APPLICATION NOTE

nicoya

Relatively Speaking: Secondary Structure Decomposition and HOS Comparison Analysis for CD Comparability Studies

Introduction

Circular dichroism (CD) spectroscopy is a well-established technique for the structural characterization of chiral molecules and has become an indispensable part in the biophysical toolbox for biotherapeutics development. Particularly, CD spectroscopy provides information about both secondary and tertiary structure of antibody-based therapeutics, including biosimilars, and as such is essential for controlling higher order structure (HOS) as a critical quality attribute (CQA) within a totality-of-evidence approach.

However, CD data are often still assessed only qualitatively by subjecting spectra to visual evaluation. To eliminate user bias, spectral differences should instead be quantified by a reproducible mathematical approach which eliminates guesswork and scrutinizes results against a well-defined acceptance criterion that informs whether these differences are statistically relevant or not.[1-3]

While HOS comparison analysis has been developed with this dedicated purpose in mind and is increasingly accepted by the industry, many practitioners of CD spectroscopy still rely on traditional secondary structure decomposition analysis of far-UV CD data, which primarily yields absolute numbers on the fractions of secondary structure. As regulatory authorities still demand secondary structure decomposition in support of a totality-of-evidence approach as well, analysts may be tempted to derive conclusions about small relative differences in HOS from this type of analysis despite its limited precision. The uncertainty for secondary structure content obtained by this analysis approach is typically on the order of 5–10% or higher, e.g., due to spectral contributions that cannot be attributed to secondary structure.[4,5]

This begs the question how a direct approach using HOS comparison analysis, which yields values for similarity assessment without intermediate steps, and an indirect approach, which derives conclusions about relative differences from absolute results obtained by secondary structure decomposition, compare to each other in their ability to reveal minor spectral differences between a reference and samples of high similarity. To shed light on this premise, we examine here an exemplary dataset for a biotherapeutic molecule spiked at different relative amounts with either BSA or higher molecular weight species to represent contamination with a foreign molecule and increased aggregate content, respectively.

Results & Discussion

Averaged and baseline-corrected far-UV spectra for all sample replicates are shown in Figure 1, including CD (Figure 1a) and absorbance spectra (Figure 1b). Absorbance showed some variation between about 0.8 A.U. and 1.2 A.U., indicating slight differences in concentration between sample replicates. To account for these differences for subsequent analysis, CD data were normalized by absorbance, using average absorbance in the range (191 ± 2) nm.

Overall, spectral differences for some samples were immediately apparent by visual assessment, with spectral differences spanning up to 2 mdeg around the main peak at ~216 nm. Quantitative data analysis confirmed this initial visual inspection but provided more detail to reveal differences that elude even the eyes of a highly experienced data analyst (Figure 2).

Differences between reference and sample data were quantitatively assessed in terms of Z-scores for both HOS comparison analysis and secondary structure decomposition, where Z-scores >2 generally indicate statistically significant differences as such values do not fall inside a ±2xSD quality range, i.e., an interval that corresponds to two times the reference standard deviation above and below the mean.

