

Tracking the Stability of Clinically Relevant Blood Plasma Proteins with Delta-S-Cys-Albumin—A Dilute-and-Shoot LC/MS-Based Marker of Specimen Exposure to Thawed Conditions

Authors

Erandi P. Kapuruge, Nilojan Jehanathan, Stephen P. Rogers, Stacy Williams, Yunro Chung, and Chad R. Borges

Correspondence

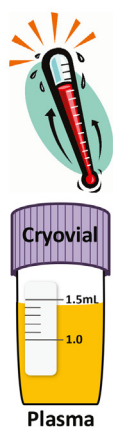
chad.borges@asu.edu

Graphical Abstract

In Brief

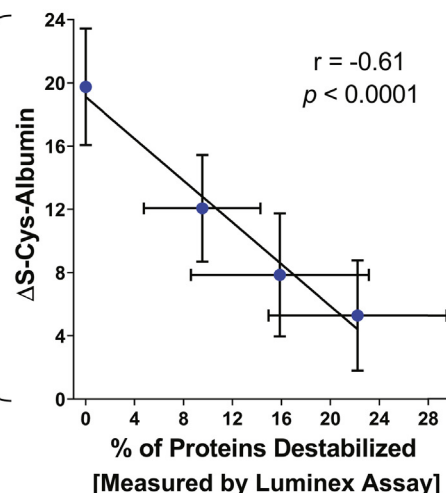
Δ S-Cys-Albumin is an endogenous marker of plasma and serum exposure to thawed conditions ($> -30^{\circ}\text{C}$). Baseline values of Δ S-Cys-Albumin in gastrointestinal cancer patients and cancer-free donors were determined. Time courses for 12 patients from each cohort were acquired at 23°C , 4°C , and -20°C . Δ S-Cys-Albumin decay kinetics were then empirically linked to the stability of 21 proteins measured *via* Luminex assay. Δ S-Cys-Albumin was linearly correlated to the percentage of destabilized plasma proteins, regardless of exposure temperature.

Plasma Protein Destabilization as the Δ S-Cys-Albumin Thawed-State Exposure Marker Decreases



Low
 Δ S-Cys-Albumin
Indicates
Exposure
to $> -30^{\circ}\text{C}$

Combined Data from 23°C , 4°C ,
and -20°C Exposures; $n = 240$ points



Highlights

- Baseline Δ S-Cys-Albumin in GI cancer patients and control donors determined.
- Twenty-four time courses for matched plasma/serum samples at 23°C , 4°C , and -20°C .
- Δ S-Cys-Albumin linked to protein stability independent of temperature.
- Minimal differences existed between GI cancer patients and cancer-free donors.

Tracking the Stability of Clinically Relevant Blood Plasma Proteins with Delta-S-Cys-Albumin—A Dilute-and-Shoot LC/MS-Based Marker of Specimen Exposure to Thawed Conditions

Erandi P. Kapuruge^{1,2,†}, Nilojan Jehanathan^{1,2,†}, Stephen P. Rogers², Stacy Williams², Yunro Chung^{2,3}, and Chad R. Borges^{1,2,*}

Biomolecular integrity can be compromised when blood plasma/serum (P/S) specimens are improperly handled. Compromised analytes can subsequently produce erroneous results—without any indication of having done so. We recently introduced an LC/MS-based marker of P/S exposure to thawed conditions called Δ S-Cys-Albumin which, aided by an established rate law, quantitatively tracks exposure of P/S to temperatures greater than their freezing point of -30°C . The purposes of this study were to (1) evaluate Δ S-Cys-Albumin baseline values in gastrointestinal cancer patients and cancer-free control donors, (2) empirically assess the kinetic profiles of Δ S-Cys-Albumin at 23°C , 4°C , and -20°C , and (3) empirically link Δ S-Cys-Albumin to the stability of clinically relevant proteins. Δ S-Cys-Albumin was measured at ≥ 9 different time points per exposure temperature in serum and K_2EDTA plasma samples from 24 separate donors in aliquots kept separately at 23°C , 4°C , and -20°C . Twenty-one clinically relevant plasma proteins were measured at four time points per temperature *via* a multiplexed immunoassay on the Luminex platform. Protein stability was assessed by mixed effects models. Coordinated shifts in stability between Δ S-Cys-Albumin and the unstable proteins were documented by repeated measures and Pearson correlations. Plasma Δ S-Cys-Albumin dropped from approximately 20% to under 5% within 96 h at 23°C , 28 days at 4°C , and 65 days at -20°C . On average, 22% of the 21 proteins significantly changed in apparent concentration at each exposure temperature ($p < 0.0008$ with $>10\%$ shift). A linear inverse relationship was found between the percentage of proteins destabilized and Δ S-Cys-Albumin ($r = -0.61$; $p < 0.0001$)—regardless of the specific time/temperature of exposure. Δ S-Cys-Albumin tracks cumulative thawed-state

exposure. These results now enable Δ S-Cys-Albumin to approximate the percentage of clinically relevant proteins that have been compromised by incidental plasma exposure to thawed-state conditions.

As the acellular component of blood, plasma/serum (P/S) exchanges biological information in the form of biomolecules with every organ system in the body. Because it carries this wealth of information and can be noninvasively collected, P/S is one of the most common biospecimens employed in biomedical research. Moreover, P/S samples collected during clinical studies frequently carry potential research value well beyond that intended by the original research design and are therefore archived for future research purposes. Millions of P/S samples are currently stored in biobanks around the world awaiting withdrawal and subsequent queries on the molecular information they contain. The answers to these queries, however, reflect both *in vivo* and *ex vivo* biochemistry—meaning that if the latter has impacted the former, incorrect and potentially misleading information will be obtained. This can readily occur when pre-analytical variables (PAVs) are inadequately controlled (1–8).

During sample collection, processing, transport, and storage, every biospecimen is exposed to PAVs. While the number of PAVs can be quite large (7–13), proper sample handling techniques can keep them tightly controlled, protecting *in vivo* biochemistry from *ex vivo* modulation. Many, if not most PAVs, present themselves during collection and initial processing, providing a single window of opportunity to either handle them correctly or not. For example, P/S PAVs such as the type of blood collection tube, proper collection tube filling,

From the ¹School of Molecular Sciences, Arizona State University, Tempe, Arizona, USA; ²The Biodesign Institute at Arizona State University, Tempe, Arizona, USA; ³College of Health Solutions, Arizona State University, Phoenix, Arizona, USA

[†]These authors contributed equally to this work.

*For correspondence: Chad R. Borges, chad.borges@asu.edu.

number of collection tube inversions, pre-centrifugation delay, and post-centrifugation delay are all completed in a single pass and the record of these PAVs will not ever have an opportunity to change during the life of the P/S specimen. Other PAVs, however, are constantly subject to change.

Of these, the PAV that is arguably the most difficult to control and track over the life of an archived specimen is exposure to thawed conditions. P/S does not fully freeze until the temperature is -30°C or colder (14, 15). When thawed, a wide range of *ex vivo* biochemical reactions can take place, distorting the portrait of *in vivo* biochemistry that the samples are supposed to reflect. While many if not most researchers who handle P/S know that it should be stored at -80°C or colder, the facts that not all labs possess a -80°C freezer, -20°C freezers are common, P/S appears (visually) to be frozen at -20°C , and many clinical research protocols allow at least temporary storage of newly collected specimens to be stored at -20°C for days to weeks before transfer to a -80°C freezer means that many P/S samples become compromised by temporary storage at -20°C . Moreover, even after placement in a -80°C freezer, an ever-present risk of exposure to thawed conditions remains in place for the life of every aliquot of P/S. For example, -80°C freezer failures occur, delays during sample shipments may result in the loss of dry ice and induce thawing, and there may be need(s) to re-aliquot samples. Thus, thawing remains a constant threat to the integrity of all P/S samples until they have been analyzed.

For any given collection of samples, the scale of the problem depends on the degree to which specimen biomolecules no longer reflect *in vivo* reality. But it is not immediately obvious that for common biomarker discovery and related investigations, even a very low percentage of molecules deviating from *in vivo* reality presents a major *false discovery trap*: several studies over the past decade have estimated the percentages of different types of biomolecules that are quantitatively altered upon P/S exposures to common thawed conditions (16–21). They have found that generally low percentages (*i.e.*, 2–20%) of biomolecules are unstable when the specimens in which they reside are exposed to common thawed conditions. On its surface, this gives the appearance that the problem is insignificant. A quick thought experiment, however, demonstrates the opposite: imagine a set of serum samples from stage I lung cancer patients that have been paired with age/gender/smoking-matched samples from at-risk but cancer-free donors to comprise a case/control study. If each sample contains 50,000 different measurable biomolecules (*i.e.*, 50,000 qualitatively different metabolites, proteins, miRNAs, etc.), we might optimistically estimate that five of these might serve as effective markers of stage I lung cancer. If, however, a portion of the case specimens were exposed to a $> -30^{\circ}\text{C}$ thaw event while the control samples were not (or vice versa—a rather common scenario as we have shown (22, 23)), resulting in statistically significant shifts in the concentration of *only* 2% of exposed-sample

biomolecules, this would introduce a 200-fold excess ((50,000 * 0.02)/5 = 200) of falsely altered biomolecules relative to the number of bona fide biomarkers waiting to be discovered. Hence the *false discovery trap*: the seemingly minor thaw event has introduced a clinical reality—hiding minefield of false/irreproducible “discoveries” waiting to mislead investigators and waste time and money in the process (Fig. 1).

