Waters™ | WYATT TECHNOLOGY

WHITE PAPER

WP3001: SEC-MALS and CG-MALS: Complementary Techniques to Characterize Protein-DNA Complexes

Sophia Kenrick, Ph.D., Waters | Wyatt Technology

Summary

Interactions between proteins and nucleic acids often result in binding stoichiometries greater than 1:1 and may exhibit cooperativity, allosteric hindrance, or other complex phenomena. Quantifying the affinity and stoichiometry of these biomolecular assemblies is key to understanding the mechanism of interaction and to developing therapeutics that target and modulate them.

Here, we present the characterization of protein-DNA complexes using two complementary multi-angle light scattering (MALS) techniques: 1) in combination with size exclusion chromatography (SEC-MALS) and protein conjugate analysis to measure the overall complex molar mass and fraction of DNA and 2) composition-gradient MALS (CG-MALS) to quantify the stoichiometry and affinity at each binding site in the complex formation.

Introduction

Protein-nucleic acid interactions are among the most fundamental biological processes. Examples include the integration of viral genes into the host DNA and DNA recombination during meiosis. Biophysical characterization of these molecules and their interactions are central to understanding how viruses invade our bodies and replicate within our cells, and to developing effective antiviral, anticancer and genetic therapeutics.

SEC-MALS

Multi-angle light scattering (MALS) enables direct measurement of the molar mass of biomolecules and complexes of biomolecules in solution. When included downstream of size exclusion chromatography (SEC), the combination of a MALS detector, UV detector, and differential refractive index (dRI) detector provides the absolute molar mass and molar mass distribution of each eluting peak, irrespective of conformation or non-ideal column interactions.

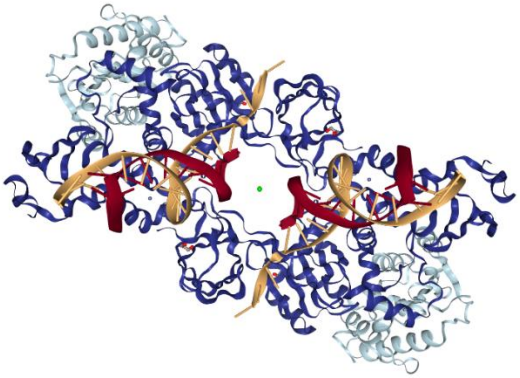


Figure 1: Top: PFV intasome (PDB ID: 3L2Q). Bottom: Cre-*loxP* synapse tetramer (PDB ID: 3MGV)

SEC-UV-MALS-dRI is especially useful for determining the molar mass of proteins conjugated with other macromolecules (e.g., pegylated proteins) or in noncovalent complexes with other species, such as protein-DNA complexes. The protein conjugate analysis in Wyatt's ASTRA® software for SEC-MALS determines the molar mass of the entire complex and the amount of modifier (e.g., DNA) associated with the protein monomer or oligomer.

CG-MALS

Composition-gradient MALS (CG-MALS) is a non-fractionating technique for quantifying the affinity and stoichiom-

etry of macromolecular interactions, without sample tagging, immobilization, or other modifications that could interfere with these phenomena. In a CG-MALS experiment, different ratios of protein and DNA are combined, and the resulting weight-average molar mass values are analyzed to determine the absolute stoichiometry and quantify the affinity at each binding site. This powerful technique is especially useful when the stoichiometry of the complexes is beyond 1:1 or when the molecules exhibit self-association in addition to hetero-association. CG-MALS can readily distinguish equivalent from non-equivalent binding sites or cooperative behavior.

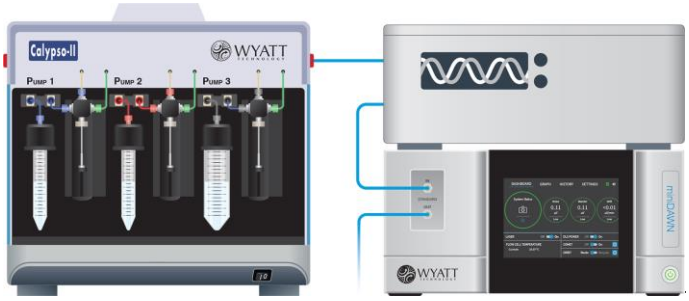


Figure 2: Schematic of CG-MALS hardware, including Calypso, inline concentration (UV) detector, and MALS detector

In this white paper, we highlight the application of SEC-MALS protein conjugate analysis for characterizing a complex of prototype foamy virus integrase (PFV IN) bound to U5 viral DNA. The complex was loaded on the SEC column, and the eluting peak confirms the protein-DNA stoichiometry in the SAXS/SANS structure of the PFV intasome (Figure 1, top). The original data were published in reference [1].

