Identification and Functional Analysis of SEDL-binding and Homologue Proteins by Immobilized GST Fusion and Motif Based Methods

Ji-Man Hong,[a] Mi Suk Jeong,[a,†] Jae Ho Kim,[‡] Boo Gil Kim,[§] Stephen R. Holbrook,[#] and Se Bok Jang[*]

Department of Molecular Biology, Pusan National University, Busan 609-735, Korea. [*]E-mail: sbjang@pusan.ac.kr
[†]Research Center for Advanced Science and Technology, Dongseo University, Busan 617-716, Korea
[‡]Department of Physiology, College of Medicine, Pusan National University, Busan 602-739, Korea
[§]Division of Architecture & Civil Engineering, Dongseo University, Busan 617-716, Korea
[#]Structural Biology Department, Physical Biosciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA
Received October 29, 2007

An X-linked skeletal disorder, SEDT (spondyloepiphyseal dysplasia tarda) is a genetic disease characterized by a disproportionately short trunk and short stature caused by mutations in the SEDL gene. This gene is evolutionarily conserved from yeast to human. The yeast SEDL protein ortholog, Trs20p, has been isolated as a member of a large multi-protein complex called the transport protein particle (TRAPP), which is involved in endoplasmic reticulum (ER)-to-Golgi transport. The interaction between SEDL and partner proteins is important in order to understand the molecular mechanism of SEDL functions. We isolated several SEDL-binding proteins derived from rat cells by an immobilized GST-fusion method. Furthermore, the SEDL-homologue proteins were identified using motif based methods. Common motifs between SEDL-binding proteins and SEDL-homologue proteins were classified into seven types and 78 common motifs were revealed. Sequence similarities were contracted to seven types using phylogenetic trees. In general, types I-III and VI were classified as having the function of acetyl-CoA carboxylase, glycogen phosphorylase, isocitrate dehydrogenase, and enolase, respectively, and type IV was found to be functionally related to the GST protein. Types V and VII were found to contribute to TRAPP vesicle trafficking.

Key Words : SEDL-binding proteins, Immobilized GST fusion, Motif analysis

Introduction

Mutations in the human SEDL gene cause spondyloepiphyseal dysplasia tarda (SEDT), a rare X-linked chondrodysplasia. SEDL mRNA is widely distributed in a variety of tissues, including heart, liver, and placenta. SEDL is highly conserved from yeast to human.[1] Its ortholog in budding yeast, Trs20p, is a subunit of the transport protein particle (TRAPP) complex that plays key roles in vesicle docking/fusion processes during endoplasmic reticulum (ER)-to-Golgi transport.[2-4] The yeast TRAPP complex consists of seven different proteins (Bet5p, Bet3p, Trs20p, Trs23p, Trs31p, Trs33p, and Trs85p). Recent studies indicate that the TRAPP complex is a nucleotide exchanger for Ypt1 and Ypt31/32, which are GTPases essential for membrane trafficking.[5,6] A 2.4 Å resolution structure of SEDL revealed an unexpected similarity to the structure of the N-terminal regulatory domain of two SNAREs, Ykt6p and Sec22b.[7-9] This finding suggests a possible interaction between subunits of the TRAPP complex and SNAREs. However, direct binding of TRAPP to SNAREs has not been documented. Bet3p, a component of TRAPP, was shown to interact genetically with yeast SNAREs and is required for vesicle targeting and function.[10] SEDL may serve as an adaptor protein in the constitution of the TRAPP complexes via multiple protein-protein interactions. An unusual property of SEDL is that its surface is apparently well suited for multiple protein-protein interactions involving different contact areas.

SEDL also binds to MBP-1, a transcription repressor for the c-myc promoter.[11] Repression of c-myc expression by MBP-1 results in apoptosis. Binding of SEDL appears to sequester MBP-1 in the cytoplasm, thus releasing its repression effects. These data indicate that SEDL is also involved in cell growth regulation. The intracellular chloride channel proteins CLIC-1 and CLIC-2 have been reported to associate with SEDL by yeast two-hybrid screening.[12] Since both SEDL and CLIC proteins are membrane-associated proteins, their interactions might be important either for function of the TRAPP complex or the proper targeting and function of the intracellular chloride channels.

Understanding the interactions between SEDL and its partner proteins is important to clarify the molecular mechanism of SEDL functions. Protein-protein interactions play a critical role in biological processes such as signal transduction pathways. We isolated several SEDL-binding proteins derived from rat cells with immobilized GST-SEDL. Furthermore, the homologues of SEDL proteins were identified by motif and domain approach methods. The homologues can be proposed to have a function as SEDL binding activity. We developed methods for predicting interaction motifs from the sequences. The results provide valuable information for the prediction of metabolic pathways from

[a]These authors contributed equally to this article.

SEDL-related protein-protein interactions.

Materials and Methods

Plasmid Constructs. Full-length mouse SEDL was amplified by PCR with oligonucleotides incorporating the *Eco*RI and the *Xho*I sites on the 5' primer and 3' primer containing stop codons as follows.