Of course, P/S thawing results in multiple different types of *ex vivo* biomolecular change. Proteins are susceptible to artifactual *ex vivo* posttranslational modifications such as glycation (24–26), oxidation (both S-cysteinylation (22, 27) and methionine sulfoxidation (27)), and proteolytic degradation (28–31). Modifications such as these can potentially impact protein quantification without any indication that they have

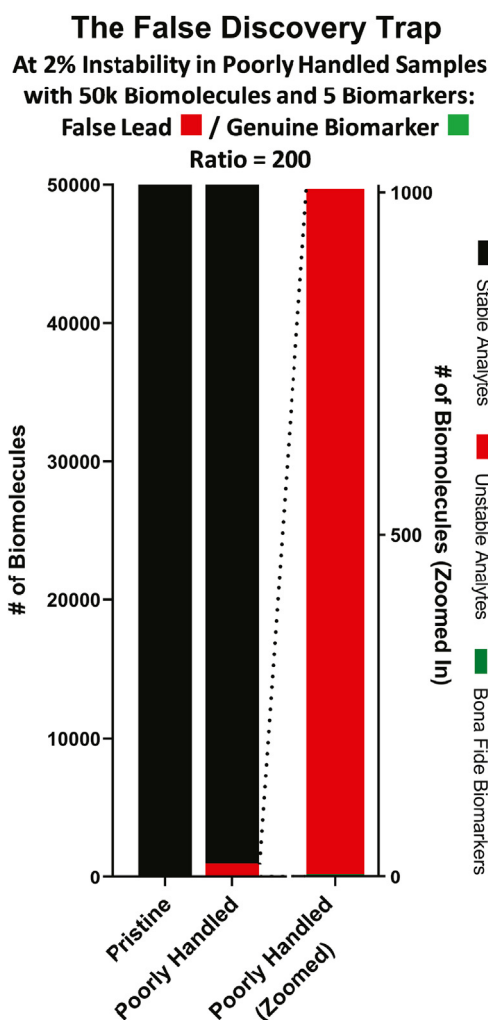


Fig. 1. The False Discovery Trap. Unstable biomolecules within a poorly handled cohort will change relative to any other cohort with which they are compared, giving them the appearance of biomarkers. Under modest cohort mishandling that introduces quantitative instability into just 2% of analytes, a severe false lead-to-genuine biomarker ratio is introduced, creating a minefield of false discoveries that can be difficult if not impossible to avoid.

occurred—particularly when quantification is dependent on specific molecular interactions that may be silently disrupted by these modifications (e.g., immunoassays). Numerous examples of apparent changes in protein concentrations occurring due to P/S exposure to thawed conditions are summarized elsewhere (18, 23, 32, 33). Unfortunately, enzymatic, oxidative, and/or other *ex vivo* chemical processes do not solely impact proteins; they impact every major class of biomolecule, including small molecules/metabolites (34–36), lipids (37, 38), cholesterol (38), peptides (7), nucleic acids such as miRNAs (39), and, to a modest degree, glycans (40–42). This problem should be of substantial concern to biomedical researchers who employ archived blood P/S in their research, as such evidence suggests that without comprehensive documentation on how these archived P/S samples have been handled and stored, it may be impossible to properly determine their suitability for specific projects.

Unfortunately, documentation alone does not always suffice as sufficient evidence of P/S integrity. We have recently shown *via* two independent incidents that empirical molecular evidence—above and beyond paper trails—may be required to accurately document the integrity of archived P/S samples (22, 23). These discoveries were enabled by our recent development of a biomarker of P/S exposure to thawed conditions known as Δ S-Cys-Albumin (22), which quantifies cumulative exposure of P/S to thawed conditions (*i.e.*, temperatures $> -30^{\circ}\text{C}$ (14, 15)). In summary, it is a 10- μL , dilute-and-shoot, intact-protein LC/MS-based assay of the relative abundances of albumin proteoforms. The assay is based on the fact that the relative abundance of S-cysteinylated (oxidized) albumin in P/S increases inexorably but to a maximum value under 100% when samples are exposed to temperatures above -30°C . The difference in the relative abundance of S-cysteinylated albumin (S-Cys-Albumin) before and after an intentional incubation that drives this proteoform to its maximum level is denoted as Δ S-Cys-Albumin. Δ S-Cys-Albumin in fully expired samples is zero. The range (with mean $\pm 95\%$ CI) observed for Δ S-Cys-Albumin in freshly collected plasma is 12 to 29% ($20.9 \pm 0.75\%$; $n = 97$), and in matched serum, it is 10 to 24% ($15.5 \pm 0.64\%$; $n = 97$) (22). Cumulative exposure is calculated *via* a multi-reaction biochemical rate law that we and others have established (22, 43) and that we have shown is applicable to actual plasma and serum samples, enabling back-calculation of the time at which unknown P/S specimens have been exposed to the equivalent of room temperature (22).

Rate law-based linkage of Δ S-Cys-Albumin to the equivalence of exposure time at room temperature provides an intrinsic connection between Δ S-Cys-Albumin and any clinical biomarker with a known stability profile at room temperature. But for candidate clinical biomarkers without known stability profiles or P/S samples that may be known to have only been exposed to refrigeration or -20°C storage, a definitive connection between Δ S-Cys-Albumin and marker stability

remains poorly established. Thus, the purpose of this study was to concurrently evaluate the stability of proteins of interest to pre-clinical research at -20°C , 4°C , and 23°C in conjunction with Δ S-Cys-Albumin measurements in order to begin to forge an empirical linkage between protein stability at all major storage temperatures and Δ S-Cys-Albumin.

EXPERIMENTAL PROCEDURES

Experimental Design and Statistical Rationale

The aspects of this study that required substantial planning with regard to methodological design and the logistics of execution included patient enrollment; collection of matched serum, lithium heparin plasma and K_2EDTA plasma, processing and short-term handling/storage; and execution of thawed-state stability studies at -20°C , 4°C , or 23°C .

The most important details on how these aspects of the study were executed as well as brief descriptions of the analytical procedures are described below. Further detailed information on donor inclusion and exclusion criteria, blood collection and processing protocols, Δ S-Cys-Albumin measurements, and protein measurements is provided in [Supplemental Data](#).

Donor Information and Plasma and Serum Specimen Collection—Matched serum, lithium heparin plasma, and K_2EDTA plasma were collected (in that order) from gastrointestinal (GI) cancer patients and cancer-free donors under informed consent and local IRB approval by the Cooperative Human Tissue Network or Valleywise Health. Specimens were collected in compliance with the Declaration of Helsinki principles. Analysis of the specimens as described in this article was approved by the Arizona State University IRB. Basic donor demographics and disease status information is provided in [Table 1](#). Patients with compromised kidney function (*i.e.*, $\text{eGFR} < 60 \text{ mL/min per } 1.73 \text{ m}^2$) were excluded due to the possibility that poor kidney function can, in theory, result in abnormally high Δ S-Cys-Albumin measurements (22).

In total, 84 patients were enrolled in this study. Thirty seven were GI cancer patients and 47 were cancer-free control donors. Δ S-Cys-Albumin in their matched K_2EDTA plasma, lithium heparin plasma, and serum was measured at baseline. Of these patients, matched K_2EDTA plasma and serum from 24 of them were employed for the Δ S-Cys-Albumin time courses. K_2EDTA plasma from these 24 patients was employed for measurements of the clinically relevant proteins. The data shown in [Figure 2](#) from 97 non-acute cardiac patients were originally described elsewhere (22) and were simply included in [Figure 2](#) for the sake of comparison.

A strict blood collection and processing protocol was followed that included the following: plasma tubes were pre-chilled to 0 to 4°C . Collection tubes were properly filled (any partially filled tubes were rejected). Immediately after collection, serum tubes were inverted (never shaken) five times and plasma tubes were inverted eight times. Serum was allowed to clot at room temperature for 45 min. Matched plasma was placed on ice while serum clotted, then all tubes were centrifuged at 4°C . Plasma and serum were then immediately aliquoted on ice. Aliquots were placed in a -80°C freezer within 2 h from the time of initial draw. To verify the timing of all processing steps, time stamps were recorded at (1) the time of initial draw, (2) time of centrifugation completion, and (3) the time at which aliquots were placed at -80°C . P/S with a visually estimated degree of hemolysis $> 250 \text{ mg/dL}$ were excluded.

Plasma and serum samples were shipped to Arizona State University overnight on dry ice. Upon receipt, specimens were verified as frozen and then unpacked into a -80°C freezer equipped with

TABLE 1
Patient characteristics for the baseline Δ S-Cys-Albumin and protein stability studies

Patient demographics	Patient samples for determination of baseline Δ S-Cys-Albumin values	Patient samples for stability study
	n (%)	n (%)
Number of patients	84	24
Sex		
Males	35 (41.7)	13 (54.2)
Females	49 (58.3)	11 (45.8)
Age (years)		
<40	16 (19.1)	3 (12.5)
40–60	44 (52.3)	11 (45.8)
>60	24 (28.6)	10 (41.7)
Ethnicity/race		
White	75 (89.3)	20 (83.3)
Black	7 (8.3)	4 (16.7)
Other	2 (2.4)	0 (0)
Disease status		
Cancer-Free	47 (56.0)	12 (50.0)
Cancer	37 (44.0)	12 (50.0)
Stage I	1 (2.70)	1 (8.3)
Stage III	1 (2.70)	1 (8.3)
Stage III	4 (10.8)	1 (8.3)
Stage IV	10 (27.0)	3 (25.0)
Stage	21 (56.8)	6 (50.0)
undetermined		

continuous temperature monitoring. All specimens were allowed to sit in the -80°C for at least 7 days prior to thawing to ensure that any residual CO_2 in the headspace had been exchanged for air in order to avoid any CO_2 -induced sample acidification effects (44).

Initial Analysis and Thawed-State Stability Studies—All samples were randomized prior to analysis using a random number generator for run order assignment. Δ S-Cys-Albumin in pristine, never-thawed aliquots was measured within a time range of 12 to 20 months after initial specimen collection, during which time specimens were kept continually at -80°C .

Following collection of baseline Δ S-Cys-Albumin values, K_2EDTA plasma and matched serum from 12 GI cancer patients and 12 cancer-free control donors were selected for inclusion in thawed-state stability studies based on their initial Δ S-Cys-Albumin values being equally distributed across the entire range of initial Δ S-Cys-Albumin values measured. Thawed-state stability studies involved incubation of separate aliquots of these plasma and serum samples for up to 65 days at -20°C , 28 days at 4°C , or 96 h at 23°C . Δ S-Cys-Albumin was measured in both plasma and serum; proteins of clinical interest were only measured in K_2EDTA plasma.

A separate aliquot for each individual nonbaseline thawed-state time point and protein assay (either 20 μl for Δ S-Cys-Albumin or 100 μl for protein measurements) was created from a parent aliquot that had never previously been thawed. To create these aliquots, the parent sample was thawed and kept on ice with strict time tracking for a period of 3.4 ± 1.1 (SD) min for the 23°C -exposed samples, 8.5 ± 2.0 min for the 4°C -exposed samples, and 14.2 ± 3.0 min for the -20°C -exposed samples—by which times all temperature exposure time courses were started. Time course aliquots were

immediately placed at -80°C upon completion of their time/temperature exposure period. Once all time courses were completed, time course aliquots were randomized and Δ S-Cys-Albumin or the clinically relevant proteins were measured in them. Clinically relevant proteins at all stability time points (including baseline) were measured as a randomized set of 240 K_2EDTA plasma samples by multiplexed immunoassay on either the Luminex 100 or 200 platform by MyriadRBM. Each protein was measured once in each sample. Given that patient specimens were collected over a period of about 1 year, clinically relevant proteins were analyzed within a time range of 18 to 30 months after initial specimen collection. With the exception of intentional thawed-state incubation periods, all specimens were kept continually at -80°C prior to analysis. After all Δ S-Cys-Albumin and clinically relevant protein time course specimens were analyzed, Δ S-Cys-Albumin was once again measured in a residual never-thawed aliquot of each sample in order to verify the long-term stability of Δ S-Cys-Albumin at -80°C . This occurred one year after the initial Δ S-Cys-Albumin measurements in never-thawed samples were made.

