somalogic

Characterization of the Binding Specificity of SOMAmer® Reagents used in the SomaScan® Assay

"Slow Off-rate Modified Aptamer" (SOMAmer) reagents are identified via the SELEX process against epitopes on a single target protein in buffer. While this shape recognition is highly specific, some proteins share structural and/or functional properties, and thus could be bound by a SOMAmer reagent originally selected against another protein. In order to test the specificity of SOMAmer reagents used in the SomaScan Assay for their respective initial target proteins, we perform a variety of characterization steps. These steps include:

- In silico selection, procurement (when available), and direct SOMAmer reagent binding experiments in buffer with "relevant relative" proteins.
- "Pulldown" assays followed by mass spectrometry-based and SDS gel-based analysis of the protein(s) bound by the SOMAmer reagent from biological matrices.

All affinity reagents (antibodies, traditional aptamers, etc.) are subject to specificity issues. Recognizing how critical the accuracy of the SomaScan Assay is for both research and clinical purposes, SomaLogic is committed to regular assessment of SOMAmer specificity, and to transparency in communicating the results of those ongoing efforts.

Introduction

The key to reliable protein measurement using affinity reagents of any kind is the continuous evaluation of the specificity of those reagents for their target proteins.

Protein contaminants in the target preparation (even at very low levels), cross-reactivity due to protein homologies, protein preparation artifacts, etc. are all known issues that can compromise the utility of any protein-binding reagent. For example, it has been estimated that ~50% of commercially available antibodies are either incorrect or insufficiently specific for their putative target protein.[1]

"Slow Off-rate Modified Aptamer" (SOMAmer) reagents – the protein-specific binding reagents at the heart of the SomaScan Assay – are initially identified via the SELEX process by their ability to strongly bind to a single purified protein.[2-4] Each SOMAmer reagent selected for the SomaScan Assay has a unique three-dimensional folded structure that binds a specific structural epitope on the surface of its target protein with high affinity. This high-affinity binding between a SOMAmer reagent and its target protein is mediated both by electrostatic contacts contributed by the DNA backbone (similar to unmodified aptamers) as well as by hydrophobic contacts enabled by the modified bases in the SOMAmer reagent sequence.[5,6]

Because many proteins share structural and functional features, it is possible that the structural epitope to which a SOMAmer reagent binds is present on proteins other than the one initially used to select the SOMAmer reagent. Indeed, we have observed that a minority of SOMAmer reagents are able to bind with some degree of affinity to highly similar proteins, presumably through such a shared structural epitope, although not always with the same high affinity. Because the SomaScan Assay is performed in a complex biological sample containing thousands of different proteins, experimentally determining which reagents may also target other proteins to some degree can be extremely valuable in interpreting biomarker discovery data derived from the assay.

Methods

We first analyzed publicly available databases of known human protein sequences[7] using sequence alignment tools (e.g., BLAST)[8] to identify those "relevant relative" proteins that share significant homology with proteins used to select the SOMAmer reagents. Proteins with significant homology to the SOMAmer target protein (i.e., proteins with greater than 40% amino acid sequence identity with the target protein) were tested experimentally, if available in the SomaLogic inventory or commercially available as folded proteins from reliable vendors.

Available related proteins were analyzed with affinity capture experiments similar to immunoprecipitation protocols. SOMAmer reagents were immobilized on streptavidin coated beads and then incubated with either the target protein or the identified related protein. The SOMAmer-protein complexes were then washed, and the proteins labeled with a fluorophore.

The complexes were then eluted and the recovery of bound protein vs. input protein was analyzed by SDS-PAGE and fluorescent imaging. When SOMAmer reagent binding to proteins other than the SELEX target was observed, we performed solution affinity measurements to determine whether the SOMAmer reagent has similar or different affinities for the target protein and related protein. If the solution K_d was within 10-fold of that for the SELEX target, the reagent was reported to bind the SELEX target and other proteins with "similar affinity." If the measured affinity differed by greater than 10-fold, we reported that the reagent binds to the protein(s) other than the SELEX target with "at least 10-fold weaker affinity."