5'-CCGGAATTCCGGATGTCTGGGAGCTTCTACTTCG-3'
5'-CCGCTCGAGCGGTTAGCTTAAAAGGTGTTTCTTCC-3'

The amplifications were performed using a procedure of 25 cycles of reaction with denaturing at 94 °C for 1 min, annealing at 54 °C for 1 min, and extension at 72 °C for 1 min. After digestion with endonucleases *Eco*RI and *Xho*I, the SEDL fragment was inserted into the GST-fusion expression vector pGEX-4T-1 (Pharmacia Biotech) and was linked by T4 ligase, producing a recombinant vector pGEX-4T-1-SEDL. The positive GST-SEDL fusion expression plasmid was identified by restriction endonuclease digestion and further verified by DNA sequencing on an automatic DNA sequencer by Macrogen. This was expressed as glutathione *S*-transferase (GST)-tagged fusion proteins in BL-21 cells (Stratagene), and affinity-purified with glutathione-Sepharose as suggested by the manufacturer.

Protein Identification by Peptide Mass Fingerprinting Analysis. Rat liver tissues were lysed in buffer A (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 5 mM $MgCl_2$, 0.1 mM phenylmethylsulfonyl fluoride, 5 μg/mL aprotinin, and 1 μM pepstatin); unbroken cells were removed by centrifugation at 1,000 × *g* for 10 min followed by centrifugation at 100,000 × *g* for 40 min. The pellets were treated with buffer A containing 1% Triton X-100 to solubilize membrane-associated proteins and centrifuged at 100,000 × *g* for 40 min to remove insoluble pellets. Immobilized GST fusion proteins were reacted with aliquots (5 mg) of the membrane extracts and washed three times with buffer A containing 0.1% Triton X-100 prior to SDS-electrophoresis. After staining with Coomassie Brilliant Blue, the candidate band was excised from the gel and digested with trypsin as described.[13] The masses of the tryptic peptides were measured with a Voyager DE time-of-flight mass spectrometer (Perspective Biosystems, Inc., Framingham, MA) at Pohang University of Science and Technology. Matrix-assisted laser desorption/ionization was performed with α-cyano-4-hydroxycinnamic acid as the matrix. Comparison of the mass values against the Swiss-Prot database was performed using Peptide Search.[14]

Basic Local Alignment Search Tool (BLAST). Homologues of SEDL proteins were searched for using SEDL query sequence and various species related to vertebrate in the NCBI (National Center for Biotechnology Information) Basic Local Alignment Search Tool.[15,16] BLAST is a sequence similarity search program that can be used via a web interface.

Pfam Domain Database and Prosite Motif Database. Common domains and motifs of SEDL-homologue proteins were searched for using Pfam (protein families database of alignments) and Prosite databases, respectively.[17-20] The process starts from seed alignments (either a multiple alignment or an alignment from other databases such as Prosite) of a non-redundant representative set of known members.

Multiple Sequence Alignment and Phylogenetic Tree ClustalW. Multiple sequence alignment using the CCBB mirroring service of ClustalW server was executed on the fasta format file of 100 homologues of SEDL proteins.[21,22] The output file format of ClustalW is PHYLIP. Clustal program series are widely used in molecular biology for the multiple alignment of both nucleic acid and protein sequences and for preparing phylogenetic trees.

PHYLIP and Treeview. Bootstrapping analysis using the PHYLIP program was executed.[23,24] One hundred bootstraps, 5 random seeds, and the Dayhoff PAM matrix and distance were used with neighbor joining. POWER uses an open-source LAMP structure and infers genetic distances and phylogenetic relationships using well-established algorithms (ClustalW and PHYLIP). Treeview was used as a viewer program.[25] To compare sequences three forms: radial (unrooted) tree, rectangular cladogram, and phylogram were used.

Results and Discussion

Protein-protein interactions play a critical role in biological processes such as signal transduction pathways. Identification and functional analysis of the SEDL homologue proteins were shown (Figure 3 and Table 5). SEDL-homologue proteins were extracted using the SEDL sequence in the NCBI BLAST. SEDL-homologue proteins were extracted by multiple sequence alignment and phylogenetic analysis using ClustalW and the PHYLIP programs.

Nine SEDL-binding proteins were identified by peptide mass fingerprinting analysis (MALDI-TOF) derived from rat liver tissue with immobilized GST-SEDL (Table 1 and Figure 1). Five SEDL-binding proteins excluding the redundant proteins of nine SEDL-binding proteins were used with NCBI VAST to search for structural similarity. One (GI 4929901) of five SEDL-binding proteins has had its three dimensional structure determined by X-ray crystallography (PDB ID: 1B8X). 1B8X is a structure of NMTS (the nuclear matrix targeting signal) in the transcription factor AML-1 (acute myelogenous leukemia) fused to GST (Tang *et al.*, 1999). The structure was identified to having structurally similarity with 2GST by NCBI VAST search. To examine the evolutionary relationships between 1B8X and 2GST, the phylogenetic analysis was performed using multiple alignments. PHYLIP and Treeview were used for phylogenetic analysis. The result showed that both proteins are in the same cluster in accord with structure and sequence (Figure 2a). It can be proposed that GST-fusion SEDL interacts with 1B8X. SEDL has also been reported to associate with CLIC1 and CLIC2 by yeast two hybrid screening.[12] The three-dimensional structure of CLIC1 is similar to the GST structure.[26] Therefore, bioinformatics analysis suggests that SEDL may interact with CLIC1.