Statistical Analysis

To evaluate the stability of clinically established protein biomarkers in K_2EDTA plasma, the concentration measured for each analyte at each condition was compared with its respective control aliquot kept continuously at -80°C . Data consisted of $n = 24$ subject samples measured longitudinally over time at temperatures of 23°C , 4°C , and -20°C , respectively. First, we performed descriptive statistics and analyzed Δ S-Cys-Albumin at the baseline data (at time = 0), comparing the effect of health status (cancer versus normal), gender (male versus female), and race (white versus black) using a two-sample t test. The age effect was tested using simple linear regression. Second, we analyzed longitudinal Δ S-Cys-Albumin data using linear mixed effects models to account for the repeated measures within subjects. Specifically, we consider the full model as follows:

$$y_{ij} = \beta_0 + \beta_1 \text{age}_i + \beta_2 \text{status}_i + \sum_{j=1}^q \beta_{3j} t_{ij} + \sum_{j=1}^q \beta_{4j} \text{age}_i t_{ij} + \sum_{j=1}^q \beta_{5j} \text{status}_i t_{ij} + \epsilon_{ij}$$

where y_{ij} is the concentration of Δ S-Cys-Albumin for the i^{th} subject at j^{th} time point, age_i is the age for the i^{th} subject, status_i is 1 if i^{th} subject's status is cancer and 0 otherwise, $t_{ij}=1$ if y_{ij} is observed at the j^{th} time point, and $t_{ij}=0$ otherwise, ϵ_{ij} is an error term, and q is the number of time points observed excluding the baseline. We employed a step-down approach that dropped the interaction term of status and time if they were nonsignificant and used reduced models with the main effects only.

The same approach was used to analyze the 21 clinically relevant proteins, with natural logarithm-transformed y values to improve normality. Using the reduced model, the least square means of the y values at each time point were computed and compared between the baseline and any of the other time points using Wald tests. Bonferroni corrections were applied to adjust the calculated p -values of the 21 proteins for pairwise comparisons between time points. Bonferroni-adjusted p -values less than 0.05 were regarded as statistically significant. All the analyses were conducted at each temperature separately. Serum and K_2EDTA plasma samples were additionally separated when analyzing Δ S-Cys-Albumin.

Finally, we conducted repeated measures correlation (45) to quantify the association between Δ S-Cys-Albumin and each unstable protein. Pearson correlation coefficients were calculated to quantify the association between Δ S-Cys-Albumin and the percentage of destabilized proteins. We used GraphPad Prism software (version

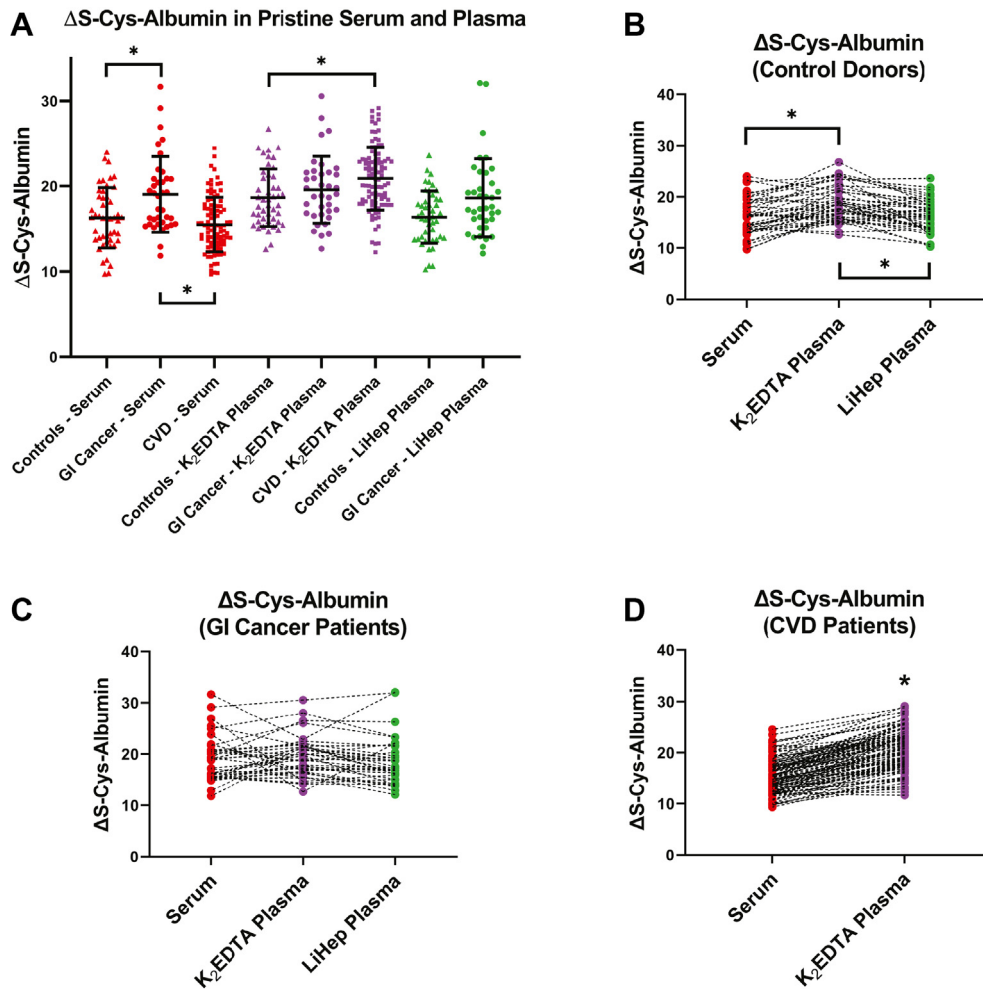


FIG. 2. Δ S-Cys-Albumin in fresh, rapidly processed human serum and matched plasma samples. A, comparisons between cancer-free control donors ($n = 47$), GI cancer patients ($n = 37$), and non-acute cardiac (CVD) patients ($n = 97$) for each matrix. Error bars represent mean \pm SD; from left to right these values are: 16.3 ± 3.5 , 19.0 ± 4.4 , 15.5 ± 3.2 , 18.6 ± 3.4 , 19.6 ± 3.9 , 20.9 ± 3.7 , 16.4 ± 3.1 , and 18.6 ± 4.6 . * indicates a significant difference between means of indicated groups based on a one-way ANOVA with Tukey's posthoc test or a t test. Results were compared to a Bonferroni-corrected p -value threshold to account for multiple comparisons (i.e., $p < 0.05/7$ or 0.0071). B–D, matched collections within each patient group, including (B) cancer-free control donors, (C) GI cancer patients, and (D) CVD patients. * indicates a significant difference between matched sets with $p < 0.00625$; repeated measures (RM) ANOVA with Tukey's posthoc test or paired t test. CVD, cardiovascular disease; GI, gastrointestinal.

9.3.1) for descriptive statistics, t -tests, and repeated measures-ANOVA; MIXED procedure in SAS 9.4 (SAS Inst. Inc.) for linear mixed effects models and R 4.1.0 (R foundation, Vienna, Austria); and R rmcrr package for correlation analyses.

Laboratory Procedures

Measurement of Δ S-Cys-Albumin—P/S samples were prepared and the percentage of albumin in the S-cysteinylation form (S-Cys-Albumin) was measured as previously described (22, 27). Briefly, 1 μ l of plasma or serum was diluted 1000-fold in 0.1% (v/v) TFA and injected onto an LC-MS instrument where albumin was concentrated and desalted on a protein cap-trap, then eluted directly into the mass spectrometer for measurement of the intact protein and relative quantification of its proteoforms. Nine microliters of the same plasma or serum sample were then placed in a 0.6-ml polypropylene Eppendorf snap-cap tube and incubated in a dry oven at 37 °C for

24 h. One microliter of this sample was then diluted 500-fold in 0.1% TFA and injected onto the LC-MS for analysis. Δ S-Cys-Albumin is defined as the difference between the percentage of albumin in the S-cysteinylation form before and after the overnight incubation at 37 °C that drives the percentage of S-Cys-Albumin to its maximum value (22). Details of the LC-MS method, including data processing are provided in Supplemental Data.

Measurement of Clinically Relevant Proteins—Twenty-five clinically relevant proteins were measured via three multi-analyte profile assays (MAPs) in each 100- μ l K₂EDTA plasma aliquot from the thawed-state stability studies. MAPs were preconfigured and validated by MyriadRBM and included “HCANCER2”, “HMP8”, and “HMPC38”.

The HCANCER2 MAP included amphiregulin, epidermal growth factor (EGF), epidermal growth factor receptor (EGF-R), epiregulin, heparin-binding EGF-like growth factor (HB-EGF), placenta growth factor (PGF), platelet-derived growth factor BB (PDGF-BB), and tenascin-C.

The HMP8 MAP included adiponectin, alpha-2-macroglobulin (A2M), ferritin, myoglobin, plasminogen activator inhibitor-1 (PAI-I), T-cell-specific protein RANTES, tissue inhibitor of metalloproteinases-1 (TIMP-1), tumor necrosis factor receptor-2 (TNF-R2), vascular cell adhesion molecule-1 (VCAM-1), EN-RAGE, and pulmonary and activation-regulated chemokine (PARC).

The HMPC38 MAP included alpha-fetoprotein (AFP), cancer antigen 125 (CA-125), cancer antigen 19-9 (CA-19-9), carcinoembryonic antigen (CEA), human chorionic gonadotropin beta (hCG-b), and neuron-specific enolase (NSE).

Samples were measured in random order by trained MyriadRBM personnel who thawed samples for the minimal required time and kept them on ice when thawed prior to analysis. Comprehensive assay validation data are available from MyriadRBM or the corresponding author upon request.

