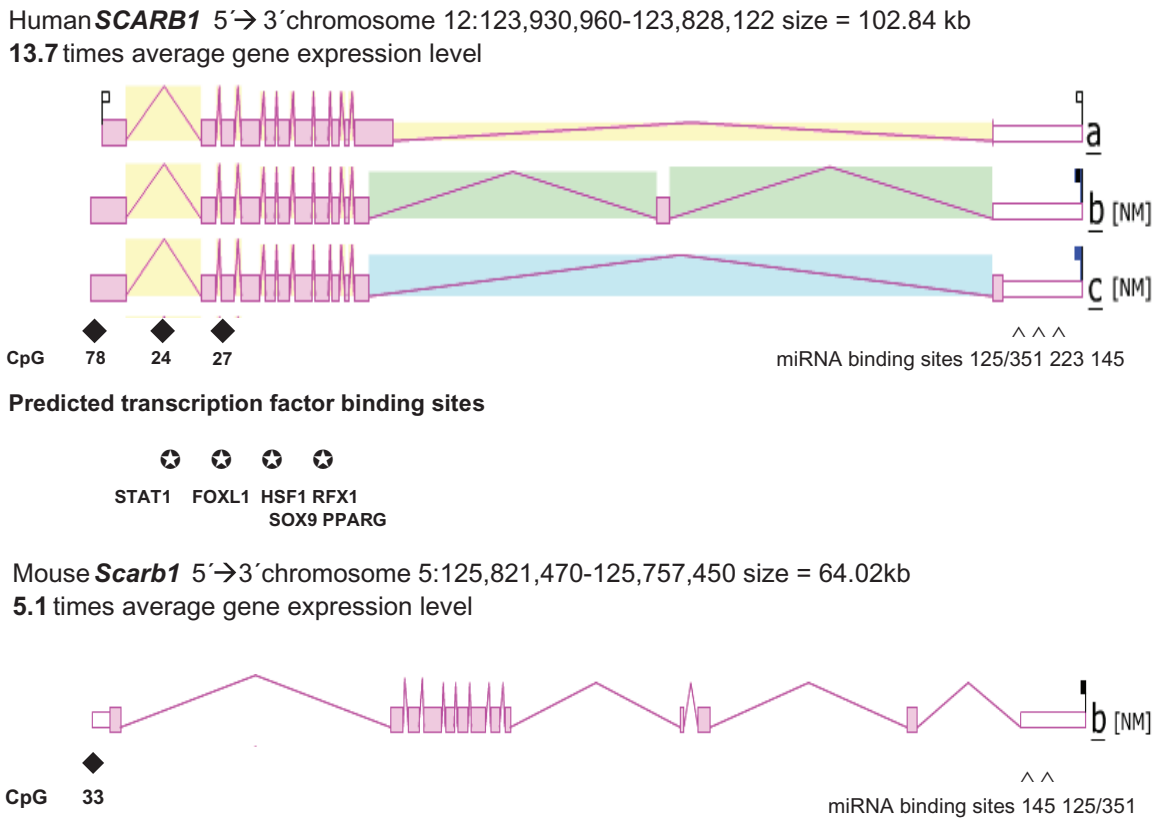

Figure 4 Gene structures and major splicing transcripts for the human and mouse *SCARB1* genes.

Notes: Derived from the AceView website (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/)[31] mature isoform variants are shown with capped 5′ and 3′ ends for the predicted mRNA sequences; NM refers to the National Center for Biotechnology Information reference sequence; exons are in pink; the directions for transcription are shown as 5′ → 3′; black squares show predicted CpG island sites at or near the 5′ untranslated regions of the genes; the symbol ∧ shows predicted microRNA binding sites observed at or near the *SCARB1* 3′ untranslated regions; sizes of mRNA sequences are shown in kilobases; predicted transcription factor binding sites for human *SCARB1* are shown as stars.

Abbreviations: SCARB1, scavenger receptor class B type 1 protein; STAT1, signal transducer and activator of transcription; FOXL1, forkhead related transcription factor; HSF1, heat shock transcription factor; RFX1, MHC class II regulatory factor; SOX9, transcription factor; PPARG, peroxisome proliferator-activated receptor γ; SCARB1, scavenger receptor class B type 1.

compared with the CpG islands within housekeeping gene promoters expressed in most tissues.[89] Of particular significance among the transcription factor binding sites observed is the peroxisome proliferator-activated receptor-γ site, which has been shown to stimulate liver *SCARB1* expression[90] and to participate in regulating cholesterol uptake and efflux from macrophages, and to promote uptake of oxidized low-density lipoproteins and subsequent differentiation of monocytes into foam cells.[91–93] The prediction of multiple miRNA target sites within the 3′-untranslated regions of human *SCARB1* and mouse *Scarb1* is also potentially of major significance because of the role that small noncoding RNAs play in regulating mRNA and protein expression during embryonic development.[94]