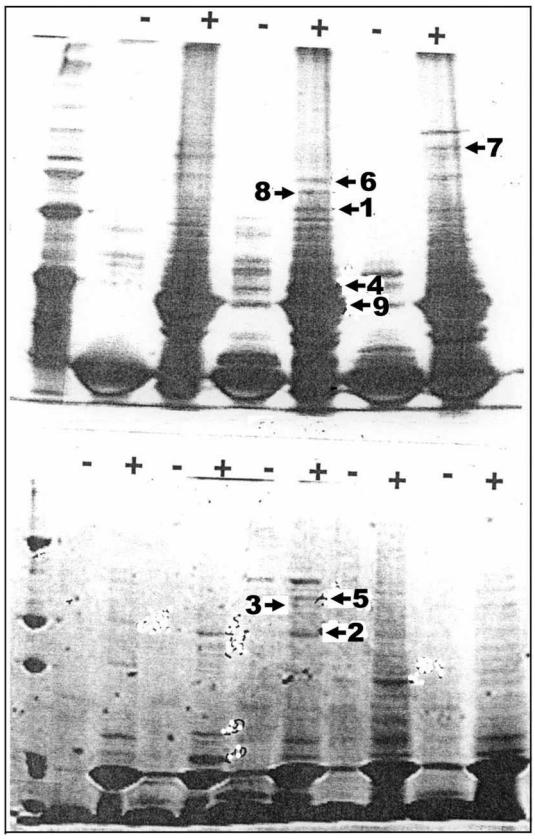

Figure 1. Identification of SEDL-binding proteins. 5 mg extract was incubated with GST-SEDL immobilized to GSH beads. Bound proteins were resolved by SDS-PAGE and stained with Coomassie Brilliant Blue dye. The protein bands from 1 to 9 were excised and digested with trypsin, and the resulting peptides were analyzed by MALDI-TOF mass spectroscopic analysis. The matched peaks are shown. (−) : GST controls, and (+) : binding proteins.

Table 1. SEDL-binding proteins derived from rat cells with immobilized GST fusion

GI	Name
① 4929901	Chain A, glutathione S-transferase fused with the nuclear matrix targeting signal of the transcription factor AML-1
② 11560087	liver glycogen phosphorylase
③ 31543464	pyruvate carboxylase
④ 1708406	IDHP_MOUSE isocitrate dehydrogenase [NADP], mitochondrial precursor
⑤ 8393186	carbamoyl-phosphate synthetase 1, mitochondrial
⑥ 28525872	muscle glycogen phosphorylase
⑦ 8393186	carbamoyl-phosphate synthetase 1, mitochondrial
⑧ 4929901	Chain A, glutathione S-transferase fused with the nuclear matrix targeting signal of the transcription factor AML-1
⑨ 4929901	Chain A, glutathione S-transferase fused with the nuclear matrix targeting signal of the transcription factor AML-1

GI: Genbank accession number. ①-⑨: Total SEDL-binding proteins. ①-⑤: Five SEDL-binding proteins excluding redundant proteins. ⑥-⑨: Redundant proteins

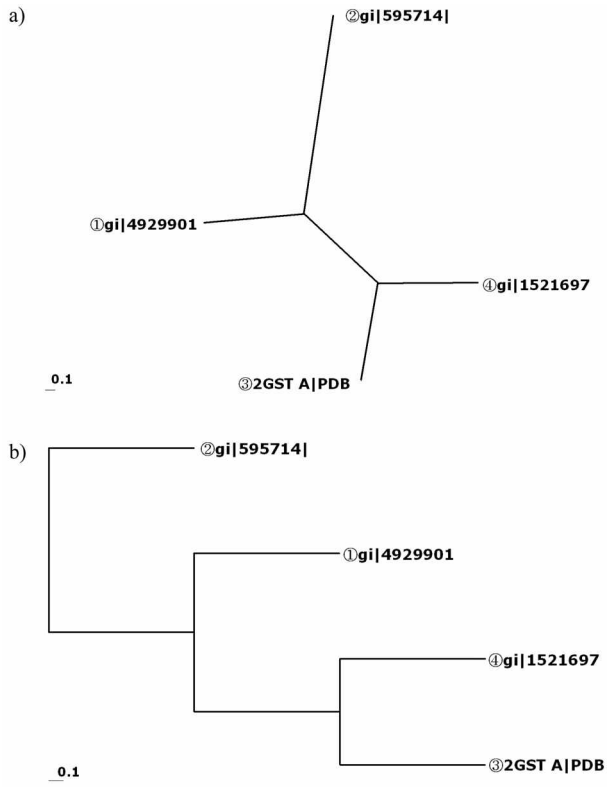

Figure 2. a) It is a radial tree among homologues such as GI 4929901(PDB: 1B8X_A), PDBID 2GST (highest structural homologue using VAST search tool) and GI 595714 and GI 15216974 (homologues of two with highest Bit score using BLAST search tool) b) Multiple sequence alignment (MSA) tree of 1B8X with SEDL-binding proteins. The tree was constructed by multiple sequence alignment of 5 SEDL-binding proteins excluding redundancy and NCBI deletion among 9 SEDL-binding proteins using the PHYLIP program. Two clusters, ③-④ and ①-⑤, are closely related evolutionarily. ① Chain A, glutathione S-transferase fused with the nuclear matrix targeting signal of transcription factor Aml-1, ② liver glycogen phosphorylase [Rattus norvegicus], Pyruvate carboxylase [Rattus norvegicus], ④ IDHP_MOUSE, isocitrate dehydrogenase [NADP], mitochondrial precursor, ⑤ carbamoyl-phosphate synthetase 1, mitochondrial [Rattus norvegicus].

The similarity between the five SEDL-binding proteins has been calculated. The multiple alignments were executed in order to understand the evolutionary relationship using the CCBB mirroring service of ClustalW server. From the phylogenetic analysis, one cluster was found in the liver glycogen phosphorylase and carbamoyl-phosphate (#2. gi/11560087 and #5. gi/8393186) (Figure 2b).

