clade.io

# CLADE™ QA Scan – Rapid analysis of pharmaceutical protein formulations

Pharmaceutical protein formulations (e.g., biologics) play a pivotal role in the treatment of a wide variety of diseases. Thus, the global biologics market has seen substantial growth, with its value reaching approximately USD 460 billion in 2022. Furthermore, forecasts indicate a projected compound annual growth rate of approximately 10% from 2023 to 2030.[1] Keeping pace with this growth and ensuring product quality necessitates effective monitoring of several critical quality attributes (CQA) of the product, critical material attributes (CMA) of materials in contact with the product, and key performance indicators (KPI) or critical process parameters (CPP) of the production process (for simplicity, these are summarized hereafter by the term Quality Attributes (QA)).

## Introduction

Most presently employed analytical methods for monitoring these parameters are complex to operate or offer limited insights. Typically, several different analytical methods must be used to reliably monitor all essential parameters. This application note demonstrates how CLADE™ QA Scan overcomes these limitations and analyzes many parameters simultaneously. Following a two-point calibration, CLADE™ QA Scan assesses protein, excipients, and polysorbate content, protein secondary structure, protein identity, and pH with a single measurement of a mixture, e.g., a protein formulation.

CLADE™ QA Scan acts as a complement to the MIRA Analyzer. Together with MIRA and AquaSpec™ technology, CLADE™ QA Scan performs multiple data-base-driven quantitative and qualitative anal-yses, based on a transparent workflow of machine learning algorithms.

This overall principle is illustrated in abstract form in **Figure 1**.

By eliminating the need for sample preparation, CLADE™ QA Scan enables rapid analysis in a single mea-surement, reducing time to just four minutes. Moreover, the advanced chemometric model creation capability empowers non-experts with the ability to develop accurate models with-in a single working day. Leveraging proprietary technology based on a comprehensive data-base of digital twins, CLADE™ QA Scan delivers reliable and reproducible results. This makes it suitable for widespread application in buffer and media control, formulation development, release, and stability testing. This application note illustrates the significant impact of this combined system on overall efficiency and quality within the an-alytical workflow.

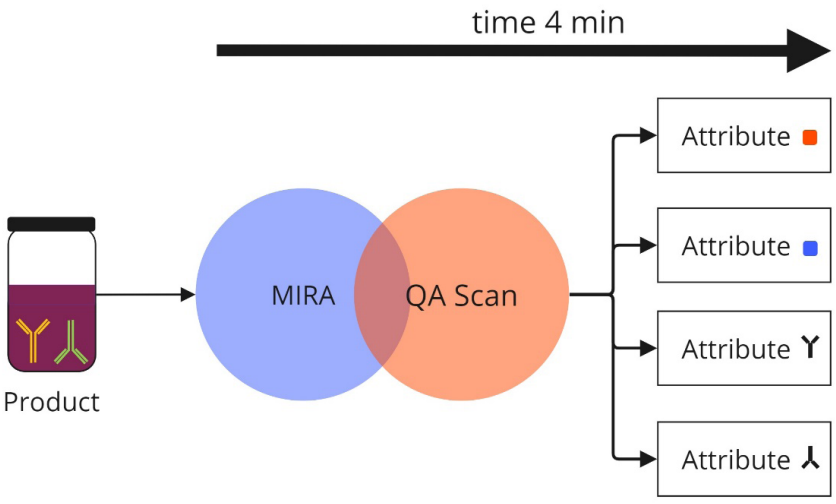


**Figure 1:** The CLADE™ QA Scan principle. First, a product comprising of various compo-nents undergoes measurement with the MIRA Analyzer. Subsequently, spectral evaluation using the CLADE™ QA Scan enables the determination of various quality attributes within the mixture.