Comparative human and mouse *SCARB1* tissue expression

Figure 5 presents "heat maps" showing comparative gene expression for various human and mouse tissues obtained from GNF Expression Atlas Data using the U133 A and GNF1H (human) and GNF1M (mouse) chips (http://genome.ucsc.edu; http://biogps.gnf.org).[68] These data supported a broad and high level of tissue expression for human and mouse *SCARB1*, particularly for the adrenal gland, liver, lymphocytes, ovary and placenta, which is consistent with previous reports for these genes.[1,8,47,74] Overall however, human and mouse *SCARB1* tissue expression levels were 3–4 times the average level of gene expression, which supports the key role played by this enzyme in lipid metabolism, especially in the liver, adrenal glands, and lymphocytes, and during embryonic development (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/).[31]

Phylogeny and divergence of SCARB1 and other vertebrate CD36-like sequences

A phylogenetic tree (Figure 6) was calculated by the progressive alignment of 19 vertebrate SCARB1 amino acid

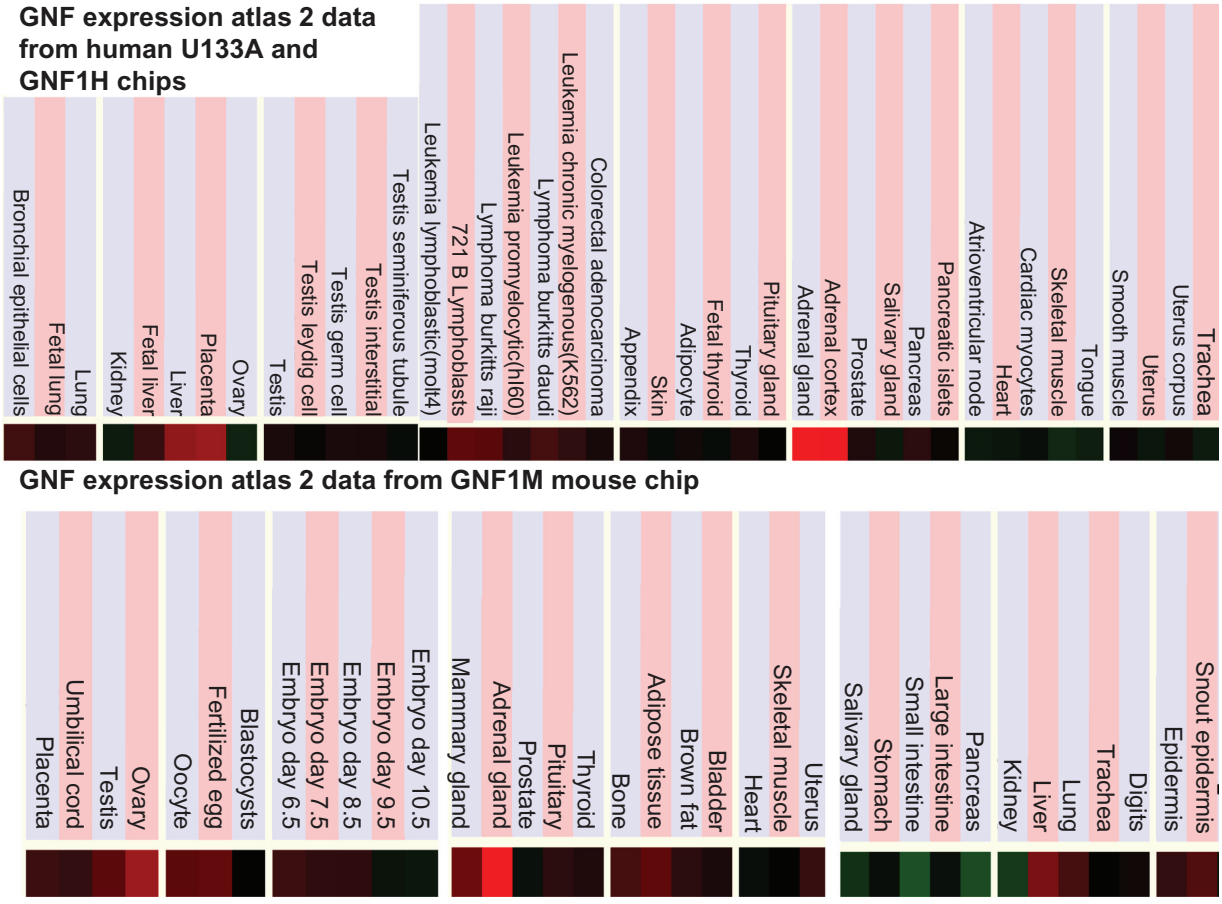

Figure 5 Comparative tissue expression for human and mouse *SCARB1* genes.
Notes: Expression "heat maps" (GNF Expression Atlas 2 data) (http://biogps.gnf.org)[68] were examined for comparative gene expression levels among human and mouse tissues for *SCARB1* genes showing high (red); intermediate (black); and low (green) expression levels. Derived from human and mouse genome browsers (http://genome.ucsc.edu).[65]
Abbreviation: SCARB1, scavenger receptor class B type 1.