Although this is a broad statement regarding specific affinity, we do not report exact K_d values because of the high variability observed in both the quality and the reported concentrations of commercially obtained purified proteins.

From this (and related) work we have identified a set of SOMAmer reagents that have been flagged for further performance review. For example, the original

SELEX target protein may no longer be commercially available, or pulldown results were ambiguous due to impurities (even small ones) in the original protein sample. SomaLogic is committed to providing specificity characterization results for these SOMAmer reagents as soon as unambiguous target binding data are available and reviewed.

Another method that we use to test SOMAmer specificity is to identify, by liquid chromatography and tandem mass spectrometry (LC-MS/MS), the proteins bound by each reagent after incubation with human plasma. We have begun employing this mode of specificity characterization on SOMAmer reagents selected to bind to proteins with high endogenous concentrations in human plasma.

Results

Highly specific target binding

We were able to identify and acquire closely related human proteins for 1612 out of 2362 SELEX targets, which we used to assess the specificity of those reagents. When tested, 73% (or 1172) demonstrate binding of the SOMAmer reagent to the specific SELEX target only. For example, a SOMAmer reagent selected to bind the protein tissue inhibitor of metalloproteinase-1 (TIMP-1), was also tested against the related proteins TIMP-2 (60% identical), TIMP-3 (31% identical), and TIMP-4 (40% identical) (Figure 1A).

When this same TIMP-1 SOMAmer reagent was used in affinity enrichment from human plasma, four unique peptides corresponding to endogenous TIMP-1 were identified by LC-MS/MS in the enriched sample and no peptides corresponding to any other member of

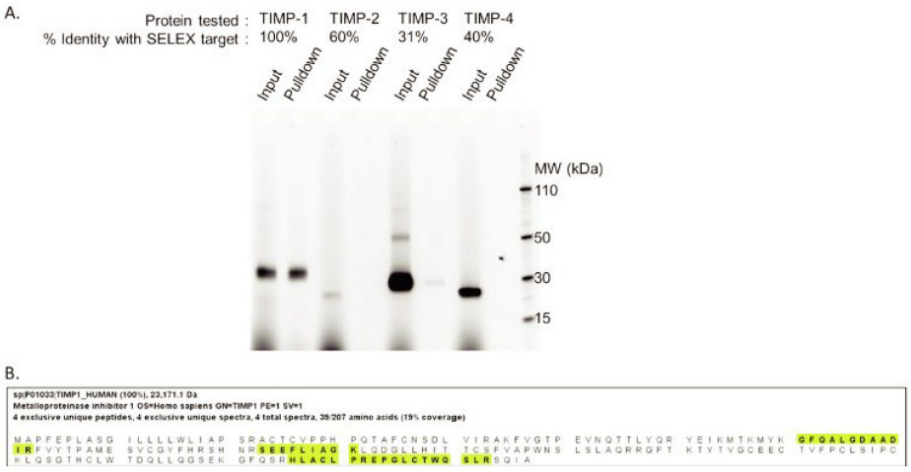

FIGURE 1 High specificity of a TIMP-1 SOMAmer reagent. A. Binding to purified TIMP-1, but not to TIMP-2, TIMP-3, or TIMP-4 in a pulldown experiment. Pulldown eluate is shown in the lanes labeled "Pulldown" and a sample of protein input is shown in lanes labeled "Input." B. TIMP-1-specific peptide sequences enriched from human plasma using a TIMP-1 SOMAmer reagent. No binding was observed with TIMP-2 or TIMP-4, and very weak binding was observed with TIMP-3 that could not be measured by solution-affinity methods (at protein concentrations up to 100 nM).