In addition, we characterize the interaction between Cre recombinase and a 34 bp *loxP* DNA sequence in two different pH buffers using CG-MALS. We confirm that at pH 7.5, two Cre recombinases bind cooperatively to the DNA recognition site, and this 2:1 complex dimerizes to form the synapse tetramer (Figure 1, bottom) while at pH 9.5, synapsis is lost.

Materials and Methods

Size-exclusion chromatography with multi-angle light scattering detection (SEC-MALS)

Size-exclusion chromatography was performed using a Superdex 200 10/300 GL column (GE Healthcare). The effluent of the column flowed through an inline UV detector, DAWN™ MALS detector, and Optilab™ differential refractive index detector (Figure 3).

Data were analyzed using the Protein Conjugate Analysis method in ASTRA. To determine the extinction coefficient of PFV IN, the protein was injected on the SEC-UV-MALS-RI system. ASTRA's UV Extinction from RI Peak method was applied using a *dn/dc* of 0.185 mL/g for the protein alone, resulting in a measured extinction coefficient of 0.75 $(g/L)^{-1}cm^{-1}$ at 280 nm for the protein. For the protein conjugate analysis, an extinction coefficient of 13.2 $(g/L)^{-1}cm^{-1}$ at 280 nm and a *dn/dc* value of 0.170 mL/g were used for the DNA ("modifier").

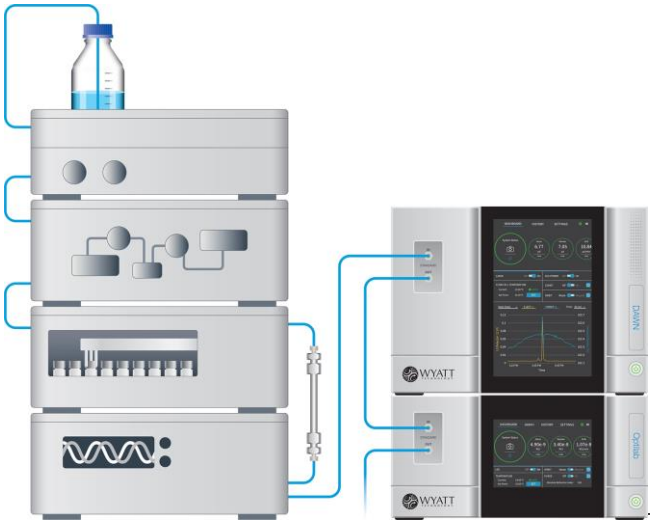


Figure 3: Schematic of SEC-UV-MALS-RI hardware for protein conjugate analysis, concentration (UV) detector, and MALS detector

Composition-gradient multi-angle light scattering (CG-MALS)

A Calypso™ composition-gradient system (Figure 2) was used to deliver protein-DNA mixtures to a DAWN MALS detector and inline UV absorbance detector (Dionex). UV absorbance data were measured at 280 nm and 260 nm simultaneously to confirm the stock concentration of each species. Each experiment consisted of two, single-component concentration gradients, to assess self-association of Cre and *loxP* individually, and a dual-component "crossover" gradient to assess hetero-association. After each composition, the flow was stopped for 1 minute (self-association) or up to 20 minutes (hetero-association) to allow the reactions to come to equilibrium within the MALS flow cell.

A typical experiment is shown in Figure 4 Cre and *loxP* were diluted in buffer to approximately 0.02 mg/mL and

0.025 mg/mL, respectively, and filtered to 0.02 µm prior to loading on the Calypso. Experiments at pH 7.5 were performed with HEPES buffer (20 mM HEPES, 150 mM NaCl, 2 mM DTT), and experiments at pH 9.5 were performed in bis-tris-phosphate (BTP) buffer (20 mM BTP, 150 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT). All buffers were filtered to 0.1 µm.