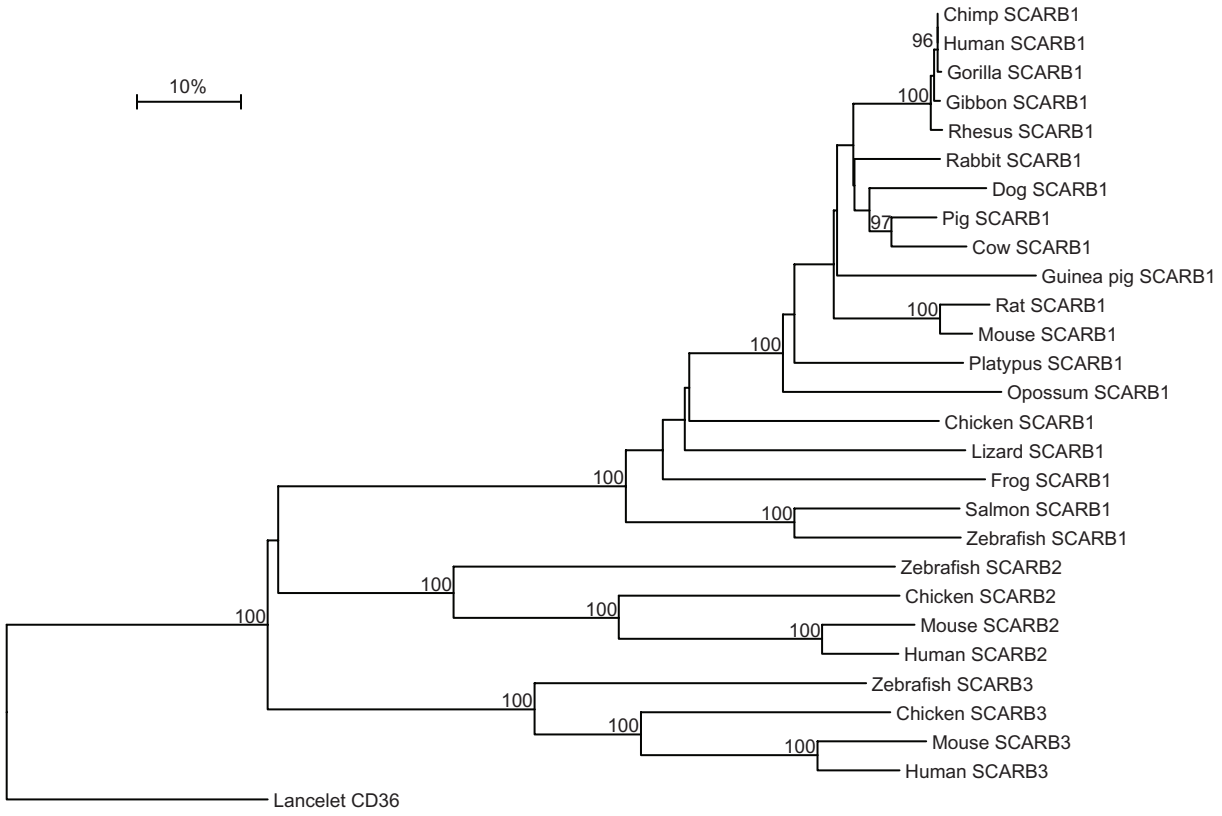

Figure 6 Phylogenetic tree of vertebrate SCARB1 amino acid sequences with human, mouse, chicken, and zebrafish SCARB2 and SCARB3 (CD36) sequences.
Notes: The tree is labeled with the SCARB-like name and the name of the animal and is "rooted" with the lancelet CD36 sequence. Note the three major clusters corresponding to the *SCARB1, SCARB2*, and *SCARB3 (CD36)* gene families. A genetic distance scale is shown. The number of times a clade (sequences common to a node or branch) occurred in the bootstrap replicates are shown. Only replicate values of 95 or more, which are highly significant, are shown, with 100 bootstrap replicates performed in each case. A proposed sequence of *CD36* gene duplication events is shown.
Abbreviation: SCARB, scavenger receptor class B.

sequences with human, mouse, chicken, and zebrafish SCARB2 and SCARB3 (CD36) sequences, which was "rooted" with the lancelet (*B. floridae*) CD36 sequence (see Table 1). The phylogram showed clustering of the SCARB1 sequences into groups, which was consistent with their evolutionary relatedness as well as groups for human, mouse, chicken, and zebrafish SCARB2 and SCARB3 (CD36) sequences, which was distinct from the lancelet CD36 sequence. These groups were significantly different from each other (with bootstrap values of about 100/100). This is suggestive of a sequence of *CD36*-like gene duplication events: ancestral *CD36* (*SCARB3*) gene duplication → *SCARB1* and *CD36* genes; followed by a further *CD36* duplication, generating the *SCARB2* and *CD36* genes found in all vertebrate species examined (Figure 6). It is apparent from this study of vertebrate CD-like genes and proteins that this is an ancient protein for which a proposed common ancestor for the *SCARB1, SCARB2*, and *SCARB3 (CD36)* genes may have predated the appearance of fish more than 500 million years ago.[95]