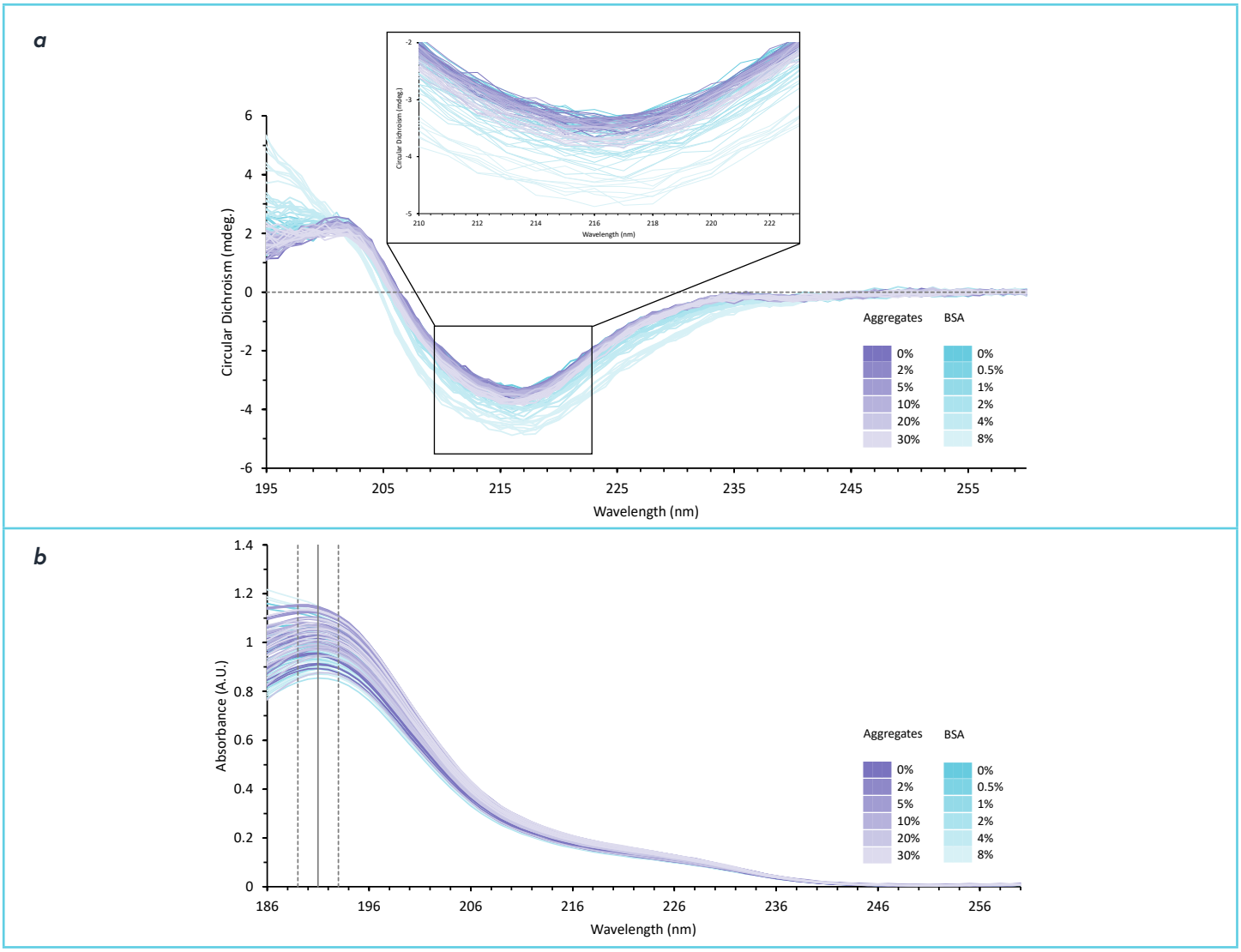

Figure 1: *Averaged and baseline-corrected far-UV data for samples spiked with different amounts of BSA (blue) or aggregate content (purple). Each panel shows a total of 108 traces for 12 samples with 9 replicates each; a) normalized circular dichroism spectra and b) absorbance spectra. Absorbance averaged across (191 ± 2) nm (black lines) was used for normalization of CD spectra.*

Similarity values from HOS comparison analysis were obtained using multiple alternative comparison methods, and results from secondary structure decomposition analysis were obtained based on different combinations of algorithms and sets of reference CD spectra (for more details see Methods section).

Highly Similar Samples

As expected, the control sample, which was identical to the reference, yielded Z-scores below 2, i.e., within the chosen ±2xSD quality range, independent of the analysis approach. Similarly, differences introduced by an aggregate content of up to 5% were too small to be captured by any of the analysis methods, making the corresponding sample spectra virtually indistinguishable from the reference spectrum, both by visual and quantitative assessment.

Highly Different Samples

On the other hand, Z-scores above 2, i.e., outside the chosen quality range, were obtained with all analysis approaches for 8% and 4% BSA and, with one exception (secondary structure decomposition with SELCON3 and SDP42 reference set), for 2% BSA. The differences between the reference and these samples were large enough that a visual gap of up to ~1 mdeg is evident at the main CD peak in the spectral data. However, spectral differences of this magnitude might still not prompt some scientists to conclude a structural difference, as can be evidenced from examples in the literature.[6-9]