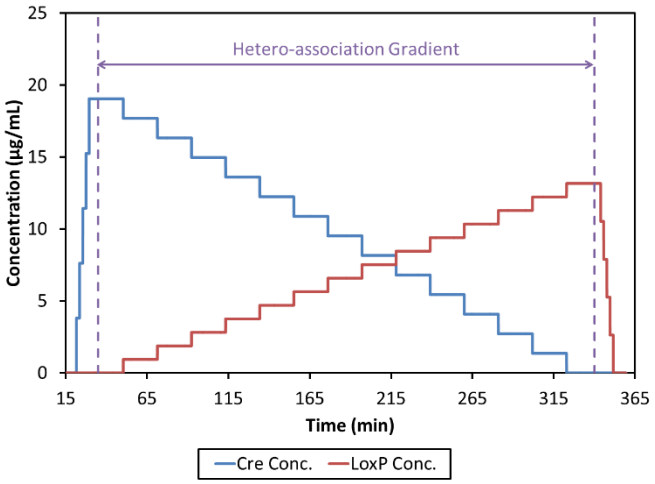


Figure 4: CG-MALS method including self-association gradients for Cre and *loxP* and hetero-association (crossover) gradient

Light scattering and composition data were fit to the appropriate model using the CALYPSO™ software to determine stoichiometry and equilibrium dissociation constant (K_d) at each binding site.

Results and Discussion

Conjugate analysis by SEC-MALS

ASTRA's protein conjugate analysis utilizes MALS, UV and dRI signals to quantify two-component systems. The combination of two simultaneous concentration detectors—UV and dRI—provides the weight fractions of protein and modifier at each eluting slice of a chromatogram. This analysis is possible when the two conjugated species have different extinction coefficients (ε), different specific refractive index increments (dn/dc), or both.

The addition of MALS data determines the molar mass of each constituent (protein and modifier) as well as the entire complex at each eluting slice. This analysis has been successfully applied to covalent protein-carbohydrate systems [2- 4] and noncovalently-bound complexes, such as lipid micelles associated with membrane proteins [5, 6]. Here, protein conjugate analysis was used to confirm that the protein-DNA stoichiometry seen in the crystal structure for the PFV intasome was representative of the complex in solution.

Figure 5 shows results of the measurement: the molar masses of the protein-DNA complex and of each component. The molar mass varies slightly across the peak, suggesting some on-column dissociation of the complex. Near the leading edge of the peak, the measured molar mass of the protein is 170 kDa, consistent with the crystal structure showing four PFV IV monomers in the intasome [7]. In the same region, the molar mass of the DNA is 22 kDa, consistent with two 19 bp DNA oligonucleotides.

ASTRA's protein conjugate analysis also reports the weight-average protein fraction across the peak. This value (89.6% protein) results in an average dn/dc of 0.183 mL/g across the peak, consistent with the estimated 0.182 mL/g used in the original analysis [1]. However, using the protein conjugate analysis means that no a priori knowledge about the amount of protein and DNA is required. As long as the extinction coefficients and dn/dc of each component are known, the measurement is valid and independent of such assumptions.

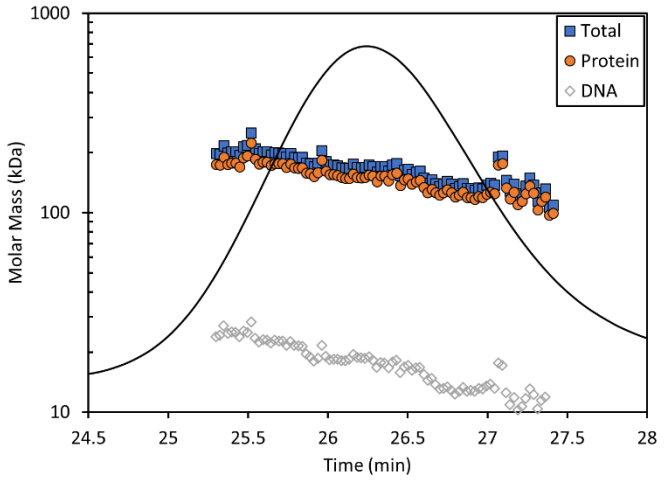


Figure 5: Protein conjugate analysis of PFV IN bound to U5 DNA. The molar mass of the protein fraction (orange circles), DNA fraction (open diamonds), and total complex molar mass (blue squares) are overlaid on the UV chromatogram.