Four of the above proteins were measured but were below the limit of quantification (LOQ) in the vast majority of samples and were therefore not reported here. These included amphiregulin, epi-regulin, PGF, and hCG-b. For a few of the 21 proteins reported on here, one or more patients had concentrations that were at or below the LOQ for all stability time points. These included alpha-fetoprotein (14 patients always below LOQ), CA-19-9 (6 such patients), and CEA (1 such patient). Data on these proteins from such patients were excluded from statistical analysis. There were also instances in which at least one but not all protein concentrations from a given patient were reported as being at or below the LOQ. This occurred at least once for AFP, CA-125, CA-19-9, CEA, EGF, HB-EGF, and PDGF-BB. (supplemental Fig. S3 provides a graphical view of the overall minor extent of these occurrences.) These data points were included in the statistical analysis as being at the LOQ because they were known to not be higher than the LOQ and, for baseline measurements (time = 0), this information alone was potentially helpful in facilitating the detection of a protein instability when protein concentrations apparently increased over time—which constituted the vast majority of instabilities detected (such as that of EGF). In all such cases, had the LOQ been lower, it would have enhanced rather than diminished the magnitude of an instability finding. Consistency in this practice, however, meant that including nonbaseline measurements at the LOQ tended to otherwise blunt the ability to statistically detect unstable proteins.

RESULTS

Δ S-Cys-Albumin Baseline Values in Fresh Plasma and Serum Samples

In 2019, we published baseline values for fresh, rapidly processed, matched K₂EDTA plasma and serum samples from 97 non-acute patients with cardiovascular disease (CVD) (22). The clinical characteristics of these donors with CVD are described elsewhere (22) but their baseline Δ S-Cys-Albumin data are provided here for comparison alongside new data from 37 GI cancer patients and 47 cancer-free control donors. (Illustrative raw and charge deconvoluted mass spectral data are provided elsewhere (27).) Matched LiHep plasma was also collected from the GI cancer patients and cancer-free control donors (Fig. 2). An elevation of Δ S-Cys-Albumin in the serum of GI cancer patients compared to the other two patient cohorts was observed. Similarly, Δ S-Cys-Albumin in K₂EDTA plasma was higher in CVD patients than in the cancer-free control donors. The difference in Δ S-Cys-Albumin between LiHep plasma groups did not quite reach statistical significance (Fig. 2A). Within CVD patients and cancer-free controls,

Δ S-Cys-Albumin in K₂EDTA plasma was higher than that of serum in matched collections; and in the cancer-free controls, it was also higher than LiHep plasma (Fig. 2, B–D). No differences in matched collections were observed, however, between the three matrices in GI cancer patients. Baseline Δ S-Cys-Albumin values in chemotherapy patients (not annotated in Fig. 2) did not differ from cancer patients who were not on chemotherapy at the time of blood draw (*t* test; *p* > 0.5). The range (with mean \pm 95% CI) observed for Δ S-Cys-Albumin in fresh (baseline) GI cancer patient serum was 12 to 32% (19.0% \pm 1.5%). In their K₂EDTA plasma, it was 13 to 31% (19.6% \pm 1.3%). And in their lithium heparin plasma, it was 12 to 32% (18.6% \pm 1.5%). For the cancer-free controls, these values for serum were 10 to 24% (16.3% \pm 1.0%), for K₂EDTA plasma were 13 to 27% (18.6% \pm 1.0%), and for lithium heparin plasma were 10 to 24% (16.4% \pm 0.92%). For comparison and as described elsewhere (22), non-acute CVD patient serum was 10 to 24% (15.5% \pm 0.64%) and their K₂EDTA plasma was 12 to 29% (20.9% \pm 0.75%).

Effects of age, gender, and race on baseline Δ S-Cys-Albumin values in the GI cancer patients and cancer-free control donors were also assessed. No significant differences were observed between genders or races for any matrix (*t*-tests; *p* > 0.05; see Table 1 for *n*-values). Δ S-Cys-Albumin was slightly but consistently linearly correlated with age in all three matrices (supplemental Fig. S1). In serum, slope = 0.079, Δ S-Cys-Albumin units/yr, *p* = 0.015, Pearson correlation coefficient (*r*) = 0.26; in K₂EDTA plasma, slope = 0.074 Δ S-Cys-Albumin units/yr, *p* = 0.013, *r* = 0.27; and in LiHep plasma, slope = 0.064 Δ S-Cys-Albumin units/yr, *p* = 0.039, *r* = 0.23. After adjusting for age, the differences between GI cancer patients and cancer-free controls for serum and LiHep plasma (Fig. 2A) were diminished but not eliminated (*t* test; serum, *p* = 0.020; K₂EDTA plasma, *p* = 0.63; LiHep plasma, *p* = 0.029). And as expected, adjustment for age did not alter the fact that Δ S-Cys-Albumin was significantly higher in K₂EDTA plasma relative to serum and LiHep plasma in matched collections from the cancer-free control patients.

Δ S-Cys-Albumin Time Courses at 23 °C, 4 °C, and –20 °C

A separate aliquot for each individual time point was created from parent samples that had never previously been thawed. To create these aliquots, the parent sample was thawed and kept on ice for strictly limited and documented time periods as described in the Experimental Procedures section. Δ S-Cys-Albumin was measured at 9 to 11 different time points per exposure temperature (23 °C, 4 °C, or –20 °C) in K₂EDTA plasma and serum aliquots from 12 GI cancer patients and 12 cancer-free donors.

K₂EDTA plasma and serum results are shown separately (Fig. 3) because, as mentioned above and as we have previously described, K₂EDTA plasma and serum both tend to have different initial Δ S-Cys-Albumin values in matched samples and different rates of change when the samples are exposed

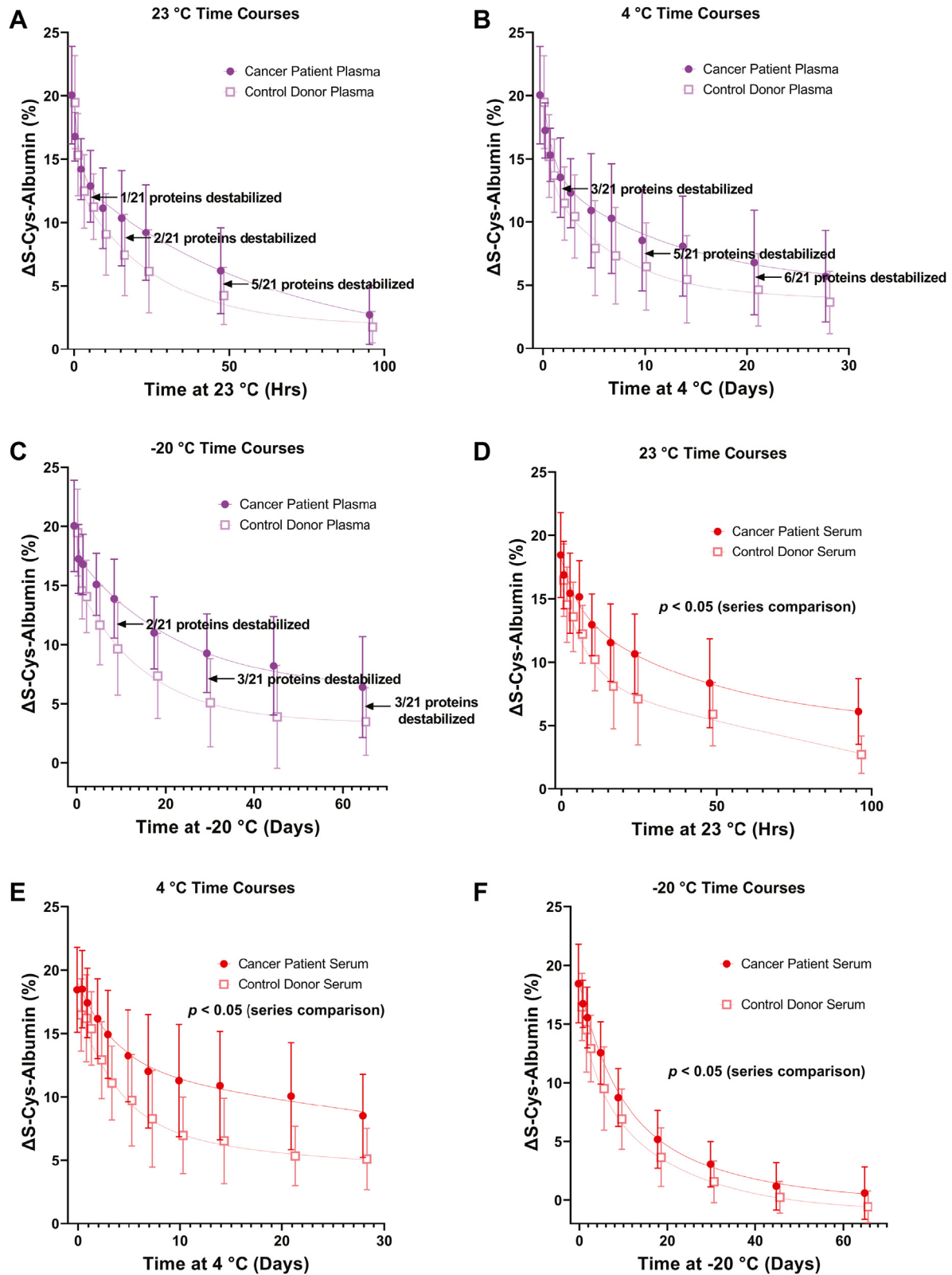


FIG. 3. Δ S-Cys-Albumin time courses at 23 °C, 4 °C, and -20 °C. Δ S-Cys-Albumin profiles over time in K₂EDTA plasma at (A) 23 °C, (B) 4 °C, and (C) -20 °C, and in serum at (D) 23 °C, (E) 4 °C, and (F) -20 °C. Jitter was added to the x-axis to prevent data point/error bar overlap. Mixed effects models indicated that the differences between GI cancer patients and cancer-free control donors were not statistically significant in K₂EDTA plasma time courses but were significant in serum time courses at all three temperatures (Bonferroni-adjusted $p < 0.05$). Modeled

to thawed conditions (22). Control measurements of Δ S-Cys-Albumin in never-thawed aliquots of these K₂EDTA plasma and serum samples that were kept continuously at -80°C for approximately 1 year after the initial baseline Δ S-Cys-Albumin measurements were made showed no significant change from their original measurements (supplemental Fig. S2). An initial mixed effects model confirmed a strong interaction between matrix type and time ($p < 0.001$ at all temperatures). As such, K₂EDTA plasma and serum time courses were subsequently analyzed separately by mixed effect models that included patient age, health status, and time at the indicated temperature as main effects and patient age \times time and health status \times time as interactions.

