Over-expression of a human CD62L ecto-domain and a potential role of RNA pseudoknot structures in recombinant protein expression

Matthew Spencer, **Nathan Max**, **Joanna Ireland**, **Zhongcheng Zou**, **Ruipeng Wang**, and **Peter Sun**[*]

Structural Immunology Section, Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 12441 Parklawn Drive, Rockville, MD 20852, United States

Abstract

L-selectin (CD62L) is an extracellular protein with a lectin-like domain that mediates rolling adhesion of leukocytes to vascular endothelial cell surfaces. Currently, there are no solved structures for the ectodomain of CD62L, nor of CD62L in complex with its ligand. We have developed a rapid mammalian recombinant protein expression system using an amplifiable glutamine synthase based vector. Here, we further developed and applied this method to express and purify the entire extracellular region of CD62L. This resulted in excess of 20 mg/liter yield of recombinant CD62L. In an attempt to understand the different expression levels among four similar CD62L constructs that differ primarily in signal sequences, we calculated the presence of potential RNA pseudoknots in their signal sequences. The results showed the presence of pseudoknots involving the start codon and between the signal sequence and gene in the mRNA of the non-expressing constructs, suggesting a potential inhibitory role of RNA pseudoknots in recombinant protein expression.

Introduction

Human L-Selectin (CD62L) is a member of the selectin family of proteins, a group that binds carbohydrates and differ mainly in the number of sushi repeat domains found after the lectin-binding and epidermal growth factor (EGF) domains [1]. This cell surface protein is involved in rolling adhesion of leukocytes to endothelial cells [2], which is used for lymphocyte homing and inflammation responses [3]. While the structures of E- and P-selectins have been determined with bound carbohydrates [4], such a structure has not been determined for L-selectin, and consequently the structural mechanism of L-selectin binding to its carbohydrate ligand remain unresolved. Furthermore, the only L-selectin structures

Corresponding author: Peter Sun, psun@nih.gov.

available are those corresponding to truncated lectin and EGF domain constructs [5, 6], leaving the role of the two sushi domains undefined.

To facilitate the structural and functional studies of L-selectin, we attempted to develop an efficient, stable recombinant over-expression system. We have previously described a method for truncated CD62L ectodomain production [7], in which a step-wise increase of methionine sulfoximine (MSX) concentration was achieved in a single- rather than multi-round selection to amplify the recombinant glutamine synthetase (GS)-containing plasmid. This reduced the time needed to establish stable recombinant-expressing clones from 6 months to less than 3 months. Here, we describe an improvement of this method to produce the full length, four-domain extracellular CD62L that includes a c-type selectin domain, an EGF calcium binding domain, and two sushi repeat domains (Figure 1). The presence of nine disulfide bonds and seven predicted glycosylation sites make this protein difficult to produce in bacteria or insect cells. We used the recently developed MSX-driven GS-system to amplify recombinant CD62L expression [8, 9], similar to the DHFR gene amplification method [10, 11].

Signal peptides are known to affect gene expression and some, such as those from rat serum albumin (RSA) and immunoglobin G (IgG), have been used to enhance recombinant protein expressions [8, 12]. In attempts to over-express CD62L, we generated recombinant CD62L expression constructs using four different signal sequences in the same GS vector. To further understand the effect of various signal sequences on recombinant CD62L expression, we analyzed the presence of stable RNA structures, such as pseudoknots, in four different signal peptide constructs of recombinant CD62L (Figure 2). Pseudoknots are tertiary structures that form in single-stranded RNA, and are found in telomerase RNA, ribozymes, ribosomal RNA, and messenger RNA [13–15]. They mainly occur when nucleotides in a hairpin loop form additional base-pairing with their complementary nucleotides some distance away on the same RNA strand [16]. These stable RNA structures are involved in gene regulation and expression in multiple ways, and pseudoknots found in viral mRNA are known to induce translational frameshifts to allow viruses to produce multiple proteins from a single mRNA [14, 15, 17, 18]. Evidence suggests that pseudoknots function as structural obstructions to induce ribosomal pause, which then initiates frameshifts through a "slippery sequence" [19, 20]. The potential involvement of RNA pseudoknots in modulating recombinant protein expressions, however, is not known. Conceivably, such pseudoknots may cause ribosomes to pause during synthesis of recombinant protein, resulting in kinetic delay in protein synthesis or even aborted translation. Here, our effort resulted in recombinant CD62L yield exceeding 20 mg/L among the high expressing clones, an approximate 5- to 10-fold increase compared to the previous attempt. Our findings suggest that the presence of RNA pseudoknots may lower recombinant protein expression yields.

Materials and Methods

Reagents

DMEM/F-12 medium, Opti-MEM I medium, CHO-S-SFM II medium, penicillin streptomycin (PS), trypsin, fetal bovine serum (FBS), phosphate-buffered saline pH 7.4 (PBS), sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) 4–12% gels,

E-Gel agarose gels, and the Lipofectamine 3000 kit were purchased from Thermo Fisher Scientific (Waltham, MA). GMEM medium, GS supplement, methionine sulfoximine (MSX), bovine serum albumin (BSA), mucin from porcine stomach type II (PSM-II), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), and DNA maxiprep kit were from Sigma-Aldrich (St. Louis, MO). The CD62L ELISA kit was from R&D Systems (St. Louis, MO). Ampicillin was from MP Biotechnologies (Santa Ana, CA). E. coli DH5α mix-and-go transformation kit was from Zymo Research (Irving, CA). HBS-P+ buffer (0.1 M HEPES, 1.5 M NaCl, 0.5% v/v Surfactant P20, pH 7.4), carboxymethylated dextran (CM5) chip, and *N*-hydroxysuccinimide/1-ethyl-3(-3-dimethyl-aminopropyl) carbodiimide hydrochloride (NHS/EDC) were from General Electric Healthcare (Little Chalfont, UK). Centrifugal filter concentrators were from EMD Millipore (Darmstadt, Germany). Anti-CD62L DREG-56 antibody was from eBioscience (San Diego, CA). RNeasy mini kit and Hot Start *Taq* were from QIAGEN (Hilden, Germany). Primers were from Integrated DNA Technologies (Coralville, IA). SuperScript III Reverse Transcriptase PRC kit was from Invitrogen (Carlsbad, CA).