Conclusion

The results of the present study indicate that vertebrate *SCARB1* genes and encoded proteins represent a distinct gene and protein family of CD36-like proteins which share key conserved sequences that have been reported for other CD36-like proteins (SCARB2 and SCARB3 [CD36]) previously studied.[19–29] SCARB1 has a unique property among these proteins in serving as a major receptor for HDL in the body and participating in reverse cholesterol transport.[10,33,46,49] HDL binding to SCARB1 stimulates endothelial nitric oxide synthase activity, apparently by increasing intracellular ceramide levels, which leads to cholesterol transfer from HDL into the target cell.[4,96] SCARB1 has also been implicated in the remodeling of large lipoprotein particles by endothelial lipase during SCARB1-mediated uptake of HDL cholesterol.[97] *SCARB1* is encoded by a single gene among the vertebrate genomes studied and is highly expressed in human and mouse tissues, particularly in adrenal glands, liver, lymphocytes, and embryonic tissues, and usually contains 12 coding exons. Predicted secondary

structures for vertebrate SCARB1 proteins show strong similarities with other CD36-like proteins, SCARB2 and SCARB3 (or CD36). Three major structural domains were apparent for vertebrate SCARB1, including the N-terminal and C-terminal cytoplasmic domain, the N-terminal and C-terminal transmembrane domain, and the exoplasmic domain, containing the "lipid-transfer active" free SH Cys384, two disulfide bridges, and several N-glycosylation sites for glycan binding. Phylogenetic studies using 19 vertebrate SCARB1 sequences with human, mouse, chicken, and zebrafish SCARB2 and SCARB3 (CD36) sequences indicated that the *SCARB1* gene has appeared early in vertebrate evolution following a gene duplication event of the ancestral *CD36* (*SCARB3*) gene, prior to the appearance of bony fish, more that 500 million years ago.