We tried to find common domains from SEDL-binding proteins and SEDL-homologue proteins using the Pfam program but were not successful. However, seven common motifs were found using the Prosite program. Two motif patterns - the casein kinase II phosphorylation site and the protein kinase C phosphorylation site - were found as relatively highly conserved motifs of the SEDL related proteins (Table 2).

SEDL-homologue and SEDL-binding proteins comprise

Table 2. Conserved motif patterns among SEDL-binding and SEDL-homologue proteins

Common motifs	Binding	Homologue
Casein kinase II phosphorylation site	Normal	High
Amidation site	Low	Normal
Protein kinase C phosphorylation site	High	Normal
Tyrosine kinase phosphorylation site	Normal	Low
Tyrosine sulfation site	Normal	Low
N-myristorylation site	High	Low
N-glycosylation site	High	Low

Binding: SEDL-binding proteins (experiment). Homologue: SEDL-homologue proteins (database). High: common motif pattern with complete agreement. Normal: common motif pattern with 2/3 agreement. Low: common motif pattern with 1/3 agreement

fourteen non-redundant sequences (Tables 3 and 4). Redundancy was determined using pairwise sequence alignment of the sequences. The reference sequence (Refseq) database from NCBI was used to extract highly homologous sequences.[27] All homologues of the fourteen sequences were extracted from the collected SEDL related proteins (Table 4). The search tool and option used were BLAST and BLASTP.[15] The sequences with E-value < 10^{-5} were extracted and divided into NP and XP. NP is the data sequence number obtained from experiments and XP is that obtained from model proteins.

SEDL-homologue and its binding proteins were divided into seven groups after checking overlap of each sequence. The multiple sequence alignments using two analysis methods were performed for the seven homologous proteins. First, a multiple sequence alignment, Muscle, was used for 309 sequences.[28] Another alignment divided them into seven forms. ClustalW was used for high accuracy alignment.[21] Common motifs of the total sequences and forms were extracted by multiple sequence alignment. The 78 conserved motifs of types II, III, V, and VI were shown in Table 5. Types I, IV, and VII have only weakly conserved motifs. Sequence similarities were classified into seven types of highly homologous proteins after multiple sequence alignment using phylogenetic trees as shown in Figure 3. The

Table 4. Homologue summary of SEDL related proteins

GI	Source	Homologues (NP+XP)	NP
8393186	experiment	144	28
31543464	experiment	156	36
11560087	experiment	60	22
1708406	experiment	27	14
4929901	experiment	165	66
13384832	database	19	10
35186898	database	13	9
5734091	database	13	9
6319731	database	19	10
1946954	database	20	11
119339	database	84	23
33468929	database	44	27
12963491	database	85	23
6755448	database	43	21

NP: data sequence number obtained from experiment (SEDL-binding proteins). XP: data sequence number obtained from model protein (SEDL-homologue proteins)

alignment tools were the Phylip 3.65 package, Treeview 1.6 (phylogram), and PAUP 4.0 (classification of species, balance tree).[29-31] Seqboot (100 boostrapping, 5 random seed, and sets of aligned sequence), Protdist (PAM matrix, protein sequence), Neighbor (neighbor joining, distance matrix), and Consense (high reliability) were used on the Phylip 3.65 program. SEDL-homologue proteins were also classified as to species distribution (Table 6).

Type I includes 10 sub-groups and six branches. The functional information of each sub-group was obtained from the KEGG pathway database, which provides information on molecular interaction networks such as pathways and complexes, information about genes and proteins generated by genome projects, and information about biochemical compounds and reactions.[32] Sub-groups 1-3 are functionally related to acetyl-CoA carboxylase, which is associated with breast cancer predisposition, and sub-groups 4-5 catalyze the synthesis of mitochondrial carbamoyl-phosphate.[33,34] Sub-group 6 was assumed to related to propionyl-CoA carboxylase in biotin metabolism, and sub-group 7, to methylcrotonoyl-

Table 3. SEDL-homologue proteins from the sequence analysis

GI	Species	SEDL-homologue proteins
13384832	M. musculus	trafficking protein particle complex 2
35186898	M. musculus	spondyloepiphyseal dysplasia tarda protein
5734091	H. sapiens	sedlin
6319731	S. cerevisiae	one of 10 subunits of the transport protein particle (TRAPP) complex of the cis-Golgi which mediates vesicle docking and fusion; mutations in the human homologue cause spondyloepiphyseal dysplasia tarda (SEDT)
1946954	C. elegans	human SEDL (spondyloepiphyseal dysplasia tarda) related protein 1
119339	H. sapiens	alpha-enolase(2-phospho-D-glyceratehydro-lyase) (Non-neural enolase) (NNE) (Enolase1) (Phosphopyruvate hydratase) (C-myc promoter-binding protein) (MBP-1)(Plasminogen-binding protein).
33468929	M. musculus	tobrevin like 1,vesicle-associated membrane protein 7; tetanus neuro toxin-insensitive vesicle-associated membrane
12963491	M. musculus	enolase 1, alpha non-neuron alpha-enolase; 2-phospho-D-glycerate hydrolase
6755448	M. musculus	SEC22 vesicle trafficking protein-like 1