Cell lines

Chinese hamster ovary cell line CHO-lec3.2.8.1 was provided by Dr. Pamela Stanley [21]. Cells were maintained in DMEM/F-12 medium with 10% FBS and 1% PS at 37 °C with 5% CO_2 for transfection and switched to GMEM with 10% FBS, 1% PS, GS supplement, and varying concentrations of MSX during selections. After cells were expanded into triple-layer culture flasks with selection medium containing 500 μM MSX, the culture medium was replaced with CHO-S-SFM II serum-free medium for protein production.

Plasmid constructs and production

The construct used is a GS-based system developed by our lab [8]. The CD62L expression construct with native 1 signal sequence was synthesized by GenScript (Piscataway, NJ), and all other constructs were synthesized by GeneCopoeia (Rockville, MD). The ectodomain of CD62L was inserted between the *BamHI* and *Xho1* restriction sites. The recombinant plasmid, pGS-CD62L, was transformed into E. coli DH5α mix-and-go cells on an LB-plate in the presence of ampicillin for maxiprep.

Transfection of CHO-lec3.2.8.1 cells and selection of stable clones

CHO-lec3.2.8.1 cells were grown in a 100 cm^2 flask until approximately 80% confluence and transfected with a mixture of 30 μg sterile-filtered pGS-CD62L plasmid DNA, 90 μL Lipofactamine 3000, 60 μL P3000 reagent, and 3 mL Opti-MEM. After 20 hours of incubation at 37 °C, cells were trypsinized, resuspended in 200 mL selection medium containing 30 μM MSX, diluted seven-fold, and plated into ten 384 well plates with 50 μL per well. MSX concentrations were increased incrementally from 30 μM to 150 μM, then to 500 μM every two weeks. This was achieved by replacing 25 μL of selection medium in the wells with fresh selection medium containing two times the target MSX concentration. After four days of incubation in the 500 μM MSX selection medium, wells were inspected every three to four days for viable colonies. Selective medium replacements were made for wells exhibiting color change and the CD62L expressions were measured by ELISA following a manufacture's instruction (R&D systems, Inc). The top 10% CD62L expression clones were

trypsinized and transferred to 48 well plates, and their expressions were evaluated again by ELISA upon confluence. The eight best CD62L expressing clones were expanded to T-75 cm^2 flasks, and the top two producers were further expanded into triple-layer flasks for recombinant CD62L production.

Purification of recombinant CD62L

A total of ~2.5 liters of spent medium harvests from triple-layer flasks were centrifuged and filtered prior to loading onto an Ni-NTA column prepared with 50 mM NiCl$_2$ and equilibrated with deionized water. Bound protein was eluted using a 1 M imidazole, 0.5 M NaCl, 20 mM HEPES, pH 7.4 elution buffer in a gradient between 0–100% elution buffer over 10 column volumes. Peaks from the Ni-NTA elution were evaluated by SDS-PAGE. Fractions containing CD62L were combined, concentrated to approximately 2 mL, and further purified using a Superdex 200 HR16/60 size exclusion column (GE Healthcare) that was equilibrated with 100 mM NaCl and 20mM HEPES. Fractions containing CD62L were concentrated to ~8mg/mL concentration in centrifugal filter concentrators. The peptide sequence of purified recombinant CD62L was confirmed by N-terminal peptide sequencing.

BIAcore binding assays

Binding experiments were performed by measuring surface plasmon resonance (SPR) using a BIAcore 3000 instrument (GE Healthcare). Recombinant CD62L and BSA (control) were immobilized to different flow cells on a CM5 surface-based sensor chip using NHS/EDC crosslinking in sodium acetate buffer, pH 5.0. For binding assays, the analyte consisted of a serial dilution of anti-CD62L antibody DREG-56 from 40 to 5 nM or 44 μM PSM-II in HBS-P buffer. PSM-II was not soluble above 44 μM, and gave low signal below this concentration, so three repeats were done at 44 μM to calculate an average. The error produced from SPR is typically as large as the constant derived from the curves.

Pseudoknot calculations and visualization

A total of four recombinant CD62L expression constructs using native 1, native 2, RSA, and IgG signal sequences were investigated (Figure 2). The constructs using native 1 or 2 signal sequences produced clones that expressed recombinant CD62L protein, while the RSA and IgG constructs did not. Since the four constructs differ primarily in their signal sequences, the presence of pseudoknots in the signal sequence (beginning 12 nucleotides 5′ of the start codon) and between the signal sequence and the coding region were calculated using DotKnot from the University of Western Australia [22, 23]. Dot-bracket visualization was performed using the Visualization Applet for RNA (VARNA) [24]. Many viral pseudoknots involve between 20–50 nucleotides; thus, pseudoknot calculations were performed using a sliding 50 nucleotide-window with 25 nucleotide overlaps. Since each signal sequence differs in length, pseudoknots were calculated to include 50 nucleotides from the CD62L structural region.

qRT-PCR comparison of mRNA levels

CHO-lec cells were transiently transfected with each of the four constructs and the CD62L mRNA was quantified by qRT-PCR. β-actin was used as a control. CHO-lec cells were

grown in 6 well plates in DMEM to approximately 50% confluent. Transfection media consisted of a mixture of 2 mL of serum-free Opti-MEM, 6 μL of PEI, and 2 μg of either RSA, IgG, native 1, or native 2 leader constructs and was incubated for 20 minutes at room temperature before adding to the CHO-lec cells. After 3 days at 37 °C, total mRNA was isolated using a Qiagen RNeasy Mini Kit. First-strand cDNA was synthesized using the Invitrogen SuperScript III reverse transcriptase kit, followed by PCR reaction using Hot Start *Taq* from Qiagen. Primers used were: CD62L forward 5′-GAACAAGGAGGACTGCGTGGAGA-3′, CD62L reverse 5′-CTTTAGTTTGTGGCAGGCGTCATC-3′, β-actin forward 5′-CCAAGGCCAACCGTGAAAAGAT-3′, and β-actin reverse 5′-CGACCAGAGGCATACAGGCACAG-3′. A 30-cycle PCR was carried out at 98 °C for 15 minutes, 98 °C for 30 seconds, 50 °C for 30 seconds, 72 °C for 1 minute, repeating steps 2–4 30 times, 72 °C for 2.5 minutes, 4 °C hold.