Acknowledgments

This project was supported by the National Institutes of Health (grants P01 HL028972 and P51 RR013986). In addition, this investigation was conducted in facilities constructed with support from Research Facilities Improvement Program (Grants 1 C06 RR13556, 1 C06 RR15456, and 1 C06 RR017515). We also acknowledge the expert assistance of Dr Bharet Patel of Griffith University with the phylogeny studies.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Acton S, Rigotti A, Landschulz KT, Xu S, Hobbs HH, Krieger M. Identification of scavenger receptor SR-BI as a high density lipoprotein receptor. *Science*. 1996;271(5248):518–520.
2. Ikemoto M, Arai H, Feng D, et al. Identification of a PDZ-domain-containing protein that interacts with the scavenger receptor class B type I. *Proc Natl Acad Sci U S A*. 2000;97(12):6538–6543.
3. Bultel-Brienne S, Lestavel S, Pilon A, et al. Lipid free apolipoprotein E binds to the class B type I scavenger receptor I (SR-BI) and enhances cholesteryl ester uptake from lipoproteins. *J Biol Chem*. 2002; 277(39):36092–36099.
4. Li XA, Titlow WB, Jackson BA, et al. High density lipoprotein binding to scavenger receptor, Class B, type I activates endothelial nitric-oxide synthase in a ceramide-dependent manner. *J Biol Chem*. 2002;277(13):11058–11063.
5. Osgood D, Corella D, Demissie S, et al. Genetic variation at the scavenger receptor class B type I gene locus determines plasma lipoprotein concentrations and particle size and interacts with type 2 diabetes: the Framingham study. *J Clin Endocrinol Metab*. 2003;88(6): 2869–2879.
6. Marsche G, Zimmermann R, Horiuchi S, Tandon NN, Sattler W, Malle E. Class B scavenger receptors CD36 and SR-BI are receptors for hypochlorite-modified low density lipoprotein. *J Biol Chem*. 2003;278(48):47562–47570.
7. Levy E, Ménard D, Suc I, et al. Ontogeny, immunolocalisation, distribution and function of SR-BI in the human intestine. *J Cell Sci*. 2004;117(Pt 2):327–337.
8. Connelly MA, Williams DL. Scavenger receptor BI: A scavenger receptor with a mission to transport high density lipoprotein lipids. *Curr Opin Lipidol*. 2004;15(3):287–295.
9. Out R, Kruijt JK, Rensen PCN, et al. Scavenger receptor BI plays a role in facilitating chylomicron metabolism. *J Biol Chem*. 2004;279(18):18401–18406.
10. Kent AP, Stylianou IM. Scavenger receptor class B member 1 protein: hepatic regulation and its effects on lipids, reverse cholesterol transport and atherosclerosis. *Hepatic Medicine:Evidence and Research*. 2011;3:29–44.
11. Scarselli E, Ansuini H, Cerino R, et al. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J*. 2002;21(19):5017–5025.
12. Bartosch B, Vitelli A, Granier C, et al. Cell entry of hepatitis C virus requires a set of co-receptors that include the CD81 tetraspanin and the SR-B1 scavenger receptor. *J Biol Chem*. 2003;278(43): 41624–41630.
13. Dorner M, Horwitz JA, Robbins JB, et al. A genetically humanized mouse model for hepatitis C virus infection. *Nature*. 2011;474(7350): 208–211.
14. Thi DVL, Dreux M, Cosset FL. Scavenger receptor class B type I and the hypervariable region-1 of hepatitis C virus in cell entry and neutralisation. *Expert Rev Mol Med*. 2011;13:e13.
15. Nofer JR, van Eck M. HDL scavenger receptor class B type 1 and platelet function. *Curr Opin Lipidol*. 2011;22(4):277–282.
16. Kawasaki Y, Nakagawa A, Nagaosa K, Shiratsuchi A, Nakanishi Y. Phosphatidylserine binding of class B scavenger receptor type I, a phagocytosis receptor of testicular Sertoli cells. *J Biol Chem*. 2002; 277(30):27559–27566.
17. Nishiuchi T, Murao K, Imachi H, et al. Scavenger receptor class BI mediates the anti-apoptotic effect of erythropoietin. *Ann Med*. 2010; 42(2):151–160.
18. Feng H, Guo L, Wang D, et al. Deficiency of scavenger receptor BI leads to impaired lymphocyte homeostasis and autoimmune disorders in mice. *Arterioscler Thromb Vasc Biol*. 2011;31(11):2543–2551.
19. Fujita H, Takata Y, Kono A, et al. Isolation and sequencing of a cDNA clone encoding the 85 kDa human lysosomal sialoglycoprotein (hLGP85) in human metastatic pancreas islet tumor cells. *Biochem Biophys Res Commun*. 1992;184(2):604–611.
20. Ogata S, Fukuda M. Lysosomal targeting of Limp II membrane glycoprotein requires a novel Leu-Ile motif at a particular position in its cytoplasmic tail. *J Biol Chem*. 1994;269(7):5210–5217.
21. Tabuchi N, Akasaki K, Sasaki T, Kanda N, Tsuji H. Identification and characterization of a major lysosomal membrane glycoprotein, LGP85/LIMP II in mouse liver. *J Biochem*. 1997;122(4):756–763.
22. Kuronita T, Eskelinen EL, Fujita H, Saftig P, Himeno M, Tanaka Y. A role for the lysosomal membrane protein LGP85 in the biogenesis and maintenance of endosomal and lysosomal morphology. *J Cell Sci*. 2002;115(Pt 21):4117–4131.
23. Lin YW, Lin HY, Tsou YL, et al. Human SCARB2-mediated entry and endocytosis of EV71. *PLoS One*. 2012;7(1):e30507.
24. Simantov R, Silverstein. CD36: a critical anti-angiogenic receptor. *Front Biosci*. 2003;8:S874–S882.
25. Adachi H, Tsujimoto M. Endothelial scavenger receptors. *Prog Lipid Res*. 2006;45(5):379–404.
26. Collot-Teixeira S, Martin J, McDermott-Roe C, Poston R, McGregor JL. CD36 and macrophages in atherosclerosis. *Cardiovasc Res*. 2007;75(3):468–477.
27. Martin CA, Longman E, Wooding C, et al. CD36, a class B scavenger receptor, functions as a monomer to bind acetylated and oxidized low-density lipoproteins. *Protein Sci*. 2007;16(11):2531–2541.
28. Gautam S, Banerjee M. The macrophage Ox-LDL receptor, CD36 and its association with type II diabetes mellitus. *Mol Genet Metab*. 2011; 102(4):389–398.
29. Ren Y. Peroxisome proliferator-activator receptor γ: a link between macrophage CD36 and inflammation in malaria infection. *PPAR Res*. 2012;2012:640769.