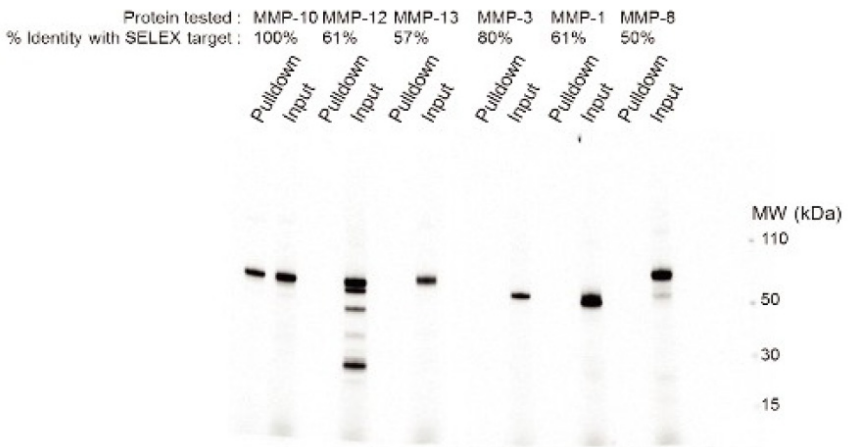

FIGURE 2 Exclusive binding of an MMP-10 specific SOMAmer reagent to MMP-10 but not to MMP-12, MMP-13, MMP-3, MMP-1, or MMP-8. Pulldown eluate is shown in the lanes labeled "Pulldown" and a sample of protein input is shown in lanes labeled "Input."

the TIMP protein family were identified (Figure 1B). Additionally, no peptides corresponding to TIMP-1 were identified in any other plasma pulldown samples performed using 142 different SOMAmer reagents, including a TIMP-2-specific SOMAmer reagent.

Another example of highly specific binding is shown in (Figure 2). A SOMAmer reagent specific for matrix metalloproteinase-10 (MMP-10) does not bind MMP-12 (61% identical), MMP-13 (57% identical), MMP-3 (80% identical), MMP-1 (61% identical), or MMP-8 (50% identical).

Binding to shared epitopes with similar affinity

When we did observe binding to proteins other than the SELEX target (27% or 440 of the reagents tested) in initial pulldown tests, we followed up with measurements of solution affinity.

We typically measure the association of radiolabeled SOMAmer reagent with protein and then capture the complex using a protein-affinity chromatography medium. Saturation binding curves are then generated by titrating increasing amounts of protein in the presence of a constant, limiting amount of SOMAmer reagent.

The K_d is determined to be the protein concentration at which the half maximal binding is observed. For example, initial pulldown tests indicated that one SOMAmer reagent binds not only to its original SELEX target (pyrophosphatase 1 (PPA1)), but also the related protein PPA2, which shares 68% amino acid sequence identity (Figure 3A). However, solution affinity measurements determined that the SOMAmer reagent affinity was greater than 10-fold stronger for PPA1 than for PPA2 (Figure 3B).

Binding to shared epitopes with similar affinity

We observed that 13% (213) of the reagents tested bound to members of a protein family with similar affinities. As previously noted, this recognition most often occurs when proteins share extensive sequence identity. Presumably, the structural epitope to which the SOMAmer reagent was selected is highly conserved and biochemically indistinguishable by solution equilibrium binding affinities.

In fact, of the reagents that could bind a related target, ~6% were products of the same gene with a common epitope (e.g., splice variants like vascular endothelial growth factor (VEGF) 121 and 165 isoforms) or shared subunits in a multi-subunit complex (e.g., cyclin-dependent kinase 1/cyclin B1 complex, in which the SOMAmer reagent binds to the cyclin B1 subunit).

The remaining ~7% appear to bind to epitopes shared amongst highly related families of proteins. For example, a SOMAmer reagent that binds to its SELEX target, calcium/calmodulin-dependent protein kinase II delta (CAMK2D), also binds the closely related proteins CAMK2A (91% identical) and CAMK2B (87% identical) (Figure 4A). Solution affinity comparisons determined that this reagent has a similar binding affinity, of approximately 2 nM, for all three proteins (Figure 4B).

As expected, the amino acid sequence identity tended to be greater for those pairs that exhibited cross-reactivity: 48% mean for pairs that exhibited no cross-reactivity (no positive pulldown results), 62% for pairs with >10-fold lower affinity, but positive pulldown results, and 70% for pairs with similar affinity.[9] In cases where a SomaScan Assay signal results from a reagent with documented binding to multiple proteins, the SomaScan Assay end user may use this information in follow-on studies to further dissect the biological pathways being affected.