Affinity and stoichiometry by CG-MALS

Under both pH conditions, the hetero-association between Cre and *loxP* is evident in the CG-MALS data. Figure 6 shows the measured weight-average molar mass during the CG-MALS hetero-association gradient. During each self-association gradient (Figure 4), the molar mass

of Cre and *loxP* alone were constant as a function of concentration at 39 kDa and 23 kDa respectively, indicating that neither species self-associated under these conditions. Therefore, the increase in M_w upon binding results from the interaction between Cre and *loxP* alone.

Differences in binding affinity and stoichiometry as a function of buffer pH are clearly seen in the graph. In both conditions, the maximum M_w occurs at an overall molar ratio of [Cre] = 2[*loxP*] suggesting that two Cre proteins bind each *loxP* recognition site. However, the maximum molar mass at pH 7.5 is ~40% higher than that measured at pH 9.5, indicating higher order complex formation. Fitting to the CG-MALS data at each condition provides the absolute stoichiometry and equilibrium association constants, for each complex that is formed under each buffer condition.

Binding at pH 7.5

As shown in Figure 6, the measured M_w reaches a maximum at the composition [Cre] = 2[*loxP*], suggesting an overall stoichiometric ratio of 2 Cre proteins per *loxP* site. However, the maximum measured M_w (113 kDa) is significantly greater than the molar mass of the $(Cre)_2$(*loxP*) complex (101 kDa). This suggests that higher order species formed when these two molecules interact.

In addition, at pH 7.5, CG-MALS data during the hetero-association gradient indicate the presence of slow association kinetics (Figure 7). After the flow is stopped, the protein-DNA solution comes to equilibrium in the MALS detector flow cell. The rise in light scattering signal corresponds to the increase in molar mass over time as Cre-*loxP* complexes are being formed. This slow association eventually reaches equilibrium over the course of ~15 minutes, timescales that are typical for some higher-order assembly processes.

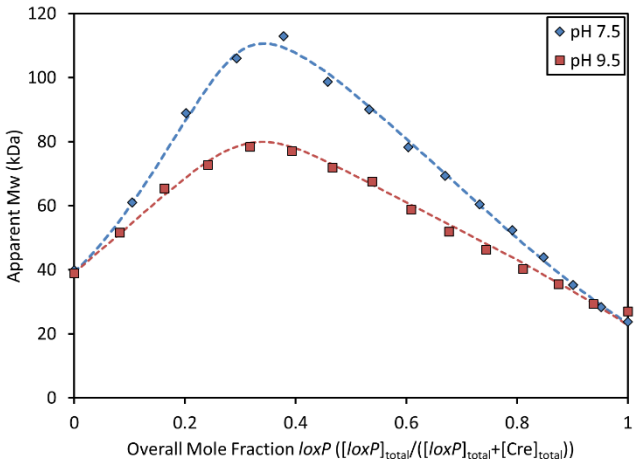
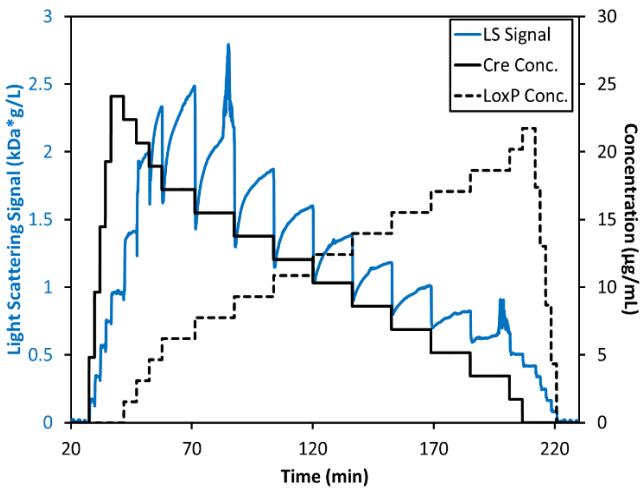


Figure 6: Overlay of measured weight-average molar mass and best fit curves (dashed lines) at pH 7.5 and pH 9.5. In both cases, the M_w reaches a maximum where [Cre] = 2[*loxP*] or an overall mole fraction of *loxP* = 0.33.