30. McLean MP, Sandhoff TW. Expression and hormonal regulation of the high-density lipoprotein (HDL) receptor scavenger receptor class B type I messenger ribonucleic acid in the rat ovary. *Endocrine*. 1998;9(3):243–252.
31. Thierry-Mieg D, Thierry-Mieg J. AceView: A comprehensive cDNA-supported gene and transcripts annotation. *Genome Biol*. 2006;7 Suppl 1: S12.1– S12.14.
32. Van Eck M, Twisk J, Hoekstra M, et al. Differential effects of scavenger receptor BI deficiency on lipid metabolism in cells of the arterial wall and in the liver. *J Biol Chem*. 2003;278(26):23699–23705.
33. Vergeer M, Korporaal SJ, Franssen R, et al. Genetic variant of the scavenger receptor BI in humans. *N Engl J Med*. 2011;364(2):136–145.
34. Tai ES, Adiconis X, Ordovas JM, et al. Polymorphisms at the SRBI locus are associated with lipoprotein levels in subjects with heterozygous familial hypercholesterolemia. *Clin Genet*. 2003;63(1):53–58.
35. Roberts CG, Shen H, Mitchell BD, Damcott CM, Shuldiner AR, Rodriguez A. Variants in scavenger receptor class B type I gene are associated with HDL cholesterol levels in younger women. *Hum Hered*. 2007;64(2):107–111.
36. Liu Y, Ordovas JM, Gao G, et al. The SCARB1 gene is associated with lipid response to dietary and pharmacological interventions. *J Hum Genet*. 2008;53(8):709–717.
37. Cerda Á, Genvigir FD, Rodrigues AC, et al. Influence of polymorphisms and cholesterol-lowering treatment on SCARB1 mRNA expression. *J Atheroscler Thromb*. 2011;18(8):640–651.
38. Acton S, Osgood D, Donoghue M, et al. Association of polymorphisms at the SR-BI gene locus with plasma lipid levels and body mass index in a white population. *Arterioscler Thromb Vasc Biol*. 1999;19(7):1734–1743.
39. Constantineau J, Greason E, West M, et al. A synonymous variant in scavenger receptor, class B, type I gene is associated with lower SR-BI protein expression and function. *Atherosclerosis*. 2010; 210(1):177–182.
40. Pérez-Martínez P, Pérez-Jiménez F, Bellido C, et al. A polymorphism exon 1 variant at the locus of the scavenger receptor class B type I (SCARB1) gene is associated with differences in insulin sensitivity in healthy people during the consumption of an olive oil-rich diet. *J Clin Endocrinol Metab*. 2005;90(4):2297–3000.
41. Ritsch A, Sonderegger G, Sandhofer A, et al. Scavenger receptor class B type I polymorphisms and peripheral arterial disease. *Metabolism*. 2007;56(8):1135–1141.
42. Rodrigues CD, Hannus M, Prudêncio M, et al. Host scavenger receptor SR-BI plays a dual role in the establishment of malaria parasite liver infection. *Cell Host Microbe*. 2008;4(3):271–278.
43. Gerhard DS, Wagner L, Feingold EA, et al; MGC Project Team. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Res*. 2004; 14(10B):2121–2127.
44. Zhang X, Merkler KA, McLean MP. Characterization of regulatory intronic and exonic sequences involved in alternative splicing of scavenger receptor class B gene. *Biochem Biophys Res Commun*. 2008;372(1):173–178.
45. Welch CL, Xia YR, Gu LJ, et al. Srb1 maps to mouse chromosome 5 in a region harboring putative QTLs for plasma lipoprotein levels. *Mamm Genome*. 1997;8(12):942–944.
46. Johnson MS, Svensson PA, Helou K, et al. Characterization and chromosomal localization of rat scavenger receptor class B type I, a high density lipoprotein receptor with a putative leucine zipper domain and peroxisomal targeting sequence. *Endocrinology*. 1998;139(1):72–80.
47. Rajapaksha WR, McBride M, Robertson L, O'Shaughnessy PJ. Sequence of the bovine HDL-receptor (SR-BI) cDNA and changes in receptor mRNA expression during granulosa cell luteinization in vivo and in vitro. *Mol Cell Endocrinol*. 1997;134(1):59–67.
48. Sundvold H, Helgeland H, Baranski M, Omholt SW, Våge DI. Characterisation of a novel paralog of scavenger receptor class B member I (SCARB1) in Atlantic salmon (*Salmo salar*). *BMC Genet*. 2011;12:52.
49. Webb NR, Connell PM, Graf GA, et al. SR-BII, an isoform of the scavenger receptor BI containing an alternate cytoplasmic tail, mediates lipid transfer between high density lipoprotein and cells. *J Biol Chem*. 1998;273(24):15241–15248.
50. Eckhardt ER, Cai L, Shetty S, et al. High density lipoprotein endocytosis by scavenger receptor SR-BII is clathrin-dependent and requires a carboxyl-terminal dileucine motif. *J Biol Chem*. 2006;281(7):4348–4353.
51. Naj AC, West M, Rich SS, et al. Association of scavenger receptor class B type I polymorphisms with subclinical atherosclerosis: the Multi-Ethnic Study of Atherosclerosis. *Circ Cardiovasc Genet*. 2010;3(1):47–52.
52. Altschul F, Vyas V, Cornfield A, et al. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–410.
53. Lander ES, Linton LM, Birren B, et al; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860–921.
54. Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*. 2005;437(7055):69–87.
55. Gibbs RA, Rogers J, Katze MG, et al. Evolutionary and biomedical insights from the rhesus macaque genome. *Science*. 2007; 316(5822):222–234.
56. Elsik CG, Tellam RL, Worley KC, et al; Bovine Genome Sequencing and Analysis Consortium. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*. 2009; 324(5926):522–528.
57. Waterston RH, Lindblad-Toh K, Birney E, et al; Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature*. 2002;420(6915):520–562.
58. Gibbs RA, Weinstock GM, Metzker ML; Rat Genome Sequencing Project Consortium. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature*. 2004;428(6982): 493–521.
59. Mikkelsen TS, Wakefield MJ, Aken B, et al. Genome of the marsupial *Monodelphis domestica* reveals innovation in noncoding sequences. *Nature*. 2007;447(7141):167–175.
60. Warren WC, Hillier LW, Marshall Graves JA, et al. Genome analysis of the platypus reveals unique signatures of evolution. *Nature*. 2008;453(7192):175–183.
61. International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*. 2004;432(7018):695–716.
62. Eckalbar WL, Lasku E, Infante CR, et al. Somitogenesis in the anole lizard and alligator reveals evolutionary convergence and divergence in the amniote segmentation clock. *Dev Biol*. 2012;363(1): 308–319.
63. Hellsten U, Harland RM, Gilchrist MJ, et al. The genome of the western clawed frog *Xenopus tropicalis*. *Science*. 2010;328(5978):633–636.
64. Sprague J, Bayraktaroglu L, Clements D, et al. The zebrafish information network: the zebrafish model organism database. *Nucleic Acids Res*. 2006;34(Database Issue):D581–D585.
65. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2002;12(6):994–1006.
66. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000;16(4):404–405.
67. Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput*. 2002; 7:310–322.
68. Su AI, Wiltshire T, Batalov S, et al. A gene atlas of the human and mouse protein encoding transcriptomes. *Proc Natl Acad Sci U S A*. 2004;101(16):6062–6067.
69. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser*. 1999;41:95–99.
70. Kimura M. *The Neutral Theory of Molecular Evolution*. Cambridge, UK: Cambridge University Press; 1983.