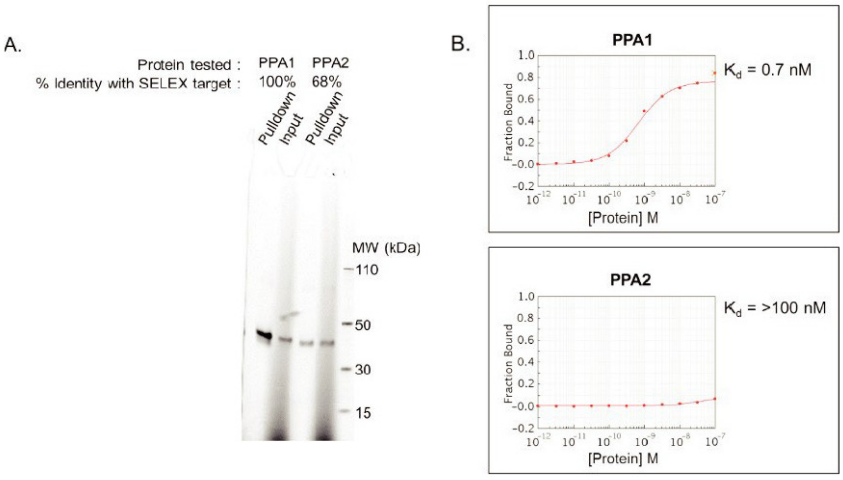

FIGURE 3 The PPA1-specific SOMAmer reagent binds to PPA1 with greater than 10-fold stronger affinity than PPA2. A. The PPA1 SOMAmer reagent binds to purified PPA1 and PPA2 in a pulldown assay. Pulldown eluate is shown in the lanes labeled "Pulldown" and a sample of protein input is shown in lanes labeled "Input". B. PPA1 SOMAmer reagent binds to PPA1 with greater than 10x stronger affinity (K_d = 0.7 nM) than PPA2 (K_d = > 100 nM). Protein concentration is shown in units of molarity (M) on the x-axis and the fraction of reagent bound, as measured by radioactivity, is shown on the y-axis.

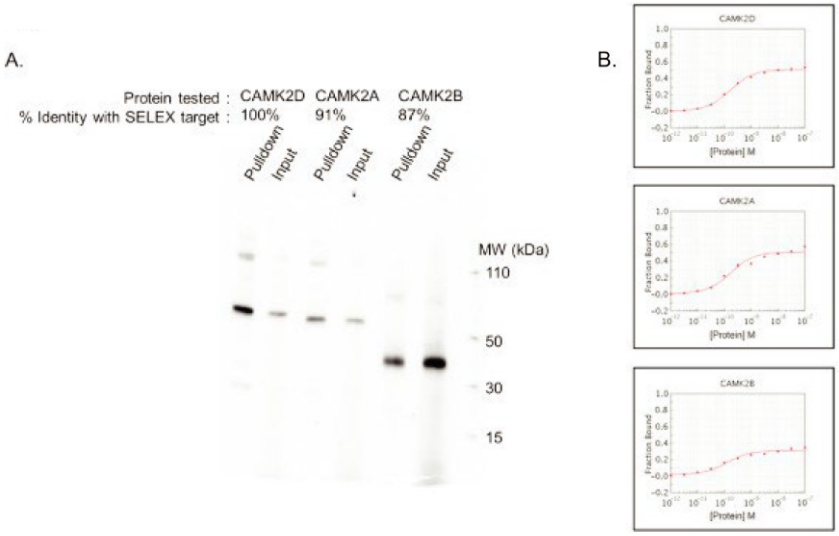

FIGURE 4 The SOMAmer reagent selected against CAMK2D binds to CAMK2A and CAMK2B with similar affinities. A. CAMK2D SOMAmer reagent binds to purified CAMK2D, CAMK2A, and CAMK2B in a pulldown assay. Pulldown eluate is shown in the lanes labeled "Pulldown" and a sample of protein input is shown in lanes labeled "Input." B. CAMK2D SOMAmer reagent binds to CAMK2D, CAMK2A, and CAMK2B with a similar affinity (K_d = 2 nM). Protein concentration is shown in units of molarity (M) on the x-axis and the fraction of reagent bound, as measured by radioactivity, is shown on the y-axis.