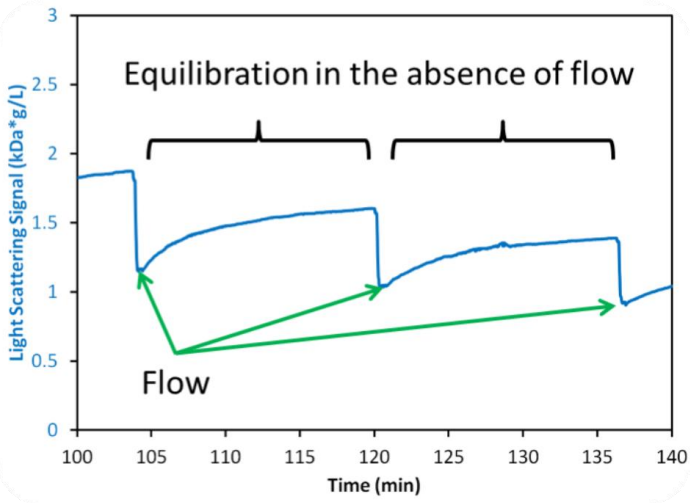


Figure 7: Left: Measured light scattering signal for the interaction between Cre and *loxP* at pH 7.5. Right: Light scattering intensity for two mixtures of Cre and *loxP*, showing the time during which the flow is stopped and the hetero-association reactions come to equilibrium within the MALS flow cell over ~15 minutes.

Multi-order cooperativity

In order to obtain a proper fit to the CG-MALS data, it was found that the association model requires three different Cre-*loxP* complexes in equilibrium with 1:1, 2:1, and 4:2 stoichiometries. The analysis shows that the binding sites are not equivalent; rather, Cre binds *loxP* with positive cooperativity. The first Cre binds each palindromic *loxP* recognition site with K_d = 170 nM, while the binding for the second Cre molecule occurs with approximately ten-fold higher affinity of K_d = 19 nM.

This 2:1 complex further self-assembles (synapsis) with K_d = 400 nM to form a final 4:2 stoichiometry. Figure 8 shows the distribution of species at pH 7.5. The concentrations of both $(Cre)_2(loxP)$ and $(Cre)_4(loxP)_2$ reach a maximum when [Cre] = 2[*loxP*] = 462 nM. As seen in the following section, synapsis does not occur at pH 9.5, indicating a further degree of cooperativity beyond the increase of affinity upon the first Cre-*loxP* binding event.

Benefits of CG-MALS

Unlike gel shift assays or other surface-based techniques, CG-MALS was able to measure all three complexes forming in solution, sampling the full binding mechanism. Other methods required controlling the concentrations of Cre and *loxP* to favor only the first binding of Cre to the palindromic DNA (assuming no synapsis occurs) or to ensure all the DNA was saturated with protein in order to focus on synapsis [7]. This type of piecemeal approach can miss key aspects of the interaction.

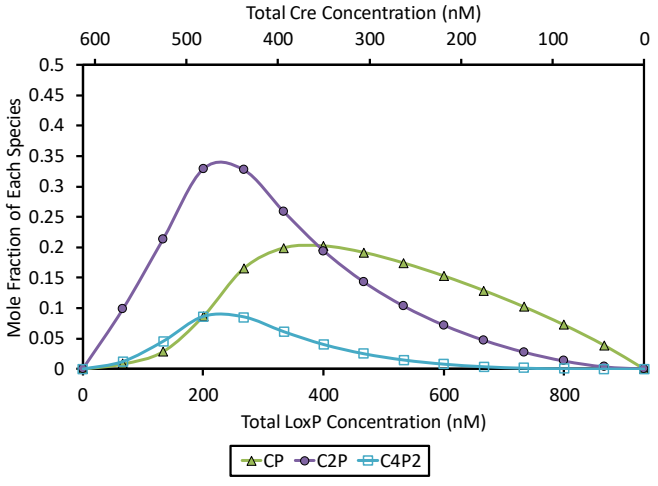


Figure 8: Molar composition of Cre-*loxP* complexes formed at pH 7.5—(Cre)(*loxP*) (CP, green triangles), $(Cre)_2(loxP)$ (C2P, purple circles), and $(Cre)_4(loxP)_2$ (open squares). The composition of free Cre and *loxP* monomers has been left off for clarity.