71. Van De Peer Y, de Wachter R. TreeCon for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput Appl Sci*. 1994;10(5):569–575.
72. Saitou N, Nei M. The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4(4):406–411.
73. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 1985;39:783–791.
74. Calvo D, Vega MA. Identification, primary structure, and distribution of CLA-1, a novel member of the CD36/LIMPII gene family. *J Biol Chem*. 1993;268(25):18929–18935.
75. Silver DL. A carboxyl-terminal PDZ-interacting domain of scavenger receptor B, Type I is essential for cell surface expression in liver. *J Biol Chem*. 2002;277(37):34042–34047.
76. Gaidukov L, Nager AR, Xu S, Penman M, Krieger M. Glycine dimerization motif in the N-terminal transmembrane domain of the high density lipoprotein receptor SR-BI required for normal receptor oligomerization and lipid transport. *J Biol Chem*. 2011;286(21):18452–18464.
77. Yua M, Romer KA, Nieland TJ, et al. Exoplasmic cysteine Cys384 of the HDL receptor SR-BI is critical for its sensitivity to a small-molecule inhibitor and normal lipid transport activity. *Proc Natl Acad Sci U S A*. 2011;108(30):12243–12248.
78. Papale GA, Hanson PJ, Sahoo D. Extracellular disulfide bonds support scavenger receptor class B type I-mediated cholesterol transport. *Biochem*istry. 2011;50(28):6245–6254.
79. Chen R, Jiang X, Sun D, et al. Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. *J Proteome Res*. 2009:8(2):651–661.
80. Junyent M, Arnett DK, Tsai MY, et al. Genetic variants at the PDZ-interacting domain of the scavenger receptor class B type I interact with diet to influence the risk of metabolic syndrome in obese men and women. *J Nutr*. 2009;139(5):842–848.
81. Kocher O, Birrane G, Yesilaltay A, et al. Identification of the PDZ3 domain of the adaptor protein PDZK1 as a second, physiologically functional binding site for the C terminus of the high density lipoprotein receptor scavenger receptor class B type I. *J Biol Chem*. 2011;286(28):25171–25186.
82. Ranganathan R, Ross EM. PDZ domain proteins: Scaffolds for signaling complexes. *Curr Biol*. 1997;7(12):R770–R773.
83. Lee H-J, Zheng JI. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun Signal*. 2010;8:8.
84. MacArthur MW, Thornton JM. Influence of proline residues on protein conformation. *J Mol Biol*. 1991;218(2):397–412.
85. Slepkov ER, Rainey JK, Sykes BD, Fliegel L. Structural and functional analysis of the Na^+/H^+ exchanger. *Biochem J*. 2007;401(3):623–633.
86. Kay BK, Williamson MP, Sudol M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J*. 2000;14(2):231–241.
87. Coleman MD, Bass RB, Mehan RS, Falke JJ. Conserved glycine residues in the cytoplasmic domain of the aspartate receptor play essential roles in kinase coupling and on-off switching. *Biochem*. 2005;44(21):7687–7695.
88. Krieger, A Möglich A, Kiefhaber T. Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide F. *J Am Chem Soc*. 2005;127(10):3346–3352.
89. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A*. 2006;103(10):1412–1417.
90. Malerød L, Sporstøl M, Juvet LK, Mousavi A, Gjøen T, Berg T. Hepatic scavenger receptor class B, type I is stimulated by peroxisome proliferator-activated receptor gamma and hepatocyte nuclear factor 4alpha. *Biochem Biophys Res Commun*. 2003;305(3):557–565.
91. Merkel M, Eckel RH, Goldberg IJ. Lipoprotein lipase: genetics, lipid uptake, and regulation. *J Lipid Res*. 2002;43(12):1997–2006.
92. SteinY, Stein O. Lipoprotein lipase and atherosclerosis. *Atherosclerosis*. 2003;170(1):1–9.
93. Ashok-Kumar M, Subhashini NGV, Kanthimathi S, et al. Associations for lipoprotein lipase and peroxisome proliferator-activated receptor-γ gene and coronary heart disease in an Indian population. *Arch Med Res*. 2010;41(1):19–25.
94. Stefani G, Slack FJ. Small non-coding RNAs in animal development. *Nat Rev Mol Cell Biol*. 2008;9(3):219–230.
95. Donoghue PCJ, Benton MJ. Rocks and clocks: calibrating the tree of life using fossils and molecules. *Trends Ecol Evol*. 2007;22(8):424–431.
96. Assanasen C, Mineo C, Seetharam D, et al Cholesterol binding, efflux, and a PDZ-interacting domain of scavenger receptor-BI mediate HDL-initiated signaling. *J Clin Invest*. 2005;115(4):969–977.
97. Nijstad N, Wiersma H, Gautier T, van der Giet M, Maugeais C, Tietge UJ. Scavenger receptor BI-mediated selective uptake is required for the remodeling of high density lipoprotein by endothelial lipase. *J Biol Chem*. 2009;284(10):6093–6100.