Results

Recombinant CD62L expression constructs

The extracellular region of human CD62L consists of 345 amino acids, including a signal peptide (1–41), a propeptide (42–51), and the mature peptide (52–345) (Gene accession: NP_000646 and P14151) (Figure 1). Structurally, mature CD62L contains a C-type lectin domain (52–170), an EGF domain (178–205), and two sushi domains (210–268, and 272–330). A six-histidine tag was added to the C-terminus of each construct to facilitate purification. In attempts to over-express CD62L, four constructs with different signal sequences (RSA, IgG, and two different native signal sequences) were generated in the same GS vector (Figure 2). The two native signal sequences, referred to as native 1 (NP_000646) and native 2 (P14151), are taken from two separate Gene bank entries of human CD62L, with native 1 containing an additional 13 amino acids at the 5′-end of the gene. In addition, the construct with native 1 signal sequence also includes the two C-terminal sushi domains, making it a full extracellular construct with a total of 345 amino acids. While using an RSA signal sequence improved recombinant TGF-β1 expression [8], to our surprise, neither RSA nor IgG leader constructs produced detectable levels of recombinant CD62L by ELISA. As all constructs contain the epitope region for ELISA antibodies, the failure of RSA and IgG signal sequences to produce CD62L suggests the effect of the signal sequence in recombinant protein expression depends on its 3′ coding sequences. Both native signal sequences resulted in measurable expressions of CD62L and the current manuscript primarily describes the expression and purification of the native 1 construct, as the expression using the native 2 signal sequences has been reported previously [7].

Production of stable CD62L expressing CHO cell line

Approximately 10^7 CD62L-transfected cells were distributed into ten 384-well plates, resulting in ~2500 cells per well. MSX concentration was increased in three steps to select for stable CD62L-expressing clones. To optimize the recovery of high-expression clones, multiple ELISA assays were carried out to screen a large number of asynchronous stable clones in transfected 384 well plates, as well as to quantify the recombinant CD62L expression among the selectively expanded CD62L-expressing clones for protein

production. Of 3840 wells, 630 (~16%) produced viable colonies, suggesting the clones were derived from single cells. Among them, 44 were expanded to 48 and 24 wells, and the eight highest-expressing clones were expanded to T-75 cm^2 flasks for final yield analysis. Of the 630 clones tested by ELISA, 457 (~70%) expressed < 1mg/L, 112 (~20%) expressed between 1–5 mg/L, 54 (9%) expressed between 6–15 mg/L, and the top seven clones (1%) expressed between 16–25 mg/L (Figure 3). Clone Cl14 produced ~ 24mg/L of CD62L prior to affinity and size-exclusion purification and was expanded for recombinant protein production. In comparison, the expression using the native 2 signal sequence resulted in 1–5 mg/L recombinant CD62L [7]. Thus, the current construct using the native 1 signal sequence yielded approximately 5- to 10-fold more CD62L protein.

Expression and purification of recombinant CD62L

A total of 2.5 L serum-free culture medium of clone CI14 was harvested and loaded onto a 20 mL Ni-NTA column. The bound CD62L was eluted with a 0–1 M imidazole gradient elution. Peak fractions were analyzed by SDS-PAGE and the recombinant CD62L was found in the elution between 50–300 mM imidazole concentrations. All CD62L-containing fractions were combined, concentrated, and further purified through a Superdex 200 HR16/60 size exclusion column. CD62L eluted with an apparent molecular weight of 50 kD, most likely due to the glycosylation of the multiple predicted glycosylation sites (Figure 4). The N-terminal peptide sequencing resulted in WTYHYSEKPMNWQRARRF, corresponding to the beginning of the mature CD62L sequence after the cleavage of pro-CD62L peptide.

To confirm the recombinant CD62L function, binding assays were carried out using SPR. Recombinant CD62L was immobilized on CM5 chip with immobilization levels of 1200, 5000, and 6700 response units (RU), respectively. BSA was immobilized at 13000 RU as a control. Serial dilutions of DREG-56 or 44 μM PSM-II were injected as analytes over the immobilized CD62L. The recombinant CD62L bound to its antibody DREG-56 and PSM-II with affinities of 17 ± 20 nM and 55 ± 54.5 μM, respectively (Figure 5). Similar binding to mucin and anti-CD62L antibody was observed compared to previous publication [7, 25]

Evaluation of the contribution of RNA pseudoknots in recombinant CD62L expression

The over-expression of current recombinant CD62L prompted us to further examine the four similar recombinant CD62L constructs that resulted in drastically different expressions of CD62L. In particular, the failure of both the RSA and IgG signal sequences to express CD62L suggested the previously-observed enhancing effect of the RSA and IgG signal sequences to protein expression is gene-dependent, and that the RSA and IgG signal sequences may interact with downstream CD62L-coding nucleotides to inhibit its expression. To rule out a potential difference in transcription of the four recombinant CD62L constructs, we compared the CD62L mRNA levels of the four expression constructs using qRT-PCR in transiently transfected CHO-lec cells. All four expression plasmids resulted in similar amount of CD62L mRNA regardless of their signal sequences (Figure 6). This suggests that the differences in protein productions are not results of their transcriptions.

We then examined potential influence of RNA pseudoknots in the expression of the four recombinant CD62L constructs. While viral RNA is known to contain pseudoknots that are capable of stalling ribosomes and inducing translational frameshifts to benefit viral gene synthesis, it is not clear if such stable RNA structures are also present in normal mRNA to modulate recombinant protein expression. A survey of published viral RNA pseudoknots showed most of them consist of 20–35 nucleotides with calculated Gibbs free energies of −10 to −30 kcal/mol (Figure 7). For an example of a known pseudoknot found in viral mRNA, Figure 7 shows the 2D representation of dot-bracket format next to the 3D NMR structure. To investigate if RNA pseudoknots contributed to the observed yield differences in CD62L production, we calculated the potential pseudoknot compositions in the signal sequences and within 50 nucleotides of the structural coding region of CD62L in all four recombinant CD62L expression constructs. Interestingly, RNA pseudoknots with similar free energies to those of known viral pseudoknots are present in the signal sequence regions of all four CD62L expression constructs (Figure 8). Second, both the IgG and RSA leader constructs contain potential pseudoknots over the translational initiation AUG codon (and the ribosome binding site in IgG), suggesting a potential barrier in ribosome binding and translation initiation of the two constructs. Third, both RSA and IgG signal sequences appear to form more stable pseudoknots than the two native signal sequences with the 5′-CD62L coding sequences immediately followed (Figure 8). The presence of stable RNA pseudoknots between the signal and coding sequences may explain the failure of RSA and IgG signal sequences in recombinant CD62L expression.

Discussion

Here we have shown that our previously published method for the generation of stable mammalian cell lines for production of recombinant protein is efficient and reproducible. Our method enables the generation of high recombinant protein expressing cell lines in approximately 2 months, much shorter than the 4–6 months required by the conventional multiple-round selection method.