[1] Biologics Market Size, Share & Trends Analysis Report, 2023 – 2030 (https://www.grandviewresearch.com/industry-analysis/biologics-market#:~:text=Report%20Overview,10.3%25%20from%202023%20to%202030)

## Material and Methods

The purpose of this study was solely to demon-strate the capabilities of CLADE™ QA Scan. A user will not need to perform DoE or similar studies. CLADE™ QA Scan capabilities were assessed using five different formulation buffers and six different proteins. In total ten formulations were prepared. The com-position of the formulations is shown in **Table 1**. These formulations were selected to represent different complexity levels for different ana-lytical questions. For protein and polysorbate quantification and protein secondary structure complexity levels were selected according to overlapping spectral features of different com-ponents in the sample spectrum. Overlapping spectral features can always be found in FTIR spectra of complex mixtures. Nonetheless each compound has its unique spectral fingerprint and with the high resolution, precision, and sensitivity of MIRA differentiation between com-pounds is possible using CLADE™ QA Scan.

| No. | Proteins | Excipients | pH |
|---|---|---|---|
| 1.1 | BSA | Citric Acid, PS80, Trehalose | 5.5 |
| 1.2 | Concanavalin A | Citric Acid, PS80, Trehalose | 5.5 |
| 2.1 | BSA | Phosphoric Acid, NaCl | 7 |
| 2.2 | Alcohol Dehydrogenase | Phosphoric Acid, NaCl | 7 |
| 3.1 | BSA | Histidine, PS20, Methionine, Sucrose, Trehalose, NaCl | 6 |
| 3.2 | Conalbumin | Histidine, PS20, Methionine, Sucrose, Trehalose, NaCl | 6 |
| 4.1 | BSA | Acetic Acid, PS20, Sorbitol | 5.2 |
| 4.2 | Lysozyme | Acetic Acid, PS20, Sorbitol | 5.2 |
| 5.1 | BSA | Histidine, Glycine, Arginine, Sucrose | 6.4 |
| 5.2 | Ovalbumin | Histidine, Glycine, Arginine, Sucrose | 6.4 |

**Table 1** – Prepared formulations to show capabilities of CLADE™ QA Scan

The concentrations of the formulation compo-nents were selected to represent a wide range of products and are listed in **Table 2**.

| Component | Range |
|---|---|
| Proteins | 8.0 – 110 mg/mL |
| Acids | 2.0 – 4.0 mg/mL |
| Sugars / Polyols | 20 – 85 mg/mL |
| Polysorbates | 0.1 – 0.6 mg/mL |
| Amino acids | 1.5 – 9.0 mg/mL |
| pH | 5.2 – 7.0 |

**Table 2** – Concentrations of formulation components

For each formulation calibration samples and validation samples were prepared. The validation set consisted of samples with the same com-position as the calibration samples, samples representing variation of the formulation with concentration variations between 0.2 % and 5.5 %, and samples representing variation of the pH value with variations up to 4 %.

All samples were measured in triplicates. Mea-surements with MIRA require no sample prepa-rations. The measurement time is as short as four minutes.

## Calibration

The outstanding feature of CLADE™ QA Scan is that only two samples are required for model calibration (e.g., the final formulation and the associated formulation buffer). Therefore, the calibration process for CLADE™ QA Scan is much more efficient and less time consuming than the prevailing algorithms often used in the field of infrared spectroscopy for quantitative analyses (e.g., PLSR, PCR, etc.).

What makes CLADE™ QA Scan so powerful is the use of spectra databases (based on MIRA data) employing innovative machine learning tech-nology. The construction of such databases is only possible due to the outstanding precision and therefore unmatched reproducibility of MIRA. These spectral databases contain a large number of digital twins of commonly used ex-cipients, buffers, and proteins. The databases are designed to account for naturally occur-ring variations in the manufacturing process (e.g., concentration fluctuations, pH, changes in secondary structure, etc.). In seamlessly expanded to include not commonly used components as needed.

When applied (after calibration), CLADE™ QA Scan only requires a single measurement of the protein formulation. Consequently, the model remains inherently robust and can be reliably applied to unknown data, within the calibration range. The spectra of four different formulations of BSA are shown in **Figure 2 on the left**. Together with the spectra of the respective formulation buffers (not shown), four individual CLADE™ QA scan models were created. Only two samples were needed for each individual model creation.