For example, although the analyte name (which was derived from the name provided for the original SELEX target protein) is listed as CAMK2D, it is possible that the SomaScan Assay signal is a result of that SOMAmer reagent binding to these three proteins separately or cumulatively, depending on their relative concentrations in the sample.

Based on these results, we have begun to perform our initial SELEX experiments with competition strategies to further ensure that we select SOMAmer reagents that can distinguish between such highly related family members for use in future SomaScan Assay versions. One such strategy was utilized to develop reagents that could distinguish between the 90% identical proteins growth differentiation factor (GDF)-11 and GDF-8. The first SOMAmer reagent selected against GDF-11 binds both proteins with comparable affinities (K_d~ 100 pM).

With such high amino acid sequence identity, this cross-reactivity was not surprising. Thus, to identify specific reagents for GDF-11 and GDF-8, special selection conditions were employed in order to isolate reagents that bind unique epitopes on the two proteins.

The results include several reagents that bind GDF-11 with very high affinity and with minimal, if any, binding to GDF-8. An example of one of the reagents is shown in Figure 5. Binding to GDF-11 is shown in red and binding to GDF-8 is shown in green. Using similar

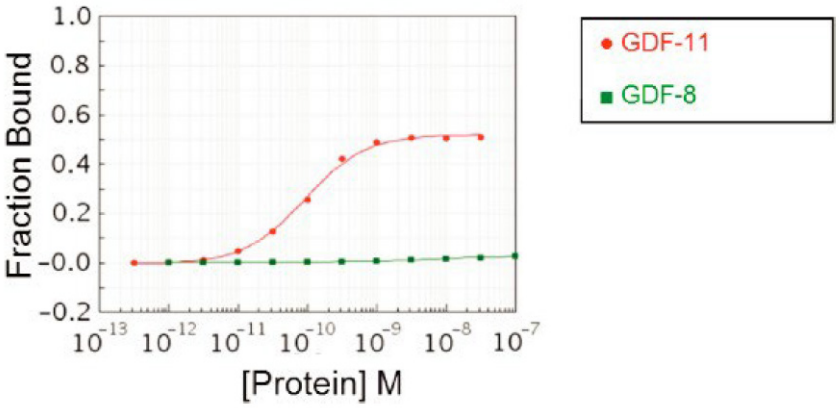

FIGURE 5 Affinity binding curves demonstrating binding of a GDF-11 specific SOMAmer reagent to GDF-11 (red) and to GDF-8 (green). Protein concentration is shown in units of molarity (M) on the x-axis and the fraction of reagent bound, as measured by radioactivity, is shown on the y-axis.[10]

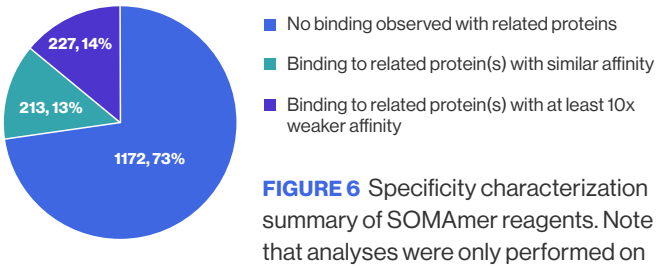

FIGURE 6 Specificity characterization summary of SOMAmer reagents. Note that analyses were only performed on reagents for which related proteins were available for testing.

selection-counter selection strategies, we have also identified GDF-8 specific SOMAmer reagents. In summary, we have tested binding to related proteins for 1612 SOMAmer reagents to date. We were unable to detect binding to any related proteins for 73% of those tested (Figure 6).