Table 5. Conserved motifs of types II, III, V, and VI

Type	Position	Sequence of Motif	Type	Position	Sequence of Motif
Type II	13-15	QIS		115-123	KCATITPDE
	17-18	RG		125-129	RVEEF
	31-46	FNRHLHFTLVKDRNV		131-132	LK
	49-51	RDY		134-135	MW
	55-56	AL		137-152	SPNGTIRNILGGTVFR
	59-73	TVRDHLVGRWIRTQQ		155-164	IICKNIPRLV
	80-82	PKR		165-167	GW
	87-91	SLEFY		173-177	IGRHA
	93-100	GRTLQNTM		179-182	GDQY
	102-103	NL		184-188	ATDFV
	109-112	CDEA		219-220	GG
	114-119	YQLGLD		224-225	GM
	127-149	EEDAGLGNGGLGRLAACFLDSMA		231-232	SI
	151-163	LGLAAYGYGIRYE		235-238	FAHS
	165-169	GIFNQ		240-241	FQ
	174-176	GWQ		248-249	WP
	178-183	EEADDW		253-258	STKNTI
	187-192	GNPWEK		262-271	YDGRFKDIFQ
	194-196	RPE		278-279	YK
	200-203	PVHF		289-300	YEHRLIDDMVAQ
	219-221	DTQ		303-304	KS
	224-226	VLA		310-322	WACKNYDGDVQSD
	228-232	PYDTP		325-327	AQG
	234-236	PGY			
	879-880	ML	Type V	60-61	YL
	883-889	HDRFKVF		76-77	VT
	891-893	DYE		116-112	NPFY
	898-899	CQ		130-131	AF
	924-932	GKFSSDRTI			
	935-936	YA	Type VI	84-85	PT
	939-940	IW		135-136	AV
	943-944	EP		244-247	VSLA
				250-251	KA
Type III	53-56	VVEM		259-261	PLY
	58-63	GDEMTR		445-447	DPF
	65-66	IW		538-539	DL
	69-74	IKEKLI		548-549	IK
	86-87	DL		556-557	ER
	92-93	RD		560-562	KYN
	99-100	VT		567-569	IEE

CoA carboxylase 1 in valine, leucine, and isoleucine degradation.[35] Sub-groups 8-10 are related to pyruvate carboxylase, which requires biotin and ATP to catalyse the carboxylation of pyruvate to oxaloacetate.[36] Type I was classified as having the function of acetyl-CoA carboxylase and carbamoyl-phosphate synthetase.

Type II includes four sub-groups. Those sub-groups are related to the functions of brain, liver, and muscle glycogen phosphorylase in the insulin-signaling pathway, starch, and sucrose metabolism.[37] The sub-groups of type III are functionally related to isocitrate dehydrogenase 2 ($NADP^+$), they play a role in high-glucose-induced apoptosis and contribute to various pathologies associated with the long-term complications of diabetes.[38]

The sub-groups of type IV are functionally related to glutathione S-transferase. Eight distinct classes of the soluble cytoplasmic mammalian glutathione S-transferases have been identified: alpha, kappa, mu, omega, pi, sigma, theta, and zeta.[39] GSTs are cytosolic dimeric proteins involved in the cellular detoxification of electrophilic compounds, including carcinogens, therapeutic drugs, environmental toxins and products of oxidative stress. Sub-groups 1-8 are of the mu classes of enzymes, sub-groups 9-10 are of the pi class, sub-group 11 is of the sigma class, and sub-groups 12-16 are of the alpha classes.

