Technical Note
Characterization of SOMAmer® Reagents Binding Specificity in the SOMAscan® 1.3k Assay

Summary

“Slow Off-Rate Modified Aptamer” (SOMAmer) reagents are identified via the SELEX process against epitopes on a single target protein in buffer. However, many 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 each SOMAmer reagent used in the SOMAscan 1.3k assay for its initial target protein, we performed a series of characterization steps. These steps included selection of similar proteins via standard \textit{in silico} methods, obtaining any identified “relevant relative” proteins (where available), direct experimental SOMAmer binding experiments with those related proteins in buffer, and/or “pull-down” assays followed by mass spectrometry-based analysis of the protein(s) bound by the SOMAmer reagent from biological matrices. In the current SOMAscan assay 51% of the SOMAmer reagents bind only their original target, and another 9% bind related proteins with >10x weaker affinity than their target protein (24% of the SOMAmer reagents in the menu cannot be tested, as no highly related proteins are available for testing) (See Figure 6).

This Technical Note describes these characterization studies in more detail.

Introduction

“Slow Off-Rate Modified Aptamer” (SOMAmer) reagents - the protein-specific binding reagents that enable the SOMAscan assay – are initially identified via the SELEX process by their ability to strongly bind to a single purified protein.

Each modified DNA-based SOMAmer 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 hydrophobic contacts enabled by the modified bases in the SOMAmer sequence.

Because there are many proteins that share structural and functional features, it is possible that the conformational epitope to which a SOMAmer reagent binds is present on other proteins related to the original protein 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 target a shared epitope within a protein species may be extremely valuable in interpreting biomarker discovery data derived from the SOMAscan assay.
As a first step, we used publicly available databases of known human protein sequences\(^6\) and sequence alignment tools (e.g., BLAST)\(^7\) to identify those “relevant relative” proteins that share significant homology with proteins used to select the SOMAmer reagents in the current version of the SOMAscan assay (SOMAscan 1.3k). Any proteins with significant homology to the SOMAmer target protein (i.e., proteins with greater than 40% sequence identity with the target protein) were obtained for direct experimental testing. It should be noted that the breadth of related proteins that can be tested is limited to those that are available in the SomaLogic inventory or commercially available as full-length proteins from reliable vendors. In cases where a protein was not available or where a sequence homology search returned no significant results (~20% of the 1.3k menu), we confirmed SOMAmer reagent pull-down of the target protein in buffer but performed no additional testing.

If related proteins were available, affinity capture experiments similar to immunoprecipitation were performed. 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 protein labeled by conjugation of NHS-AlexaFluor\(^{TM}\) 647. The complexes were then eluted and the recovery of bound protein vs. input protein was analyzed by SDS-PAGE and fluorescent imaging. When any binding to proteins other than the SELEX target was observed (at 100 nM in buffer), we performed solution affinity measurements to determine whether the SOMAmer reagent has similar or disparate 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 have chosen not to report exact $K_d$ values because of the high variability observed in both the quality and the reported concentrations of commercially obtained purified proteins.

We have also identified a small number of SOMAmer reagents (76 out of 1305) 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 protein sample impurity. Specificity characterization results for these SOMAmer reagents will be communicated as soon as unambiguous target binding data are available and reviewed.

Another method that we have used to test SOMAmer reagent 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. We have noted specific enrichment of the intended endogenous target protein in the SOMAscan menu annotation separately from the specificity results generated using purified proteins, when observed.
Results

Highly specific target binding

For 73% of cases in which proteins related to the SELEX target were available for testing, we observed binding of the SOMAmer reagent to the specific SELEX target and not to any of the related proteins. For example, SOMAmer 2211-9 was tested against the intended target, tissue inhibitor of metalloproteinase-1 (TIMP-1), and the related proteins TIMP-2 (60% identical), TIMP-3 (31% identical), and TIMP-4 (40% identical) (Figure 1A).