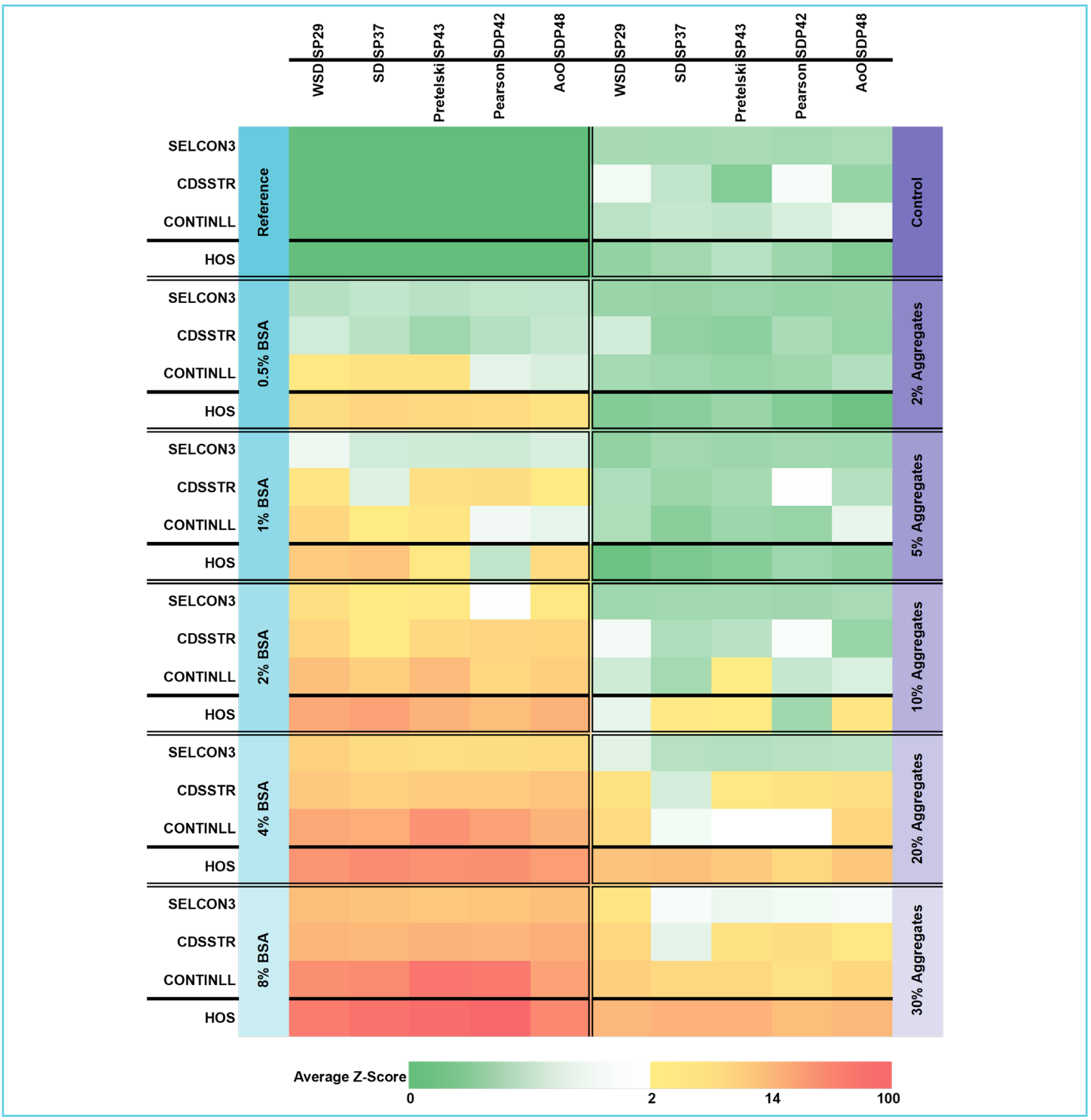

Figure 2: *Heat map showing mean Z-scores for samples spiked with different amounts of BSA (blue) or aggregate content (purple). For each sample, top three rows show values obtained with secondary structure decomposition algorithms (SELCON3, CDSSTR, CONTINLL), where columns correspond to the reference set used (SP29, SP37, SP43, SDP42, SDP48); bottom row shows values obtained with HOS comparison analysis, where columns correspond to the comparison method used (Weighted Spectral Difference (WSD), Spectral Difference (SD), Prestrelski Correlation Coefficient, Pearson Correlation Coefficient, Area of Overlap (AoO)). Colouring on a linear scale from green to white indicates values within the quality range of ±2xSD, and colouring on a logarithmic scale from yellow to red indicates values outside the quality range; i.e., white fields indicate values that barely fail the criterion for spectral difference (i.e., Z-scores barely within the quality range) whereas yellow fields indicate values that barely meet the criterion (i.e., Z-scores barely outside the quality range).*

Minor Differences

A review of the quantitative results becomes more interesting when looking at samples with minor differences that yielded Z-scores closer to the limits of the quality range, such as 0.5% BSA. Here, all HOS comparison methods yielded Z-scores >2, clearly indicating that samples were significantly different to the reference. Secondary structure decomposition, however, in most cases yielded Z-scores <2; both SELCON3 and CDSSTR failed to reveal differences independent of the reference set used, and CONTINLL only yielded Z-scores >2 for three out of the five reference sets tested.

For 1% BSA, all HOS comparison methods except Pearson Correlation Coefficient yielded Z-scores >2, while secondary structure decomposition algorithms CONTINLL and CDSSTR did so for three and four of the five tested reference sets, respectively. SELCON3 did not indicate a difference for any spectral reference set.

In the case of 10% aggregate content, three of the five HOS comparison methods (Prestrelski Correlation Coefficient, Spectral Difference, and Area of Overlap) indicated a difference, but only one combination of secondary structure decomposition algorithm and reference set (CONTINLL with SP43) indicated a difference.