The validation results for protein quantification are summarized and shown in **Figure 2 on the right** where the values measured by CLADE™ QA Scan are plotted against the reference values (weighed-in concentration). For simplicity, the spectra and validation results of formulation one of BSA are not shown here as its protein content is sig-nificantly higher and therefore would lead to a distortion of the graph.

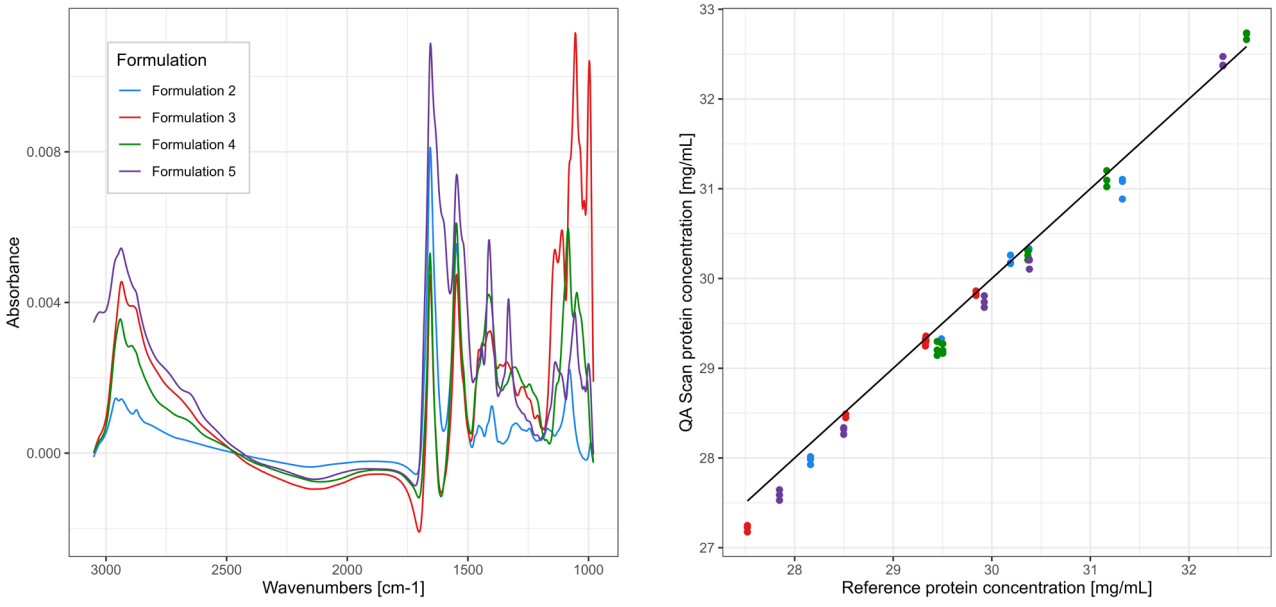


Figure 2 – left – spectra of formulations with BSA used for calibration – right – results of applying the calibrated CLADE™ QA Scan to the validation sets. For simplicity, the spectra and validation results of formulation 1 are not shown here (higher protein concentration would lead to a distortion of the graph).

## Results

Proteins as pharmaceutically active ingredients require extensive monitoring. Excipients in protein formulations stabilize the protein and guarantee its functionality. Therefore, the protein and excipients concentration must be monitored to guarantee quality and safety of the product. Mean validation results for protein, excipient (grouped into sub-groups), as well as polysor-bate quantification by CLADE™ QA Scan are illustrated in **Figure 3** and listed in **Table 3**.