Neither interaction was found to be statistically significant in either K₂EDTA plasma or serum—meaning that there was no statistical evidence that patient age or health status impacted the manner in which Δ S-Cys-Albumin decayed over time. Data were re-analyzed after eliminating these interactions from the model in order to robustly identify significant main effects: as expected, exposure time at all three temperatures resulted in strongly significant changes in Δ S-Cys-Albumin values (Bonferroni-adjusted $p < 1 \times 10^{-8}$). Patient health status did not significantly impact overall K₂EDTA plasma time course results at any temperature (Fig. 3, A–C), but it did significantly impact overall serum time course results at all three temperatures (Bonferroni-adjusted $p < 0.05$; Fig. 3, D–F). This indicated that GI cancer patient serum samples tended to have slight to modestly higher Δ S-Cys-Albumin values than the cancer-free control donors regardless of any thawed-state exposure that may have occurred. Overall, these results are consistent with the fact that Δ S-Cys-Albumin in the serum of cancer patients started (at baseline) at significantly higher values than in the serum of cancer-free donors. (Fig. 2A). The mixed effects model also revealed that patient age had a slight but statistically significant impact on overall K₂EDTA plasma time course results at 4°C and -20°C and significantly impacted overall serum time course results at 23°C (Bonferroni-adjusted $p < 0.05$; time courses stratified by age are not shown). These observations were consistent with the weak correlations of Δ S-Cys-Albumin with patient age observed in fresh K₂EDTA plasma and serum samples. As noted below, however, the vast majority of clinically relevant proteins in fresh K₂EDTA plasma samples had much stronger correlations with patient age than did Δ S-Cys-Albumin.

Clinically Relevant Protein Time Courses at 23°C , 4°C , and -20°C

Twenty-one clinically relevant proteins (supplemental Table S1 and listed above) were measured by Luminex

assay in the K₂EDTA plasma samples from both the GI cancer patients and cancer-free controls at baseline plus three additional time points per temperature (supplemental Fig. S3). At baseline, there were no significant differences in any clinically relevant protein based on gender, race, or health status after correcting for multiple comparisons. Notably, however, all of these proteins except α -fetoprotein (AFP) and carcinoembryonic antigen (CEA) were strongly significantly correlated with patient/donor age (supplemental Table S2).

Several proteins exhibited instability in K₂EDTA plasma over time at 23°C , 4°C , and -20°C (Figs. 4 and S3). Repeated measures-based mixed effects models were used to determine which proteins changed in a statistically significant manner over time at each temperature. Interactions of patient age \times exposure time and patient health status \times exposure time at each temperature were considered (along with patient age, health status, and exposure time as main effects) but no interactions were found to be statistically significant. As such, the mixed effects model was simplified to include only patient age, health status, and exposure time at each temperature as main effects. Following the identification of significantly altered proteins, Wald tests were applied to identify which time point(s) were significantly different from baseline. Here, criteria of raw p -value less than 0.05/63 (due there being 21 proteins compared at 3 time points per temperature) or 7.9×10^{-4} and fold change $>10\%$ were both required to consider a protein as significantly altered by a given time-temperature exposure (Fig. 4). The proteins most strongly and consistently impacted by exposure of K₂EDTA plasma to 23°C , 4°C , or -20°C were neuron-specific enolase (NSE), epidermal growth factor (EGF), plasminogen activator inhibitor-1 (PAI-1), and RANTES (also known as CCL5).

Relationships Between Δ S-Cys-Albumin and Unstable Clinically Relevant Proteins

Concurrent quantitative changes in Δ S-Cys-Albumin and each of these four unstable proteins were analyzed for consistency by repeated measures correlation (45). Strong correlations were found at most temperatures ($p < 0.001$, with $p < 0.05/12$ or 0.0042 considered significant due to multiple comparisons) (Fig. 5). EGF, PAI-1, and RANTES all exhibited apparent increases in concentration as Δ S-Cys-Albumin decreased (i.e., with increased exposure time) at all three temperatures. NSE was unique in that at 23°C and 4°C , it initially increased but then decreased at longer exposure times (Figs. 5 and S3). At -20°C , however, it only exhibited an apparent increase over the time span monitored.

Finally, the relationship between Δ S-Cys-Albumin and the total percentage of destabilized proteins in K₂EDTA plasma

interactions of patient health status \times time and patient age \times time were not statistically significant in any matrix or at any temperature (Bonferroni-adjusted $p < 0.05$) — meaning that there was no statistical evidence that patient age or health status impacted the manner in which Δ S-Cys-Albumin decayed over time. $n = 12$ patient samples per data point. Error bars indicate SD. Lines are fitted 2-phase exponential decay curves that are intended to serve only as visual guides. GI, gastrointestinal.

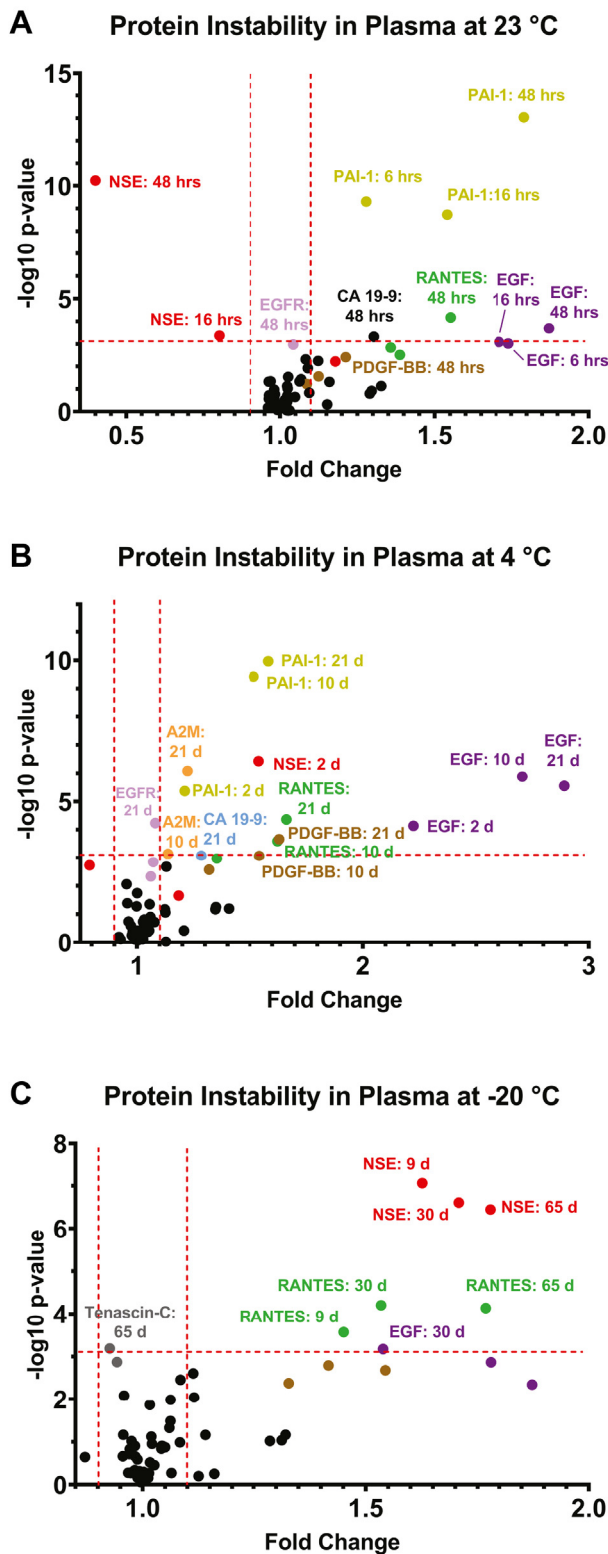


FIG. 4. Clinically relevant protein instability in K_2EDTA plasma exposed to thawed conditions. Volcano plots are shown for (A) 23 °C, (B) 4 °C, and (C) -20 °C. Raw p -values were determined using repeated measures-based mixed effects models on natural log-transformed data. Each data point represents the $-\log_{10} p$ -value and

was evaluated, considering the data from all three temperatures together (Fig. 6). Proteins were considered significantly changed based on the p -value and fold-change criteria described above for the data in Figure 4. In addition, to account for the increase-then-decrease behavior of NSE (or similarly behaving proteins), once a protein reached the threshold for significant change (destabilization), all future time points for the protein at that temperature were also considered points at which the protein was destabilized. A clear inverse linear relationship was found ($r = -0.61$ and $p < 0.0001$).

For the 23 °C time course, the once-changed-then-always-subsequently-changed policy noted above pertaining to Figure 6 had no effect. For the 4 °C time course, the policy kept only NSE significant when it might not otherwise have been. But this is appropriate because based on the 23 °C time course (Figs. 4 and S3), as NSE had an excursion in the positive direction then proceeded downward, transitioning through a false negative range. No other proteins in the 4 °C time course were affected by this policy. For the -20 °C time course, the policy kept only EGF at 65 days positive. Without the policy, EGF at 65 days would not have been considered statistically significant within Figure 6 because it had an elevated p -value (though its fold-change was actually increased to 1.9-fold at this time point). If the EGF point at 65 days was not considered significant, this would shift the mean data point in Figure 6 that is lowest and furthest to the right on its x-axis to about 20.6%, which would actually bring it closer to the regression line—a result that shows that as it relates to proteins other than NSE, this policy actually caused a slight departure from the linear correlation shown in Figure 6.

DISCUSSION

Δ S-Cys-Albumin was reported by our group as an endogenous marker of blood plasma and serum exposure to thawed conditions (>-30 °C) in 2019 (22). It is unique among candidate markers for this purpose because it has been extensively characterized and validated: its mechanism of formation is understood; the multi-reaction rate law that governs albumin S-cysteinylation at 23 °C was established and can be used to approximate exposure times of unknown samples; the population reference range for fresh samples from CVD patients was determined in 2019 (and has now been estimated here for GI cancer patients and cancer-free controls); it has passed both group-wise and individual sample-level blind challenges; and it has been employed with “real life” samples to detect previously undisclosed thawed-state exposures of nominally

mean of measurements made on 24 separate donor samples (including 12 from GI cancer patients and 12 from cancer-free control donors). Proteins in the upper left and upper right regions delineated by dashed red lines representing a $-\log_{10} p$ -value of $-\log_{10}(0.05/63) = 3.10$, and fold change $>10\%$ were considered significantly altered by the indicated exposure conditions. GI, gastrointestinal.