For both 20% and 30% aggregate content, all HOS comparison analysis methods indicated a difference, but secondary structure decomposition did so only for certain combinations of algorithm and reference set. Again, SELCON3 performed most poorly, yielding a Z-score >2 for only the samples with 30% aggregate content with SP29 and suggesting no difference for 20% aggregate content no matter which reference set was used.

Nature of structural changes

As secondary structure decomposition primarily quantifies secondary structure content, the nature of structural differences detected was also examined. Figure 3 shows for each combination of algorithm and reference set for which of the different secondary structure elements Z-scores >2 were obtained and whether secondary structure content increased or decreased.

As can be expected for spiking with BSA, which has predominantly α-helical structure, an increase in regular or distorted helix (green) yielded the largest Z-score in most cases. For BSA contents between 1% and 4%, changes in sheet content (red) gave the largest Z-scores in some cases, particularly for 2% BSA using CDSSTR. This was mostly due to a reduction in β-sheet content, which is in line with a reduced relative content of the antibody molecule in the sample. However, SELCON3 suggested an increase in sheet structure for 4% and 8% BSA, and both an increase (2% BSA, CDSSTR and CONTINLL) and decrease (8% BSA, SELCON3 and CONTINLL) was suggested for turns and unordered structure with some reference sets.

Conclusion

Here, an exemplary data set including far-UV CD spectra for biotherapeutics samples spiked with different amounts of BSA and aggregate content was analysed by both HOS comparison analysis and secondary structure decomposition. Results illustrate that HOS comparison analysis is superior to secondary structure decomposition analysis in the detection of minor spectral differences and, in turn, in providing an objective, statistically validated assessment of relative changes in HOS to support informed decisions in biotherapeutics development. Namely, HOS comparison analysis yielded overall larger Z-scores across different comparison methods than secondary structure decomposition analysis, demonstrating its unmatched ability to detect small differences between samples.

However, HOS comparison analysis remains the preferred approach for similar comparative studies for several additional reasons. HOS comparison analysis is independent of spectral range and does not depend as much on low-wavelength data like secondary structure decomposition does. It can also be applied to, e.g., near-UV CD data informing about changes in tertiary structure, which is essential in a totality of evidence approach required for biosimilar approval by regulatory authorities,[11-22] or even to data obtained with techniques other than CD spectroscopy, e.g., infrared spectroscopy.[23,24]

Moreover, HOS comparison analysis is less dependent on accurate input information such as sample concentration and path length because an internal reference is used and normalization accounts for small differences in concentration.

HOS comparison analysis is also less prone to user bias because comparison methods perform similarly, whereas there are greater differences resulting from the choice of algorithm and reference set for secondary structure decomposition.

Finally, the mathematical foundation of comparison methods is simpler and is easily implemented as described in the literature,[1,2-23-27] whereas some tools for secondary structure decomposition are less transparent, most are only available online, and tools are almost always used in a black-box approach.