Three sub-groups of type V are related to trafficking

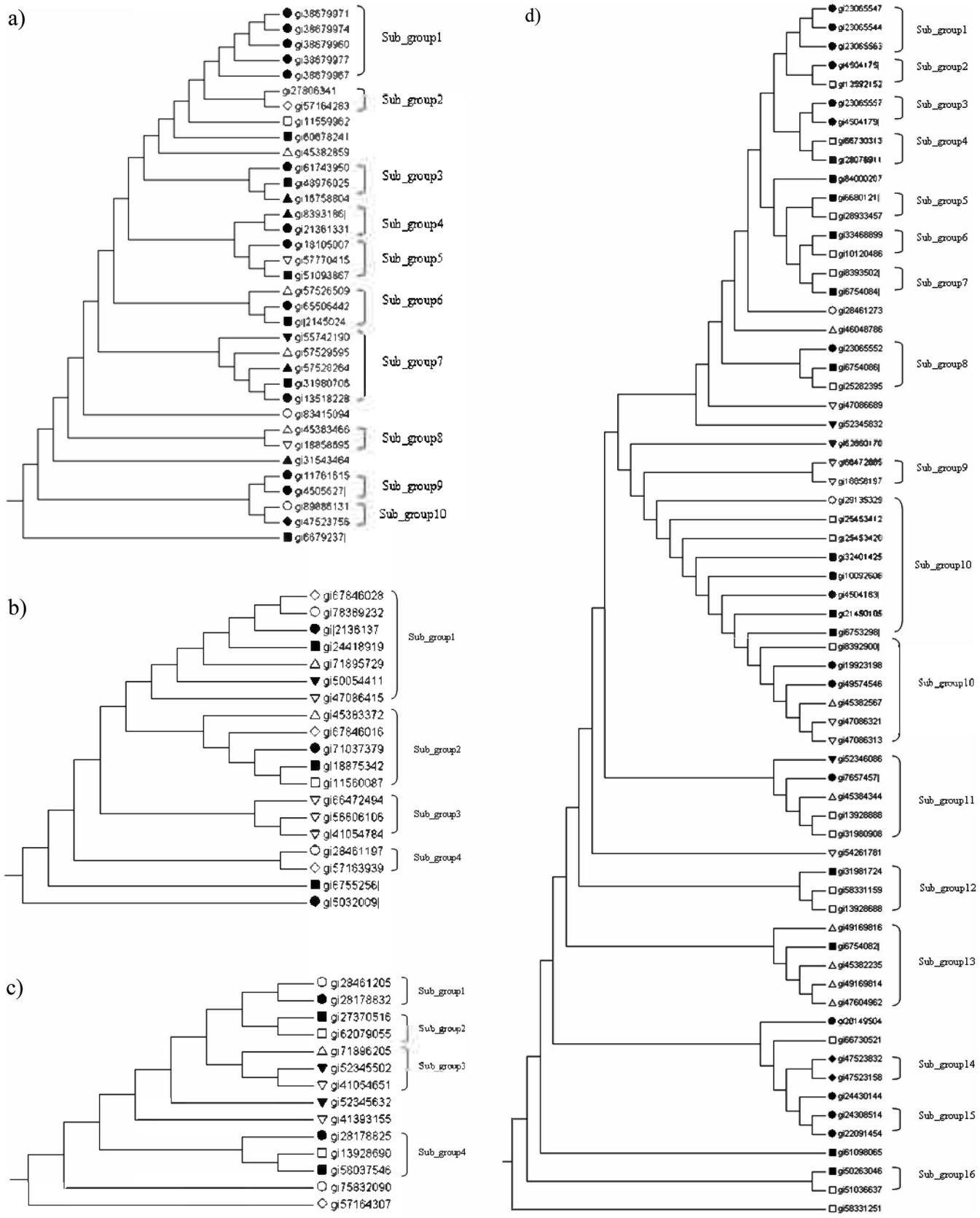

Figure 3. Classification of SEDL-homologue proteins using phylogenetic trees. a) Type I: acetyl-CoA carboxylase and carbamoyl-phosphate synthetase, b) type II: glycogen phosphorylase, c) type III: isocitrate dehydrogenase, d) type IV: glutathione S-transferase, e) type V: TRAPP, f) type VI: enolase, and g) type VII: vesicle trafficking.

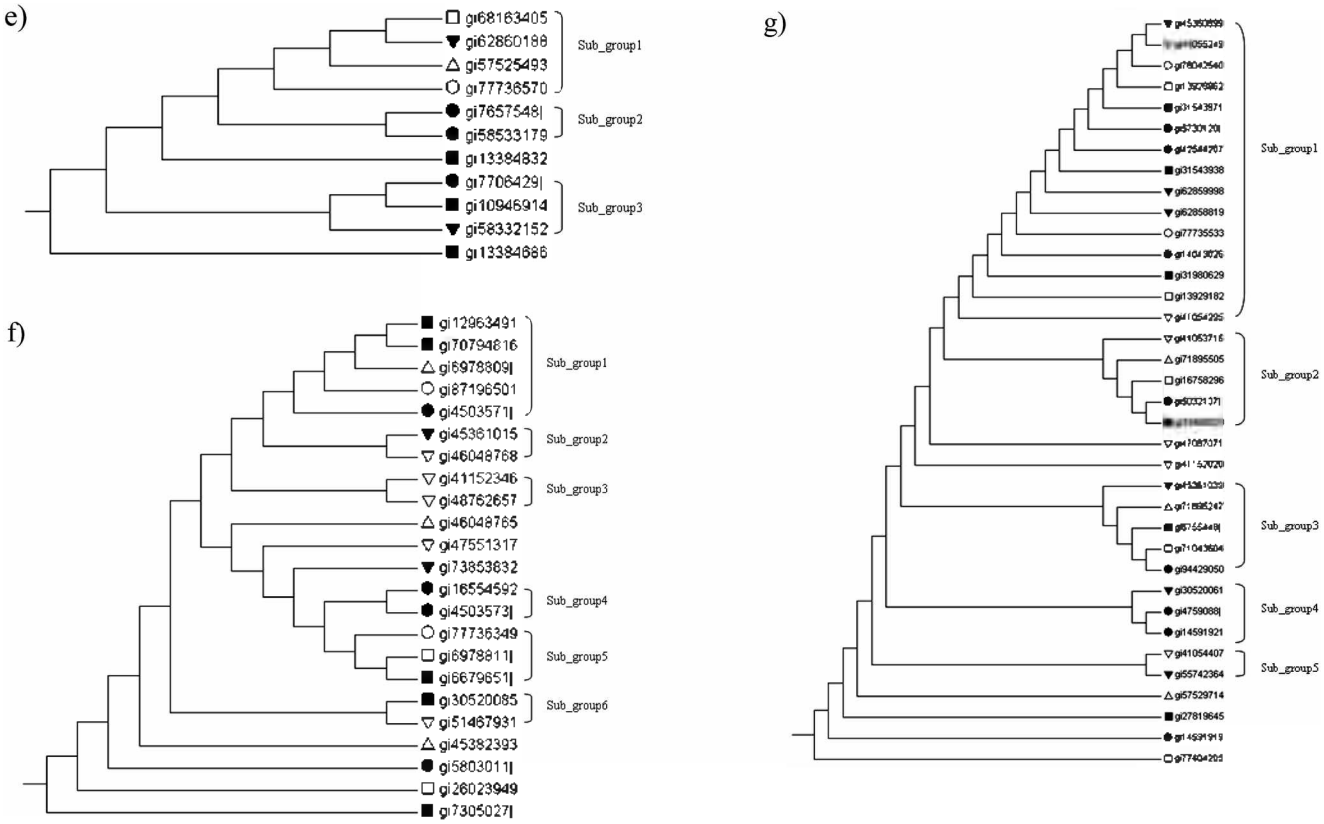

Figure 3. Continued.