Supplementary figure

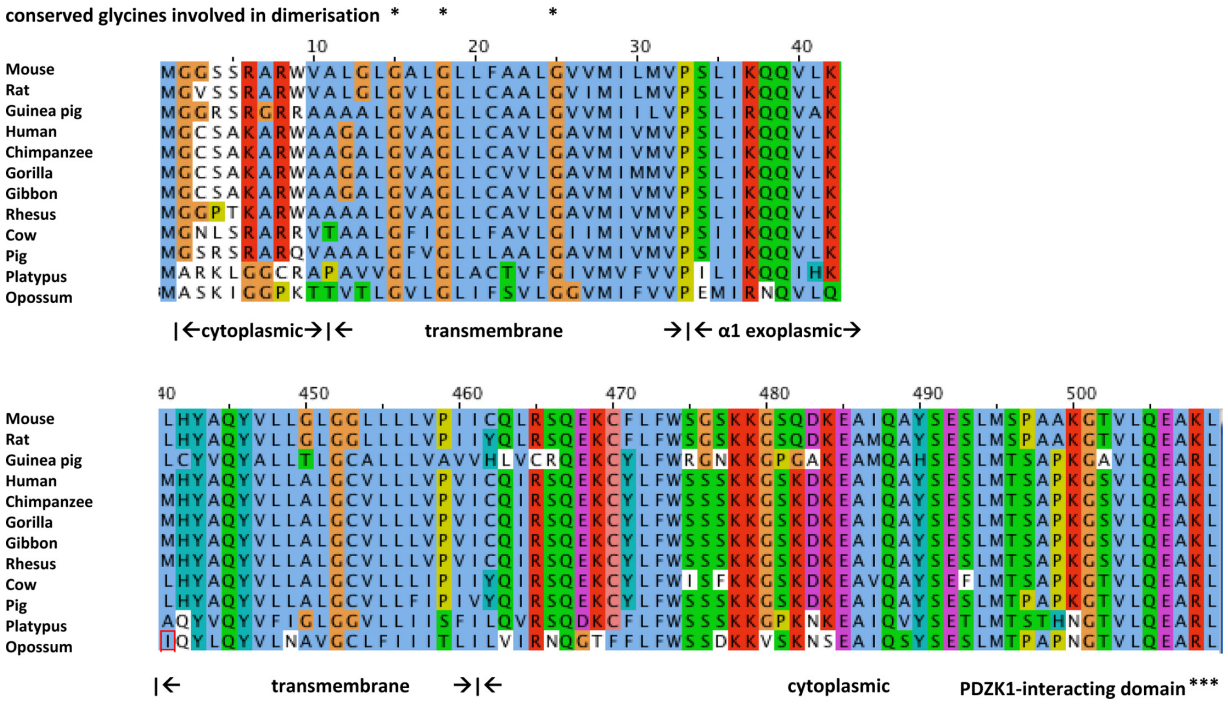

Figure S1 Amino acid sequence alignments for mammalian SCARB1 N-terminal and C-terminal sequences.
Notes: Amino acids are color-coded: yellow for proline (P); S (serine); green for hydrophilic amino acids, S (serine), Q (glutamine), N (asparagine), and T (threonine); brown for glycine (G); light blue for hydrophobic amino acids, L (leucine), I (isoleucine), V (valine), M (methionine), W (tryptophan); dark blue for amino acids, T (tyrosine) and H (histidine); purple for acidic amino acids, E (glutamate) and D (aspartate); and red for basic amino acids, K (lysine) and R (arginine); conserved N-terminal glycines are shown as*; the last three C-terminal amino acids (PDZK1-interacting zone) are also shown as*.
Abbreviation: SCARB1, scavenger receptor class B type 1.