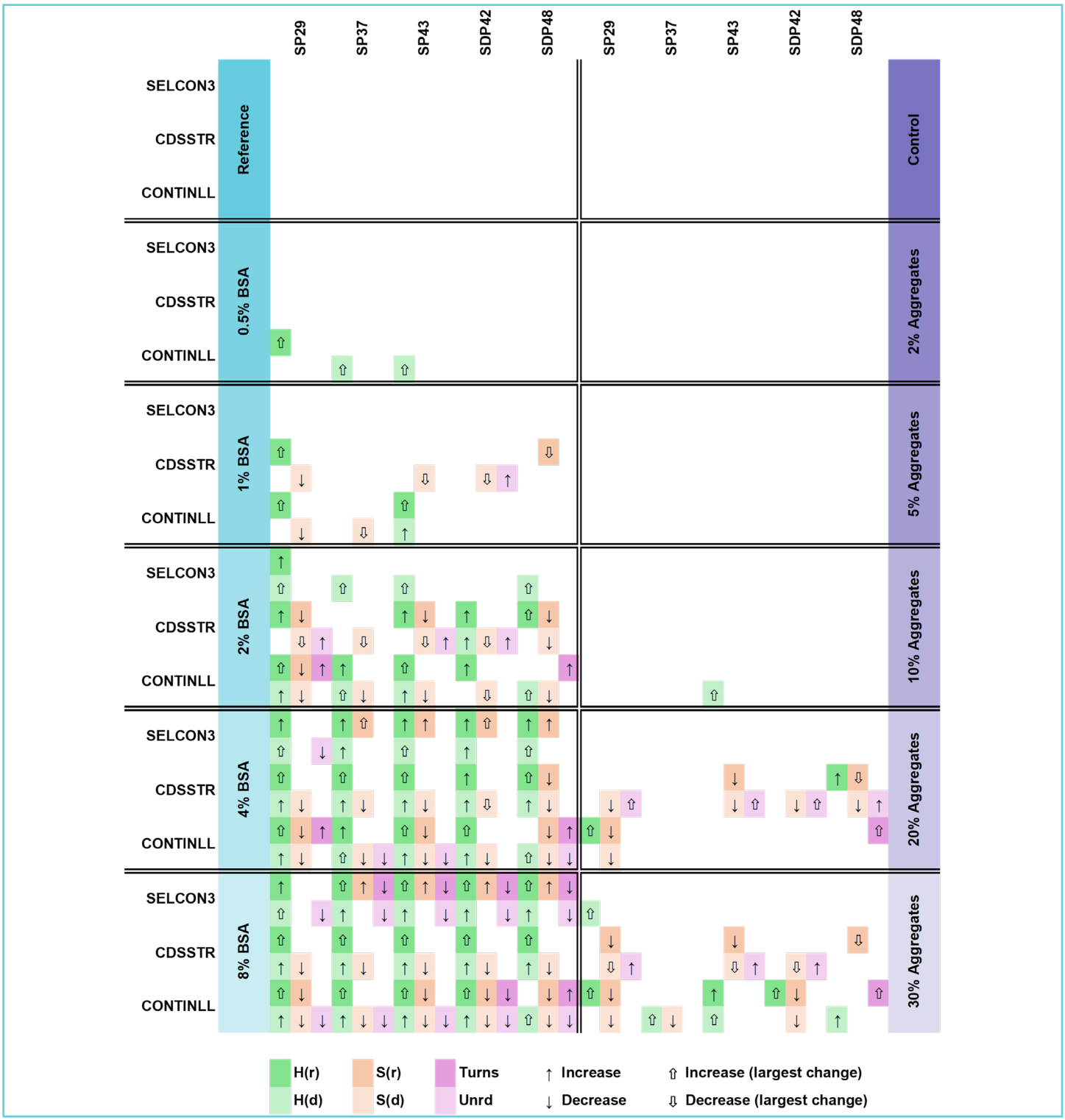

Figure 3: *Changes in secondary structure suggested by secondary structure decomposition for samples spiked with different amounts of BSA (blue) or aggregate content (purple). For each sample, rows indicate results obtained with secondary structure decomposition algorithms (SELCON3, CDSSTR, CONTINLL), where columns correspond to the reference set used (SP29, SP37, SP43, SDP42, SDP48). Changes are only indicated where the absolute Z-score was >2. Arrows indicate increase (↑) or decrease (↓) in secondary structure elements (H(r): regular helix, dark green; H(d): distorted helix, light green; S(r): regular sheet, dark red; S(d): distorted sheet, light red; Turn: turns, dark pink; and Unrd: unordered, light pink). Thick arrows (⇧ or ⇩) indicate largest changes across the six secondary structure elements for each combination of sample, algorithm, and reference set.*

Methods

Spectra were measured using a Chirascan Q100 automated CD spectrometer (formerly Applied Photophysics®, now part of Nicoya) equipped with a flow cell with an optical path length of 0.1 mm.

Samples used to obtain the data set contained a biotherapeutic molecule spiked with either BSA or higher molecular weight aggregates. All samples were used as-is at a nominal concentration of 1 mg/mL and transferred to a 96-well microwell plate for automated runs.

Sample Preparation

Before preparation of microplates, samples and buffer were subjected to a vacuum (~-95 kPa) for ~30 min to remove excess dissolved gas. After preparation of a microplate, it was centrifuged at 3000 rpm for 5 min (Heraeus Megafuge 11, Rotor T41) to remove bubbles.

Microplates were prepared in a layout with alternating buffer/sample pairs to account for potential effects of different sample storage times before acquisition and to prevent cross-correlation in the statistical analysis. Replicates were arranged in separate blocks of wells, and buffer/sample pairs were arranged in a randomized order. Samples of BSA and aggregates are listed in Tables 1 & 2, respectively. The layout of the microplate is shown in Figure 4.

Data Acquisition

Absorbance and CD spectra were obtained simultaneously with acquisition settings as specified in Table 3 for the 12 different samples, with a total of 9 replicate measurements per sample obtained in groups of 3 across 3 separate automated runs. For each replicate, an individual buffer blank was obtained to be used for baseline correction, and 3 repeat scans were obtained in each acquisition. In consequence, the raw data set contained a total of 648 single scans.

Between consecutive measurements, the flow cell was cleaned thoroughly using the default cleaning protocol including flushing with detergent (10% Decon 90), water, and acetone, followed by drying with a vacuum pump.