Table 6. Species distribution of SEDL-homologue proteins

Species	Symbol	Type I	Type II	Type III	Type IV	Type V	Type VI	Type VII
Homo sapiens	●	12	3	2	15	3	4	8
Bos taurus	○	1	2	2	2	1	2	2
Mus musculus	■	6	3	2	15	3	5	7
Rattus norvegicus	□	5	1	2	15	1	3	5
Gallus gallus	△	3	2	1	7	1	3	3
Danio rerio	▽	4	4	2	6	0	4	6
Xenopus tropicalis	▼	1	1	2	3	2	2	5
Ovis aries	◥	1	3	1	0	0	0	0
Sus scrofa	◆	1	0	0	2	0	0	0
acetyl-coA carboxylase	none	1	0	0	0	0	0	0
Total		35	19	14	65	11	23	36

protein particle complex proteins involved in the targeting and fusion of ER-to-Golgi transport vesicles with their acceptor compartment.[40] Six sub-groups of type VI are related to enolase isoenzymes, which show tumor suppressor activity and growth inhibitory effects, and which also function in glycolysis and the gluconeogenesis pathway.[41] Type VII includes five sub-groups. Sub-group 1 is functionally related to vesicle-associated membrane, prenylated SNARE, and YKT 6v-SNARE proteins. Sub-group 2 is related to synaptobrevin-like 1 in intracellular trafficking and secretion, and sub-groups 3-5 are related to SEC22 vesicle trafficking proteins, which complex with SNARE and play a role in the ER-to-Golgi protein trafficking.[42]

Conclusion

The SEDL protein is well suited for multiple protein-protein interactions, which can include the binding of a SNARE domain in the process of ER-to-Golgi vesicle transport. Nevertheless, the mechanism by which SEDL mutations cause rare osteochondrodysplasia is not known.[43,44] To further investigate the functions of SEDL in cells, we performed a SEDL binding experiment and motif searching for SEDL-associated proteins. Five SEDL-binding proteins and nine SEDL-homologue proteins were found, and those proteins were classified into seven common motifs. However, among those motifs, types I, IV, and VII were found to

have weakly conserved common motifs. We further investigated the sequence similarities using phylogenetic tree. based on the results, types I-VII were reclassified. In general. types I-III and VI were classified as having the functions of acetyl-CoA carboxylase, glycogen phosphorylase, isocitrate dehydrogenase, and enolase, respectively, and type IV was found to be functionally related to the GST protein. Types V and VII were found to contribute to TRAPP vesicle trafficking. We isolated rat-cells-derived SEDL-binding proteins by an immobilized GST-fusion method. and developed methods for predicting interaction motifs from the sequence. The results can provide valuable information for the prediction of metabolic pathways from SEDL-related protein-protein interactions.

Acknowledgements. This study was supported by a grant of the National R&D Program for Cancer Control. Ministry of Health & Welfare. Republic of Korea (0720110) to S.B.J. and by the Korea Research Foundation Grants funded by the Korean Government (MOEHRD) to M.S.J. (KRF-2005-075-C00019).