The results show that with CLADE™ QA Scan, proteins and excipients can be predicted with low errors and great repeatability. Polysorbate is usually used in low concentrations. Here, the results show larger errors as formulations also contain excipients and proteins with sometimes large overlapping spectral bands with polysorbates. Further, polysorbates in formulations sometimes have ultralow concentrations down to 0.1 mg/mL. Nevertheless, CLADE™ QA Scan manages to meet accuracy requirements, as error rates of around 10% are still generally accepted in the industry.[2]

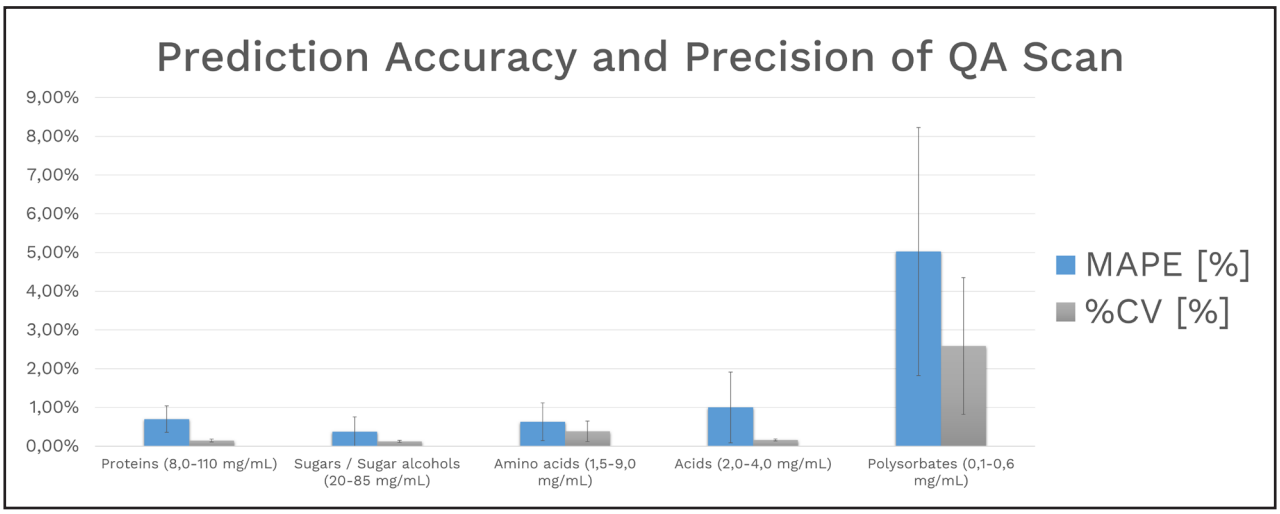


Figure 3 – Validation results for protein, excipients (grouped into sub-groups), and polysorbates quantification by CLADE™ QA Scan. Shown are the mean MAPE (mean average percent error) and the mean %CV for the determination of proteins, polysorbates and excipients. For simplification, the excipients were summarized in individual subgroups. In addition, the distribution of these parameters (MAPE and %CV) is shown by the respective error bars.

[2] Development and Validation of a Polysorbate 20 Assay in a Therapeutic Antibody Formulation by RP-HPLC and Charged Aerosol Detector (CAD) by Rowel Tobias, Alan Leslie, William Hanshaw, Paul Lightner, Glenn Petrie, Mike Whalon; 2016.

Proteins have specific and individual pH ranges in which they are pharmaceutically active. Therefore, the pH value is a crucial quality attribute and needs to be monitored.

Mean validation results for pH measurement by CLADE™ QA Scan are shown in **Table 3**. CLADE™ QA Scan can determine pH with low prediction error and high repeatability.

| Analysis | Range | MAPE [%] | %CV [%] |
|---|---|---|---|
| **Proteins** | | | |
| Protein content | 8.0 – 110 mg/mL | 0.70 | 0.14 |
| **Excipients** | | | |
| Sugars / Sugar alcohols | 20 – 85 mg/mL | 0.38 | 0.12 |
| Amino acids | 1.5 – 9.0 mg/mL | 0.63 | 0.38 |
| Acids | 2.0 – 4.0 mg/mL | 1.00 | 0.16 |
| **Surfactants** | | | |
| PS20/PS80 | 0.1 – 0.6 mg/mL | 5.03 | 2.59 |
| | | **MAE [-]** | **SD (standard deviation) [-]** |
| **pH** | 5.2 – 7.0 | < 0.1 | < 0.1 |

Table 3 – Mean validation results for simultaneous determination of several parameters in a protein formulation by CLADE™ QA Scan.