While secondary structures-such as simple hairpins-in mRNA that are close to the start codon are known to inhibit protein translation [26], the potential role of RNA pseudoknots in recombinant protein production is novel and could potentially be used to improve expression yields. Though pseudoknots have been found to be beneficial in gene regulation in viruses, they are also known to act as "ribosomal roadblocks". The more stable a pseudoknot, the more difficult it becomes for a ribosome to break the pseudoknot, and thus more likely to lead to aborted protein synthesis [27]. Here, we examined the potential differences in the pseudoknot structures among the four recombinant CD62L constructs. Our analyses showed that pseudoknots are not only present in viral genes, they are likely present in non-viral genes, such as in the coding region of CD62L, thus could potentially regulate the expression of recombinant genes. Second, the formation of significant pseudoknots between the signal sequences of IgG and RSA and the 5′-CD62L coding sequences suggests expression enhancement from signal sequences is not universal, but rather gene dependent.

It is worth noting that the predicted pseudoknots may differ from those present on mRNA *in vivo* in the absence of experimental validation. Likewise, the calculated Gibbs free energy

may not directly correlated to this "roadblocking" effect, which is also determined by the interactions between the different pseudoknot components [27]. While RNA pseudoknot analyses of the four recombinant CD62L expression constructs are consistent with the notion of stable RNA structures impeding the efficiency of protein translation, further experimental evidence is needed to determine if pseudoknots contribute to the expression of other recombinant proteins. Further, it is not clear if the current improvement in recombinant CD62L production of native 1 over native 2 signal sequences is solely due to their RNA pseudoknot structural differences. Other differences, such as the inclusion of the two sushi domains in the native 1 construct may also affect the recombinant CD62L production.

Acknowledgments

Abbreviations

SPR	surface plasmon resonance
PS	penicillin streptomycin
FBS	fetal bovine serum
PBS	phosphate-buffered saline pH 7.4
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
BSA	bovine serum albumin
PSM-II	mucin from porcine stomach type II
HEPES	4-2-hydroxyethyl-1-piperazineethanesulfonic acid
CM5	carboxymethylated dextran
NHS/EDC	*N*-hydroxysuccinimide/1-ethyl-3-3-dimethyl-aminopropyl carbodiimide hydrochloride
MSX	methionine sulfoximine
GS	glutamine synthetase
EGF	epidermal growth factor
IgG	Immunoglobin G
RSA	rat serum albumin
PCR	polymerase chain reaction
PEI	polyethylenimine

References

1. Kansas GS. Selectins and their ligands: Current concepts and controversies. Blood. 1996; 88:3259–3287. [PubMed: 8896391]

2. Ley K, Laudanna C, Cybulsky MI, Nourshargh S. Getting to the site of inflammation: the leukocyte adhesion cascade updated. Nat Rev Immunol. 2007; 7:678–689. [PubMed: 17717539]

3. Ley K. The role of selectins in inflammation and disease. Trends Mol Med. 2003; 9:263–268. [PubMed: 12829015]

4. Somers WS, Shaw GD, Camphausen RT. Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to SLe(x) and PSGL-1 (vol 103, pg 467, 2001). Cell. 2001; 105:971–971.

5. Structure of lectin and EGF domains of L-selectin. TO BE PUBLISHED.

6. Gifford JL, Ishida H, Vogel HJ. Structural insights into calmodulin-regulated L-selectin ectodomain shedding. J Biol Chem. 2012; 287:26513–26527. [PubMed: 22711531]

7. Brown HA, Roth G, Holzapfel G, Shen S, Rahbari K, Ireland J, Zou Z, Sun PD. Development of an improved mammalian overexpression method for human CD62L. Protein Expr Purif. 2015; 105:8–13. [PubMed: 25286402]

8. Zou Z, Sun PD. Overexpression of human transforming growth factor-beta1 using a recombinant CHO cell expression system. Protein Expr Purif. 2004; 37:265–272. [PubMed: 15358346]

9. Zou Z, Sun PD. An improved recombinant mammalian cell expression system for human transforming growth factor-beta2 and -beta3 preparations. Protein Expr Purif. 2006; 50:9–17. [PubMed: 16901717]

10. Wigler M, Perucho M, Kurtz D, Dana S, Pellicer A, Axel R, Silverstein S. Transformation of mammalian cells with an amplifiable dominant-acting gene. Proc Natl Acad Sci U S A. 1980; 77:3567–3570. [PubMed: 6251468]

11. Schimke RT. Methotrexate Resistance and Gene Amplification - Mechanisms and Implications. Cancer. 1986; 57:1912–1917. [PubMed: 3513938]

12. Coloma MJ, Hastings A, Wims LA, Morrison SL. Novel Vectors for the Expression of Antibody Molecules Using Variable Regions Generated by Polymerase Chain-Reaction. J Immunol Methods. 1992; 152:89–104. [PubMed: 1640112]

13. Cash DD, Cohen-Zontag O, Kim NK, Shefer K, Brown Y, Ulyanov NB, Tzfati Y, Feigon J. Pyrimidine motif triple helix in the Kluyveromyces lactis telomerase RNA pseudoknot is essential for function in vivo. P Natl Acad Sci USA. 2013; 110:10970–10975.

14. Brierley I, Pennell S, Gilbert RJC. Viral RNA pseudoknots: versatile motifs in gene expression and replication. Nat Rev Microbiol. 2007; 5:598–610. [PubMed: 17632571]

15. Peselis A, Serganov A. Structure and function of pseudoknots involved in gene expression control. Wires Rna. 2014; 5:803–822. [PubMed: 25044223]

16. Pleij CWA, Rietveld K, Bosch L. A New Principle of Rna Folding Based on Pseudoknotting. Nucleic Acids Res. 1985; 13:1717–1731. [PubMed: 4000943]

17. Brierley I, Gilbert RJC, Pennell S. RNA pseudoknots and the regulation of protein synthesis. Biochem Soc T. 2008; 36:684–689.