In addition, the assumption that only one interaction is happening at a time may be invalid and cause misinterpretation of the measured affinity. CG-MALS allows sampling of compositions where Cre is in excess and where *loxP* is in excess, thus allowing all parts of the interaction to be measured. Since the light scattering data are directly proportional to molar mass, data interpretation is straightforward, and allows direct measurement of the complexes being formed and their equilibrium constants.

Finally, CG-MALS provides a window into association kinetics in solution. The molar mass (light scattering intensity) as a function of time gives insight into the amount of time to reach equilibrium or if self-assembly continues indefinitely.

Binding at pH 9.5

The interaction measured at pH 9.5 is significantly different from that measured at pH 7.5. In this case, we do not observe the slow association kinetics that occurred at pH 7.5. Moreover, the weight-average molar mass is significantly lower than that measured at pH 7.5, indicating that synapsis does not occur (Figure 6). At this pH, two Cre proteins bind each *loxP* DNA, but this complex does not further associate into the synapse tetramer.

At pH 9.5, not only is synapsis abolished, but cooperativity is lost, and two Cre proteins bind each *loxP* with equivalent affinity, K_d = 24 nM. Figure 9 shows the distribution of species in solution during the hetero-association gradient.

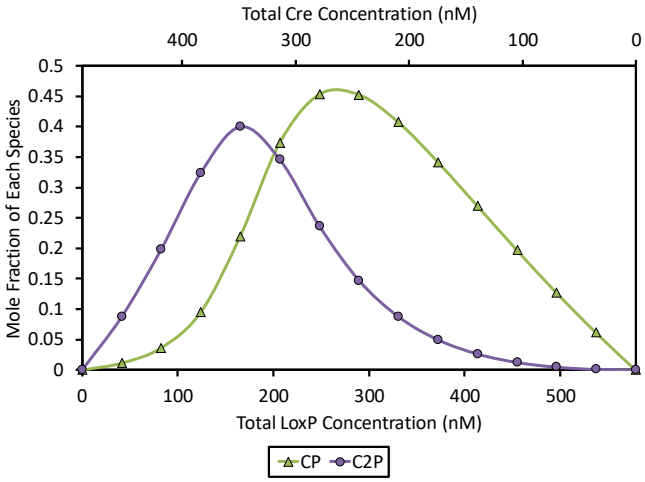


Figure 9: Molar composition of Cre-*loxP* complexes formed at pH 9.5—(Cre)(*loxP*) (CP, green triangles) and $(Cre)_2(loxP)$ (C2P, purple circles). The composition of free monomers has been left off for clarity.

As for the pH 7.5 data, the composition of the $(Cre)_2$(*loxP*) complex reaches a maximum where [Cre] = 2[*loxP*] = 343 nM. Since the *loxP* presents two equivalent binding sites at pH 9.5, the composition of (Cre)(*loxP*) complex forms the dominant species for compositions where [*loxP*] ≥ [Cre].

Conclusions

Multi-angle light scattering provides multiple, complementary measurements of protein-DNA complexes and their interactions.

- Protein conjugate analysis by SEC-MALS-UV-dRI is essential for characterizing two-component systems to assess conjugation ratio, oligomeric state and molar masses.
- CG-MALS provides rapid, in-depth quantification of protein-DNA interactions, elucidating cooperativity as well as self-assembly into higher-order structures. Microscopic equilibrium constants are determined for each binding site. Kinetics is also evident.

Utilizing the same MALS, UV and dRI detectors with distinct sample preparation and delivery front ends, SEC-MALS and CG-MALS constitute an effective and comprehensive biophysical characterization system.

These simple, robust techniques are uniquely suited for studying protein-DNA interactions, and are applicable to many other biomolecular complexes and interactions [8-19].