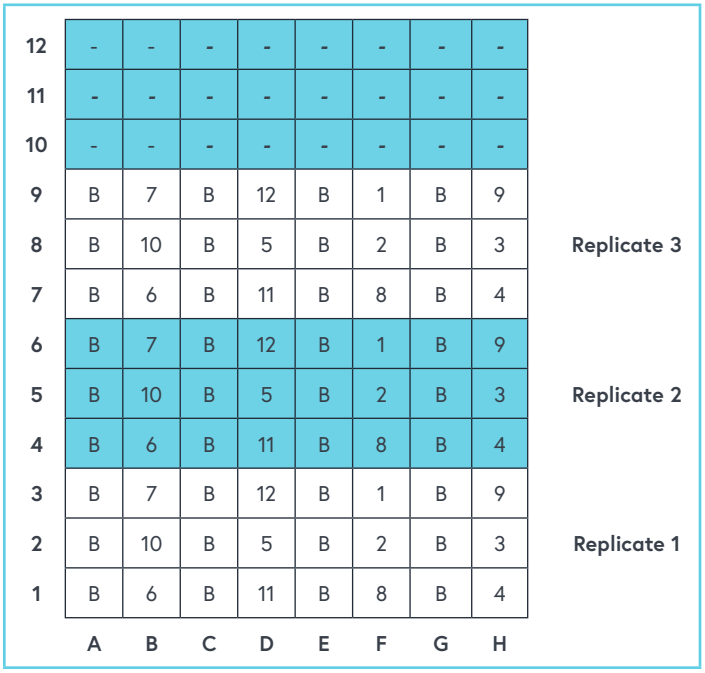

Figure 4: *Microplate layout. B: Buffer, 1 to 12: Samples as labelled in Tables 1 & 2.*

Table 1: *List of BSA samples*

Sample Number	Sample Type
1	Reference
2	0.5% BSA
3	1% BSA
4	2% BSA
5	4% BSA
6	8% BSA

Table 2: *List of aggregate samples*

Sample Number	Sample Type
7	Control
8	2% Aggregates
9	5% Aggregates
10	10% Aggregates
11	20% Aggregates
12	30% Aggregates

Table 3: Acquisition settings for CD & absorbance spectra

Parameter	Setting
Wavelength range (nm)	180 to 260
Step size (nm)	1
Bandwidth (nm)	1
Time-per-point (s)	1
Measurement temperature (°C)	20
Repeat scans	3

Data Processing

Basic data processing was carried out using the default processing script of the autosampler software, which includes averaging of repeat scans and subtraction of the buffer average from the corresponding sample average for each buffer/sample pair.

Before data were subjected to either HOS comparison analysis or secondary structure decomposition, the far-UV CD spectra were normalized by absorbance to account for slight differences in sample concentrations. To this end, CD values were divided by the average absorbance across (191 ± 2) nm, i.e., making use of a small interval around the main carbonyl absorbance peak to minimise the effect of system noise. Moreover, CD data below 195 nm were characterized by elevated signal noise due to detector voltages >750 V and were thus excluded from further analysis.

For analysis with CDPro, spectra were first converted from millidegrees to mean residue molar absorbance, $\Delta\varepsilon$ (M^{-1} cm^{-1}) based on a nominal concentration of 1 mg/mL, an optical path length of 0.01 cm, and a molecular weight of 150 kDa.

Data Analysis

HOS comparison analysis was carried out using qBiC software version 1.1.2 to yield similarity values for five different comparison methods including Pearson and Prestrelski Correlation Coefficients, Spectral Difference and Weighted Spectral Difference, as well as Area of Overlap.

Secondary structure decomposition was carried out using CDPro to yield relative contents of six different secondary structure elements (H(r): regular helix; H(d): distorted helix; S(r): regular sheet; S(d): distorted sheet; Turn: turns; and Unrd: unordered).[28] Analysis with CDPro was carried out for the 12 samples with 9 replicates each using all three included algorithms (SELCON3, CONTINLL, CDSSTR) and five different reference sets (SP29, SP37, SP43, SDP42, SDP48); a custom script was used to automate generation of input files and execution of a total of 1620 secondary structure decomposition calculations.

To enable a direct comparison between results from HOS comparison analysis and secondary structure decomposition, results were expressed in terms of Z-scores. For HOS comparison analysis, Z-scores were directly obtained from the qBiC software. For secondary structure decomposition, Z-scores were calculated from results expressing fractions of secondary structure. That is, for each secondary structure element, mean and standard deviation were calculated from reference replicates. Then, for any sample replicate, the reference mean was subtracted from the sample replicate secondary structure fraction and the resulting difference divided by the reference standard deviation to obtain the sample replicate Z-score.