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). 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 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, in which 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).
Figure 2. Exclusive binding of a 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".

Binding to shared epitopes with disparate affinity

In any instance where we observed binding to proteins other than the SELEX target (20% of the 1,305 reagents) 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 Kd is determined to be the protein concentration at which the half maximal binding is observed. For example, initial pulldown tests indicated that SOMAmer 5021-13 binds not only to the SELEX target pyrophosphatase 1 (PPA1), but also the related protein PPA2, which shares 68% amino acid sequence identity (Figure 3A). Solution affinity measurements determined that the SOMAmer affinity was greater than 10-fold stronger for PPA1 than for PPA2 (Figure 3B).
Figure 3. PPA1-specific SOMAmer reagent binds to PPA1 with greater than 10-fold stronger affinity than PPA2. A. 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 10× stronger affinity (K<sub>d</sub> = 0.7 nM) than PPA2 (K<sub>d</sub> = > 100 nM). Protein concentration is shown in M units on the x-axis and the fraction of reagent bound, as measured by radioactivity, is shown on the y-axis.

Binding to shared epitopes with similar affinity

We observed that 10% of the 1,305 reagents bound to members of a protein family with highly similar affinities. As previously noted, this recognition most often occurs when proteins share extensive sequence identity. Presumably, the binding epitope to which the SOMAmer reagent was selected is highly conserved and biochemically indistinguishable by solution equilibrium binding affinities. Because the reagents in the SOMAscan assay are typically selected only against a single target protein in the absence of protein competitor strategies, it is not surprising that certain reagents bind to epitopes shared amongst a highly related family of proteins. For example, SOMAmer 3419-49 binds to the SELEX target calcium/calmodulin dependent protein kinase II delta (CAMK2D) as well as the 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).
Figure 4. 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 M units on the x-axis and the fraction of reagent bound, as measured by radioactivity, is shown on the y-axis.

In cases where a SOMAscan signal results from a reagent with documented binding to multiple proteins, the SOMAscan end user may use this information in follow-on studies to further dissect the biological pathways being affected. For example, although the analyte name (which was derived from the name of the original SELEX target protein) is listed as CAMK2D, it is possible that the SOMAscan signal is a result of SOMAmer 3419-49 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 SELEX experiments with competition strategies to 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 the 90% identical proteins GDF-11 and GDF-8. The initial reagent selected against GDF-11 binds to both proteins with comparable affinities (K_d ~ 100 pM). With such high amino acid sequence identity, this cross-reactivity was not surprising. To identify specific reagents for GDF-11 and GDF-8, special selection conditions were employed in order to compete for sequence binding between GDF-11 and GDF-8 SOMAmer reagent libraries. 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 selection-counter selection strategies, we have also identified GDF-8 specific SOMAmer reagents.
In summary, we were able to test binding to related proteins for 76% of the 1,305 SOMAmer reagents on the current SOMAscan 1.3K menu. We were unable to detect binding to any related proteins for 51% of the menu (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 from human plasma has been confirmed for 115 of the SOMAmer reagents on the current menu. We were either unable to identify or obtain related proteins to test for shared epitope binding for 24% of the SOMAmer reagent menu.
Conclusions

In our ongoing commitment to improving the quality of the SOMAscan multiplexed proteomic assay, we have cataloged the specificity of all SOMAmer reagents used in the current SOMAscan 1.3K Assay by performing protein sequence homology searches, followed by experimentally comparing the ability of SOMAmer reagents to bind to their cognate versus related protein targets in buffer. 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 have also begun to query SOMAmer reagent specificity in a complex biological sample using tandem mass spectrometry. This has resulted in a richer set of characterization criteria for all reagents used in the SOMAscan 1.3K 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.

References


All trademarks, service marks, trade names and product names are the property of their respective owners, including SomaLogic® and SOMAmer®, which are registered trademarks of SomaLogic, Inc.