18. Staple DW, Butcher SE. Pseudoknots: RNA structures with diverse functions. Plos Biol. 2005; 3:956–959.

19. Plant EP, Jacobs KL, Harger JW, Meskauskas A, Jacobs JL, Baxter JL, Petrov AN, Dinman JD. The 9-A solution: how mRNA pseudoknots promote efficient programmed −1 ribosomal frameshifting. Rna. 2003; 9:168–174. [PubMed: 12554858]

20. Giedroc DP, Theimer CA, Nixon PL. Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. J Mol Biol. 2000; 298:167–185. [PubMed: 10764589]

21. Patnaik SK, Stanley P. Lectin-resistant CHO glycosylation mutants. Methods Enzymol. 2006; 416:159–182. [PubMed: 17113866]

22. Sperschneider J, Datta A. DotKnot: pseudoknot prediction using the probability dot plot under a refined energy model. Nucleic Acids Res. 2010; 38

23. Sperschneider J, Datta A, Wise MJ. Heuristic RNA pseudoknot prediction including intramolecular kissing hairpins. Rna. 2011; 17:27–38. [PubMed: 21098139]

24. Darty K, Denise A, Ponty Y. VARNA: Interactive drawing and editing of the RNA secondary structure. Bioinformatics. 2009; 25:1974–1975. [PubMed: 19398448]

25. Klopocki AG, Yago T, Mehta P, Yang J, Wu T, Leppanen A, Bovin NV, Cummings RD, Zhu C, McEver RP. Replacing a lectin domain residue in L-selectin enhances binding to P-selectin glycoprotein ligand−1 but not to 6-sulfo-sialyl Lewis x. J Biol Chem. 2008; 283:11493–11500. [PubMed: 18250165]

26. Kozak M. Circumstances and Mechanisms of Inhibition of Translation by Secondary Structure in Eukaryotic Messenger-Rnas. Mol Cell Biol. 1989; 9:5134–5142. [PubMed: 2601712]

27. Tholstrup J, Oddershede LB, Sorensen MA. mRNA pseudoknot structures can act as ribosomal roadblocks. Nucleic Acids Res. 2012; 40:303–313. [PubMed: 21908395]

28. Holland JA, Hansen MR, Du Z, Hoffman DW. An examination of coaxial stacking of helical stems in a pseudoknot motif: the gene 32 messenger RNA pseudoknot of bacteriophage T2. Rna. 1999; 5:257–271. [PubMed: 10024177]

29. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF chimera - A visualization system for exploratory research and analysis. J Comput Chem. 2004; 25:1605–1612. [PubMed: 15264254]

Highlights

- Optimized GS-based protein expression system for stable cell clone production

- Increased yield of the full ectodomain of a human selectin protein, CD62L

- Investigated effect of tertiary mRNA structures on protein expression levels

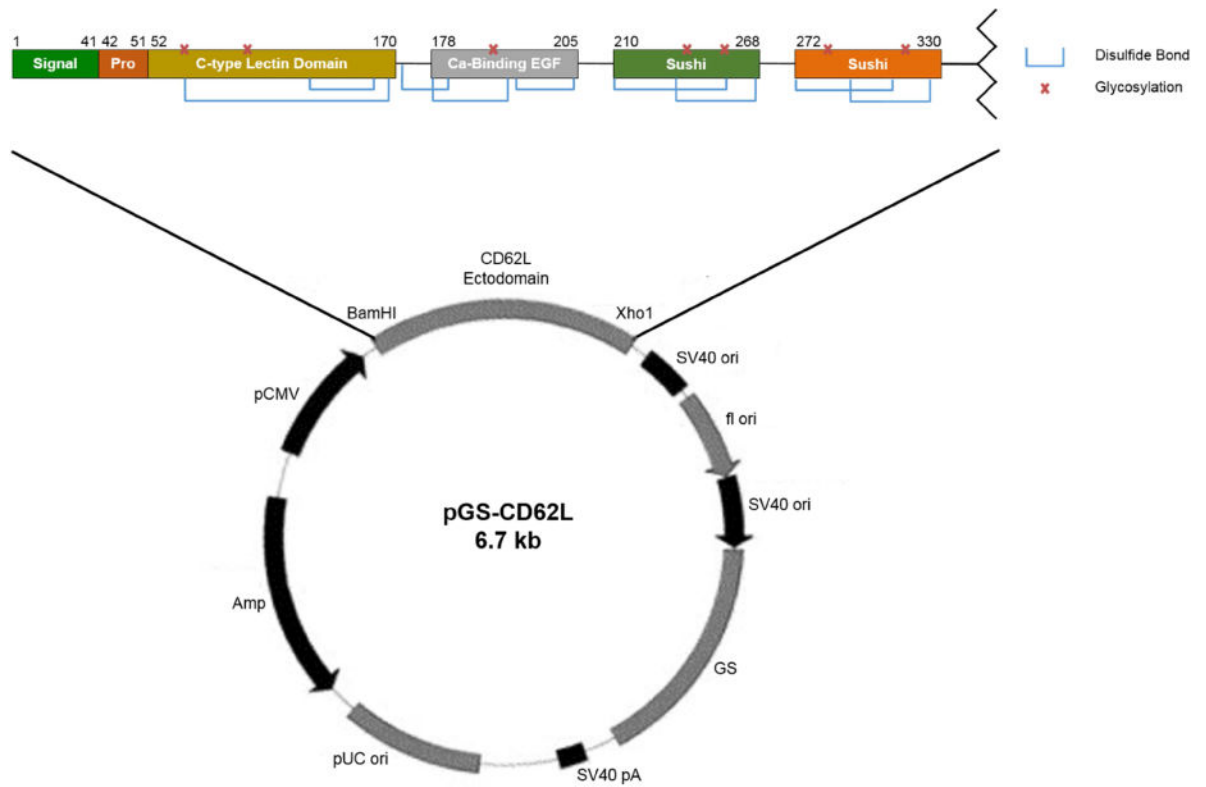- Pseudoknots involving start codon and signal sequences affect protein expression