As either an increase or decrease in secondary structure may indicate a significant conformational change, only absolutes of Z-scores from secondary structure decomposition were used for subsequent comparison.

To simplify data interpretation, means of Z-scores were obtained by averaging across the replicates of each sample. For results from secondary structure decomposition, only the largest mean Z-score obtained across the six different secondary structure elements was further used for any given combination of sample, algorithm, and reference set.

References

1. Jones, C. "Impact of Imperfect Data on the Performance of Algorithms to Compare Near-Ultraviolet Circular Dichroism Spectra." Applied Spectroscopy 75, no. 7 (2021): 857–66. https://doi.org/10.1177/0003702821992370.
2. Jones, C. "Glycoconjugate Vaccine Batch Consistency Assessed by Objective Comparison of Circular Dichroism Spectra." Journal of Pharmaceutical and Biomedical Analysis 191 (2020): 113571. https://doi.org/10.1016/j.jpba.2020.113571.
3. Jones, C. "Circular Dichroism of Biopharmaceutical Proteins in a Quality-Regulated Environment." Journal of Pharmaceutical and Biomedical Analysis 219 (2022): 114945.
4. Nagy, G., and H. Grubmüller. "How Accurate Is Circular Dichroism-Based Model Validation?" European Biophysics Journal 49, no. 6 (2020): 497–510. https://doi.org/10.1007/s00249-020-01457-6.
5. Micsonai, A., F. Wien, L. Kernya, Y.-H. Lee, Y. Goto, M. Réfrégiers, et al. "Accurate Secondary Structure Prediction and Fold Recognition for Circular Dichroism Spectroscopy." Proceedings of the National Academy of Sciences 112, no. 24 (2015): E3095–E3103. https://doi.org/10.1073/pnas.1500851112.
6. Bandyopadhyay, S., M. Mahajan, T. Mehta, A. K. Singh, A. K. Gupta, A. Parikh, et al. "Physicochemical and Functional Characterization of a Biosimilar Adalimumab ZRC-3197." Biosimilars 5 (2015): 1–18.
7. Al-Kinani, K. K., Z. E. Jassim, S. S. Taher, and A. A. Hussein. "Comparative Biosimilar Quality Studies Between a Rituximab Product and MabThera." Journal of Advanced Pharmacy Education and Research 11, no. 4 (2021): 41–49.
8. Liu, Y., J. Xie, Z. Li, X. Mei, D. Cao, S. Li, et al. "Demonstration of Physicochemical and Functional Similarity of the Biosimilar BAT1806/BIIB800 to Reference Tocilizumab." BioDrugs 38 (2024): 571–88.
9. Lee, H., J. Huh, D. Kim, S. Lee, J. Lee, J. Lee, et al. "Analytical Characterization for Similarity Assessment Between an Aflibercept Biosimilar SB15 and Reference Product (Eylea®)." Ophthalmology and Therapy 13 (2024): 2209–25.
10. Micsonai, A., É. Moussong, F. Wien, E. Boros, H. Vadászi, N. Murvai, et al. "BeStSel: Webserver for Secondary Structure and Fold Prediction for Protein CD Spectroscopy." Nucleic Acids Research 50, no. W1 (2022): W90–W98. https://doi.org/10.1093/nar/gkac330.
11. Barnett, G. V., G. Balakrishnan, N. Chennamsetty, L. Hoffman, J. Bongers, L. Tao, et al. "Probing the Tryptophan Environment in Therapeutic Proteins: Implications for Higher Order Structure on Tryptophan Oxidation." Journal of Pharmaceutical Sciences 108, no. 6 (2019): 1944–52. https://doi.org/10.1016/j.xphs.2019.01.004.
12. Barnett, G. V., G. Balakrishnan, N. Chennamsetty, B. Meengs, J. Meyer, J. Bongers, et al. "Enhanced Precision of Circular Dichroism Spectral Measurements Permits Detection of Subtle Higher Order Structural Changes in Therapeutic Proteins." Journal of Pharmaceutical Sciences 107, no. 10 (2018): 2559–69. https://doi.org/10.1016/j.xphs.2018.06.020.
13. Hutterer, K. M., A. Polozova, S. Kuhns, H. J. McBride, X. Cao, and J. Liu. "Assessing Analytical and Functional Similarity of Proposed Amgen Biosimilar ABP 980 to Trastuzumab." BioDrugs 33, no. 3 (2019): 321–33. https://doi.org/10.1007/s40259-019-00350-9.