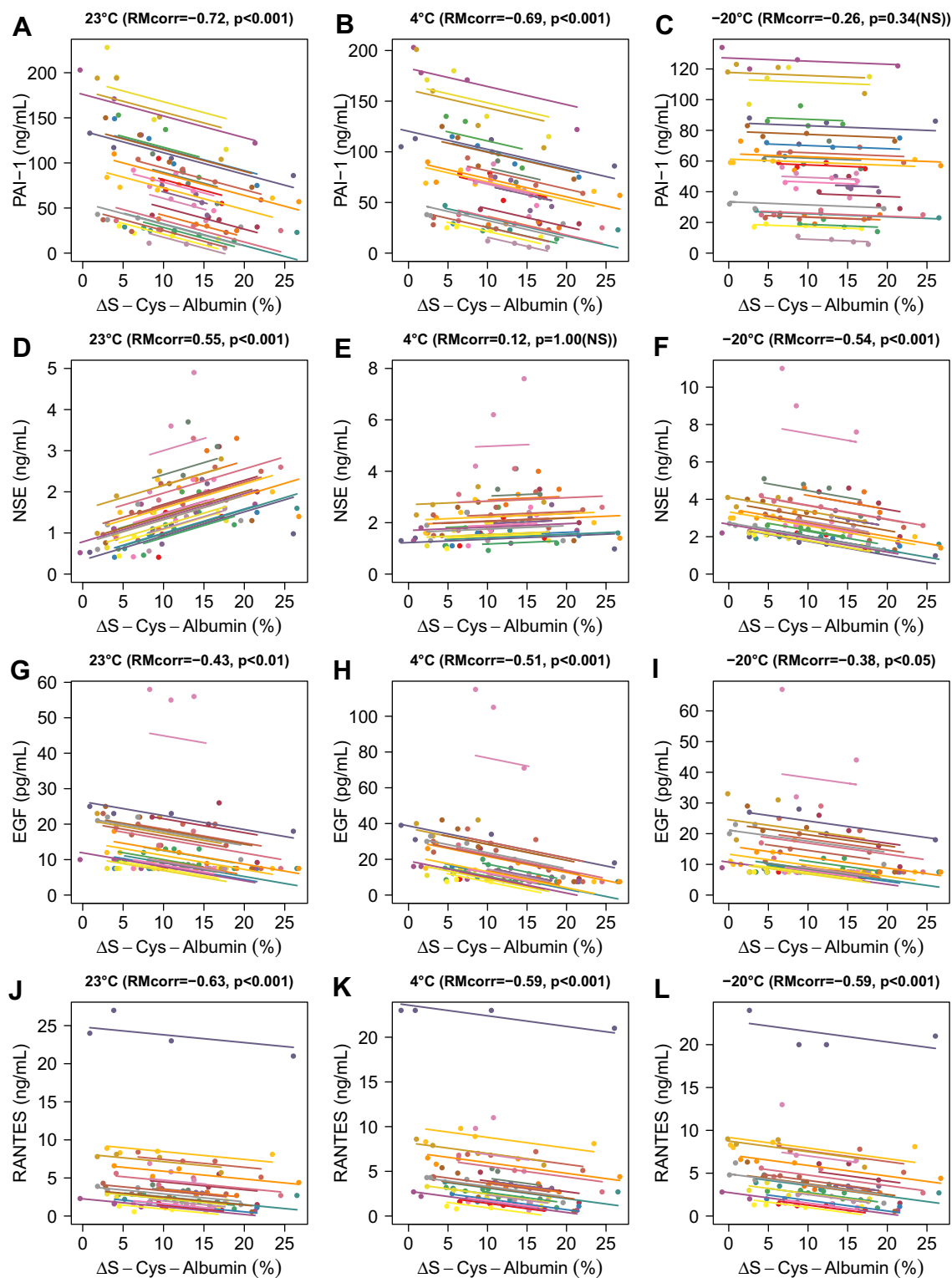


FIG. 5. **Relationships between Δ S-Cys-Albumin and specific destabilized proteins.** Repeated measures (RM) correlation plots for plasminogen activator inhibitor-1 (PAI-1) (A-C), neuron-specific enolase (NSE) (D-F), epidermal growth factor (EGF) (G-I), and RANTES (J-L). The threshold for statistical significance was placed at $p < 0.05/12$ or 0.0042 due to multiple comparisons.

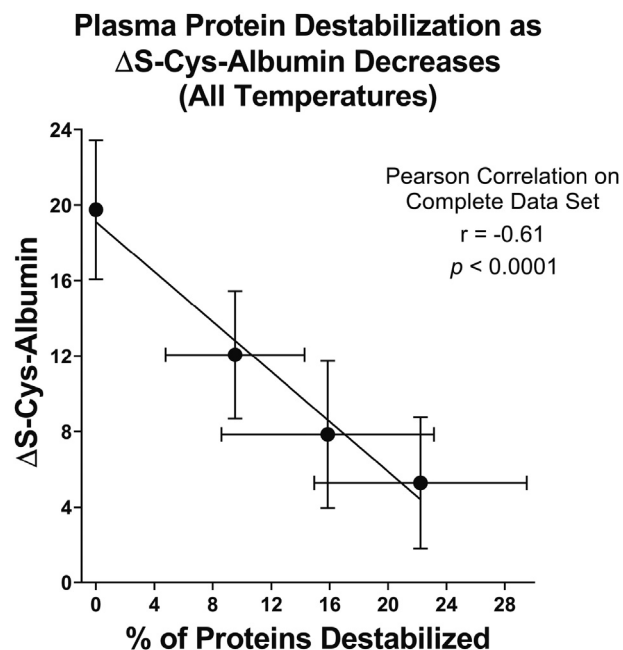


FIG. 6. Relationship between Δ S-Cys-Albumin and the percentage of proteins destabilized in K_2 EDTA plasma regardless of specific time-temperature exposure. Each point represents the mean Δ S-Cys-Albumin value at the 0th ($n = 24$), first ($n = 72$), second ($n = 72$), and third ($n = 72$) exposure time points from all three temperatures (23 °C, 4 °C, and -20 °C) combined. Error bars represent SD. Pearson correlation results shown ($r = -0.61$, $p < 0.0001$) were based on all 240 individual data points, not just the four mean values shown.

pristine samples being employed for biomarker discovery and validation purposes (22, 23). The 2019 paper (22) provides the information needed for laboratories to set guidelines on how Δ S-Cys-Albumin values in unknown samples can be used to evaluate the quality of P/S samples prior to starting a proteomic analysis. Here, we have extended the potential utility of Δ S-Cys-Albumin by generating empirical stability time course profiles from two dozen individual patients for both serum and K_2 EDTA plasma at -20 °C, 4 °C, and 23 °C. Moreover, we have empirically linked Δ S-Cys-Albumin to the stability of 21 clinically relevant proteins (as measured by Luminex assay) and have shown how drops in plasma Δ S-Cys-Albumin below the range observed in pristine samples can serve as a surrogate indicator to estimate the percentage of immunoassay-measured proteins that have been destabilized in mishandled specimens (Fig. 6).

Such empirical linkage of Δ S-Cys-Albumin to protein stability is useful for situations in which the stability of (pre)clinically important protein(s) of interest at room temperature is not known and therefore cannot be linked to a Δ S-Cys-Albumin cutoff threshold *via* its established rate law. This linkage, combined with the established mechanism of instability, known rate law, and other validation criteria described above make Δ S-Cys-Albumin the most thoroughly characterized and

validated marker of blood P/S exposure to thawed conditions to date—by a substantial margin.

In this study, we extended measurements of Δ S-Cys-Albumin in freshly collected and processed (baseline) samples from matched serum and K_2 EDTA plasma to include matched LiHep plasma as well. Δ S-Cys-Albumin in LiHep samples was the same as in serum samples, but Δ S-Cys-Albumin in K_2 EDTA plasma of healthy and CVD patients tended to run a bit higher than matched serum (and LiHep) plasma. Δ S-Cys-Albumin in freshly collected and processed samples from cancer patients, however, was the same across all three matched matrices. The mechanism behind this discrepancy is not clear but seems to be driven by modestly elevated levels of Δ S-Cys-Albumin in cancer patient serum *versus* serum from the other two patient cohorts (Fig. 2A). In fact, though it may not have quite reached statistical significance, Δ S-Cys-Albumin was modestly elevated in every matrix from cancer patients relative to the cancer-free control donors.

Elevated Δ S-Cys-Albumin could be caused by high cysteine/cystine concentration and/or low albumin concentrations. The most likely cause of elevated cysteine/cystine is poor renal function (22, 46), but patients with compromised renal function (eGFR <60 ml/min per 1.73 m²) were excluded from this study. Hypoalbuminemia, however, is common in cancer patients (47) and can be associated with chemotherapy (48) as well as diminished patient survival rates (49). Most cancer patients were not undergoing chemotherapy at the time of blood collection, and those that were did not have different Δ S-Cys-Albumin values. Absolute concentrations of albumin were not determined, leaving modest hypoalbuminemia as the most likely explanation for slightly elevated Δ S-Cys-Albumin in the cancer patients relative to controls. (Notably, even after correction for age, the significant difference in baseline serum Δ S-Cys-Albumin between cancer patients and cancer-free donors remained.) The source of the difference in Δ S-Cys-Albumin between serum and plasma was investigated in 2019 (22). No definitive explanation was identified, but the pre-centrifugation time difference between plasma and serum as well as serum clotting time were ruled out as contributors to this phenomenon.

The slight increase in baseline Δ S-Cys-Albumin observed with patient age was consistent with known decreases in P/S albumin concentration and modest increases in cysteine/cystine with age (50, 51). Notably, most of the clinically relevant proteins in baseline specimens were more strongly correlated with patient/donor age than was Δ S-Cys-Albumin (supplemental Table S2).

Variability among individual patient time course profiles is expected (22). The results observed here (Fig. 3) are in line with the theoretical Δ S-Cys-Albumin decay curves at 23 °C based on the previously established rate law and known population reference ranges for starting concentrations of reactants and products (22). Serum tends to run at the upper limit predicted, but this probably has to do with the difficulty in

estimating the concentration of catalytically available copper in the original model (22). This estimation is difficult because approximately 95% of copper is bound to ceruloplasmin in serum (52) where it is much less catalytically available relative to K_2EDTA plasma (22) where it has been extracted from ceruloplasmin and is mostly bound to EDTA.