Secondary structure (alpha-helix and beta-sheet content) provides information regarding the three-dimensional structure of the protein. Differences between the secondary structure and that of reference values can be an indication of degradation, incorrect folding of a protein or contamination with other proteins. Validation results for protein secondary structure determined by CLADE™ QA Scan are shown in **Table 4**. The secondary structure (including changes) for all proteins can be measured with low standard deviations. Prediction results for BSA in different formulations are equal for alpha-helix for all formulations with 59 % and for beta-sheet on average 1.5 % with small deviations.

| No. | Protein | α-helix | | β-sheet | |
|---|---|---|---|---|---|
| | | Ratio [%] | SD [%] | Ratio [%] | SD [%] |
| 1.1 | BSA | 59 | << 1 | 1.5 | < 0.1 |
| 1.2 | Concanavalin A | 0.03 | 0.02 | 50 | << 1 |
| 2.2 | Alcohol Dehydrogenase | 25 | << 1 | 24 | << 1 |
| 3.2 | Conalbumin | 35 | << 1 | 15 | << 1 |
| 4.2 | Lysozyme | 44 | << 1 | 1.8 | < 0.1 |
| 5.2 | Ovalbumin | 26 | << 1 | 22 | << 1 |

Table 4 – Validation results for protein secondary structure with mean of predicted ratio of secondary structure and SD (standard deviation regarding repeated measurements)

The identification of proteins (protein ID) is crucial to make sure the right protein was produced, shipped, and/or used to prepare the formulation. For this, QA Scan compares the protein spectrum in a measured sample to a reference generated via calibration. Validation results for protein identification by QA Scan are shown in **Table 5**. Protein ID can be measured for all proteins and formulations with small standard deviation regarding repeated measurements. For simplification, BSA is listed only once because all results were comparably good despite the degree of complexity.

| No. | Protein | Mean ID [-] | SD [-] |
|---|---|---|---|
| 1.1 | BSA | 0.999 | < 0.001 |
| 1.2 | Concanavalin A | 0.998 | < 0.001 |
| 2.2 | Alcohol Dehydrogenase | 0.997 | < 0.001 |
| 3.2 | Conalbumin | 0.998 | < 0.001 |
| 4.2 | Lysozyme | 0.998 | < 0.001 |
| 5.2 | Ovalbumin | 0.998 | < 0.001 |

Table 5 – Validation results for protein ID with mean ID and SD (standard deviation regarding repeated measurements)

CLADE™ QA Scan enables easy measurement of quality attributes of pharmaceutical protein formula-tions. Current analytical workflows usually re-quire a wide range of elaborate and time-con-suming analytical steps. By contrast, MIRA and CLADE™ QA Scan significantly facilitate the analysis of protein, excipients, and polysorbate content, protein secondary structure, protein identity, and pH is significantly facilitated. It enables the user to get results for all these attributes in one sample measurement, with low prediction errors and high reproducibility. Thus, CLADE™ QA Scan simultaneously addresses currently established techniques such as UV-VIS spectrophotome-try, peptide mapping (HPLC), circular dichroism spectroscopy, IC, and pH meter. CLADE™ QA Scan users perform more efficient analysis, shorten devel-opment cycles, benefit from rapid release test-ing, and ultimately get products to market faster.

clade.io

CLADE

# Contact us to learn more!


✔ https://clade.io/

✔ Gürkan Korkmaz
Head of Sales & Marketing
Email: guerkan.korkmaz@clade.io
Tel: +49 176 126 67 004

✔ CLADE GmbH
Schelztorstraße 54-56
73728 Esslingen am Neckar
Germany