14. Jeong, Y. R., R. U. Jeong, J. H. Son, J. C. Kwon, S. Jung, M. A. Song, et al. "Comprehensive Physicochemical and Biological Characterization of the Proposed Biosimilar Darbepoetin Alfa, LBDE, and Its Originator Darbepoetin Alfa, NESP®." BioDrugs 32, no. 2 (2018): 153–68. https://doi.org/10.1007/s40259-018-0272-7.
15. Lee, J., H. A. Kang, J. S. Bae, K. D. Kim, K. H. Lee, K. J. Lim, et al. "Evaluation of Analytical Similarity Between Trastuzumab Biosimilar CT-P6 and Reference Product Using Statistical Analyses." mAbs 10, no. 4 (2018): 547–71. https://doi.org/10.1080/19420862.2018.1440170.
16. Lee, K. H., J. Lee, J. S. Bae, Y. J. Kim, H. A. Kang, S. H. Kim, et al. "Analytical Similarity Assessment of Rituximab Biosimilar CT-P10 to Reference Medicinal Product." mAbs 10, no. 3 (2018): 380–96. https://doi.org/10.1080/19420862.2018.1433976.
17. Li, M., V. A. Beaumont, S. Akbar, H. Duncan, A. Creasy, W. Wang, et al. "Comprehensive Characterization of Higher Order Structure Changes in Methionine Oxidized Monoclonal Antibodies via NMR Chemometric Analysis and Biophysical Approaches." mAbs 16, no. 1 (2024): 2292688.
18. Liu, J., T. Eris, C. Li, S. Cao, and S. Kuhns. "Assessing Analytical Similarity of Proposed Amgen Biosimilar ABP 501 to Adalimumab." BioDrugs 30, no. 4 (2016): 321–38. https://doi.org/10.1007/s40259-016-0184-3.
19. McAvan, B. S., L. A. Bowsher, T. Powell, J. F. O'Hara, M. Spitali, R. Goodacre, et al. "Raman Spectroscopy to Monitor Post-Translational Modifications and Degradation in Monoclonal Antibody Therapeutics." Analytical Chemistry 92, no. 15 (2020): 10381–89. https://doi.org/10.1021/acs.analchem.0c00627.
20. Miao, S., L. Fan, L. Zhao, D. Ding, X. Liu, H. Wang, etal. "Physicochemical and Biological Characterization of the Proposed Biosimilar Tocilizumab." BioMed Research International 2017 (2017): 1–13. https://doi.org/10.1155/2017/4926168.
21. Nair, R., N. Krishnan, V. Shenoy, and R. N. Seetharam. "Evaluation of Methods Employed in Establishing Preclinical Similarity of Adalimumab Biosimilars." Advances in Pharmacological and Pharmaceutical Sciences 2025, no. 1 (2025).
22. Saricay, Y., B. J. de Kort, H. Yigit-Gercek, and E. H. C. Dirksen. "A Multi-Angular View on the Impact of Protein Unfolding on Biophysical Structural Data." Analytical Biochemistry 630 (2021): 114331. https://doi.org/10.1016/j.ab.2021.114331.
23. Jiang, Y., C. Li, X. Nguyen, S. Muzammil, E. Towers, J. Gabrielson, et al. "Qualification of FTIR Spectroscopic Method for Protein Secondary Structural Analysis." Journal of Pharmaceutical Sciences 100, no. 11 (2011): 4631–41. https://doi.org/10.1002/jps.22620.
24. Kendrick, B. S., A. Dong, S. D. Allison, M. C. Manning, and J. F. Carpenter. "Quantitation of the Area of Overlap Between Second-Derivative Amide I Infrared Spectra to Determine the Structural Similarity of a Protein in Different States." Journal of Pharmaceutical Sciences 85 (1996): 155–58.
25. Li, C. H., X. Nguyen, L. Narhi, L. Chemmalil, E. Towers, S. Muzammil, et al. "Applications of Circular Dichroism (CD) for Structural Analysis of Proteins: Qualification of Near- and Far-UV CD for Protein Higher Order Structural Analysis." Journal of Pharmaceutical Sciences 100, no. 11 (2011): 4642–54. https://doi.org/10.1002/jps.22621.
26. Teska, B. M., C. Li, B. C. Winn, K. K. Arthur, Y. Jiang, and J. P. Gabrielson. "Comparison of Quantitative Spectral Similarity Analysis Methods for Protein Higher-Order Structure Confirmation." Analytical Biochemistry 434, no. 1 (2013): 153–65. https://doi.org/10.1016/j.ab.2012.11.024.
27. Dinh, N. N., B. C. Winn, K. K. Arthur, and J. P. Gabrielson. "Quantitative Spectral Comparison by Weighted Spectral Difference for Protein Higher Order Structure Confirmation." Analytical Biochemistry 464 (2014): 60–62. https://doi.org/10.1016/j.ab.2014.07.011.
28. Sreerama, N., and R. W. Woody. "Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set." Analytical Biochemistry 287, no. 2 (2000): 252–60. https://doi.org/10.1006/abio.2000.4880.