The observation that Δ S-Cys-Albumin in serum decreases at least at fast, if not faster, at -20°C compared to its rate of decrease at 4°C was unexpected. Plasma and serum are partially (noneutectically) frozen at -20°C (14, 15, 53). This creates a non-native system of reactants and products that are likely differently concentrated compared to when all components of plasma or serum are in the liquid state. The rate laws for the relevant reactions have not been determined at 4°C or -20°C , but regardless of whether or not the reactions exhibit Arrhenius behavior, it remains possible that this differently concentrated state may lead to a situation in which the reactions run faster despite the colder temperature. Of course, this may also be reflected in the rates of other biochemical reactions—hence the need for an *empirical* linkage between Δ S-Cys-Albumin and proteins of clinical interest.

Additionally, as we have previously observed, copper is less catalytically available in serum than it is in EDTA plasma—which is likely due to the fact that it is mostly bound to ceruloplasmin in serum (22). This may account for the fact that this effect is only observed in serum—that is, in a partially frozen state at -20°C , ceruloplasmin's structure may be altered, allowing copper to become abnormally catalytically available. At this point, however, this conjecture is strictly speculative.

The mechanisms behind the documented protein instabilities are only partially understood at this point. PAI-I is known to be released from platelets during freeze-thaw cycles (54) and in the presence of residual thrombin activity (55). For the other proteins that exhibited instability, it is logical to conclude that biomolecular changes that took place upon exposure to thawed conditions resulted in disruption of protein epitope(s) that were involved in mediating the protein interaction-based quantification. The mechanisms behind these disruptions may include artifactual oxidation (especially of free cysteine and/or methionine residues (56–58), cf. supplemental Table S1), proteolysis, or other forms of *ex vivo* posttranslational modification (24)—even, theoretically, disturbances to tertiary or quaternary protein structure that disrupt epitopes involved the quantification process. Interestingly, NSE has a particularly large number of free cysteine residues (supplemental Table S1) that may be susceptible to *ex vivo* oxidation by mechanisms akin to those that drive albumin S-cysteinylation (22, 27, 59). A list of proteins measured in this study alongside details of the clonality (*i.e.*, polyclonal or monoclonal) of capture and detection antibodies that were employed is provided in supplemental Table S1. While one might expect assays that rely monoclonal antibodies to be more susceptible to disruption in target protein epitopes, no

clear pattern emerged from this information that would suggest that the clonality of antibody reagents employed had a major impact on the observed stability of plasma proteins.

CONCLUSIONS

The work presented here reinforces the fact that the exposure of human plasma to thawed-state conditions (*i.e.*, $> -30^\circ\text{C}$) can lead to apparent quantitative changes in nontrivial percentages of clinically important proteins as measured by molecular interaction-based assays (*e.g.*, antibody-based protein assays) (17, 60–63). Concurrent measurements of Δ S-Cys-Albumin and the concentrations of 21 proteins of interest to (pre)clinical research provided a direct empirical linkage between these clinically important proteins and Δ S-Cys-Albumin—allowing Δ S-Cys-Albumin to be used as a surrogate indicator of the suitability of plasma samples for analysis of the clinically relevant protein(s). Additionally, the data presented here show how Δ S-Cys-Albumin measurements can approximate the overall extent to which the proteome of plasma sample(s) has been altered (as detected by molecular interaction-based assays) due to exposure to thawed-state conditions. Therefore, Δ S-Cys-Albumin can now be used as *both* a marker of cumulative exposure to thawed-state conditions as well as an estimator of cumulative damage to clinically important proteins present in plasma. This latter feature, facilitated by the data in Figure 6, may serve as an effective means of using Δ S-Cys-Albumin measurements to empirically gauge the risk of false discovery when employing archived specimens for research purposes.

DATA AVAILABILITY

All data are presented within the article and/or supplemental data in graphical and/or tabular format. More detailed versions of the data, including limited numbers of raw mass spectra are available upon request from the corresponding author. All raw mass spectrometry files have been uploaded to the MassIVE database at <https://massive.ucsd.edu/> and made publicly accessible. The data set was assigned the identifier “MSV000090061” with the name “Delta-S-Cys-Albumin Baseline and Time Courses for Stability Linkage to Other Proteins”, <https://doi.org/10.25345/C5K35MJ29> (for viewing without downloading), and FTP download link: <ftp://massive.ucsd.edu/MSV000090061/>. The data set is 1.02 TB in size.

Supplemental data—This article contains [supplemental data](#). References cited in the supplemental data include reference (22).

Acknowledgments—We are grateful to Dr Stephanie M. F. Thibert, Dr Jesús Aguilar Díaz de León, and Ms Kazi Waheeda for their assistance with specimen processing and other time course logistics. We would also like to thank the NCI funded Cooperative Human Tissue Network (CHTN; Nashville, TN)

and Valleywise Health (Phoenix, AZ) for providing bio-specimens. Other investigators may have received aliquots from the same plasma/serum specimens provided by CHTN.

Funding and additional information—The research reported here was supported in part by the National Cancer Institute of the National Institutes of Health under award number R33CA217702 (to C. R. B.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions—E. P. K., N. J., S. P. R., and C. R. B. conceptualization; E. P. K., N. J., S. P. R., Y. C., and C. R. B. methodology; E. P. K. and N. J. validation; E. P. K., N. J., and S. P. R. investigation; E. P. K., Y. C., and C. R. B. writing—original draft; N. J., S. P. R., and C. R. B. data curation; E. P. K., N. J., S. P. R., S. W., Y. C., and C. R. B. writing—review and editing; S. W. and Y. C. software; S. W., Y. C., and C. R. B. formal analysis; S. W., Y. C., and C. R. B. visualization; Y. C. and C. R. B. supervision; C. R. B. project administration; C. R. B. funding acquisition.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: CVD, cardiovascular disease; EGF, Epidermal Growth Factor; GI, gastrointestinal; MAP, multi-analyte profile; P/S, plasma/serum; PAV, pre-analytical variable.

Received March 25, 2022, and in revised form, August 12, 2022
Published, MCPRO Papers in Press, September 28, 2022, <https://doi.org/10.1016/j.mcpro.2022.100420>

REFERENCES

- Poste, G., Compton, C. C., and Barker, A. D. (2015) The national biomarker development alliance: confronting the poor productivity of biomarker research and development. *Exp. Rev. Mol. Diagn.* **15**, 211–218
- Lim, M. D., Dickherber, A., and Compton, C. C. (2011) Before you analyze a human specimen, think quality, variability, and bias. *Anal. Chem.* **83**, 8–13
- Lippi, G., Chance, J. J., Church, S., Dazzi, P., Fontana, R., Giavarina, D., et al. (2011) Preanalytical quality improvement: from dream to reality. *Clin. Chem. Lab. Med.* **49**, 1113–1126
- Ransohoff, D. F., and Gourlay, M. L. (2010) Sources of bias in specimens for research about molecular markers for cancer. *J. Clin. Oncol.* **28**, 698–704
- Compton, C. (2007) Getting to personalized cancer medicine: taking out the garbage. *Cancer* **110**, 1641–1643
- Lippi, G., Guidi, G. C., Mattiuzzi, C., and Plebani, M. (2006) Preanalytical variability: the dark side of the moon in laboratory testing. *Clin. Chem. Lab. Med.* **44**, 358–365
- McLerran, D., Grizzle, W. E., Feng, Z., Bigbee, W. L., Banez, L. L., Cazares, L. H., et al. (2008) Analytical validation of serum proteomic profiling for diagnosis of prostate cancer: sources of sample bias. *Clin. Chem.* **54**, 44–52
- Ellervik, C., and Vaught, J. (2015) Preanalytical variables affecting the integrity of human biospecimens in biobanking. *Clin. Chem.* **61**, 914–934
- Khan, J., Lieberman, J. A., and Lockwood, C. M. (2017) Variability in, variability out: best practice recommendations to standardize pre-analytical variables in the detection of circulating and tissue microRNAs. *Clin. Chem. Lab. Med.* **55**, 608–621
- Hewitt, S. M., Badve, S. S., and True, L. D. (2012) Impact of preanalytic factors on the design and application of integral biomarkers for directing patient therapy. *Clin. Cancer Res.* **18**, 1524–1530
- Agrawal, L., Engel, K. B., Greytak, S. R., and Moore, H. M. (2018) Understanding preanalytical variables and their effects on clinical biomarkers of oncology and immunotherapy. *Semin. Cancer Biol.* **52**, 26–38
- Moore, H. M., Compton, C. C., Alper, J., and Vaught, J. B. (2011) International approaches to advancing biospecimen science. *Cancer Epidemiol. Biomarkers Prev.* **20**, 729–732
- Betsou, F., Barnes, R., Burke, T., Coppola, D., Desouza, Y., Eliason, J., et al. (2009) Human biospecimen research: experimental protocol and quality control tools. *Cancer Epidemiol. Biomarkers Prev.* **18**, 1017–1025
- Bravo, M. I., Grancha, S., and Jorquera, J. I. (2006) Effect of temperature on plasma freezing under industrial conditions. *Pharmeuropa Sci. Notes* **2006**, 31–35
- Human plasma for fractionation, monograph 0853. (2005) *European Pharmacopoeia 5.0, Suppl 5.3, 1746-1747* Strasbourg. Council of Europe, France
- Daniels, J. R., Cao, Z., Maisha, M., Schnackenberg, L. K., Sun, J., Pence, L., et al. (2019) Stability of the human plasma proteome to pre-analytical variability as assessed by an aptamer-based approach. *J. Proteome Res.* **18**, 3661–3670
- Cao, Z., Kamlage, B., Wagner-Golbs, A., Maisha, M., Sun, J., Schnackenberg, L. K., et al. (2019) An integrated analysis of metabolites, peptides, and inflammation biomarkers for assessment of preanalytical variability of human plasma. *J. Proteome Res.* **18**, 2411–2421
- Oddoze, C., Lombard, E., and Portugal, H. (2012) Stability study of 81 analytes in human whole blood, in serum and in plasma. *Clin. Biochem.* **45**, 464–469
- Kamlage, B., Maldonado, S. G., Bethan, B., Peter, E., Schmitz, O., Liebenberg, V., et al. (2014) Quality markers addressing preanalytical variations of blood and plasma processing identified by broad and targeted metabolite profiling. *Clin. Chem.* **60**, 399–412
- Anton, G., Wilson, R., Yu, Z. H., Prehn, C., Zukunft, S., Adamski, J., et al. (2015) Pre-analytical sample quality: metabolite ratios as an intrinsic marker for prolonged room temperature exposure of serum samples. *PLoS One* **10**, e0121495
- Kamlage, B., Neuber, S., Bethan, B., Gonzalez Maldonado, S., Wagner-Golbs, A., Peter, E., et al. (2018) Impact of prolonged blood incubation and extended serum storage at room temperature on the human serum metabolome. *Metabolites* **8**, 6
- Jeffs, J. W., Jehanathan, N., Thibert, S. M. F., Ferdosi, S., Pham, L., Wilson, Z. T., et al. (2019) Delta-S-Cys-Albumin: a lab test that quantifies cumulative exposure of archived human blood plasma and serum samples to thawed conditions. *Mol. Cell Proteomics* **18**, 2121–2137
- Hu, Y., Mulot, C., Bourreau, C., Martin, D., Laurent-Puig, P., Radoi, L., et al. (2020) Biochemically tracked variability of blood plasma thawed-state exposure times in a multisite collection study. *Biopreserv. Biobank* **18**, 376–388
- Jeffs, J. W., Ferdosi, S., Yassine, H. N., and Borges, C. R. (2017) Ex vivo instability of glycated albumin: a role for autooxidative glycation. *Arch. Biochem. Biophys.* **629**, 36–42
- Hunt, J. V., Dean, R. T., and Wolff, S. P. (1988) Hydroxyl radical production and autooxidative glycosylation. Glucose autooxidation as the cause of protein damage in the experimental glycation model of diabetes mellitus and ageing. *Biochem. J.* **256**, 205–212
- Wolff, S. P., and Dean, R. T. (1987) Glucose autooxidation and protein modification. The potential role of 'autooxidative glycosylation' in diabetes. *Biochem. J.* **245**, 243–250
- Borges, C. R., Rehder, D. S., Jensen, S., Schaab, M. R., Sherma, N. D., Yassine, H., et al. (2014) Elevated plasma albumin and apolipoprotein A-I oxidation under suboptimal specimen storage conditions. *Mol. Cell Proteomics* **13**, 1890–1899
- Niederkofer, E. E., Kiernan, U. A., O'Rear, J., Menon, S., Saghir, S., Protter, A. A., et al. (2008) Detection of endogenous B-type natriuretic peptide at very low concentrations in patients with heart failure. *Circ. Heart Fail.* **1**, 258–264
- Craft, D., Yi, J., and Gelfand, C. A. (2009) Time-dependent and sample-to-sample variations in human plasma peptidome are both minimized through use of protease inhibitors. *Anal. Lett.* **42**, 1398–1406