When binding to related proteins was detected, about half of these SOMAmer reagents exhibited binding to at least one related protein with similar affinity while the other half bound to related proteins, but with at least 10-fold weaker affinity. Specific target enrichment by pulldowns from human plasma has been confirmed for 123 of the SOMAmer reagents.

Additional ongoing characterization work

We have recently added to our in-house protein mass spectrometry capabilities, and current characterization efforts for each reagent have been expanded to include assessing the target accuracy using buffer pulldown assays.

The SOMAmer reagents are used to capture the recombinant protein employed during SELEX followed by LC-MS analysis of the captured protein ligand. This “bottom-up” approach has thus far improved our platform by identifying a small number of reagents that were selected to bind proteins that were either mislabeled by the provider or were present as a contaminant in the protein preparation.

Conclusion

In our ongoing commitment to enhancing the high quality of our proprietary SomaScan Assay, we continue to characterize the binding of the SOMAmer reagents we have selected for use in the assay.

We have performed protein sequence homology searches, followed by experimentally comparing the ability of SOMAmer reagents to bind to their cognate versus related protein targets in buffer for 1612 reagents to date.

SomaScan Assay end users may use these data to determine which, if any, proteins other than the indicated SELEX target could result in a positive signal in the SomaScan Assay. We are also now querying SOMAmer reagent specificity in a complex biological sample using tandem mass spectrometry, resulting in a richer set of characterization criteria for all reagents used in the SomaScan Assay.

Many of the analytes detected in the SomaScan Assay are expected to be at relatively low concentrations in human plasma, thus making it challenging to purify adequate amounts of these low abundance proteins for peptide identification by LC-MS/MS.

Through ongoing method development in affinity-capture as well as exploration of other biological matrices, such as cell lysate, that may have higher endogenous concentrations of certain proteins relative to plasma, we plan to expand our mass spectrometry-based characterization of SOMAmer reagents to cover a greater portion of the SomaScan menu.

To date, our efforts have kept the number of reagents with uncertain binding profiles to a very low percentage in the SomaScan Assay and helped us further improve each subsequent version of the assay.

In addition to our commitment to the quality of our assay, we are also committed to transparency about the specificity of the reagents in the assay. Therefore, specificity characterization results for SOMAmer reagents will be provided regularly as they are obtained and analyzed from our continuous evaluation efforts. The analyses we have performed and will perform continue to give us high confidence in the overall reliability of protein information generated by the SomaScan Assay.

References

1. Berglund, L. *et al*. A Genecentric Human Protein Atlas for Expression Profiles Based on Antibodies. *Mol. Cell. Proteomics* 7 (10) 2019-2027 (2008).
2. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249, 505–10 (1990).
3. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* 346, 818–22 (1990).
4. Ozer, A., Pagano, J. M. & Lis, J. T. New Technologies Provide Quantum Changes in the Scale, Speed, and Success of SELEX Methods and Aptamer Characterization. *Mol. Ther. Nucleic Acids* 3, e183 (2014).
5. Gelinas, A. D., Davies, D. R. & Janjic, N. Embracing proteins: structural themes in aptamer-protein complexes. *Curr. Opin. Struct. Biol*. 36, 122–32 (2016).
6. Jarvis, T. C. *et al*. Non-helical DNA Triplex Forms a Unique Aptamer Scaffold for High Affinity Recognition of Nerve Growth Factor. *Structure* 23, 1293–304 (2015).
7. UniProt Consortium, T. U. UniProt: a hub for protein information. *Nucleic Acids Res*. 43, D204–12 (2015).
8. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol*. 215, 403–10 (1990).
9. Sun, B. B. *et al*. Genomic atlas of the human plasma proteome. *Nature* 558, 73-9 (2018).
10. Ochsner, U. A. et al. "Targeting Unique Epitopes on Highly Similar Proteins GDF-11 and GDF-8 with Modified DNA Aptamers." Biochemistry 58, 4632-4640 (2019).

somalogic

SL00000430 Rev 3: 2024-02

Characterization of the Binding Specificity of SOMAmer® Reagents used in the SomaScan® Assay