Figure 1.
Plasmid map of pGS-CD62L and primary structure of CD62L ectodomain containing nine disulfide bonds and seven predicted glycosylation sites. The construct used contained all six domains; the signal and propeptide were cleaved post-translationally before purification.
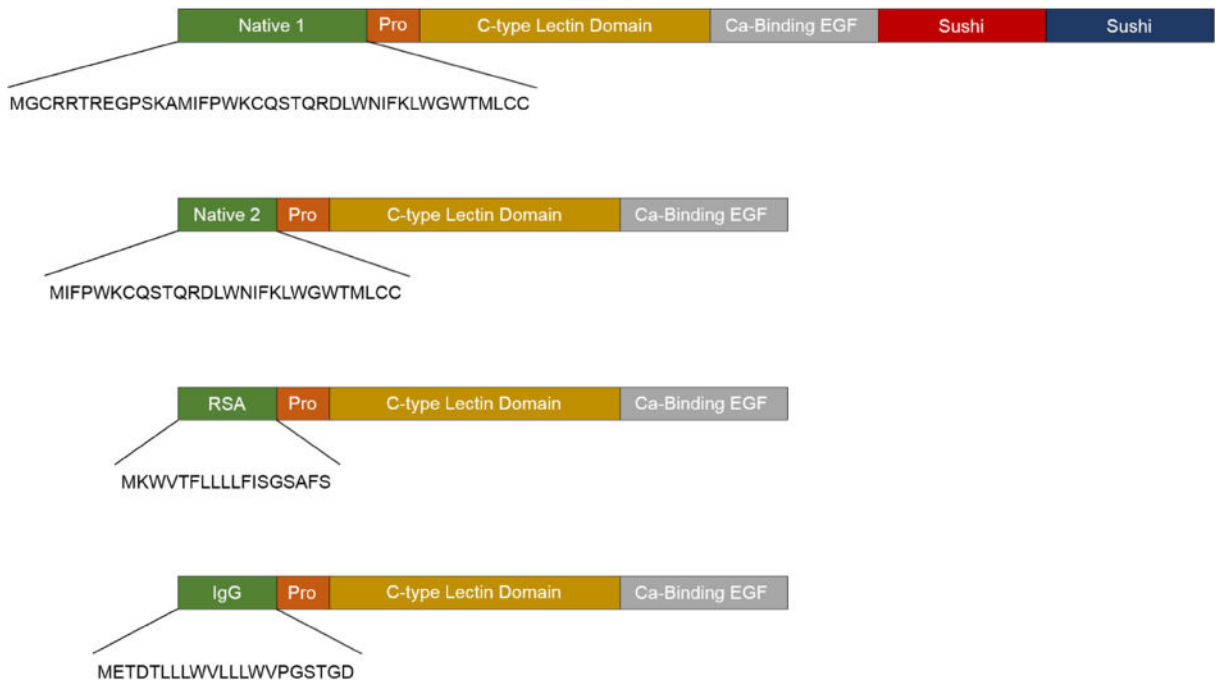
Figure 2.
The four recombinant CD62L expression constructs with differing signal sequences. All four are cloned into the pGS-CD62L vectors. Native 1 and native 2 expressed recombinant CD62L, while RSA and IgG did not.
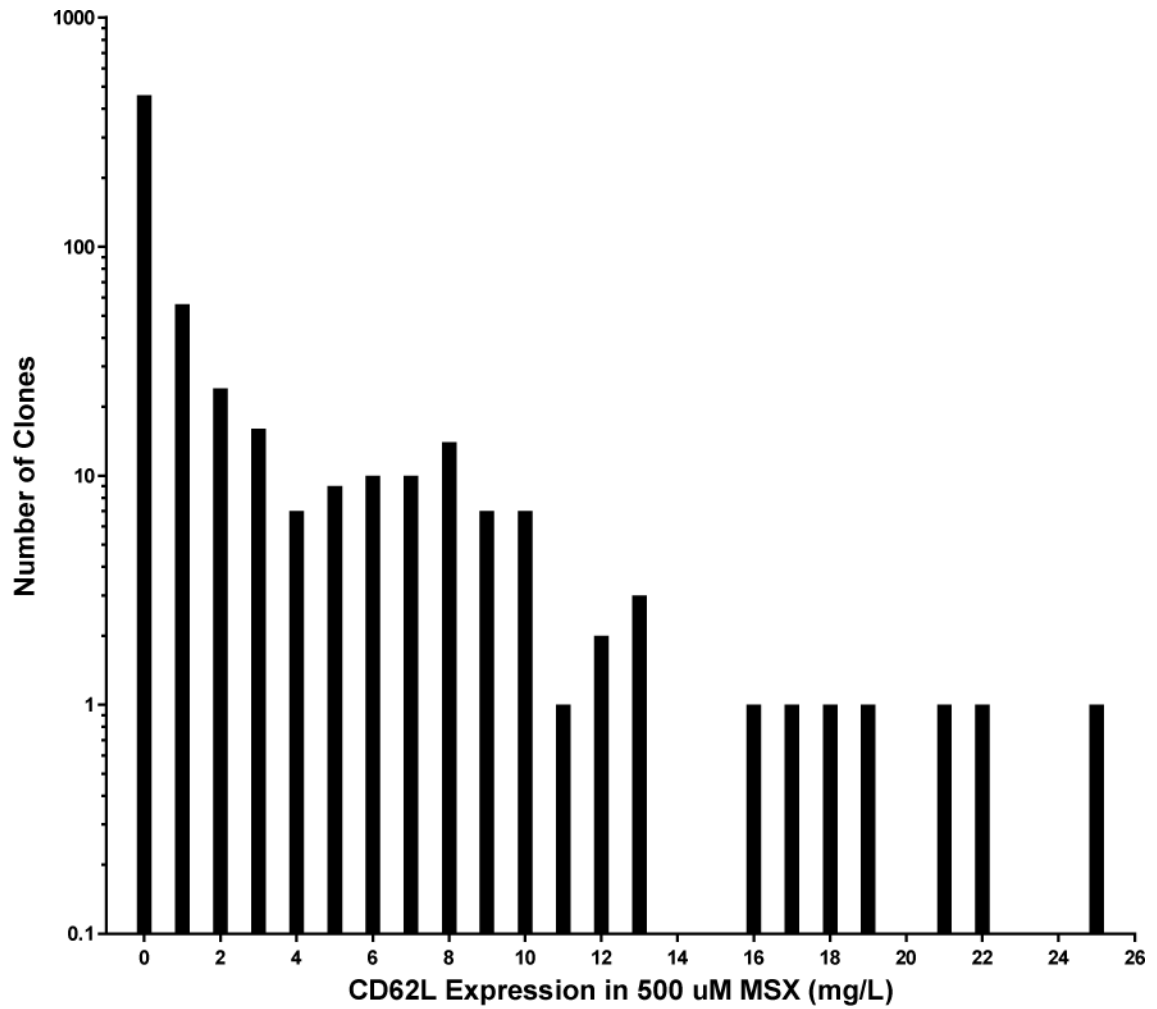
Figure 3.
Frequency distribution of CD62L-producing clones from single-round gene amplification, analyzed by ELISA.
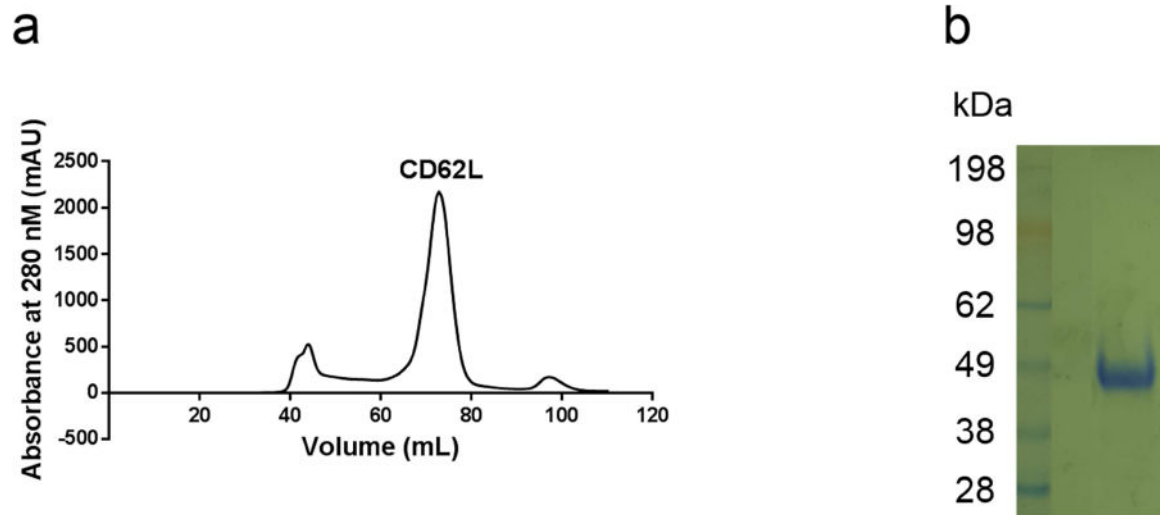
Figure 4.