References

1. Gedeon, A. K.; Colley, A.; Jamiesion. R.; Thompson. E. M.: Rogers, J.; Sillence. D.; Tiller. G. E.; Mulley, J. C.: Gecz, J. *Nat. Genet.* **1999**, *22*, 400-404.
2. Sacher. M.; Jiang. Y.; Barrowman. J.; Scarpa. A.; Burston. J.; Zhang. L.; Schieltz. D.; Yates. J. R.; Abeliovich. H.; Ferro-Novick, S. *EMBO J.* **1998**, *17*. 2494-2503.
3. Sacher, M.; Barrowman. J.; Wang. W.; Horecka. J.; Zhang. Y.; Pypaert. M.; Ferro-Novick. S. *Mol. Cell* **2001**. *7*, 433-442.
4. Barrowman. J.; Sacher. M.; Ferro-Novick. S. *EMBO J.* **2000**. *19*. 862-869.
5. Jones. S.; Newman. C.; Liu. F.; Segev. N. *Mol. Biol. Cell* **2000**. *11*. 4403-4411.
6. 6. Wang. W.; Sacher, M.; Ferro-Novick. S. *J. Cell Biol.* **2000**, *151*. 289-296.
7. Tochio. H.; Tsui. M. M.; Banfield. D. K.; Zhang. M. *Science* **2001**. *293*. 698-702.
8. Gonzalez. L. C.; Jr. Weis. W. I.; Scheller. R. H. *J. Biol. Chem.* **2001**, *276*, 24203-24211.
9. Jang. S. B.; Kim. Y. G.; Cho. Y.-S.; Suh. P. G.; Kim. K. H.; Oh. B. H. *J. Biol. Chem.* **2002**, *277*, 49863-49869.
10. Rossi. G.; Kolstad. K.; Stone. S.; Palluault. F.; Ferro-Novick. S. *Mol. Biol. Cell* **1995**. *6*. 1769-1780.
11. Ghosh. A. K.; Majumder. M.; Steele. R.; White. R. A.; Ray. R. B. *Mol. Cell Biol.* **2001**. *21*. 655-662.
12. Fan, L.; Yu. W.; Zhu, X. *FEBS Letters* **2003**. *540*, 77-80.
13. Jensen. O. N.; Vorm. O.; Mann, M. *Electrophoresis* **1996**, *17*. 938-944.
14. Jensen. O. N.; Podtelejnikov. A. V.; Mann. M. *Anal. Chem.* **1997**. *69*. 4741-4750.
15. Altschul. S. F.; Gish. W.; Miller. W.; Myers. E. W.; Lipman. D. J. *J. Mol. Biol.* **1990**. *215*(3). 403-410.
16. McGinnis. S.; Madden. T. L. *Nucleic Acids Research* **2004**. *32*. W20-W25.
17. Sonnhammer. E. L. L.; Eddy. S. R.; Birney. E.; Bateman. A.; Durbin. R. *Nucleic Acids Research* **1998**. *26*(1). 320-322.
18. Finn. R. D.; Mistry. J.; Schuster-Bockler. B.; Griffiths-Jones. S.; Hollich. V.; Lassmann. T.; Moxon. S.; Marshall. M.; Khanna. A.; Durbin. R.; Eddy. S. R.; Sonnhammer. E. L. L.; Bateman. A. *Nucleic Acids Research* **2006**. *34*. D247-D251.
19. Bairoch. A. *Nucleic Acids Research* **1991**. *19*. 2241-2245.
20. Hulo. N.; Bairoch. A.; Bulliard. V.; Cerutti. L.; Castro. E. D.; Langendijk-Genevaux. P. S.; Sigrist. C. J. A. *Nucleic Acids Research* **2006**. *34*. D227-D230.
21. Thompson. J. D.; Higgins. D. G.; Gibson. T. J. *Nucleic Acids Research* **1994**. *22*(22). 4673-4680.
22. Chenna. R.; Sugawara. H.; Koike. T.; Lopez. R.; Gibson. T. J.; Higgins. D. G.; Thompson. J. D. *Nucleic Acids Research* **2003**. *31*(13). 3497-3500.
23. Baum. B. R. *PHYLIP: Phylogeny Inference Package*. Version 3.2. Quarterly Review: **1989**. *64*. 539-541.
24. Lim. A.; Zhang. L. *WebPHYLIP: a Web Interface to PHYLIP. BIOINFORMATICS APPLICATIONS NOTE* **1999**. *15*(12). 1068-1069.
25. Saldanha. A. J. *Bioinformatics* **2004**. *20*(17). 3246-8.
26. Harrop. S. J.; DeMaere. M. Z.; Fairlie. W. D.; Reztsova. T.; Valenzuela. S. M.; Warton. K.; Bauskin. A. R.; Wu. W. M.; Pankhurst. S.; Campbell. T. J.; Breit. S. N.; Curmi. P. M. G. *J. Biol. Chem.* **2001**. *276*. 44993-45000.
27. Kim. D. P.; Tatiana. T.; Donna. R. M. *Nucleic Acids Research* **2005**. *33*. 501-504.
28. Robert. C. E. *Nucleic Acids Research* **2004**. *32*. 1792-1797.
29. Felsenstein. J. *PHYLIP (phylogeny inference package)*. version 3.57c: Department of Genetics, University of Washington: Seattle. 1995.
30. Page. R. D. *Comput. Appl. Biosci.* **1996**. *12*(4). 357-358.
31. Swofford. D. L. *PAUP. Phylogenetic Analysis Using Parsimony*. Version 4.0; Sinauer Associates: Sunderland. MA. 1998.
32. Kanehisa. M.; Goto. S.; Kawashima. S.; Okuno. Y.; Hattori. M. *Nucleic Acids Research* **2004**. *32*. 277-280.
33. Ray. H.; Moreau. K.; Dizin. E.; Callebaut. I.; Venezia. N. D. *J. Mol. Biol.* **2006**. *359*. 973-982.
34. Tygstrup. N.; Bangert. K.; Ott. P.; Gisgaard. H. C. *Biochem. Biophys. Res. Commun.* **2002**. *290*. 518-525.
35. Link. V.; Shevchenko. A.; Heisenberg. C. P. *BMC Dev. Biol.* **2006**. *6*. 1-9.
36. Fahien. L. A.; Davis. J. W.; Laboy. J. *J. Biol. Chem.* **1993**. *268*. 17935-17942.
37. Gevaert. K.; Goethals. M.; Martens. L.; van Damme. J.; States. A.; Thomas. G. R.; Vandekerckhove. J. *Nat. Biotechnol.* **2003**. *21*. 566-569.
38. Shin. A. H.; Kil. I. S.; Yang. E. S.; Huh. T. L.; Yang. C. H.; Park. J. W. *Biochem. Biophys. Res. Commun.* **2004**. *325*. 32-38.
39. Campbell. E.; Takahashi. Y.; Abramovitz. M.; Peretz. M.; Listowsky. I. *J. Biol. Chem.* **1990**. *265*. 9188-9193.
40. Adams. M. D.; Soares. M. B.; Kerlavage. A. R.; Fields. C.; Venter. J. C. *Nat. Genet.* **1993**. *4*, 373-380.
41. Ejeskar. K.; Krona. C.; Caren. J.; Zaibak. F.; Li. L.; Martinsson. T.; Ioannou. P. A. *BMC Cancer* **2005**. *5*. 161-174.
42. Parlati. F.; McNew. J. A.; Fukuda. R.; Miller. R.; Sollner. T. H.; Rothman. J. E. *Nature* **2007**. *407*. 194-198.
43. Park. J. S.; Lee. K. M.; Jeong. M. S.; Jin. G.; Jang. S. B. *Bull. Korean Chem. Soc.* **2007**. *28*, 574-580.
44. Jeong. M. S.; Jang. S. B. *Bull. Korean Chem. Soc.* **2006**, *27*, 87-92.