Acknowledgements

We would like to thank Kushol Gupta and Gregory Van Duyne at the University of Pennsylvania for kindly providing the Cre and *loxP* samples and buffers and SEC-MALS data.

To learn more about the theory, technology and applications of CG-MALS, please visit wyatt.com/CG-MALS.

Click the button below to request information on Calypso and DAWN instruments.

Request information

Bibliography

[1] K. Gupta et al., "Solution conformations of prototype foamy virus integrase and its stable synaptic complex with U5 viral DNA," *Structure* **20**, 1918-1928 (2012).

[2] J. Morimoto et al., "Dextran as a generally applicable multivalent scaffold for improving immunoglobulin-binding affinities of peptide and peptidomimetic ligands," *Bioconjugate Chemistry* **25**, 1479-1491 (2014).

[3] G. Y. Berguig et al., "Intracellular delivery and trafficking dynamics of a lymphoma-targeting antibody-polymer conjugate," *Molecular Pharmaceutics* **9**(12), 3506-3514 (2012).

[4] Y. Peng and L. Zhang, "Characterization of a polysaccharide-protein complex from Ganoderma tsugae mycelium by size-exclusion chromatography combined with laser light scattering," *J. Biochemical and Biophysical Methods* **30,** 243-252 (2003).

[5] S. Kenrick and M. Chen, "A light scattering toolbox for characterizing transport proteins and their interactions," 2014. [Online]. Available: http://www.wyatt.com/files/literature/app-notes/sec-mals-proteins/transport-protein-sec-cg-mals-poster.pdf.

[6] D. J. Slotboom et al., "Static light scattering to characterize membrane proteins in detergent solution," *Methods* **46**(2), 73-82 (2008).

[7] K. Ghosh et al., "Synapsis of *loxP* sites by Cre recombinase," *The Journal of Biological Chemistry* **282**(33), 24004-24016 (2007).

[8] J. Pallesen et al., "Structures of Ebola virus GP and sGP in complex with therapeutic antibodies," *Nature Microbiology* **9**(1), 2016 (2016).

[9] K. Hastie and E. O. Saphire, "Understanding Antibody and Viral Glycoprotein Interactions Using CG-MALS," 2015. [Online]. Available: http://www.wyatt.com/files/literature/app-notes/cg-mals/scripps-vgp+igg.pdf. [Accessed 18 4 2018].

[10] M. Zhao et al., "Mechanistic insights into the recycling machine of the SNARE complex," *Nature* **518**, 61-67 (2015).

[11] D. Some, "Light-scattering-based analysis of biomolecular interactions," *Biophysical Reviews* **5**(2), 147-158 (2013).

[12] J. Arora et al., "Charge-mediated Fab-Fc interactions in an IgG1 antibody induce reversible self-association, cluster formation, and elevated viscosity," *mAbs* **8**, 1561-1574 (2016).

[13] D. B. Halling et al., "Calcium-dependent stoichiometries of the KCa2.2 (SK) intracellular domain/calmodulin complex in solution," *J. Gen. Physiology* **143**(2), 231-252 (2014).

[14] A. K. Attri et al., "pH-dependent self-association of zinc-free Insulin characterized by concentration-gradient static light scattering," *Biophys Chem* **148**(1-3), 28-33 (2010).

[15] A. Martos et al., "Characterization of Self-Association and Heteroassociation of Bacterial Cell Division Proteins FtsZ and ZipA in Solution by Composition Gradient-Static Light Scattering," *Biochemistry* **49**(51), 10780-10787 (2010).

[16] S. L. Mitchell et al., "The VieB auxiliary protein negatively regulates the VieSA signal transduction system in Vibrio Cholerae," *BMC Microbiology* **15**(59), 1-16 (2015).

[17] H. L. Zhu et al., "Quantitative Characterization of Heparin Binding to Tau Protein," *J. Biolog Chem* **285**(6), 3592-3599 (2010).

[18] J. C. Gaulding et al., "Reversible Inter- and Intra-Microgel Cross-Linking Using Disulfides," *Macromolecules* **45**(1), 39-45 (2012).

[19] G. V. Crichlow et al., "Dimerization of FIR upon FUSE DNA binding suggest a mechanism of c-myc inhibition," *The EMBO Journal* **27**, 277-289 (2008).