Gel filtration of CD62L. (a) Size exclusion chromatography profile with CD62L peak labeled. (b) Coomassie blue-stained SDS-PAGE of fraction from gel filtration showing a band corresponding to recombinant CD62L.
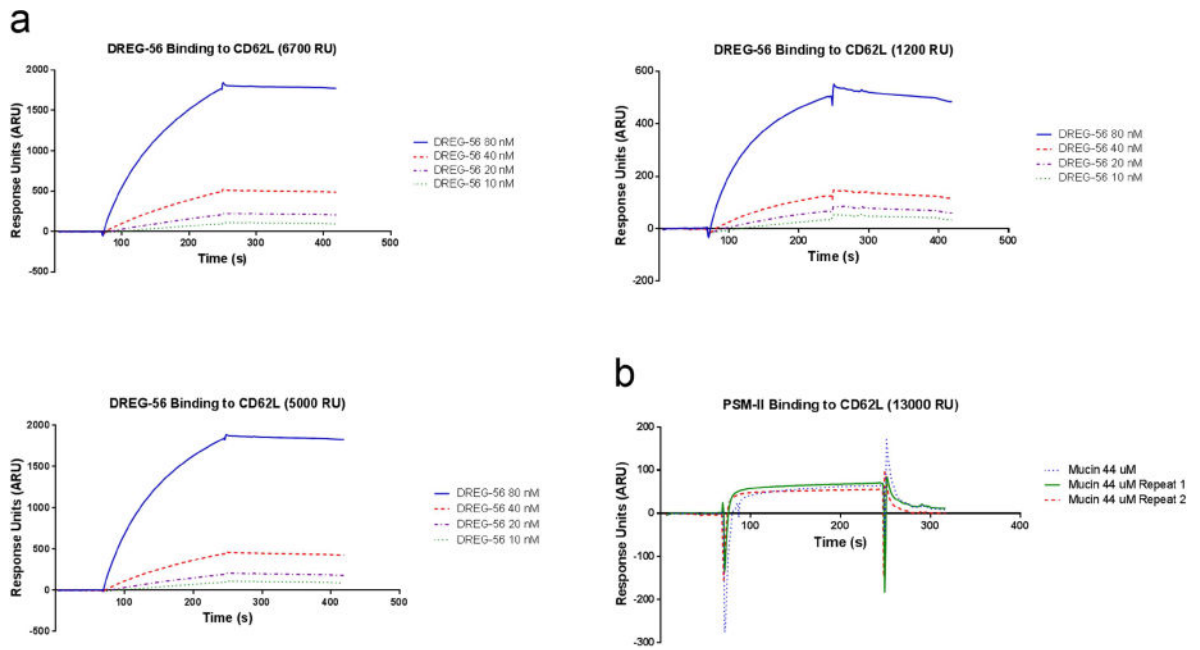
Figure 5.
Biding analysis on BIAcore showing CD62L binding to (a) an anti-CD62L antibody (DREG-56) and (b) PSM-II. BSA was used as a control to subtract background binding.
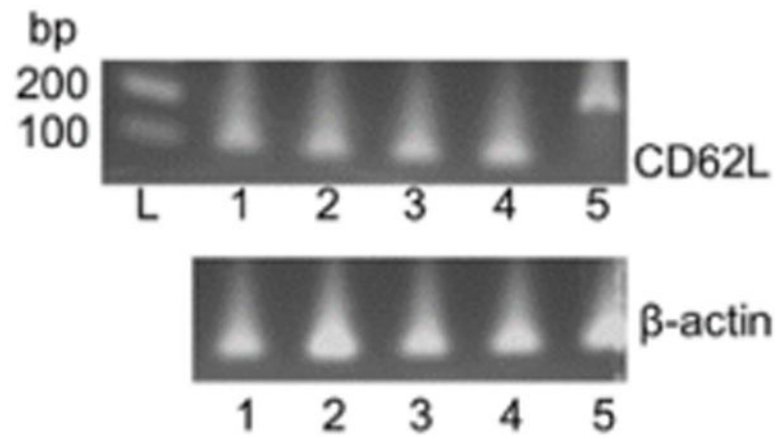
Figure 6.
Agarose gel electrophoresis showing the RT-PCR bands for CD62L (top) and β-actin (bottom). mRNA samples were prepared from either CD62L-transfected or non-transfected CHO-lec cells. The lanes are designated as L, ladder; 1, native 1; 2, native 2; 3, IgG; 4, RSA; and 5, untransfected.
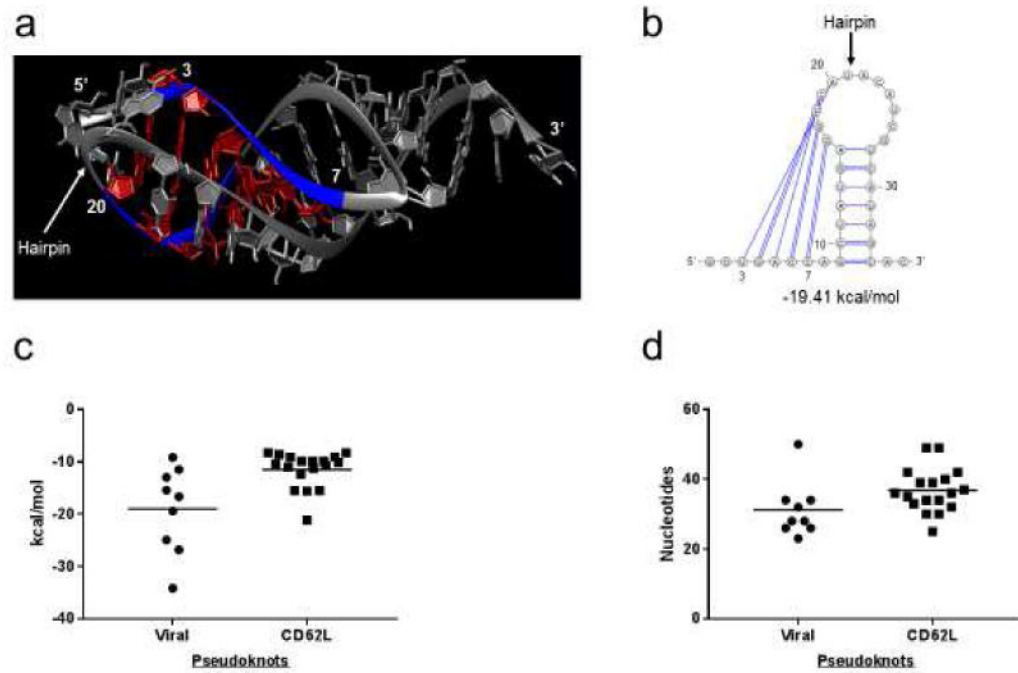
Figure 7.