30. Yi, J., Kim, C., and Gelfand, C. A. (2007) Inhibition of intrinsic proteolytic activities moderates preanalytical variability and instability of human plasma. *J. Proteome Res.* **6**, 1768–1781
31. Yi, J., Liu, Z., Craft, D., O'Mullan, P., Ju, G., and Gelfand, C. A. (2008) Intrinsic peptidase activity causes a sequential multi-step reaction (SMSR) in digestion of human plasma peptides. *J. Proteome Res.* **7**, 5112–5118
32. Hubel, A., Spindler, R., and Skubitz, A. P. (2014) Storage of human biospecimens: selection of the optimal storage temperature. *Biopreserv. Biobank* **12**, 165–175
33. Betsou, F., Gunter, E., Clements, J., DeSouza, Y., Goddard, K. A., Guadagni, F., et al. (2013) Identification of evidence-based biospecimen quality-control tools: a report of the international society for biological and environmental repositories (ISBER) biospecimen science working group. *J. Mol. Diagn.* **15**, 3–16
34. Boomsma, F., Alberts, G., van Eijk, L., Man in 't Veld, A. J., and Schalekamp, M. A. (1993) Optimal collection and storage conditions for catecholamine measurements in human plasma and urine. *Clin. Chem.* **39**, 2503–2508
35. Lawrence, J. M., Umekubo, M. A., Chiu, V., and Petitti, D. B. (2000) Split sample analysis of serum folate levels after 18 days in frozen storage. *Clin. Lab.* **46**, 483–486
36. Karlsen, A., Blomhoff, R., and Gundersen, T. E. (2007) Stability of whole blood and plasma ascorbic acid. *Eur. J. Clin. Nutr.* **61**, 1233–1236
37. Barden, A. E., Mas, E., Croft, K. D., Phillips, M., and Mori, T. A. (2014) Minimizing artifactual elevation of lipid peroxidation products (F2-isoprostanes) in plasma during collection and storage. *Anal. Biochem.* **449**, 129–131
38. Devanapalli, B., Bermingham, M. A., and Mahajan, D. (2002) Effect of long-term storage at -80 degrees C on the various lipid parameters in stored plasma samples. *Clin. Chim. Acta* **322**, 179–181
39. Sourvinou, I. S., Markou, A., and Lianidou, E. S. (2013) Quantification of circulating miRNAs in plasma effect of preanalytical and analytical parameters on their isolation and stability. *J. Mol. Diagn.* **15**, 827–834
40. Dedova, T., Grunow, D., Kappert, K., Flach, D., Tauber, R., and Blanchard, V. (2018) The effect of blood sampling and preanalytical processing on human N-glycome. *PLoS One* **13**, e0200507
41. Hu, Y., Ferdosi, S., Kapuruge, E. P., Diaz de Leon, J. A., Stucker, I., Radoi, L., et al. (2019) Diagnostic and prognostic performance of blood plasma glycan features in the women epidemiology lung cancer (WELCA) study. *J. Proteome Res.* **18**, 3985–3998
42. Amezcua Martin, M., Wuhler, M., and Falck, D. (2021) Serum and plasma immunoglobulin G Fc N-glycosylation is stable during storage. *J. Proteome Res.* **20**, 2935–2941
43. Kachur, A. V., Koch, C. J., and Biaglow, J. E. (1999) Mechanism of copper-catalyzed autooxidation of cysteine. *Free Radic. Res.* **31**, 23–34
44. Murphy, B. M., Swarts, S., Mueller, B. M., van der Geer, P., Manning, M. C., and Fitchmun, M. I. (2013) Protein instability following transport or storage on dry ice. *Nat. Met.* **10**, 278–279
45. Bakdash, J. Z., and Marusich, L. R. (2017) Repeated measures correlation. *Front. Psychol.* **8**, 456
46. Nakanishi, T., Hasuike, Y., Otaki, Y., Hama, Y., Nanami, M., Miyagawa, K., et al. (2003) Free cysteine is increased in plasma from hemodialysis patients. *Kidney Int.* **63**, 1137–1140
47. Nazha, B., Moussaly, E., Zaarour, M., Weerasinghe, C., and Azab, B. (2015) Hypoalbuminemia in colorectal cancer prognosis: nutritional marker or inflammatory surrogate? *World J. Gastrointest. Surg.* **7**, 370–377
48. Abudayyeh, A., Abdelrahim, M., and Salahudeen, A. (2014) Fluid and electrolyte abnormalities in patients with cancer. In: Finkel, K. W., ed. *Renal Disease in Cancer Patients*, Academic Press, NY: 167–182
49. Gupta, D., and Lis, C. G. (2010) Pretreatment serum albumin as a predictor of cancer survival: a systematic review of the epidemiological literature. *Nutr. J.* **9**, 69
50. Weaving, G., Batstone, G. F., and Jones, R. G. (2016) Age and sex variation in serum albumin concentration: an observational study. *Ann. Clin. Biochem.* **53**, 106–111
51. Jones, D. P., Mody, V. C., Jr., Carlson, J. L., Lynn, M. J., and Sternberg, P., Jr. (2002) Redox analysis of human plasma allows separation of pro-oxidant events of aging from decline in antioxidant defenses. *Free Radic. Biol. Med.* **33**, 1290–1300
52. Hellman, N. E., and Gitlin, J. D. (2002) Ceruloplasmin metabolism and function. *Annu. Rev. Nutr.* **22**, 439–458
53. Farrugia, A., Hill, R., Douglas, S., Karabagias, K., and Kleinig, A. (1992) Factor VIII/von Willebrand factor levels in plasma frozen to -30 degrees C in air or halogenated hydrocarbons. *Thromb. Res.* **68**, 97–102
54. Brogren, H., Wallmark, K., Deinum, J., Karlsson, L., and Jern, S. (2011) Platelets retain high levels of active plasminogen activator inhibitor 1. *PLoS One* **6**, e26762
55. Huebner, B. R., Moore, E. E., Moore, H. B., Stettler, G. R., Nunns, G. R., Lawson, P., et al. (2018) Thrombin provokes degranulation of platelet alpha-granules leading to the release of active plasminogen activator inhibitor-1 (PAI-1). *Shock (Augusta, GA)* **50**, 671–676
56. Gaza-Bulseco, G., Faldut, S., Hurkmans, K., Chumsae, C., and Liu, H. (2008) Effect of methionine oxidation of a recombinant monoclonal antibody on the binding affinity to protein A and protein G. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **870**, 55–62
57. Gao, J., Yao, Y., and Squier, T. C. (2001) Oxidatively modified calmodulin binds to the plasma membrane Ca-ATPase in a nonproductive and conformationally disordered complex. *Biophys. J.* **80**, 1791–1801
58. Nabuchi, Y., Fujiwara, E., Ueno, K., Kuboniwa, H., Asoh, Y., and Ushio, H. (1995) Oxidation of recombinant human parathyroid hormone: effect of oxidized position on the biological activity. *Pharm. Res.* **12**, 2049–2052
59. Rehder, D. S., and Borges, C. R. (2010) Cysteine sulfenic acid as an intermediate in disulfide bond formation and nonenzymatic protein folding. *Biochemistry* **49**, 7748–7755
60. Jung, K., Lein, M., Brux, B., Sinha, P., Schnorr, D., and Loening, S. A. (2000) Different stability of free and complexed prostate-specific antigen in serum in relation to specimen handling and storage conditions. *Clin. Chem. Lab. Med.* **38**, 1271–1275
61. de Jager, W., Bourcier, K., Rijkers, G. T., Prakken, B. J., and Seyfert-Margolis, V. (2009) Prerequisites for cytokine measurements in clinical trials with multiplex immunoassays. *BMC Immunol.* **10**, 52
62. Kisand, K., Kerna, I., Kumm, J., Jonsson, H., and Tamm, A. (2011) Impact of cryopreservation on serum concentration of matrix metalloproteinases (MMP)-7, TIMP-1, vascular growth factors (VEGF) and VEGF-R2 in Biobank samples. *Clin. Chem. Lab. Med.* **49**, 229–235
63. Kugler, K. G., Hackl, W. O., Mueller, L. A., Fiegl, H., Graber, A., and Pfeiffer, R. M. (2011) The impact of sample storage time on estimates of association in biomarker discovery studies. *J. Clin. Bioinform.* **1**, 9