Example of a known pseudoknot from bacteriophage T2 (PDB: 2TPK [28, 29]) shown in (a) a 3D structure derived from NMR with corresponding labeled nucleotides, and (b) a 2D representation with the calculated Gibbs free energy. Published sequences from viral pseudoknots were compared both in Gibbs free energy (c) and length (d) to the predicted pseudoknots in the first 200 nucleotides of all four recombinant CD62L constructs.
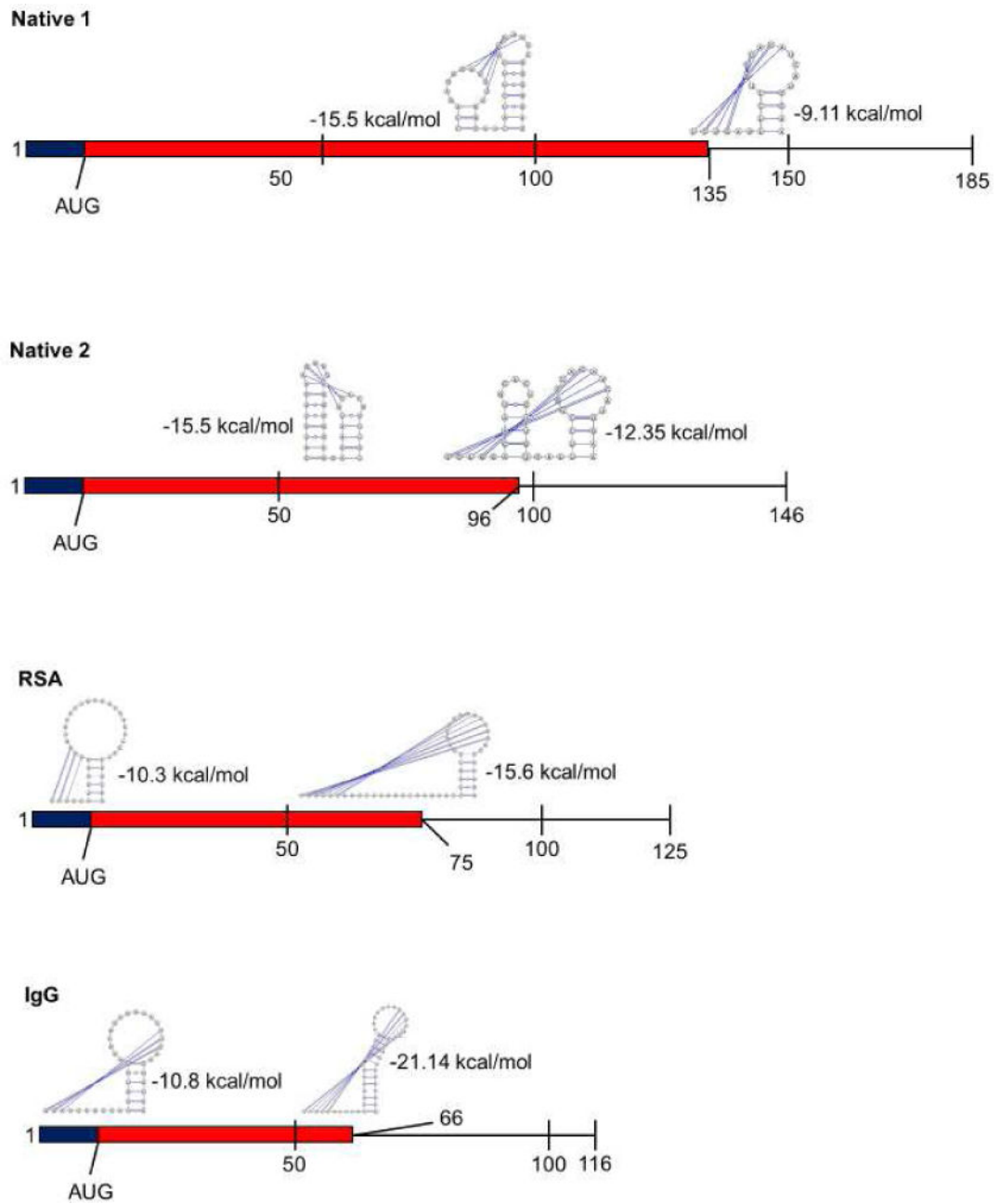
Figure 8.
Predicted most stable pseudoknots present in the four recombinant CD62L expression constructs. Kozak and start codons are within the first 15 nucleotides of each construct and are shown in blue. Pseudoknots were calculated including both the signal sequence and the first 50 nucleotides of CD62L structural region. Common pseudoknots present solely in CD62L structural regions are not included.